ČESKÉ VYSOKÉ
UČENÍ TECHNICKÉ
V PRAZE

FAKULTA STROJNÍ

HABILITAČNÍ PRÁCE

NOMINÁLNÍ ANALÝZA ZOBRAZENÍ A NÁVRH KOMBINOVANÝCH OPTICKÝCH PRVKŮ A SOUSTAV

PETR POKORNÝ

PRAHA 2021

Předmluva

V rámci svého působení na katedře fyziky Fakulty stavební ČVUT v Praze jsem se mohl výzkumně podílet na řadě témat z oblasti aplikované optiky. Zejména se jedná o analýzu a použití aktivních optických prvků v zobrazovacích a měřických systémech, teoretickou analýzu difrakčních jevů a další oblasti optického zobrazení, primární návrh optických soustav s korigovanými aberacemi nebo optické (laserové) skenování.

Tato práce je průvodním textem a souhrnným představením témat publikací v impaktovaných mezinárodních časopisech, na kterých jsem se autorsky podílel a které významným způsobem přispěly k rozšíření znalostí daného oboru. V příloze jsou kopie těchto prací uvedeny a v průběhu textu se na ně průběžně odkazuji.

Tematicky jsem práci rozdělil do tří základních kapitol. První z nich představuje analýzu a použití aktivních kapalinových membránových čoček při návrhu kombinovaných optických členů. Nejprve jsou aktivní prvky představeny obecně, poté je diskutován postup modelování deformace membrán zmíněných čoček a ten je následně využit k nominálním propočtům kombinovaných členů. Druhá z kapitol se zabývá vybranými oblastmi optického zobrazení, jmenovitě skalární teorií difrakce a jejím aproximativním řešením v případě difrakčních mřížek a kruhových apertur a nominálním návrhem optických soustav s korigovanými aberacemi. Jako poslední téma, kapitolu, jsem zařadil analýzu určení přesnosti polohy bodu při pozemním laserovém skenování.

Rád bych zde za odbornou spolupráci na zmíněných tématech a publikaci prací poděkoval svým kolegům ze Skupiny aplikované optiky, která na katedře fyziky Fakulty stavební ČVUT v Praze působí, jmenovitě zejména prof. RNDr. Antonínu Mikšovi, CSc., prof. Ing. Jiřímu Novákovi, Ph.D., doc. Ing. Pavlovi Novákovi, Ph.D., a studentům Ing. Filipovi Šmejkalovi a Ing. Jindřichovi Brzobohatému. Dále zaslouží poděkování kolegové prof. Ing. Milan Jirásek, DrSc., a Ing. Martin Horák, Ph.D., z katedry mechaniky na téže fakultě.

V Praze dne 1. listopadu 2021

Petr Pokorný

Obsah

1 Úvod

Tato práce je průvodním textem a souhrnným představením témat publikací v impaktovaných mezinárodních časopisech, na kterých se autor podílel, které jsou stěžejní součástí této práce a podle nichž jsou děleny jednotlivé kapitoly do tří základních celků.

Prvním tématem, které je publikováno v pracích [1–7], je analýza aktivních optických prvků a jejich využití v zobrazovacích a měřických systémech. Nejprve jsou aktivní prvky představeny obecně, poté je diskutován postup modelování deformace membrán kapalinových čoček a ten je následně využit k nominálním propočtům kombinovaných členů. Také je zmíněna možnost využití aktivních čoček v optickém skenování.

Druhá kapitola se zabývá vybranými tématy optického zobrazení a souhrnem prací [8–16]. Jedná se zejména o témata skalární teorie difrakce a jejího aproximativního řešení v případě difrakčních mřížek a kruhových apertur a dále nominální návrh optických soustav s korigovanými aberacemi.

Poslední oblastí, které je věnován text této práce, je geometrický popis a charakteristika přesnosti měření optickými (laserovými) skenery a souhrn výstupů publikací [17, 18].

Kopie výše zmíněných článků jsou uvedeny postupně v Příloze A, kde čtenář najde úplné informace a reference.

Oblast, které se autor také věnoval a publikace [19–21] jsou uvedeny v Příloze A jako výstup jeho vědecké činnosti, avšak vzhledem k rozsahu není pojednána v této práci, je měření vybraných parametrů optických soustav.

2 Aktivní optické prvky a jejich využití

Aktivní optické prvky se v posledních letech staly cílem zájmu velké řady výzkumníků a výzkumných týmů po celém světě a stále nacházejí čím dál větší uplatnění v zobrazovacích a měřických systémech. Jejich největší předností je možnost velmi rychlé a kontrolované změny zobrazovacích parametrů. Dříve bylo nutné pro modifikaci parametrů optických systémů zpravidla měnit vzájemné pozice jednotlivých optických členů, ze kterých byla soustava sestavena. S pomocí aktivních prvků je tato nutnost výrazným způsobem potlačena, jelikož ty jsou samy o sobě schopné měnit například optickou mohutnost, polarizaci nebo tvarové parametry.

Tématu analýzy a aplikace aktivních optických prvků se autor intenzivně zabýval a výsledky jeho výzkumné činnosti jsou publikovány v řadě mezinárodních impaktovaných časopisů [1–7]. Kopie těchto publikací jsou uvedeny v Příloze A.

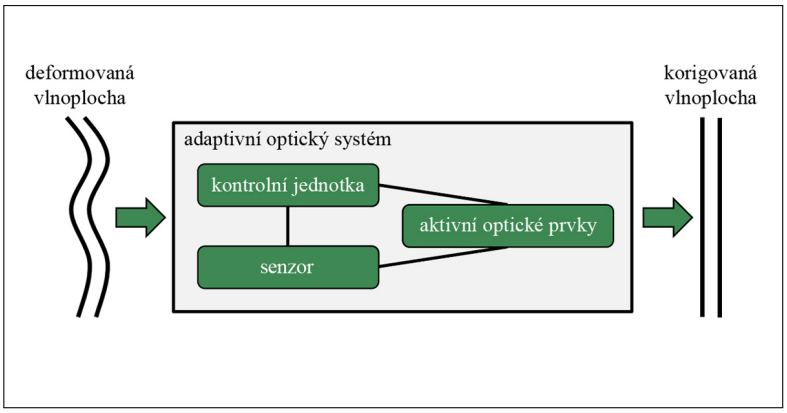
První část této kapitoly představí základní principy fungování aktivních optických prvků, jejich základní dělení a současně komerčně dostupné produkty. Následuje souhrnný popis modelování deformace membrán kapalinových čoček a návrh nerovnoměrné tloušťky membrány pro aplikace v optice, kterému se autor věnoval a podílel se na publikaci prací [1–3]. Jednotlivé prvky lze kombinovat do hybridních optických systémů nebo navrhovat nové komponenty specifických zobrazovacích vlastností (např. čočky dvojí křivosti, zoom systémy nebo hybridní čočky [4–6]). Možnost návrhu startovních parametrů membránových čoček pro tato využití je představena v další části kapitoly. Poslední část kapitoly ukazuje použití aktivních membránových kapalinových čoček v oblasti optického skenování [7].

2.1 Základní principy fungování aktivních optických prvků

Základní princip aktivní změny optického zobrazení daného prvku můžeme popsat schopností prostorově měnit geometrický tvar funkčních ploch (například ploch čoček, zrcadel, optických rozhraní a podobně), nebo jako možnost měnit rozložení indexu lomu materiálu, ze kterého je prvek, nebo jeho část, vyroben. Tím je přímo ovlivňováno světelné záření, které daným optickým členem, prostředím, prochází, nebo které se od něj odráží.

Zmínku o aplikaci aktivních prvků lze datovat do 60. let 20. století. H. W. Babcock r. 1953 publikoval teoretickou práci, ve které se zabýval korekcí vlnoploch pomocí adaptivní optiky v astronomii [22]. Další obecný přehled týkající se adaptivní optiky může čtenář nalézt například v publikacích [23–27].

Základní fungování adaptivního systému je ukázáno na obr. 2.1. Ten se obecně skládá



Obrázek 2.1: Princip adaptivního optického systému

ze tří komponent:

1. ze senzoru (například senzoru vlnoplochy, CCD senzoru apod.), který registruje stav dopadajícího vlnového pole vstupujícího do systému (v případě senzoru vlnoplochy například její deformovaný tvar, v případě CCD senzoru například rozložení energie obrazu bodu);

2. z kontrolní jednotky, která kvantifikuje příchozí vlnění a určuje vhodnou korekci pro zajištění zobrazení požadovaných parametrů;

3. z kombinované (hybridní) optické soustavy s aktivními optickými prvky, která má za úkol korigovat záření a předávat ho dál systému.

Výše popsané tři základní kroky (registrace, kvantifikace, korekce) jsou iterativně opakovány tak, aby bylo zajištěno zobrazení s co nejvyšší kvalitou. Příchozí vlnění je tak dynamicky korigováno a optické zobrazení adaptivním systémem může mít takové parametry, kterých by za pomoci klasické optiky s výlučně fixními prvky (klasické čočky, hranoly a zrcadla) nebylo možné dosáhnout.

Od prvního uvedení myšlenky adaptivní optiky a aktivních prvků došlo a stále dochází k jejich prudkému rozvoji v oblastech výzkumu, vývoje a aplikací. V současnosti jsou komponenty uplatňovány běžně v oblastech inspekční techniky, spotřební elektroniky, záznamu dat, dále poté v medicíně (korekce zraku, zrakové simulátory, zobrazovací a diagnostické systémy), optické mikroskopii, laserových technologiích, holografii nebo komunikačních systémech ve volném prostoru [28–36].

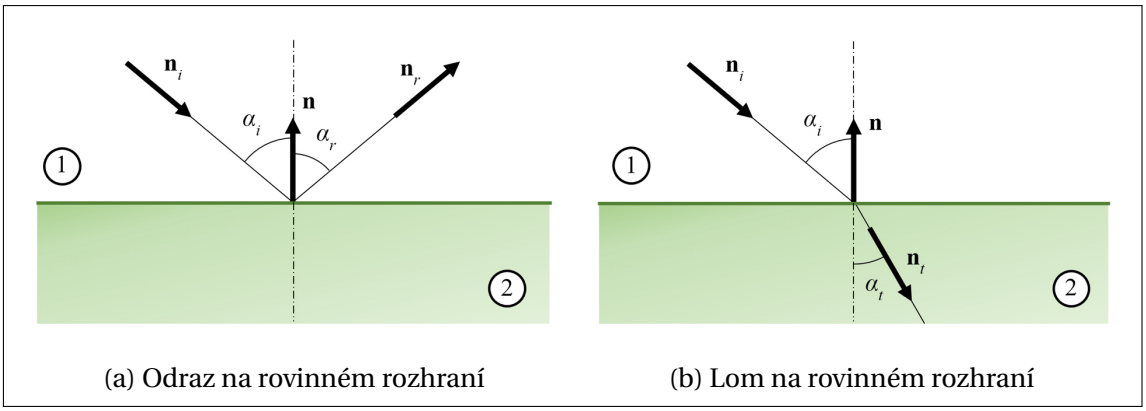
Jak už bylo zmíněno v úvodu této kapitoly, můžeme principy změny optického zobrazení rozdělit do dvou kategorií:

1. pomocí změny geometrického tvaru funkčních ploch,

2. prostorovou změnou indexu lomu materiálu.

Zabývejme se nyní v krátkosti prvním případem. Rozhraní mezi dvěma optickými prostředími, jejichž indexy lomu označíme n_i a n_t můžeme obecně vyjádřit pomocí implicitní funkce $F(\mathbf{r},\mathbf{p}(\xi)) = 0$ [37–39], kde $\mathbf{r}$ značí polohový vektor, a tedy funkční závislost plochy na poloze, a $\mathbf{p}$ je vektor dalších parametrů závislých na proměnné ξ (např. čas). Jednotkový normálový vektor daného rozhraní lze poté snadno určit jako [37–39]

$$\mathbf{n}(\xi) = \frac{\nabla F(\mathbf{r},\mathbf{p}(\xi))}{\|\nabla F(\mathbf{r},\mathbf{p}(\xi))\|}. \tag{2.1}$$



(a) Odraz na rovinném rozhraní (b) Lom na rovinném rozhraní

Obrázek 2.2: Zákon odrazu a lomu

Rovnice (2.1) popisuje geometrickou proměnlivost rozhraní v závislosti na proměnné ξ. Dosazením (2.1) do známého zákona odrazu nebo lomu nebo s použitím Fresnelových vztahů můžeme snadno vyhodnotit vlastnosti záření, které se bude od daného rozhraní odrážet, nebo jím bude procházet. Například zákon odrazu ve vektorovém tvaru můžeme psát jako [40–43]

$$\mathbf{n}_r = \mathbf{n}_i - 2\mathbf{n}(\mathbf{n}\cdot\mathbf{n}_i) \tag{2.2}$$

a zákon lomu formou

$$\mathbf{n}_t = \frac{n_i}{n_t}\mathbf{n}_i - \mathbf{n}\left\{\sqrt{1-\left(\frac{n_i}{n_t}\right)^2\left[1-(\mathbf{n}\cdot\mathbf{n}_i)^2\right]} + \frac{n_i}{n_t}(\mathbf{n}\cdot\mathbf{n}_i)\right\}, \tag{2.3}$$

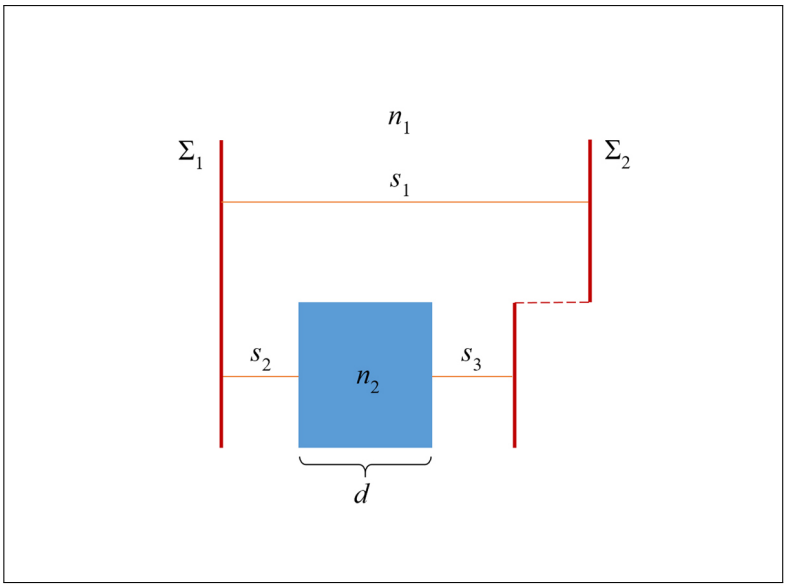
kde $\mathbf{n}_i$, $\mathbf{n}_r$ a $\mathbf{n}_t$ značí postupně jednotkové normálové vektory dopadající, odražené a prošlé vlnoplochy ve zkoumaném bodě, $\mathbf{n}$ značí jednotkový normálový vektor plochy rozhraní (orientovaný směrem k příchozí vlnoploše viz obr. 2.2) a n_i a n_t jsou indexy lomu prostředí, ze kterého vlna přichází a do kterého prochází.

Změny geometrického tvaru funkčních ploch obecně využívají aktivní kapalinové čočky [26, 28, 29] nebo deformovatelná zrcadla [30, 31].

Druhou kategorii změny optického zobrazení, tj. pomocí prostorové změny indexu lomu materiálu, můžeme demonstrovat nejlépe pomocí optické dráhy OPD, která je definována jako [40–43]

$$OPD = \int_{\mathrm{A}}^{\mathrm{B}} n(s)\,\mathrm{d}s, \tag{2.4}$$

kde $n(s)$ je index lomu daného prostředí závislý na poloze s a ds je element křivky. Optická dráha je jednoznačně závislá na indexu lomu prostředí v daném místě, v principu je to veličina odpovídající vzdálenosti, kterou by světlo urazilo ve vakuu. Dále je známa tzv. věta o stálosti optických drah [40–43], která říká, že optická dráha mezi dvěma vlnoplochami téhož svazku paprsků je v izotropním prostředí konstantní. Uvažujme dále jednoduchý diskrétní případ na obr. 2.3 (tj. integrál nahraďme sumou součinů elementů dráhy paprsku a hodnot indexu lomu v daném místě). Vlnoplocha Σ_1 prochází homogenním izotropním prostředím indexu lomu n_1, až dorazí k variaci materiálu charakterizované indexem lomu n_2 (předpokládejme opět homogenní izotropní prostředí). Optická dráha odpovídající úsečce s_1 bude $OPD_1 = n_1 s_1$. Optická dráha v místech změny bude $OPD_2 = n_1(s_2 + s_3) + n_2 d$. Budeme-li znát vzdálenosti s_1, s_2 a d, dostaneme z podmínky rovnosti optických drah mezi dvěma vlnoplochami příslušejících jednomu svazku,



Obrázek 2.3: Šíření vlnoplochy prostředím s variací indexu lomu materiálu

tj. $OPD_1 = OPD_2$, pro hodnotu s_3 výraz $s_3 = s_1 - s_2 - d\frac{n_2}{n_1}$. Jinými slovy to znamená, že bude-li v cestě části paprsku stát prostorová změna indexu lomu, dojde k prostorové deformaci výsledné vlnoplochy.

Tohoto principu využívají zejména tzv. prostorové modulátory světla pracující na bázi tekutých krystalů [32, 33, 44, 45].

2.1.1 Aktivní čočky

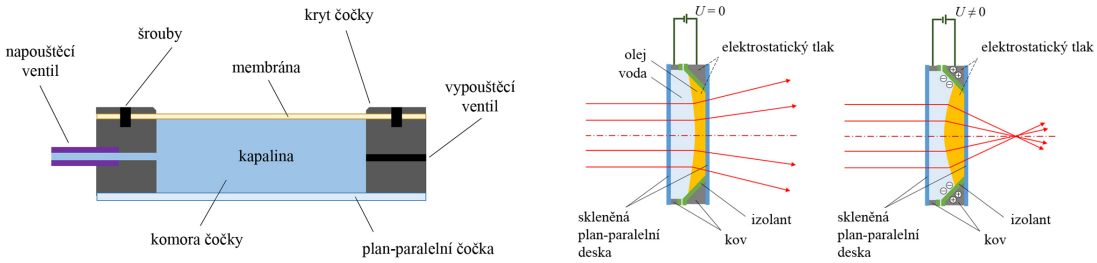
Představme nyní stručně základní funkční principy aktivních čoček, které, jak bylo uvedeno výše, modifikují průchozí optické záření prostorovou změnou rozhraní mezi dvěma či více optickými prostředími. Jako základní kategorie aktivních čoček můžeme jmenovat například:

- membránové kapalinové čočky,
- elektrosmáčivé čočky,
- dielektroforetické čočky,
- akustooptické čočky.

Princip fungování membránových kapalinových čoček je ukázán na obr. 2.4a. Základem čočky je komora, která je z jedné strany tvořena planparalelní deskou a z druhé pružnou membránou z vhodného materiálu, která je připevněna ke stěnám komory krytem čočky a šrouby. Do komory je napouštěcím ventilem přivedena optická kapalina. Změnou objemu kapaliny (např. ještě s použitím vypouštěcího ventilu) lze poté měnit hydrostatický tlak uvnitř komory čočky. Je-li uvnitř systému přetlak, dojde k vyboulení membrány a zvýšení optické mohutnosti čočky. V opačném případě dojde ke klenutí membrány a zmenšení optické mohutnosti. Výše popsané schéma konstrukce je rozdílné v závislosti na konkrétním výrobci čoček.

Membrány čoček musí splňovat řadu parametrů:

- vysoká světelná propustnost,



(a) Kapalinová membránová čočka (b) Elektrosmáčivá aktivní čočka [29]

Obrázek 2.4: Principy fungování aktivních čoček

- malá disperze pro požadované světelné spektrum,
- dostatečně hladký povrch pro všechny použité stavy napnutí,
- vysoká návratná deformovatelnost,
- vhodné chemické vlastnosti (např. nereaktivnost se zvolenou kapalinou),
- vhodné fyzikální vlastnosti (homogenita materiálu, teplotní stálost apod.).

Zvolená optická kapalina poté musí splňovat například tato kritéria:

- vysoká světelná propustnost,
- širokopásmovost, tj. propustnost pro požadované vlnové délky (např. pro IR oblasti mají některé kapaliny jinou spektrální propustnost než pro viditelné světlo),
- odpovídající index lomu (menší deformace způsobí větší modifikaci procházejícího záření),
- malá hustota (vzhledem ke gravitačním účinkům na kapalinu),
- bez chemické reakce s membránou,
- nevypařující se v použitých podmínkách,
- malá viskozita (pro dosažení vysokofrekvenčních změn zobrazení).

Zmiňme nyní princip fungování čoček založených na elektrosmáčivosti [26, 46], což je elektrokapilární jev, kdy vodivá kapalina (kapka elektrolytu) umístěná na nevodivé dielektrické vrstvě zformuje kapénku s kontaktním úhlem θ_0 na okrajích a po přivedení elektrického napětí je možné tento úhel měnit. Změna kontaktního úhlu θ na přivedeném napětí U je popsána např. Young-Lippmannovou rovnicí [26, 46]

$$\cos\theta = \cos\theta_0 + \frac{\varepsilon_0 \varepsilon}{2d\gamma} U^2, \tag{2.5}$$

kde θ_0 je kontaktní úhel při $U = 0$ V, ε_0 je permitivita vakua, ε je dielektrická konstanta izolantu, d je tloušťka izolantu, γ je povrchové napětí mezi elektrolytem a okolní atmosférou

(povrch kapaliny nepřipadající izolantu) a U je přivedené napětí. Velikost změny kontaktního úhlu je tedy nepřímo úměrná tloušťce izolantu a kvadraticky úměrná přivedenému napětí.

Elektrosmáčivé čočky jsou zpravidla menší než výše zmíněné membránové čočky a jejich rozsah lámavosti je do 15 dpt. Možná praktická realizace je v jednoduchosti ukázána na obr. 2.4b. Čočka je konstruována pomocí dvou nemísitelných kapalin, které jsou umístěny mezi dvěma krycími sklíčky (planparalelními deskami). Přivedené napětí má za následek změnu tvaru rozhraní mezi těmito kapalinami a je tak možné provádět vysokofrekvenční změny zobrazení daného optického elementu. Výrobně jsou tyto čočky poměrně nenáročné a není třeba mechanických pohybů jako například u čoček membránových.

Kapaliny elektrosmáčivých čoček by měly splňovat následující parametry:

- nemísitelnost,
- vhodná rozdílnost indexu lomu,
- malý rozdíl hustot (vzhledem ke gravitačním silám),
- jedna z kapalin musí být vodivá, zatímco druhá je izolantem,
- dostatečná transparentnost pro použité záření,
- chemická nereaktivnost se světlem.

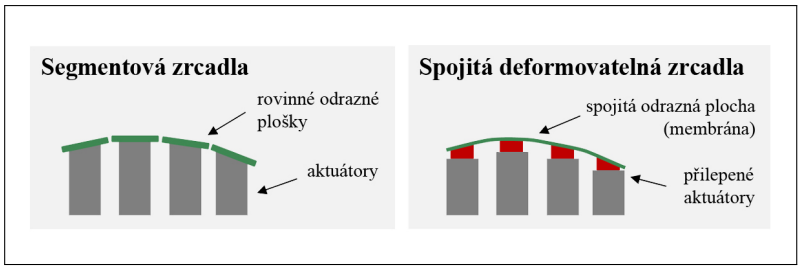
Čočky dielektroforetické využívají tzv. dielektroforetické síly (dielektroforeze) [26, 47]. To je síla působící na dielektrické částice (neutrální částice s rovnoměrně rozloženými kladnými a zápornými náboji) v nehomogenním elektrickém poli, které lze vytvářet například použitím zakřivených elektrod nebo tzv. proužkovaných nebo dírkovaných elektrod. Malé množství dielektrické kapaliny v nehomogenním poli tak může aktivně a kontrolovaně měnit svůj tvar a tím vytvořit aktivní optický prvek.

Konstrukčně jsou pak tyto čočky velmi podobné elektrosmáčivým. Mají však rozdílné vlastnosti použitých kapalin:

- kapaliny musí být nevodivé s různými dielektrickými konstantami,
- působící elektrické pole musí být nehomogenní.

Další parametry jako propustnost, nemísitelnost nebo teplotní stabilita kapalin apod. jsou obdobné. V porovnání s elektrosmáčivými čočkami jsou dielektroforetické mnohem méně energetický náročné (zejména díky malé generaci tepla) a stabilnější. Ovšem vývoj dielektrických kapalin a dielektroforetických čoček je zatím v počátcích v porovnání s čočkami elektrosmáčivými. Je však příslibem aplikací nejen v zobrazovacích systémech v mikro a makro rozměrech, ale například i v oblastech aktivních děličů svazků, aktivních clon nebo v difuzérech svazků.

Na závěr této kapitoly zmiňme čočky akustooptické [26]. Ty využívají akustických vln ke změně rozložení indexu lomu v materiálu působením na jeho molekuly. Čočka poté může být konstruována například pomocí válcové komory se dvěma nemísitelnými kapalinami (např. voda a silikonový olej), kdy jedno z rozhraní komory je osazeno konkávním akustickým převodníkem. Pomocí něho lze cílit akustické vlny na rozhraní kapalin, které mají rozdílné energetické hustoty (rychlosti šíření zvuku) a indexy lomu, čímž dochází k vzniku stojaté akustické vlny a zakřivení rozhraní. Změnou napětí na převodníku lze poté měnit amplitudu stojaté vlny – optickou mohutnost čočky.



Obrázek 2.5: Princip fungování segmentových a mikromechanických membránových zrcadel

2.1.2 Deformovatelná zrcadla

Deformovatelná zrcadla jsou aktivní optické prvky, které, jak již název napovídá, modifikují záření pomocí změny tvaru reflexní funkční plochy. V praxi se využívá velká řada principů, jak deformace docílit. Jako příklad zmiňme základní tři typy deformovatelných zrcadel:

- zrcadla mechanicky deformovatelná pomocí aktuátorů (zpravidla piezoelektrických),
- mikromechanická membránová zrcadla,
- mikro-elektro-mechanická zrcadla (tzv. MEMS z angl. Micro Electro Mechanical Systems).

Způsob provedení deformace uvedený výše je volen s ohledem na zamýšlené použití zrcadel a jejich konkrétní aplikaci. Princip fungování mechanicky deformovatelných zrcadel je velmi snadný. Na obr. 2.5 je schematicky ukázána konstrukce vybraných zrcadel. Aktuátory jsou připevněny k jednotlivým segmentům reflexní plochy nebo ke spojité membráně. Změnou prostorové polohy a orientace aktuátorů dochází k směřování segmentů nebo deformaci membrány, a tím ke změně zobrazovacích vlastností zrcadla.

Na levé části obr. 2.5 je ukázáno schéma segmentových zrcadel, která byla často využívána např. v astronomii [23–25, 27]. Membránová deformovatelná zrcadla, jejichž princip je schematicky ukázán na pravé části obr. 2.5, jsou využívána v aplikacích, kde není třeba tak velkého rozsahu deformace vzhledem k použití membrány.

Prakticky jsou používána různá rozložení aktuátorů, např. ve čtvercových nebo hexagonálních polích. Jelikož je v praxi nejvíce používáno kruhových apertur, jeví se hexagonální rozložení jako nejefektivnější [25]. Velmi často jsou používány aktuátory feroelektrické, vyráběné z keramických materiálů a založené na inverzním piezoelektrickém nebo elektrostrikčním efektu [25]. Inverzní piezoelektrický jev popisuje lineární závislost deformace elementu na působícím elektrickém poli, zatímco elektrostrikční jev je charakterizován deformací nevodiče nebo dielektrika v závislosti kvadratické.

MEMS deformovatelná zrcadla využívají obecného Coulombova zákona elektrostatiky, který popisuje velikost síly působící mezi dvěma bodovými náboji jako nepřímou kvadratickou závislost na vzdálenosti mezi danými náboji. Budou-li mít tedy dvě elektrody (elektrostatické aktuátory) náboje, bude mezi nimi působit síla, a může tak dojít k deformaci, tj. ke změně jejich vzájemné vzdálenosti.

Komerčně je nabízena celá řada deformovatelných zrcadel založených na různých principech [30, 31, 35, 36]. Například firma Imagine Optic [35] je distributor zrcadel o průměrech od 7 do 500 mm s deformací do ± 50 μm. Společnost Flexible Optical [30] nabízí

piezoelektrická a mikromechanická membránová deformovatelná zrcadla o aperturách 15 mm až 50 mm s rozsahem deformace do 12 μm. Komerčně dostupná zrcadla firmy Thorlabs [36] jsou k dispozici s 10mm až 14mm aperturami, v případě MEMS konstrukce nabízí tato společnost maximální deformace 3.5 μm se sub-nanometrovou opakovatelností. Firma Alpao [31] nabízí zrcadla do průměru 240 mm s maximálními deformacemi vlnoplochy až 80 μm. Zájemce může ale nalézt celou řadu dalších společností, které deformovatelná zrcadla nabízí.

2.1.3 Prostorové modulátory světla

Prostorové modulátory světla se řadí mezi aktivní optické prvky na bázi tekutých krystalů [23, 44, 45, 48], což jsou elementy na pomezí krystalické a tekuté struktury charakteristické svou optickou a elektrickou anizotropií. Toto pomezí krystalické a tekuté formy je dáno molekulární kompaktností a orientací takovou, která se při větším množství materiálu vyznačuje krystalickými vlastnostmi (jako např. dvojlom), ale zároveň natolik slabou, že je možné ji snadno měnit vnějšími vlivy (např. elektrickým polem).

Obecně je známa celá řada typů tekutých krystalů [23, 44, 45], ovšem v oblasti aktivních optických prvků se používají zpravidla tzv. nematické tekuté krystaly. Jedná se o modulátory polarizace, jejichž natočená molekulární struktura rotuje rovinu polarizace světla, které skrze krystaly prochází. Uspořádáme-li pole buněk tekutých krystalů do pravidelné struktury pixelů (obdoba CCD senzoru), poté můžeme prostorově modulovat průchozí záření (na vstupu zpravidla lineárně polarizované) v místě každého pixelu a docílit tak modulace vlnoplochy.

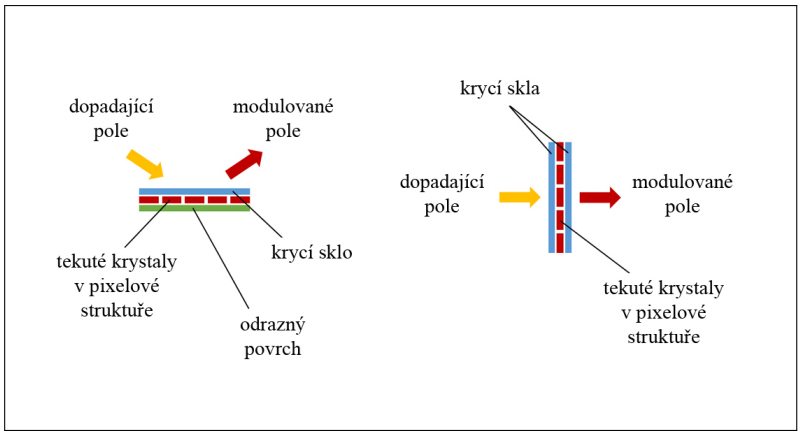
Princip prostorové modulace lze snadno demonstrovat následujícím způsobem. Je-li přítomna různá orientace molekul krystalů v rámci struktury pixelů, potom bude každý z pixelů charakterizován různou dielektrickou konstantou (permitivitou) ε. Je široce známo, že pro rychlost šíření v v daném prostředí platí $v = 1/\sqrt{\varepsilon\mu}$, kde μ značí permeabilitu, pro index lomu takového prostředí platí $n = c/v$, kde c je rychlost světla ve vakuu, a optický dráhový rozdíl je dán vztahem $OPD = \int n\,\mathrm{d}s$. Kombinací těchto jednoduchých vztahů lze snadno usoudit, jak bude vlnoplocha procházející strukturou pixelů tekutých krystalů modulována.

Zařízením pracujícím na tomto principu se říká prostorové modulátory světla (SLM z angl. Spatial Light Modulator) [32, 33].

Jak bylo zmíněno, modulátory na bázi tekutých krystalů jsou zpravidla pixelové struktury. To má za následek několik jevů, se kterými je nutno při návrhu optické soustavy s danými prvky počítat. Mezi každým z pixelů se bude nacházet slepé místo, kde nebude docházet k modulaci (rámeček pixelu). Velikost aktivní plochy pixelu může být popsána tzv. faktorem zaplnění (z angl. fill factor), který vyjadřuje poměr součtu aktivních ploch všech pixelů vůči ploše celého senzoru. Modulátory se tedy kromě primární modulace projeví jako difrakční mřížky. Dále povrch prostorových modulátorů není zpravidla rovinný (jako následek výrobního procesu), a tak je třeba modulovanou vlnoplochu průběžně kontrolovat a případně zavádět vhodné korekce. Omezujícím faktorem je také to, že mezi sousedními pixely nemůže zpravidla dojít k příliš velké změně indexu lomu. Rozsah modulace je tak určitou měrou omezen.

Prostorové modulátory světla svou konstrukcí můžeme dělit na tři základní kategorie:

- reflexní (odrazné),
- transmisní (propustné),



Obrázek 2.6: Schéma odrazných a transmisních prostorových modulátorů světla

- difraktivní.

Na obr. 2.6 je schéma reflexního a transmisního modulátoru naznačeno. Odrazný modulátor je v principu konstruován tak, že tekuté krystaly jsou umístěny mezi planparalelní a reflexní desku. Příchozí záření tedy prochází skrze krystaly, odráží se od reflexní plochy a pokračuje zpět skrze krystaly druhým průchodem. V případě propustných modulátorů je struktura krystalů umístěna mezi dvě planparalelní desky a záření po průchodu skrz krystaly pokračuje modulované dál. Difraktivní modulátory mohou být konstruovány obojím způsobem, jak reflexním, tak transmisním. Jejich princip je založen na variabilním uzavření průchodu světla některými pixely. Odražené nebo prošlé záření poté podléhá předem definované difrakci na struktuře displeje modulátoru.

Komerčně jsou prostorové modulátory nabízeny např. firmou Hamamatsu [33], která nabízí fázové reflexní modulátory. Firma Holoeye [32] je výrobcem a prodejcem fázových nebo amplitudových modulátorů (případně kombinace) reflexního i transmisního typu.

Existuje celá řada aplikací, ve kterých prostorové modulátory nachází uplatnění. Jako příklad jmenujme digitální holografii, optické pasti a mikro-manipulace, tvarování laserových svazků nebo optickou metrologii [32, 33].

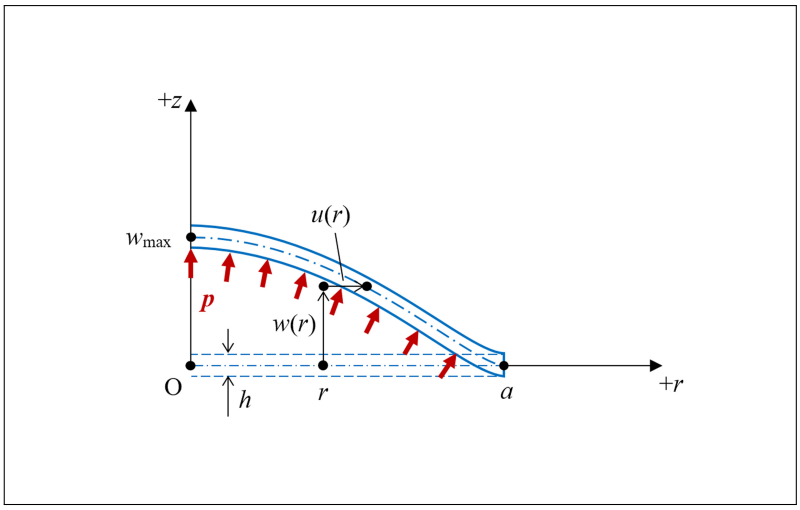
2.2 Deformace membrán kapalinových čoček

V této kapitole budou shrnuty závěry z tématu výpočtu deformace membrán kapalinových čoček a návrhu nerovnoměrné tloušťky membrány pro aplikace v optice, kterému se autor věnoval a podílel se na publikaci prací [1–3].

Zabývejme se nejprve zjednodušeným případem – deformace kruhové předpjaté osově symetrické membrány konstantní tloušťky [2, 3].

Vzhledem k tomu, že deformace membrány v aktivních čočkách několikanásobně převyšují její tloušťku, není možné s dostatečnou přesností použít klasické řešení výpočtu založené na lineární teorii pružnosti [49, 50]. Vliv radiální deformace membrány také nemůže být pominut. Model uvažovaný v pracích [2, 3] je založen na přesných geometrických rovnicích a Saint Venantově-Kirchhoffově materiálovém modelu, který předpokládá lineární vztah mezi Greenovým-Lagrangeovým přetvořením a druhým Piolovým-Kirchhoffovým napětím [51].

Předpokládejme tedy kruhovou osově symetrickou membránu kapalinové čočky, jejíž osa symetrie je totožná se z-ovou souřadnou osou a radiální osa r je dána dle obr. 2.7.



Obrázek 2.7: Schéma radiálního řezu kruhovou osově symetrickou membránou kapalinové čočky při působení rovnoměrného tlaku p (čárkovaná modrá linie – počáteční tvar při nulovém tlaku a počátečním předpětí, $w(r)$ a $u(r)$ – vertikální a horizontální výchylka bodu o počátečních souřadnicích $[r,0]$, a – poloměr membrány, h – tloušťka membrány, $w_{\max}$ – maximální výchylka) [3]

Membrána je charakterizována poloměrem a a konstantní tloušťkou h takovou, že ohybová tuhost může být zanedbána, a je fixována po svém obvodu. Působením tlaku optické kapaliny se deformuje, přičemž obecný bod o počáteční poloze $[r,0]$ je vychýlen do nové pozice $[r+u(r), w(r)]$, kde $w(r)$ značí vychýlení ve směru osy z a $u(r)$ je radiální vychýlení. Výsledný deformovaný tvar membrány poté může být popsán funkcí $g(r)$ implicitně dané vztahem $g(r+u(r)) = w(r)$. Předpokládejme dále, že membrána je ve svém počátečním stavu (nepůsobí-li tlak p) předpjatá a počáteční protažení (angl. stretch) je ve všech směrech $\lambda_0 > 1$. Dále uvažujme, že změna tloušťky membrány je zanedbatelná, a tedy výsledný tvar může být charakterizován středovou křivkou, viz silná čerchovaná modrá křivka na obr. 2.7. Nominální geometrickou tloušťku membrány před předpětím označme $\tilde{h}$ a poloměr $\tilde{a} = a/\lambda_0$.

Jak je v publikacích [2, 3] podrobně odvozeno, výslednou deformaci střednicové plochy membrány lze popsat pomocí rovnic rovnováhy

$$\begin{aligned} \tilde{h}\left[r\sigma_r(1+u')\right]' - \tilde{h}\sigma_t\left(1+\frac{u}{r}\right) &= p(r+u)w', \\ -\tilde{h}(r\sigma_r w')' &= p(r+u)(1+u'), \end{aligned} \tag{2.6}$$

kde

$$\begin{aligned} \sigma_r &= \frac{E\lambda_0^2}{1-\nu^2}(\tilde{\varepsilon}_r + \nu\tilde{\varepsilon}_t) + \sigma_0, \\ \sigma_t &= \frac{E\lambda_0^2}{1-\nu^2}(\tilde{\varepsilon}_t + \nu\tilde{\varepsilon}_r) + \sigma_0, \end{aligned} \tag{2.7}$$

E značí Youngův modul pružnosti, ν Poissonův poměr, $\sigma_0 = \frac{E}{1-\nu}\varepsilon_0 = \frac{E}{2(1-\nu)}(\lambda_0^2-1)$ představuje počáteční předpětí a $\varepsilon_0 = \frac{1}{2}(\lambda_0^2-1)$ počáteční přetvoření a dále

$$\tilde{\varepsilon}_r = u' + \tfrac{1}{2}\left(u'^2 + w'^2\right), \quad \tilde{\varepsilon}_t = \frac{u}{r} + \frac{u^2}{2r^2} \tag{2.8}$$

jsou přetvoření s uvážením předpjatého stavu jako referenčního (vhodný pro výpočet v rámci optických simulací, kdy známe rozměry vstupní pupily – ukotvení membrány v předpjatém stavu). Dále jsou v práci [3] definovány a diskutovány hraniční podmínky řešení, musí platit

$$w(a) = 0\,, \quad u(0) = 0\,, \quad u(a) = 0\,, \quad w'(0) = 0\,. \tag{2.9}$$

Jinou z možností je vyšetření tvaru membrány pomocí stavu ekvilibria, tj. minimalizací potenciální energie systému. Celková potenciální energie E_p uvažovaného systému bude dána vztahem [3]

$$E_p = E_{\text{int}} + E_{\text{ext}}\,, \tag{2.10}$$

tedy jako součet vnitřní energie přetvoření E_{int} a energie vnějších sil E_{ext}, kde

$$E_{\text{int}} = \frac{\pi \tilde{h} E a^2}{1-\nu^2}\left[\lambda_0^2 \int_0^1 (\tilde{\varepsilon}_r^2 + 2\nu\tilde{\varepsilon}_r\tilde{\varepsilon}_t + \tilde{\varepsilon}_t^2)\,\rho\,\mathrm{d}\rho \right. \tag{2.11}$$
$$\left. + 2(1+\nu)\varepsilon_0 \int_0^1 (\tilde{\varepsilon}_r + \tilde{\varepsilon}_t)\,\rho\,\mathrm{d}\rho + \frac{(1+\nu)\varepsilon_0^2}{\lambda_0^2}\right],$$

$$E_{\text{ext}} = -2\pi p a^3 \int_0^1 \tilde{w}(\rho+\tilde{u})\left(1 + \frac{\mathrm{d}\tilde{u}}{\mathrm{d}\rho}\right)\mathrm{d}\rho\,. \tag{2.12}$$

Cílem matematických modelací je nalézt tvar funkcí u a w, známe-li počáteční předpětí σ_0, resp. protažení λ_0, a rozměry předpjaté membrány $\tilde{h}$ a $\tilde{a}$.

Jako jeden z vhodných způsobů se jeví použití metody řad (výše uvedené diferenciální rovnice (2.6) charakterizující vztah mezi výchylkami nemají analytické řešení, a tak je třeba volit řešení numerické).

Očekáváme-li regulární řešení zmíněného problému, využijeme polynomickou aproximaci funkcí výchylek. Dále je vhodné vyjádřit úlohu v bezrozměrném tvaru tak, že je oblast řešení transformována na interval $[0,1]$. Dostáváme tak problém pro bezrozměrné výchylky $\tilde{u} = u/a$ a $\tilde{w} = w/a$ závislé na bezrozměrné souřadnici $\rho = r/a$. Výchylky poté lze aproximovat polynomy

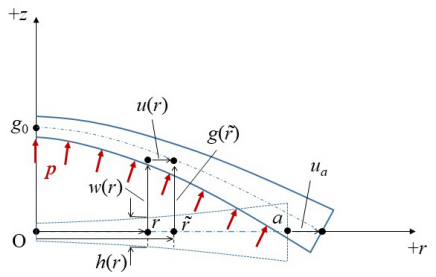
$$\tilde{w}(\rho) = \sum_{i=1}^{N} b_i(1-\rho^{2i})\,, \quad \tilde{u}(\rho) = \sum_{j=1}^{M} c_j(\rho - \rho^{2j+1})\,, \tag{2.13}$$

kde b_i $(i = 1,2,\ldots,N)$ a c_j $(j = 1,2,\ldots,M)$ jsou koeficienty řad. Všimněme si, že řady (2.13) automaticky splňují hraniční podmínky (2.9).

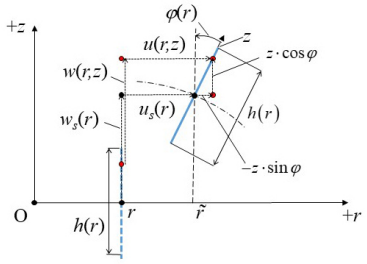
S vyžitím zmíněných řad (2.13) dále dostáváme vztahy pro přetvoření, platí

$$\tilde{\varepsilon}_r = \frac{\mathrm{d}\tilde{u}}{\mathrm{d}\rho} + \frac{1}{2}\left[\left(\frac{\mathrm{d}\tilde{u}}{\mathrm{d}\rho}\right)^2 + \left(\frac{\mathrm{d}\tilde{w}}{\mathrm{d}\rho}\right)^2\right] \tag{2.14}$$
$$= \sum_{j=1}^{M} c_j\left[1-(2j+1)\rho^{2j}\right] + \frac{1}{2}\left\{\sum_{j=1}^{M} c_j\left[1-(2j+1)\rho^{2j}\right]\right\}^2$$
$$+ 2\left(\sum_{i=1}^{N} i\,b_i\rho^{2i-1}\right)^2,$$

$$\tilde{\varepsilon}_t = \frac{\tilde{u}}{\rho} + \frac{\tilde{u}^2}{2\rho^2} \tag{2.15}$$
$$= \sum_{j=1}^{M} c_j(1-\rho^{2j}) + \frac{1}{2}\left[\sum_{j=1}^{M} c_j(1-\rho^{2j})\right]^2.$$



(a) Osově symetrická membrána proměnné tloušťky zatížená rovnoměrným tlakem



(b) Průřez membránou při vychýlení a rotaci

Obrázek 2.8: Deformace osově symetrické membrány proměnné tloušťky [1]

Po dosazení do (2.10) a (2.11) snadno vyjádříme úlohu jako optimalizační problém hledání koeficientů b_i a c_j řad (2.13). Tím je problematika určení tvaru deformované membrány jednoznačně formulována.

Výše představený postup výpočtu deformace membrány určí tvar střednicové plochy. Za předpokladu, že membrána je velmi tenká, bude rozdíl oproti vnějším plochám velmi malý a pro většinu praktických použití zanedbatelný. Nicméně pro velmi přesné optické aplikace je třeba znát tvar deformovaných vnějších ploch membrány. Tato problematika je velmi podrobně řešena v práci [1], na které se autor podílel. Článek prezentuje metodiku výpočtu proměnné tloušťky membrány takové, aby byl docílen požadovaný tvar její vnější nebo vnitřní plochy.

Postup výpočtu vychází z rovnic pro výpočet proměnné tloušťky membrány, jestliže zadáme požadovaný tvar střednicové plochy po deformaci. Označíme-li $g(\tilde{r})$ jako funkci reprezentující předepsaný tvar střednicové plochy deformované membrány a $\tilde{r} = r + u(r)$ jako deformovanou radiální souřadnici (viz obr. 2.8a), poté lze nalézt optimální profil tloušťky membrány jako funkci $h(r)$ pomocí rovnic [52]

$$A(u,r)\,\eta^2 + B(u,r)\,\eta + C(u,r) = 0 \tag{2.16}$$

a

$$h = \frac{p}{E}\left[-\frac{(r+u)^2}{r\,g'\sqrt{\eta}}\right]\frac{1}{\left(1+g'^2\right)\eta - 1}\left(1+\frac{\nu\eta}{\alpha}\right), \tag{2.17}$$

kde

$$\begin{aligned}
\eta &= \left(1+u'\right)^2, \\
A(u,r) &= 1+g'^2, \\
B(u,r) &= \frac{2\nu u}{r} + \frac{\nu u^2}{r^2} - 1 + \nu\left(1-g'^2\right)\alpha(u,r), \\
C(u,r) &= \left(\frac{2u}{r} + \frac{u^2}{r^2} - \nu\right)\alpha(u,r), \\
\alpha(u,r) &= \frac{\left(1+\frac{u}{r}\right)^2 g'}{(r+u)\,g'' - 2g' - 2g'^3}.
\end{aligned}$$

Poznamenejme, že první a druhá derivace funkce g v uvedených rovnicích je uvažována vzhledem k deformované radiální souřadnici $\tilde{r}$ (na rozdíl od funkcí $u(r)$ a $w(r)$, které jsou

vzhledem k souřadnici r), a tudíž platí: $g' = \frac{dg(\tilde{r})}{d\tilde{r}}$ a $g'' = \frac{d^2g(\tilde{r})}{d\tilde{r}^2}$. Dále pro řešení výše uvedené rovnice je třeba uvažovat okrajové podmínky

$$u(0) = 0\,, \quad u(a) = u_a\,, \tag{2.18}$$

kde u_a představuje počáteční protažení membrány (vzniklé např. při konstrukci čočky). Rozsáhlejší analýza ukáže, že první z podmínek je splněna vždy, a tedy pro řešení výše uvedené rovnice proměnné tloušťky (nelineární diferenciální rovnice prvního řádu proměnné $u(r)$) postačuje okrajová podmínka druhá. Podrobnější rozbor rovnic může čtenář nalézt v práci [1].

Jak už bylo zmíněno výše, pro optický návrh je nezbytné znát tvar vnějších ploch membrány, resp. je nutné navrhnout proměnnou tloušťku membrány takovým způsobem, aby se do předepsané formy deformovala vnější nebo vnitřní plocha membrány.

Uvážíme-li situaci na obr. 2.8b, poté lze radiální souřadnice vnitřní a vnější plochy membrány $\tilde{r}_{out}$ a $\tilde{r}_{in}$ vyjádřit vztahy

$$\tilde{r}_{out} = \tilde{r} - \frac{1}{2}\tilde{h}\,(r)\sin\left(\varphi\right), \qquad \tilde{r}_{in} = \tilde{r} + \frac{1}{2}\tilde{h}\,(r)\sin\left(\varphi\right) \tag{2.19}$$

a odpovídající vertikální souřadnice g_{out} a g_{in} jako

$$g_{out} = g\,(\tilde{r}) + \frac{1}{2}\tilde{h}\,(r)\cos\left(\varphi\right), \qquad g_{in} = g\,(\tilde{r}) - \frac{1}{2}\tilde{h}\,(r)\cos\left(\varphi\right)\,, \tag{2.20}$$

kde φ je úhel natočení řezu membrány, který můžeme dále určit jako: $\varphi = \arctan g'$, a tudíž platí: $\sin\varphi = \frac{g'}{\sqrt{g'^2+1}}$ a $\cos\varphi = \frac{1}{\sqrt{g'^2+1}}$. V tomto případě $\tilde{h}(r)$ značí tloušťku membrány, pro kterou platí [53]

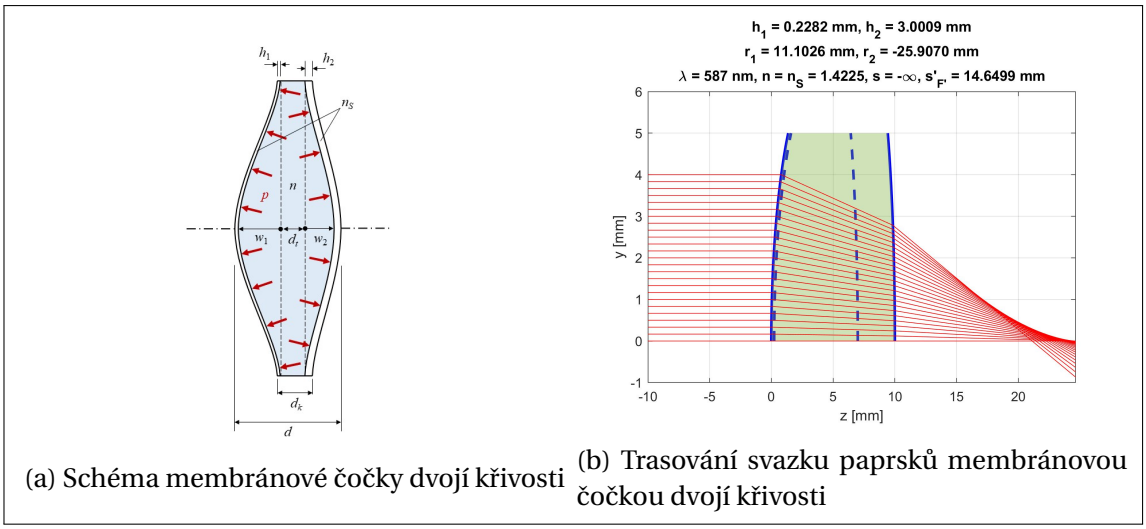
$$\tilde{h}\,(r) = h\,(r)\,\sqrt{2\varepsilon_z + 1} = h\,(r)\,\sqrt{\frac{2\nu\,(\varepsilon_r + \varepsilon_t)}{\nu - 1} + 1}\,, \tag{2.21}$$

kde $\varepsilon_r = u' + \frac{1}{2}\left(u'^2 + w'^2\right)$ a $\varepsilon_t = \frac{u}{r} + \frac{u^2}{2r^2}$ jsou radiální a tangenciální přetvoření.

Popsané rovnice nemají analytické řešení, stejně jako tomu bylo při hledání deformovaného tvaru membrány, a je třeba úlohu řešit numericky, opět například pomocí metody řad. Na vstupu výpočtu je zadán tvar vnitřní nebo vnější plochy a cílem úlohy je nalézt takovou funkci proměnné tloušťky, pro kterou bude tvar vypočtené vnitřní nebo vnější plochy po zatížení daným tlakem odpovídat ploše předepsané. Podrobnější pojednání o řešení a optické analýze může čtenář nalézt v pracích [1–3, 52, 53].

2.3 Návrh kombinovaných membránových čoček

V kapitole 2.2 byl shrnut postup výpočtu deformace membrán kapalinových čoček. S tímto nástrojem lze modelovat a analyzovat membránové čočky a použít vypočtené tvary optických rozhraní k návrhu unikátních optických systémů, které kombinují jak fixní prvky tak prvky aktivní – kapalinové membránové čočky. Výhodou kombinovaných systémů je to, že umožňují korigovat optické aberace v mnohem širší oblasti obrazového prostoru, než je tomu u samotné jednoduché čočky. Ta může totiž minimalizovat aberace jen pro velmi omezenou oblast (odpovídající jedné poloze předmětu a obrazu, resp. ohniskové vzdálenosti).



(a) Schéma membránové čočky dvojí křivosti

(b) Trasování svazku paprsků membránovou čočkou dvojí křivosti

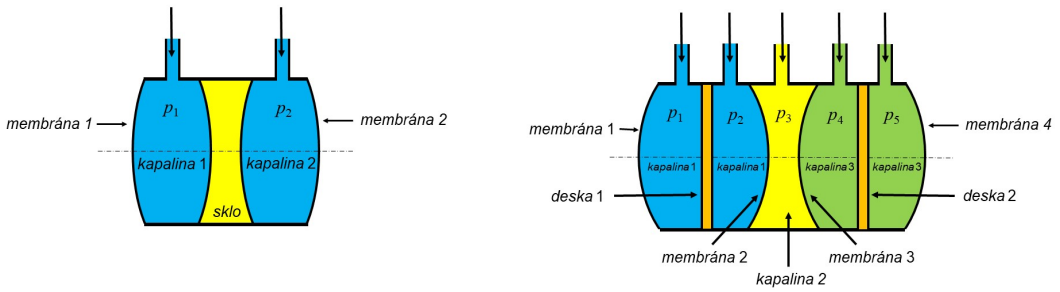
Obrázek 2.9: Membránová čočka dvojí křivosti [6]

Návrhu nominálních parametrů vybraných kombinovaných systémů se autor této práce věnoval a podílel se na publikaci výsledků v mezinárodních časopisech [4–6]. Níže budou představeny a shrnuty vybrané výsledky.

Návrh parametrů jakéhokoli optického systému je založen na modelování optických aberací, zpravidla paprskových nebo vlnových. Známe-li aberace soustavy, poté lze parametry jednotlivých členů systému optimalizovat tak (hledat jejich numerické hodnoty), aby vybrané hodnoty aberací byly minimalizovány (např. poloměr rozptylového kroužku, průběh paprskové aberace apod.). Jelikož je podrobný rozbor a představení vztahů, které se v průběhu optického návrhu používají, nad rámec této práce, čtenář může najít informace například v literatuře [40–43, 48, 54–56].

Jako první z analýz kombinovaných membránových čoček, které se autor věnoval, představme membránovou čočku dvojí křivosti [6]. Její schéma je zobrazeno na obr. 2.9a. Čočka se skládá ze dvou osově symetrických membrán různé tloušťky h_1 a h_2, které budou mít při stejném působícím tlaku p kapaliny uvnitř čočky různou osovou křivost. Změnou objemu kapaliny je možné měnit tlak, a tím pádem i optickou mohutnost čočky. V nominálním stavu, kdy tlak nedeformuje membrány, je osová tloušťka čočky označena d_k. Zvýšením tlaku dojde k její změně na d. Dále na obr. 2.9a značí n_S index lomu materiálu membrány, n index lomu kapaliny, w_1 a w_2 osovou maximální deformaci membrány a d_t je vnitřní osová tloušťka čočky při nominálním tlaku (nulové deformaci membrán). Výpočet deformace membrány lze provést řešením rovnic představených v kapitole 2.2.

Známe-li tvar membrán, lze snadno určit osové křivosti jednotlivých ploch, které oddělují jednotlivé optické materiály. To je vstupem pro výpočet např. Seidelových aberací třetího řádu [40–43, 48, 54–56]. Ve speciálním případě, bude-li index lomu materiálu membrán stejný jako index lomu kapaliny, čehož lze výrobně dosáhnout, bude prvotní analýza aberací odpovídat hodnotám tlusté čočky. Následně lze optimalizovat geometrické parametry čočky (tloušťky membrán, jejich počáteční nominální vzdálenost a působící tlak) tak, aby byly aberace minimalizovány. Další z možností je provést trasování paprsků optickou soustavou čočky (viz obr. 2.9b) a analýzu paprskových aberací společně s následnou optimalizací. Podrobnosti analýzy může čtenář nalézt v práci [6]. Je zde podrobně vysvětlen postup výpočtu nominálních parametrů čočky s minimalizovanou sférickou aberací třetího řádu pro předmět v nekonečnu pro různé indexy lomů kapaliny



(a) Zoom systém ze tří členů, kdy dva vnější jsou aktivní a prostřední fixní

(b) Obecný hybridní zoom systém

Obrázek 2.10: Schémata vybraných kombinovaných zoom systémů [5]

a působící tlaky. Dále je např. ukázána analýza parametrů čočky, které zajistí stejnou hodnotu ohniskové vzdálenosti pro různé tloušťky membrán.

Další z užitečných aplikací aktivních čoček je konstrukce hybridních zoom systémů. Cílem práce [5], na které autor spolupracoval, bylo vytvořit metodiku počátečního návrhu zoom systému, který bude složen z tenkých optických členů (tenkých membránových čoček) s proměnnou ohniskovou vzdáleností.

V případě klasických zoom systémů se jednotlivé optické komponenty během změny ohniskové vzdálenosti systému nebo jeho zvětšení pohybují podél optické osy. Vnitřní struktura jednotlivých parametrů (poloměry křivosti ploch čoček, vrcholové vzdálenosti a indexy lomu) je fixní. Základní rozdíl těchto zoom systémů oproti těm s členy proměnné ohniskové vzdálenosti je ten, že aktivní prvky zůstávají ve stejné pozici a jejich vnitřní struktura (parametry) se mění. V případě klasických systémů nemůže být aberace nulová pro celý rozsah ohniskových vzdáleností, může být pouze minimalizována. V případě hybridních zoom systémů s aktivními prvky lze dosáhnout situace, kdy budou Seidelovy koeficienty (aberace) splňovat některou z následujících podmínek:

- systém bude mít požadované hodnoty aberací pro celý rozsah ohniskových vzdáleností zoom systému,
- aberace se budou měnit pouze minimálně v rámci rozsahu ohniskových vzdáleností zoom systému.

Hybridní zoom systémy lze konstruovat různými způsoby. Vybrané jsou ukázány na obr. 2.10. Vzhledem ke konstrukční jednoduchosti se jeví jako vhodné takové soustavy, které jsou složeny ze dvou vnějších aktivních čoček s fixní čočkou mezi nimi (viz obr. 2.10a). U těchto konstrukcí ale není možné kompenzovat aberace pro všechny stavy systému (ohniskové vzdálenosti nebo zvětšení), jelikož jsou zde přítomny pouze dva variabilní parametry – první a poslední poloměr křivosti. Aberace tak mohou být korigovány pouze pro speciální případy, v ostatních se budou měnit. Dále je nutné zajistit fixní vzdálenost obrazové roviny od posledního komponentu zoom systému pro všechny hodnoty ohniskových vzdáleností (pro předmět v nekonečnu) nebo příčné zvětšení (pro předmět v konečné vzdálenosti). Obecnější, výhodnější, ale komplikovanější může být případ schematicky ukázaný na obr. 2.10b. První a třetí komponent (zobrazeny

modrou a zelenou barvou) jsou tvořeny jako dvě plankonvexní čočky oddělené tenkou deskou z transparentního materiálu (sklo, Sylgard apod.) Druhý komponent je tvořen druhou a třetí membránou a je použit jako kompenzační rezervoár. Poloměry křivosti jednotlivých membrán se budou měnit se změnou jednotlivých tlaků. Takovýto systém může kompenzovat aberace pro více případů než předchozí.

Představený postup v práci [5] umožňuje navrhnout jak vnější parametry systému (ohniskové vzdálenosti jednotlivých členů a jejich osové rozestupy), tak vnitřní parametry jednotlivých členů (poloměry křivosti, tloušťky a indexy lomu). Takto získané startovací parametry mohou být použity pro následný optimální optický návrh. Pro podrobnější studium a příklady odkážeme čtenáře na zmíněnou publikaci.

Jako poslední část této kapitoly zmiňme výstup práce autora publikovaný v článku [4], kde je cílem zevrubná analýza hybridní soustavy (čočky) z jedné fixní a jedné nebo dvou aktivních čoček. Je zde velmi podrobně rozebrán popis aberací těchto systémů, který je dále využit pro nominální optický návrh jako startovní bod.

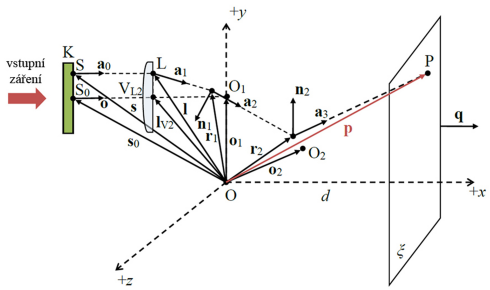
2.4 Využití aktivních prvků v optických skenerech

Jak již bylo zmíněno, aktivní optické prvky mohou nalézt a nacházejí široké uplatnění v radě praktických aplikacích. Jednou z nich je i oblast adaptivního skenování. Optické skenery jsou zařízení pro bezkontaktní, velmi rychlé a poměrně přesné měření prostorové polohy bodů na předmětu zájmu nebo k modifikaci vlastností jeho povrchu. Rozdělit můžeme skenery do tří skupin:

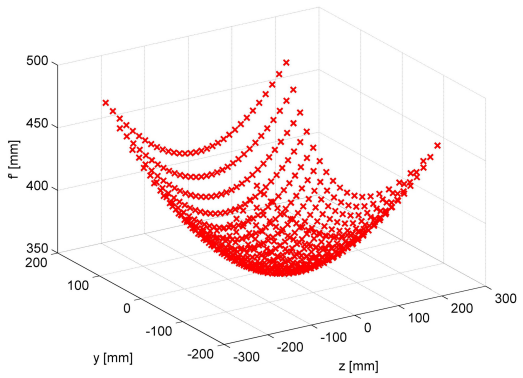
- zaměřování, prověřování nebo kontrola stávajících prvků,
- projekční technologie,
- laserové technologie při výrobě.

Zejména v projekčních technologiích a při výrobě (gravírování, řezání, svařování apod.) je žádoucí, aby stopa svazku byla zaostřena v rovině předmětu (na povrch předmětu). To je s použitím klasických optických prvků poměrně komplikovaná záležitost, zejména jedná-li se o vysokofrekvenční aplikace, kdy se stopa pohybuje velmi rychle. Zde nachází výhodné uplatnění právě aktivní optické prvky, které mohou bez změny polohy a velmi rychle modifikovat parametry zobrazení projekční optiky.

V rámci práce [7] autor studoval použití aktivní plankonvexní čočky v dvouzrcadlovém optickém skeneru, kde podrobně ukázal vztahy pro trasování paprsku optickou rozmítací soustavou a také demonstroval výpočet ohniskové vzdálenosti aktivní čočky takové, která zajistí zaostření systému do požadované polohy na předmětu. Principiální schéma trasování je ukázáno na obr. 2.11a, závislost hodnot ohniskové vzdálenosti pro zaostření svazku ve vybraných bodech detekční roviny je poté zobrazena na obr. 2.11b. Podrobné odvození vztahů a jejich souhrnný popis je nad rámec této práce, a proto pro více informací odkážeme čtenáře na zmíněnou publikaci.



(a) Principiální schéma trasování optického svazku soustavou dvou-zrcadlového skeneru a plankonvexní aktivní čočkou



(b) Závislost hodnot ohniskové vzdálenosti pro zaostření svazku ve vybraných bodech detekční roviny

Obrázek 2.11: Použití aktivních čoček v optickém skenování [7]

3 Vybrané oblasti optického zobrazení

V této kapitole budou shrnuty a představeny vybrané výstupy publikací [8–16], na kterých se autor podílel v rámci analýzy různých témat optického zobrazení. Kopie těchto publikací jsou uvedeny v Příloze A.

První část se zabývá problémem skalární difrakce a jejího řešení na mřížce, kruhové pupile a mezikruží a analýzou rozptylové funkce bodu a osové rozptylové funkce bodu [8–12]. Následuje část věnovaná aberaci optických soustav, a to zejména sférické aberaci a jejímu vlivu na hloubku ostrosti a návrhu systému s jednou nebo dvěma asférickými plochami s korigovanou sférickou aberací [13, 14]. Kapitola je zakončena tématem teoretické analýzy návrhu základních parametrů dvoučlenných optických systémů [15].

3.1 Skalární teorie difrakce

3.1.1 Skalární teorie difrakce a její aproximativní řešení

V optické teorii je známo, že komplexní amplituda pole v libovolném bodě P prostoru za rovinnou aperturou (nebo řadou apertur) může být vypočtena pomocí vztahu [40, 41, 57–60]

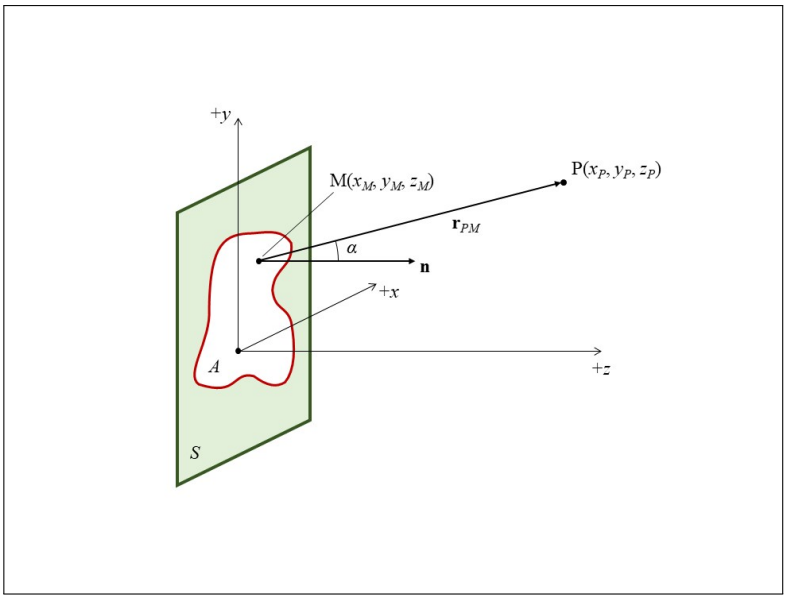
$$U(P) = -\frac{\mathrm{i}}{\lambda} \iint_A U(M) \frac{\exp(\mathrm{i}kr_{PM})}{r_{PM}} \cos(\mathbf{n}, \mathbf{r}_{PM}) \, \mathrm{d}A, \tag{3.1}$$

kde je integrace prováděna přes plochu A apertury, M značí bod na této ploše apertury, r_{PM} značí vzdálenost mezi body P a M a $\cos(\mathbf{n}, \mathbf{r}_{PM})$ značí cosinus úhlu, který svírá vnitřní jednotkový normálový vektor $\mathbf{n}$ k ploše S se směrem vektoru $\mathbf{r}_{PM}$. Uvažovaná situace je ukázána na obr. 3.1. Je běžnou praxí, že směr šíření pole je volen ve směru osy $+z$, tj. platí $\mathbf{n} = (0, 0, 1)$.

Rovnice (3.1) je tzv. Sommerfeldovo řešení difrakční úlohy (Sommerfeldův difrakční integrál nebo také Rayleigh-Sommerfledovo řešení druhého druhu). Pomocí tohoto vztahu můžeme vypočítat vlastnosti pole $U(P)$ (komplexní amplitudu) v bodě P ohraničeném plochou S, jestliže je známa komplexní amplituda $U(M)$ na oblasti apertury. Intenzitu pole $I(P)$ v bodě P poté můžeme určit ze vztahu

$$I(P) = |U(P)|^2 . \tag{3.2}$$

Vztah (3.1) obecně nemá analytické řešení, a proto je vhodné (možné) použít aproximativní přístupy, které v řadě praktických případů analytické řešení už poskytují, nebo alespoň numerické řešení výrazně zjednodušují.



Obrázek 3.1: Difrakce na rovinné apertuře

Vzdálenost r_{PM} mezi body P a M je dána vztahem (viz obr. 3.1, $z_M = 0$)

$$\begin{aligned} r_{PM} &= \sqrt{(x_P - x_M)^2 + (y_P - y_M)^2 + z_P^2} \\ &\approx z_P + \frac{(x_P - x_M)^2 + (y_P - y_M)^2}{2z_P}, \end{aligned} \tag{3.3}$$

kde je použit Taylorův rozvoj [37–39] se zanedbáním vyšších řádů. V případě, že předpokládáme malé difrakční úhly (budeme vyšetřovat pole v blízkosti optické osy), tj. můžeme s dostatečnou přesností položit $\cos(\mathbf{n}, \mathbf{r}_{PM}) \approx 1$, poté lze rovnici (3.1) vyjádřit jako

$$\begin{aligned} U(P) &= C \iint_A U(M) \exp\left\{\frac{\mathrm{i}k}{2z_P}\left[(x_P - x_M)^2 + (y_P - y_M)^2\right]\right\} \mathrm{d}x_M \mathrm{d}y_M, \\ C &= -\frac{\mathrm{i}}{\lambda}\frac{\exp(\mathrm{i}kz_P)}{z_P}, \end{aligned} \tag{3.4}$$

kde r_{PM} bylo v exponentu (3.1) nahrazeno rovnicí (3.3) a jmenovatel je položen $r_{PM} = z_P$. Rovnice (3.4) představuje tzv. Fresnelovu aproximaci Sommerfeldova difrakčního integrálu [40, 41, 57–60].

Dalšího zjednodušení dosáhneme v případě, že difrakce nastává na apertuře, jejíž rozměry jsou výrazně menší než vzdálenost z_P. Poté můžeme předpokládat

$$\exp\left[\frac{\mathrm{i}k}{2z_P}(x_M^2 + y_M^2)\right] \approx 1. \tag{3.5}$$

Druhá možnost obdobného zjednodušení nastane v případě difrakce konvergentní sférické (nebo alespoň přibližně sférické) vlny se středem v bodě P nebo v jeho blízkosti, kdy platí

$$U(M) = T(M) \exp\left[-\frac{\mathrm{i}k}{2z_P}(x_M^2 + y_M^2)\right], \tag{3.6}$$

kde funkce $T(M)$ charakterizuje vlastnosti této vlny v rovině apertury.

Vztahy (3.5) a (3.6) vedou po dosazení do (3.4) a snadné úpravě na tzv. Fraunhoferovu aproximaci difrakční úlohy [40, 41, 57–60]. Označíme-li

$$u = x_P/z_P \quad \text{a} \quad v = y_P/z_P\,, \tag{3.7}$$

komplexní amplitudu v bodě P můžeme vyjádřit pomocí vztahů

$$U(P) = C \iint_A U(M) \exp\left[-\mathrm{i}k(ux_M + vy_M)\right] \mathrm{d}x_M \mathrm{d}y_M\,, \tag{3.8}$$

$$U(P) = C \iint_A T(M) \exp\left[-\mathrm{i}k(ux_M + vy_M)\right] \mathrm{d}x_M \mathrm{d}y_M\,. \tag{3.9}$$

Ze vztahů (3.8) a (3.9) je patrné, že komplexní amplituda $U(P)$ je úměrná Fourierově transformaci [37–39] pole v rovině apertury. Jelikož je numerický výpočet Fourierovy transformace velmi dobře realizován na současných počítačích, je toto jednoduché řešení pro analýzy difrakčních jevů často používáno.

3.1.2 Difrakce na mřížce

Pomocí vztahů ukázaných v předchozí kapitole lze analyzovat řadu difrakčních problémů. Jedním z nich je difrakce na mřížce, neboli šíření pole za stínítkem s periodickou soustavou otvorů. Tímto tématem se autor zabýval a publikoval analýzu difrakce na dokonalé a nedokonalé mřížce, kdy jsou její hrany popsány harmonickými funkcemi, v práci [8]. V publikaci [9] je poté analyzován vliv konečných rozměrů mřížky na tzv. Talbotův jev – samozobrazování periodické struktury za mřížkou. Vztahy odvozené v těchto pracích jsou velmi užitečným a přitom jednoduchým nástrojem pro analýzu zmíněného difrakčního problému.

Pro ilustraci ukažme postup výpočtu stavu pole za nedokonalou difrakční mřížkou s použitím Fraunhoferovy aproximace. Situaci lze analyzovat způsobem, kdy je každá jednotlivá štěrbina mřížky popsána individuální difrakcí na apertuře. Pole za takovou nedokonalou mřížkou poté bude charakterizováno superpozicí příspěvků jednotlivých subapertur. Obecně lze funkci propustnosti mřížky popsat vztahem [8]

$$f(\xi,\eta) = A \sum_{n=-\infty}^{\infty} \mathrm{rect}\left[\frac{\xi - \xi_{n,0}(\eta)}{b_n(\eta)}\right], \tag{3.10}$$

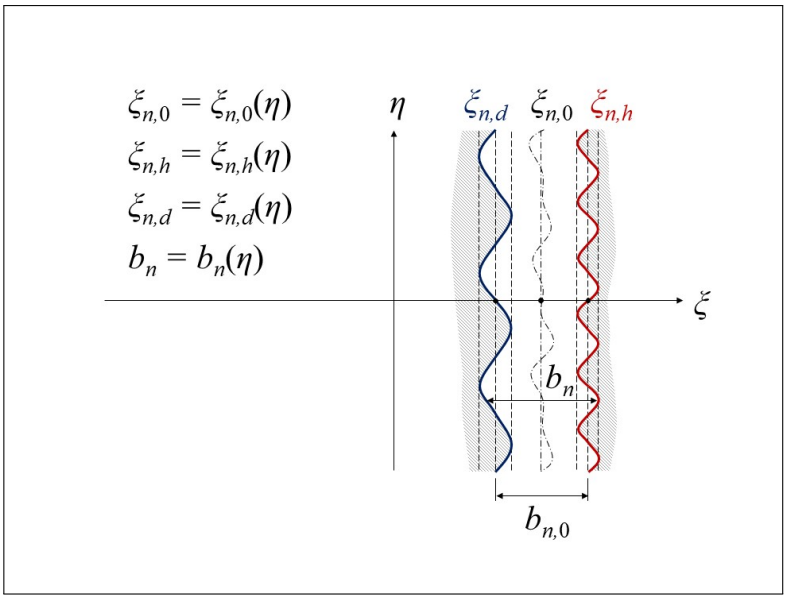
kde A je konstanta, $\xi_{n,0}(\eta)$ označuje funkci centrální linie n-té štěrbiny (subapertury) a $b_n(\eta)$ je funkce popisující její šířku. Obr. 3.2 zobrazuje uvažovanou situaci.

Dále lze psát

$$\xi_{n,0}(\eta) = \frac{1}{2}\left[\xi_{n,d}(\eta) + \xi_{n,h}(\eta)\right], \quad b_n(\eta) = \xi_{n,h}(\eta) - \xi_{n,d}(\eta)\,, \tag{3.11}$$

kde $\xi_{n,d}(\eta)$ a $\xi_{n,h}(\eta)$ jsou funkce jednotlivých hran n-té subapertury (štěrbiny). Tyto hranové funkce mohou být vyjádřeny řadou způsobů (aproximací), například rozvojem v mocninnou řadu, Legendreovy polynomy, Fourierovu trigonometrickou řadu apod. [37].

Výše uvedený obecný formalismus zahrnuje různé typy mřížek, které mohou mít štěrbiny různých tvarů (popsaných individuálními hranovými funkcemi $\xi_{n,d}(\eta)$ a $\xi_{n,h}(\eta)$). Centrální linie štěrbin mřížek (rovnice (3.11)) také nemusí být vzájemně rovnoběžné, jelikož každá z hran může mít různý trend.



Obrázek 3.2: Schéma n-té štěrbiny (subapertury) nedokonalé mřížky

Zmiňme dále podmínky existence mřížky. Ta může být fyzicky realizovatelná (vyrobena), pokud platí $\max[\xi_{n-1,h}(\eta)] < \min[\xi_{n,d}(\eta)]$ pro každé η.

Předpokládejme nyní, že mřížka má N štěrbin, tj. $n \in [1, N]$. Poté pro komplexní amplitudu $U(x,y,z)$ a intenzitu $I(x,y,z)$ v rovině z za mřížkou s použitím principu superpozice dostáváme

$$U(x,y,z) = \sum_{n=1}^{N} U_{S,n}\,, \quad I(x,y,z) = \left| \sum_{n=1}^{N} U_{S,n} \right|^2, \tag{3.12}$$

kde $U_{S,n}$ označuje příspěvek n-té subapertury mřížky. Uvážíme-li dále pro ilustraci, že na mřížku dopadá rovinná vlna, lze v případě Fraunhoferovy aproximace Sommerfeldova difrakčního integrálu, rovnice (3.8), pro mřížky výšky $2d$ psát

$$U_{S,n} = CK \int_{-d}^{d} \int_{\xi_{n,d}(y_M)}^{\xi_{n,h}(y_M)} \exp\left[-\mathrm{i}k(ux_M + vy_M)\right] \mathrm{d}x_M \mathrm{d}y_M\,, \tag{3.13}$$

kde $K = konst.$, $u = x/z$, $v = y/z$, $C = -\frac{\mathrm{i}}{\lambda}\frac{\exp(\mathrm{i}kz)}{z}$, $k = \frac{2\pi}{\lambda}$.

Uvažujme dále hrany štěrbin mřížky charakterizovány pomocí harmonických funkcí způsobem

$$\begin{aligned} \xi_{n,d}(\eta) &= \xi_{n,0}(0) - \frac{b_{n,0}}{2} - A_{n,1}\sin(\Omega_{n,1}\eta + \varphi_{n,1})\,, \\ \xi_{n,h}(\eta) &= \xi_{n,0}(0) + \frac{b_{n,0}}{2} + A_{n,2}\sin(\Omega_{n,2}\eta + \varphi_{n,2})\,, \end{aligned} \tag{3.14}$$

kde $\xi_{n,0}(0)$ a $b_{n,0}$ jsou pozice centrálních linií a šířky subapertury pro souřadnici $\eta = 0$, $A_{n,1}$ a $A_{n,2}$ jsou amplitudy hranových funkcí, $\Omega_{n,1}$ a $\Omega_{n,2}$ jsou úhlové frekvence hranových funkcí a $\varphi_{n,1}$ a $\varphi_{n,2}$ jsou jejich počáteční fázové posuny. Označíme-li dále $U_{S,n} =$

$CKU_{S,n,xy}$, poté po provedení vhodných zjednodušení dostáváme

$$U_{S,n,xy} = \mathrm{i}D_{S,n,2}\int_{-d}^{d}\left\{\exp(-\mathrm{i}kvy_M)\exp[-\mathrm{i}\beta_{n,2}\sin(\Omega_{n,2}y_M+\varphi_{n,2})]\right\}\,\mathrm{d}y_M \tag{3.15}$$
$$-\mathrm{i}D_{S,n,1}\int_{-d}^{d}\left\{\exp(-\mathrm{i}kvy_M)\exp[\mathrm{i}\beta_{n,1}\sin(\Omega_{n,1}y_M+\varphi_{n,1})]\right\}\,\mathrm{d}y_M\,,$$

kde

$$\beta_{n,2} = A_{n,2}ku\,,\quad \beta_{n,1} = A_{n,1}ku\,, \tag{3.16}$$
$$D_{S,n,1} = \frac{1}{ku}\exp\left[-\mathrm{i}ku\left(\xi_{n,0}(0)-\frac{b_{n,0}}{2}\right)\right]\,,$$
$$D_{S,n,2} = \frac{1}{ku}\exp\left[-\mathrm{i}ku\left(\xi_{n,0}(0)+\frac{b_{n,0}}{2}\right)\right]\,.$$

Pro limitní případ $u = 0$ následně platí

$$U_{S,n,xy}\big|_{u=0} = \int_{-d}^{d}\exp(-\mathrm{i}kvy_M)\left[b_0 + A_{n,1}\sin(\Omega_{n,1}y_M+\varphi_{n,1}) + A_{n,2}\sin(\Omega_{n,2}y_M+\varphi_{n,2})\right]\,\mathrm{d}y_M\,, \tag{3.17}$$

pro $v = 0$

$$U_{S,n,xy}\big|_{v=0} = \mathrm{i}D_{S,n,2}\int_{-d}^{d}\exp[-\mathrm{i}\beta_{n,2}\sin(\Omega_{n,2}y_M+\varphi_{n,2})]\,\mathrm{d}y_M \tag{3.18}$$
$$-\mathrm{i}D_{S,n,1}\int_{-d}^{d}\exp[\mathrm{i}\beta_{n,1}\sin(\Omega_{n,1}y_M+\varphi_{n,1})]\,\mathrm{d}y_M$$

a pro $u = v = 0$ dostáváme

$$U_{S,n,xy}\big|_{u=v=0} = 2b_{n,0}d + \frac{2A_{n,2}}{\Omega_{n,2}}\sin(\Omega_{n,2}d)\sin\varphi_{n,2} + \frac{2A_{n,1}}{\Omega_{n,1}}\sin(\Omega_{n,1}d)\sin\varphi_{n,1}\,. \tag{3.19}$$

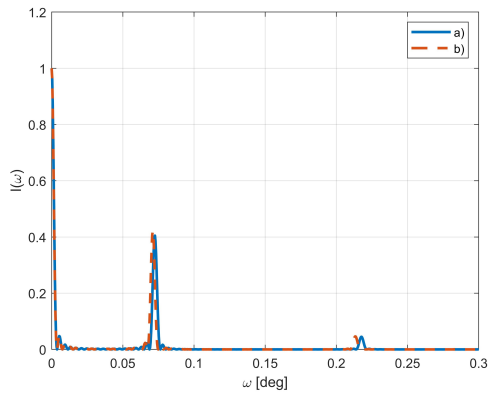
Použitím rovnic (3.13) až (3.19) ve vztahu (3.12) lze vyjádřit stav pole (komplexní amplitudu) v rovině za mřížkou. Pro intenzitu pole poté platí

$$I(x,y,z) = \frac{K^2}{\lambda^2 z^2}\left|\sum_{n=1}^{N}U_{S,n,xy}\right|^2\,. \tag{3.20}$$

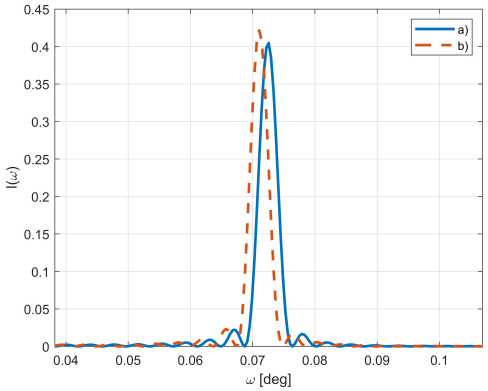
V limitním případě $I(0,0,z)$ dostáváme

$$I(0,0,z) = 4\frac{K^2}{\lambda^2 z^2}\left|\sum_{n=1}^{N}\left[b_{n,0}d + \frac{A_{n,2}}{\Omega_{n,2}}\sin(\Omega_{n,2}d)\sin\varphi_{n,2} + \frac{A_{n,1}}{\Omega_{n,1}}\sin(\Omega_{n,1}d)\sin\varphi_{n,1}\right]\right|^2\,. \tag{3.21}$$

V publikaci [8] autoři dále analyzují a odvozují vztahy pro charakteristiku mřížek s malými defekty jednotlivých vrypů (štěrbin), malými amplitudami hranových funkcí nebo proměnnou periodou mřížky. Detailnější informace může čtenář nalézt v Příloze A.



(a) Řez v rovině $y = 0$



(b) Detail řezu v rovině $y = 0$ v oblasti prvního difrakčního maxima

Obrázek 3.3: Srovnání difrakce na dokonalé a nedokonalé mřížce – normalizovaná intenzita pro a) dokonalou a b) nedokonalou mřížku s proměnnou periodou [8]

Na obr. 3.3a je pro příklad ukázán řez normalizovanou intenzitou při Fraunhoferově difrakci v rovině $y = 0$ pro dokonalou mřížku s periodou p_0 a nedokonalou mřížku s proměnnou periodou, která je charakterizována vztahy

$$\begin{aligned} \xi_{n,0}(0) &= n_p(1 + A_p)p_0\,, \quad n_p \in [-N_p, N_p]\,, \\ b_{n,0} &= b_0\,, \quad A_{n,1} = A_1\,, \quad A_{n,2} = A_2\,, \\ \Omega_{n,1} &= \Omega_1\,, \quad \Omega_{n,2} = \Omega_2\,, \quad \varphi_{n,1} = \varphi_1\,, \quad \varphi_{n,2} = \varphi_2\,, \end{aligned} \tag{3.22}$$

kde $A_p = 0.02$, $p_0 = 0.5$ mm, $N_p = 9$, $b_0 = 0.25$ mm, $A_1 = A_2 = \Omega_1 = \Omega_2 = \varphi_1 = \varphi_2 = 0$. Dále jsou předpokládány rozměry mřížky $c = d = 5$ mm, parametry dopadajícího pole $A = K = 1$ a $\lambda = 633$ nm a rovina detekční roviny je v pozici $z = 150$ mm. Tato mřížka má tedy lineární hrany ale proměnnou periodu, která se lineárně mění se vzdáleností od centra mřížky. Obr. 3.3b následně ukazuje výřez kolem prvního difrakčního maxima. Horizontální osa grafů je uvedena v úhlové souřadnici ω, kdy platí $\tan\omega = \frac{x}{z}$. Z obrázků je zřejmé, že díky nepravidelnosti mřížky dojde jak k posunu maxim difrakčních řádů, tak ke změně intenzity.

Jak již bylo zmíněno v úvodu kapitoly, v publikaci [9], na které se autor této práce podílel, je analyzován vliv konečných rozměrů mřížky na tzv. Talbotův jev. Jestliže pole dopadá na periodickou strukturu (např. právě na difrakční mřížku nebo pole děr v neprůhledném stínítku), poté existují vzdálenosti, ve kterých bude struktura intenzity odpovídat přesně struktuře stínítka (difrakční mřížky). Toto samozobrazení má aplikaci v řadě metrologických aplikacích, interferometrii, mikroelektronice apod.

V publikaci [9] autoři provedli podrobnou analýzu tohoto jevu a zejména poté ukázali jednoduchý vztah pro odhad šířky tzv. přechodové funkce Talbotova jevu, která charakterizuje vliv konečných rozměrů mřížky na vzniklé Talbotovy obrazy (obrazy nebudou ostré, ale dojde k mírnému rozmazání jejich hran). Šířku přechodové funkce Δ (horizontální vzdálenost mezi body na dolní a horní hraně pulzu obrazu normované intenzity, které vytíná přímka procházející inflexním bodem hrany pulzu) lze odhadnout ze vztahu [9]

$$\Delta \approx 0.75\frac{p^2}{c}\,, \tag{3.23}$$

kde p je perioda mřížky a c značí šířku mřížky.

3.1.3 Difrakce na kruhové pupile a mezikruží

Jak je zřejmé v úvodní části této kapitoly, stav pole za optickou soustavou je závislý na tvaru výstupní pupily soustavy, přes jejíž plochu probíhá integrace ve vztahu (3.1), na vlnové délce použitého záření a na vzdálenosti, ve které je stav pole vyšetřován. Obraz bodu je poté charakterizován rozložením intenzity (energie) – tzv. rozptylovou funkcí bodu (PSF z angl. Point Spread Function) [40, 41, 54, 60]. Jestliže je optická soustava rotačně symetrická a bez aberací, maximální intenzita pole v obrazovém prostoru leží v rovině, která je totožná s geometricko-optickým obrazem bodu na předmětu.

V optické praxi jsou nejčastěji analyzovány pupily kruhové. Problematice kruhových pupil a pupil tvaru mezikruží se autor této práce zabýval a podílel se na publikaci výsledků v pracích [10–12]. Následující pasáž představí stěžejní myšlenky a výstupy zmíněných publikací.

Výpočet stavu pole za aperturou je historicky dobře znám a existuje celá řada možností, jak rozptylovou funkci bodu určit (např. numerická integrace, Fourierova transformace a další). V publikaci [10] se autor této práce podílel na formulaci explicitního výpočtu rozptylové funkce bodu v případě optické soustavy s kruhovou aperturou pro zobrazení osového bodu předmětu. Tento přístup rozšířil dosud známé možnosti výpočtů o další alternativu, ve které je integrace nahrazena rozvojem ve vhodnou řadu, a je tedy velmi snadno implementovatelná.

Uvažujme rotačně symetrický optický systém s kruhovou výstupní pupilou, který zobrazuje osový bod předmětu. Vlnoplocha za výstupní pupilou bude v případě systému bez aberací konvergentní sférická se středem v obraze osového bodu. V takovém případě lze položit $U(M) = A(M)\exp(-ikR)/R$, kde $A(M)$ je komplexní amplituda vlnoplochy v bodě M (na sférické vlnoploše) a R značí poloměr vlnoplochy. Nechť jsou dále hodnoty x_P a y_P výrazně menší než poloměr R a můžeme tak předpokládat následující zjednodušení ve vztahu (3.1): $\cos(\mathbf{n}, \mathbf{r}_{PM}) \approx 1$, $\exp(ikr_{MP})/r_{MP} \approx \exp(ikr_{MP})/R$. Komplexní amplitudu v bodě P (pro $U(M) \approx A(\rho)$ a kruhovou pupilu) lze poté psát ve tvaru [40, 41, 54, 60]

$$U(P) = -\mathrm{i}\frac{2\pi a^2}{\lambda R}\left[\frac{\exp(\mathrm{i}k\bar{R})}{R}\right]\int_0^1 A(\rho)J_0\left(\tau\rho\right)\,\rho\,\mathrm{d}\rho\,, \tag{3.24}$$

kde

$$\tau = \frac{2\pi}{\lambda}\left(\frac{a}{R}\right)\sqrt{x_P^2+y_P^2} = \frac{\pi\sqrt{x_P^2+y_P^2}}{\lambda c} = \frac{\pi t}{\lambda c}\,, \tag{3.25}$$

$t = \sqrt{x_P^2+y_P^2}$, $g = \pi t/c$, $\rho = r/a$, $\bar{R} = r_{OP} - R$, r_{OP} je vzdálenost mezi středem výstupní pupily a bodem P, $a = r_{\max}$ je maximální hodnota r, J_0 je Besselova funkce prvního druhu a $c = R/(2a)$ je clonové číslo optického systému v obrazovém prostoru.

V případě, kdy je optický systém zatížený aberacemi, můžeme vztah (3.24) vyjádřit jako [10]

$$U(P) = L\int_0^1 P(\rho)J_0\left(\tau\rho\right)\,\rho\,\mathrm{d}\rho\,, \tag{3.26}$$

kde

$$L = -\mathrm{i}\frac{2\pi a^2}{\lambda R}\left[\frac{\exp(\mathrm{i}k\bar{R})}{R}\right] = -\mathrm{i}\left(\frac{\pi}{2\lambda c^2}\right)\exp(\mathrm{i}k\bar{R})\,, \tag{3.27}$$

$$P(\rho) = T(\rho)\exp\left[\mathrm{i}kW(\rho)\right] = T(\rho)\{\cos[kW(\rho)] + i\sin[kW(\rho)]\}\,, \tag{3.28}$$

kde $T(\rho)$ charakterizuje funkci amplitudové propustnosti pupily a $W(\rho)$ je vlnová aberace optického systému [40, 41, 54, 60].

Intenzita $I(P)$ může poté být vyjádřena jako

$$I(\mathrm{P}) = |U(\mathrm{P})|^2 = K \left| \int_0^1 P(\rho) J_0\left(\tau\rho\right)\ \rho \,\mathrm{d}\rho \right|^2 , \tag{3.29}$$

kde

$$K = |L|^2 = \left(\frac{\pi}{2\lambda c^2}\right)^2 . \tag{3.30}$$

Lze ukázat, že výpočet rozptylové funkce osového bodu systému (vyčíslení rovnice (3.29)) lze provést pomocí tzv. Soninova integrálu [61, 62]

$$J_{m+n+1}(x) = \frac{x^{n+1}}{2^n \Gamma(n+1)} \int_0^{\pi/2} J_m(x\sin t)\sin^{m+1} t \cos^{2n+1} t \,\mathrm{d}t . \tag{3.31}$$

Jestliže v rovnici (3.31) položíme $m = 0$, $x = \tau$ a $\sin t = \rho$, platí

$$\int_0^1 J_0(\tau\rho)(1-\rho^2)^n \rho \,\mathrm{d}\rho = 2^n \Gamma(n+1) \frac{J_{n+1}(\tau)}{\tau^{n+1}} , \tag{3.32}$$

kde $\Gamma(n+1)$ je Gamma funkce [37]. Je-li dále možné vyjádřit funkce $T(\rho)\sin[kW(\rho)]$ a $T(\rho)\cos[kW(\rho)]$ rozvojem v řady

$$\begin{aligned} T(\rho)\sin[kW(\rho)] &= \sum_s p_s (1-\rho^2)^s , \\ T(\rho)\cos[kW(\rho)] &= \sum_s q_s (1-\rho^2)^s , \end{aligned} \tag{3.33}$$

lze vyjádřit komplexní amplitudu $U(P)$ pomocí vztahu [10]

$$U(\mathrm{P}) = L \int_0^1 P(\rho) J_0\left(\tau\rho\right)\ \rho \,\mathrm{d}\rho = L \sum_s 2^s (q_s + i p_s) \Gamma(s+1) \frac{J_{s+1}(\tau)}{\tau^{s+1}} . \tag{3.34}$$

Pomocí rovnice (3.34) lze explicitně vypočítat komplexní amplitudu pole v případě zobrazení osového bodu optickým systémem, který je zatížen aberacemi, jestliže je funkce $P(\rho)$ rozvinuta v řadu bázových funkcí $\{(1-\rho^2)^s\}$, kde $s = 0, 1, 2, \ldots$. Následně lze snadno určit rozptylovou funkci bodu (PSF) pomocí (3.29), tj. $I(\mathrm{P}) = |U(\mathrm{P})|^2$. Normalizované rozložení intenzity (za předpokladu $I_n(\tau = 0) = 1$) je tak dáno vztahem [10]

$$I_n(\mathrm{P}) = K_n \left| \sum_s 2^s (q_s + i p_s) \Gamma(s+1) \frac{J_{s+1}(\tau)}{\tau^{s+1}} \right|^2 , \tag{3.35}$$

kde

$$K_n = 4 \left| \sum_s q_s \right|^{-2} . \tag{3.36}$$

Osová rozptylová funkce bodu (APSF z angl. Axial Point Spread Function), tj. rozložení energie v podélném směru osy šíření, lze určit jako limitní případ pro $\tau = 0$. V případě komplexní amplitudy pole platí [10]

$$U(\tau = 0) = L \sum_s \frac{(q_s + i p_s)}{2(s+1)}, \tag{3.37}$$

a normalizovaná intenzita poté bude dána vztahem [10]

$$\mathrm{APSF}_n = I_n(\tau = 0) = K_n \left| \sum_s \frac{(q_s + i p_s)}{2(s+1)} \right|^2. \tag{3.38}$$

Výše uvedený aparát je základem pro explicitní výpočet rozptylové funkce bodu. V práci [10] autoři dále podrobně analyzují případ výpočtu koeficientů rozvoje funkce $P(\rho)$ metodou nejmenších čtverců a pomocí Taylorova rozvoje. Metodu dále prezentují na příkladech optického systému zatíženého aberacemi do pátého řádu. Podrobnější informace může čtenář nalézt v kopii publikace [10] v Příloze A této práce.

Jedna ze základních složek optických aberací, která ovlivní zobrazení osového bodu rotačně symetrickou optickou soustavou, je tzv. defokusace. Jedná se o případ, kdy se poloha maximální intenzity pole v obrazovém prostoru nenachází v geometricko-optickém obraze osového bodu – je posunuta o hodnotu Δ. Analýzu rozptylové funkce bodu a osové rozptylové funkce bodu v případě defokusace pro optickou soustavu s kruhovou pupilou a pupilou tvaru mezikruží se autor věnoval a podílel se na publikaci práce [11]. V ní je podrobně odvozen výpočet komplexní amplitudy a intenzity pole pro danou situaci a jsou také odvozeny vztahy, pomocí kterých lze jednoduše určit základní charakteristiky rozptylových funkcí, např. polohy prvních minim v podélném a příčném směru. Je zde dále také ukázáno, jakým způsobem volit tvar mezikruží (poloměry ohraničujících kružnic), aby byla odpovídajícím způsobem ovlivněna hloubka ostrosti optické soustavy.

V případě rovnoměrně rozložené (konstantní) amplitudy v rovině výstupní pupily lze pro normalizovanou intenzitu $\mathrm{APSF}_n(\Delta) = I_n(\Delta)$ v podélném směru a středu difrakčního obrazce psát (tj. uvážíme $\tau = 0$, $J_0(0) = 1$, $A(\rho) = A_0$, $I_n(0,0) = 1$ a normalizované poloměry kružnic ohraničující mezikruží ρ_1 a ρ_2) [11]

$$\mathrm{APSF}_n(\Delta) = I_n(\Delta) = \frac{4}{(\rho_2^2 - \rho_1^2)^2} \left[\frac{\sin[k\beta(\rho_2^2 - \rho_1^2)/2]}{k\beta} \right]^2, \tag{3.39}$$

kde

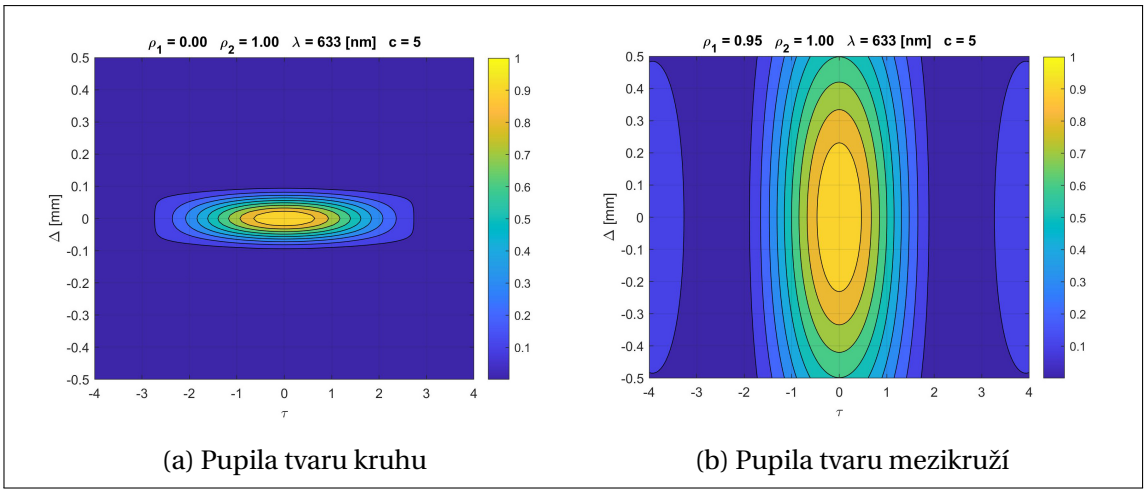
$$\beta = \frac{\Delta}{8c^2}. \tag{3.40}$$

Chceme-li určit hodnotu Δ_m, pro kterou normalizovaná intenzita na optické ose klesne na $I_m = I_n(\Delta_m)$, lze pro výpočet použít přibližný vztah [11]

$$\Delta_m(I_m) = \pm \frac{16c^2}{k\Delta_\rho} \sqrt{10 - 2\sqrt{30\sqrt{I_m} - 5}}, \tag{3.41}$$

kde reálné řešení dostáváme pro minimální hodnotu $I_m = (5/30)^2 \approx 0.028$.

Pomocí rovnic (3.39) a (3.41) lze studovat, jaký vliv bude mít geometrie mezikruží na zobrazení bodu optickou soustavou, na osovou rozptylovou funkci, a tím pádem na hloubku ostrosti optické soustavy. Větší hodnota Δ_m totiž jinými slovy říká, že bod bude



(a) Pupila tvaru kruhu (b) Pupila tvaru mezikruží

Obrázek 3.4: Rozložení normalizované intenzity kolem obrazu osového bodu optické soustavy s konstantní amplitudou v rovině výstupní pupily tvaru kruhu a mezikruží [11]

zobrazitelný ostře ve větší podélné oblasti kolem geometricko-optické polohy svého ideálního obrazu. Například na umístění senzoru za optickou soustavou mohou tak být kladeny nižší požadavky na přesnost. Na obr. 3.4a je pro ilustraci ukázáno rozložení intenzity kolem obrazu osového bodu optické soustavy s konstantní amplitudou v rovině výstupní pupily tvaru kruhu ($\rho_1 = 0$, $\rho_2 = 1$) a na obr. 3.4b poté rozložení intenzity pro pupilu tvaru mezikruží ($\rho_1 = 0.95$, $\rho_2 = 1$). Další příklady a analýzy (např. pro Gaussovské svazky [40, 41]) může čtenář nalézt v kopii článku v Příloze A.

Řešení úlohy defokusace v obráceném pojetí bylo publikováno v práci [12], kde se autor podílel na tvorbě metodiky určení prostorové polohy částice z měření dvojrozměrného rozložení intenzity v mikroskopii. Ze zaznamenaného rozložení normalizované intenzity bodu je tak možné určit hodnotu defokusace, resp. podélnou polohu částice, ve které se nachází vzhledem k rovině senzoru.

Lze ukázat, že hodnota malé defokusace může být vyjádřena jako funkce normalizovaného osvětlení $L_n(\tau_m, \Delta)$ na vybrané kruhové oblasti. Platí [12]

$$\Delta = \frac{4c^2\tau_m^2}{k}\sqrt{\frac{1 - J_0^2(\tau_m) - J_1^2(\tau_m) - L_n(\tau_m, \Delta)}{\tau_m^2\left[J_0^2(\tau_m) + J_1^2(\tau_m)\right] - 4J_1^2(\tau_m)}}, \tag{3.42}$$

kde hodnotu osvětlení $L_n(\tau_m, \Delta)$ určíme z experimentálního měření a pomocí integrace

$$L_n(\tau_m, \Delta) = \frac{1}{2}\int_0^{\tau_m} I_n(\tau, \Delta)\tau \, \mathrm{d}\tau, \tag{3.43}$$

τ_m značí poloměr kruhu v rovině obrazu, na kterém integraci provádíme a jehož střed je umístěn v místě maxima registrované intenzity, J_0 a J_1 jsou Besselovy funkce prvního druhu (řádu nula a jedna).

Pro velké hodnoty defokusací neexistuje jednoduché řešení, ale alternativní způsoby jsou diskutovány v práci [12]. Dále jsou zde uvedeny vztahy pro výpočet defokusace v případě čtvercové apertury a řešení jsou demonstrována na příkladech a experimentálním měření. Podrobnosti může čtenář najít v kopii publikace v Příloze A.

3.2 Návrh optických soustav s korigovanými aberacemi

Porozumění vzniku a možnostem eliminace optických aberací patří mezi základní úkoly během návrhu optických systémů. Obecně se tomuto tématu věnovala řada autorů a neustále jsou na danou problematiku publikovány nové práce. Komplexní analýza aberací a všech možností jejich minimalizace pro optický návrh je nad rámec této práce, nicméně čtenář snadno najde řadu knih a článků, které se těmto oblastem věnují [40, 41, 48, 54–56].

Obecně můžeme postup optického návrhu charakterizovat následujícími kroky:

1. zadání základních zobrazovacích požadavků na optický systém,
2. analýza průchodu paprsků (vlnového pole) optickým systémem,
3. určení zájmových charakteristik (např. rozptylová funkce bodu, Seidelovy sumy apod.),
4. optimalizace parametrů charakterizujících optické členy (optická rozhraní, materiály, geometrické pozice apod.), aby zájmové charakteristiky dosahovaly požadovaných hodnot (např. minimum poloměru rozptylového kroužku).

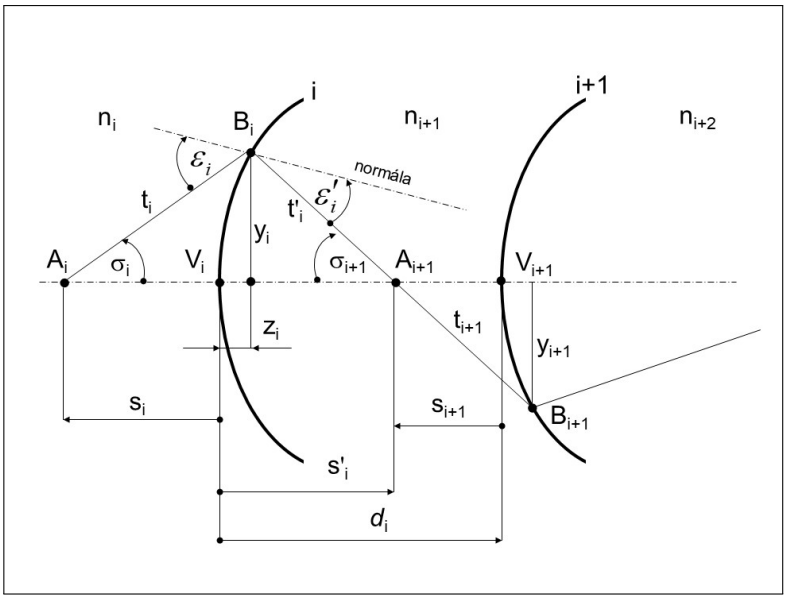
Popsaný postup zahrnuje jak případ nominálního paraxiálního návrhu, tak i následné reálné optimalizace parametrů systému. Použijeme-li obecné formulace zobrazovacích rovnic optické teorie, můžeme vyjádřit vliv parametrů optických členů systému na požadovanou zájmovou charakteristiku. Následně lze určit parametry systému tak, aby zájmová charakteristika splňovala požadovaná kritéria.

Jako jednoduchý příklad uveďme princip výpočtu tvaru rozhraní jednoduché čočky takové, aby zobrazovala osový bod předmětu s co nejlepší kvalitou. Základní požadavky na optický systém budou v daném případě vzdálenost předmětu a obrazu od vrcholů čočky. Předpokládejme, že první plocha čočky bude rovina a druhá rotačně symetrická asférická plocha určená např. rozvojem v mocninnou řadu a že tloušťka čočky je fixně zvolena. Zájmovou charakteristiku zvolíme pro jednoduchost jako kvadratický průměr vzdáleností mezi optimální polohou obrazu v obrazové rovině a průsečíky, ve kterých protínají obrazovou rovinu paprsky svazku vycházející z osového bodu předmětu po průchodu čočkou. Pomocí zobrazovacích rovnic vyjádříme funkční závislost zájmové charakteristiky na parametrech mocninné řady, která popisuje zadní asférickou plochu čočky. Tímto definujeme optimalizační úlohu, kdy dále hledáme takové hodnoty parametrů asférické plochy, aby hodnota zájmové charakteristiky byla minimální.

Vybraným oblastem základního optického návrhu se autor této práce věnoval a podílel se na publikacích [13–15], ve kterých prezentuje zejména postup minimalizace sférické aberace a její vliv na hloubku ostrosti optické soustavy, a dále základní teoretickou analýzu návrhu dvoučlenných optických systémů.

3.2.1 Korekce sférické aberace zobrazení osového bodu

Uveďme nyní jako ukázku výstupů publikovaných v práci [13] postup návrhu optického systému s jednou nebo dvěma asférickými plochami, který minimalizuje sférickou aberaci při zobrazení osového bodu předmětu. Úloha předpokládá, že známe polohy jednotlivých optických rozhraní (osové vrcholové vzdálenosti), poloměry křivosti sférických



Obrázek 3.5: Lom paprsků na optických plochách [13]

ploch a parametry první asférické plochy. Cílem je nalézt takové parametry druhé asférické plochy, kdy obraz osového bodu předmětu bude stigmatický, tj. osový bod předmětu se zobrazí opět jako bod.

Pro úvod do problematiky se zabývejme lomem paprsků na řadě optických rotačně symetrických rozhraních. Na obr. 3.5 je zobrazeno schéma lomu meridionálního aperturního paprsku na i-té a $(i+1)$-ní ploše (sférické nebo asférické). Jednotlivé symboly v obrázku značí: n_i – index lomu prostředí za i-tým optickým rozhraním, σ_i – úhel mezi paprskem dopadajícím na i-tou plochu a optickou osou soustavy (osou z), σ_{i+1} – úhel mezi paprskem dopadajícím na $(i+1)$-ní plochu a optickou osou, s_i – osovou vzdálenost mezi bodem A_i a vrcholem V_i i-té plochy, y_i – příčnou vzdálenost bodu B_i od optické osy, d_i – vzdálenost mezi vrcholem V_{i+1} $(i+1)$-ní plochy a vrcholem V_i i-té plochy, z_i – z-ovou souřadnici bodu B_i měřenou od vrcholu V_i, t_i – vzdálenost mezi body A_i a B_i. Význam dalších symbolů je patrný z obr. 3.5. Pro ujasnění znaménkové konvence je vzdálenost s_i předpokládána negativní, jestliže bod A_i leží vlevo od vrcholu V_i a kladná v opačném případě. Úhly σ jsou měřeny od optické osy a jsou kladné, jestliže jsou měřeny ve směru hodinových ručiček.

Obecně lze průchod meridionálního paprsku skrze asférickou plochu, která je popsána funkcí $z = f(y)$, charakterizovat pomocí vztahů $(i = 1,2,3,\ldots,m)$ [63]

$$y_i = (s_i - z_i)\tan\sigma_i\,, \quad \tan\omega_i = \left(\frac{\mathrm{d}z}{\mathrm{d}y}\right)_i\,, \quad \varepsilon_i = \omega_i - \sigma_i\,, \tag{3.44}$$

$$\sin\varepsilon_i' = \frac{n_i}{n_{i+1}}\sin\varepsilon_i\,, \quad \sigma_{i+1} = \sigma_i + \varepsilon_i - \varepsilon_i'\,, \quad s_i' = \frac{y_i}{\tan\sigma_{i+1}} + z_i\,,$$

kde ω_i je úhel mezi normálou v bodě (y_i, z_i) a osou z a ε_i a ε_i' značí úhly dopadajícího a lomeného paprsku vzhledem k normále i-tého rozhraní.

Předpokládejme dále, že je optický systém složen z m optických ploch a zobrazuje bod $A \equiv A_1$ do bodu $A' \equiv A_{m+1}$. Nechť $[AA']$ značí optickou dráhu obecného meridionálního paprsku (sumu násobků geometrických vzdáleností a indexů lomů v jednotlivých optických prostředích mezi rozhraními) a nechť $[AA']_0$ značí optickou dráhu paprsku procházejícího podél optické osy. Rozdíl optických drah těchto paprsků poté bude dán vztahem

$$\delta = [AA'] - [AA']_0\,. \tag{3.45}$$

S ohledem na obr. 3.5 lze dále pro rotačně symetrické sférické a asférické plochy psát

$$\delta = \sum_{i=1}^{m} \delta_i = \sum_{i=1}^{m} \left[n_{i+1} \left(t'_i - s'_i \right) - n_i \left(t_i - s_i \right) \right] \tag{3.46}$$
$$= \sum_{i=1}^{m} \left[y_i \left(n_{i+1} \tan \frac{\sigma_{i+1}}{2} - n_i \tan \frac{\sigma_i}{2} \right) - \left(n_{i+1} - n_i \right) z_i \right] .$$

Je známo, že bude-li bod A stigmaticky zobrazen do bodu A', poté musí být splněna podmínka $\delta = 0$ pro všechny paprsky svazku. Navržení optického systému, který bude nulovat všechny výše zmíněné optické dráhové rozdíly δ povede k optimálnímu optickému návrhu.

Pro jednoduchost a demonstraci principu dalšího postupu uvažujme nyní jednoduchou čočku ve vzduchu ($m = 2$, $n_1 = 1$, $n_2 = n$, $n_3 = 1$), kde bude tvar první plochy znám, např. sférický nebo asférický. Tudíž lze velmi snadno určit parametry paprsku po průchodu touto první plochou. Vzhledem k obr. 3.5 dále platí

$$\tan \sigma_2 = \frac{y_2}{s_2 - z_2}, \quad \tan \sigma_3 = \frac{y_2}{s'_2 - z_2} . \tag{3.47}$$

Řešení soustavy (3.47) pro y_2 a z_2 vede na

$$y_2 = \frac{(s'_2 - s_2) \tan \sigma_2 \tan \sigma_3}{\tan \sigma_2 - \tan \sigma_3}, \quad z_2 = \frac{s_2 \tan \sigma_2 - s'_2 \tan \sigma_3}{\tan \sigma_2 - \tan \sigma_3} . \tag{3.48}$$

Pro stigmaticky zobrazený bod A do bodu A', tj. $\delta = \delta_1 + \delta_2 = 0$, dále platí

$$\delta_1 + y_2 \left(\tan \frac{\sigma_3}{2} - n \tan \frac{\sigma_2}{2} \right) - (1 - n) z_2 = 0 , \tag{3.49}$$

kde

$$\delta_1 = y_1 \left(n \tan \frac{\sigma_2}{2} - \tan \frac{\sigma_1}{2} \right) - (n - 1) z_1 \tag{3.50}$$

je optický dráhový rozdíl zavedený první plochou čočky. Dosazením rovnic (3.48) do (3.49) poté dostáváme pro úhel σ_3 následující rovnici, platí

$$\alpha \sin \sigma_3 + \beta \cos \sigma_3 + \gamma = 0 , \tag{3.51}$$

kde

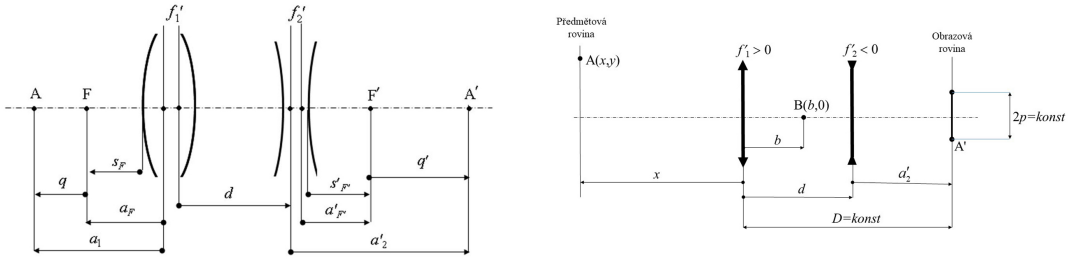
$$\begin{aligned} \alpha &= (\delta_1 - s'_2 + n s_2) \cos \sigma_2 - n (s_2 - s'_2) , \\ \beta &= -(\delta_1 - s'_2 + n s_2) \sin \sigma_2 , \\ \gamma &= (s_2 - s'_2) \sin \sigma_2 . \end{aligned} \tag{3.52}$$

Řešení rovnice (3.51) vede na vztah

$$\sin \sigma_3 = - \frac{\alpha \gamma + \beta \sqrt{\alpha^2 + \beta^2 - \gamma^2}}{\alpha^2 + \beta^2} . \tag{3.53}$$

Dosazením (3.53) do vztahů (3.48) lze po snadné úpravě získat parametrické souřadnice bodu na druhém povrchu čočky (asférickém). Tato čočka bude zobrazovat bod A stigmaticky do bodu A'. Tuto množinu bodů lze poté aproximovat rozvojem v mocninou řadu nebo jiné vhodné vyjádření pro další optické analýzy.

V práci [13] dále autoři prezentují další analýzy a nástroje pro jiné možnosti trasování mimoosových paprsků obecnými asférickými plochami. Jaký má vliv sférická aberace na hloubku ostrosti optické soustavy a tvar kaustiky při zobrazení osového bodu může čtenář nalézt v publikaci [14]. Kopie těchto článků jsou k dispozici v Příloze A.



(a) Schéma dvoučlenného optického systému (b) Schéma analaktického dalekohledu

Obrázek 3.6: Základní schéma dvoučlenné optické soustavy a analaktického dalekohledu [15]

3.2.2 Návrh dvoučlenných optických systémů

V optické praxi je používána celá řada typů optických systémů, jejichž konstrukce se liší v závislosti na zamýšleném použití a požadavcích na kvalitu optického zobrazení. Jedněmi z velmi často používaných jsou dvoučlenné optické systémy skládající se z objektivu a okuláru. Objektiv i okulár jsou samostatné soustavy tvořené řadou optických členů, které jsou spjaty takovým způsobem, aby poskytovaly požadované zobrazení.

V práci [15] se autor podílel na základní teoretické analýze dvoučlenných optických systémů, kdy první člen má kladnou lámavost a druhý zápornou. Jedná se zejména o optické systémy Petzvalova objektivu, teleobjektivu, reversního teleobjektivu a objektivů analaktického typu. Tyto optické systémy jsou velmi často používány v praxi, zejména nacházejí uplatnění ve fotografii a měřických zařízeních (např. teodolity nebo nivelační přístroje), kde se používá zpravidla analaktický typ dalekohledu s vnitřním zaostřováním (délka tubusu dalekohledu zůstává stálá). V publikaci je prezentována metoda návrhu základních parametrů objektivu, tj. ohniskové vzdálenosti jednotlivých optických komponentů a jejich vzájemné vzdálenosti a poloměry křivosti jednotlivých rozhraní čoček, jestliže jsou pro konstrukci použity tmelené dublety. Následně je provedena důkladná analýza optických aberací těchto systémů, pomocí které lze provést prvotní návrh systému takový, kdy budou volené aberace minimalizovány.

V následující části budou představeny základní vztahy pro charakteristiku dvoučlenného analaktického dalekohledu.

Předpokládejme, že je optická soustava dalekohledu složena ze dvou komponentů. Optickým komponentem máme na mysli soustavu jedné nebo více jednoduchých čoček (dublet, triplet apod.). Na obr. 3.6a je zobrazeno základní schéma této soustavy. Význam jednotlivých symbolů na obrázku je následující: f_1' a f_2' jsou obrazové ohniskové vzdálenosti prvního a druhého komponentu, d je jejich vzájemná vzdálenost, a_1 je vzdálenost mezi osovým bodem na předmětu A a objektovou hlavní rovinou prvního optického komponentu, a_2' je vzdálenost mezi obrazem A' osového bodu A a obrazovou hlavní rovinou druhého optického komponentu, a_F je vzdálenost mezi objektovým ohniskem F objektivu a hlavní rovinou prvního optického komponentu, $a'_{F'}$ je vzdálenost mezi obrazovým ohniskem objektivu F' a obrazovou hlavní rovinou druhého komponentu, q je vzdálenost mezi bodem A a objektovým ohniskem F a q' je vzdálenost mezi bodem A' a obrazovým ohniskem F'.

Předpokládejme dále, že optická soustava je ve vzduchu. Použitím základních zobra-

zovacích rovnic [40, 41, 63] poté dostáváme vztahy [15]

$$qq' = -f'^2\,, \quad a_F = -f'\left(1 - d/f_2'\right), \quad a_{F'}' = f'\left(1 - d/f_1'\right), \quad f' = \frac{f_1' f_2'}{f_1' + f_2' - d}, \tag{3.54}$$

kde f' je obrazová ohnisková vzdálenost optického systému. Příčné zvětšení m optického systému může být vypočteno ze vztahu

$$m = \frac{y'}{y} = \frac{f'}{q} = -\frac{q'}{f'} = \frac{a_1' a_2'}{a_1 a_2}, \tag{3.55}$$

kde y značí velikost předmětu a y' je velikost obrazu. Rovnice (3.54) a (3.55) charakterizují základní zobrazovací vlastnosti dvoučlenného optického systému ve vzduchu.

Nyní se zabývejme výše zmíněným analaktickým dalekohledem. Jedná se o optický systém ze dvou komponentů, kdy první má kladnou lámavost a druhý (vnitřní komponent) zápornou. Schéma takového systému je ukázáno na obr. 3.6b. Bez újmy na obecnosti jsou ukázány jednotlivé komponenty ve formě tenkých čoček. V obrazové rovině se nachází záměrný kříž na skleněné destičce. Na ní jsou vyryté dvě linie o vzájemné vzdálenosti $2p$. Ostření na různě vzdálené objekty se provádí posunem druhého komponentu, zatímco vzdálenost D zůstává neměnná – tzv. vnitřní ostření. Neměnnost vzdálenosti D je výhodná v řadě aplikací, kde není žádoucí měnit mechanickou délku tubusu dalekohledu.

Ze zobrazovacích rovnic lze odvodit vztahy [15]

$$a_1' = \frac{a_1 f_1'}{a_1 + f_1'}, \quad a_2 = a_1' - d, \quad a_2' = \frac{a_2 f_2'}{a_2 + f_2'}, \quad m = \frac{a_1' a_2'}{a_1 a_2} = \frac{p}{y}, \tag{3.56}$$

kde a_i a a_i' $(i = 1,2)$ značí předmětové a obrazové osové vzdálenosti a m je příčné zvětšení soustavy. Nechť $a_1 = x$. S použitím podmínky $D = d + a_2 = konst.$ a s rovnicí příčného zvětšení poté z rovnice (3.56) dostáváme

$$\begin{gathered} d - D + \frac{f_2'\left(\frac{f_1' x}{f_1' + x} - d\right)}{f_2' + \left(\frac{f_1' x}{f_1' + x} - d\right)} = 0\,, \\ \frac{f_1' f_2' y}{\left(f_1' + x\right)\left(f_2' + \frac{f_1' x}{f_1' + x} - d\right)} - p = 0\,. \end{gathered} \tag{3.57}$$

Eliminací vzdálenosti d z rovnic (3.57) poté dostáváme

$$F(x,y) = Ax^2 + Bxy + Cy^2 + Ex + Fy + G\,, \tag{3.58}$$

kde

$$\begin{aligned} A &= p^2 f_2'\,, \quad B = p f_1'\left(D - f_1' - 2f_2'\right), \quad C = f_1'^2 f_2'\,, \\ E &= 2p^2 f_1' f_2'\,, \quad F = f_1'^2 p\left(D - 2f_2'\right), \quad G = p^2 f_1'^2 f_2'\,. \end{aligned} \tag{3.59}$$

Rovnice (3.58) popisuje křivku, na které bod $A(x,y)$ leží, jestliže se mění vzdálenost x. Snadno lze poté z této rovnice ukázat, že křivkou je hyperbola s asymptotami

$$y_{a1} = k_1 x + q_1\,, \quad y_{a2} = k_2 x + q_2\,, \tag{3.60}$$

jejichž parametry lze vyjádřit postupem uvedeným např. v [37]. Po úpravě dostáváme [15]

$$k_{1,2} = \frac{p\left[f_1' + 2f_2' - D \mp \sqrt{(f_1' - D)(f_1' + 4f_2' - D)}\right]}{2f_1'f_2'}\,, \tag{3.61}$$
$$q_i = \frac{p(2f_2'p - 2f_1'f_2'k_i + f_1'k_iD)}{p(f_1' + 2f_2' - D) - 2f_1'f_2'k_i}\,, \quad i = 1,2\,.$$

Úhel β mezi asymptotami může být vypočten ze vztahu [37]

$$\tan\beta = \left|\frac{k_2 - k_1}{1 + k_1k_2}\right|\,. \tag{3.62}$$

Předpokládejme dále, že je soustava zaostřena na nekonečno a následující veličiny jsou známé: ohnisková vzdálenost f_0' soustavy, délka dalekohledu D, pozice analaktického bodu $b = -q_2/k_2$, vzájemná vzdálenost vrypů na destičce $2p$ a parametr k_2 (obvykle volen $k_2 = -0.01$). Následně mohou být hodnoty ohniskových vzdáleností f_1' a f_2' a vzdálenosti d_0 (pro předmět v nekonečnu) určeny ze vztahů [15]

$$G = b^2p^2 + (2k_2b^2p + 6bp^2)D + (p - bk_2)^2D^2\,, \tag{3.63}$$
$$f_1' = -\frac{bp - pD - bk_2D + \sqrt{G}}{4p + 2bk_2 - 2k_2D}\,, \quad f_2' = \frac{f_1'^2k_2p - f_1'k_2pD}{f_1'^2k_2^2 - 2f_1'k_2p + p^2}\,,$$
$$d_0 = f_1' + f_2' - f_1'f_2'/f_0'\,.$$

Při přeostřování na konečnou vzdálenost x od prvního optického komponentu soustavy můžeme určit požadovanou vzdálenost d_x mezi komponenty ze vztahů [15]

$$d_x = \frac{H + f_1'x + D(f_1' + x)}{2(f_1' + x)}\,,\; H = \sqrt{[f_1'x - D(f_1' + x)][4f_1'f_2' + (f_1' + 4f_2')x - D(f_1' + x)]}\,. \tag{3.64}$$

Zabývejme se nyní rozdílem mezi asymptotami a křivkou polohy bodu $A(x,y)$. Souřadnice y bodu $A(x,y)$ může být vypočtena ze vztahu [15]

$$y = \frac{p}{2f_1'f_2'}\left[(f_1' + x)(2f_2' - D) + f_1'x - H\right]\,. \tag{3.65}$$

Výška obrazu y' následně bude dána vztahem

$$y' = \frac{f_1'f_2'}{f_1'x - (f_1' + x)(d_x - f_2')}\,y\,. \tag{3.66}$$

Bude-li bod $A(x,y)$ ležet na druhé asymptotě, jeho obrazová výška y_a' bude dána vztahem

$$y_a' = U - a'\left(\frac{U}{f_2'} - k_2 + \frac{q_2}{f_1'}\right)\,, \quad U = q_2 + d_x\left(k_2 - \frac{q_2}{f_1'}\right)\,. \tag{3.67}$$

Rozdíl mezi obrazovými výškami $\Delta y' = y' - y_a'$ lze snadno určit pomocí rovnic (3.66) a (3.67). Vzdálenost x_B mezi bodem $A(x,y)$ a analaktickým bodem B může být vyjádřena vztahem $x_B = -y_a/k_2$. Lze ukázat, že pro delší vzdálenosti x_B je rozdíl $\Delta y'$ velmi malý, a tedy výpočet vzdálenosti x při změření y může být snadno proveden ze znalosti parametrů asymptoty. Na krátké vzdálenosti je ovšem potřeba zavádět vhodné numerické korekce.

V práci [15] autoři dále podrobně představují postup výpočtu parametrů jednotlivých členů optického dvoučlenného systému, který bude mít korigovány vybrané aberace. Odvozené vztahy následně prezentují na několika příkladech. Pro další podrobnosti odkážeme čtenáře na zmíněnou publikaci, jejíž kopie je uvedena v Příloze A.

4 Optické (laserové) skenery

Optické (laserové) skenování [64, 65] se během posledních let stalo velmi populární technikou kvantitativní charakteristiky geometrie povrchu a vlastností objektů. A to jak v rámci měření velmi malých předmětů (např. v mikroskopii a strojírenství), tak pro charakteristiku velmi rozsáhlých oblastí (např. pozemní skenování v geodézii). Hlavním důvodem úspěchu této techniky je zejména rozvoj a dostupnost komerčních zařízení, která jsou schopna v rámci zlomků vteřin zaznamenat informace o milionech bodů. Také výpočetní kapacity značně narostly a analýza mračen bodů, která jsou uložena jako velmi velké datové soubory, je snazší.

Tématu optického (laserového) skenování se autor této práce aktivně věnoval a je spoluautorem několika článků, které se danou problematikou zabývají [17, 18]. V publikaci [18] je podrobně provedena analýza polohy stopy optického svazku a její přesnosti pro jednozrcadlové a dvouzrcadlové skenery. Práce [17] je zaměřena zejména na základní charakteristiku vlivů (matematických i fyzikálních) na přesnost pozemního skenování na dlouhé vzdálenosti. Kopie obou publikací jsou uvedeny v Příloze A, kde čtenář může nalézt podrobné informace a reference.

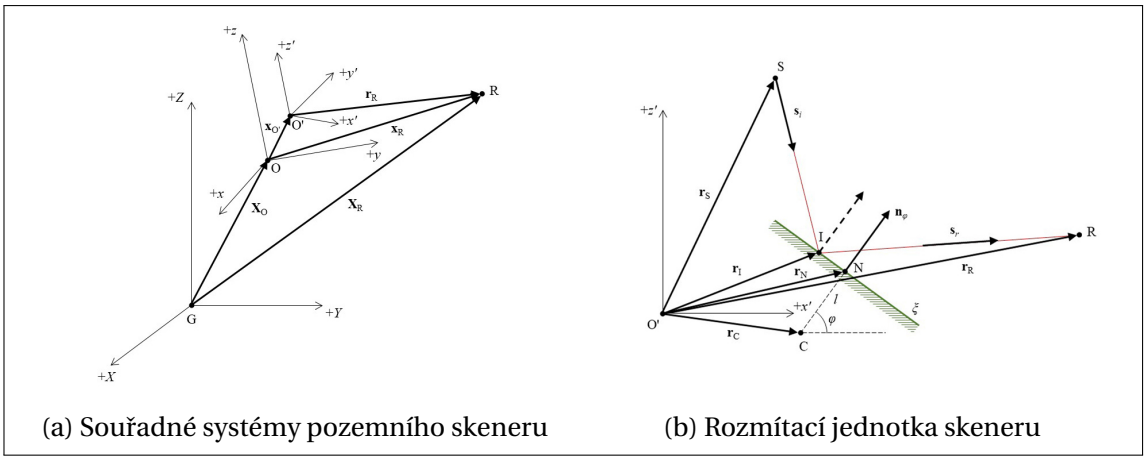
Představme nyní základní formulaci apriorní analýzy přesnosti pro pozemní laserové skenování. Kvalitu měření bodů ovlivňují v první řadě geometrické nejistoty v konstrukci skeneru. V principu pozemní skener určuje prostorové sférické souřadnice (horizontální a vertikální úhly a vzdálenosti) vzhledem k vhodně zvolenému souřadnému systému. Na obr. 4.1a je ukázán pravotočivý kartézský souřadný systém pro matematický popis polohy bodu určeného z měření. Globální souřadný systém (X, Y, Z) je určen svým počátkem G. V tomto systému se nachází lokální souřadný systém skeneru (x, y, z), jehož počátek O je dán polohovým vektorem $\mathbf{X}_O$ v globálním systému a osy systému skeneru jsou rotovány o hodnoty *Roll*, *Pitch* a *Yaw*. V systému skeneru je dále umístěna rozmítací jednotka, jejíž interní souřadný systém (x', y', z') je charakterizován počátkem O', který je dán polohovým vektorem $\mathbf{x}_{O'}$ v rámci systému skeneru, a osy systému jednotky jsou otočeny o úhly α, β a γ. Měřený bod R je určen polohovým vektorem $\mathbf{X}_R$ v globálním systému souřadnic, $\mathbf{x}_R$ v systému skeneru a $\mathbf{r}_R$ v systému rozmítací jednotky.

Transformační vztahy mezi jednotlivými souřadnými systémy mohou být charakterizovány následujícími vztahy, platí [17]

$$\mathbf{X}_R = \mathbf{X}_O + \mathbf{S}\mathbf{x}_R\,, \quad \mathbf{x}_R = \mathbf{x}_{O'} + \mathbf{R}\mathbf{r}_R\,, \tag{4.1}$$

kde

$$\mathbf{X}_R = [X_R, Y_R, Z_R]^T\,, \quad \mathbf{X}_O = [X_O, Y_O, Z_O]^T\,, \quad \mathbf{x}_R = [x_R, y_R, z_R]^T$$
$$\mathbf{x}_{O'} = [x_{O'}, y_{O'}, z_{O'}]^T\,, \quad \mathbf{r}_R = [x'_R, y'_R, z'_R]^T\,,$$



(a) Souřadné systémy pozemního skeneru (b) Rozmítací jednotka skeneru

Obrázek 4.1: Základní schémata souřadných systémů a rozmítací jednotky pozemního laserového skeneru [17]

$$\mathbf{S} = \mathbf{R}_z(Yaw)\mathbf{R}_y(Pitch)\mathbf{R}_x(Roll)\,,$$
$$\mathbf{R} = \mathbf{R}_z(\gamma)\mathbf{R}_y(\beta)\mathbf{R}_x(\alpha)\,,$$

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

V pozemních skenovacích systémech je rozmítací jednotka obecně tvořena jednou odraznou plochou (zrcadlem, monogonem, optickým, polygonálním nebo pyramidálním hranolem), která rozmítá laserový svazek ve vertikálním směru relativně vzhledem k tělu skeneru. Dále se celé tělo skeneru otáčí a je tak zajištěno rozmítání v horizontálním směru. Jinou z možností je použití např. dvouzrcadlových skenerů [18], tou se ale v této práci dále zabývat nebudeme (v praxi pozemních skenerů se příliš nevyužívá). Uvažme tedy systém jedné reflexní plochy, která rozmítá svazek záření ve vertikální rovině vzhledem k tělu skeneru. Matematické schéma je zobrazeno na obr. 4.1b.

Předpokládejme, že střed otáčení C reflexní plochy ξ je dán polohovým vektorem $\mathbf{r}_C$ a nechť se tato plocha dále otáčí kolem osy rovnoběžné s osou y' souřadného systému jednotky o úhel φ. Vzdálenost mezi odraznou plochou ξ a bodem C je označena jako l. Rovnice roviny odrazné plochy může být určena pozicí bodu N, který je dán polohovým vektorem $\mathbf{r}_N$, a jednotkovým normálovým vektorem $\mathbf{n}$. Dále předpokládejme, že osa laserového svazku (záměrná přímka) dopadajícího na odraznou plochu je určena bodem S (např. zdrojem záření) charakterizovaným polohovým vektorem $\mathbf{r}_S$ a jednotkovým směrovým vektorem $\mathbf{s}_i$. Průsečík záměrné přímky s odraznou rovinou je označen I a je dán polohovým vektorem $\mathbf{r}_I$. Záření se v tomto bodě odráží dle zákona odrazu a pokračuje ve směru daném jednotkovým směrovým vektorem $\mathbf{s}_r$. Tento vektor může být určen z vektorové formy zákona odrazu, platí [40, 41]

$$\mathbf{s}_r = \mathbf{s}_i - 2\mathbf{n}(\mathbf{s}_i \cdot \mathbf{n})\,, \tag{4.2}$$

kde $\mathbf{s}_i$ značí jednotkový směrový vektor osy dopadajícího svazku, $\mathbf{n}$ je jednotkový normálový vektor odrazné plochy (orientovaný do směru k dopadajícímu záření), $\mathbf{s}_r$ je jednotkový směrový vektor osy odraženého svazku. Vztah pro rotaci jednotkového normálového

vektoru $\mathbf{n}_0$ okolo osy dané jednotkovým směrovým vektorem $\mathbf{c}$ o úhel φ, která protíná počátek pomocného souřadného systému, lze psát jako [18]

$$\mathbf{n}_\varphi = \mathbf{n}_0 \cos\varphi + \mathbf{c}\,(\mathbf{c}\cdot\mathbf{n}_0)\,(1-\cos\varphi) + (\mathbf{c}\times\mathbf{n}_0)\sin\varphi\,. \tag{4.3}$$

Pozice bodu R vzhledem k rozmítací jednotce lze vyjádřit po úpravě jako [17]

$$\mathbf{r}_R = \mathrm{r}_S + \frac{(\mathbf{r}_C - \mathbf{r}_S + l\,\mathbf{n}_\varphi)\cdot\mathbf{n}_\varphi}{\mathbf{s}_i\cdot\mathbf{n}_\varphi}\mathbf{s}_i + p_r[\mathbf{s}_i - 2\mathbf{n}_\varphi(\mathbf{s}_i\cdot\mathbf{n}_\varphi)]\,, \tag{4.4}$$

kde p_r je parametr numericky rovný prostorové vzdálenosti mezi body I a R.

Pro teoretické analýzy (např. nejistoty polohy stopy svazku apod.) lze hodnotu parametru p_r určit z nominální definice pozice detekční roviny. Jestliže je detekční rovina ρ, na kterou dopadá záření a odráží se, dána bodem D určeným polohovým vektorem $\mathbf{r}_D$ a jednotkovým normálovým vektorem $\mathbf{n}_D$, poté pro parametr p_r platí [17]

$$p_r = \frac{\left[\mathbf{r}_D - \mathbf{r}_S - \frac{(\mathbf{r}_C-\mathbf{r}_S+l\,\mathbf{n}_\varphi)\cdot\mathbf{n}_\varphi}{\mathbf{s}_i\cdot\mathbf{n}_\varphi}\mathbf{s}_i\right]\cdot\mathbf{n}_D}{[\mathbf{s}_i - 2\mathbf{n}_\varphi(\mathbf{s}_i\cdot\mathbf{n}_\varphi)]\cdot\mathbf{n}_D}. \tag{4.5}$$

V praktických aplikacích je nejčastěji poloha cílového bodu určována pomocí měření tranzitního času (metoda time-of-flight), který charakterizuje dobu trvání mezi vysláním a přijmutím pulzu, kdy je svazek záření vyslán ze zdroje, projde rozmítací jednotkou, putuje k cíli, odrazí se a stejnou cestou putuje zpět do přijímače skeneru. Označíme-li rychlost pulzu v daném prostředí $v = c/n$ ($c = 299\,792\,458$ m/s [40, 41] je rychlost světla ve vakuu, n značí index lomu prostředí), poté vzdálenost d měřeného bodu od referenčního bodu (zdroje pulzu) je dána jako $d = v\tau/2$, kde τ označuje tranzitní čas mezi vysláním a přijmutím signálu. V našem případě je vzdálenost mezi zdrojem a odrazným bodem dána jako $d = p_i + p_r$. Po úpravě poté můžeme parametr p_r vyjádřit s použitím měřeného tranzitního času τ jako

$$p_r = v\frac{\tau}{2} - \frac{(\mathbf{r}_C - \mathbf{r}_S + l\,\mathbf{n}_\varphi)\cdot\mathbf{n}_\varphi}{\mathbf{s}_i\cdot\mathbf{n}_\varphi}\,. \tag{4.6}$$

Pomocí výše uvedených vztahů lze analyzovat vliv nejistot konstrukce skeneru nebo pozice skeneru na výsledné globální souřadnice bodu R. Pro tyto potřeby je možné použít simulační metody, ve kterých jsou jednotlivé nejistoty dány relevantními pravděpodobnostními modely a přidány k nominálním modelovaným hodnotám. Následné statistické zpracování může sloužit k vyhodnocení prostorové nejistoty bodu R. V některých případech je možné vyjádřit vztahy pro polohu bodu R analyticky a použít k odhadu nejistot např. zákon přenášení variancí [18, 37].

Další ze základních vlivů na přesnost určení polohy bodů při laserovém skenování má přesnost elektronického měření vzdálenosti. V rámci pozemního laserového skenování se používá měření bez odrazných zařízení, protože na cílovém objektu nejsou umístěny odrazné značky. Obecně existuje celá řada typů elektronického měření vzdáleností. Základní a nejjednodušší metodou je tzv. metoda tranzitního času. Při ní je měřen čas putování světelného pulzu od vyslání skenerem, přes odraz na cílovém objektu až k následnému přijetí skenerem. Druhou kategorií mohou být metody založené na měření fázové změny – metody amplitudově nebo fázově modulované spojité vlny [66]. Tyto metody dosahují lepší přesnosti, ale jsou konstrukčně náročnější, a proto se v oblasti pozemního laserového skenování příliš nepoužívají. Dále tedy budeme uvažovat pouze metodu tranzitního času.

Jak již bylo zmíněno, metoda tranzitního času [64–66] je založena na velmi přesném měření času šíření elektromagnetického pulzu skrze prostředí mezi zdrojem záření (skenerem) a cílem. Puls je ve skeneru generován zpravidla pulzní diodou. Na dělícím zařízení (např. polopropustném zrcadle) je jeho část odkloněna na spínač měření času. Další část pokračuje k cílovému objektu, od kterého se odráží, a putuje zpět do zařízení, kde je časový spínač vypnut. Časová měřící jednotka je elektronické zařízení, které čítá sumu period generovaných na oscilátoru. Ta je poté technikami šíření signálu přepočtena na tranzitní čas τ. Pro měřenou vzdálenost d poté dostáváme vztah [17]

$$2d = v\,\tau\,, \tag{4.7}$$

kde $v = c/n$ je rychlost šíření pulzu prostředím, c je rychlost světla ve vakuu a n je index lomu prostředí.

Je-li požadována milimetrová přesnost určení vzdálenosti, musí být tranzitní čas τ měřen s pikosekundovou přesností. Přibližný odhad lze provést jednoduchou úvahou – diferenciací vztahu (4.7), kde pro jednoduchost předpokládáme $n = 1$ (přibližná hodnota indexu lomu vzduchu). Po dosazení dostáváme aproximaci požadované přesnosti měření tranzitního času $\delta\tau = 2\delta d/c$. Po vyčíslení s hodnotou rychlosti světla a předpokládanou chybou vzdálenosti $\delta d = 1$ mm dostáváme $\delta t = 6.67$ ps.

Jestliže označíme hodnotu registrované amplitudy nebo počtu pulzů $I(\tau)$, můžeme odhadnout měřený tranzitní čas $\overline{\tau}$ určený z měření z registrovaných dat například pomocí těžiště (prvním momentem) podle vztahu [17]

$$\overline{\tau} = \frac{1}{P}\int_{-\infty}^{+\infty} \tau\, I(\tau)\,\mathrm{d}\tau\,, \quad P = \int_{-\infty}^{+\infty} I(\tau)\,\mathrm{d}\tau\,, \tag{4.8}$$

kde integraci provedeme v reálném případě numericky na diskrétní množině dat. Nejistotu σ_τ následně lze odhadnout jako druhou odmocninu variance σ_τ^2, pro kterou platí

$$\sigma_\tau^2 = \left(\frac{1}{P}\int_{-\infty}^{+\infty} \tau^2\, I(\tau)\,\mathrm{d}\tau\right) - \overline{\tau}^2\,. \tag{4.9}$$

Vzdálenost $d_{\overline{\tau}}$ a její nejistota σ_d odpovídající měřenému tranzitnímu času $\overline{\tau}$ a jeho nejistotě σ_τ budou následně dány vztahy

$$d_{\overline{\tau}} = \frac{v}{2}\overline{\tau}\,, \quad \sigma_d = \frac{v}{2}\sigma_\tau\,. \tag{4.10}$$

Předpokládejme dále, že čítač času je sám ovlivněn náhodnou chybou, tj. registrované hodnoty tranzitních časů jsou zatíženy náhodným šumem. Bez další znalosti čítače můžeme předpokládat, že chyba čítače respektuje rovnoměrné rozdělení pravděpodobnosti na intervalu $\Delta\tau$. Nejistota $u_{\Delta\tau}$ následně může být odhadnuta jako směrodatná odchylka rovnoměrného rozdělení a výsledná nejistota měření tranzitního času u_τ bude dána kvadratickým součtem zmíněných dílčích nejistot, platí tedy

$$u_\tau = \sqrt{\sigma_\tau^2 + u_{\Delta\tau}^2}\,, \quad u_{\Delta\tau} = \frac{\Delta\tau}{\sqrt{12}}\,. \tag{4.11}$$

Nejistotu u_d měřené vzdálenosti d lze poté odhadnout souhrnným vztahem

$$u_d = \frac{v}{2}\sqrt{\sigma_\tau^2 + u_{\Delta\tau}^2}\,. \tag{4.12}$$

Výše uvedený odhad nejistot měření tranzitního času a určení vzdálenosti může být snadno implementován do modelu určení polohy bodu například pomocí kovariančních matic při použití zákona přenášení variancí [18, 37]. Lze tak odhadnout celkový efekt zmíněných nejistot na výsledky měření.

Na závěr této kapitoly zmiňme výčet fyzikálních aspektů použitého laserového záření a jeho prostupu okolním prostředím na výsledky laserového skenování. Podrobnou analýzu může čtenář nalézt v kopii publikace [17] v Příloze A, zde se omezíme na základní charakteristiky a popis.

Fyzikální vlivy ovlivňující přesnost laserového skenování můžeme rozdělit do čtyř základních kategorií:

- vliv odrazivosti přírodních materiálů,
- vliv variace intenzity a velikosti stopy laserového svazku na skloněné odrazné ploše na přesnost měření tranzitního času,
- vliv nerovinné geometrie cíle na registrované intenzity odraženého pulzu,
- vliv nehomogenního rozložení atmosféry na šíření laserového svazku prostorem a polohu určeného bodu.

Odrazivost je optická vlastnost materiálů, která vyjadřuje poměr odražené intenzity vůči množství intenzity záření, která na daný povrch dopadá [41]. Koeficient odrazivosti, který kvantitativně tuto vlastnost charakterizuje, je závislý na materiálu (chemickém složení, struktuře, teplotě, drsnosti a barvě), na kterém k odrazu dochází a na typu dopadajícího záření (vlnová délka, polarizace apod.). Pro bezkontaktní měření, jakým laserové skenování je, je odrazivost kritická. Jestliže jsou vlastnosti povrchu takové, že se dopadající záření odráží zpět přímo do skeneru, detektor může pulz registrovat. V jiných případech může dojít k úplné ztrátě signálu a bezkontaktní metody selhávají. Odrazivost lze rozdělit do tří základních kategorií vzhledem k typu povrchu, kde k odrazu dochází:

- odraz na Lambertovském povrchu (ideálně matný, difúzní povrch), který odráží energii do všech směrů [40, 41] – tyto povrchy ve skutečnosti neexistují, ale některé materiály se svými vlastnostmi k nim blíží, např. cihlové stěny nebo školní tabule;
- odraz na zrcadlovém povrchu, kdy se dopadající svazek záření odráží dle zákona odrazu;
- odraz na lesklém povrchu, kde dochází k rozptýlení záření, ale dominantním směrem zůstává ten odpovídající zákonu odrazu – tyto povrchy jsou v praxi poměrně časté (např. mokrá vozovka) a mají negativní vliv na měření laserovým skenováním, kdy se zpět do registračního zařízení vrací jen malý zlomek záření a analýza signálu je tak velmi ztížená.

U laserových svazků je známo, že rozložení intenzity v řezu kolmo na směr jejich šíření není rovnoměrné. Základní popis laserového záření poskytují tzv. Gaussovské svazky [40, 41]. Tato charakteristika také říká, že svazek nelze v příčném směru zcela omezit a jeho stopa bude dána určitou ploškou, která je rozbíhavá se vzrůstající vzdáleností od skeneru. Velikost této stopy má poté vliv na registrovanou intenzitu, zejména jestliže se bude záření odrážet od skloněné plochy. V práci [17] je provedena podrobná analýza a jsou v ní uvedeny například vztahy pro definici požadované frekvence registrovaného tranzitního času

takového, aby čítač vliv skloněné plochy vůbec registroval. Vzhledem k tomu, že v současnosti dostupné a používané čítače času registrují s velmi vysokou frekvencí, sklon odrazné plochy může v kritických případech významným způsobem ovlivnit vypočtený tranzitní čas.

Jestliže bude docházet k měření na velké vzdálenosti, může stopa laserového svazku být už poměrně velká. Registrovaný průběh intenzity odraženého pulzu neponese informaci pouze o středu stopy, ale bude v ní zahrnuta celá plocha stopy svazku. Jestliže bude odrazná plocha skloněná nebo prostorově členěná (např. měření do rohu budovy nebo na roh budovy, na hranu stěny apod.) může být výsledek vypočtené vzdálenosti cíle od skeneru zkreslen. V práci [17] je provedena důkladná analýza problému, která je formulována na základě základních vztahů vlnové optiky, a je zde uvedena metodika, jak daný problém studovat pro libovolné geometrické parametry zkoumané situace. Na příkladě je poté ukázáno, že při cílení na rohovou budovu ve vzdálenosti 50 m může dojít ke zkreslení v řádu milimetrů, což je pro přesné práce výrazná chyba.

Posledním vlivem, který zde zmíníme, je vliv variací atmosféry, ve které laserové skenování probíhá, na výsledky měření. Šíření laserového pulzu obecně probíhá v prostředí, které je prostorově i časově nestálé. Zabývejme se zejména prostorovou nehomogenitou prostředí, kdy fyzikální parametry atmosféry jsou funkcemi prostorové polohy, jelikož ta má na geometrické určení cíle zásadní vliv (v krátkém časovém intervalu měření). Například nad silně slunečně osvětleným asfaltovým povrchem bude výrazně jiná teplota než nad vodní plochou. Obdobně variuje i atmosférický tlak, vlhkost a další parametry. Souhrnně se vlivy projeví na změnách indexu lomu prostředí, ve kterém se pulz vyslaný ze skeneru šíří. Ten se následně nepohybuje po přímce, ale po zakřivené trajektorii. Tvar této trajektorie nelze v obecném případě určit analyticky a úlohu je třeba řešit numericky. V práci [17] jsou vybrané postupy numerického trasování představeny. Vliv na laserové skenování je poté takový, že pulz putuje prostředím po delší trajektorii než po přímé, která se při geometrických přepočtech uvažuje. Tím je zaprvé ovlivněna měřená vzdálenost (registrovaný tranzitní čas je delší, než by měl být). Ukazuje se ale, že hodnoty, o které jsou rekonstruované vzdálenosti chybné, jsou zanedbatelné a pro běžné práce není třeba korekce uvažovat. Zásadní vliv má ale zakřivená dráha svazku na registrované úhly. Pulz totiž přichází po zakřivené trajektorii do skeneru ve směru, který neodpovídá skutečné poloze cíle, ale je k této trajektorii tečný. V práci [17] je na konkrétním realistickém příkladě ukázáno, že chyba v určení směru může být na vzdálenosti 100 m až 0.0035 deg (odpovídá přibližně 6 mm na 100 m), což je pro přesné práce nepřijatelné a je třeba provádět vhodné korekce měření.

5 Závěr

Tato práce souhrnně představila vybraná témata výzkumné činnosti, na kterých se autor podílel v rámci působení ve Skupině aplikované optiky na katedře fyziky Fakulty stavební ČVUT v Praze. Obsahově je práce průvodním textem k publikacím [1–21] v impaktovaných mezinárodních časopisech, jejichž kopie jsou uvedeny v Příloze A.

Nejprve bylo představeno téma analýzy aktivních optických prvků a jejich použití v zobrazovacích a měřických systémech. Vědecké práce provedené v této oblasti významným způsobem přispěly k rozvoji základních nástrojů pro modelování průběhu deformací membrán kapalinových čoček. Dále jsou například využitelné pro návrh takových parametrů membrán, pro které při zatížení odpovídajícím tlakem dojde k požadované deformaci do předepsaného tvaru. Před vydáním zmíněných publikací řada prací studovala danou problematiku, ale zpravidla s využitím zjednodušujícího aparátu, který ne zcela přesně vystihl velké deformace tenkých elastických membrán používaných v aktivních kapalinových čočkách. Publikované nové vztahy tak slouží k přesnějšímu nominálnímu návrhu a analýze, a významně tak doplňují znalosti v dané oblasti.

S použitím zmíněných nástrojů byly dále představeny výstupy publikací zabývajících se postupy návrhu a optické analýzy kombinovaných aktivních soustav, které jsou složeny z několika členů s proměnnými a fixními parametry. Jmenovitě se jedná o hybridní čočky s jedním fixním a jedním nebo dvěma aktivními prvky, zoom systémy složené z aktivních prvků, membránové čočky dvojí křivosti a optické skenery s vloženým aktivním prvkem. Tyto modely a optické soustavy nebyly do té doby představeny a analyzovány. Daná kombinovaná zařízení umožňují například kompenzovat aberace a zajistit proměnnost parametrů zobrazení s minimalizovaným mechanickým pohybem, což u klasických systémů nebylo možné.

Druhá kapitola shrnula výstupy zabývající se vybranými tématy optického zobrazení, například skalární teorií difrakce a nominálním návrhem optických soustav s korigovanými aberacemi.

Byl prezentován nový univerzální postup analýzy difrakce na dokonalé a nedokonalé mřížce a vliv konečných rozměrů mřížky na tzv. hranovou přechodovou funkci v difrakčních obrazcích v Talbotově vzdálenosti. S danými nástroji je možné například snadno analyzovat vliv výrobních nedokonalostí mřížek na jejich konkrétní aplikaci. Pro velmi rychlé použití jsou v publikacích odvozeny jednoduché analytické vztahy, pomocí kterých lze vybrané charakteristiky zobrazení kvantifikovat.

Dále práce ukázala výstupy publikací zabývajících se důkladnou analýzou rozptylové funkce bodu a osové rozptylové funkce bodu při zobrazení optickými soustavami s kruhovými aperturami a aperturami ve tvaru mezikruží. Stěžejním výstupem daných prací jsou jednoduché analytické vztahy, které slouží k charakteristice parametrů obrazů osových bodů, a významně tak přispívají svou užitečností v praktických aplikacích například

v mikroskopii.

V závěru kapitoly byly ukázány výstupy publikací, které významně rozšiřují fundamentální analýzu návrhu optických soustav s korigovanými aberacemi. Je zde prezentována velmi jednoduchá ale velmi užitečná metoda výpočtu parametrů optických ploch pro případ soustav s jednou nebo dvěma asférickými rozhraními stigmaticky zobrazující osové body. V publikacích je dále ukázán užitečný postup, pomocí kterého lze určit koeficienty aberací takovým způsobem, kdy bude rozptylová funkce bodu ve vybrané oblasti obrazového prostoru v důsledku splňovat požadované charakteristiky. Dále je také prezentována základní teoretická analýza dvoučlenných optických systémů, kdy první člen má kladnou lámavost a druhý zápornou (zejména optické systémy Petzvalova objektivu, teleobjektivu, reversního teleobjektivu a objektivů analaktického typu).

Obecně jsou výstupy druhé kapitoly, která představuje výčet z celkem devíti publikací v mezinárodních impaktovaných časopisech, hodnotným přínosem v oblasti optického zobrazení a nominálního návrhu optických soustav.

Závěrečná kapitola práce byla věnována analýze geometrické přesnosti poloh bodů určených pomocí optického (laserového) skenování.

V publikacích zabývajících se tímto tématem se autor podílel na formulaci komplexního aparátu pro nominální analýzu přesnosti, kterou lze od optického (laserového) skenování očekávat. Takto souhrnný rozbor nebyl do té doby publikován a významným způsobem tak přispěl k dané oblasti.

Pomocí představených vztahů lze snadno modelovat řadu optických soustav skenerů a analyzovat konstrukční nejistoty a jejich vliv na určovanou polohu bodů. Dále lze simulovat působení fyzikálních vlivů na měření, jako například vliv odrazivosti skloněných cílových ploch různých materiálů, šíření laserového svazku nehomogenní atmosférou a podobně.

Zmíněné publikace nejsou kompletním výčtem autorovy činnosti, ale výběrem nejvýznamnějších příspěvků v dané oblasti aplikované optiky. Mimo těchto výstupů se autor podílel například na publikaci téměř dvaceti příspěvků ve sbornících mezinárodních konferencí, více než třiceti článků v českých recenzovaných časopisech, byl více než třicetkrát citován v prestižních mezinárodních časopisech, je spoluautorem dvou užitných vzorů a dvou funkčních vzorků, vytvořil řadu výzkumných zpráv v průběhu řešení několika výzkumných grantů základního i aplikovaného výzkumu, nebo recenzí na články a odborné knihy. Kompletní výčet publikací a vědeckých výstupů, který svým rozsahem přesahuje možný rámec této práce, lze dohledat například ve veřejně dostupných databázích, kde autor figuruje.

Literatura

[1] F. Šmejkal, A. Mikš, and P. Pokorný. Design of membrane thickness for a focus-variable liquid lens with a defined shape of the membrane's outer surface. *Optik*, 244:167507, 2021.

[2] P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek. Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure. *Appl. Opt.*, 56(26):5939–5947, 2017.

[3] P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek. Deformation of a prestressed liquid lens membrane. *Appl. Opt.*, 56(34):9368–9376, 2017.

[4] A. Mikš and P. Pokorný. Hybrid liquid-membrane lenses made by one fixed and one or two active optical components. *Journal of the Optical Society of America A*, 38(1):99–107, 2021.

[5] A. Mikš and P. Pokorný. Design of zoom systems composed of lenses with variable focal length. *Applied Optics*, 59(34):10838–10845, 2020.

[6] A. Mikš and P. Pokorný. Double curvature membrane lens. *Applied Optics*, 59(32):9924–9930, 2020.

[7] P. Pokorny and A. Miks. 3D optical two-mirror scanner with focus-tunable lens. *Appl. Opt.*, 54(22):6955–6960, 2015.

[8] A. Mikš and P. Pokorný. Diffraction by perfect and imperfect amplitude grating. *Applied Optics*, 59(30):9368–9375, 2020.

[9] A. Mikš and P. Pokorný. Edge spread function of Talbot phenomenon. *Optik - International Journal for Light and Electron Optics*, 127(19):8065–8069, 2016.

[10] A. Mikš and P. Pokorný. Explicit calculation of point spread function of optical system. *Optik*, 239:166885 (pp 10), 2021.

[11] A. Mikš and P. Pokorný. Influence of circular and annular pupil function on axial point spread function of optical system. *Optik*, 230:166317 (pp 11), 2021.

[12] A. Mikš and P. Pokorný. Analytic formulas for calculation of axial defocus in microscopy. *Optik*, 201:163483 (pp 11), 2020.

[13] A. Mikš and P. Pokorný. Calculation of a lens system with one or two aspherical surfaces having corrected spherical aberration. *Journal of the Optical Society of America A*, 37(9):1390–1397, 2020.

[14] A. Mikš and P. Pokorný. Spherical aberration of an optical system and its influence on depth of focus. *Appl. Opt.*, 56(17):5099–5105, 2017.

[15] A. Mikš and P. Pokorný. Fundamental design parameters of two-component optical systems: theoretical analysis. *Applied Optics*, 59(7):1998–2003, 2020.

[16] A. Miks and P. Pokorny. Analytical expressions for the circle of confusion induced by plane-parallel plate. *Optics and Lasers in Engineering*, 50(11):1517–1521, 2012.

[17] J. Brzobohatý, F. Šmejkal, and P. Pokorný. Fundamental characterisation of a-prior

measurementaccuracy of terrestrial laser scanning. *Applied Optics*, 59(33):10243–10252, 2020.

[18] P. Pokorny. One-mirror and two-mirror three-dimensional optical scanners - position and accuracy of laser beam spot. *Appl. Opt.*, 53(12):2730–2740, 2014.

[19] P. Pokorný, F. Šmejkal, and A. Mikš. Deformation of a spherical lens caused by its own weight. *Appl. Opt.*, 56(36):9984–9991, 2017.

[20] A. Mikš and P. Pokorný. Simple method for determination of parameters of cemented doublet. *Appl. Opt.*, 55(20):5456–5458, 2016.

[21] A. Miks and P. Pokorny. Use of diffraction grating for measuring the focal length and distortion of optical systems. *Appl. Opt.*, 54(34):10200–10206, 2015.

[22] H. W. Babcock. The possibility of compensating astronomical seeing. *Publications of the Astronomical Society of the Pacific*, 65(386):229, 1953.

[23] R. Tyson. *Adaptive Optics Engineering Handbook*. Optical Science and Engineering Series. CRC Press, 1999.

[24] R. K. Tyson. *Introduction to Adaptive Optics*. SPIE tutorial texts. SPIE Press, 2000.

[25] R. K. Tyson and B. W. Frazier. *Field Guide to Adaptive Optics*. SPIE field guides. Society of Photo Optical, 2012.

[26] H. Ren and S. T. Wu. *Introduction to Adaptive Lenses*. Wiley Series in Pure and Applied Optics. Wiley, 2012.

[27] R. K. Tyson. *Principles of Adaptive Optics, Fourth Edition*. CRC Press, 2015.

[28] Optotune. `http://www.optotune.com`.

[29] Cornig Varioptic Lenses. `https://www.corning.com/worldwide/en/products/advanced-optics/product-materials/corning-varioptic-lenses.html`.

[30] OKO - Flexible optical. `http://www.okotech.com`.

[31] ALPAO. `http://www.alpao.com`.

[32] Holoeye. `http://holoeye.com`.

[33] Hamamatsu - LCOS-SLM (Optical Phase Modulator). `https://www.hamamatsu.com/eu/en/product/optical-components/lcos-slm/index.html`.

[34] Imagine eyes. `http://www.imagine-eyes.com`.

[35] Imagine optic. `http://www.imagine-optic.com`.

[36] Thorlabs - Active optical devices. `https://www.thorlabs.de/navigation.cfm?guide_id=2218`.

[37] K. Rektorys. *Survey of Applicable Mathematics*. Kluwer Academic Publishers, 1994.

[38] G. A. Korn and T. M. Korn. *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review*. Dover Civil and Mechanical Engineering. Dover Publications, 2013.

[39] G. B. Arfken, H. J. Weber, and F. E. Harris. *Mathematical Methods for Physicists: A Comprehensive Guide*. YBP Print DDA. Elsevier, 2012.

[40] A. Mikš. *Aplikovaná optika*. ČVUT v Praze, 2009.

[41] M. Born, E. Wolf, and A. B. Bhatia. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 2000.

[42] B. E. A. Saleh and M. C. Teich. *Fundamentals of Photonics*. Wiley Series in Pure and Applied Optics. Wiley, 2007.

[43] F. Träger. *Springer Handbook of Lasers and Optics*. Springer Handbooks. Springer Berlin Heidelberg, 2012.

[44] P. Yeh and C. Gu. *Optics of Liquid Crystal Displays*. Wiley Series in Pure and Applied Optics. Wiley, 2010.

[45] R. H. Chen. *Liquid Crystal Displays: Fundamental Physics and Technology*. Wiley Series in Display Technology. Wiley, 2011.

[46] M. G. Lippmann. Relations entre les phenomenes electriques et capillaries. *Ann. Chim. Phys.*, 5(494), 1875.

[47] H. A. Pohl. *Dielectrophoresis: the behavior of neutral matter in nonuniform electric fields*. Cambridge University Press, 1978.

[48] M. Bass, C. DeCusatis, J. M. Enoch, V. Lakshminarayanan, G. Li, C. MacDonald, V. N. Mahajan, and E. Van Stryland. *Handbook of Optics, Third Edition Volume I*. Handbook of Optics. McGraw-Hill Education, 2009.

[49] S. Timoshenko and S. Woinowsky-Krieger. *Theory of plates and shells*. Engineering societies monographs. McGraw-Hill, 1959.

[50] A. S. Volmir. *Flexible Plates and Shells*. Technical report. Air Force Flight Dynamics Laboratory, Research and Technology Division, Air Force Systems Command, 1967.

[51] B. Audoly and Y. Pomeau. *Elasticity and Geometry: From Hair Curls to the Non-linear Response of Shells*. OUP Oxford, 2010.

[52] M. Jirásek, F. Šmejkal, and M. Horák. Pressurized axisymmetric membrane deforming into a prescribed shape. *International Journal of Solids and Structures*, 198:1–16, 2020.

[53] A. Mikš and F. Šmejkal. Dependence of the imaging properties of the liquid lens with variable focal length on membrane thickness. *Appl. Opt.*, 57(22):6439–6445, Aug 2018.

[54] W. T. Welford. *Aberrations of Optical Systems*. CRC Press, 2017.

[55] G. G. Slyusarev. *Aberration and Optical Design Theory,*. Taylor & Francis, 1984.

[56] J. Sasián. *Introduction to Aberrations in Optical Imaging Systems*. Cambridge University Press, 2013.

[57] J. D. Gaskill and John Wiley & Sons. *Linear Systems, Fourier Transforms, and Optics*. Wiley Series in Pure and Applied Optics. Wiley, 1978.

[58] A. Papoulis. *Systems and Transforms with Applications in Optics*. McGraw-Hill series in systems science. R.E. Krieger, 1981.

[59] R. N. Bracewell and R. Bracewell. *The Fourier Transform and Its Applications*. Electrical engineering series. McGraw Hill, 2000.

[60] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill physical and quantum electronics series. W. H. Freeman, 2005.

[61] Sonin integral. `https://encyclopediaofmath.org/wiki/Sonin_integral`.

[62] G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge Mathematical Library. Cambridge University Press, 1995.

[63] B. Havelka. *Geometrical optics I, II*. NČSAV, 1956.

[64] G. F. Marshall and G. E. Stutz. *Handbook of Optical and Laser Scanning*. Optical Science and Engineering. CRC Press, 2004.

[65] G. Vosselman and H. G. Maas. *Airborne and Terrestrial Laser Scanning*. Whittles Publishing, 2010.

[66] T. Schulz. *Calibration of a terrestrial laser scanner for engineering geodesy*. PhD thesis, 2008.

A Kopie vybraných publikací autora

Publikace odkazované v Kapitole 2:

[1] Design of membrane thickness for a focus-variable liquid lens with a defined shape of the membrane's outer surface
[2] Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure
[3] Deformation of a prestressed liquid lens membrane
[4] Hybrid liquid-membrane lenses made by one fixed and one or two active optical components
[5] Design of zoom systems composed of lenses with variable focal length
[6] Double curvature membrane lens
[7] 3D optical two-mirror scanner with focus-tunable lens

Publikace odkazované v Kapitole 3:

[8] Diffraction by perfect and imperfect amplitude grating
[9] Edge spread function of Talbot phenomenon
[10] Explicit calculation of Point Spread Function of optical system
[11] Influence of circular and annular pupil function on Axial Point Spread Function of optical system
[12] Analytic formulas for calculation of axial defocus in microscopy
[13] Calculation of a lens system with one or two aspherical surfaces having corrected spherical aberration
[14] Spherical aberration of an optical system and its influence on depth of focus
[15] Fundamental design parameters of two-component optical systems: theoretical analysis
[16] Analytical expressions for the circle of confusion induced by plane-parallel plate

Publikace odkazované v Kapitole 4:

[17] Fundamental characterization of *a priori* measurement accuracy of terrestrial laser scanning
[18] One-mirror and two-mirror three-dimensional optical scanners – position and accuracy of laser beam spot

Ostatní publikace:

[19] Deformation of a spherical lens caused by its own weight
[20] Simple method for determination of parameters of cemented doublet
[21] Use of diffraction grating for measuring the focal length and distortion of optical systems

Optik - International Journal for Light and Electron Optics 244 (2021) 167507

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.com/locate/ijleo

Original research article

Design of membrane thickness for a focus-variable liquid lens with a defined shape of the membrane's outer surface

Filip Šmejkal *, Antonín Mikš, Petr Pokorný

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thakurova 6, 166 29 Prague, Czech Republic

ARTICLE INFO

Keywords:
Liquid lens
Membrane
Aberration
Thickness optimization

ABSTRACT

The paper presents an improved method for membrane liquid lens thickness optimization, such that for given value of optical liquid pressure, the membrane's outer surface is deformed to the prescribed shape. Furthermore, the influence of liquid pressure change on deformed membrane shape, as well as on imaging properties of the lens, is investigated on the example of plano-hyperbolic lens.

1. Introduction

Membrane liquid lenses have been increasingly popular topics in optics during past few years. These optical elements are usually composed of constant thickness axisymmetric membrane clamped at its edges, which encloses optical liquid filled chamber, see Fig. 1. (a). The change of volume of the optical liquid inside the lens induces membrane deformation and therefore change of imaging properties of the lens, see Fig. 1. (b).

This basic principle is in several variations usually present in almost every commercially available [1–4] or experimental [5–28] designs of these devices. The main benefit of using membranes in active optical elements is their capability to change shape by means of adjusting optical liquid pressure. However, one of the drawbacks of this property is the fact, that optical aberrations [29,30] introduced by deformed membrane shape (although they might be corrected e.g. by other optical elements) will differ for different values of liquid pressure.

Calculation of deformed shape of a constant thickness membrane actuated by hydrostatic pressure is described in detail in previous publications of the authors [31,32]. It has been discovered that to calculate accurately the deformed shape of a membrane, one can't simply accept simplifications commonly used in theory of elasticity [33,34], i.e. small deflections, small cross-sectional rotations, or small strains. One of the approaches to significantly reduce, or completely remove, (some) of the optical aberrations, is to use membrane with optimized variable thickness. This ensures that for given value of hydrostatic pressure, the membrane deforms exactly to the prescribed shape. The problem of thickness optimization to reach prescribed shape of a deformed membrane's midsurface is discussed in detail in previous paper of the authors [35].

The main goal of the present paper is to extend the aforementioned theory to allow optimization of variable thickness of membrane lens such that the prescribed shape of the membrane's outer surface is reached. The difference between shape of the midsurface and the outer surface, where the refraction actually occurs, has significant impact on imaging properties, especially for large deformations of the membrane.

In principle, the refraction on the membrane's inner surface can be neglected in the analysis because it is possible, and practical, to

* Corresponding author.
E-mail address: filip.smejkal@fsv.cvut.cz (F. Šmejkal).

https://doi.org/10.1016/j.ijleo.2021.167507
Received 23 January 2021; Received in revised form 24 May 2021; Accepted 19 June 2021
Available online 24 June 2021

use optical liquid with almost identical value of refractive index as the membrane's material (e.g., Sylgard [36] in liquid and solid form).

The second area of focus of the paper is the investigation of the influence of pressure change on deformed membrane shape and imaging properties. Prescribed shape can be achieved only for one specific value of pressure. If we change the value of pressure from the one, which was used to design an optimal thickness profile, the deformed shape changes as well, and it results in increase of aberrations of the optical system.

To our best knowledge, the derivations presented in this paper have not been published so far and the proposed formulas can significantly contribute in analysis and design procedures of membrane liquid lenses.

2. Membrane variable thickness optimization

2.1. Solution for prescribed midsurface shape

This section briefly summarizes the theory presented in authors' previous publication [35] which can be used to find an optimal thickness profile for membrane whose midsurface is supposed to deform to the prescribed shape. Suppose we have a situation illustrated in Fig. 2 [35]. We consider a circular axisymmetric membrane of a liquid lens with axis of symmetry z and radial coordinate r. In undeformed state (stress-free), the membrane is characterized by its radius a and variable thickness $h(r)$ with h so small, that its flexural stiffness can be neglected. Along its circumference, the membrane can be generally prestressed in radial direction by prescribed displacement u_a, which can be understood as a distance the membrane has to be stretched by in radial direction before clamping into the liquid lens chamber with radius $a + u_a$. After applying the pressure of optical liquid, p, the membrane deforms and general point on its midsurface with original coordinates $[r,0]$ is displaced to the new position $[r + u(r),w(r)]$, where $u(r)$ is the radial displacement and $w(r)$ is the vertical displacement (displacement in direction of axis z). The deformed shape of the membrane can be described by function $g(r)$, which can be implicitly defined as $g(r + u(r)) = g(\tilde{r}) = w(r)$. The membrane thickness h is considered to be very small compared to the vertical displacement and therefore the deformed shape is characterized by the midsurface.

To find an optimal membrane thickness profile represented by the function $h(r)$, one can apply the following equations [35].

$$A(u,r)\eta^2 + B(u,r)\eta + C(u,r) = 0 \tag{1}$$

and

$$h = \frac{p}{E}\left(-\frac{(r+u)^2}{rg'\sqrt{\eta}}\right)\frac{1}{(1+g'2)\eta - 1}\left(1+\frac{\nu\eta}{\alpha}\right), \tag{2}$$

where

$$\eta = (1+u')^2, \tag{3}$$

$$A(u,r) = 1 + g'2, \tag{4}$$

$$B(u,r) = \frac{2\nu u}{r} + \frac{\nu u^2}{r^2} - 1 + \nu\left(1 - g'2\right)\alpha(u,r), \tag{5}$$

$$C(u,r) = \left(\frac{2u}{r} + \frac{u^2}{r^2} - \nu\right)\alpha(u,r), \tag{6}$$

$$\alpha(u,r) = \frac{\left(1+\frac{u}{r}\right)^2 g'}{(r+u)g'' - 2g' - 2g'3}, \tag{7}$$

$g(\tilde{r})$ is known function representing prescribed shape of a deformed membrane midsurface, $\tilde{r} = r + u(r)$ (see Fig. 2) is the deformed radial coordinate, E and ν are the Young's modulus and Poisson's ratio of the membrane material.

It is important to note, that the first and second derivative of function g, which occurs in (2–7), is with respect to the deformed radial

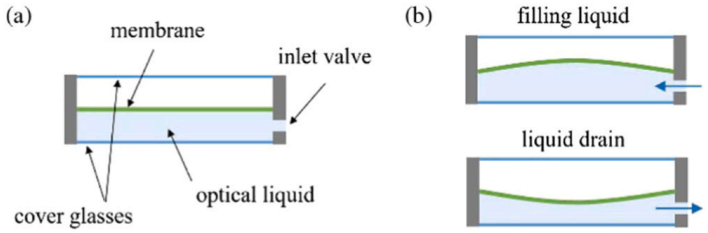


Fig. 1. (a) Schematic drawing of the membrane liquid lens and (b) result of optical liquid volume change.

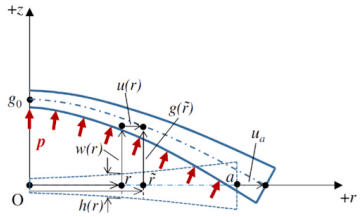


Fig. 2. Axisymmetric membrane with variable thickness actuated by constant hydrostatic pressure.

coordinate $\widetilde{r}$ (unlike the derivatives of functions $u(r)$ and $w(r)$, which are with respect to the radial coordinate r), therefore it reads $g' = \frac{dg(\widetilde{r})}{d\widetilde{r}}$ and $g'' = \frac{d^2g(\widetilde{r})}{d\widetilde{r}^2}$.

The expression (1) represents first order nonlinear differential equation with one unknown function $u(r)$, which has to satisfy two boundary conditions

$$u(0) = 0, \quad u(a) = u_a. \tag{8}$$

At the first glance, one could argue, that the problem is over constrained, since we have two boundary conditions for the first order differential equation. However, detailed analysis shows that the first boundary condition is always satisfied and therefore it does not have to be taken into account.

Solving the Eq. (1) and substituting its solution, function $u(r)$, to (2), we get function for variable thickness $h(r)$ optimized, such that membrane midsurface deforms to the prescribed shape $g(\widetilde{r})$ for given value of pressure.

Eq. (1) can't be generally analytically solved, thus it is necessary to use some of the available numerical methods to obtain the solution. This procedure is described in detail in [35], where shooting method [41] is applied to find the solution of Eq. (1). This paper further assumes, that by means of the expressions (1–8) we are able to find function of optimal variable thickness $h(r)$ based on the given input parameters p, a, u_a, E, ν and $g(\widetilde{r})$.

To further clarify the limitations of the presented theory, the following points should be noted:

- Although this theory takes into account the possibility of very large deformations (geometrical nonlinearity), it is based on the Saint Venant-Kirchhoff constitutive law [37], which postulates a linear relation between the Green-Lagrange strain and second Piola-Kirchhoff stress and is the simplest hyperelastic material law defined only by two parameters. Furthermore, presented formulas assume only static state, i.e. no dynamic actions are taken into account. This implies that the presented mechanical model can't in its current state take into account material nonlinearities like creep (deformation increasing in time while maintaining constant pressure), cycling (changing material parameters due to the repetitive loading), fatigue (material failure due to the repetitive loading before reaching its ultimate strength), viscoelastic behavior (deformation also depends on the driving speed), plasticity etc.
- In general, the membrane deformation would be also affected by the gravity (optical liquid self-weight), which is not taken into account in the presented mechanical model either. Strictly speaking, the gravity effect is always present, no matter in which orientation the lens is placed. However, with z axis positioned in vertical direction (upwards or downwards), one can imagine that the gravity effect will be less important than in case, where z axis is in horizontal position (the worst scenario). Placing the liquid lens z axis into different than vertical position and taking the liquid self-weight into account would cause non-symmetric deformation which would already require two dimensional mechanical model, because the presented mechanical model postulates only axisymmetric actions. This problem was briefly addressed in [10] where authors used commercial finite element software to investigate it. One can also conclude form works [5,10,39,40] that the effect of gravity can be neglected for large applied inner pressures. Although this is a very interesting and important topic for general analysis, it deserves its own study and it is beyond the scope of this paper.
- The presented mechanical model assumes idealized conditions at the membrane clamped edge. In fact, as mentioned earlier, it neglects the membrane flexural stiffness, which results in absence of points of inflection near to the edge of the membrane. In reality, those will be present, especially in the case of thick membranes. However, the presented model is very useful to apply and use in the case of thin membranes (usually used in optics) where the flexural stiffness is very small and points of inflection are very close to the clamped edge. And for both thick and thin membranes, while designing the optical system, one usually consider areas close to the optical axis because of optical aberrations in zones far from the axis. Another effect which affects idealized condition of the model is practical clamping of the membrane to the lens body during fabrication, as it is challenging to obtain uniform prestress.

The aforementioned limitations of the theory clearly described some challenging topics for usage of the presented model. However, for static initial analysis and design of most practically used liquid lenses the model gives satisfactory results and useful tool for analysis, as will be presented in the following parts of the paper.

2.2. Extension of the theory for membrane's outer surface

The above described theory for numerical design of the membrane midsurface using the optimization of the membrane thickness

can be successfully used for practical optical applications where small deformations occur. However, it is necessary to realize that the refraction of the light passing through the lens occurs mainly on the interface membrane-air. And this interface lays on the outer surface of the membrane and not on its midsurface, see Fig. 3 for plano-convex membrane lens. In optical practice it is therefore desirable to design variable thickness of the membrane, such that for above mentioned given input parameters, the membrane's outer surface deforms to the prescribed shape, because here, the refraction occurs. The difference between shape of the midsurface and the outer surface can have significant impact on imaging properties especially for large deformations of the membrane. Without loss of generality, the refraction on the membrane's inner surface can be neglected in the analysis because it is possible to use optical liquid with almost identical value of refractive index as the membrane material (e.g. Sylgard [36] in liquid and solid form).

To solve the aforementioned problem of designing the shape of outer membrane's surface, consider the situation in Fig. 4. Radial coordinates of the outer and inner surface of the membrane, $\tilde{r}_{out}$ and $\tilde{r}_{in}$, respectively, can be expressed as:

$$\tilde{r}_{out} = \tilde{r} - \frac{1}{2}\tilde{h}(r)\sin(\varphi), \qquad \tilde{r}_{in} = \tilde{r} + \frac{1}{2}\tilde{h}(r)\sin(\varphi) \tag{9}$$

and the corresponding vertical coordinates, g_{out} and g_{in} as:

$$g_{out} = g(\tilde{r}) + \frac{1}{2}\tilde{h}(r)\cos(\varphi), \qquad g_{in} = g(\tilde{r}) - \frac{1}{2}\tilde{h}(r)\cos(\varphi), \tag{10}$$

where φ is cross-section rotation angle, which can be further expressed as $\varphi = \arctan(g')$, therefore $\sin\varphi = \frac{g'}{\sqrt{g'2+1}}$ and $\cos\varphi = \frac{1}{\sqrt{g'2+1}}$. Here, $\tilde{h}(r)$is the notation for the membrane's thickness, deformed due to the plane stress, and for this mechanical model it can be expressed as [38].

$$\tilde{h}(r) = h(r)\sqrt{2\varepsilon_z + 1} = h(r)\sqrt{\frac{2\nu(\varepsilon_r + \varepsilon_t)}{\nu - 1} + 1}, \tag{11}$$

where $\varepsilon_r = u' + \frac{1}{2}(u'2 + w'2)$ and $\varepsilon_t = \frac{u}{r} + \frac{u^2}{2r^2}$ are the radial and tangential strains.

Similarly, as in the situation for the midsurface calculation summarized in the previous section, the procedure to find an optimal thickness profile of the membrane such that its outer surface deforms to the prescribed shape, has to be solved numerically.

Suppose we have an even function $g(\tilde{r})$, where $g(0) = 0$, which represents desired shape of the membrane liquid lens outer surface. Furthermore, the input parameters p, a, u_a, E and ν are known. To get the finite thickness at the edge of the membrane $h(a)$, one has to prescribe nonzero, positive value of u_a [35]. Let us define functions $g_i(\tilde{r}_i)$, $u_i(r)$ and $h_i(r)$, where $i = 1, 2, 3\ldots n$ corresponds to the iteration count. For $i = 1$ it follows $g_1(\tilde{r}_1) = g(\tilde{r})$ and by solving Eq. (1) and substituting its solution to (2) we get functions $u_1(r)$ and $h_1(r)$. Proceeding in the calculation we have $w_1(r) = g_1(\tilde{r}_1)$ and substituting into expressions (9–11) for the membrane's outer surface to get functions $\tilde{h}_1(r)$,$\tilde{r}_{out,1}(\tilde{r}_1, r)$ and $g_{out,1}(\tilde{r}_1, r)$. In the end of the iteration it is necessary to check the difference between the current shape of the membrane's outer surface and the prescribed shape. An arbitrary point of the membrane's outer surface in the current iteration is so far expressed by means of parameters $\tilde{r}_1$ and r as $[\tilde{r}_{out,1}(\tilde{r}_1, r), g_{out,1}(\tilde{r}_1, r)]$ and to find the above mentioned difference we need to express it by means of a single parameter $\tilde{r}_1$ as $[\tilde{r}_1, \bar{g}_{out,1}(\tilde{r}_1)]$.

Considering that the whole process is solved numerically and therefore the only function we know in analytical form is the function of prescribed shape $g(\tilde{r})$, one of the suitable ways to obtain an analytical form of function$\bar{g}_{out,1}(\tilde{r}_1)$is to use an approximation with even power series in the sense of least squares fit:

$$\bar{g}_{out,1}(\tilde{r}_1) = \sum_{j=1}^{N} A_{2j,1}\tilde{r}_1^{2j}, \tag{12}$$

where $A_{2j,1}$ are even coefficients and N is the number of coefficients.

Now, we can express deviation $\delta g_1(\tilde{r}_1)$ of the membrane's outer surface in the current iteration from the prescribed shape. For

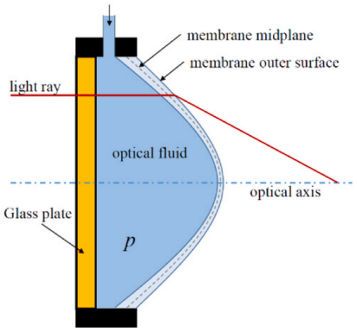


Fig. 3. Schematic drawing of light refraction on plano-convex membrane liquid lens.

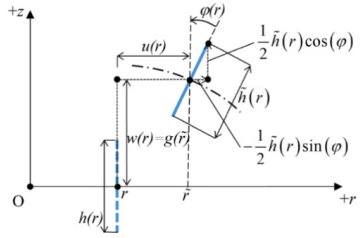


Fig. 4. Schematic drawing of membrane cross-section displacement and rotation.

$\delta g_1(0) = 0$ it yields:

$$\delta g_1(\tilde{r}_1) = g(\tilde{r}) - \overline{g}_{out,1}(\tilde{r}_1) + \overline{g}_{out,1}(0). \tag{13}$$

Thus, for input prescribed shape for the next iteration, we can write:

$$g_2(\tilde{r}) = g(\tilde{r}) + \frac{1}{2}\delta g_1(\tilde{r}_1). \tag{14}$$

For $i = 2, 3 \ldots n$, the procedure stays the same, until the condition $\delta g_n(0) \approx 0$ is satisfied with sufficient accuracy. Membrane with variable thickness $h_n(r)$ is then deformed for given input parameters, such that its outer surface exactly corresponds to the originally prescribed shape $g(\tilde{r})$.

3. Calculation of optical aberrations induced by the change of actuating pressure

Suppose the optical system of a liquid lens whose variable thickness membrane deforms exactly to the defined shape (e.g. spherical) for certain value of pressure. The change of the pressure induces the change in the vertex radius of curvature (thus change in paraxial focal length) and in general, change in the whole deformed shape (it won't be spherical anymore). If the deformed shape of the membrane was spherical even after the pressure change, wave aberration $W_{2,s}$ would be introduced to the system. However, due to the change of the membrane's shape from spherical to generally aspherical, the wave aberration $W_{2,a}$ will be introduced to the system. The difference between the aberrations is therefore expressed as $\delta W_2 = W_{2,a} - W_{2,s}$.

Let us investigate the above described situation for the case of aspherical surface, which can be expressed as

$$z = \overline{z} + \sum_{i=2}^{N} \alpha_{2i} r^{2i}, \quad \overline{z} = \frac{cr^2}{1 + \sqrt{1 - (1-\varepsilon^2)c^2 r^2}}, \tag{15}$$

where $\overline{z}$ represents the conic section of the surface, $c = 1/R$, R is the vertex radius of curvature, r is the radial coordinate (the surface is axisymmetric along the $z$axis), ε is the numerical eccentricity [41], α_{2i} are the aspherical coefficients, and N is the number of power series coefficients. Let us now approximate the conic section in Eq. (15) by Taylor series [41]. For 10th order we get:

$$\overline{z} = \frac{c}{2}r^2 - \frac{c^3(\varepsilon^2-1)}{8}r^4 + \frac{c^5(\varepsilon^2-1)^2}{16}r^6 - \frac{5c^7(\varepsilon^2-1)^3}{128}r^8 + \frac{7c^9(\varepsilon^2-1)^4}{256}r^{10}. \tag{16}$$

Maximal error $\Delta z_{\max}$ of the approximation (16) can be estimated with the following member of higher order of the Taylor series, one gets [41]:

$$\Delta z_{\max} \approx -\frac{21\, c^{11}(\varepsilon^2-1)^5}{1024} r^{12}. \tag{17}$$

In the following chapter, we will be investigating the difference between the ideal hyperboloidical shape of the membrane's outer surface that would correspond to the focal length 2 and the actual shape which developed from the ideal hyperboloidical shape optimized for the focal length 1 by changing the actuating pressure such that the paraxial focal length corresponds to the focal length 2. Because the actual shape is general (it is not an exact conic section anymore), it is necessary to approximate it by an appropriate function.

For many purposes, as will be shown in the following chapter, it is more convenient to use parabolic approximation instead of spherical approximation with aspherical coefficients expressed by formula (15). One can simply obtain the parabolic approximation of the axially symmetric optical surface, also commonly used in optical practice, by substituting $\varepsilon = 1$ into (15). The membrane shape optimized for given pressure p can then be expressed as:

$$z(p) = \sum_{i=1}^{N} A_{2i} r^{2i}. \tag{18}$$

For further analysis, suppose that the value of actuating pressure p will be changed to p_1, therefore the shape is to become not ideal (as it was optimized for the pressure p). Then the corresponding shape of the membrane surface changes from (18) to:

$$z_1(p_1) = \sum_{i=1}^{N} B_{2i} r^{2i}, \tag{19}$$

Where B_{2i} are the coefficients corresponding to the parabolic approximation of the surface that developed from (18) by changing the actuating pressure from p to p_1. Considering that the optimal shape of the membrane for the pressure p_1 would be given by:

$$z_0(p_1) = \sum_{i=1}^{N} C_{2i} r^{2i}, \tag{20}$$

where C_{2i} are (in analogy to A_{2i}) the coefficients corresponding to the parabolic approximation of the optimal shape of the membrane for the actuating pressure p_1, the difference between the real and ideal shape of the membrane surface for pressure p_1 can be therefore expressed as:

$$\delta z(p_1) = \sum_{i=1}^{N} \left(B_{2i} - C_{2i}\right) r^{2i} = \sum_{i=1}^{N} D_{2i} r^{2i}. \tag{21}$$

Ideally, the coefficients D_{2i} would be zero. However, this scenario can't occur, because the optimal thickness is designed only for one specific value of pressure. It is expected, that the more the actuating pressure differs from the value for which the optimal thickness was designed, the more the resulting deformed shape differs from the one, which was originally prescribed (spherical, parabolic, hyperbolical etc.). As a result of this deviation, the membrane introduces certain wave aberration [29,30], which depends on the actuating pressure.

Without the loss of generality, the refraction on the membrane's inner surface can be neglected in the analysis because it is possible to use optical liquid with almost identical value of refractive index as the membrane material (e.g. Sylgard [36] in liquid and solid form). Afterwards, wave aberration $\delta W(p_1)$ introduced to the system due to the difference of the real and the ideal surface $\delta z(p_1)$ can be approximately expressed as [29,30].

$$\delta W(p_1) = \left(n' \cos\varepsilon' - n \cos\varepsilon\right) \delta z(p_1) = \left(n' \cos\varepsilon' - n \cos\varepsilon\right) \sum_{i=1}^{N} D_{2i} r^{2i} \approx \left(n' - n\right) \sum_{i=1}^{N} D_{2i} r^{2i}, \tag{22}$$

wheren', n are the refractive indexes of the substance outside and inside of the liquid lens, and ε, ε' are the angles of incident and refracted ray on the membrane's outer surface. Expression (22) is sufficiently accurate for our purpose, but reader can find more accurate formulas in [42,43].

By means of the first derivative with respect to r of the expression (22), we get transverse ray aberration $\delta y'$ [29,30].

$$\delta y' \approx 2R' \left(n' - n\right) \sum_{i=1}^{N} i D_{2i} r^{2i-1}, \tag{23}$$

whereR'denotes the radius of the reference sphere in the image space.

Denoting the transverse ray aberration$\delta y'_k$, which corresponds to the value of radial coordinate $r = r_k$, we can estimate the radius ρ of the circle of confusion as [29,30].

$$\rho = \sqrt{\frac{1}{K} \sum_{k=1}^{K} (\delta y'_k - \delta \bar{y}')^2}, \quad \delta \bar{y}' = \frac{1}{K} \sum_{k=1}^{K} \delta y'_k, \tag{24}$$

where K is the number of rays.

4. Examples

4.1. Example 1

The first example demonstrates the effect of difference between the outer surface and midsurface of the membrane, which are considered for a membrane's thickness and shape optimization, on imaging properties of the focus-variable liquid membrane lens.

Suppose that the lens is composed of a glass plate and a hyperboloidical membrane surface filled with optical liquid, see Fig. 3. From geometrical optics [30] it is known, that axisymmetric plano-hyperboloidic lens is spherical aberration free for the light beam parallel to its optical axis and entering from the side of its planar surface. Formula (25) expresses the general equation of the second-order curve (conic section):

$$r^2 = 2Rz + \left(\varepsilon^2 - 1\right) z^2, \tag{25}$$

where R is the vertex radius of curvature. From geometrical optics, for this hyperboloidical lens it further yields:

$$R = -f'(n-1), \qquad \varepsilon = n. \tag{26}$$

Substituting (26) into (25), expressing the variable z and replacing radial coordinate r for $\widetilde{r}$(Section 2), we get for the prescribed hyperboloidic shape of the outer surface of the membrane:

$$z(\widetilde{r}) = \frac{\widetilde{r}^2}{f'(1-n)\left(1+\sqrt{1-\frac{(1+n)\widetilde{r}^2}{f'2(1-n)}}\right)}. \tag{27}$$

Suppose next that the membrane radius after introducing prestress is $a_p = 10$ mm, prestressing displacement is $u_a = 1$ mm and therefore the radius of the membrane before introducing prestress (before clamping into the lens chamber) is $a = a_p\text{-}u_a = 9$ mm. The membrane is made of the material Sylgard 184 [36], whose Young's modulus and Poisson's ratio is assumed to be $E = 1.97$ MPa and $\nu = 0.4$. It is further demanded that for actuating pressure $p = 0.001E = 1.97$ kPa, the paraxial focal length is $f' = 100$ mm. The refractive index of the optical liquid in this example is assumed to be $n = 1.4118$, which corresponds to the material Sylgard 184 in liquid form for the wavelength 633 nm. This kind of lens can be widely used in the field of optical scanners [44], for example.

In the next part of this example it is convenient to use parabolic approximation of the expression (27). Let us investigate, how the exact formula for hyperbola $z(\widetilde{r})$ (27) differs for this specific example from its approximation with Taylor series $\overline{z}(\widetilde{r})$ (16) presented in the previous chapter. Using the formula (17) one gets the approximate value of maximal error of this approximation, $\Delta z_{\max} \approx 3.4 \times 10^{-8}$ mm, which tells us that the approximation accuracy is sufficient.

To demonstrate the effect of difference between the outer surface and midsurface on optical imaging properties, consider the following scenario. Let the membrane thickness is optimized for the desired hyperbolic shape of the midsurface, as was presented in previous publication of the authors [35]. Therefore, the paraxial focal plane is supposed to lay at a distance behind the vertex of the midsurface which corresponds to desired focal length f'. Let z denotes coordinate behind the membrane in a direction of ray propagation. Further, let z_m is the coordinate of the maximal deformation of the midsurface (i.e., $z_m = 0$ for zero applied pressure), then the paraxial focal plane will intersect the optical axis at coordinate $z_{Fm} = z_m + f'$. However, as the difference between optimized midsurface and real outer surface of the membrane is neglected, the outer shape of the membrane will differ from supposed hyperbolic shape, and it will affect imaging quality in the image (detection) plane. Fig. 5 shows transverse ray aberration $\delta y'$ in the detection plane of z-coordinate $z_d = z_{Fm}$ for the supposed situation as a function of impinging height y on the planar surface of the lens. Fig. 6 then shows transverse ray aberration in the detection plane where the gyration radius ρ (the radius of the circle of confusion), for the real imaging by the outer surface of the supposed lens, is minimal. In the title of Fig. 6, z_o is the vertex coordinate of the outer surface of the membrane, z_{Fo} denotes coordinate of the effective focal point, $\Delta s = z_{Fo} - z_{Fm}$, s'_{Fo} is the axial distance from the outer surface of the membrane to the detection plane, and ρ is the value of the gyration radius.

From the presented figures, one can simply see that neglecting the difference between the midsurface and the outer surface of the membrane during the thickness optimization can result in enormous errors in optical imaging. Therefore, the proposed optimization of the outer membrane's shape, presented in this paper, has to be performed for correct designing of the membrane liquid lenses.

4.2. Example 2

This example demonstrates the influence of pressure change in liquid membrane lens on deformed membrane shape and imaging properties of the lens. Let the nominal value of pressure, the one the lens membrane thickness and the shape of outer surface was optimized for, is p, and it varies.

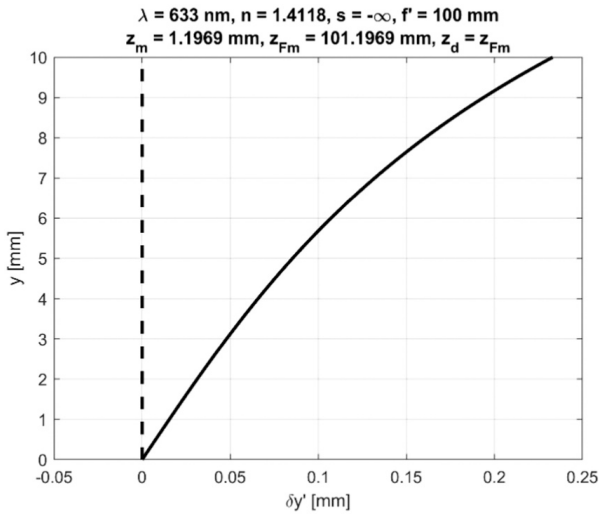


Fig. 5. Difference in transverse ray aberration in the paraxial focal plane for the case of optimized hyperbolic midsurface of the membrane lens (dashed line – supposed compensated aberration for hyperbolic shape of the midsurface, solid line – real aberration induced by neglecting the effect of membrane thickness and incorrect shape of the outer surface).

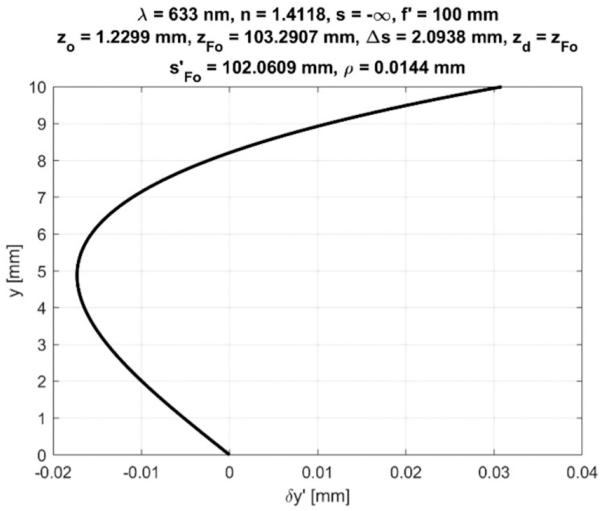


Fig. 6. Real transverse ray aberration in the plane with minimal radius of circle of confusion ρ.

Consider the material and geometrical parameters the same as in the previous example. If we change the value of the pressure from $p = 1.97$ kPa to $p_1 = 1.5249$ kPa, the paraxial focal length changes from $f' = 100$ mm to $f'_1 = 125$ mm, and the vertex radius of curvature from $R = -41.18$ mm to $R_1 = -51.475$ mm. If we further change the pressure to $p_2 = 1.2502$ kPa, the paraxial focal length changes to $f'_2 = 150$ mm and the vertex radius of curvature to $R_2 = -61.77$ mm. The deformed shape of the membrane after the change of actuating pressure can be determined by minimizing of the total potential energy by means of using various numerical methods. This procedure is described in detail in [31,32,38]. For the purpose of this paper, the power series method [41] combined with optimization of its coefficients is applied to minimize the total potential energy. The corresponding shapes of the membrane's outer surface $z_{1,1}(\tilde{r})$ and $z_{1,2}(\tilde{r})$ (for actuating pressures p_1 and p_2, respectively) can be described using the formula (19), for $N = 5$ it yields

$$z_1(\tilde{r}) = B_2\tilde{r}^2 + B_4\tilde{r}^4 + B_6\tilde{r}^6 + B_8\tilde{r}^8 + B_{10}\tilde{r}^{10}, \tag{28}$$

where the values of the coefficients B_{2i} for the functions $z_{1,1}(\tilde{r})$ and $z_{1,2}(\tilde{r})$ are specified in Table 1.

To assess the difference between the functions $z_{1,1}(\tilde{r})$, $z_{1,2}(\tilde{r})$ and the corresponding ideal hyperbolical shapes for given values of paraxial focal lengths f'_1 and f'_2, it is necessary to express the ideal shapes in the form of the power series as well. The approximation of the general conic section with Taylor series (16) can be applied for this purpose, thus for functions $\bar{z}_{0,1}(\tilde{r})$ and $\bar{z}_{0,2}(\tilde{r})$ we can write:

$$\bar{z}_0(\tilde{r}) = C_2\tilde{r}^2 + C_4\tilde{r}^4 + C_6\tilde{r}^6 + C_8\tilde{r}^8 + C_{10}\tilde{r}^{10}, \tag{29}$$

where the values of the coefficients $C_2 = \frac{1}{2R}$, $C_4 = -\frac{\varepsilon^2-1}{8R^3}$, $C_6 = \frac{(\varepsilon^2-1)^2}{16R^5}$, $C_8 = -\frac{5(\varepsilon^2-1)^3}{128R^7}$ and $C_{10} = \frac{7(\varepsilon^2-1)^4}{256R^9}$ after substituting (26) for focal lengths f'_1 and f'_2 are specified in Table 2.

Using the values of coefficients specified in Tables 1 and 2 we can proceed to the calculation of the diameter of the circle of confusion [30] for the given focal lengths of the membrane liquid lens $f' = 100$ mm, $f'_1 = 125$ mm and $f'_2 = 150$ mm. The calculation was performed using approximate formulas (21–26) and compared with the results obtained from OSLO software [45]. Fig. 7 shows the dependence of the focal length on the diameter of circle of confusion generated by liquid lens due to the change of the shape of the membrane's outer surface from the ideal hyperboloid to the general axisymmetric aspherical surface. The calculation is performed for two values of the diameter of the light beam entering the lens from the side of its planar surface, $D_1 = 15$ mm and $D_2 = 20$ mm.

4.3. Example 3

Suppose we have membrane liquid lens with identical parameters as in example 1. The membrane is, again, made of the material

Table 1
Values of coefficients B_{2i}.

Coeff.	Function	
	$z_{1,1}(\tilde{r})$	$z_{1,2}(\tilde{r})$
B_2	$-9.712788414425 \times 10^{-3}$	$-8.094344629507 \times 10^{-3}$
B_4	$2.305369478 \times 10^{-6}$	$2.411960994 \times 10^{-6}$
B_6	2.963174×10^{-9}	2.310320×10^{-9}
B_8	3.6524×10^{-11}	-3.0162×10^{-11}
B_{10}	1.43×10^{-13}	1.19×10^{-13}

Table 2
Values of coefficients C_{2i}.

Coeff.	Function	
	$\bar{z}_{0,1}(\bar{r})$	$\bar{z}_{0,2}(\bar{r})$
C_2	$-9.713453132589\times10^{-3}$	$-8.094544277157\times10^{-3}$
C_4	9.10224624×10^{-7}	5.26750361×10^{-7}
C_6	-1.70590×10^{-10}	-6.8556×10^{-11}
C_8	4.0×10^{-14}	1.1×10^{-14}
C_{10}	0	0

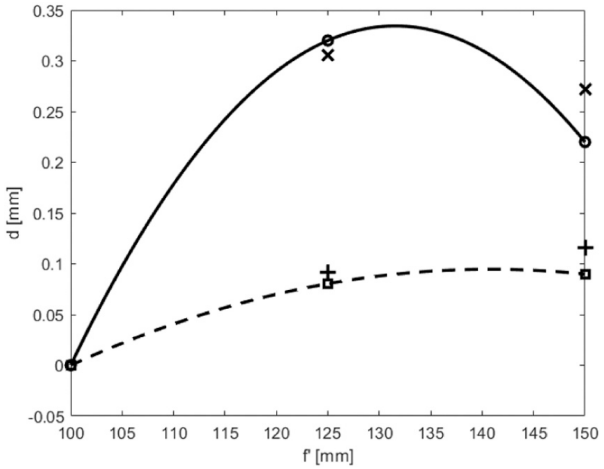


Fig. 7. Dependence of the diameter*d*of the circle of confusion on the focal length (solid – parabolic fit of values obtained from OSLO for $D_2 = 20$ mm, circle - values obtained from OSLO for $D_2 = 20$ mm, cross – values obtained from approximate formulas for $D_2 = 20$ mm, dashed – parabolic fit of values obtained from OSLO for $D_1 = 15$ mm, square - values obtained from OSLO for $D_1 = 15$ mm, plus sign – values obtained from approximate formulas for $D_1 = 15$ mm).

Sylgard 184 [36], whose refractive index is $n_1 = 1.4118$. However, this time the optical liquid inside the lens is distilled water with refractive index $n = 1.3318$. Design of the desired shape of the membrane's outer surface for focal length $f' = 100$ mm is again performed using formula (27) and the thickness profile for this specific shape for actuating pressure $p = 1.97$ kPa is determined by means of using an iterative procedure derived in Section 2. In this example, all the surfaces will be expressed by means of formula (19) for $N = 5$. Coefficients B_{2i} for the membrane's inner and outer surface, $z_{0,i}(\bar{r})$ and $z_{0,o}(\bar{r})$ for the focal length $f' = 100$ mm are specified in Table 3.

If we change the value of the pressure from $p = 1.97$ kPa to $p_1 = 1.4975$ kPa, the paraxial focal length changes from $f' = 100$ mm to $f'_1 = 125$ mm and the vertex radius of curvature from $R = -33.181$ mm to $R_1 = -41.475$ mm. If we further change the pressure to $p_2 = 1.2156$ kPa, the paraxial focal length changes to $f'_2 = 150$ mm and the vertex radius of curvature to $R_2 = -49.77$ mm. Coefficients B_{2i} for the membrane's inner and outer surface, $z_{1,i}(\bar{r})$ and $z_{1,o}(\bar{r})$ and $z_{2,i}(\bar{r})$ and $z_{2,o}(\bar{r})$ for the focal lengths $f'_1 = 125$ mm and $f'_2 = 150$ mm are specified in Tables 4 and 5.

Using the values of coefficients specified in Tables 3–5 we can proceed to the calculation of the diameter of the circle of confusion for the given focal lengths of the membrane liquid lens $f' = 100$ mm, $f'_1 = 125$ mm and $f'_2 = 150$ mm. The results are obtained from OSLO software.

Fig. 8 shows the dependence of the diameter d of the circle of confusion on the focal length, which was generated by liquid lens due to the change of the shape of the membrane's outer surface from the ideal hyperboloid to the general axisymmetric aspherical surface. The calculation is, again, performed for two values of the diameter of the light beam entering the lens from the side of its planar surface, $D_1 = 15$ mm and $D_2 = 20$ mm.

From Figs. 7 and 8 it is clear, that if we change the actuating pressure from its original value, for which the optimal thickness profile of the membrane was designed and for which the membrane's outer surface deforms exactly to the hyperboloidic shape, the deformed shape changes to the general axisymmetric aspherical surface. Consequently, the optical aberrations introduced to the system by the membrane increase. Considering that example 2 represents more complex lens, which composes of more optical substances (refractive indexes of distilled water and membrane differ) and one of its surfaces has hyperboloidic shape, it can be expected, that the value of the diameter of circle of confusion won't intersect zero even for the original focal length $f' = 100$ mm, see Fig. 8.

5. Conclusions

The paper presents the problem of optimal design of the variable thickness profile of the membrane in membrane liquid lens.

Table 3
Values of coefficients B_{2i} of membrane's inner and outer surface for $f' = 100$ mm, membrane's vertex deformed thickness $\widetilde{h}(0) = 46.5 \times 10^{-3}$ mm.

	Function	
Coeff.	$z_{0,i}(\widetilde{r})$ - inner	$z_{0,o}(\widetilde{r})$ – outer
B_2	-1.5427681909898 $\times 10^{-2}$	-1.5069111713259 $\times 10^{-2}$
B_4	1.796443969 $\times 10^{-6}$	2.633718250 $\times 10^{-6}$
B_6	-1.876658 $\times 10^{-9}$	-5.72136 $\times 10^{-10}$
B_8	-4.836 $\times 10^{-12}$	-3.528 $\times 10^{-12}$
B_{10}	1.2 $\times 10^{-14}$	1.5 $\times 10^{-14}$

Table 4
Values of coefficients B_{2i} of membrane's inner and outer surface for $f'_1 = 125$ mm, membrane's vertex deformed thickness $\widetilde{h}(0) = 47.4 \times 10^{-3}$ mm.

	Function	
Coeff.	$z_{1,i}(\widetilde{r})$ - inner	$z_{1,o}(\widetilde{r})$ – outer
B_2	-1.2406949245625 $\times 10^{-2}$	-1.2055457579851 $\times 10^{-2}$
B_4	3.236978306 $\times 10^{-6}$	3.982282655 $\times 10^{-6}$
B_6	-1.629885 $\times 10^{-9}$	-6.18083 $\times 10^{-10}$
B_8	-2.080 $\times 10^{-12}$	-1.228 $\times 10^{-12}$
B_{10}	1.0 $\times 10^{-14}$	1.1 $\times 10^{-14}$

Table 5
Values of coefficients B_{2i} of membrane's inner and outer surface for $f'_2 = 150$ mm, membrane's vertex deformed thickness $\widetilde{h}(0) = 47.85 \times 10^{-3}$ mm.

	Function	
Coeff.	$z_{2,i}(\widetilde{r})$ - inner	$z_{2,o}(\widetilde{r})$ – outer
B_2	-1.0393807675511 $\times 10^{-2}$	-1.0046215531050 $\times 10^{-2}$
B_4	3.519069108 $\times 10^{-6}$	4.216893073 $\times 10^{-6}$
B_6	-1.820727 $\times 10^{-9}$	-9.25025 $\times 10^{-10}$
B_8	-9.09 $\times 10^{-13}$	-2.12 $\times 10^{-13}$
B_{10}	0.7 $\times 10^{-14}$	0.8 $\times 10^{-14}$

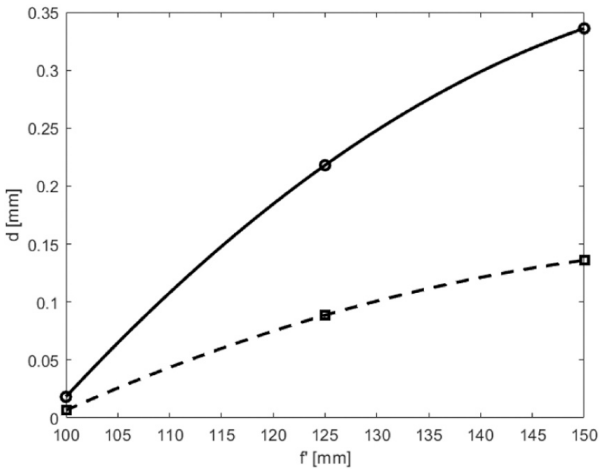


Fig. 8. Dependence of the diameter d of the circle of confusion on the focal length (solid – parabolic fit of values obtained from OSLO for $D_2 = 20$ mm, circle - values obtained from OSLO for $D_2 = 20$ mm, dashed – parabolic fit of values obtained from OSLO for $D_1 = 15$ mm, square - values obtained from OSLO for $D_1 = 15$ mm).

Firstly, the general formulas allowing us to design variable thickness profile, such that for specific value of actuating pressure the membrane's midsurface deforms exactly to the prescribed shape, are introduced. Secondly, an iterative procedure, which can be applied to obtain variable thickness profile, such that the membrane's outer surface deforms to the prescribed shape, is derived.

Furthermore, the influence of the pressure change on the deformed membrane shape and on its optical aberrations is investigated. In the end, the example for a specific plano-hyperbolic lens demonstrates the effect of difference between the outer surface and mid-surface of the membrane, which are considered for a membrane's thickness and shape optimization, on imaging properties of the focus-variable liquid membrane lens. One can see that neglecting the difference between the midsurface and outer surface can result in enormous errors in optical imaging. Further, two examples that show how the diameter of circle of confusion of the liquid lens depends on the focal length when the actuating pressure is changed from the value for which the optimal thickness was designed, are presented.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS20/093/OHK1/2T/11.

References

[1] Varioptic, ⟨http://www.varioptic.com⟩.
[2] Optotune, ⟨http://www.optotune.com⟩.
[3] Flexible optical, ⟨http://www.okotech.com⟩.
[4] TAG optics, ⟨http://www.tag-optics.com⟩.
[5] N. Sugiura, S. Morita, Variable-focus liquid-filled optical lens, Appl. Opt. 32 (1993) 4181–4186.
[6] A.H. Rawicz, I. Mikhailenko, Modeling a variable-focus liquid-filled optical lens, Appl. Opt. 35 (1996) 1587–1589.
[7] D.Y. Zhang, N. Justis, V. Lien, Y. Berdichevsky, Yu-Hwa Lo, High-performance fluidic adaptive lenses, Appl. Opt. 43 (2004) 783–787.
[8] H. Ren, S.-T. Wu, Variable-focus liquid lens, Opt. Express 15 (2007) 5931–5936.
[9] L. Wang, H. OkuM. Ishikawa, Development of Variable-Focus Lens with Liquid-Membrane-Liquid Structure and 30 mm Optical Aperture, 2013.
[10] S.T. Choi, B.S. Son, G.W. Seo, S.-Y. Park, K.-S. Lee, Optomechanical analysis of nonlinear elastomer membrane deformation under hydraulic pressure for variable-focus liquid-filled microlenses, Opt. Express 22 (2014) 6133–6146.
[11] Q. Yang, P. Kobrin, C. Seabury, S. Narayanaswamy, W. Christian, Mechanical modeling of fluid-driven polymer lenses, Appl. Opt. 47 (2008) 3658–3668.
[12] L. Wang, H. Oku, M. Ishikawa, An improved low-optical-power variable focus lens with a large aperture, Opt. Express 22 (2014) 19448–19456.
[13] D. Shaw, C.-W. Lin, Design and analysis of an asymmetrical liquidfilled lens, Opt. Eng. 46 (2007), 123002.
[14] N. Hasan, A. Banerjee, H. Kim, C.H. Mastrangelo, Tunable-focus lens for adaptive eyeglasses, Opt. Express 25 (2017) 1221–1233.
[15] H. Choi, D.S. Han, Y.H. Won, Adaptive double-sided fluidic lens of polydimethylsiloxane membranes of matching thickness, Opt. Lett. 36 (2011) 4701–4703.
[16] L. Li, Q.-H. Wang, W. Jiang, Liquid lens with double tunable surfaces for large power tunability and improved optical performance, J. Opt. 13 (2011), 115503.
[17] Y.-K. Fuh, M.-X. Lin, S. Lee, Characterizing aberration of a pressure-actuated tunable biconvex microlens with a simple spherically corrected design, Opt. Lasers Eng. 50 (2012) 1677–1682.
[18] D. Liang, X.-Y. Wang, A bio-inspired optical system with a polymer membrane and integrated structure, Bioinspir. Biomim. 11 (2016), 066008.
[19] J.-W. Du, X.-Y. Wang, D. Liang, Bionic optical imaging system with aspheric solid–liquid mixed variable-focus lens, Opt. Eng. 55 (2016), 023105.
[20] K. Mishra, C. Murade, B. Carreel, I. Roghair, J.M. Oh, G. Manukyan, D. Van, Optofluidic lens with tunable focal length and asphericity, Sci. Rep. 4 (2014) 6378.
[21] D. Shaw, T.E. Sun, Optical properties of variable-focus liquid-filled optical lenses with different membrane shapes, Opt. Eng. 46 (2007), 024002.
[22] H. Yu, G. Zhou, H.M. Leung, F.S. Chau, Tunable liquid-filled lens integrated with aspherical surface for spherical aberration compensation, Opt. Express 18 (2010) 9945–9954.
[23] G.-H. Feng, Y.-C. Chou, Flexible meniscus/biconvex lens system with fluidic-controlled tunable-focus applications, Appl. Opt. 48 (2009) 3284–3290.
[24] F. Schneider, J. Draheim, R. Kamberger, P. Waibel, U. Wallrabe, Optical characterization of adaptive fluidic silicone-membrane lenses, Opt. Express 17 (2009) 11813–11821.
[25] A. Santiago-Alvarado, J. González-Garcíaa, F. Itubide-Jiméneza, M. Campos-Garcíab, V. Cruz-Martineza, P. Raffertyc, Simulating the functioning of variable focus length liquid-filled lenses using the finite element method (fem), Optik 124 (2013) 1677–1682.
[26] H. Huang, K. Wei, Q. Wang, Y. Zhao, Improved optical resolution for elastomer-liquid lens at high diopter using varied thickness membrane, in: Proceedings of the SPIE 9705, Microfluidics, BioMEMS, and Medical Microsystems XIV, 24 March, 2016, 970504.
[27] Z. Ding, C. Wang, Z. Hu, Z. Cao, Z. Zhou, X. Chen, H. Chen, W. Qiao, Surface profiling of an aspherical liquid lens with a varied thickness membrane, Opt. Express 25 (2017) 3122–3132.
[28] P. Zhao, A. Çaglar, H. Zappe, Spherical aberration free liquidfilled tunable lens with variable thickness membrane, Opt. Express 23 (2015) 21264–21278.
[29] M. Born, E. Wolf. Principles of Optics, seventh ed., Cambridge University Press, 1999.
[30] W.T. Welford, Aberrations of the Symmetrical Optical Systems, Academic Press, London, 1974.
[31] P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, M. Jirásek, Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure, Appl. Opt. 56 (2017) 5939–5947.
[32] P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, M. Jirásek, Deformation of a prestressed liquid lens membrane, Appl. Opt. 56 (2017) 9368–9376.
[33] S. Timoshenko, S. Woinowsky-Krieger, Theory of plates and shells. Engineering Societies Monographs, McGraw-Hill, 1959.
[34] A.S. Volmir, Flexible Plates and Shells, Technical Report, Air Force Flight Dynamics Laboratory, Research and Technology Division, Air Force Systems Command, 1967.
[35] M. Jirásek, F. Šmejkal, M. Horák, Pressurized axisymmetric membrane deforming into a prescribed shape, Int. J. Solids Struct. 198 (2020) 1–16.
[36] Sylgard 184 Silicone Elastomer Kit. ⟨http://www.dowcorning.com/applications/search/products/Details.aspx?prod=01064291⟩.
[37] B. Audoly, Y. Pomeau, Elasticity and Geometry: from Hair Curls to the Non-linear Response of Shells, OUP, Oxford, 2010.
[38] A. Mikš, F. Šmejkal, Dependence of the imaging properties of the liquid lens with variable focal length on membrane thickness, Appl. Opt. 57 (2018) 6439–6445.
[39] Kevin Newman, Kyle Stephens, Analysis of gravitational effects on liquid lenses (ANGEL), in: Proceedings of the SPIE 8450, Modern Technologies in Space- and Ground-based Telescopes and Instrumentation II, 84500G,13 September, 2012. ⟨https://doi.org/10.1117/12.926520⟩.
[40] Pengpeng Zhao, Çağlar Ataman, Hans Zappe, Gravity-immune liquid-filled tunable lens with reduced spherical aberration, Appl. Opt. 55 (2016) 7816–7823, https://doi.org/10.1364/AO.55.007816.
[41] K. Rektorys, Survey of Applicable Mathematics, Springer, 1994.
[42] M. Rimmer, Analysis of perturbed lens systems, Appl. Opt. 9 (1970) 533–537.

[43] H.H. Hopkins, H.J. Tiziani, A theoretical and experimental study of lens centring errors and their influence on optical image quality, Br. J. Appl. Phys. 17 (1966) 33–54.
[44] P. Pokorny, A. Miks, 3D optical two-mirror scanner with focus-tunable lens, Appl. Opt. 54 (2015) 6955–6960.
[45] OSLO, Lambda Research Corporation, Littleton, USA, 2020. ⟨http://www.lambdares.com/oslo⟩.

applied optics

Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure

Petr Pokorný,[1,*] Filip Šmejkal,[1] Pavel Kulmon,[1] Pavel Novák,[1] Jiří Novák,[1] Antonín Mikš,[1] Martin Horák,[2] and Milan Jirásek[2]

[1]Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague, Czech Republic
[2]Czech Technical University in Prague, Faculty of Civil Engineering, Department of Mechanics, Thákurova 7, 166 29 Prague, Czech Republic
*Corresponding author: petr.pokorny@fsv.cvut.cz

Received 25 April 2017; revised 28 June 2017; accepted 28 June 2017; posted 29 June 2017 (Doc. ID 294575); published 18 July 2017

The paper discusses a numerical calculation of deformation of a circular axisymmetric membrane of a liquid lens caused by the pressure of an optical liquid. Since such deflections of the membrane are many times larger than the membrane thickness, a nonlinear model is applied and generalized relationships are derived that characterize the resulting shape with a high precision and permit an accurate analysis of imaging properties of the lens and of optical aberrations. By comparison with experimental data, it is shown that the presented model is suitable to describe the deformation of the membrane of the lens.

OCIS codes: (120.4640) Optical instruments; (220.1080) Active or adaptive optics; (350.4600) Optical engineering.

https://doi.org/10.1364/AO.56.005939

1. INTRODUCTION

In recent years, a rapid development and application of so-called active lenses with tunable parameters have been noticed [1]. Such lenses can change their internal or external parameters in a predefined way and correspondingly modify the generated optical image. There are many ways that the lens parameters can be varied. One of the most widespread types are membrane liquid lenses [1–23]. At present, some of these optical elements are produced commercially [2,3] and applied in various imaging applications. The fundamental component of the lens is a membrane clamped at its edges, covering a chamber with an optical liquid, as is schematically shown in Fig. 1(a). By changing the volume of the liquid, the shape of the membrane can be changed and the optical properties of the lens can be controlled [Fig. 1(b)]. Various authors analyzed both theoretically (numerically) and experimentally the deflection of membrane elements under uniform fluid pressure incorporated in different types of membrane fluidic lenses using different mechanical models of elastic membranes [4–23]. The accurate modeling of the shape of the deflected membrane of a fluidic lens is crucial for optical properties of such lenses.

To precisely predict imaging properties of membrane liquid lenses, one has to know the shape of the membrane surface after deformation very accurately. Considering large deformations, compared to the membrane thickness, a classical linear theory [24,25] typically used in mechanics or civil engineering does not provide sufficient accuracy. Large deformations and stresses in plates were studied by Hencky [26] and Chein [27]. Campbell [28] generalized Hencky's solution to the case with initial stress. Numerical procedures based on iterative calculations of derived differential equations and series solutions were presented by Goldberg and Pifko [29,30]. A modified method of finite differences for a system of nonlinear differential equations describing strains and stresses was presented by Kao and Perrone [31,32]. Pettit solved a system of three differential equations using the Runge–Kutta method of the fourth order [33]. A solution based on stress determination by the Newton–Raphson method was carried out by Kelkar *et al.* [34]. A solution based on series and analytical expressions for the series coefficients was published by Fichter [35]. Allman presented examples of variational solutions for the nonlinear deflection of an annular membrane [36]. Shepload and Dugundji presented solutions for clamped circular plates under initial tension with transition to membrane behavior [37]. Zhao [38] showed nonlinear models for prestretched and postheated membrane solved by both the Ritz method and the Galerkin method. Stanford and Ifju studied the validity range of low-fidelity structural membrane models [39]. A generalized solution for large deflections was derived by Mikš and Novák [40] with the use of optimization algorithms.

In the aforementioned works, the authors built their solutions on various assumptions and boundary conditions, and therefore the nonlinear differential equations and their solutions are different. This paper presents an approach based on precise

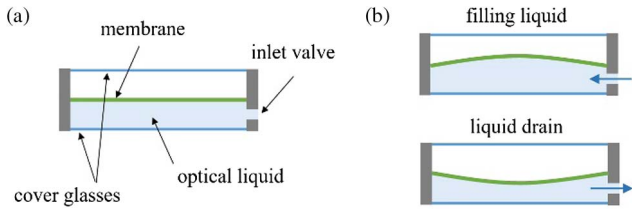


Fig. 1. (a) Scheme of liquid membrane lens and (b) result of a change of liquid volume.

geometrical equations and the Saint Venant–Kirchhoff material model, which postulates a linear relationship between the Green–Lagrange strain and the second Piola–Kirchoff stress [41]. Since the problem does not have an exact analytical solution, optimization algorithms [42,43] are applied to series expansions. The presented derived model is illustrated by an example and verified by a laboratory experiment. It is shown that such an approach gives very accurate results, and it can be used for an accurate optical design of optical systems with incorporated membrane liquid lenses because it permits the analysis of their imaging properties.

2. VARIATIONALLY CONSISTENT DERIVATION OF GOVERNING EQUATIONS FOR LARGE DEFORMATIONS OF MEMBRANE OF LIQUID LENSES

Consider a membrane of a liquid lens to be circular and axisymmetrical around axis z, which intersects the center of the lens as is shown in Fig. 2. The radius of the membrane is denoted as a, the vertical deflection of the membrane is described by function $w(r)$, and the horizontal (radial) displacement is described by $u(r)$. The membrane is assumed to have a constant thickness h. Any chosen point on the middle surface of the membrane will move to the position $[r + u(r), w(r)]$. The deformed shape of the membrane is described by a function $g(r)$ implicitly defined by the relation $g(r + u(r)) = w(r)$. Let us also suppose that the membrane is clamped at its edge, and its bending stiffness is negligible.

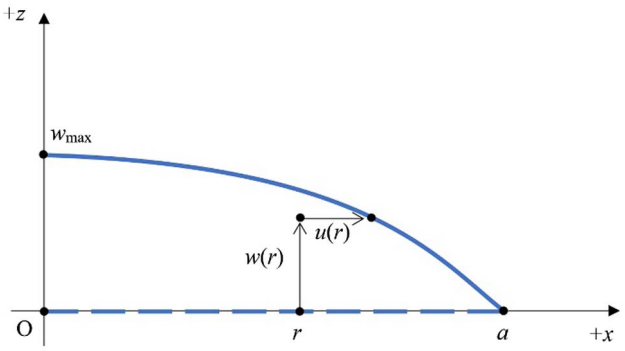


Fig. 2. Scheme of deformation of circular axis-symmetrical membrane under constant pressure (dashed line—membrane under zero pressure, $w(r)$ and $u(r)$–displacements of a point with initial position $[r, 0]$, a–membrane radius, w_{max}–maximal deflection).

As was already mentioned in the introduction, the constitutive behavior is described by the Saint Venant–Kirchhoff material model, which deals with the Green–Lagrange strain and the second Piola–Kirchhoff stress [41]. Normal strains in the radial direction, ε_r, and in the tangential direction, ε_t, are then expressed as

$$\varepsilon_r = \frac{1}{2}(\lambda_r^2 - 1) = u' + \frac{1}{2}(u'^2 + w'^2),$$
$$\varepsilon_t = \frac{1}{2}(\lambda_t^2 - 1) = \frac{u}{r} + \frac{u^2}{2r^2}, \tag{1}$$

where u' and w' are derivatives of displacements with respect to the radial coordinate r, and

$$\lambda_r = \sqrt{(1 + u')^2 + w'^2},$$
$$\lambda_t = 1 + \frac{u}{r} \tag{2}$$

are stretches in the radial and tangential directions. Since the material is in a plane-stress state and the shear strain γ_{rt} vanishes, the strain density energy Λ_{int} (per unit initial volume) is expressed as [24,25]

$$\Lambda_{\text{int}}(\varepsilon_r, \varepsilon_t) = \frac{E}{2(1 - \nu^2)}(\varepsilon_r^2 + 2\nu\varepsilon_r\varepsilon_t + \varepsilon_t^2), \tag{3}$$

where E is the Young modulus and ν is the Poisson ratio. These elastic constants are characteristics of the chosen material of the membrane.

Differentiation of the elastic potential, Eq. (3), leads to the stress-strain equations [24,25]

$$\sigma_r = \frac{\partial\Lambda_{\text{int}}}{\partial\varepsilon_r} = \frac{E}{1 - \nu^2}(\varepsilon_r + \nu\varepsilon_t),$$
$$\sigma_t = \frac{\partial\Lambda_{\text{int}}}{\partial\varepsilon_t} = \frac{E}{1 - \nu^2}(\nu\varepsilon_r + \varepsilon_t). \tag{4}$$

For a constant pressure p, one can express the total potential energy E_p of the pressurized membrane as

$$E_p = E_{\text{int}} + E_{\text{ext}}, \tag{5}$$

where

$$E_{\text{int}} = 2\pi h \int_0^a \Lambda_{\text{int}} r\,\mathrm{d}r \tag{6}$$

is the stored elastic energy (work done by internal forces), and

$$E_{\text{ext}} = -2\pi p \int_0^a w(r + u)(1 + u')\,\mathrm{d}r \tag{7}$$

is the potential energy of external forces.

The stable equilibrium state of the membrane corresponds to the minimum of potential energy E_p, and so the first variation of potential energy at this state must vanish. The first variation of the stored elastic energy can be expressed as

$$\delta E_{\text{int}} = 2\pi h \int_0^a \left(\frac{\partial\Lambda_{\text{int}}}{\partial\varepsilon_r}\delta\varepsilon_r + \frac{\partial\Lambda_{\text{int}}}{\partial\varepsilon_t}\delta\varepsilon_t\right) r\,\mathrm{d}r$$
$$= 2\pi h \int_0^a (\sigma_r\delta\varepsilon_r + \sigma_t\delta\varepsilon_t) r\,\mathrm{d}r. \tag{8}$$

As follows from Eq. (1), the strain variations are linked to the displacement variations by

$$\delta\varepsilon_r = \delta u' + u'\delta u' + w'\delta w',$$
$$\delta\varepsilon_t = \frac{\delta u}{r} + \frac{u\delta u}{r^2}. \tag{9}$$

Substituting these expressions into Eq. (8), making use of Eq. (4) and integrating by parts, we obtain

$$\begin{aligned}\delta E_{\text{int}} &= 2\pi h \int_0^a \left(r\sigma_r(\delta u' + u'\delta u' + w'\delta w') + \sigma_t\left(\delta u + \frac{u\delta u}{r}\right)\right)\mathrm{d}r \\ &= 2\pi h[r\sigma_r(\delta u + u'\delta u + w'\delta w)]_{r=0}^a \\ &\quad - 2\pi h \int_0^a ((r\sigma_r(1+u'))'\delta u + (r\sigma_r w')'\delta w)\mathrm{d}r \\ &\quad + 2\pi h \int_0^a \sigma_t\left(\delta u + \frac{u\delta u}{r}\right)\mathrm{d}r. \end{aligned} \tag{10}$$

In a similar fashion, the variation of the potential energy of external forces is expressed as

$$\begin{aligned}\delta E_{\text{ext}} &= -2\pi p \int_0^a \delta w(r+u)(1+u')\mathrm{d}r - 2\pi p \int_0^a w\delta u(1+u')\mathrm{d}r \\ &\quad - 2\pi p \int_0^a w(r+u)\delta u'\mathrm{d}r \\ &= -2\pi p \int_0^a \delta w(r+u)(1+u')\mathrm{d}r - 2\pi p[w(r+u)\delta u]_{r=0}^a \\ &\quad + 2\pi p \int_0^a w'(r+u)\delta u\mathrm{d}r. \end{aligned} \tag{11}$$

Admissible displacement functions must satisfy boundary conditions

$$u(0) = 0, \qquad u(a) = 0, \qquad w(a) = 0, \tag{12}$$

and analogous conditions must be satisfied by the variations δu and δw. Consequently, most of the boundary terms in Eqs. (10), (11) vanish and the variation of total potential energy can be written as

$$\begin{aligned}\delta E_p &= \delta E_{\text{int}} + \delta E_{\text{ext}} \\ &= -2\pi h r\sigma_r w'\delta w|_{r=0} \\ &\quad + 2\pi \int_0^a \left(h\sigma_t\left(1+\frac{u}{r}\right) - h(r\sigma_r(1+u'))' \right. \\ &\quad \left. + pw'(r+u)\right)\delta u\mathrm{d}r \\ &\quad - 2\pi \int_0^a (h(r\sigma_r w')' + p(r+u)(1+u'))\delta w\mathrm{d}r. \end{aligned} \tag{13}$$

Since the values of variations δu and δw in the open interval $(0, a)$ are arbitrary, the terms that multiply these variations in the integrals must vanish (almost everywhere). The corresponding optimality conditions

$$h[r\sigma_r(1+u')]' - h\sigma_t\left(1+\frac{u}{r}\right) = p(r+u)w',$$
$$-h(r\sigma_r w')' = p(r+u)(1+u') \tag{14}$$

represent the strong form of equilibrium equations. In the case of small radial displacements, we have $u' \ll 1$ and $u \ll r$, and Eq. (14) reduce to

$$h(r\sigma_r)' - h\sigma_t = prw',$$
$$-h(r\sigma_r w')' = pr. \tag{15}$$

Such simplified equilibrium equations are considered in papers [24–40]. The present paper deals with the more accurate equilibrium Eq. (14).

It is also interesting to look at the structure of the boundary conditions. At $r = a$, conditions $u(a) = 0$ and $w(a) = 0$ describe the constraints imposed by fixing the physical boundary of the membrane. On the other hand, point $r = 0$ physically corresponds to the center of the membrane and becomes a part of the (mathematical) boundary only when the domain of analysis is reduced to the interval $[0, a]$ based on axial symmetry. Continuity of the radial displacement implies that $u(0) = 0$, but continuity of the deflection does not lead to any constraint on $w(0)$. Consequently, the variation δw at $r = 0$ is arbitrary, and the missing boundary condition would normally be obtained by setting the term that multiplies $\delta w(0)$ in Eq. (13) to zero. In the present case, this term, given by $-2\pi h r\sigma_r w'$, seems to vanish automatically at $r = 0$. However, one should realize that this reasoning is based on the implicit assumption that $\sigma_r w'$ remains bounded as $r \to 0$. In principle, a concentrated force F_0 could be applied at the center of the membrane, and then the stress would be unbounded. The potential energy would have to be enriched by the term $-F_0 w(0)$, and the resulting boundary condition would read

$$\lim_{r\to 0^+}(r\sigma_r w') = -\frac{F_0}{2\pi h}. \tag{16}$$

Note that $Q(r) = 2\pi h r\sigma_r(r)w'(r)$ corresponds to the transversal component of the specific internal force integrated along a circle of radius r, and Eq. (16) can be interpreted as $Q(0) = -F_0$. The second equilibrium equation from Eq. (14) can be rewritten as

$$-Q' = 2\pi p(r+u)(1+u'). \tag{17}$$

Integrating and taking into account condition $Q(0) = -F_0$, we obtain

$$Q(r) = -F_0 - \pi p(r+u(r))^2. \tag{18}$$

This relation represents the equilibrium condition written for the part of the membrane which, in the undeformed configuration, has the shape of a disk of radius r.

While the boundary conditions of Eq. (12) are kinematic and, from the mathematical point of view, are essential, the fourth boundary condition of Eq. (16) is static, and it does not need to be imposed *a priori* on trial functions that approximate the exact solution. In the absence of a concentrated force, it can be expected that the deflection is continuously differentiable over the entire membrane. Combined with axial symmetry, continuous differentiability implies that

$$w'(0) = 0. \tag{19}$$

This condition is not essential, but if it is satisfied by the numerical approximation of the exact solution, faster convergence

can be expected. Therefore, in the following section, we will consider Eqs. (12) and (19) as the appropriate boundary conditions.

3. NUMERICAL SOLUTION OF CALCULATION OF MEMBRANE'S DEFORMATION

Since Eq. (14) cannot be solved analytically, a numerical procedure based on series expansion will be presented in this paper. Instead of solving the differential equations directly, we can exploit the fact that the solution minimizes potential energy among all kinematically admissible states. The displacement functions will be approximated by suitable polynomial series, the potential energy will be expressed as a function of the series coefficients, and an appropriate optimization algorithm will be invoked [42,43]. Direct minimization of the potential energy given by Eq. (5) turns out to be more robust than, e.g., minimization of the residual in Eq. (14). It is convenient to introduce a dimensionless spatial coordinate $\rho = r/a$ and solve the problem on the interval [0, 1]. In terms of the dimensionless coordinate, the potential energy and strains can be expressed as

$$\begin{aligned} E_p &= \frac{a^2 E\pi h}{(1-\nu^2)}\int_0^1 (\varepsilon_r^2 + 2\nu\varepsilon_r\varepsilon_t + \varepsilon_t^2)\rho \mathrm{d}\rho \\ &\quad - 2\pi p a^2 \int_0^1 w\left(\rho + \frac{u}{a}\right)\left(1 + \frac{u'}{a}\right)\mathrm{d}\rho, \\ \varepsilon_r &= \frac{u'}{a} + \frac{1}{2a^2}(u'^2 + w'^2), \\ \varepsilon_t &= \frac{u}{a\rho} + \frac{u^2}{2a^2\rho^2}. \end{aligned} \tag{20}$$

To keep the notation simple, functions u and w are denoted by the same symbols even when they are considered as functions of the dimensionless variable ρ, and their derivatives with respect to ρ are in Eq. (20) denoted by primes.

As explained in detail in the previous section, the displacements should satisfy boundary conditions of Eqs. (12) and (19). From the symmetry of the problem it is clear that the deflection w should be an even function, and the radial displacement u should be an odd function. Therefore, let us use polynomial approximations in the form

$$w(\rho) = \sum_{i=1}^{N} b_i(1-\rho^{2i}), \qquad u(\rho) = \sum_{j=1}^{M} c_j(\rho - \rho^{2j+1}), \tag{21}$$

where $b_i (i = 1, 2, \ldots N)$ and $c_j (j = 1, 2 \ldots M)$ are coefficients to be found. Such approximations automatically satisfy the boundary conditions. The maximal deflection

$$w_{\max} = w(0) = \sum_{i=1}^{N} b_i \tag{22}$$

corresponds to the sum of all coefficients b_i. To express the resulting displacements as functions of the variable r, the following formulas can be applied [as can be obtained from the substitution $\rho = r/a$ and Eq. (21)], as follows:

$$w(r) = \sum_{i=1}^{N} b_i\left(1 - \frac{r^{2i}}{a^{2i}}\right), \qquad u(r) = \frac{r}{a}\sum_{j=1}^{M} c_j\left(1 - \frac{r^{2j}}{a^{2j}}\right). \tag{23}$$

The procedure can be summarized as follows:

1. Choose an appropriate series expansion that approximates displacement functions $w(\rho)$ and $u(\rho)$, and satisfies boundary conditions given by Eqs. (12) and (19). In the present study, Eq. (21) was used.
2. Substitute the approximation series [Eq. (21) in the present case] into Eq. (20) and find the unknown coefficients based on the condition of minimum of potential energy E_p, using an appropriate optimization algorithm.
3. Calculate the resulting displacements according to Eq. (21) or Eq. (23), and the final shape of the membrane given by the function $g(r + u(r)) = w(r)$.

4. EXAMPLE

To illustrate the aforementioned procedure, let us analyze a membrane using the same parameters as Mikš and Novák in their paper [40]: radius $a = 10$ mm, thickness $h = 0.1$ mm, Young modulus $E = 1.97$ MPa, Poisson ratio $\nu = 0.4$, and pressure $p = 0.001 \cdot E = 1.97$ kPa.

The calculation was done in the MATLAB software, exploiting the large-scale interior-point algorithm with Hessian update by the Broyden–Fletcher–Goldfarb–Shanno method [42,43]. As a check, the residuals in equilibrium Eq. (14) were calculated. The maximal absolute value of residuals (i.e., differences between the left- and the right-hand side) was $2.8 \cdot 10^{-3}$ Pa for the first equation and $1.3 \cdot 10^{-3}$ Pa for the second equation (for equations rewritten in terms of the dimensionless coordinate $\rho = r/a$).

The calculated displacements $w(r)$ and $u(r)$ are shown in Fig. 3. The red-dashed line indicates the approximate solution obtained with a simplified model based on assumptions $u' \ll 1$ and $u \ll r$ [Eq. (15)]. Figure 4 shows the resulting shape of

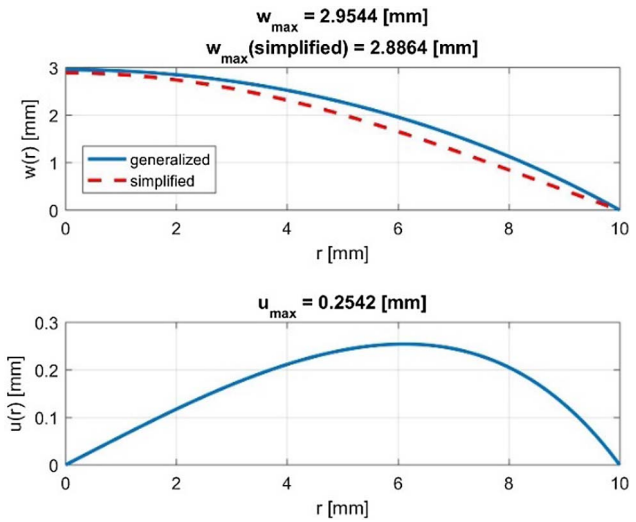


Fig. 3. Calculated displacements $w(r)$ and $u(r)$ and comparison with simplified model [Eq. (15), red-dashed line].

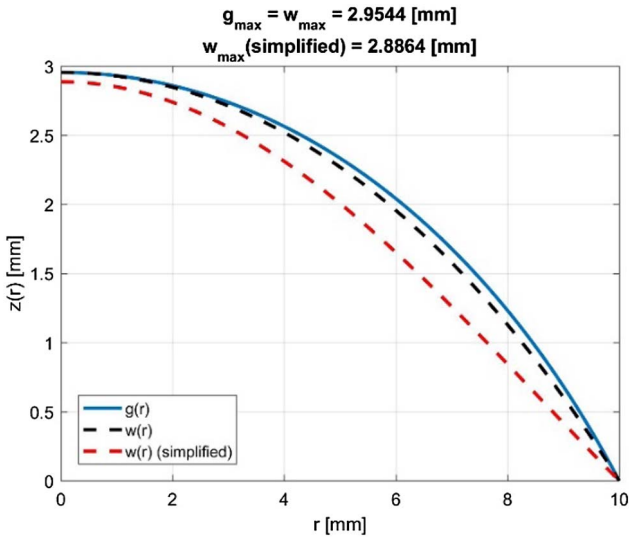


Fig. 4. Final membrane shape (blue line), vertical deflection (black dashed line) and comparison with simplified model [Eq. (15), red-dashed line].

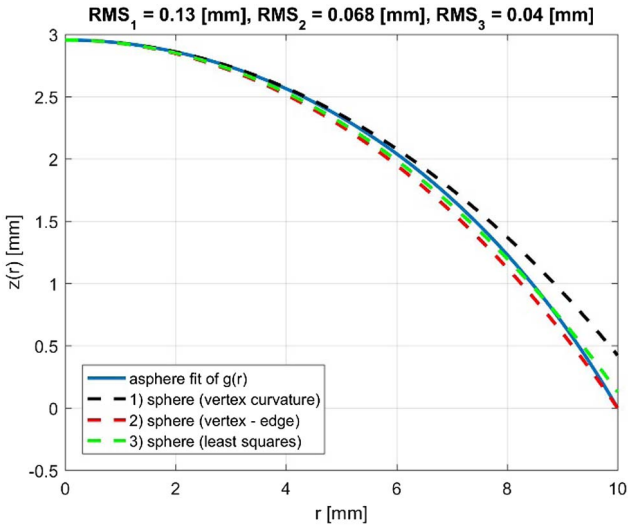


Fig. 5. Approximation of resulting membrane shape by different spheres: (1) spherical part of aspherical expression, (2) sphere matching the deflection at $r = 0$ and the edge of the membrane ($r = a$), (3) sphere matching the deflection at $r = 0$ and minimizing in the least-squares sense the deviation from the calculated membrane shape.

the membrane $g(r)$, which was afterward approximated by an aspherical expression given by the formula

$$z(r) = z_{\max} + \frac{cr^2}{1 + \sqrt{1 - c^2 r^2}} + \sum_{i=2}^{5} A_{2i} r^{2i}, \qquad (24)$$

where $z_{\max}$ is the maximal deflection of the surface, $c = 1/R_0$ is the axial curvature, and $A_{2i} (i = 2, \ldots 5)$ are aspherical coefficients that characterize deviations from the sphere of radius R_0. The results of the approximation are summarized in Table 1 for both the generalized and the simplified solution.

The final shape of the deformed membrane was compared to spherical approximations. The differences between the exact shape and its approximations by spheres are shown in Figs. 5 and 6. Three approximations are considered: 1) the spherical part of aspherical expression in Eq. (24) (i.e., aspherical coefficients $A_{2i} = 0$ for every i); 2) a sphere with the same maximal deflection at $r = 0$, passing through the edge of the membrane ($r = a$); and 3) a sphere with the same maximal deflection at $r = 0$ min imizing in the least-squares sense the differences between the calculated approximation and the exact membrane shape. RMS error values are specified in the figure. As is shown, the difference between the resulting membrane shape and a sphere is not negligible. Focusing on the differences between approximations shown in Fig. 6, one can say that the least average deflection is obtained in the case of approximation in the least-squares sense (case 3). However, considering the spherical shape as close as possible to the real one measured from the central part to the maximal distance, the best results are obtained with the spherical part of the aspherical expression

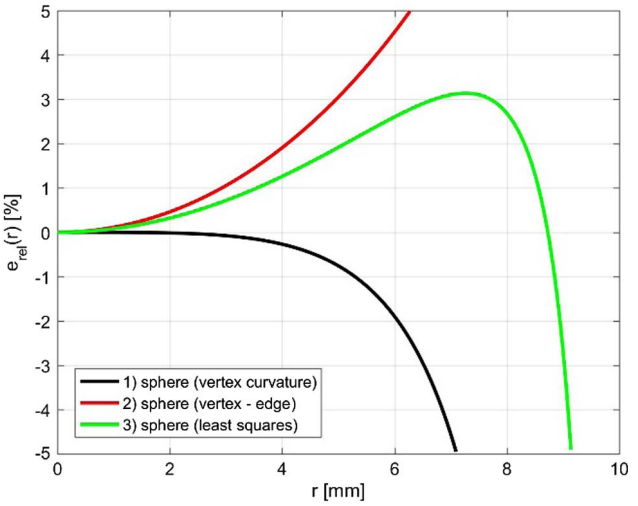


Fig. 6. Relative error of spherical approximations of the membrane shape: (1) spherical part of aspherical expression, (2) sphere matching the deflection at $r = 0$ and the edge of the membrane ($r = a$), (3) sphere matching the deflection at $r = 0$ and minimizing in the least-squares sense the deviation from the calculated membrane shape.

[spherical part of Eq. (24)]. Therefore, if one wants to use the region close to the optical axis for imaging purposes ($r \to 0$), the best approximation of the membrane shape will be given by the spherical part of Eq. (24). In the presented

Table 1. Approximation of Deformed Membrane Shape by Aspherical Expression

	$z_{\max}$ [mm]	R_0[mm]	A_4[mm^{-3}]	A_6[mm^{-5}]	A_8[mm^{-7}]	A_{10}[mm^{-9}]	RMS [mm]
Complete model	2.9544	−21.0163	−2.4120e-05	−1.8963e-07	1.2167e-09	−1.1365e-11	4.9e-05
Simplified model	2.8864	−13.3843	1.3661e-04	1.7797e-07	−2.3620e-10	8.1189e-12	2.6e-05

example, the relative error will be less than 1% in one half of the studied region.

To illustrate imaging properties of the membrane lens, the results from the previous part were used for a simulation of aberrations. Calculated aspherical surfaces were used for optical ray-tracing in OSLO software. Figs. 7 a) and b) show the influence of the membrane's shape on the transversal spherical aberration of the membrane lens with thickness $h = 0.1$ mm. DX and DY denote the transversal ray aberrations, and FX and FY denote the height of an incident ray in the interval −5 mm to 5 mm. The diameter of the lens was chosen as 10 mm, and the calculation was made for the wavelength of light $\lambda = 587$ nm. The first figure [Fig. 7(a)] shows the transversal spherical aberration of the membrane lens with radius of curvature $R = 21.016$ mm and axial thickness $d = 5$ mm if the aspherical coefficients of approximation (24) are neglected. The liquid of the lens was distilled water. In the second figure [Fig. 7(b)], the transversal spherical aberration is shown for an aspherical membrane lens with the same vertex radius $R = 21.016$ mm and the same axial thickness, with the following values of aspherical coefficients: $A_4 = 2.41198\text{e}-05\ \text{mm}^{-3}$, $A_6 = 1.89627\text{e}-07\ \text{mm}^{-5}$, $A_8 = -1.21673\text{e}-09\ \text{mm}^{-7}$, and $A_{10} = 1.13654\text{e}-11\ \text{mm}^{-9}$. Both lenses have the same focal length $f' = 63.105$ mm. As is obvious from the figures, the membrane lens cannot be replaced by a spherical lens, because the difference in the aberrations is too large.

Let us now analyze the influence of the approximate solution of membrane deformation, which can be calculated from Eqs. (15), on the transversal spherical aberration of the lens with thickness $h = 0.1$ mm, radius of curvature $R = 21.016$ mm, and axial thickness $d = 5$ mm. The optical liquid is distilled water. The aspherical coefficients of this lens are $A_4 = -6.75618\text{e}-05\ \text{mm}^{-3}$, $A_6 = -1.49495\text{e}-08\ \text{mm}^{-5}$, $A_8 = -2.37860\text{e}-11\ \text{mm}^{-7}$, and $A_{10} = -3.77091\text{e}-14\ \text{mm}^{-9}$. Note that the shape of the membrane was calculated from Eq. (15) by varying the applied pressure to get the required radius of curvature. Figure 7(c) shows the transversal spherical aberrations. Comparing the results to Fig. 7(b) (solution based on the complete model), one can see that the spherical aberration is "undercorrected" by the solution based on Eq. (14), while the approximate solution of Eq. (15) gives an "overcorrected" spherical aberration. From the presented results, it is obvious that an accurate analysis of membrane lenses with variable focus must be based on the complete model, because the error in imaging properties caused by using the approximate approach is too large.

A very important parameter, which has a crucial effect on the calculated membrane deflection of the membrane lens and thus also on the imaging characteristics, is the membrane thickness. Therefore, the influence of thickness on imaging properties was studied in more detail. First, an initial calculation was done with a membrane of the same parameters as in the previous part, i.e., radius $a = 10$ mm, thickness $h = 0.1$ mm, Young modulus $E = 1.97$ MPa, Poisson ratio $\nu = 0.4$, and pressure $p = 0.001 \cdot E = 1.97$ kPa. The results of the approximation with Eq. (24) are presented in Table 1. Afterward, the thickness was varied and the pressure was optimized to get the same spherical part of the approximation of Eq. (24) as for the initial

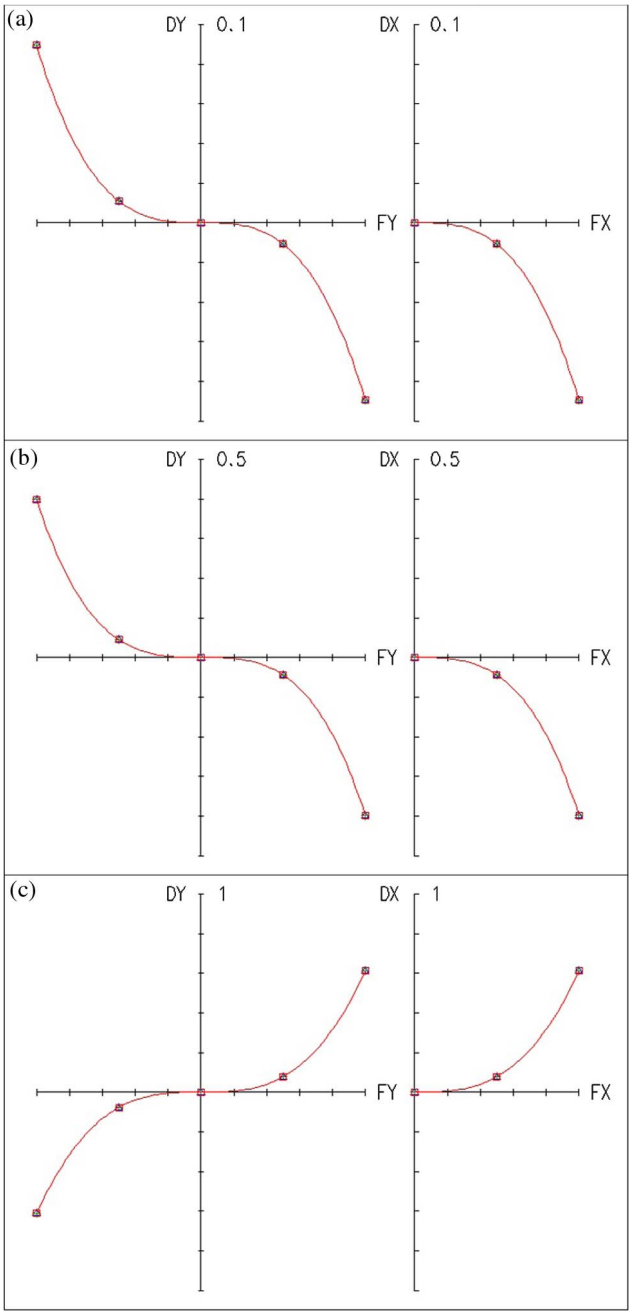


Fig. 7. Comparison of transversal spherical aberration of the membrane lens (a) with neglected aspherical coefficients of approximation (24), (b) with aspherical coefficients taken into account, (c) with aspherical coefficients taken into account but deformation calculated from simplified Eq. (15) and with the applied pressure adjusted to obtain the same paraxial properties (the same focal length) as with the general Eq. (14); DX and DY denote the transversal ray aberrations, and FX and FY denote the height of the incident ray in the interval from −5 mm to 5 mm.

parameters. Therefore, paraxial properties of the lens remain the same. However, aspherical coefficients and therefore aberrations of optical imaging change. Table 2 presents the results of the analysis. It is obvious that the thickness of the membrane affects the aspherical coefficients. A different pressure has to be

Table 2. Comparison of Applied Pressure and Aspherical Coefficients of Lenses with the Same Optical Paraxial Properties and Different Thicknesses

h[mm]	p[kPa]	z_{max}[mm]	R_0[mm]	A_4[mm^{-3}]	A_6[mm^{-5}]	A_8[mm^{-7}]	A_{10}[mm^{-9}]
0.3	5.91	2.9544	−21.0163	−2.4123e-05	−1.8954e-07	1.2157e-09	−1.1361e-11
0.5	9.85			−2.4132e-05	−1.8927e-07	1.2125e-09	−1.1348e-11
0.7	13.79			−2.4149e-05	−1.8879e-07	1.2069e-09	−1.1326e-11

applied to obtain the same paraxial optical properties, and therefore various lenses can image differently.

5. EXPERIMENTAL VERIFICATION OF THEORETICAL RESULTS

The experimental measurement of a liquid membrane lens was done to check whether the theoretical prediction of membrane deflections corresponds to the real behavior. The scheme of the lens is shown in Fig. 8. The basic part is a chamber. The liquid (distilled water in our case) is injected by a precise pressure pump through an inlet valve. The cover of the lens is attached by screws and fixes the membrane (made of Sylgard 184 [44]), which is deformed depending on the amount of liquid carried into the chamber. The second part of the lens is a flat plane-parallel plate. The amount of liquid in the chamber can be reduced by a bleed valve. For the experiment, the chamber diameter was $D = 30$ mm and the membrane thickness $h = 0.5$ mm. Because mechanical parameters of the membrane change for different recipes of mixtures and are affected by thermal conditions during measurement, optimal values of mechanical parameters of the membrane used for the calculation (Young's modulus $E = 1.1$ MPa and Poisson's ratio $\nu = 0.45$) were obtained by minimizing the difference between the theoretically calculated and experimentally measured maximal membrane deflection.

The shape of the membrane was determined by a device for measurement of surface topography (see Fig. 9) with 4 degrees of freedom. The basic component of the device is a chromatic confocal sensor [45], which can be positioned using step motors in two mutually orthogonal directions (axis x and axis z) and rotated (around axis y). The sample can then be rotated using the measurement table, and data characterizing the complete topography can be obtained.

In view of the rotational symmetry of the lens, only the meridian $z = f(x)$ was measured. The range of measurement was from −5 mm to 5 mm because of mechanical construction of the device. The estimated uncertainty of the deflection measurement using the aforementioned laboratory device is 0.02 mm (combination of uncertainties in the measurement with the confocal sensor and mechanical errors in the device).

The results of the measurement are shown in Fig. 10 and Table 3 for three pressures—$p = 3.0$, 3.5, and 4.0 kPa—and the experimental values are compared to theoretical predictions calculated using the procedure presented in Section 3. RMS errors (differences between theoretically predicted and measured values) are $RMS_{3.0} = 0.033$ mm, $RMS_{3.5} = 0.024$ mm, and $RMS_{4.0} = 0.014$ mm. It is confirmed that the presented theoretical model corresponds very well to the real behavior.

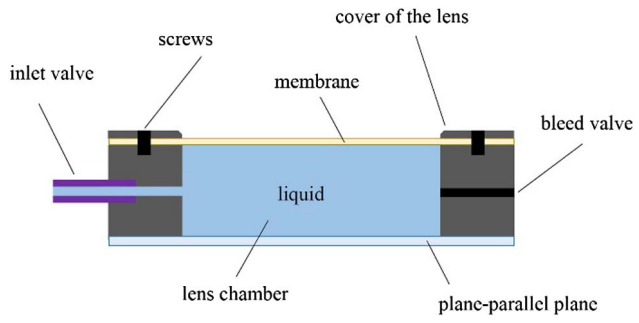


Fig. 8. Scheme of laboratory sample of liquid membrane lens.

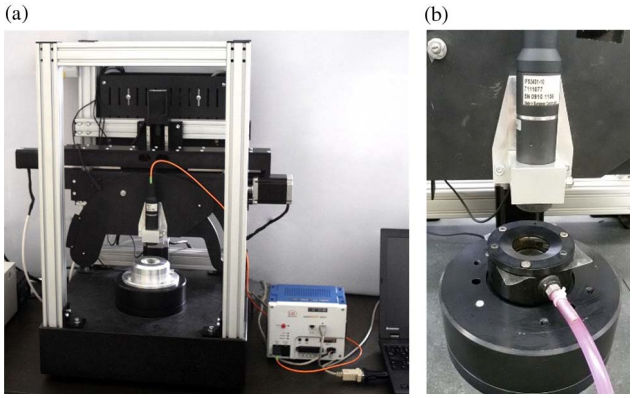


Fig. 9. (a) Laboratory device for surface topography measurement and (b) measured sample of liquid membrane lens of the measurement table.

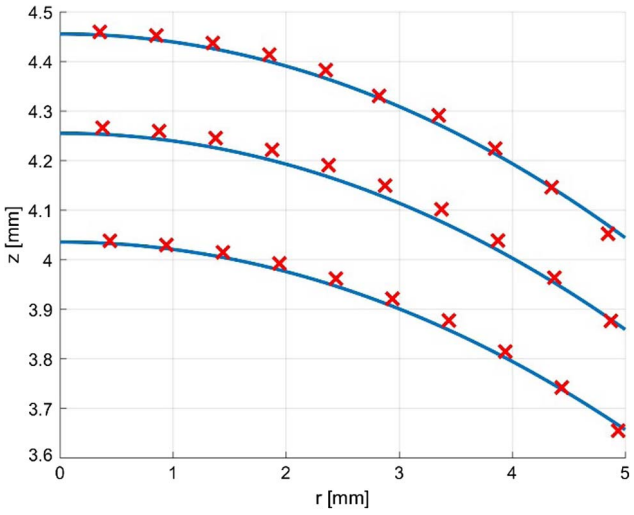


Fig. 10. Measured lens deformation (red crosses) and comparison to theoretical prediction (blue lines) for pressures $p = 3.0$, 3.5, and 4.0 kPa.

Table 3. Approximation of Measured Membrane Shape by Aspherical Expression [Eq. (24)]

p[kPa]	z_{max}[mm]	R_0[mm]	A_4[mm^{-3}]	A_6[mm^{-5}]	A_8[mm^{-7}]	A_{10}[mm^{-9}]	RMS [mm]
3.0	4.05	−32.24	4.483e-05	−3.879e-06	1.453e-08	−7.486e-12	1.5e-02
3.5	4.28	−28.60	1.657e-04	−5.835e-06	1.438e-08	3.673e-11	1.2e-02
4.0	4.47	−29.22	3.730e-05	−1.438e-06	−2.639e-09	3.630e-11	9.7e-03

Therefore, the method can be used for modeling the shape of membrane liquid lenses.

6. CONCLUSION

The paper presented a mathematical model and numerical solution for the calculation of the deformation of an axisymmetrical membrane liquid lens. The general formulas for deflections were derived based on precise geometrical equations and the Saint Venant–Kirchhoff material model, which assumes a linear relationship between the Green–Lagrange strain and the second Piola–Kirchoff stress. The numerical solution was obtained by minimizing the potential energy. As one possible approach, an approximation using power series that satisfy the boundary conditions of the problem was presented. As is shown in an example and verified by an experiment, the derived general formulas and the numerical procedure can be used for accurate modeling of membrane liquid lenses. The results can be exploited in optical system design with incorporated liquid lenses, and the developed method will be used in the future for precise analysis of the imaging properties of such systems.

Funding. Ceské Vysoké Ucení Technické v Praze (CVUT) (SGS17/004/OHK1/1T/11).

REFERENCES

1. H. Ren and S. T. Wu, *Introduction to Adaptive Lenses* (Wiley, 2012).
2. Optotune, http://www.optotune.com.
3. Holochip, http://www.holochip.com.
4. A. H. Rawicz and I. Mikhailenko, "Modeling a variable-focus liquid-filled optical lens," Appl. Opt. **35**, 1587–1589 (1996).
5. D. Y. Zhang, N. Justis, V. Lien, Y. Berdichevsky, and Y.-H. Lo, "High-performance fluidic adaptive lenses," Appl. Opt. **43**, 783–787 (2004).
6. N. Sugiura and S. Morita, "Variable-focus liquid-filled optical lens," Appl. Opt. **32**, 4181–4186 (1993).
7. H. Ren, D. Fox, P. A. Anderson, B. Wu, and S.-T. Wu, "Tunable-focus liquid lens controlled using a servo motor," Opt. Express **14**, 8031–8036 (2006).
8. Q. Yang, P. Kobrin, C. Seabury, S. Narayanaswamy, and W. Christian, "Mechanical modeling of fluid-driven polymer lenses," Appl. Opt. **47**, 3658–3668 (2008).
9. G.-H. Feng and Y.-C. Chou, "Flexible meniscus/biconvex lens system with fluidic-controlled tunable-focus applications," Appl. Opt. **48**, 3284–3290 (2009).
10. H. W. Ren and S. T. Wu, "Variable-focus liquid lens," Opt. Express **15**, 5931–5936 (2007).
11. D. Shaw and T. E. Sun, "Optical properties of variable-focus liquid-filled optical lenses with different membrane shapes," Opt. Eng. **46**, 024002 (2007).
12. D. Shaw and C.-W. Lin, "Design and analysis of an asymmetrical liquid-filled lens," Opt. Eng. **46**, 123002 (2007).
13. H. Choi, D. S. Han, and Y. H. Won, "Adaptive double-sided fluidic lens of polydimethylsiloxane membranes of matching thickness," Opt. Lett. **36**, 4701–4703 (2011).
14. L. Li, Q. H. Wang, and W. Jiang, "Liquid lens with double tunable surfaces for large power tunability and improved optical performance," J. Opt. **13**, 115503 (2011).
15. F. Schneider, J. Draheim, R. Kamberger, P. Waibel, and U. Wallrabe, "Optical characterization of adaptive fluidic silicone-membrane lenses," Opt. Express **17**, 11813–11821 (2009).
16. Y.-K. Fuh, M.-X. Lin, and S. Lee, "Characterizing aberration of a pressure-actuated tunable biconvex microlens with a simple spherically-corrected design," Opt. Lasers Eng. **50**, 1677–1682 (2012).
17. A. Miks, J. Novak, and P. Novak, "Algebraic and numerical analysis of imaging properties of thin tunable-focus fluidic membrane lenses with parabolic surfaces," Appl. Opt. **52**, 2136–2144 (2013).
18. L. H. Wang, H. Oku, and M. Ishikawa, "Development of variable focus lens with liquid-membrane-liquid structure and 30 mm optical aperture," Proc. SPIE **8617**, 861706 (2013).
19. L. H. Wang, H. Oku, and M. Ishikawa, "An improved low-optical-power variable focus lens with a large aperture," Opt. Express **22**, 19448–19456 (2014).
20. S. T. Choi, B. S. Son, G. W. Seo, S. Y. Park, and K. S. Lee, "Opto-mechanical analysis of nonlinear elastomer membrane deformation under hydraulic pressure for variable-focus liquid-filled microlenses," Opt Express **22** 6133–6146 (2014).
21. D. Liang and X.-Y. Wang, "A bio-inspired optical system with a polymer membrane and integrated structure," Bioinsp. Biomim **11**, 066008 (2016).
22. J.-W. Du, X.-Y. Wang, and D. Liang, "Bionic optical imaging system with aspheric solid-liquid mixed variable-focus lens," Opt. Eng. **55**, 023105 (2016).
23. N. Hasan, A. Banerjee, H. Kim, and C. H. Mastrangelo, "Tunable-focus lens for adaptive eyeglasses," Opt. Express **25**, 1221–1233 (2017).
24. S. P. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells* (McGraw Hill, 1964).
25. A. S. Volmir, *Flexible Plates and Shells* (State Publishing House of Technical-Theoretical Literature, 1956), in Russian.
26. H. Hencky, "On the stress state in circular plates with vanishing bending stiffness," Zeitschrift für Mathematik und Physik **63**, 311–317 (1915).
27. W. Z. Chein, "Asymptotic behavior of a thin clamped circular plate under uniform normal pressure at very large deflection," Sci. Rep. Natn. Tsinghua Univ. **A5**, 71–94 (1948).
28. J. D. Campbell, "On the theory of initially tensioned circular membranes subjected to uniform pressure," Quart. J. Mech. Appl. Math. **9**, 84–93 (1956).
29. M. A. Goldberg and A. B. Pifko, "Large deflection analysis of uniformly loaded annular membranes," AIAA J. **1**, 2111–2115 (1963).
30. M. A. Goldberg and A. B. Pifko, "Iterative and power series solutions for the large deflection of an annular membrane," AIAA J. **2**, 2234 (1964).
31. N. Perrone and R. Kao, "A general nonlinear relaxation technique for solving nonlinear problems in mechanics," J. Appl. Mech. **38**, 371–376 (1971).
32. R. Kao and N. Perrone, "Large deflections of axisymmetric circular membranes," Int. J. Solids Structures **7**, 1601–1612 (1971).
33. R. G. Pettit, "Geometric and optical properties of inflatable point-focusing reflections," Master thesis (Brigham Young University, 1985).
34. A. Kelkar, W. Elber, and I. S. Rajut, "Large deflections of circular isotropic membranes subjected to arbitrary axisymmetric loading," Comput. Struct. **21**, 413–421 (1985).
35. W. B. Fichter, "Some Solutions for the Large Deflections of Uniformly Loaded Circular Membranes," NASA Technical Paper 3658 (1997).

36. D. J. Allman, "Variational solutions for the nonlinear deflexion of an annular membrane under axial load," Int. J. Mech. Sci. **24**, 749–753 (1982).
37. M. Sheploak and J. Dugundji, "Large deflections of clamped circular plates under initial tension and transitions to membrane behavior," J. Appl. Mech. **65**, 107 (1998).
38. F. Zhao, "Nonlinear solutions for circular membranes and thin plates," Proc. SPIE **6926**, 69260W (2008).
39. B. Stanford and P. Ifju, "The validity range of low fidelity structural membrane models," Exp. Mech. **48**, 697–711 (2008).
40. A. Mikš and P. Novák, "Calculation of a surface shape of a pressure actuated membrane liquid lens," Opt. Lasers Eng. **58**, 60–66 (2014).
41. B. Audoly and Y. Pomeau, *Elasticity and Geometry: From Hair Curls to the Nonlinear Response of Shells* (Oxford University, 2010).
42. M. Aoki, *Introduction to Optimization Techniques: Fundamentals and Applications of Nonlinear Programming* (Macmillan, 1971).
43. L. E. Scales, *Introduction to Non-Linear Optimization* (Macmillian, 1985).
44. Sylgard 184 Silicone Elastomer Kit http://www.dowcorning.com/applications/search/products/Details.aspx?prod=01064291.
45. Micro Epsilon, http://www.micro-epsilon.co.uk/displacement-position-sensors/confocal-sensor/.

applied optics

Deformation of a prestressed liquid lens membrane

Petr Pokorný,[1,*] Filip Šmejkal,[1] Pavel Kulmon,[1] Pavel Novák,[1] Jiří Novák,[1] Antonín Mikš,[1] Martin Horák,[2] and Milan Jirásek[2]

[1]Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
[2]Czech Technical University in Prague, Faculty of Civil Engineering, Department of Mechanics, Thákurova 7, 166 29 Prague 6, Czech Republic
*Corresponding author: petr.pokorny@fsv.cvut.cz

Received 29 August 2017; accepted 23 October 2017; posted 25 October 2017 (Doc. ID 305850); published 22 November 2017

This paper presents a complete model for analysis of the deformed shape of a prestressed circular axisymmetric membrane of a liquid lens. The governing equations are derived using the exact relation between displacements and the Green–Lagrange strains combined with the Saint Venant–Kirchhoff material law, which postulates a linear relation between the Green–Lagrange strains and the second Piola–Kirchoff stresses. A numerical solution based on minimization of potential energy is illustrated by an example, and the dependence of the maximum membrane deflection on material properties and initial prestress is analyzed. The theoretical model is then experimentally validated. It is shown that the model is suitable for large-strain analysis of liquid lens membranes and provides sufficiently accurate results that can be used in further analyses and simulations of imaging properties of active optical elements based on liquid lenses.

OCIS codes: (120.4640) Optical instruments; (220.1080) Active or adaptive optics; (350.4600) Optical engineering.

https://doi.org/10.1364/AO.56.009368

1. INTRODUCTION

Active optical elements represent a modern direction in optics, which in recent years has been under rapid development. They include well-known membrane liquid lenses [1–23], which are commercially available for imaging purposes [2,3]. The basic element of such lenses is a membrane fixed along its circumference (the specific type of fixation varies depending on the manufacturer). The lens chamber is filled by an optical liquid. When the volume of the liquid is changed, the liquid inserts uniform pressure on the membrane and the membrane deflects.

In the past, many authors performed theoretical and experimental analyses of deformed membranes loaded by a uniform pressure using various mechanical models [4–23] and various numerical techniques [24–37].

In a previous paper [38], the authors analyzed in detail the deformation of a circular axisymmetric liquid lens membrane. This work extends the previous analysis by including the effect of initial prestress, which is inevitably induced during installation of the liquid lens. The membrane is again considered to be fixed along its circumference and subjected to uniform pressure of the liquid.

Since the deflections of the membrane substantially exceed its thickness, solutions based on standard linear elasticity [39,40] are not sufficiently accurate. The influence of radial deflections is also nonnegligible. The model presented in the next section is based on exact strain-displacement relations defining the Green–Lagrange strain and on the Saint Venant–Kirchhoff material law that assumes a linear relation between the Green–Lagrange strain and the second Piola–Kirchoff stress [41]. This model properly takes into account geometrically nonlinear effects and includes the membrane's radial displacements, which are in Refs. [4–23] and considered as small or totally neglected.

In the next section, the governing equations are derived and a numerical solution based on polynomial series and optimization algorithms [42,43] is presented. An example dealing with specific parameters of a membrane lens illustrates the numerical solution and compares it to results obtained by finite elements. The dependence of the maximum membrane deflection on the material properties and initial prestress is studied. In the last section, the proposed model is validated by comparison to deflections measured in experiments. It is confirmed that the model is suitable for an accurate description of the given problem and can be used in further simulations and analyses related to optical design requiring a highly accurate description of optical interfaces.

2. THEORETICAL PREDICTION OF PRESTRESSED MEMBRANE SHAPE

A. Derivation of Governing Equations

Let us consider a circular axisymmetric membrane of a liquid lens, with the axis of symmetry denoted as z and the radial axis as r; see Fig. 1. The membrane is characterized by constant thickness h and radius a, with h so small that the bending stiffness can be neglected. Along its circumference, the membrane is fixed. Under the pressure of the optical liquid in the

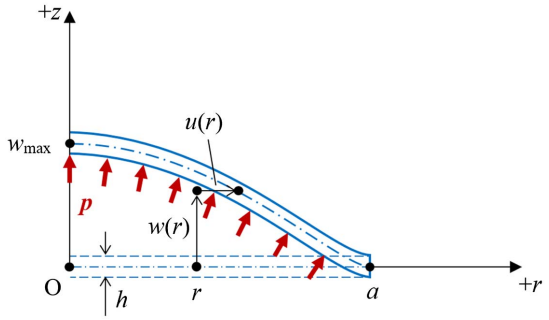


Fig. 1. Sketch of a circular axisymmetric membrane deformed by constant liquid pressure p (dashed blue lines—initial straight shape of the prestressed membrane at zero applied pressure, $w(r)$ and $u(r)$—vertical and horizontal displacements of a point with initial coordinates $[r, 0]$, a—membrane radius, h—membrane thickness, w_{max}—maximum deflection).

lens, the membrane deforms and a general point on its midplane with initial coordinates $[r, 0]$ is displaced to a new position $[r + u(r), w(r)]$, where $w(r)$ denotes the deflection (displacement in the z direction) and $u(r)$ is the radial displacement. The deformed shape of the membrane can be described by a certain function $g(r)$, implicitly defined by the relation $g(r + u(r)) = w(r)$.

Let us further assume that the membrane in its initial state (i.e., at vanishing applied pressure p) is prestressed, and its initial stretch in all in-plane directions is $\lambda_0 > 1$ (the stretch is defined with respect to the stress-free reference configuration). Further, the effect of the change of thickness h on the membrane shape is negligible, and so the deformed shape is described by the midsurface.

The radial stretch λ_r and the tangential (circumferential) stretch λ_t in the deformed state are easily expressed as

$$\lambda_r = \lambda_0\sqrt{(1+u')^2 + w'^2},$$
$$\lambda_t = \lambda_0\left(1 + \frac{u}{r}\right), \tag{1}$$

where u' and w' are the derivatives of displacements u and w with respect to the radial coordinate, r. Let us assume that the material can be described by the Saint Venant–Kirchhoff model, which postulates a linear relation between the Green–Lagrange strain and the second Piola–Kirchhoff stress [41]. Based on the definition of Green–Lagrange strain, the inplane normal strains are evaluated from the stretches as

$$\varepsilon_r = \frac{1}{2}(\lambda_r^2 - 1) = \lambda_0^2\left[u' + \frac{1}{2}(u'^2 + w'^2)\right] + \frac{1}{2}(\lambda_0^2 - 1)$$
$$= \lambda_0^2\tilde{\varepsilon}_r + \varepsilon_0, \tag{2}$$

$$\varepsilon_t = \frac{1}{2}(\lambda_t^2 - 1) = \lambda_0^2\left(\frac{u}{r} + \frac{u^2}{2r^2}\right) + \frac{1}{2}(\lambda_0^2 - 1)$$
$$= \lambda_0^2\tilde{\varepsilon}_t + \varepsilon_0, \tag{3}$$

where

$$\varepsilon_0 = \frac{1}{2}(\lambda_0^2 - 1) \tag{4}$$

is the initial strain and

$$\tilde{\varepsilon}_r = u' + \frac{1}{2}(u'^2 + w'^2), \tag{5}$$

$$\tilde{\varepsilon}_t = \frac{u}{r} + \frac{u^2}{2r^2} \tag{6}$$

are strains that would be obtained by taking the initial prestressed state as the reference configuration [while Eqs. (2) and (3) use the stress-free state as the reference configuration].

Since the state of the material corresponds to plane stress, the strain energy density $\mathcal{E}_{int}$ (per unit volume in the stress-free state) is given by [39,40]

$$\mathcal{E}_{int}(\varepsilon_r, \varepsilon_t) = \frac{E}{2(1-\nu^2)}(\varepsilon_r^2 + 2\nu\varepsilon_r\varepsilon_t + \varepsilon_t^2), \tag{7}$$

where E is the Young modulus and ν is the Poisson ratio characterizing the membrane material. Differentiating the strain energy density with respect to strains, we get the corresponding work-conjugate stresses

$$\sigma_r = \frac{\partial\mathcal{E}_{int}}{\partial\varepsilon_r} = \frac{E}{1-\nu^2}(\varepsilon_r + \nu\varepsilon_t), \tag{8}$$

$$\sigma_t = \frac{\partial\mathcal{E}_{int}}{\partial\varepsilon_t} = \frac{E}{1-\nu^2}(\varepsilon_t + \nu\varepsilon_r), \tag{9}$$

and substituting from Eqs. (2) and (3), we can express them as

$$\sigma_r = \frac{E\lambda_0^2}{1-\nu^2}(\tilde{\varepsilon}_r + \nu\tilde{\varepsilon}_t) + \sigma_0, \tag{10}$$

$$\sigma_t = \frac{E\lambda_0^2}{1-\nu^2}(\tilde{\varepsilon}_t + \nu\tilde{\varepsilon}_r) + \sigma_0, \tag{11}$$

where

$$\sigma_0 = \frac{E}{1-\nu}\varepsilon_0 = \frac{E}{2(1-\nu)}(\lambda_0^2 - 1) \tag{12}$$

is the initial prestress.

Since the strain energy density has been differentiated with respect to Green–Lagrange strains, the resulting stresses are the second Piola–Kirchhoff stresses, with the stress-free configuration taken as the referential one. It is important to realize that symbols a and h denote the membrane radius and thickness in the state before application of the liquid pressure but after application of the prestress. The corresponding dimensions in the stress-free state will be denoted as $\tilde{a}$ and $\tilde{h}$, with $\tilde{a} = a/\lambda_0$ and with $\tilde{h}$ considered as a primary geometric characteristic of the membrane.

The equilibrium state after application of pressure p can be found by exploiting the principle of minimum potential energy. The total potential energy,

$$E_p = E_{int} + E_{ext}, \tag{13}$$

is the sum of the strain energy, E_{int}, and the energy of external forces, E_{ext}. The state of minimum potential energy can be attained only if the variation of functional E_p vanishes for all admissible variations of displacements u and w. Since $\mathcal{E}_{int}$ represents the strain energy density per unit volume in the stress-free reference configuration, the strain energy

$$E_{\text{int}} = 2\pi\tilde{h}\int_0^{\tilde{a}} \mathcal{E}_{\text{int}}\tilde{r}\,\mathrm{d}\tilde{r} \tag{14}$$

must be evaluated by integrating over a cylinder of radius $\tilde{a} = a/\lambda_0$ and height $\tilde{h}$. The integration variable $\tilde{r}$ that varies from 0 to $\tilde{a}$ corresponds to the radial coordinate r (in the prestressed state) divided by λ_0. Therefore, we can transform Eq. (14) into

$$E_{\text{int}} = \frac{2\pi\tilde{h}}{\lambda_0^2}\int_0^{a} \mathcal{E}_{\text{int}} r\,\mathrm{d}r. \tag{15}$$

The energy of external forces can be expressed as minus the applied pressure multiplied by the volume between the initial midplane and the deformed midsurface, leading to

$$E_{\text{ext}} = -2\pi p\int_0^{a} w(r+u)(1+u')\mathrm{d}r. \tag{16}$$

The first variations of the strain energy and of the energy of external forces can now be evaluated as

$$\begin{aligned}
\delta E_{\text{int}} &= \frac{2\pi\tilde{h}}{\lambda_0^2}\int_0^{a}\left(\frac{\partial\mathcal{E}_{\text{int}}}{\partial\varepsilon_r}\delta\varepsilon_r + \frac{\partial\mathcal{E}_{\text{int}}}{\partial\varepsilon_t}\delta\varepsilon_t\right) r\mathrm{d}r \\
&= \frac{2\pi\tilde{h}}{\lambda_0^2}\int_0^{a}(\sigma_r\delta\varepsilon_r + \sigma_t\delta\varepsilon_t) r\mathrm{d}r \\
&= 2\pi\tilde{h}\int_0^{a}(\sigma_r\delta\tilde{\varepsilon}_r + \sigma_t\delta\tilde{\varepsilon}_t) r\mathrm{d}r \\
&= 2\pi\tilde{h}\int_0^{a}(r\sigma_r(\delta u' + u'\delta u' + w'\delta w') \\
&\quad + \sigma_t(\delta u + u\delta u/r))\mathrm{d}r \\
&= 2\pi\tilde{h}[r\sigma_r(\delta u + u'\delta u + w'\delta w)]_{r=0}^{a} \\
&\quad - 2\pi\tilde{h}\int_0^{a}((r\sigma_r(1+u'))\delta u + (r\sigma_r w')'\delta w)\mathrm{d}r \\
&\quad + 2\pi\tilde{h}\int_0^{a}\sigma_t\left(\delta u + \frac{u\delta u}{r}\right)\mathrm{d}r,
\end{aligned} \tag{17}$$

and

$$\begin{aligned}
\delta E_{\text{ext}} &= -2\pi p\int_0^{a}\delta w(r+u)(1+u')\mathrm{d}r \\
&\quad -2\pi p[w(r+u)\delta u]_{r=0}^{a} \\
&\quad + 2\pi p\int_0^{a} w'(r+u)\delta u\mathrm{d}r.
\end{aligned} \tag{18}$$

As shown in Fig. 1, admissible functions u and w are constrained by boundary conditions

$$w(a) = 0,\ u(0) = 0,\ u(a) = 0, \tag{19}$$

and analogous conditions must be satisfied by their variations δu and δw. Summing Eqs. (17) and (18) and making use of the boundary conditions, we obtain the variation of total potential energy in the form

$$\begin{aligned}
\delta E_p &= -2\pi\tilde{h}r\sigma_r w'\delta w|_{r=0} \\
&\quad + 2\pi\int_0^{a}\left(\tilde{h}\sigma_t\left(1+\frac{u}{r}\right) - \tilde{h}(r\sigma_r(1+u'))'\right. \\
&\quad \left. + pw'(r+u)\right)\delta u\mathrm{d}r \\
&\quad -2\pi\int_0^{a}(\tilde{h}(r\sigma_r w')' + p(r+u)(1+u'))\delta w\mathrm{d}r.
\end{aligned} \tag{20}$$

The corresponding strong form of equilibrium equations reads

$$\tilde{h}(r\sigma_r(1+u'))' - \tilde{h}\sigma_t\left(1+\frac{u}{r}\right) = p(r+u)w', \tag{21}$$

$$-\tilde{h}(r\sigma_r w')' = p(r+u)(1+u'). \tag{22}$$

In many publications mentioned in the introduction [24–37], the radial displacements u and their derivatives u' are considered as negligible. Based on such an assumption, Eq. (21) would reduce to

$$\tilde{h}(r\sigma_r)' - \tilde{h}\sigma_t = prw', \tag{23}$$

$$-\tilde{h}(r\sigma_r w')' = pr. \tag{24}$$

Since the variation δw at $r = 0$ is completely arbitrary, the first term in Eq. (20) leads to the boundary condition

$$r\sigma_r w' = 0 \quad \text{at } r = 0. \tag{25}$$

At a first glance, the condition seems to be satisfied automatically. Indeed, if $\sigma_r w'$ has a finite value at $r = 0$, then multiplication by zero leads to $r\sigma_r w' = 0$. In a general setting, a concentrated vertical force F_0 could be applied at $r = 0$, and then the resulting boundary condition would read $\lim_{r\to 0^+}(r\sigma_r w') = F_0/(2\pi\tilde{h})$. For the problem studied here, no such concentrated force is present, and a refined analysis leads to the conclusion that the deflection w considered as a function of inplane coordinates x and y should be continuously differentiable. Consequently, the derivative of w with respect to r at $r = 0$ should vanish, and this can be imposed as the fourth boundary condition,

$$w'(0) = 0, \tag{26}$$

which supplements Eq. (19) conditions.

B. Numerical Approximation by Power Series

The governing differential equations written in terms of displacements u and w could be constructed by substituting the strain-displacement Eqs. (5) and (6) and the stress-strain Eqs. (10) and (11) into the equilibrium Eqs. (21) and (22). However, the resulting set of two partial differential equations with the boundary conditions of Eqs. (19) and (26) cannot be solved analytically. An approximate numerical solution can be constructed by minimizing the energy functional E_p over a finite-dimensional space of suitable approximation functions.

Since the solution is expected to be highly regular, polynomial approximations seem to be a good choice. It is also convenient to reformulate the problem in terms of dimensionless displacements $\tilde{u} = u/a$ and $\tilde{w} = w/a$, and dimensionless spatial coordinate $\rho = r/a$, so that the domain of analysis is transformed to the interval [0, 1]. The displacements can then be approximated by the polynomial series

$$\tilde{w}(\rho) = \sum_{i=1}^{N} b_i(1-\rho^{2i}), \qquad \tilde{u}(\rho) = \sum_{j=1}^{M} c_j(\rho - \rho^{2j+1}), \tag{27}$$

where b_i $(i = 1, 2, \ldots, N)$ and c_j $(j = 1, 2, \ldots, M)$ are arbitrary coefficients. Note that these approximations automatically satisfy the boundary conditions of Eq. (19) as well as of Eq. (26).

In terms of the dimensionless quantities, Eqs. (5) and (6) can be rewritten as

$$\tilde{\varepsilon}_r = \frac{\mathrm{d}\tilde{u}}{\mathrm{d}\rho} + \frac{1}{2}\left[\left(\frac{\mathrm{d}\tilde{u}}{\mathrm{d}\rho}\right)^2 + \left(\frac{\mathrm{d}\tilde{w}}{\mathrm{d}\rho}\right)^2\right] = \sum_{j=1}^{M} c_j[1-(2j+1)\rho^{2j}] + \frac{1}{2}\left(\sum_{j=1}^{M} c_j[1-(2j+1)\rho^{2j}]\right)^2 + 2\left(\sum_{i=1}^{N} i b_i \rho^{2i-1}\right)^2, \tag{28}$$

$$\tilde{\varepsilon}_t = \frac{\tilde{u}}{\rho} + \frac{\tilde{u}^2}{2\rho^2} = \sum_{j=1}^{M} c_j(1-\rho^{2j}) + \frac{1}{2}\left(\sum_{j=1}^{M} c_j(1-\rho^{2j})\right)^2. \tag{29}$$

Substituting from Eqs. (7), (2), and (3) into Eq. (15) and transforming the integration variable from r to ρ, we obtain the strain energy

$$E_{\mathrm{int}} = \frac{2\pi\tilde{h}}{\lambda_0^2}\int_0^a \frac{E}{2(1-\nu^2)}(\varepsilon_r^2 + 2\nu\varepsilon_r\varepsilon_t + \varepsilon_t^2) r\mathrm{d}r = \frac{\pi\tilde{h}Ea^2}{1-\nu^2}\left(\lambda_0^2\int_0^1(\tilde{\varepsilon}_r^2 + 2\nu\tilde{\varepsilon}_r\tilde{\varepsilon}_t + \tilde{\varepsilon}_t^2)\rho\mathrm{d}\rho + 2(1+\nu)\varepsilon_0\int_0^1(\tilde{\varepsilon}_r+\tilde{\varepsilon}_t)\rho\mathrm{d}\rho + \frac{(1+\nu)\varepsilon_0^2}{\lambda_0^2}\right). \tag{30}$$

In a similar spirit, the load energy of Eq. (16) can be expressed as

$$E_{\mathrm{ext}} = -2\pi p a^3\int_0^1 \tilde{w}(\rho+\tilde{u})\left(1+\frac{\mathrm{d}\tilde{u}}{\mathrm{d}\rho}\right)\mathrm{d}\rho. \tag{31}$$

Substituting Eqs. (28), (29) into Eq. (30) and Eq. (27) into Eq. (31), it is possible to express the potential energy $E_p = E_{\mathrm{int}} + E_{\mathrm{ext}}$ as a function of coefficients b_i and c_j. This objective function is then minimized by invoking a selected optimization algorithm [42,43]. Note that the last term in Eq. (30), $(1+\nu)\varepsilon_0^2/\lambda_0^2$, is constant and therefore does not need to be included in the objective function.

3. NUMERICAL EXAMPLE

A. Comparison of Power Series Solution with Finite Element Method

In this section, the numerical solution is constructed for a specific example of a prestressed liquid lens membrane. Consider a circular membrane characterized by initial thickness $\tilde{h} = 0.1$ mm, radius (in prestressed state) $a = 10$ mm, Young's modulus $E = 1.97$ MPa, and Poisson's ratio $\nu = 0.4$. The membrane is prestressed by biaxial inplane stress $\sigma_0 = 60$ kPa and then loaded by constant pressure $p = 0.001E = 1.97$ kPa.

Figure 2 shows the membrane displacements and the deformed shape computed using the optimization algorithm described in the previous section. Deflections $w(r)$ and radial displacements $u(r)$ were approximated by the series in Eq. (27) with $N = M = 5$, and the objective function to be minimized was the potential energy E_p. Optimization was performed in MATLAB [44] using the quasi-Newton algorithm with Hessian updates based on the Broyden–Fletcher–Goldfarb–Shanno method [42,43].

As seen in Fig. 2(c), the resulting shape of the deformed membrane is not spherical. To quantify the deviation from a perfect sphere, let us approximate function $g(r)$ that describes the deformed shape by

$$z(r) = z_{\max} + \frac{cr^2}{1+\sqrt{1-c^2r^2}} + \sum_{i=2}^{5} A_{2i}r^{2i}, \tag{32}$$

where $z_{\max}$ is the maximum deflection, $c = 1/R_0 = z''/(1+z'^2)^{3/2}$ is the apex curvature (R_0 is the apex radius of curvature, z' and z'' are the first and second derivatives of the approximated function $g(r)$ at $r = 0$), and A_{2i} are aspherical coefficients that characterize the deviation from an apex sphere in the vertical direction. For our example, the resulting parameters have the following values: $z_{\max} = 2.5590$ mm, $R_0 = +22.5530$ mm, $A_4 = +1.8528\cdot10^{-5}$ mm^{-3}, $A_6 = +2.1047\cdot10^{-8}$ mm^{-5}, $A_8 = +4.0574\cdot10^{-10}$ mm^{-7}, and $A_{10} = +1.3916\cdot10^{-12}$ mm^{-9}. The corresponding mean

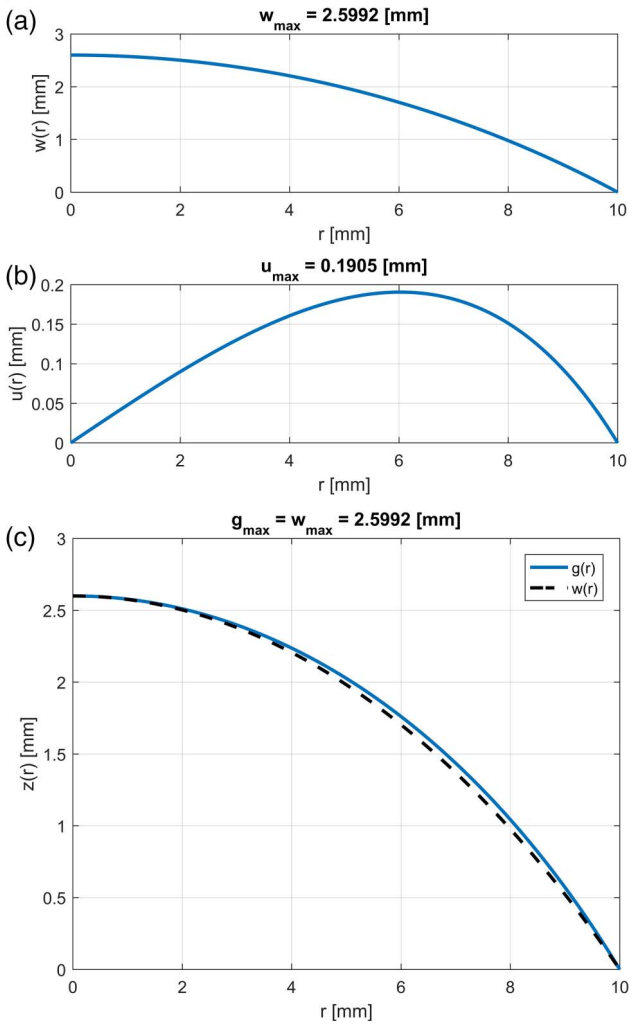


Fig. 2. Deformed membrane approximated by power series: (a) deflections $w(r)$, (b) radial displacements $u(r)$, (c) membrane shape $g(r)$.

quadratic approximation error is RMS $= 3.7 \cdot 10^{-6}$ mm. These results clearly show that the deformed shape indeed deviates from a sphere.

In a foregoing paper [38] the authors studied in detail the difference between the solutions of the general equations of Eq. (21) and their simplified form in Eq. (23). For the specific example considered here, a similar difference would arise. As clearly seen in Fig. 2(b), the radial displacement $u(r)$ is definitely not negligible while the simplified equations in Eq. (23) were derived using the assumptions that $u(r) \ll r$ and $u'(r) \ll 1$ for all $r \in [0, a]$.

Similar results were obtained in Ref. [38] for a membrane with no prestress. In that work, the authors studied the influence of membrane thickness on the deformed shape. The effect was found to be strong, and thus the imaging properties can be determined in a reliable way only if the thickness is known with high accuracy.

For comparison, the deformed shape of the membrane was also computed using the finite element method. Axisymmetric membrane elements with linear interpolation of both displacement components were implemented by the authors into the open-source simulation platform OOFEM [45]. The geometrical and material properties were the same as in the polynomial series approximation. Figure 3(a) shows the difference between the values of function $g(r)$ obtained using the approximation by polynomial series with $M = N = 4$ and by the finite element method (FEM) using a mesh consisting of 3200 linear finite elements. The root mean square of the difference, RMS $= 3.5$ nm, confirms that the results are comparable and both methods can be used. Approximation by polynomial series provides a good accuracy even with a relatively low number of polynomial terms, but the solution obtained by FEM converges in a more regular manner as the mesh is refined.

The diagrams in Fig. 3(b) indicate that FEM leads to a quadratic rate of convergence in terms of the maximum deflection (blue crosses) as well as the volume under the deformed membrane (orange crosses). The error was evaluated by comparing the results obtained with various meshes to those obtained with an extremely fine mesh consisting of 50,000 elements ($w_{\max} = 2.5992560$ mm, $V = 431.5885857$ mm^3). For 100 elements, the maximum deflection is determined with an error of about 270 nm, and then the error decreases 4 times whenever the number of elements is doubled. On the other hand, approximation by polynomials gives the maximum deflection with an error of about 13 nm already for $N = M = 3$, but as the order of the approximating polynomial increases, the error is not reduced in a regular manner and for higher orders it can even grow.

Let us now study the influence of the prestress on imaging properties of a plano-convex liquid membrane lens, in which

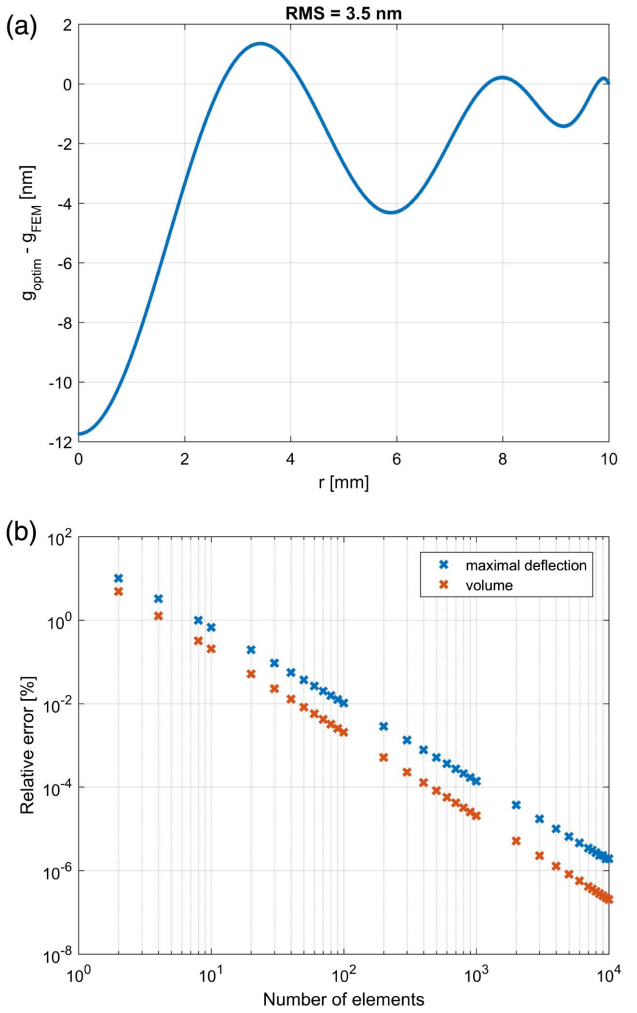


Fig. 3. (a) Difference between the deformed membrane shapes obtained by a power series approximation (g_{optim}) and by finite elements (g_{FEM}), and (b) dependence of the relative error on the number of finite elements.

one side of the lens chamber is formed by a plane-parallel plate and the other by a membrane. The membrane is supposed to have the same properties as in the previous example. The objective is to achieve an apex radius of curvature $R_0 = +22$ mm. Table 1 shows how the pressure p needed to deform the membrane to this desired shape depends on the prestress σ_0. For different values of the initial prestress, the pressure can be adjusted such that the paraxial imaging properties remain the same, but the general imaging properties will be different; this is documented by the variation of

Table 1. Dependence of Pressure and Membrane Shape on Initial Prestress, Provided that the Same Apex Radius $R_0 = +22$ mm is Maintained

σ_0 [kPa]	p [kPa]	$z_{\max}$ [mm]	A_4 [10^{-5} mm^{-3}]	A_6 [10^{-8} mm^{-5}]	A_8 [10^{-10} mm^{-7}]	A_{10} [10^{-12} mm^{-9}]
0	1.611	2.7600	+2.0279	+0.1181	−3.8856	+7.3839
30	1.851	2.7117	+1.8545	+8.8108	−1.2752	+4.6678
60	2.106	2.6745	+1.6965	+7.0029	−0.0666	+3.1490
90	2.372	2.6460	+1.5587	+5.8806	+0.4570	+2.2675

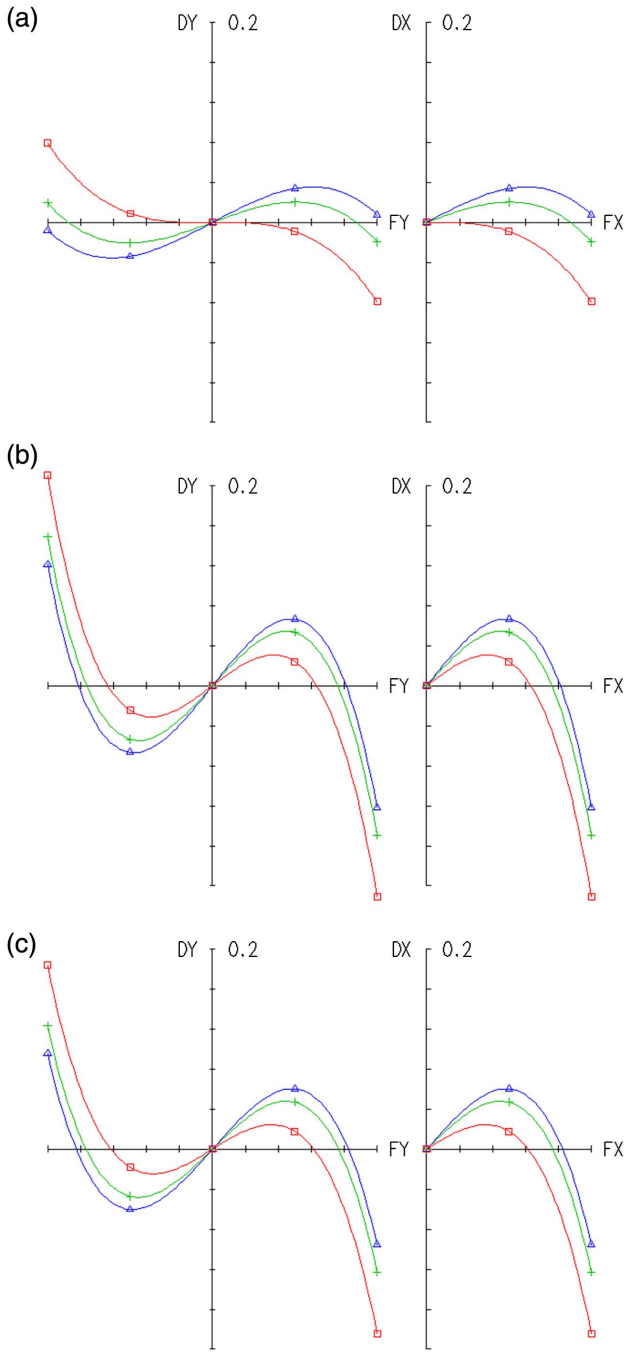


Fig. 4. Ray aberrations for (a) a spherical membrane of radius $R_0 = 22$ mm, (b) an aspherical membrane obtained when no initial prestress is applied, (c) an aspherical membrane obtained when the initial prestress is set to $\sigma_0 = 60$ kPa; DX a DY denote transversal ray aberrations, and FX a FY denote the height of an incident ray in the interval between −5 mm and 5 mm. Individual curves correspond to values obtained for wavelengths 486 nm (red), 587 nm (green), and 656 nm (blue).

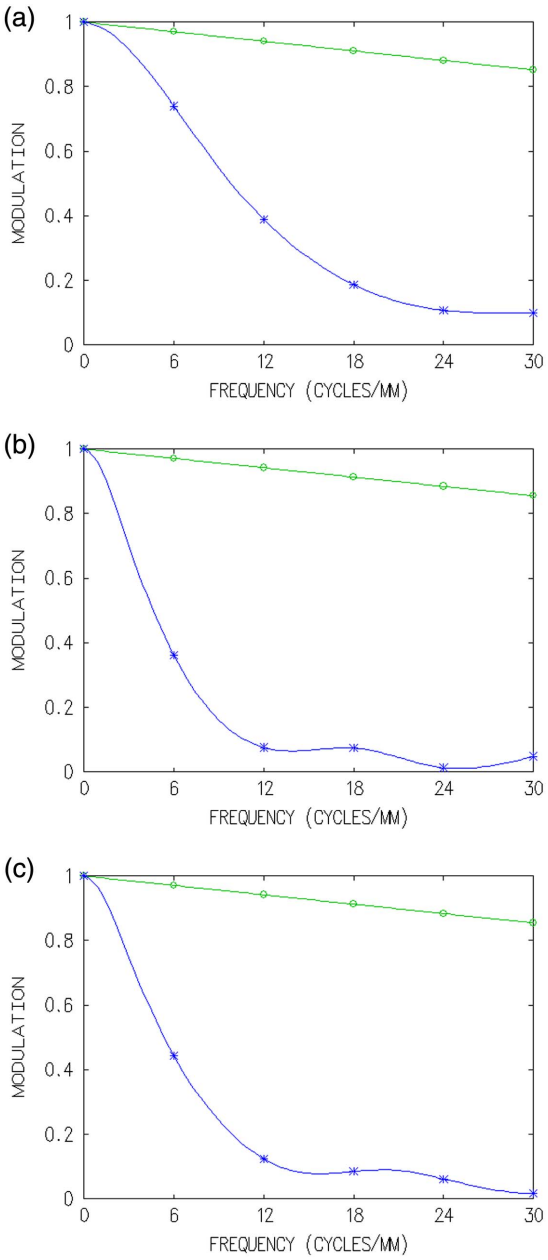


Fig. 5. Polychromatic MTF for (a) a spherical membrane of radius $R_0 = 22$ mm, (b) an aspherical membrane obtained when no initial prestress is applied, (c) an aspherical membrane obtained when the initial prestress is set to $\sigma_0 = 60$ kPa.

aspherical coefficients. Figure 4 shows the ray aberrations and Fig. 5 shows the polychromatic modulation transfer function (MTF) for (a) a spherical membrane of radius $R_0 = 22$ mm, (b) an aspherical membrane shape obtained if no prestress is applied, and (c) spherical membrane shape obtained if the prestress is set to $\sigma_0 = 60$ kPa. In Fig. 4, DX a DY denote transversal ray aberrations, and FX a FY denote the height of an incident ray in the interval between −5 mm and 5 mm.

The results indicate that the imaging properties of aspherical shapes obtained with different prestress values are almost the same; in other words, the prestress has a negligible influence on the imaging properties provided that the applied pressure is properly adjusted. On the other hand, deviations from the imaging properties of a spherical lens are non-negligible.

B. Dependence of Maximum Membrane Deflection on Material Parameters and Initial Prestress

In experiments, a unique evaluation of membrane material properties and of the initial prestress is often difficult. Therefore, it is useful to exploit the measured deflection for parameter identification.

Figure 6 shows the dependence of the maximum deflection $w_{\max}$ on pressure p for a fixed prestress value σ_0 and for selected values of Young's modulus E for a membrane with other parameters taken by the same values as in the previous examples. Also shown is the dependence of the maximum deflection $w_{\max}$ on pressure p for a fixed value of Young's modulus E and selected values of prestress σ_0. By measuring the actual dependence of the maximum deflection on applied pressure and comparing it to the plotted curves, membrane parameters E and σ_0 can be determined. This approach was exploited for parameter identification, as will be described in the following section.

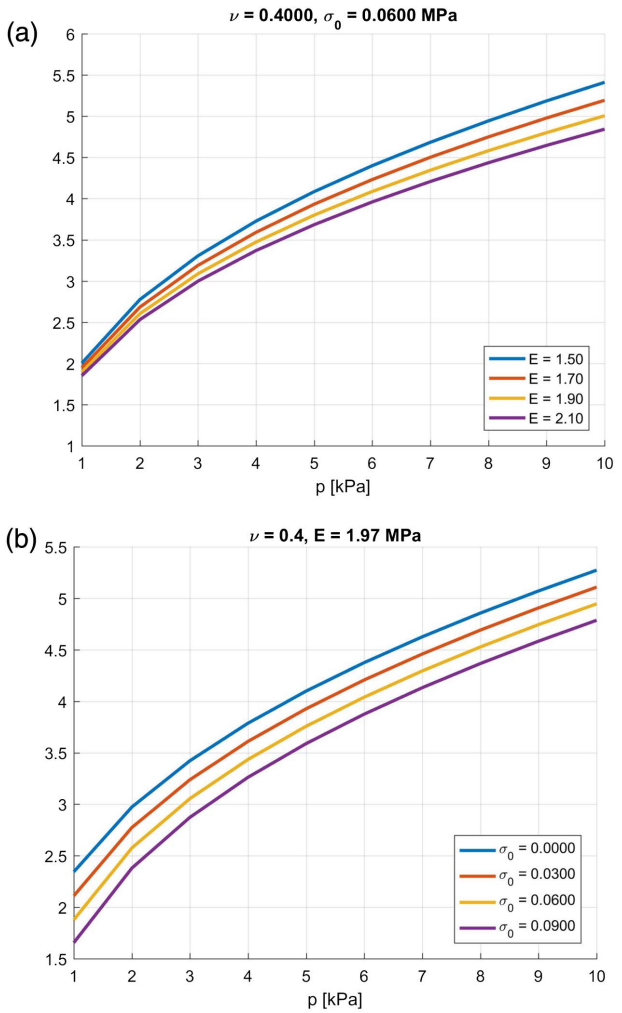


Fig. 6. (a) Dependence of the maximum deflection $w_{\max}$ on pressure p for a fixed value of prestress $\sigma_0 = 60$ kPa and selected values of Young's modulus E, (b) dependence of the maximum deflection $w_{\max}$ on pressure p for a fixed value of Young's modulus $E = 1.97$ MPa and selected values of prestress σ_0.

4. EXPERIMENTAL CONFIRMATION OF GOVERNING EQUATIONS

For validation of the theoretical results by real data, the deformed shape of a membrane of a liquid lens was measured experimentally, using a coordinate scanning device with a chromatic confocal sensor [46]. Figure 7 depicts the experimental setup for topography measurements of optical surfaces and the sample of a liquid lens. The liquid lens is formed by a chamber filled by an optical liquid (in our case distilled water), which is covered by an elastic membrane. The membrane is fixed by screws. By injecting (or removing) the liquid, the volume of the chamber is changed and the membrane is deformed. The membrane shape is then measured using the scanning device, which can determine the position of a measured point with an error of 0.02 mm (combining the mechanical uncertainty with the uncertainty induced by the sensor).

The membrane was made of Sylgard 184 [47]. The measured initial membrane thickness was $\tilde{h} = 0.45$ mm, and the radius measured from the support in the prestressed state was $a = 14.45$ mm. The Poisson ratio was taken as $\nu = 0.41$, and the Young modulus $E = 1.078$ MPa and prestress $\sigma_0 = 0.0583$ MPa were determined using the identification procedure described in the previous section (by optimizing the agreement between the measured and computed dependence of maximum deflection on applied pressure). Owing to the high slope of the membrane surface near the support and in view of the limitations of the measurements by a chromatic confocal sensor, permitting only small deviations of the measured direction from the normal to the deformed surface, the measurements focused on the region near the apex. Since the membrane remains axially symmetric, only the meridian was measured.

Figure 8 compares the measured and simulated shapes for pressures $p = 1.00$, 2.00, 3.00, and 4.00 kPa. The root mean square of the difference between the experiments and numerical simulations is not greater than 0.021 mm (as indicated in the figure legend), which corresponds to the estimated uncertainty of measurements, and thus the theoretical model can be considered as validated. Consequently, the results of simulations can be used in the design of optical systems and analysis of their imaging properties, e.g., of the spherical aberration, the point spread function, and the modulation transfer function.

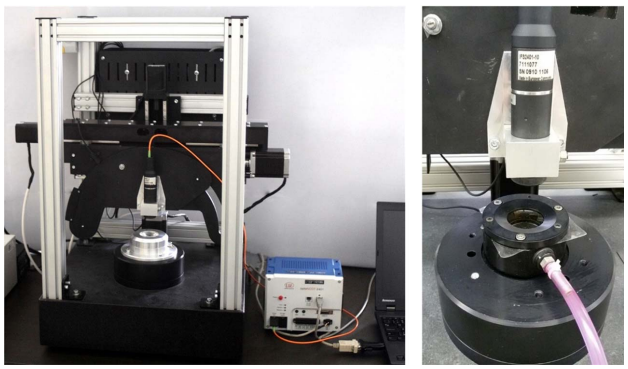


Fig. 7. Device for measurement of the shape of optical surfaces and an experimental specimen of a membrane liquid lens.

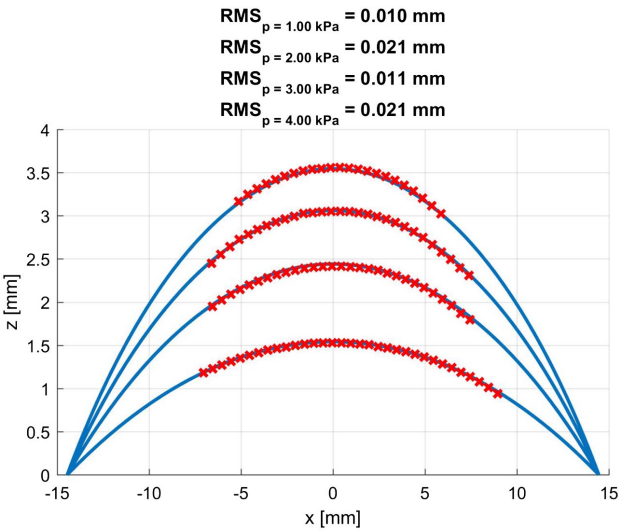


Fig. 8. Comparison of the experimentally measured and numerically simulated shapes of a membrane liquid lens for applied pressures $p = 1.00$, 2.00, 3.00, and 4.00 kPa.

5. CONCLUSION

In this paper, we have presented the theoretical description and numerical solution of an axisymmetric circular membrane of a liquid lens with an initial prestress. The model is based on the exact relation between displacements and the Green–Lagrange strain, on the Saint Venant–Kirchhoff material law, and on the principle of minimum potential energy. The corresponding governing equations have been derived, and a numerical method based on polynomial approximations of displacement functions and minimization of potential energy has been developed. The solution using polynomial approximations has been verified by comparing the results to those obtained by the finite element method, and the model has been validated by comparison of simulations to experimental data. A simple parametric study has shown how the maximum deflection depends on the elastic modulus, initial prestress, and applied pressure.

The presented numerical results and experimental measurements confirm that the selected model of a prestressed membrane is suitable for an accurate description of deformed liquid lens membranes up to the large-strain regime. Therefore, the model can be used for further analyses and simulations of imaging properties of these optical elements. A method of parameter identification has been presented that can be useful for characterization of mechanical properties of membrane samples in practical applications. It has also been documented by an example that the initial prestress has almost no influence on the imaging properties of a plane-convex membrane lens.

Funding. Ceské Vysoké Ucení Technické v Praze (CVUT) (SGS17/004/OHK1/1T/11).

REFERENCES

1. H. Ren and S. Wu, *Introduction to Adaptive Lenses*, Wiley Series in Pure and Applied Optics (Wiley, 2012).
2. Optotune, http://www.optotune.com.
3. Holochip, http://www.holochip.com.
4. A. H. Rawicz and I. Mikhailenko, "Modeling a variable-focus liquid-filled optical lens," Appl. Opt. **35**, 1587–1589 (1996).
5. D.-Y. Zhang, N. Justis, V. Lien, Y. Berdichevsky, and Y.-H. Lo, "High-performance fluidic adaptive lenses," Appl. Opt. **43**, 783–787 (2004).
6. N. Sugiura and S. Morita, "Variable-focus liquid-filled optical lens," Appl. Opt. **32**, 4181–4186 (1993).
7. H. Ren, D. Fox, P. A. Anderson, B. Wu, and S.-T. Wu, "Tunable-focus liquid lens controlled using a servo motor," Opt. Express **14**, 8031–8036 (2006).
8. Q. Yang, P. Kobrin, C. Seabury, S. Narayanaswamy, and W. Christian, "Mechanical modeling of fluid-driven polymer lenses," Appl. Opt. **47**, 3658–3668 (2008).
9. G.-H. Feng and Y.-C. Chou, "Flexible meniscus/biconvex lens system with fluidic-controlled tunable-focus applications," Appl. Opt. **48**, 3284–3290 (2009).
10. H. Ren and S.-T. Wu, "Variable-focus liquid lens," Opt. Express **15**, 5931–5936 (2007).
11. D. Shaw and T. E. Sun, "Optical properties of variable-focus liquid-filled optical lenses with different membrane shapes," Opt. Eng. **46**, 024002 (2007).
12. D. Shaw and C.-W. Lin, "Design and analysis of an asymmetrical liquid-filled lens," Opt. Eng. **46**, 123002 (2007).
13. H. Choi, D. S. Han, and Y. H. Won, "Adaptive double-sided fluidic lens of polydimethylsiloxane membranes of matching thickness," Opt. Lett. **36**, 4701–4703 (2011).
14. L. Li, Q.-H. Wang, and W. Jiang, "Liquid lens with double tunable surfaces for large power tunability and improved optical performance," J. Opt. **13**, 115503 (2011).
15. F. Schneider, J. Draheim, R. Kamberger, P. Waibel, and U. Wallrabe, "Optical characterization of adaptive fluidic silicone-membrane lenses," Opt. Express **17**, 11813–11821 (2009).
16. Y.-K. Fuh, M.-X. Lin, and S. Lee, "Characterizing aberration of a pressure-actuated tunable biconvex microlens with a simple spherically-corrected design," Opt. Lasers Eng. **50**, 1677–1682 (2012).
17. A. Miks, J. Novak, and P. Novak, "Algebraic and numerical analysis of imaging properties of thin tunable-focus fluidic membrane lenses with parabolic surfaces," Appl. Opt. **52**, 2136–2144 (2013).
18. L. Wang, H. Oku, and M. Ishikawa, "Development of variable-focus lens with liquid-membrane-liquid structure and 30 mm optical aperture," Proc. SPIE **8617**, 861706 (2013).
19. L. Wang, H. Oku, and M. Ishikawa, "An improved low-optical-power variable focus lens with a large aperture," Opt. Express **22**, 19448–19456 (2014).
20. S. T. Choi, B. S. Son, G. W. Seo, S.-Y. Park, and K.-S. Lee, "Opto-mechanical analysis of nonlinear elastomer membrane deformation under hydraulic pressure for variable-focus liquid-filled microlenses," Opt. Express **22**, 6133–6146 (2014).
21. D. Liang and X.-Y. Wang, "A bio-inspired optical system with a polymer membrane and integrated structure," Bioinspiration Biomimetics **11**, 066008 (2016).
22. J.-W. Du, X.-Y. Wang, and D. Liang, "Bionic optical imaging system with aspheric solid-liquid mixed variable-focus lens," Opt. Eng. **55**, 023105 (2016).
23. N. Hasan, A. Banerjee, H. Kim, and C. H. Mastrangelo, "Tunable-focus lens for adaptive eyeglasses," Opt. Express **25**, 1221–1233 (2017).
24. H. Hencky, "On the stress state in circular plates with vanishing bending stiffness," Z. Math. Phys. **63**, 311–317 (1915).
25. W. Z. Chein, "Asymptotic behavior of a thin clamped circular plate under uniform normal pressure at very large deflection," Sci. Rep. **A5**, 71–94 (1948).
26. J. D. Campbell, "On the theory of initially tensioned circular membranes subjected to uniform pressure," Q. J. Mech. Appl. Math. **9**, 84–93 (1956).
27. M. A. Goldberg and A. B. Pifko, "Large deflection analysis of uniformly loaded annular membranes," AIAA J. **1**, 2111–2115 (1963).
28. A. B. Pifko and M. A. Goldberg, "Iterative and power series solutions for the large deflection of an annular membrane," AIAA J. **2**, 2234 (1964).
29. N. Perrone and R. Kao, "A general nonlinear relaxation technique for solving nonlinear problems in mechanics," J. Appl. Mech. **38**, 371–376 (1971).

30. R. Kao and N. Perrone, "Large deflections of axisymmetric circular membranes," Int. J. Solids Struct. **7**, 1601–1612 (1971).
31. A. Kelkar, W. Elber, and I. Raju, "Large deflections of circular isotropic membranes subjected to arbitrary axisymmetric loading," Comput. Struct. **21**, 413–421 (1985).
32. W. Fichter, "Some solutions for the large deflections of uniformly loaded circular membranes," Technical paper 3658 (National Aeronautics and Space Administration, Langley Research Center, 1997).
33. D. Allman, "Variational solutions for the nonlinear deflexion of an annular membrane under axial load," Int. J. Mech. Sci. **24**, 749–753 (1982).
34. M. Sheploak and J. Dugundji, "Large deflections of clamped circular plates under initial tension and transitions to membrane behavior," J. Appl. Mech. **65**, 107–115 (1998).
35. F. Zhao, "Nonlinear solutions for circular membranes and thin plates," Proc. SPIE **6926**, 69260W (2008).
36. B. Stanford and P. Ifju, "The validity range of low fidelity structural membrane models," Exp. Mech. **48**, 697–711 (2008).
37. A. Mikš and P. Novák, "Calculation of a surface shape of a pressure actuated membrane liquid lens," Opt. Lasers Eng. **58**, 60–66 (2014).
38. P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek, "Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure," Appl. Opt. **56**, 5939–5947 (2017).
39. S. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells*, Engineering Societies Monographs (McGraw-Hill, 1959).
40. A. Volmir, "Flexible plates and shells," Technical Report (Air Force Flight Dynamics Laboratory, Research and Technology Division, Air Force Systems Command, 1967).
41. B. Audoly and Y. Pomeau, *Elasticity and Geometry: From Hair Curls to the Non-linear Response of Shells* (Oxford University, 2010).
42. M. Aoki, *Introduction to Optimization Techniques: Fundamentals and Applications of Nonlinear Programming*, Macmillan Series in Applied Computer Science (Macmillan, 1971).
43. L. E. Scales, *Introduction to Non-Linear Optimization*, Macmillan Computer Science Series (Springer-Verlag, 1985).
44. MathWorks, http://www.mathworks.com.
45. B. Patzák, "OOFEM—an object-oriented simulation tool for advanced modeling of materials and structures," Acta Polytech. **52**, 59–66 (2012).
46. Micro Epsilon, http://www.micro-epsilon.co.uk/displacement-position-sensors/confocal-sensor/.
47. "Sylgard 184 silicone elastomer kit," http://www.dowcorning.com/applications/search/products/Details.aspx?prod=01064291.

Research Article Vol. 38, No. 1 / January 2021 / Journal of the Optical Society of America A 99

Journal of the Optical Society of America A

OPTICS, IMAGE SCIENCE, AND VISION

Hybrid liquid-membrane lenses made by one fixed and one or two active optical components

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 23 September 2020; revised 25 November 2020; accepted 25 November 2020; posted 30 November 2020 (Doc. ID 410892); published 17 December 2020

The paper presents a methodology of calculation of the inner structure of two- and three-component hybrid liquid-membrane lenses with variable focal length that have corrected spherical aberration and coma. Specifically, the formulas for calculation of initial-design inner parameters (radii of curvatures of individual surfaces, axial thickness, and refractive indices of a material of the lens) of a thin-lens system are derived for a hybrid two-component system (doublet) made by one glass and one liquid-membrane lens, and a hybrid three-component lens (triplet) made by one glass lens and two liquid-membrane lenses, which both have variable focal length and corrected spherical aberration and coma for an object at infinity. As optimization during the optical design process requires the starting point be very close to the optimal solution, the presented approach can be successfully used for its calculation, as it is based on fundamental proven formulas of optical aberrations.

https://doi.org/10.1364/JOSAA.410892

1. INTRODUCTION

Liquid-membrane lenses have found many practical applications in many fields of science and technology, due especially to a possibility to change their imaging properties (focal length) in a prescribed way with variation of the inner pressure p in liquid, which is not possible with classic glass lenses [1–37]. The change in inner pressure p can be easily realized with a variation of a volume of the liquid in between membranes. These systems can find applications in scanning devices, for example, where the scanning unit is placed behind the active lens with variable focal length, and one demands similar properties of the laser beam spot among whole field of view.

While designing optical systems composed of many members having variable focal length, first, the outer parameters of the system are calculated (i.e., focal lengths and axial distances in between individual components). This topic is studied in detail in previous authors' works, e.g., [19–21]. The next step of the design is to determine the inner structure (parameters) of individual components in such a way that the optical system fulfills conditions required on optical imaging quality. In a previous paper [28], the authors present only the case for a specific lens (plano-convex lens with variable focal length), from which the optical system can be constructed, while the surfaces of individual lenses can be both spherical or aspherical. Detailed analysis and calculation of individual membrane deformation was presented by the authors in [30–32].

The goal of this paper is to derive formulas that can be used for calculation of inner parameters (i.e., radii of curvatures of individual surfaces, axial thickness, and refractive indices of a material of the lens) of a hybrid two-component system (doublet) made by one glass and one liquid-membrane lens that has variable focal length and corrected spherical aberration and coma [38,39] for an object at infinity. Further, the paper presents formulas for a hybrid three-component lens (triplet) made by one glass lens and two liquid-membrane lenses of similar properties, i.e., with corrected spherical aberration and coma.

While the liquid-membrane lenses are used for the mentioned hybrid systems, one should discuss the effect of gravity, which can affect deformation of the membrane, and hence the optical aberrations, especially if the lens is oriented vertically. However, as one can conclude from earlier works [7,24], this effect can be neglected for large applied pressures inside the lenses. Therefore, the influence of gravity on the inner-parameters design, which is presented in this paper, will not be considered.

During the optical design process, an initial guess of the optimization process has to be estimated or calculated as close as possible to the optimal global solution to the problem. Results of this paper, as they are based on analytic derivations coming from fundamental formulas of optical aberrations, can be used as a very strong tool for calculation of the starting point of further optimization processes.

2. LIQUID-MEMBRANE LENSES AND THEIR PROPERTIES

A scheme of a simple liquid-membrane lens with both surfaces made by membranes is shown in Fig. 1. In between the membranes, an optical liquid is filled. Many liquids can be used for optical applications (see, e.g., [40,41]). Thicknesses and shapes of surfaces of those membranes can in general be different, or one can be replaced by a glass plane–parallel plate (Fig. 2).

A two-component hybrid liquid-membrane lens (doublet) can be made by one glass lens and one membrane, as shown in Fig. 3. In the case of the hybrid doublet, the spherical aberration and coma can be corrected for a specific value of the focal length, which is not possible in the cases shown in Figs. 1 and 2. However, for a different value of the focal length, the correction of the spherical aberration and coma will be affected, as one has only one free parameter for the design—the radius of curvature of the membrane.

To correct the spherical aberration, coma, and a longitudinal chromatic aberration for a specific value of the focal length, one can design a hybrid three-component lens with two membranes (triplet), schematically shown in Fig. 4. For a different value of the focal length of this lens, the correction of coma and chromatic aberration will be affected; the lens has two free parameters only—the radii of curvatures of the first and second membranes.

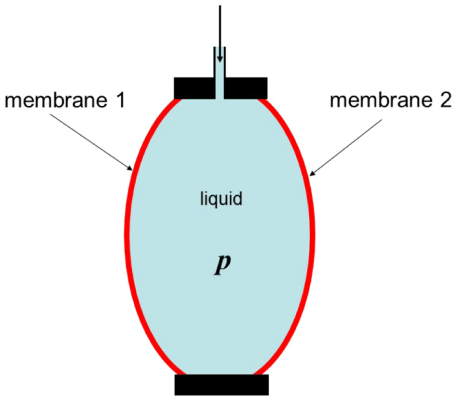


Fig. 1. Scheme of liquid-membrane lens with two membranes.

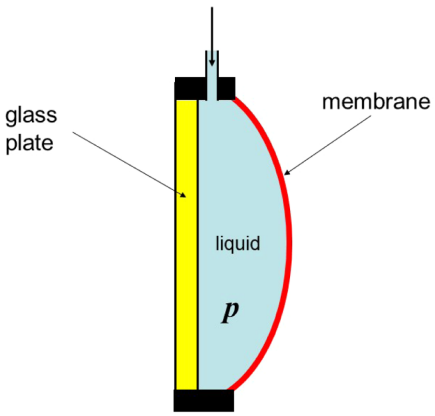


Fig. 2. Scheme of liquid-membrane lens with one membrane.

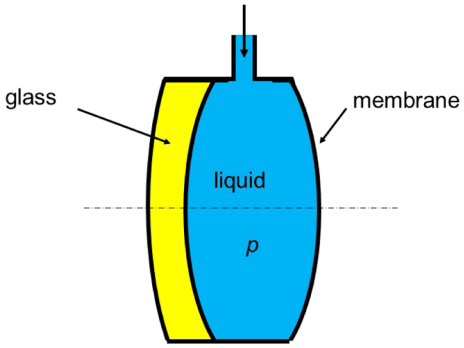


Fig. 3. Scheme of two-component hybrid liquid-membrane lens (doublet) with one membrane.

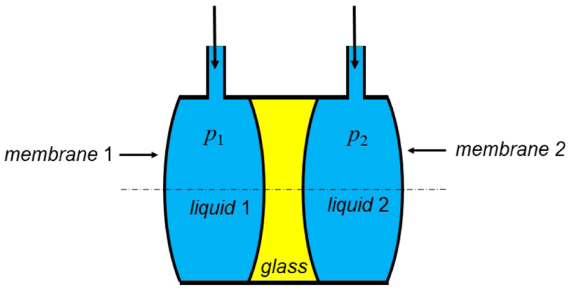


Fig. 4. Scheme of three-component hybrid liquid-membrane lens (triplet) with two membranes.

3. SEIDEL ABERRATION COEFFICIENTS

Assume a thin-lens optical system in air that consists of K lenses. Seidel aberration coefficients (Seidel sums) [38,39] S_{I}, S_{II}, S_{III}, S_{IV}, and S_{V} are then given by the following formulas:

$$S_{\mathrm{I}} = \sum_{i=1}^{K} h_i^4 M_i, \tag{1}$$

$$S_{\mathrm{II}} = \sum_{i=1}^{K} h_i^3 \bar{h}_i M_i + \sum_{i=1}^{K} h_i^2 N_i, \tag{2}$$

$$S_{\mathrm{III}} = \sum_{i=1}^{K} h_i^2 \bar{h}_i^2 M_i + 2\sum_{i=1}^{K} h_i \bar{h}_i N_i + \sum_{i=1}^{K} \varphi_i, \tag{3}$$

$$S_{\mathrm{IV}} = \sum_{i=1}^{K} \frac{\varphi_i}{n_i}, \tag{4}$$

$$S_{\mathrm{V}} = \sum_{i=1}^{K} h_i \bar{h}_i^3 M_i + 3\sum_{i=1}^{K} \bar{h}_i^2 N_i + \sum_{i=1}^{K} \frac{\bar{h}_i}{h_i}\left(3 + \frac{1}{n_i}\right)\varphi_i, \tag{5}$$

where

$$M_i = \varphi_i^3 (A_i X_i^2 + B_i X_i Y_i + C_i Y_i^2 + D_i),$$

$$N_i = \varphi_i^2 (E_i X_i + G_i Y_i), \tag{6}$$

where

$$A_i = \frac{n_i + 2}{4n_i(n_i - 1)^2}, \quad B_i = \frac{n_i + 1}{n_i(n_i - 1)}, \quad C_i = \frac{3n_i + 2}{4n_i},$$

$$D_i = \frac{{n_i}^2}{4(n_i - 1)^2}, \quad E_i = B_i/2, \quad G_i = \frac{2n_i + 1}{2n_i},$$

$$\varphi_i = (n_i - 1)\left(\frac{1}{r_i} - \frac{1}{r_i'}\right), \quad X_i = \frac{r_i' + r_i}{r_i' - r_i},$$

$$Y_i = \frac{s_i' + s_i}{s_i' - s_i} = \frac{m_i + 1}{m_i - 1} = -1 - \frac{2}{s_i\varphi_i} = 1 - \frac{2}{s_i'\varphi_i},$$

$$Y_{i+1} = \frac{h_i\varphi_i}{h_{i+1}\varphi_{i+1}}(Y_i - 1) - 1. \tag{7}$$

In the aforementioned formulas, h_i denotes the incidence height of a paraxial aperture ray (auxiliary aperture ray) at the ith lens, $\bar{h}_i$ is the incidence height of a paraxial principal ray (auxiliary principal ray) at the ith lens, r_i, r_i' are the vertex radii of curvatures of the ith lens, s_i, s_i' are, respectively, object and image distances from the ith lens, n_i is the refractive index of the ith lens, φ_i is the optical power of the ith lens, S_{I} is the Seidel sum for spherical aberration, S_{II} is the Seidel sum for coma, S_{III} is the Seidel sum for astigmatism, S_{IV} is the Seidel sum for field curvature (Petzval sum), and S_{V} is the Seidel sum for distortion.

One can calculate radii of curvatures of the ith lens, in the case when the lens power φ_i, refractive index n_i, and shape parameters X_i are known. Using Eq. (7), it holds:

$$r_i = \frac{2(n_i - 1)}{\varphi_i(X_i + 1)}, \quad r_i' = \frac{2(n_i - 1)}{\varphi_i(X_i - 1)}. \tag{8}$$

Transverse ray aberrations $\delta y'$ and $\delta x'$ in the image plane of the optical system composed by K thin lenses in air can be calculated from the formulas

$$\delta y' = -\frac{y_P(y_P^2 + x_P^2)}{2(s_1 - \bar{s}_1)^3 u_1^3 u_K'} S_{\mathrm{I}} + \frac{y_1(3y_P^2 + x_P^2)}{2(s_1 - \bar{s}_1)^3 u_1^2 u_K' \bar{u}_1} S_{\mathrm{II}}$$
$$- \frac{y_1^2 y_P}{2(s_1 - \bar{s}_1)^3 u_1 u_K' \bar{u}_1^2}(3S_{\mathrm{III}} + I^2 S_{\mathrm{IV}}) + \frac{y_1^3}{2(s_1 - \bar{s}_1)^3 u_K' \bar{u}_1^3} S_{\mathrm{V}},$$

$$\delta x' = -\frac{x_P(y_P^2 + x_P^2)}{2(s_1 - \bar{s}_1)^3 u_1^3 u_K'} S_{\mathrm{I}} + \frac{2y_1 y_P x_P}{2(s_1 - \bar{s}_1)^3 u_1^2 u_K' \bar{u}_1} S_{\mathrm{II}}$$
$$- \frac{y_1^2 x_P}{2(s_1 - \bar{s}_1)^3 u_1 u_K' \bar{u}_1^2}(S_{\mathrm{III}} + I^2 S_{\mathrm{IV}}), \tag{9}$$

where x_P, y_P denote coordinates of the ray in the plane of the entrance pupil of the optical system, y_1 is the distance of the object point from the optical axis, s_1 is the distance of the object plane from the first surface of the optical system, $\bar{s}_1$ is the entrance pupil distance from the first lens of the optical system, $u_1 = h_1/s_1$ is the angle of paraxial aperture ray in the object space, $\bar{u}_1 = \bar{h}_1/\bar{s}_1$ is the angle of principal ray in the object space, $u_K' = u_1/m$ is the angle of paraxial aperture ray in the image space, and m is the transverse magnification of the optical system. Furthermore, I denotes the Lagrange–Helmholtz invariant, which is defined for the optical system in air as follows:

$$I = h_1\bar{h}_1\left(\frac{1}{s_1} - \frac{1}{\bar{s}_1}\right) = u_1\bar{h}_1 - \bar{u}_1 h_1. \tag{10}$$

Without loss of generality, one can suppose unit Lagrange–Helmholtz invariant, i.e., $I = 1$, and one can set $h_1 = 1$. Afterwards, it follows that

$$\bar{h}_1 = \frac{s_1\bar{s}_1}{\bar{s}_1 - s_1}. \tag{11}$$

The so-called Seidel difference formula (in the case of $I = 1$ and $h_1 = 1$) between h and $\bar{h}$ is then given with the formula

$$\bar{h}_j = h_j\left(\bar{h}_1 + \sum_{i=2}^{j}\frac{d_{i-1}}{h_{i-1}h_i}\right), \tag{12}$$

where d_i denotes the axial distance between ith and $(i + 1)$th lenses.

In a general case, one can express Seidel aberration coefficients for the centered optical system of J spherical surfaces as follows [38,39]:

$$S_{\mathrm{I}} = \sum_{i=1}^{J} U_i, \quad S_{\mathrm{II}} = \sum_{i=1}^{J} U_i V_i, \quad S_{\mathrm{III}} = \sum_{i=1}^{J} U_i V_i^2,$$

$$S_{\mathrm{IV}} = \sum_{i=1}^{J} P_i, \quad S_{\mathrm{V}} = \sum_{i=1}^{J}(U_i V_i^2 + I^2 P_i) V_i, \tag{13}$$

where

$$U_i = h_i\left(\frac{u_{i+1} - u_i}{1/n_{i+1} - 1/n_i}\right)^2\left(\frac{u_{i+1}}{n_{i+1}} - \frac{u_i}{n_i}\right),$$

$$V_i = \frac{\bar{u}_{i+1} - \bar{u}_i}{u_{i+1} - u_i}, \quad P_i = \frac{1}{h_i}\left(\frac{u_{i+1}}{n_i} - \frac{u_i}{n_{i+1}}\right), \tag{14}$$

where u_i and $\bar{u}_i$ denote paraxial angles of the aperture and chief rays incident at the ith surface of the optical system, n_i is the index of refraction in front of the ith surface, and h_i is the incident height of the paraxial aperture ray at the ith surface. Furthermore, it holds: $\bar{u}_1 = \bar{h}_1/\bar{s}_1 = y_0/(\bar{s}_1 - s_1)$, where y_0 is the object height. It is evident from the previous equations that one can use an arbitrary choice of input parameters (h_1, $u_1 = h_1/s_1$, $\bar{h}_1$, $\bar{u}_1 = \bar{h}_1/\bar{s}_1$) for calculation of the third-order aberration coefficients for a given object distance s_1 and a position $\bar{s}_1$ of the entrance pupil.

4. TWO-COMPONENT HYBRID LIQUID-MEMBRANE LENS

Consider now the hybrid liquid-membrane lens (doublet) composed of two lenses of optical powers φ_1 and φ_2, which are made by materials (glasses or liquids) with refractive indices and Abbe numbers n_1, ν_1, n_2, and ν_2. With an appropriate choice of materials, this doublet can have the Seidel coefficient S_{I} of spherical aberration, Seidel coefficient S_{II} of coma, and coefficient C_{I} of longitudinal chromatic aberration of required values for one specific focal length. However, the chromatic aberration cannot

be corrected in every case, as there is only a limited amount of types of optical liquid appropriate for those lenses.

One can see in Fig. 3 that a change in focal length can be realized by modifying the vertex radius of curvature of the membrane. If the doublet has corrected spherical aberration and coma for one specific focal length (one can choose this value), the aberrations will be affected for its different value.

Suppose now the first approximation of the design, i.e., the doublet is composed of two thin lenses and has optical power $\varphi = 1$. Further, let the doublet image an object at infinity ($s_1 = \infty$) and the entrance pupil be at vertex of the first surface ($\bar{s}_1 = 0$). With Eqs. (1) and (2) for $K = 2$, one can, after modification, derive the following polynomial of the fifth order for the optical power φ_1 of the first lens, and it holds:

$$a_5\varphi_1^5 + a_4\varphi_1^4 + a_3\varphi_1^3 + a_2\varphi_1^2 + a_1\varphi_1 + a_0 = 0, \tag{15}$$

where

$$a_5 = \frac{(n_1 - n_2)^3}{4{n_1}^2{n_2}^2(n_1 - 1)^4(n_2 - 1)^2},$$

$$a_4 = -\frac{(n_1 - n_2)^2(n_2 + 2)(n_1 - n_2 + n_1n_2)}{4{n_1}^2{n_2}^2(n_1 - 1)^4(n_2 - 1)^2},$$

$$a_3 = \frac{n_1 - n_2}{4{n_1}^2{n_2}^2(n_1 - 1)^4(n_2 - 1)^2}\left\{{n_1}^3[S_{II}(n_2 - {n_2}^2) - {n_2}^2]\right.$$
$$+ {n_1}^2[4n_2 + 6{n_2}^2 + 2{n_2}^3 + 1 - S_{II}(n_2 - {n_2}^3)]$$
$$\left. - n_1[3n_2 + 7{n_2}^2 + 5{n_2}^3 - S_{II}({n_2}^2 - {n_2}^3)] + {n_2}^2 + 2{n_2}^3\right\},$$

$$a_2 = \frac{(2{n_1}^3 - 4{n_1}^2 + 2n_1)S_{II} + {n_1}^3 - 8{n_1}^2 + 15n_1 - 2}{4{n_1}^2(n_1 - 1)^3(n_2 - 1)}$$
$$- \frac{(n_1 - 1)S_I + ({n_1}^2 - 2n_1)S_{II} + {n_1}^2 - 5n_1 + 1}{4{n_1}^2(n_1 - 1)^3}$$
$$- \frac{2(1 - n_1)S_I + {n_1}^2S_{II} + n_1}{4n_1n_2(n_1 - 1)^3},$$

$$a_1 = -\frac{n_1 - n_2}{4n_1{n_2}^2(n_1 - 1)^2(n_2 - 1)^2}[(2{S_{II}}^2 + 2S_{II} - 2S_I + 2){n_2}^3$$
$$+ (-4{S_{II}}^2 - 3S_{II} + 2S_I + 1){n_2}^2$$
$$+ (2{S_{II}}^2 + S_{II} + 2S_I)n_2 - 2S_I],$$

$$a_0 = -\frac{S_I - {S_{II}}^2}{4(n_1 - 1)^2}$$
$$- \frac{1}{{n_2}^2(n_1 - 1)^2(n_2 - 1)^2}\left[\left(-\frac{{S_{II}}^2}{2} - \frac{S_{II}}{4} + \frac{S_I}{2} - \frac{1}{4}\right){n_2}^3\right.$$
$$\left. + \left({S_{II}}^2 + \frac{S_{II}}{2} - \frac{3S_I}{4}\right){n_2}^2 + \left(-\frac{{S_{II}}^2}{2} - \frac{S_{II}}{4}\right)n_2 + \frac{S_I}{4}\right].$$

The optical power φ_2 of the second lens then can be calculated with the formula

$$\varphi_2 = 1 - \varphi_1. \tag{16}$$

Shape parameters X_1 and X_2 of the lens can be calculated from the following formulas:

$$X_1 = \frac{A}{B}, \quad X_2 = \frac{\varphi_1(X_1 - 1)(n_2 - 1)}{\varphi_2(n_1 - 1)} - 1, \tag{17}$$

where

$$A = (E_2 + G_1 - G_2 - E_2H){\varphi_1}^2 + (E_2H - 2E_2)\varphi_1 + E_2 + G_2 + S_{II},$$

$$B = (E_1 - E_2H){\varphi_1}^2 + (E_2H)\varphi_1,$$

$$H = \frac{n_2 - 1}{n_1 - 1}.$$

Afterwards, the radii of curvatures can be calculated with Eq. (8). Those values then can be used as a starting point (with corresponding axial thicknesses of individual lenses) for further doublet optimization with, for example, optical design software Zemax [42] or OSLO [43].

5. THREE-COMPONENT HYBRID LIQUID-MEMBRANE LENS

Consider now the situation of a three-component hybrid liquid membrane as in Fig. 4. The inner parameters of individual components of this hybrid triplet can be calculated following three steps.

1. The optical power φ_1 of the first lens of the membrane component is set, and the overall optical power of the membrane component is set $\varphi = 1$. Further, materials of individual lenses of the membrane component are selected, i.e., refractive indices and Abbe numbers. The optical power φ_2 of the second lens and φ_3 of the third lens then can be calculated as follows:

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3, \quad C_I = \varphi_1/\nu_2 + \varphi_2/\nu_3 + \varphi_3/\nu_4. \tag{18}$$

Further, one gets the following formulas for optical powers φ_2 and φ_3:

$$\varphi_2 = \frac{\nu_3(\nu_2\varphi - \nu_2\varphi_1 + \nu_4\varphi_1 - C_I\nu_2\nu_4)}{\nu_2(\nu_3 - \nu_4)},$$

$$\varphi_3 = -\frac{\nu_4(\nu_2\varphi - \nu_2\varphi_1 + \nu_3\varphi_1 - C_I\nu_2\nu_3)}{\nu_2(\nu_3 - \nu_4)}, \tag{19}$$

where ν_2, ν_3, and ν_4 are Abbe numbers of the lenses, and C_I is the coefficient of longitudinal chromatic aberrations of the membrane component [38,39]. With the use of the paraxial imaging equation [38,39]

$$n_{i+1}u_{i+1} - n_{i+1}u_{i+1} = h_i(n_{i+1} - n_i)/r_i, \quad i = 1, 2, 3, 4, \tag{20}$$

one gets for paraxial aperture angles u_i the following formulas:

$$u_2 = \varphi_1 - u_1 + \frac{2u_1}{n_2} - Q\left(\frac{1}{n_2} - 1\right),$$

$$u_3 = \varphi_1 - u_1 + \frac{2u_1}{n_3} - Q\left(\frac{1}{n_3} - 1\right),$$

$$u_4 = \varphi_1 - u_1 + \frac{2u_1 + \varphi_2}{n_4} - Q\left(\frac{1}{n_4} - 1\right) - \frac{\varphi_2(n_4 - 1)}{n_4(n_3 - 1)},$$

$$u_5 = u_1 + 1, \tag{21}$$

where u_1 is the paraxial aperture angle in the object space of the optical system, n_2, n_3, and n_4 are the indices of refraction of components' (lenses') materials, r_i is the radius of curvature of the ith optical surface, Q is Abbe invariant [38,39] on the second surface of the system, and h_i is the paraxial impinging height. In the presented case of the optical system of thin lenses, one can choose $h_i = 1$ ($i = 1, 2, 3, 4$). In Eq. (21), the air is supposed to be the object and image space of the optical system ($n_1 = n_5 = 1$). For the object at infinity, it holds: $u_1 = 0$. Substituting Eq. (21) into Eq. (13), one gets the following formula for the Seidel coefficient of spherical aberration S_I of the thin hybrid lens:

$$S_\mathrm{I} = a_2 Q^2 + a_1 Q + a_0, \tag{22}$$

where

$$a_2 = \frac{2\varphi_2}{n_3} - 2\frac{\varphi_1 + \varphi_2 - 1}{n_4} + \frac{n_2 + 2\varphi_1}{n_2},$$

$$a_1 = A_{10} + A_{11}u_1,$$

$$A_{11} = 8\frac{\varphi_1 + \varphi_2 - 1}{n_4} - 2\left(\frac{3n_2 + 4\varphi_1}{n_2}\right) - 8\frac{\varphi_2}{n_3},$$

$$A_{10} = \frac{3{\varphi_1}^2}{n_2 - 1} - \frac{{\alpha_1}^2}{n_4 - 1} - \frac{2n_3\varphi_2}{n_3 - 1}\left[\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2}\left(1 - \frac{(n_4 - 1)}{(n_3 - 1)}\right) - \frac{\varphi_1}{n_3}\right] - \frac{2n_4\alpha_1}{n_4 - 1}\left(\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2} - \alpha_2 - 1\right)$$
$$- \frac{{\varphi_2}^2}{{n_4}^2(n_3 - 1)^2}[{n_4}^2(n_3 - 1) - {n_3}^2(n_4 - 1)],$$

$$\alpha_1 = \varphi_1 + \frac{\varphi_2}{n_4} - \alpha_2 - 1, \quad \alpha_2 = \frac{\varphi_2(n_4 - 1)}{n_4(n_3 - 1)},$$

$$a_0 = A_{02}{u_1}^2 + A_{01}u_1 + A_{00},$$

$$A_{02} = 8\frac{n_2 + \varphi_1}{n_2} - 8\frac{\varphi_1 + \varphi_2 - 1}{n_4} + 8\frac{\varphi_2}{n_3},$$

$$A_{01} = \frac{5n_3 + 6\varphi_1 + 6\varphi_2 - 6n_3\varphi_1 - 6\varphi_1\varphi_2 + {\varphi_2}^2 - 5}{n_3 - 1} - \frac{(\varphi_1 + \varphi_2 - 1)}{n_4(n_3 - 1)(n_4 - 1)}[7n_4(n_3 - 1 + \varphi_1 + \varphi_2 - n_3\varphi_1)$$
$$+ \varphi_2 n_3(n_4 - 8)] - \frac{7{\varphi_1}^2}{n_2 - 1},$$

$$A_{00} = \frac{n_2{\varphi_1}^3}{(n_2 - 1)^2} - \frac{{n_4}^2}{(n_4 - 1)^2}\left[\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2} - \frac{\varphi_2(n_4 - 1)}{{n_4}^2(n_3 - 1)} - 1\right]\left[\varphi_1 + \frac{\varphi_2}{n_4} - \frac{\varphi_2(n_4 - 1)}{n_4(n_3 - 1)} - 1\right]^2$$
$$+ \frac{n_3{\varphi_2}^2(n_3 - n_4)}{{n_4}^2(n_3 - 1)^3}[n_3\varphi_2 + n_4\varphi_1(n_3 - 1)].$$

In the case of coma, one gets for the Seidel coefficient S_II of coma for the thin three-component lens, after substitution of Eq. (21) into Eq. (13), as follows ($\bar{s}_1 = 0$),

$$S_\mathrm{II} = b_0 + Qb_1, \tag{23}$$

where

$$b_1 = \frac{\varphi_1 + \varphi_2 - 1}{n_4} - \frac{n_2 + \varphi_1}{n_2} - \frac{\varphi_2}{n_3},$$

$$b_0 = B_{00} + B_{01}u_1,$$

$$B_{01} = 3 + 2\left(\frac{\varphi_1}{n_2} + \frac{\varphi_2}{n_3}\right) - 2\frac{\varphi_1 + \varphi_2 - 1}{n_4},$$

$$B_{00} = \frac{n_4}{n_4 - 1}\left[\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2} - \frac{\varphi_2(n_4 - 1)}{{n_4}^2(n_3 - 1)} - 1\right]$$
$$\times\left[\varphi_1 + \frac{\varphi_2}{n_4} - \frac{\varphi_2(n_4 - 1)}{n_4(n_3 - 1)} - 1\right] - \frac{{\varphi_1}^2}{n_2 - 1}$$
$$+ \frac{\varphi_2(n_3 - n_4)}{{n_4}^2(n_3 - 1)^2}[n_3\varphi_2 + n_4\varphi_1(n_3 - 1)].$$

2. Setting the Seidel coefficient of spherical aberration S_I equal to a desired value, one can calculate the parameter Q with Eq. (22). Afterwards, the radii of curvatures of individual surfaces of a thin three-component lens can be enumerated from the formulas

$$\frac{1}{r_1} = Q - u_1 + \frac{n_2\varphi_1}{n_2 - 1}, \quad \frac{1}{r_2} = Q - u_1 + \varphi_1,$$

$$\frac{1}{r_3} = Q - \frac{\varphi_2 + (u_1 - \varphi_1)(n_3 - 1)}{n_3 - 1},$$

$$\frac{1}{r_4} = Q - \frac{\varphi_3}{n_4 - 1} - \frac{\varphi_2 + (u_1 - \varphi_1)(n_3 - 1)}{n_3 - 1}. \quad (24)$$

3. For completing the calculation, one should multiply the values calculated from Eq. (24) by the desired value of the focal length of the appropriate component.

If one requires a specific value of the Seidel coefficient of coma S_{II}, then an appropriate value of the optical power φ_1 has to be chosen to obtain the coefficient S_{II} in Eq. (23).

With a variation of the focal length $f' = 1/\varphi$ of the hybrid triplet, the second glass lens remains unchanged, but the first and second membranes change their radii of curvatures. Therefore, one knows the optical power φ_2 and the shape factor X_2 of the glass lens. Suppose next that the object is at infinity. Further, one can derive the following formula for the Seidel coefficient of spherical aberration S_I from Eqs. (1)–(6), and it holds:

$$S_I = e_3{\varphi_1}^3 + e_2{\varphi_1}^2 + e_1\varphi_1 + e_0, \quad (25)$$

where

$$e_3 = A_1 - A_3 - B_1 + B_3 + C_1 - C_3 + D_1 - D_3,$$

$$e_2 = -\varphi_2(3A_3 - 3B_3 - 4C_2 + 3C_3 + 3D_3) + (3A_3 - B_3 - C_3 + 3D_3)\varphi + A(2A_1 - AB_1) + B(B_3 - 2A_3),$$

$$e_1 = -{\varphi_2}^2[3(A_3 - B_3 + C_3 + D_3) - 4C_2 + 2B_2X_2] + \varphi_2[2(3A_3 - B_3 - C_3 + 3D_3)\varphi + 2B(B_3 - 2A_3)] + (C_3 - B_3 - 3A_3 - 3D_3)\varphi^2 + (4A_3B)\varphi + A^2A_1 - A_3B^2,$$

$$e_0 = -{\varphi_2}^3(-A_2{X_2}^2 + B_2X_2 + A_3 - B_3 - C_2 + C_3 - D_2 + D_3) + {\varphi_2}^2[(3A_3 - B_3 - C_3 + 3D_3)\varphi + (B_3 - 2A_3)B] - \varphi_2[(3A_3 + B_3 - C_3 + 3D_3)\varphi^2 + (B - 4\varphi)A_3B] + (A_3 + B_3 + C_3 + D_3)\varphi^3 - (2A_3 + B_3)B\varphi^2 + (A_3B^2)\varphi,$$

$$A = \frac{\varphi_2(X_2 + 1)(n_1 - 1)}{n_2 - 1}, \quad B = \frac{\varphi_2(X_2 - 1)(n_3 - 1)}{n_2 - 1}.$$

Similarly, one gets for the Seidel coefficient of coma S_{II}

$$S_{II} = g_2{\varphi_1}^2 + g_1\varphi_1 + g_0, \quad (26)$$

where

$$g_2 = E_1 - E_3 - F_1 + F_3,$$

$$g_1 = 2E_3\varphi - \varphi_2(2E_3 + 2F_2 - 2F_3) + AE_1 - BE_3,$$

$$g_0 = -{\varphi_2}^2(E_3 + F_2 - F_3 - E_2X_2) + \varphi_2(2E_3\varphi - BE_3) - (E_3 + F_3)\varphi^2 + BE_3\varphi.$$

One can calculate the optical power φ_1 of the first lens from Eq. (25) for the Seidel coefficient of spherical aberration S_I. Afterwards, the optical power φ_3 of the third lens and shape parameters X_1 and X_3 of the first and the third lenses can be calculated from the formulas

$$\varphi_3 = \varphi - \varphi_1 - \varphi_2, \quad X_1 = \frac{A}{\varphi_1} + 1, \quad X_3 = \frac{B}{\varphi_3} - 1. \quad (27)$$

The radii of curvatures then can be calculated with Eq. (8).

6. EXAMPLES

A. Example 1

Consider now a situation of a laser scanner (monochromatic light) that has a hybrid liquid-membrane lens with variable focal length (doublet in Fig. 3) as an objective. Let the objective be in front of the scanning unit (e.g., rotating mirror), which steers the beam in a field of angles $w = \pm w_{max}$. To get the beam spot aligned in a line in the image plane behind the scanner, the focal length of the objective has to vary according the formula $f'_w = a + (f'_0 - a)/\cos w$, where a denotes the distance of an axis of rotation of the mirror from an image principal plane of the objective, f'_0 is the nominal focal length for $w = 0$ deg, and f'_w is the focal length of the objective for the angle w. One can see that the objective has to be corrected from spherical aberration only for this specific case.

The aforementioned case is the hybrid liquid-membrane lens (doublet) with a variable focal lens made by three spherical surfaces. The axial thickness d_1 of the first glass lens, with radii of curvatures r_1 and r_2, is constant, while the axial thickness d_2 of the liquid-membrane lens, with radii of curvatures r_2 and r_3, is changing according to the variation of the focal length of the objective, and the edge thickness of the doublet has to be constant (the construction demand). The glass lens is made from Schott N-BK7, and both the membrane and liquid are from Sylgard 184 [40]. The index of refraction n of Sylgard 184 for the wavelength λ can be calculated from the following formula:

$$n^2 - 1 = \frac{1.0093\lambda^2}{\lambda^2 - 0.013185}, \quad (28)$$

where the wavelength λ is in micrometers.

Let one study the influence of the change in focal length of the used hybrid liquid-membrane lens on its imaging quality. With Eqs. (15)–(17), one gets the following values of the radii of curvatures of the thin aplanatic doublet ($f' = 1$ mm): $\varphi_1 = -1.148\ \text{mm}^{-1}$, $r_1 = 0.3931$ mm, $r_2 = 0.2094$ mm, $r_3 = -4.3843$ mm. After the thickness consideration and optimization in software OSLO [43], the parameters of the designed hybrid liquid-membrane lens for focal length $f' = 100$ mm and its transverse diameter $D = 20$ mm are shown in Table 1, where the Strehl ratio ($S.R.$) [38,39] for the wavelength $\lambda = 635$ nm and imaging of the axial point is presented as well. The edge thickness was chosen as 3 mm, $a = 15$ mm, and maximal view angle $w_{max} = 51$ deg. Figure 5 shows the plot of aberrations for the object at infinity for $f' = 100$ mm, and Fig. 6 shows aberrations for $f' = 150$ mm.

With the change in pressure in the back liquid-membrane lens, the focal length of the doublet can be variable. Table 2

Table 1. Parameters of Two-Component Hybrid Liquid-Membrane Lens (Doublet) with $f' = 100$ mm, $D = 20$ mm, and $S.R. = 0.924$

r [mm]	d [mm]	Material
40.611	2.000	N-BK7
21.694	5.587	Sylgard 184
−345.076		

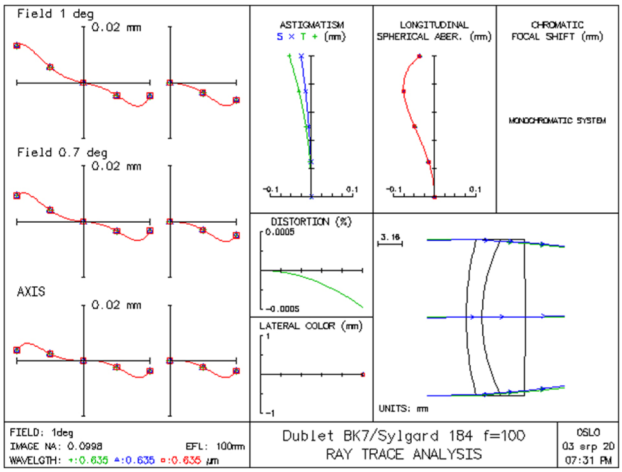


Fig. 5. Aberrations of two-component hybrid liquid-membrane lens (doublet) for $f' = 100$ mm.

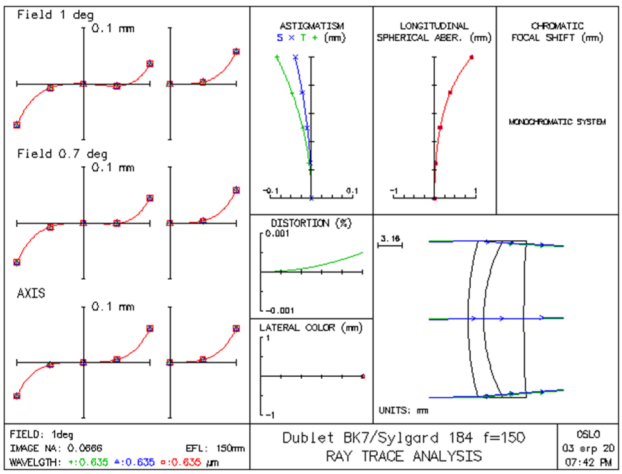


Fig. 6. Aberrations of two-component hybrid liquid-membrane lens (doublet) for $f' = 150$ mm.

presents values of the third radius of curvature r_3 of the membrane of the doublet, axial thickness d of the lens, and $S.R.$ for the wavelength 635 nm and imaging of the axial point, as a dependent variables on the focal length f' of the doublet.

One can see from the results in Table 2 that the imaging quality is very good for both focal distances. Further, one can deduct from the aforementioned situation that the F-number of the system $F = f'/D$ will vary according to the change in focal length, and, therefore, the spot size of the beam as well.

Table 2. Third Radius of Curvature r_3, Radial Thickness d, and Strehl Ratio of Two-Component Hybrid Liquid-Membrane Lens (Doublet) for $\lambda = 635$ nm for Imaging of Axial Point for Different Values of Focal Length

f' [mm]	r_3 [mm]	d [mm]	$S.R.$
100	−345.076	5.587	0.924
150	191.000	5.180	0.520

Moreover, the circular shape of the spot for the angle $w = 0$ deg becomes elliptical for different values of w [38,39].

B. Example 2

Consider now the situation of a thin three-component hybrid liquid-membrane lens (Fig. 4). The liquid in the first and third lenses is water, and the second lens is made by Schott SF6 glass. The focal length of this system is $f' = 1$ mm. Further, suppose Seidel coefficients of spherical aberration $S_{\mathrm{I}} = 0$ and coma $S_{\mathrm{II}} = 0$, and coefficient of longitudinal chromatic aberration $C_{\mathrm{I}} = 0$ as well for this focal length. After the solution to Eqs. (22)–(24), one gets (for an object at infinity) $X_1 = -1.6473\mathrm{e}-01$, $X_2 = 5.0623\mathrm{e}+00$, $X_3 = -1.9377\mathrm{e}+01$, $r_1 = 4.6090\mathrm{e}-01$ mm, $r_2 = -3.3052\mathrm{e}-01$ mm, $r_3 = -4.9325\mathrm{e}-01$ mm, $r_4 = -4.4483\mathrm{e}-01$ mm.

For different values of the focal length of the hybrid triplet, the optical power φ_2 and the shape coefficient X_2 remain constant. Optical powers and shape parameters of the first and second membrane lenses change, i.e., the radii of curvatures of the first and second membranes vary as well. Further, the recalculation can be processed with Eqs. (25)–(27).

Table 3 presents parameters of hybrid triplets for focal lengths $f' = 100$ mm, $f' = 150$ mm, and $f' = 200$ mm, for diameter of the beam $D = 20$ mm. Axial thicknesses of liquid-membrane lenses are calculated to get the edge thickness of the 3 mm lens for transverse diameter of the 22 mm lens. The $S.R.$ in the tables

Table 3. Parameters of Three-Component Hybrid Liquid-Membrane Lens (Triplet)

r [mm]	d [mm]	Material
$f' = 100$ mm, $D = 20$ mm, $S.R. = 0.990$, $\lambda = 635$ nm		
46.090	6.202	Water
−33.052	2.000	SF6
−49.325	3.139	Water
−44.483		
$f' = 150$ mm, $D = 20$ mm, $S.R. = 0.996$, $\lambda = 635$ nm		
114.070	5.416	Water
−33.052	2.000	SF6
−49.325	3.280	Water
−40.500		
$f' = 200$ mm, $D = 20$ mm, $S.R. = 0.998$, $\lambda = 635$ nm		
327.190	5.069	Water
−33.052	2.000	SF6
−49.325	3.308	Water
−39.800		

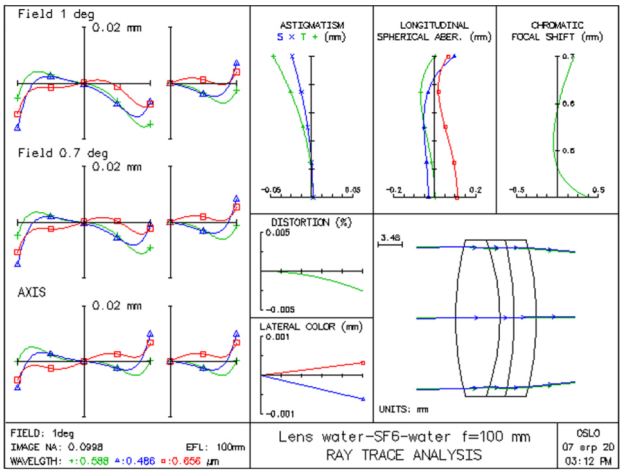


Fig. 7. Aberrations of three-component hybrid liquid-membrane lens (triplet) for $f' = 100$ mm.

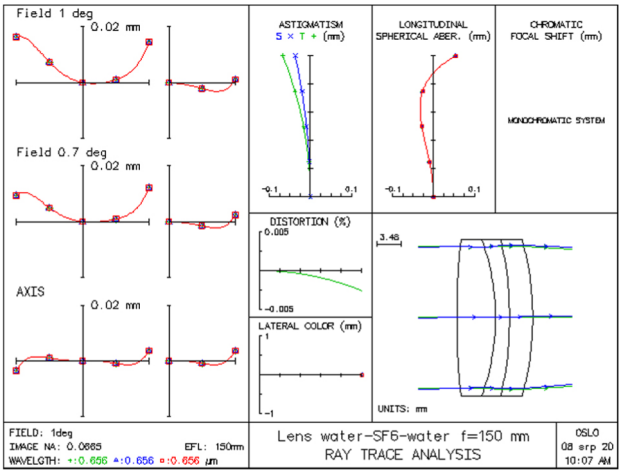


Fig. 8. Aberrations of three-component hybrid liquid-membrane lens (triplet) for $f' = 150$ mm.

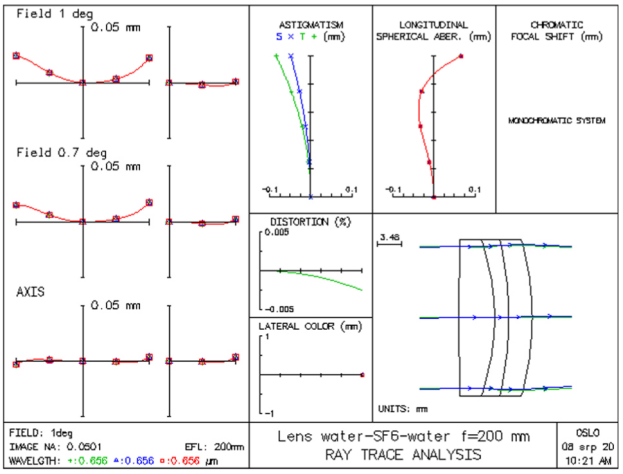


Fig. 9. Aberrations of three-component hybrid liquid-membrane lens (triplet) for $f' = 200$ mm.

is calculated for the wavelength $\lambda = 635$ nm and of imaging of the axial point. Figures 7–9 show aberrations of the hybrid triplet calculated in the software OSLO [43].

It is obvious from the aforementioned analysis that in the presented case of the hybrid liquid-membrane triplet (for the wavelength 635 nm and imaging of the axial point), one can obtain excellent imaging quality for a wide spectrum of focal length variation.

7. CONCLUSION

The paper presented formulas for calculation of parameters of hybrid doublets and triplets with variable focal length, which have the desired values of the Seidel coefficient S_{I} of spherical aberration, Seidel coefficient S_{II} of coma, and coefficient C_{I} of longitudinal chromatic aberration for one specific focal length. The procedure of calculation was demonstrated on examples of the doublet and triplet. In the case of a hybrid triplet, one can obtain excellent imaging quality for a wide spectrum of focal length variation and imaging of the axial point. The calculated values of radii of curvatures with formulas derived in the paper can be used as a very good starting point for further optimization and designing processes of the optical system in optical design software.

Funding. Czech Technical University in Prague (SGS20/093/OHK1/2T/11).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. L. Guoqiang, "Adaptive lens," in *Progress in Optics* (Elsevier, 2010), Vol. **55**, pp. 199–283.
2. H. Ren and S.-T. Wu, *Introduction to Adaptive Lenses* (Wiley, 2012).
3. http://www.optotune.com.
4. http://www.varioptic.com.
5. A. H. Rawicz and I. Mikhailenko, "Modeling a variable-focus liquid-filled optical lens," Appl. Opt. **35**, 1587–1589 (1996).
6. D.-Y. Zhang, N. Justis, V. Lien, Y. Berdichevsky, and Y.-H. Lo, "High-performance fluidic adaptive lenses," Appl. Opt. **43**, 783–787 (2004).
7. N. Sugiura and S. Morita, "Variable-focus liquid-filled optical lens," Appl. Opt. **32**, 4181–4186 (1993).
8. H. Ren, D. Fox, P. A. Anderson, B. Wu, and S.-T. Wu, "Tunable-focus liquid lens controlled using a servo motor," Opt. Express **14**, 8031–8036 (2006).
9. Q. Yang, P. Kobrin, C. Seabury, S. Narayanaswamy, and W. Christian, "Mechanical modeling of fluid-driven polymer lenses," Appl. Opt. **47**, 3658–3668 (2008).
10. G.-H. Feng and Y.-C. Chou, "Flexible meniscus/biconvex lens system with fluidic-controlled tunable-focus applications," Appl. Opt. **48**, 3284–3290 (2009).
11. H. Ren and S.-T. Wu, "Variable-focus liquid lens," Opt. Express **15**, 5931–5936 (2007).
12. D. Shaw and T. E. Sun, "Optical properties of variable-focus liquid-filled optical lenses with different membrane shapes," Opt. Eng. **46**, 024002 (2007).
13. D. Shaw and C.-W. Lin, "Design and analysis of an asymmetrical liquid-filled lens," Opt. Eng. **46**, 123002 (2007).
14. H. Choi, D. S. Han, and Y. H. Won, "Adaptive double-sided fluidic lens of polydimethylsiloxane membranes of matching thickness," Opt. Lett. **36**, 4701–4703 (2011).
15. L. Li, Q.-H. Wang, and W. Jiang, "Liquid lens with double tunable surfaces for large power tunability and improved optical performance," J. Opt. **13**, 115503 (2011).

16. F. Schneider, J. Draheim, R. Kamberger, P. Waibel, and U. Wallrabe, "Optical characterization of adaptive fluidic silicone-membrane lenses," Opt. Express **17**, 11813–11821 (2009).
17. Y.-K. Fuh, M.-X. Lin, and S. Lee, "Characterizing aberration of a pressure-actuated tunable biconvex microlens with a simple spherically-corrected design," Opt. Lasers Eng. **50**, 1677–1682 (2012).
18. A. Miks, J. Novak, and P. Novak, "Algebraic and numerical analysis of imaging properties of thin tunable-focus fluidic membrane lenses with parabolic surfaces," Appl. Opt. **52**, 2136–2144 (2013).
19. A. Mikš and J. Novák, "Three-component double conjugate zoom lens system from tunable focus lenses," Appl. Opt. **52**, 862–865 (2013).
20. A. Mikš and J. Novák, "Paraxial imaging properties of double conjugate zoom lens system composed of three tunable-focus lenses," Opt. Lasers Eng. **53**, 86–89 (2014).
21. A. Mikš and P. Novák, "Double-sided telecentric zoom lens consisting of four tunable lenses with fixed distance between object and image plane," Appl. Opt. **56**, 7020–7023 (2017).
22. L. Wang, H. Oku, and M. Ishikawa, "Development of variable-focus lens with liquid-membrane-liquid structure and 30 mm optical aperture," Proc. SPIE **8617**, 861706 (2013).
23. L. Wang, H. Oku, and M. Ishikawa, "An improved low-optical-power variable focus lens with a large aperture," Opt. Express **22**, 19448–19456 (2014).
24. S. T. Choi, B. S. Son, G. W. Seo, S.-Y. Park, and K.-S. Lee, "Opto-mechanical analysis of nonlinear elastomer membrane deformation under hydraulic pressure for variable-focus liquid-filled microlenses," Opt. Express **22**, 6133–6146 (2014).
25. D. Liang and X.-Y. Wang, "A bio-inspired optical system with a polymer membrane and integrated structure," Bioinsp. Biomim. **11**, 066008 (2016).
26. N. Hasan, A. Banerjee, H. Kim, and C. H. Mastrangelo, "Tunable-focus lens for adaptive eyeglasses," Opt. Express **25**, 1221–1233 (2017).
27. A. Miks, J. Novak, and P. Novak, "Generalized refractive tunable-focus lens and its imaging characteristics," Opt. Express **18**, 9034–9047 (2010).
28. A. Mikš and J. Novák, "Third-order aberrations of the thin refractive tunable-focus lens," Opt. Lett. **35**, 1031–1033 (2010).
29. A. Mikš and P. Novák, "Calculation of a surface shape of a pressure actuated membrane liquid lens," Opt. Lasers Eng. **58**, 60–66 (2014).
30. P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek, "Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure," Appl. Opt. **56**, 5939–5947 (2017).
31. P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek, "Deformation of a prestressed liquid lens membrane," Appl. Opt. **56**, 9368–9376 (2017).
32. A. Mikš and F. Šmejkal, "Dependence of the imaging properties of the liquid lens with variable focal length on membrane thickness," Appl. Opt. **57**, 6439–6445 (2018).
33. P. Zhao, Ç. Ataman, and H. Zappe, "Spherical aberration free liquid-filled tunable lens with variable thickness membrane," Opt. Express **23**, 21264–21278 (2015).
34. H. Yu, G. Zhou, H. M. Leung, and F. S. Chau, "Tunable liquid-filled lens integrated with aspherical surface for spherical aberration compensation," Opt. Express **18**, 9945–9954 (2010).
35. J.-W. Du, X.-Y. Wang, S. Qiang Zhu, and D. Liang, "Doublet liquid variable-focus lens for spherical aberration correction," Optik **130**, 1244–1253 (2017).
36. S. Reichelt and H. Zappe, "Design of spherically corrected, achromatic variable-focus liquid lenses," Opt. Express **15**, 14146–14154 (2007).
37. L. Li, "Zoom lens design using liquid lenses for achromatic and spherical aberration corrected target," Opt. Eng. **51**, 043001 (2012).
38. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 1999).
39. W. Welford, *Aberrations of the Symmetrical Optical Systems* (Academic, 1974).
40. "Sylgard 184 silicone elastomer kit," http://www.dowcorning.com/.
41. "Specifications of Cargille optical liquids, Cargille laboratories," https://www.cargille.com/refractive-index-liquids/.
42. https://www.zemax.com/.
43. https://www.lambdares.com/.

10838 Vol. 59, No. 34 / 1 December 2020 / Applied Optics Research Article

applied optics

Design of zoom systems composed of lenses with variable focal length

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 23 September 2020; revised 29 October 2020; accepted 30 October 2020; posted 3 November 2020 (Doc. ID 410891); published 30 November 2020

This paper develops a methodology for a design of zoom-systems, which are composed of thin optical components with a variable focal length (e.g., membrane lenses). The proposed procedure allows us to design not only the outer parameters of the system (focal lengths and separation of lenses), but the inner structure of individual components of the system can be calculated as well (radii of curvature, thicknesses, and refractive indices) i.e., the starting values of the mentioned parameters can be calculated and used for the next optimization.

https://doi.org/10.1364/AO.410891

1. INTRODUCTION

Refractive tunable-focus lenses with variable optical parameters were developed in recent years [1–6]. Such components offer a possibility to design novel optical devices with no analogy in classical systems, as they can change their focal length within a certain range. With the use of several tunable-focus lenses, the optical systems with variable optical parameters (e.g., the focal length or magnification) can be built, while the mutual positions of the inner components of the system remain without change.

Optical systems with variable optical parameters (zoom lenses) [7–29] have a wide application area in various parts of practice, for example, in microscopy or photography. Those systems have to fulfill a fundamental requirements of unchanging position of the image plane with respect to some fixed point during the change of the focal length or magnification of the zoom lens and small residual aberrations [30–36]. In contemporary zoom lenses, the change in the focal length and a fixed position of the image plane are achieved by the appropriate axial shift of individual elements of the zoom lens. However, the positions of pupils of the system have to be changed during zooming, too. In the paper [10], the authors present in detail the calculation of paraxial optical parameters in classical zoom lens systems that have fixed positions of two pairs of planes (two-conjugate zoom systems). Nevertheless, several different approaches were developed for controlling the focal length of lenses.

Tunable-focus lenses use multiple construction principles for changing their parameters, e.g., the principle of voltage-controlled liquid crystals as active optical elements, the controlled injection of fluid into chambers with deformable membranes, thermo-optical or electro-active polymers, and electro-wetting. A possibility to tune lens parameters provides another degree of freedom in the optical design process. It results in enormous advantage of the proposed solution of a zoom system using tunable-focus lenses, as the individual components can vary their parameters, and they can change the imaging properties of the system without moving their mutual position. Mechanical design of those zoom devices can be much simpler as well. Compared to the classical solutions presented in [22–29], the active focus-variable lenses provide new possibilities and many advantages.

The goal of this paper is to develop a methodology for zoom-system design, which are composed of thin optical components (e.g., membrane lenses) with a variable focal length. The proposed procedure allows us to design not only the outer parameters of the system (focal lengths and separation of lenses), but the inner structure of individual components of the system can be calculated as well (radii of curvature, thicknesses, and refractive indices)—i.e., the starting values of the mentioned parameters can be calculated and used for the next optimization.

2. ABERRATIONS OF THE OPTICAL SYSTEM

A. Seidel Aberration Coefficients of the Optical System

Aberrations are essential factors that affect the image quality of optical systems. For the purpose of the paper, consider a rotationally symmetric system of refractive lenses (see Fig. 1) consisting of K spherical surfaces [30–36]. In the case that one knows the radii of curvature of individual lenses, their thicknesses, indices of refraction, and axial distances between individual lenses, aberration coefficients of the third order (Seidel aberration coefficients) can be calculated.

Aberration coefficients of the third order (Seidel coefficients) can be expressed for the centered optical system of spherical surfaces as [30,31]

$$S_{\mathrm{I}}=\sum_{i=1}^{K}U_i,\quad S_{\mathrm{II}}=\sum_{i=1}^{K}U_iV_i,\quad S_{\mathrm{III}}=\sum_{i=1}^{K}U_iV_i^2,$$
$$S_{\mathrm{IV}}=\sum_{i=1}^{K}P_i,\quad S_{\mathrm{V}}=\sum_{i=1}^{K}\left(U_iV_i^2+H^2P_i\right)V_i, \tag{1}$$

where

$$U_i=h_i\left(\frac{\sigma_{i+1}-\sigma_i}{1/n_{i+1}-1/n_i}\right)^2\left(\frac{\sigma_{i+1}}{n_{i+1}}-\frac{\sigma_i}{n_i}\right),$$
$$V_i=\frac{\bar\sigma_{i+1}-\bar\sigma_i}{\sigma_{i+1}-\sigma_i},\quad P_i=\frac{1}{h_i}\left(\frac{\sigma_{i+1}}{n_i}-\frac{\sigma_i}{n_{i+1}}\right), \tag{2}$$

and

$$H=n_1\sigma_1y_0=n'_K\sigma'_Ky'_0=n_1\left(\bar h_1\sigma_1-h_1\bar\sigma_1\right)$$
$$=n'_K(\bar h_K\sigma'_K-h_K\bar\sigma'_K), \tag{3}$$

where σ_i and $\bar\sigma_i$ are the paraxial angles of the aperture and chief rays incident at the ith surface of the optical system, n_i is the index of refraction in front of the ith surface, and h_i is the incident height of paraxial aperture ray at ith surface. Furthermore, it holds (see Fig. 1) that $\bar\sigma_1=\bar h_1/\bar s_1=y_0/(\bar s_1-s_1)$, where y_0 is the object height. Individual aberration coefficients of the third order have the following meaning: S_{I} is the coefficient of spherical aberration, S_{II} is the coefficient of coma, S_{III} is the coefficient of astigmatism, S_{IV} is the Petzval coefficient, and S_{V} is the coefficient of distortion. It is evident from previous equations that one can use an arbitrary choice of input parameters $(h_1,\sigma_1=h_1/s_1,\bar h_1,\bar\sigma_1=\bar h_1/\bar s_1)$ for the calculation of the third-order aberration coefficients for a given object distance s_1 and a position $\bar s_1$ of the entrance pupil.

B. Seidel Aberration Coefficients for Thin Lenses

Assume a thin lens optical system in air that consists of K lenses. Seidel aberration coefficients [30–36] (Seidel sums) S_{I}, S_{II}, S_{III}, S_{IV}, and S_{V} are given by the following formulas [21]:

$$S_{\mathrm{I}}=\sum_{i=1}^{K}h_i^4M_i, \tag{4}$$

$$S_{\mathrm{II}}=\sum_{i=1}^{K}h_i^3\bar h_iM_i+\sum_{i=1}^{K}h_i^2N_i, \tag{5}$$

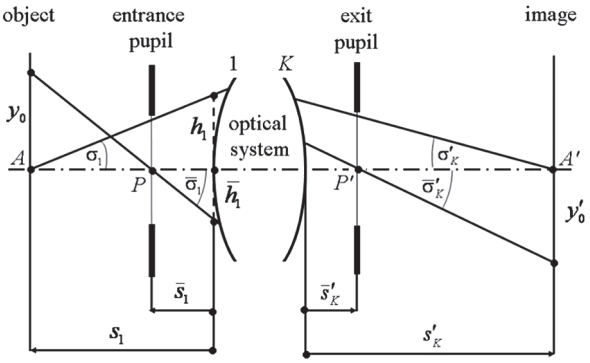


Fig. 1. General rotationally symmetric optical system composed of K spherical surfaces.

$$S_{\mathrm{III}}=\sum_{i=1}^{K}h_i^2\bar h_i^2M_i+2\sum_{i=1}^{K}h_i\bar h_iN_i+\sum_{i=1}^{K}\varphi_i, \tag{6}$$

$$S_{\mathrm{IV}}=\sum_{i=1}^{K}\frac{\varphi_i}{n_i}, \tag{7}$$

$$S_{\mathrm{V}}=\sum_{i=1}^{K}h_i\bar h_i^3M_i+3\sum_{i=1}^{K}\bar h_i^2N_i+\sum_{i=1}^{K}\frac{\bar h_i}{h_i}\left(3+\frac{1}{n_i}\right)\varphi_i, \tag{8}$$

where [21,35]

$$M_i=\varphi_i^3\left(\bar M_i+2\bar N_iY_i+1.06Y_i^2\right), \tag{9}$$

$$N_i=\varphi_i^2\left(\bar N_i+1.31Y_i\right), \tag{10}$$

$$\bar M_i=f_i'^3M_i-2f_i'^2N_iY_i+1.56Y_i^2, \tag{11}$$

$$\bar N_i=f_i'^2N_i-1.31Y_i. \tag{12}$$

The aforementioned formulas give sufficiently accurate results for all practical cases. The parameters $\bar M$ and $\bar N$ describe the spherical aberration and the coma for a system of several thin lenses in contact (cemented doublet, triplet, etc.) having unit focal length and unit magnification [35], and $\bar M_i=M_i(\varphi_i=1,m_i=-1)$, $\bar N_i=N_i(\varphi_i=1,m_i=-1)$. Parameter Y_i in Eqs. (9)–(12) has the following form:

$$Y_i=\frac{s'_i+s_i}{s'_i-s_i}=\frac{m_i+1}{m_i-1}=-1-\frac{2}{s_i\varphi_i}=1-\frac{2}{s'_i\varphi_i},$$
$$Y_{i+1}=\frac{h_i\varphi_i}{h_{i+1}\varphi_{i+1}}(Y_i-1)-1. \tag{13}$$

In previous equations, s_i and s'_i denote the object and image distances from the ith lens. Without loss of generality, one can put the Lagrange–Helmholtz invariant $H=\sigma_1\bar h_1-\bar\sigma_1h_1=1$, and $h_1=1$. Afterwards, it holds that $\bar h_1=s_1\bar s_1/(\bar s_1-s_1)$.

If one substitutes Eqs. (9) and (10) into Eqs. (4)–(8), then it holds for Seidel aberration coefficients for zoom lenses [21] that

$$S_{\mathrm{I}j}=\sum_{i=1}^{K}h_{ji}^4\varphi_i^3\bar M_i+2\sum_{i=1}^{K}h_{ji}^4\varphi_i^3Y_{ji}\bar N_i+1.06\sum_{i=1}^{K}h_{ji}^4\varphi_i^3Y_{ji}^2,$$
$$S_{\mathrm{II}j}=\sum_{i=1}^{K}h_{ji}^3\bar h_{ji}\varphi_i^3\bar M_i+\sum_{i=1}^{K}h_{ji}^2\varphi_i^2(2h_{ji}\bar h_{ji}\varphi_iY_{ji}+1)\bar N_i$$
$$+\sum_{i=1}^{K}h_{ji}^2\varphi_i^2Y_{ji}(1.06h_{ji}\bar h_{ji}\varphi_iY_{ji}+1.31),$$
$$S_{\mathrm{III}j}=\sum_{i=1}^{K}h_{ji}^2\bar h_{ji}^2\varphi_i^3\bar M_i+2\sum_{i=1}^{K}h_{ji}\bar h_{ji}\varphi_i^2(h_{ji}\bar h_{ji}\varphi_iY_{ji}+1)\bar N_i$$
$$+\sum_{i=1}^{K}h_{ji}\bar h_{ji}\varphi_i^2Y_{ji}(1.06h_{ji}\bar h_{ji}\varphi_iY_{ji}+2.62)+\sum_{i=1}^{K}\varphi_i,$$
$$S_{\mathrm{V}j}=\sum_{i=1}^{K}h_{ji}\bar h_{ji}^3\varphi_i^3\bar M_i+\sum_{i=1}^{K}\bar h_{ji}^2\varphi_i^2(2h_{ji}\bar h_{ji}\varphi_iY_{ji}+3)\bar N_i$$
$$+\sum_{i=1}^{K}\bar h_{ji}^2\varphi_i^2Y_{ji}(1.06h_{ji}\bar h_{ji}\varphi_iY_{ji}+3.93)+3.62\sum_{i=1}^{K}\frac{\bar h_{ji}}{h_{ji}}\varphi_i, \tag{14}$$

where $i = 1, 2, \ldots, K$, K is the number of zoom lens members, $j = 1, 2, \ldots, L$, and L is the number of configurations of the zoom lens system (e.g., values of focal lengths or magnifications, which are considered for correction of the optical system; note that for classic optical systems the condition $L \gg K$ should be fulfilled). By solving the system of Eq. (14) (least-squares method) one obtains the values of variables $\bar{M}$ and $\bar{N}$ for individual members of the optical system. In the case that individual components of the zoom system are composed of lenses with variable focal length, one substitutes in Eq. (14) as follows: $\varphi_i \to \varphi_{ji}$, where φ_{ji} is the optical power of the ith component in the jth configuration (state) of the zoom system.

3. DESIGN OF ZOOM SYSTEMS WITH COMPONENTS WITH VARIABLE FOCAL LENGTH

A. Zoom Systems with Variable Focal Length

In the case of classic zoom systems, individual components of the optical system move along the optical axis while changing the system's focal length or magnification, and the inner structures of individual parts (radii of curvature, vertex distances, and refractive indices) remain unchanged. The fundamental difference of those zoom systems having variable focal length is that the components remain in the same positions and the inner structure (parameters) is modified. In classic optical systems, the aberrations cannot be zero in the whole range of focal distances; they can be only minimized. However, in the case of zoom systems with components with variable focal length, one can obtain a situation in which Seidel coefficients of the zoom system (the aberrations) will fulfill one of the following:

- The system will have desired values of aberrations in the whole range of focal distances of the zoom system.
- The aberrations will change minimally in the whole range of focal distances of the zoom system.

During designing the zoom system with components of variable focal length, one follows the scheme that, first, the outer parameters of the system are given, i.e., focal lengths and mutual distances of individual components of the zoom system. In the previous papers of the authors [7–13], the mentioned problem is studied in detail, and one can use the presented methods published in this works. The second step of the design is to calculate so-called inner parameters (inner structure) of individual components of the zoom system.

Suppose now that the inner parameters of individual components of the zoom system are designed, while the components have variable focal length, and they are made by thin membrane lenses. Let S_{I}, S_{II}, and C_{I} denote the desired values of the Seidel coefficients of the spherical aberration, coma, and coefficient of longitudinal chromatic aberration of one of the components of the zoom system. Further, let the individual membrane components be made by three (or more) membrane lenses or a combination of membrane and classic glass lenses (hybrid components). Systems of three or more components are chosen, as only a limited number of optical liquids is available for membrane lenses, and three-component systems offer a possibility to reach the desired values of Seidel coefficients S_{I}, S_{II}, and C_{I}. For construction simplicity, the hybrid components made by one (middle) glass lens and two membranes in front of and behind the glass lens are advantageous (see Fig. 2).

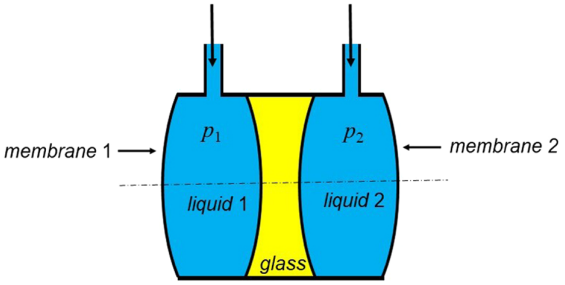


Fig. 2. Scheme of three-component hybrid lens.

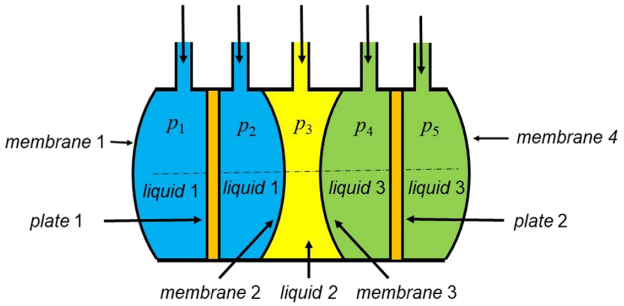


Fig. 3. Scheme of general three-component hybrid lens.

Suppose now the situation shown in Fig. 3. Individual components of the zoom system are made by three membrane lenses. The first and the third lenses (presented with blue and green color in Fig. 3) are both made by two plano-convex lenses separated by a thin plate of a transparent material (glass, Sylgard, etc.). The second lens is made by membrane 2 and membrane 3, and it is used as a compensating reservoir. The radii of the curvatures of membranes 1, 2, 3, and 4 can be in certain ranges modified by applying pressures p_1, p_2, p_4, and p_5.

In a situation when zoom systems will be made by three-component hybrid components with middle glass lens of fixed dimensions (Fig. 2), the aberrations of the system cannot be corrected for all states of the systems (focal lengths or magnifications). These systems have only two selected parameters: the first and the last radius of curvature. Aberrations can be corrected for several specific cases only. In other states, aberrations will be changing, and one has to fulfill a fixed distance of the image plane from the last component of the zoom system for all values of the focal length (for an object at infinity) or transverse magnification (for an object at finite distance). To compensate for aberrations for more states of the zoom system, its individual components have to be made by systems with more components as shown, for example, in Fig. 3.

B. Calculation of Inner Parameters of Zoom Systems with Components with Variable Focal Length

Suppose that number of components of the zoom system and their outer parameters (focal lengths and axial distances between individual components) are known. Further, let those components be made by several thin lenses (e.g., triplets). Calculation

of the inner parameters of individual thin components of the zoom system can be realized as follows:

1. One chooses a number L of states (configurations), e.g., number of focal lengths of the zoom system, for which the system has corrected or minimized aberrations.
2. The system of Eq. (14) for selected Seidel coefficients is solved.
3. By solving the system of equations Eq. (14), one obtains the values of variables $\bar{M}_i$ and $\bar{N}_i$, where $i = 1, 2, .., K$, (K is the number of zoom lens members) for an individual member of the optical system.
4. One sets $S_{\mathrm{I}} = \bar{M}_i$ and $S_{\mathrm{II}} = \bar{N}_i$.

Let the individual components of the zoom system be triplets. Afterwards, one should calculate the inner parameters of the individual component with the following three steps:

1. The optical power φ_1 of the first lens of the membrane component is set, and the overall optical power of the membrane component is set to $\varphi = 1$. Further, the materials of the individual lenses of the membrane component are selected, i.e., refractive indices and Abbe numbers. The optical power φ_2 of the second lens and φ_3 of the third lens can then be calculated as follows:

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3, \quad C_{\mathrm{I}} = \varphi_1/\nu_2 + \varphi_2/\nu_3 + \varphi_3/\nu_4. \tag{15}$$

Further, one gets the following formulas for optical powers φ_2 and φ_3:

$$\varphi_2 = \frac{\nu_3(\nu_2\varphi - \nu_2\varphi_1 + \nu_4\varphi_1 - C_{\mathrm{I}}\nu_2\nu_4)}{\nu_2\,(\nu_3 - \nu_4)},$$

$$\varphi_3 = -\frac{\nu_4(\nu_2\varphi - \nu_2\varphi_1 + \nu_3\varphi_1 - C_{\mathrm{I}}\nu_2\nu_3)}{\nu_2\,(\nu_3 - \nu_4)}, \tag{16}$$

where ν_2, ν_3, and ν_4 are Abbe numbers of lenses, and C_I is the coefficient of longitudinal chromatic aberrations of the membrane component [30,31]. With the use of the paraxial imaging equation [30,31]

$$n_{i+1}u_{i+1} - n_{i+1}u_{i+1} = h_i(n_{i+1} - n_i)/r_i, \quad i = 1, 2, 3, 4, \tag{17}$$

one gets for the paraxial aperture angles u_i the following formulas:

$$u_2 = \varphi_1 - u_1 + \frac{2u_1}{n_2} - Q\left(\frac{1}{n_2} - 1\right),$$

$$u_3 = \varphi_1 - u_1 + \frac{2u_1}{n_3} - Q\left(\frac{1}{n_3} - 1\right),$$

$$u_4 = \varphi_1 - u_1 + \frac{2u_1 + \varphi_2}{n_4} - Q\left(\frac{1}{n_4} - 1\right) - \frac{\varphi_2\,(n_4 - 1)}{n_4\,(n_3 - 1)},$$

$$u_5 = u_1 + 1, \tag{18}$$

where $u_1 = \sigma_1$ is the paraxial aperture angle in the object space of the optical system; n_2, n_3, and n_4 are the indices of refraction of the components' (lenses') materials; r_i is the radius of curvature of the ith optical surface; Q is Abbe invariant [30,31] on the second surface of the system; and h_i is the paraxial impinging height. In the presented case of the optical system with thin lenses, one can choose $h_i = 1$ ($i = 1, 2, 3, 4$). Let the object and image space of the optical system be air ($n_1 = n_5 = 1$). For the object at infinity, it holds that $u_1 = 0$. Substituting Eq. (18) in Eq. (1), one gets the following formula for the Seidel coefficient of spherical aberration S_{I} of the thin hybrid lens:

$$S_{\mathrm{I}} = a_2 Q^2 + a_1 Q + a_0, \tag{19}$$

where

$$a_2 = \frac{2\varphi_2}{n_3} - 2\frac{\varphi_1 + \varphi_2 - 1}{n_4} + \frac{n_2 + 2\varphi_1}{n_2},$$

$$a_1 = A_{10} + A_{11}u_1,$$

$$A_{11} = 8\frac{\varphi_1 + \varphi_2 - 1}{n_4} - 2\left(\frac{3n_2 + 4\varphi_1}{n_2}\right) - 8\frac{\varphi_2}{n_3},$$

$$A_{10} = \frac{3{\varphi_1}^2}{n_2 - 1} - \frac{{\alpha_1}^2}{n_4 - 1} - \frac{2n_3\varphi_2}{n_3 - 1}\left[\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2}\left(1 - \frac{(n_4 - 1)}{(n_3 - 1)}\right) - \frac{\varphi_1}{n_3}\right] - \frac{2n_4\alpha_1}{n_4 - 1}\left(\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2} - \alpha_2 - 1\right) - \frac{{\varphi_2}^2}{{n_4}^2(n_3 - 1)^2}\left[{n_4}^2\,(n_3 - 1) - {n_3}^2\,(n_4 - 1)\right],$$

$$\alpha_1 = \varphi_1 + \frac{\varphi_2}{n_4} - \alpha_2 - 1, \quad \alpha_2 = \frac{\varphi_2\,(n_4 - 1)}{n_4\,(n_3 - 1)},$$

$$a_0 = A_{02}{u_1}^2 + A_{01}u_1 + A_{00},$$

$$A_{02} = 8\frac{n_2 + \varphi_1}{n_2} - 8\frac{\varphi_1 + \varphi_2 - 1}{n_4} + 8\frac{\varphi_2}{n_3},$$

$$A_{01} = \frac{5n_3 + 6\varphi_1 + 6\varphi_2 - 6n_3\varphi_1 - 6\varphi_1\varphi_2 + {\varphi_2}^2 - 5}{n_3 - 1} - \frac{(\varphi_1 + \varphi_2 - 1)}{n_4\,(n_3 - 1)\,(n_4 - 1)}\left[7n_4\,(n_3 - 1 + \varphi_1 + \varphi_2 - n_3\varphi_1) + \varphi_2 n_3\,(n_4 - 8)\right] - \frac{7{\varphi_1}^2}{n_2 - 1},$$

$$A_{00} = \frac{n_2{\varphi_1}^3}{(n_2 - 1)^2} - \frac{{n_4}^2}{(n_4 - 1)^2}\left[\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2} - \frac{\varphi_2\,(n_4 - 1)}{{n_4}^2\,(n_3 - 1)} - 1\right] \times \left[\varphi_1 + \frac{\varphi_2}{n_4} - \frac{\varphi_2\,(n_4 - 1)}{n_4\,(n_3 - 1)} - 1\right]^2 + \frac{n_3{\varphi_2}^2\,(n_3 - n_4)}{{n_4}^2(n_3 - 1)^3}\left[n_3\varphi_2 + n_4\varphi_1\,(n_3 - 1)\right].$$

In the case of coma, one gets for the Seidel coefficient S_{II} of coma for the thin three-component lens, after substitution of Eq. (18) in Eq. (1), the following ($\bar{s}_1 = 0$):

$$S_{II} = b_0 + Qb_1, \tag{20}$$

where

$$b_1 = \frac{\varphi_1 + \varphi_2 - 1}{n_4} - \frac{n_2 + \varphi_1}{n_2} - \frac{\varphi_2}{n_3},$$

$$b_0 = B_{00} + B_{01} u_1,$$

$$B_{01} = 3 + 2\left(\frac{\varphi_1}{n_2} + \frac{\varphi_2}{n_3}\right) - 2\frac{\varphi_1 + \varphi_2 - 1}{n_4},$$

$$B_{00} = \frac{n_4}{n_4 - 1}\left[\frac{\varphi_1}{n_4} + \frac{\varphi_2}{{n_4}^2} - \frac{\varphi_2\,(n_4 - 1)}{{n_4}^2\,(n_3 - 1)} - 1\right] \times \left[\varphi_1 + \frac{\varphi_2}{n_4} - \frac{\varphi_2\,(n_4 - 1)}{n_4\,(n_3 - 1)} - 1\right] - \frac{{\varphi_1}^2}{n_2 - 1} + \frac{\varphi_2\,(n_3 - n_4)}{{n_4}^2 (n_3 - 1)^2}\left[n_3\varphi_2 + n_4\varphi_1\,(n_3 - 1)\right].$$

2. Setting the Seidel coefficient of spherical aberration S_{I} equal to a desired value, then one can calculate the parameter Q with Eq. (19). Afterwards, the radii of curvature of individual surfaces of thin three-component lens can be enumerated from the following formulas:

$$\frac{1}{r_1} = Q - u_1 + \frac{n_2\varphi_1}{n_2 - 1}, \quad \frac{1}{r_2} = Q - u_1 + \varphi_1,$$
$$\frac{1}{r_3} = Q - \frac{\varphi_2 + (u_1 - \varphi_1)\,(n_3 - 1)}{n_3 - 1},$$
$$\frac{1}{r_4} = Q - \frac{\varphi_3}{n_4 - 1} - \frac{\varphi_2 + (u_1 - \varphi_1)\,(n_3 - 1)}{n_3 - 1}. \tag{21}$$

3. To complete the calculation, one should multiply the values calculated from Eq. (21) by the desired value of the focal length of the appropriate component.

The process of the aforementioned steps is then repeated for each member of the zoom system, i.e., for $i = 1, 2, \ldots, K$, where K denotes the number of zoom lens members for an individual member of the optical system.

Suppose now the situation of the three-component hybrid lens (Fig. 2), where X_2 denotes the shape parameter, and the optical power φ_2 is fixed while changing the focal length of the lens; the variable parameters of the lens are the radii of curvature of the first and the second membrane. Further, let the optical power of the hybrid lens be $\varphi = \varphi_1 + \varphi_2 + \varphi_3$. Afterwards, one can derive the following formula for the Seidel coefficient of spherical aberration S_{I} with Eq. (1) for the object at infinity:

$$S_{\mathrm{I}} = e_3{\varphi_1}^3 + e_2{\varphi_1}^2 + e_1\varphi_1 + e_0, \tag{22}$$

where

$$e_3 = A_1 - A_3 - B_1 + B_3 + C_1 - C_3 + D_1 - D_3,$$

$$e_2 = -\varphi_2(3A_3 - 3B_3 - 4C_2 + 3C_3 + 3D_3) + (3A_3 - B_3 - C_3 + 3D_3)\,\varphi + A(2A_1 - AB_1) + B(B_3 - 2A_3),$$

$$e_1 = -{\varphi_2}^2\left[3(A_3 - B_3 + C_3 + D_3) - 4C_2 + 2B_2X_2\right] + \varphi_2[2(3A_3 - B_3 - C_3 + 3D_3)\,\varphi + 2B(B_3 - 2A_3)] + (C_3 - B_3 - 3A_3 - 3D_3)\,\varphi^2 + (4A_3B)\,\varphi + A^2A_1 - A_3B^2,$$

$$e_0 = -{\varphi_2}^3\left(-A_2{X_2}^2 + B_2X_2 + A_3 - B_3 - C_2 + C_3 - D_2 + D_3\right) + {\varphi_2}^2[(3A_3 - B_3 - C_3 + 3D_3)\,\varphi + (B_3 - 2A_3)\,B] - \varphi_2\left[(3A_3 + B_3 - C_3 + 3D_3)\,\varphi^2 + (B - 4\varphi)\,A_3B\right] + (A_3 + B_3 + C_3 + D_3)\,\varphi^3 - (2A_3 + B_3)\,B\varphi^2 + \left(A_3B^2\right)\varphi,$$

$$A = \frac{\varphi_2(X_2 + 1)\,(n_1 - 1)}{n_2 - 1}, \quad B = \frac{\varphi_2(X_2 - 1)\,(n_3 - 1)}{n_2 - 1}.$$

Similarly, one gets for the Seidel coefficient of coma S_{II} the following:

$$S_{\mathrm{II}} = g_2{\varphi_1}^2 + g_1\varphi_1 + g_0, \tag{23}$$

where

$$g_2 = E_1 - E_3 - F_1 + F_3,$$

$$g_1 = 2E_3\varphi - \varphi_2(2E_3 + 2F_2 - 2F_3) + AE_1 - BE_3,$$

$$g_0 = -{\varphi_2}^2(E_3 + F_2 - F_3 - E_2X_2) + \varphi_2(2E_3\varphi - BE_3) - (E_3 + F_3)\,\varphi^2 + BE_3\varphi,$$

where $(i = 1, 2, 3)$ [21,35],

$$A_i = \frac{n_i + 2}{4n_i(n_i - 1)^2}, \quad B_i = \frac{n_i + 1}{n_i\,(n_i - 1)}, \quad C_i = \frac{3n_i + 2}{4n_i},$$
$$D_i = \frac{{n_i}^2}{4(n_i - 1)^2}, \quad E_i = B_i/2, \quad G_i = \frac{2n_i + 1}{2n_i},$$
$$\varphi_i = (n_i - 1)\left(\frac{1}{r_i} - \frac{1}{r'_i}\right), \quad X_i = \frac{r'_i + r_i}{r'_i - r_i}, \tag{24}$$

where r_i, r'_i, n_i, and X_i denote radii of curvature, refractive indices of the materials (glass, liquid), and the shape parameters of ith lens. Further, it holds that $r_2 = r'_1, r_3 = r'_2$.

One can calculate the optical power φ_1 of the first lens from Eq. (22) for the Seidel coefficient of spherical aberration S_{I}. Afterwards, the optical power φ_3 of the third lens and shape parameters X_1 and X_3 of the first and the third lenses can be calculated from the following formulas:

$$\varphi_3 = \varphi - \varphi_1 - \varphi_2, \quad X_1 = \frac{A}{\varphi_1} + 1, \quad X_3 = \frac{B}{\varphi_3} - 1. \tag{25}$$

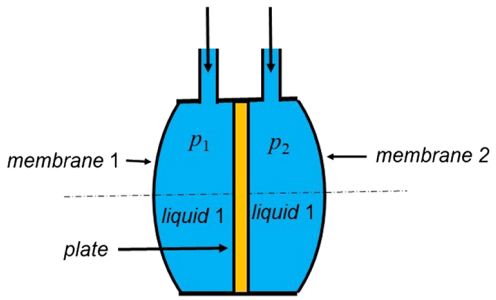


Fig. 4. General membrane lens with plane-parallel plate.

Finally, one can calculate the radius of curvature of the ith lens, in the case that the lens power φ_i, refractive index n_i, and the shape parameters X_i are known. It holds that

$$r_i = \frac{2(n_i - 1)}{\varphi_i(X_i + 1)}, \quad r_i' = \frac{2(n_i - 1)}{\varphi_i(X_i - 1)}. \tag{26}$$

To recalculate Seidel coefficients S_{I}, S_{II} for a different position of the object and the entrance pupil, and for a different value of the focal length, one can use formulas presented, for example, in papers [31,32,36].

C. Effect of Plane-parallel Plate on Imaging Properties of Membrane Lens

Consider now the situation of a general membrane lens with a plane-parallel plate placed between two plano-convex lenses (see Fig. 4) and study the influence of the plate on the imaging properties of the membrane lens used in the zoom system in Fig. 3. For simplicity, consider the spherical aberration only. For the Seidel coefficient S_{I} of spherical aberration of the aforementioned membrane lens, one gets with Eqs. (1) and (2)

$$\begin{aligned} S_{\mathrm{I}} = {} & \frac{d\varphi_1^{\,4}}{n_3}\left(\frac{1}{n_3^{\,2}} - \frac{1}{n_2^{\,2}}\right) + \frac{\varphi_1^{\,3}}{n_2^{\,2}(n_2 - 1)^2} \\ & - \frac{(n_3 - d\varphi_1)\,(n_3\varphi_1 - n_2 n_3\varphi_1 + n_2 n_3\varphi_4 - d n_2\varphi_1\varphi_4)^2}{n_2^{\,2} n_3^{\,4}(n_2 - 1)^2} \\ & \times \left(n_3\varphi_1 - n_2^{\,2} n_3\varphi_1 + n_2^{\,2} n_3\varphi_4 - d n_2^{\,2}\varphi_1\varphi_4\right), \end{aligned} \tag{27}$$

where

$$\varphi_1 = \frac{n_2 - 1}{r_1}, \qquad \varphi_4 = -\frac{n_2 - 1}{r_4} \tag{28}$$

are optical powers of the first and the second (last) surface of the lens.

For the Seidel coefficient of spherical aberration of a thin membrane lens, one can write

$$\begin{aligned} S_{I0} = {} & (\varphi_1 - \varphi_4)^3 - \frac{2(\varphi_1^{\,3} - 2\varphi_1\varphi_4^{\,2} + \varphi_4^{\,3})}{n_2 - 1} \\ & + \frac{\varphi_1^{\,3} - \varphi_4^{\,3}}{(n_2 - 1)^2} + \frac{2\varphi_1^{\,2}\,(\varphi_1 - \varphi_4)}{n_2}. \end{aligned} \tag{29}$$

Further, the difference between the Seidel coefficient of the spherical aberration of the thin membrane lens and the membrane lens having the transparent plate of thickness d in between the membranes then can be expressed with the following formula:

$$\begin{aligned} \Delta S_{\mathrm{I}} = {} & \frac{\left(\varphi_1 - n_2^{\,2}\varphi_1 + n_2^{\,2}\varphi_4\right)(\varphi_1 - n_2\varphi_1 + n_2\varphi_4)^2}{n_2^{\,2}(n_2 - 1)^2} \\ & + \frac{d\varphi_1^{\,4}\left(n_2^{\,2} - n_3^{\,2}\right)}{n_2^{\,2} n_3^{\,3}} \\ & - \frac{(n_3 - d\varphi_1)\,(n_3\varphi_1 - n_2 n_3\varphi_1 + n_2 n_3\varphi_4 - d n_2\varphi_1\varphi_4)^2}{n_2^{\,2} n_3^{\,4}(n_2 - 1)^2} \\ & \times \left(n_3\varphi_1 - n_2^{\,2} n_3\varphi_1 + n_2^{\,2} n_3\varphi_4 - d n_2^{\,2}\varphi_1\varphi_4\right). \end{aligned} \tag{30}$$

The Seidel coefficient of the spherical aberration of the plane-parallel plate of thickness d made by material of the refractive index n, surrounded by a space with the refractive index n_0, can be calculated from the formula

$$S_{\mathrm{I},plan} = -\frac{d n_0^{\,2} u_1^{\,4}\left(n^2 - n_0^{\,2}\right)}{n^3}, \tag{31}$$

where u_1 is the aperture angle of the ray impinging on the plate. As one can see from the aforementioned formula, if the material of the plate has the refractive index n, which is similar to the refractive index n_0 of the surrounding space, then one can practically neglect the effect of the plane-parallel plate in the membrane lens.

Table 1 presents the error of the Seidel coefficient S_{I} of the spherical aberration of the lens influenced by the plane-parallel plate of thickness $d_2 = 1$ mm ($d_1 = d_3 = 0$) made by glass Schott BK7 ($n = 1.5168$). The calculation is done for the object at infinity. The liquid in between membranes and the plate is water ($n = 1.333$). S_{II} denotes the Seidel coefficient of coma for the case in which the entrance pupil is at the vertex of the first surface of the lens. The values of $S_{\mathrm{I},0}$ and $S_{\mathrm{II},0}$ present Seidel coefficients for the cases in which the plate has zero thickness, r_1 and r_2 are the radii of curvature of the first and the last surface of the lens ($r_2 = r_3 = \infty$), and f' is the focal length of the lens.

Table 1. Influence of a Plane-parallel Plate on Imaging with a Membrane Lens (Length Values in Millimeters)

S_{I}	S_{II}	$S_{\mathrm{I},0}$	$S_{\mathrm{II},0}$	r_1	r_4	d_2	f'	Error [%]
1.8357e−04	−1.4925e−03	1.8951e−04	−1.5252e−03	20.00	−20.00	1.00	3.0027e+01	−3.2333e+00
2.3315e−05	−3.7720e−04	2.3689e−05	−3.8129e−04	40.000	−40.00	1.00	6.0053e+01	−1.6019e+00
6.9449e−06	−1.6825e−04	7.0189e−06	−1.6946e−04	60.000	−60.00	1.00	9.0080e+01	−1.0647e+00
2.9377e−06	−9.4811e−05	2.9611e−06	−9.5322e−05	80.000	−80.00	1.00	1.2011e+02	−7.9727e−01
1.5065e−06	−6.0744e−05	1.5161e−06	−6.1006e−05	100.00	−100.00	1.00	1.5013e+02	−6.3723e−01

Table 2. Outer Parameters of the Optical System (Length Values in Millimeters)

f'	f_1'	f_2'	f_3'	s_p	s_p'	h_1	h_2	h_3	$\bar{h}_1$	$\bar{h}_2$	$\bar{h}_3$
55.0000	730.3056	−34.8367	36.5817	20.5631	−78.9646	4.9107	4.7762	8.0357	−0.9347	0.0000	1.1684
82.5000	75.4234	−26.7195	41.3782	27.2172	−63.1605	7.3661	5.4128	8.0357	−0.8248	0.0000	1.0310
110.0000	57.9037	−27.4887	52.3328	30.5531	−47.8663	9.8214	6.4291	8.0357	−0.6944	0.0000	0.8680
137.5000	53.7200	−31.8809	78.4208	31.8624	−36.6996	12.2768	7.7061	8.0357	−0.5793	0.0000	0.7241
165.0000	55.6694	−45.2755	242.4949	31.2141	−27.8736	14.7321	9.4394	8.0357	−0.4729	0.0000	0.5912

Length values are in millimeters in Table 1. One can see that the effect of the plate is minimal, and one does not have to consider its influence for the starting design of the optical system.

4. EXAMPLE

Suppose now the three-component zoom system with members of variable focal length, which are made be three-lens components (Fig. 3). Suppose next that the object is at infinity ($s_1 = -\infty$). During the calculation, the first step is to set $\varphi = 1/f'$, d_1, d_2, and s_F', where f' is the focal length of the optical system, d_1 and d_2 are axial distances between individual components of the system, and s_F' is the image distance. Further, let one demand the Petzval sum $S_{\mathrm{IV}} \approx 0$. Afterwards, optical powers φ_1, φ_2, and φ_3 can be calculated from formulas for the optical power φ of the optical system, for the position of the image focal point s_F', and for the Petzval coefficient ($S_{\mathrm{IV}} \approx \varphi_1 + \varphi_2 + \varphi_3 = 0$), from which the following system of equations can be derived, and one can easily calculate the desired values:

$$\begin{aligned} \left(d_1 + d_2 s'_F \varphi\right)\varphi_1 &= \frac{d_2\varphi}{2}\left(\sqrt{\psi} - 1\right) - s'_F\varphi + 1, \\ 2s'_F\varphi_2 &= -\sqrt{\psi} - 1, \\ \varphi_3 &= -\varphi_1 - \varphi_2, \end{aligned} \tag{32}$$

where

$$\psi = \left(-\frac{4\varphi}{d_1 d_2}\right) s'^{3}_F + \frac{8}{d_1 d_2} s'^{2}_F + 4\frac{(d_1 + d_2)\varphi - 1}{d_1 d_2 \varphi} s'_F + 1.$$

With the aforementioned Eq. (32), one can simply recalculate focal lengths f_1', f_2', and f_3' of the individual members of the system for different values of the focal length f', and to construct zoom diagrams, for example.

Let the following values be selected as an example: $d_1 = 20$ mm, $d_2 = 25$ mm, and $s_F' = 90$ mm. Further, let the aperture stop be in the position of the second member. With Eq. (32) and paraxial imaging formulas, one can calculate the parameters presented in Table 2, i.e., focal lengths f_1', f_2', and f_3' of individual components; the impinging height h of the aperture ray; and the impinging height $\bar{h}$ of the principal ray. The value of s_p is the position of the entrance pupil, and s_p' is the position of the exit pupil. Length units are millimeters in Table 2. The focal length of the zoom system is selected in the interval $f' \in [55 - 165]$ mm, the f-number of the zoom system is $F\# = 5.6$ for $f' = 55$ mm, and the image height is 5 mm for all values of the focal length. Further, it holds [35] that

Table 3. Coefficients $\bar{M}$, $\bar{N}$, $\bar{M}_0$, and $\bar{N}_0$ of the Optical System for Different Boundary Values of the Interval for the Focal Length f'

Member	$\bar{M}$	$\bar{N}$	$\bar{M}_0$	$\bar{N}_0$
	$f'_{\min} = 55$mm, $f'_{\max} = 165$mm			
1	4.5753	1.7281	2.1792	0.4181
2	6.4526	−0.6185	8.7496	−1.9285
3	1.5348	−0.9450	4.4848	−2.2550
	$f'_{\min} = 82.5$mm, $f'_{\max} = 137.5$mm			
1	2.8988	1.2723	1.4141	−0.0377
2	1.7536	−0.2572	3.3280	−1.5672
3	1.2975	−0.7042	3.7659	−2.0142

Table 4. Starting Parameters of the Optical Zoom System (Length Values in Millimeters)

	$n_1 = 1.3330$, $n_2 = 1.8052$, $\nu_1 = 57.07$, $\nu_2 = 25.43$, $n_3 = n_1$, $\nu_3 = \nu_1$				
	$f' = 55$	$f' = 82.5$	$f' = 137.5$	$f' = 165$	
	$r/730.3056$	$r/74.4234$	$r/53.7200$	$r/55.6694$	d
1	0.4181	0.4510	0.4510	0.4181	
2	−0.4188	−0.3612	−0.3612	−0.4188	
3	−0.7195	−0.5649	−0.5649	−0.7195	
4	−0.5190	−0.4545	−0.4545	−0.5190	
	$-r/34.8367$	$-r/26.7195$	$-r/31.8809$	$-r/45.2755$	20
5	0.6656	0.7102	0.7102	0.6656	
6	−0.3808	−0.2777	−0.2777	−0.3808	
7	−0.6142	−0.3841	−0.3841	−0.6142	
8	−0.3430	−0.3323	−0.3323	−0.3430	
	$r/36.5817$	$r/41.3782$	$r/78.4208$	$r/242.4949$	25
9	1.0007	0.8843	0.8843	1.0007	
10	−0.2404	−0.2492	−0.2492	−0.2404	
11	−0.3163	−0.3318	−0.3318	−0.3163	
12	−0.2925	−0.3042	−0.3042	−0.2925	

$$\begin{aligned} \bar{M}_{0i} &= M_i(\varphi_i = 1,\ m_i = 0) = \bar{M}_i - 2\bar{N}_i + 1.06, \\ \bar{N}_{0i} &= N_i(\varphi_i = 1,\ m_i = 0) = \bar{N}_i - 1.31. \end{aligned} \tag{33}$$

After the solution of Eq. (14), for $S_{\mathrm{I}} = S_{\mathrm{II}} = S_{\mathrm{III}} = 0$, $K = 3$ and $L = 2$, and then one gets the results presented in Table 3. Afterwards, Table 4 presents the starting parameters of the zoom system for focal lengths $f' = 55$ mm, 82.5 mm, 137.5 mm, and 165 mm.

5. CONCLUSION

The paper presented a design process of a zoom system composed by members with variable focal length. The formulas for calculation of the aberration properties of individual components of the zoom system are presented, while the third-order aberration theory and thin lenses are considered. In detail, there is a solution of a calculation of the inner structure of individual components presented, especially for the case of three-lens components (triplets). Those components (Fig. 3) can be of the properties that Seidel coefficients S_{I}, S_{II}, and C_{I} can have desired values for any configuration (state) of the zoom system. Individual components are then made by liquid membrane lenses. It was shown that in the case of hybrid triplets, where one of the lenses is a classic glass lens (Fig. 2), one can obtain the desired values of the aforementioned Seidel coefficients for one configuration of the zoom system only (e.g., for one value of the focal length). For other configurations, the condition cannot be reached as one has two free parameters only (the radii of outer membranes' surfaces). The zoom system made by those hybrid components (Fig. 2) will have worse imaging quality than the one composed by liquid membrane lenses shown in Fig. 3.

It is shown in the paper that one cannot reach the same imaging quality of zoom systems composed by members of variable focal length compared to classic zoom systems of glass lenses. The next disadvantage of the membrane liquid components is that those have only a small diameter, and the mechanism maintaining the change of the imaging properties might be quite complex in comparison to simple axial movement of individual classic glass lenses.

To conclude, the zoom systems made by liquid membrane lenses (or hybrid components) only cannot compete with classic zoom systems of glass lenses from the point of imaging quality and field of view. However, those novel systems can be appropriate in many fields, e.g. scanning systems or individual components of the zoom systems. Further development of those components promises wider application areas, especially as they will improve the imaging qualities, and they can simplify the zoom systems' construction.

Funding. Czech Technical University in Prague (České Vysoké Učení Technické v Praze) (SGS20/093/OHK1/2T/11).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. L. Guoqiang, "Adaptive lens," in *Progress in Optics* (Elsevier, 2010).
2. H. Ren and S.-T. Wu, *Introduction to Adaptive Lenses* (Wiley, 2012).
3. http://www.varioptic.com.
4. http://www.optotune.com.
5. http://www.okotech.com.
6. http://www.tag-optics.com.
7. A. Miks, J. Novak, and P. Novak, "Generalized refractive tunable-focus lens and its imaging characteristics," Opt. Express **18**, 9034–9047 (2010).
8. A. Miks and J. Novak, "Analysis of two-element zoom systems based on variable power lenses," Opt. Express **18**, 6797–6810 (2010).
9. A. Miks and J. Novak, "Analysis of three-element zoom lens based on refractive variable-focus lenses," Opt. Express **19**, 23989–23996 (2011).
10. A. Mikš and J. Novák, "Three-component double conjugate zoom lens system from tunable focus lenses," Appl. Opt. **52**, 862–865 (2013).
11. A. Miks and J. Novak, "Paraxial analysis of zoom lens composed of three tunable-focus elements with fixed position of image-space focal point and object-image distance," Opt. Express **22**, 27056–27062 (2014).
12. A. Mikš and P. Novák, "Paraxial design of four-component zoom lens with fixed position of optical center composed of members with variable focal length," Opt. Express **26**, 25611–25616 (2018).
13. A. Mikš and J. Novák, "Third-order aberrations of the thin refractive tunable-focus lens," Opt. Lett. **35**, 1031–1033 (2010).
14. D.-Y. Zhang, N. Justis, and Y.-H. Lo, "Fluidic adaptive zoom lens with high zoom ratio and widely tunable field of view," Opt. Commun. **249**, 175–182 (2005).
15. H. Ren, D. Fox, P. A. Anderson, B. Wu, and S.-T. Wu, "Tunable-focus liquid lens controlled using a servo motor," Opt. Express **14**, 8031–8036 (2006).
16. H. Ren and S.-T. Wu, "Variable-focus liquid lens," Opt. Express **15**, 5931–5936 (2007).
17. A. F. Naumov, G. D. Love, M. Y. Loktev, and F. L. Vladimirov, "Control optimization of spherical modal liquid crystal lenses," Opt. Express **4**, 344–352 (1999).
18. H. Ren, Y.-H. Fan, S. Gauza, and S.-T. Wu, "Tunable-focus flat liquid crystal spherical lens," Appl. Phys. Lett. **84**, 4789–4791 (2004).
19. R. Marks, D. L. Mathine, G. Peyman, J. Schwiegerling, and N. Peyghambarian, "Adjustable fluidic lenses for ophthalmic corrections," Opt. Lett. **34**, 515–517 (2009).
20. F. C. Wippermann, P. Schreiber, A. Bräuer, and P. Craen, "Bifocal liquid lens zoom objective for mobile phone applications," Proc. SPIE **6501**, 650109 (2007).
21. A. Mikš, J. Novák, and P. Novák, "Method of zoom lens design," Appl. Opt. **47**, 6088–6098 (2008).
22. A. Clark, *Zoom Lenses* (Adam Hilger, 1973).
23. K. Yamaji, *Progress in Optics* (North-Holland, 1967), Vol. **VI**.
24. G. Wooters and E. W. Silvertooth, "Optically compensated zoom lens," J. Opt. Soc. Am. **55**, 347–351 (1965).
25. A. V. Grinkevich, "Version of an objective with variable focal length," J. Opt. Technol. **73**, 343–345 (2006).
26. K. Tanaka, "Paraxial analysis of mechanically compensated zoom lenses 1: four-component type," Appl. Opt. **21**, 2174–2183 (1982).
27. G. H. Matter and E. T. Luszcz, "A family of optically compensated zoom lenses," Appl. Opt. **9**, 844–848 (1970).
28. K. Tanaka, "Paraxial analysis of mechanically compensated zoom lenses 2: generalization of Yamaji Type V," Appl. Opt. **21**, 4045–4053 (1982).
29. K. Tanaka, "Paraxial analysis of mechanically compensated zoom lenses 3: five-component type," Appl. Opt. **22**, 541–553 (1983).
30. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 1999).
31. W. Welford, *Aberrations of the Symmetrical Optical Systems* (Academic, 1974).
32. A. Mikš and J. Novák, "Third-order aberration coefficients of a thick lens with a given value of its focal length," Appl. Opt. **57**, 4263–4266 (2018).
33. B. Havelka, *Geometrical Optics I, II* (Czech Academy of Science, 1955).
34. A. Mikš, *Applied Optics* (CTU in Prague, 2009).
35. A. Mikš, "Modification of the formulas for third-order aberration coefficients," J. Opt. Soc. Am. A **19**, 1867–1871 (2002).
36. C. G. Wynne, "Primary aberrations and conjugate change," Proc. Phys. Soc. London, Sect. B **65**, 429–437 (1952).

applied optics

Double curvature membrane lens

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
*Corresponding author: petr.pokorny@fsv.cvut.cz

Received 3 July 2020; revised 9 October 2020; accepted 9 October 2020; posted 12 October 2020 (Doc. ID 401892); published 2 November 2020

The paper presents a theoretical analysis of properties of a specific liquid membrane lens composed of two axially symmetric membranes of different thicknesses and double curvature. These membranes enclose a space where an optical liquid is filled. Mechanical and optical properties of the lens are then changed by varying the volume of the liquid. The paper presents new formulas for calculation of membrane deflections, radii of curvatures of the membranes, and axial geometry, which offer to minimize the third-order spherical aberration of the lens for an object at infinity. The presented theory is examined on specific examples.

https://doi.org/10.1364/AO.401892

1. INTRODUCTION

Active optical components with continuously varying optical properties, i.e., focal length, have found growing utilization in many fields of science, industrial applications, and medicine [1–15]. Those elements can be realized by many approaches, e.g., as mirrors with variable shapes or lenses with an active change in focal length [1–15]. There are many components available from commercial companies with a quality suitable especially for less demanding applications [13,14]. Nevertheless, development of those components is in progress, and one can expect intense propagation of more perfect samples, which will suit the most stringent requirements in the future. Afterwards, one would be able to design specific imaging lenses or objectives with variable focal lengths, for example, zoom lenses.

Recently, many papers presenting a design or experimental verification of liquid-filled tunable lenses have been published [6–8,12] that minimize especially spherical aberration under specific circumstances. For example, in Ref. [6], a pressure-actuated tunable biconvex microlens is simply designed, manufactured, and experimentally verified. However, there is no analytic description or detailed formulation of the lens's imaging properties presented. In Refs. [7,8,16], the authors design a membrane thickness to reduce the spherical aberration of a plano-convex liquid-filled lens. Similarly, Ref. [12] presents an iterative numerical design of a plano-convex membrane lens with corrected spherical aberration over a user-defined focal length range.

This paper presents a theoretical analysis of h_1 of a specific liquid membrane lens composed of two axially symmetric membranes of different thicknesses h_1 and h_2 and double curvature. A scheme of such a lens is shown in Fig. 1. The membranes enclose a space filled by optical liquid with refractive index n. By changing the volume of the liquid, the inner pressure p varies, and it is possible to deform membranes and to change the optical properties of the lens. In a nominal state, while the pressure does not deform membranes, the axial thickness of the lens is d_k. By raising the pressure, its value is changed to d (see Fig. 1). In Fig. 1, n_S denotes the refractive index of the membranes' material, w_1 and w_2 are axial deformations of the membranes, and d_t is the inner axial thickness for zero pressure. This type of lens enables two surfaces of different properties (e.g., different values of radii of curvatures) and variable axial thicknesses. A combination of those lenses can be used to design much more complex optical systems with continuous variation of optical characteristics (zoom systems).

In the first part of the paper, a calculation of membrane deformation under uniform pressure is presented. Afterwards, the paper discusses a spherical aberration of a thick lens, and formulas are then used to derive initial geometric parameters of the double curvature membrane lens (i.e., radii of curvatures, thicknesses, and distances of membranes of the lens) with minimized spherical aberration.

2. CALCULATION OF MEMBRANE DEFORMATION UNDER UNIFORM PRESSURE

This section summarizes a process of calculation of membrane deformation under uniform hydro-static pressure, and it enables to theoretically characterize imaging properties of the aforementioned lens of double curvature. The authors of this paper have presented many publications on this topic, e.g., Refs. [9–11].

Let the membrane of the liquid lens be of spherical shape with constant thickness h and radius a, which is clamped at its circumference. Without loss of generality, one can neglect the change in thickness of the membrane during deformation. It is shown in other papers of the authors [11,16] that the influence

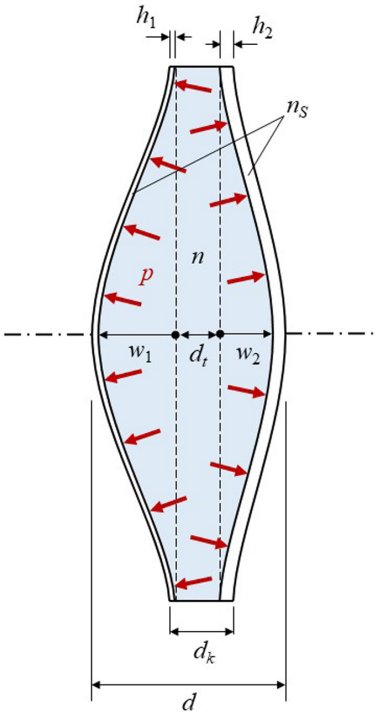


Fig. 1. Scheme of double curvature membrane lens.

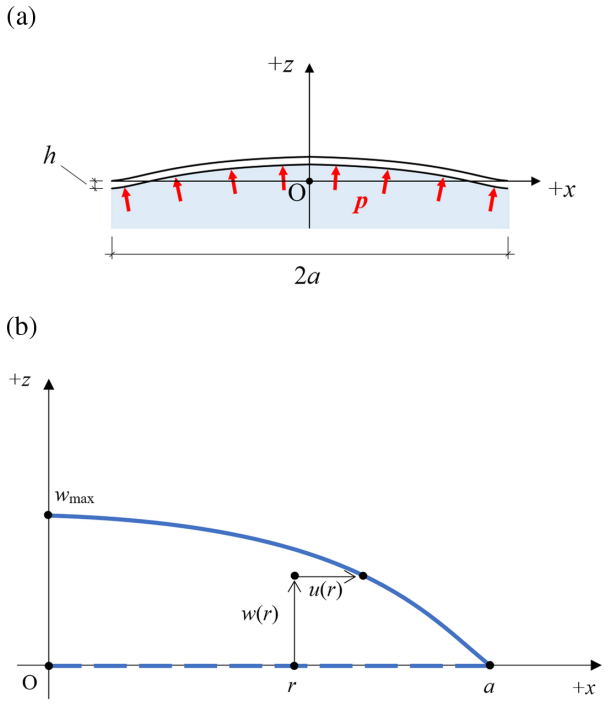


Fig. 2. (a) Scheme of circular membrane of radius a with constant thickness h loaded by constant pressure p in meridian section (x, z); (b) scheme of movement of selected point of membrane mid-plane during deformation from point $[r, 0]$ to $[r + u(r), w(r)]$, where $u(r)$ denotes the radial deformation and $w(r)$ the deformation in vertical direction.

of membrane thickness, and its variation, on wave spherical aberration is negligible for most practical purposes.

Considering the symmetry of the membrane, one can suppose a meridian section (x, z) [see Fig. 2(a)]. Under the load of constant pressure p, the membrane is deformed. Figure 2(b) shows movement of a selected point of the membrane's mid-plane from point $[r, 0]$ to $[r + u(r), w(r)]$, where $u(r)$ denotes the radial deformation and $w(r)$ the deformation in vertical direction. As one can suppose, the maximal deflection $w_{\max}$ occurs for $r = 0$.

Further, let the membrane have negligible bending stiffness. It is shown in previous papers of Pokorný *et al.* [9,10] that one can suppose a Saint-Venant–Kirchhoff material model with Green–Lagrange strain and the second Piola–Kirchhoff stress [17] for the membrane. Let the membrane in the double curvature lens be prestressed, and its initial stress in all directions is $\lambda_0 > 1$. Afterwards, one can write the following formulas for the stretch λ_r in radial direction and λ_t in tangent direction [10]:

$$\lambda_r = \lambda_0\sqrt{(1+u')^2 + w'^2}, \quad \lambda_t = \lambda_0\left(1 + \frac{u}{r}\right), \tag{1}$$

where u' and w' denote, respectively, derivations of functions $u = u(r)$ and $w = w(r)$ with respect to r. It is known from the theory of elasticity that the strain density energy W_{int} per unit volume can be expressed as follows [10,18,19]:

$$W_{\text{int}} = \frac{E}{2(1-\nu^2)}(\varepsilon_r^2 + 2\nu\varepsilon_r\varepsilon_t + \varepsilon_t^2), \tag{2}$$

where E is the Young modulus, ν is the Poisson number, and ε_r and ε_t are Green–Lagrange strains in radial and tangent directions, respectively, given by the formulas

$$\varepsilon_r = \frac{1}{2}(\lambda_r^2 - 1), \quad \varepsilon_t = \frac{1}{2}(\lambda_t^2 - 1). \tag{3}$$

Substitution of Eq. (1) into Eq. (3) then obtains [10]

$$\varepsilon_r = \lambda_0^2\left[u' + \frac{1}{2}\left(u'^2 + w'^2\right)\right] + \frac{1}{2}\left(\lambda_0^2 - 1\right),$$

$$\varepsilon_t = \lambda_0^2\left(\frac{u}{r} + \frac{u^2}{2r^2}\right) + \frac{1}{2}\left(\lambda_0^2 - 1\right). \tag{4}$$

To characterize the membrane after deformation, the potential energy can be minimized to achieve system equilibrium. If one is able to write a formula describing the potential energy of the system and to minimize it, with an optimal optimization algorithm, for example, the given problem will be solved. The potential energy E_p for constant pressure p is given as follows [10]:

$$E_p = E_{\text{int}} + E_{\text{ext}}, \tag{5}$$

where

$$E_{\text{int}} = 2\pi h\int_0^a W_{\text{int}} r\,\mathrm{d}r \tag{6}$$

is the inner elastic energy (i.e., the work of inner forces), and

$$E_{\text{ext}} = -2\pi p\int_0^a w(r+u)(1+u')\mathrm{d}r \tag{7}$$

is the outer forces' potential energy. After substitution of Eqs. (2) and (4) into Eqs. (5)–(7), one gets the final formula for potential energy E_p.

Let one study boundary conditions for the aforementioned case now. Figure 2(b) shows that the following formulas must be fulfilled for displacement in horizontal and radial directions:

$$w(a) = 0, \quad u(0) = 0, \quad u(a) = 0. \tag{8}$$

The three conditions in Eq. (8) can be completed with another one:

$$w'(0) = 0, \tag{9}$$

which is not necessary; however, it is considerable, with the symmetry of the problem and constant pressure uniformly loaded at the surface of the membrane. Moreover, it arises from numerical analyses that Eq. (9) significantly improves the convergence of calculation during an approximate solution of potential energy minimization.

One can find an equilibrium equation with variation calculus (variation of potential energy ∂E_p has to equal zero). One gets after modification and with boundary conditions in Eqs. (8) and (9) [10]

$$h[r\sigma_r(1+u')]' - h\sigma_t\left(1+\frac{u}{r}\right) = p(r+u)w',$$

$$-h(r\sigma_r w')' = p(r+u)(1+u'), \tag{10}$$

where σ_r and σ_t are stresses in radial and tangent directions, respectively, where the stress–strain law is applied ($\sigma_r = \partial W_{\text{int}}/\partial\varepsilon_r, \sigma_t = \partial W_{\text{int}}/\partial\varepsilon_t$). Therefore, it holds [10]:

$$\sigma_r = \lambda_0^2\frac{E}{1-\nu^2}(\tilde{\varepsilon}_r + \nu\tilde{\varepsilon}_t) + \frac{E}{1-\nu}\varepsilon_0,$$

$$\sigma_t = \lambda_0^2\frac{E}{1-\nu^2}(\tilde{\varepsilon}_t + \nu\tilde{\varepsilon}_r) + \frac{E}{1-\nu}\varepsilon_0, \tag{11}$$

where $\tilde{\varepsilon}_r = u' + \frac{1}{2}(u'^2 + w'^2)$, $\tilde{\varepsilon}_t = \frac{u}{r} + \frac{u^2}{2r^2}$, $\varepsilon_0 = \frac{1}{2}(\lambda_0^2 - 1)$ is the initial strain and $E\varepsilon_0/(1-\nu)$ the initial prestress σ_0.

It is obvious that Eq. (10) does not have an analytic solution. Therefore, one has to find an approximate form. In practical calculations, it is more suitable to find the minimum of the potential energy E_p given by Eq. (5). It will be used during further calculations. A specific solution can be found with the appropriate optimization algorithm [20,21] and series expansion.

It is suitable to choose the series in such a form to fulfill boundary conditions (8) and (9) directly. One of the options is to set deflections $w = w(r)$ and $u = u(r)$ in the form

$$w(r) = \sum_{i=1}^{N} b_i(a^{2i} - r^{2i}), \quad u(r) = \sum_{j=1}^{M} c_j(ra^{2j+1} - ar^{2j+1}), \tag{12}$$

where b_i $(i = 1, 2, \ldots, N)$ and c_j $(j = 1, 2, \ldots, M)$ are coefficients to be calculated. After substitution of Eq. (12) into Eqs. (2)–(7) and with the appropriate optimization algorithm, the approximate solution can be obtained for the minimum of the potential energy (5) with a goal function defined as follows:

$$M = \int_0^a \left[E_{\text{int}}(b_i, c_j, r) + E_{\text{ext}}(b_i, c_j, r)\right] \mathrm{d}r. \tag{13}$$

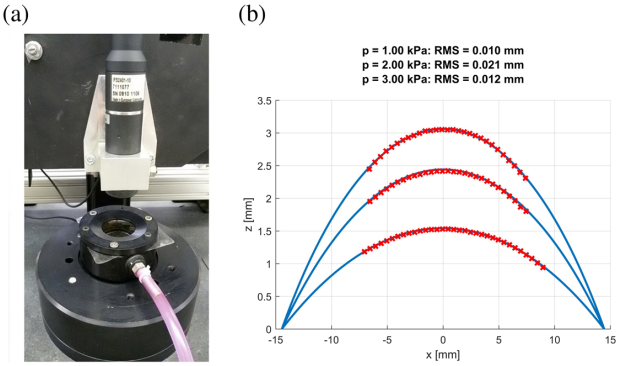


Fig. 3. (a) Device for measurement of surface topography with chromatic confocal sensor and experimental sample of plano-convex membrane lens; (b) comparison of experimentally measured data of membrane lens deformation (red crosses) and simulated data (blue curve) for pressures $p = 1.00$, 2.00 and 3.00 kPa.

The aforementioned theory of membrane deformation was experimentally verified. Figure 3(a) shows a device for measurement of surface topography with a chromatic sensor, and laboratory sample of a plano-convex membrane lens. Its main part is a chamber filled with optical liquid (distilled water in our case), which is enclosed by an elastic membrane and mechanically clamped at its edges. The liquid volume can be changed, which modifies loading pressure. After deformation, the optical parameters of the lens change. The membrane of the lens is made by Sylgard 184 [22]. Geometrical and mechanical parameters of the membrane were experimentally measured: $a = 14.45$ mm, $h = 0.45$ mm, $E = 1.078$ MPa, $\nu = 0.41$, $\sigma_0 = 0.0583$ MPa. Figure 3(b) shows a comparison between measured and predicted values of deflection. As the lens is symmetrical, the meridian section of the surface was evaluated only. The sensor has to be ideally positioned in a normal direction to the surface under test; therefore, only the vertex part of the lens was investigated. One can see that mean quadratic errors (RMS values) show very good agreement between theory and experiment, and it is in accordance with the nominal accuracy of the measurement device, estimated as 0.02 mm (considering uncertainties in a chromatic sensor measurement and in mechanical construction).

3. SPHERICAL ABERRATIONS OF THE THICK LENS IN AIR

It is generally well known that imaging with optical systems depends on their optical aberrations [23–26]. This section presents variations of aberrations of liquid lenses that are made by two membranes of different thicknesses. Without loss of generality, one can neglect the prestress of the membrane for the next calculations; therefore, $\lambda_0 = 1$. With a change in volume of the optical liquid inside the lens, the focal length of the lens will vary, and therefore its optical aberrations change as well. As the membranes have different thicknesses, and their deformation will differ. Therefore, radii of curvatures of the lens outer surfaces will not stay the same, and the so-called shape factor [23,24] of the lens changes as well as its thickness. It is known

[25] that the value of the shape factor of the thickness of the lens affects its aberrations.

It can be shown simply that if the membrane is made from material with the refractive index n_m and has thickness t, then the change in the lens surface's optical power caused by the membrane thickness is

$$\delta\varphi_m \approx (n_m - n)\frac{t}{r^2}, \tag{14}$$

where r is the radius of the surface, and n is the refractive index of the liquid. It is obvious from Eq. (14) that the change in optical power $\delta\varphi_m$ is small, and that the influence of the membrane on imaging properties of the focus variable liquid lens can be neglected.

This part discusses the effect of the shape factor and lens thickness on the spherical aberration. Let r_1 and r_2 be the vertex radii of curvature of the lens surfaces (membrane surfaces), n the refractive index of liquid between membranes, φ_1 and φ_2 the optical powers of lens surfaces, d the axial thickness of the lens, X the shape factor of the lens, and φ the optical power of the thick lens; then the following equations hold for the thick spherical lens in air:

$$\begin{aligned}
&X = \frac{r_2 + r_1}{r_2 - r_1}, \quad r_1 = \frac{(n-1)}{\varphi(X+1)}\left(\sqrt{\frac{\varphi d(X^2-1)+n}{n}} + 1\right),\\
&\varphi_1 = \frac{n-1}{r_1}, \quad \varphi_2 = \frac{\varphi - \varphi_1}{1 - d\varphi_1/n}, \quad r_2 = \frac{1-n}{\varphi_2},\\
&\varphi = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + \frac{d(n-1)^2}{nr_1r_2}.
\end{aligned} \tag{15}$$

The Seidel coefficient of spherical aberration S_{I} of a thick lens in air, for an object at infinity, can be expressed with the formula [25]

$$\begin{aligned}
\frac{S_{\mathrm{I}}}{\varphi^3} = &\frac{1}{(n-1)^2}\left[\frac{n+2}{n}\varphi_1^2 - (2n+1)\varphi_1 + n^2\right]\\
&+ d\left[\frac{\varphi_1(\varphi_1 - n)^2(\varphi_1 - n^2)}{n^3(n-1)^2}\right].
\end{aligned} \tag{16}$$

Further, the transverse, $\delta y'$, and lateral, $\delta s'$, spherical aberrations can be calculated as follows [23–25]:

$$\delta y' = -\tfrac{1}{2} f' H^3 S_{\mathrm{I}}, \quad \delta s' = -\tfrac{1}{2}(f'H)^2 S_{\mathrm{I}}, \tag{17}$$

where f' denotes the focal length of the lens, and H is the impinging height of the aperture ray at the first surface of the lens. Formulas for calculation of Seidel aberration coefficients of the coma, S_{II}, astigmatism, S_{III}, field curvature, S_{IV}, and distortion, S_{V}, of a thick lens are presented in Ref. [25]. Calculation of Seidel coefficients for different positions of the object and entrance pupil can be calculated with formulas presented in Refs. [23,24].

In the case of a classic lens in air, it is possible to realize such a situation in which the shape factor is constant while changing the focal length f_{I}' of the lens to the value $f_{\mathrm{II}}' = k f_{\mathrm{I}}'$, where k is constant. One can see from Eq. (15) that for such a case, the following formulas hold:

$$\begin{aligned}
&\frac{1}{f_{\mathrm{I}}'} = (n-1)\left(\frac{1}{r_1} - \frac{1}{r_2}\right) + \frac{d(n-1)^2}{nr_1r_2},\\
&\frac{1}{f_{\mathrm{II}}'} = \frac{1}{kf_{\mathrm{I}}'} = (n-1)\left(\frac{1}{kr_1} - \frac{1}{kr_2}\right) + \frac{kd(n-1)^2}{nkr_1kr_2}\\
&\qquad = (n-1)\left(\frac{1}{r_{1,\mathrm{II}}} - \frac{1}{r_{2,\mathrm{II}}}\right) + \frac{d_{\mathrm{II}}(n-1)^2}{nr_{1,\mathrm{II}}r_{2,\mathrm{II}}}.
\end{aligned} \tag{18}$$

Therefore, the lens with focal length f_{II}' has the radii of curvatures $r_{1,\mathrm{II}} = kr_1$, $r_{2,\mathrm{II}} = kr_2$ and axial thickness $d_{\mathrm{II}} = kd$, which implies

$$X_{\mathrm{I}} = \frac{r_2 + r_1}{r_2 - r_1}, \quad X_{\mathrm{II}} = \frac{kr_2 + kr_1}{kr_2 - kr_1} = X_{\mathrm{I}}. \tag{19}$$

If the radii of curvatures and the thickness of the lens are multiplied by the same constant value, the focal length changes in the same way; nevertheless, the shape factor of the lens remains constant.

Let one calculate parameters of a thick lens with minimized spherical aberration of the third order for an object at infinity. Without loss of generality, one can suppose the focal length of the lens $\bar{f}' = 1$. Therefore, the optical power gives $\bar{\varphi} = 1$ as well. Finding an extreme of S_{I} [Eq. (16)] with respect to φ_1 (the first derivative equals zero), one gets

$$\begin{aligned}
&4\bar{d}\bar{\varphi}_1^3 - 3\bar{d}n\,(n+2)\,\bar{\varphi}_1^2 + n^2\left[2n\left(2\bar{d}+1\right) + 2\left(\bar{d}+2\right)\right]\bar{\varphi}_1\\
&\quad - n^3\left[n\left(\bar{d}+2\right) + 1\right] = 0,
\end{aligned} \tag{20}$$

where $\bar{\varphi}_1$ and $\bar{d}$ denote, respectively, the optical power of the first surface and the lens thickness for a normalized value of $\bar{\varphi} = 1$.

To the authors' best knowledge, Eq. (20) has not been published in previous papers; therefore, it offers an original approach to calculate the lens with the aforementioned properties.

The solution to Eq. (20) gives a value of the optical power $\bar{\varphi}_1$, and one can simply calculate the first radius of curvature of the thick lens: $\bar{r}_1 = (n-1)/\bar{\varphi}_1$. The second radius then can be calculated with the following formula:

$$\bar{r}_2 = -\frac{(n-1)\left(\bar{d} - \bar{d}n + n\bar{r}_1\right)}{n\left(\bar{r}_1 - n + 1\right)}. \tag{21}$$

To calculate the parameter of the lens with a focal length value different from $\bar{f}' = 1$, but $f' = k\bar{f}'$, where k denotes a constant, it holds: $r_1 = k\bar{r}_1$, $r_2 = k\bar{r}_2$, and $d = k\bar{d}$.

Table 1 shows an example situation of lenses with different values of refractive indices n for a thin lens, $d = 0$ mm, as well as for a thick lens with axial thickness $d = 0.1$ mm. Optical powers of lenses are $\varphi = 1/f' = 1$ mm. One can simply calculate parameters for different values of the focal length f' by multiplying radii of curvatures and axial thicknesses by this value of the focal length. All the lenses calculated in this way have minimized third-order spherical aberration for an object at infinity. In Table 1, all lengths are in millimeters.

Table 1. Parameters of Lenses with Minimized Third-Order Spherical Aberrations for an Object at Infinity for Different Values of Refractive Indices

$f' = 1$ mm, $d = 0$ mm			$f' = 1$ mm, $d = 0.1$ mm		
n	r_1 [mm]	r_2 [mm]	n	r_1 [mm]	r_2 [mm]
1.500	0.5833	−3.5000	1.500	0.5891	−3.1197
1.625	0.6561	−13.1818	1.625	0.6628	−10.3311
1.750	0.7143	15.0000	1.750	0.7217	18.0191
1.875	0.7614	5.8649	1.875	0.7696	5.9985
2.000	0.8000	4.0000	2.000	0.8088	3.9673

4. EXAMPLES

A. Example 1

Suppose now an example of a liquid membrane lens composed of two membranes while the first has thickness $h_1 = 0.5$ mm and the second one has thickness $h_2 = 0.8$ mm, and study the effect of inner pressure variation on imaging properties of the lens. Let the membranes be made by Sylgard 184 [22] and the transverse radius of the lens be 15 mm. Results of calculations are presented in Table 2, where p denotes the inner pressure in kPa, X is the shape factor, f' is the focal length of the lens, d is the axial thickness, r_1 and r_2 are the radii of curvatures, S_I is Seidel coefficient of spherical aberration, $\delta y'$ denotes the transverse spherical aberration for an aperture ray impinging on the first surface at height H from the optical axis calculated with Eq. (17), $\delta y'_{\mathrm{OSLO}}$ denotes the same value calculated with the software OSLO, w_1 and w_2 are the maximal deflections of the membranes, n is the refractive index of the liquid filled between the membranes (water), and λ is the used wavelength. Lengths are shown in millimeters in Table 2.

As one can see from Table 2, changing the pressure in the lens affects its focal length f', its axial thickness d, radii of curvatures r_1, r_2, as well as its shape factor X. Dependence of the focal length f' on the pressure p can be approximated with the formula

$$f' = \frac{a_1 p + a_2}{p + a_3}, \tag{22}$$

where $a_1 = 32.466$ mm, $a_2 = 137.700$ mm · kPa, and $a_3 = 1.256$ kPa. The RMS error of approximation equals 0.14 mm for this case. Dependence of the shape factor X on the pressure p can then be expressed as

$$X = \frac{x_1 p + x_2}{p + x_3}, \tag{23}$$

where $x_1 = 0.024$, $x_2 = 1.395$ kPa, and $x_3 = 18.297$ kPa, with the RMS error of the approximation less than 2e−4.

B. Example 2

The second example presents a double curvature membrane lens with a minimized value of spherical aberration for an object at infinity. Let the initial parameters of the lens be $f' = 20$ mm and $d = 10$ mm.

Reference [10] explains that the maximal deformations w of membranes are related to the mid-plane. Imaging of the aforementioned lens will be generally realized by four refractive surfaces (see Fig. 1). Denote the axial distance between the first and second surface of the lens $t_1 = h_1$ and the refractive index between those surfaces as n_S. The index of refraction between the second and third surface is n, and the axial distance between them is given with the formula $t_2 = w_1 + w_2 + d_t$, where d_t is a correction that has to be calculated to maintain the required focal length f' of the lens (axial distance between membranes with applied zero pressure). Similar to the first and second surface, the distance between the last two surfaces is $t_3 = h_2$, and those surfaces enclose a material with the refractive index n_S. It is obvious that it holds for the axial thickness of the lens $d = t_1 + t_2 + t_3$, and that $d_t \geq 0$ has to be fulfilled to maintain the lens design possible for manufacture.

Without loss of generality, one can further suppose the same index of refraction for the membrane material and the liquid, e.g., $n_S = n = 1.4225$ (Sylgard 184 [22] in liquid and solid form), and that the radii of curvatures for both membrane surfaces are constant (change in curvature between the mid-plane and its outer surfaces is negligible). With Eqs. (20) and (21), one gets for minimized spherical aberration the following values: $\bar{d} = 0.5$ mm, $\bar{r}_1 = 0.5551$ mm, $\bar{r}_2 = -1.2953$ mm, and, therefore, $r_1 = 11.1025$ mm, $r_2 = -25.9070$ mm for the designed lens.

The further goal would be to find thicknesses of the first and second membranes for a given pressure p and required values of the focal length, i.e., to calculate corresponding radii of curvatures. Supposing the same membrane parameters as in the first example (material Sylgard 184 [22]) and the transverse radius of the lens 5 mm, one gets thicknesses h_1 and h_2 of the membranes and other lens parameters for selected pressures p, which are shown in Table 3. Units of pressures in Table 3 are kPa, and length units are in millimeters. The thicknesses were calculated to obtain corresponding radii of curvatures with error less than 1e−3mm.

Figure 4 shows a ray-tracing through the double curvature membrane lens for the second example in Table 3, where the surfaces were calculated with formulas presented in Section 2.

Table 2. Parameters of Double Curvature Membrane Lens

$H = 5$ mm, $n = 1.333$, $\lambda = 587$ nm, $h_1 = 0.5$ mm, $h_2 = 0.8$ mm, $d_t = 2$ mm										
p [kPa]	X	f' [mm]	d [mm]	r_1 [mm]	r_2 [mm]	$1000\,S_\mathrm{I}$	$\delta y'$ [mm]	$\delta y'_{\mathrm{OSLO}}$ [mm]	w_1 [mm]	w_2 [mm]
1.00	0.0735	75.4386	7.2779	45.9431	−53.2351	0.0111	−0.0524	−0.0508	2.8478	2.4301
4.00	0.0668	50.8614	10.4616	30.4838	−34.8445	0.0365	−0.1161	−0.1082	4.5723	3.8893
7.00	0.0618	44.3218	12.2750	26.2777	−29.7415	0.0554	−0.1536	−0.1390	5.5585	4.7165
10.00	0.0577	41.0056	13.6499	24.0992	−27.0491	0.0703	−0.1802	−0.1590	6.3086	5.3414

Table 3. Parameters of Double Curvature Membrane Lens for Different Inner Pressured and Fixed Focal Lengths

p [kPa]	h_1 [mm]	h_2 [mm]	r_1 [mm]	r_2 [mm]	w_1 [mm]	w_2 [mm]	d_t [mm]
1.00	0.0570	0.7503	11.1029	−25.9069	1.3676	0.5703	7.2548
4.00	0.2282	3.0009	11.1026	−25.9070	1.3677	0.5703	4.8329
7.00	0.3992	5.2520	11.1018	−25.9072	1.3678	0.5703	2.4107

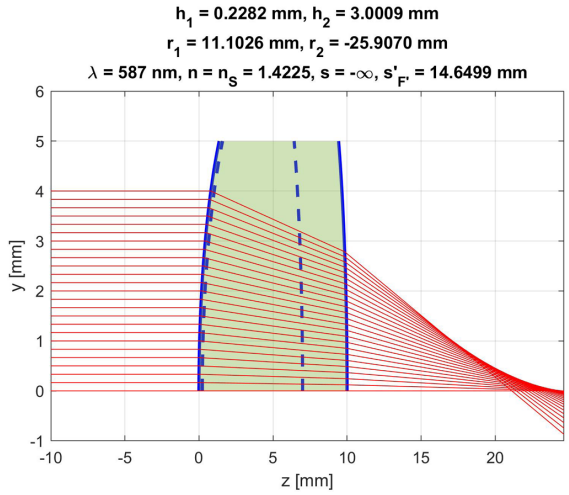


Fig. 4. Ray-tracing through the double curvature membrane lens (blue solid lines, surfaces of the lens; blue dashed lines, inner surfaces of membranes).

In the figure, s represents the axial distance of an object from the first surface of the lens, and $s'_{F'}$ is the axial distance of the focal point from the second surface of the thick spherical lens with the same vertex radii of curvatures as the double membrane lens calculated by the known formula for a thick lens in air [23]:

$$s'_{F'} = f'\left(1 - \varphi_1 \frac{d}{n}\right), \quad \varphi_1 = \frac{n-1}{r_1}, \tag{24}$$

where f' is the focal length of the lens, φ_1 is the optical power of the first surface of the lens, d denotes the axial thickness of the lens, and n is the index of refraction of the lens. To demonstrate imaging properties of this lens, Fig. 5 shows transverse spherical aberrations for the case in which an image plane is located at the distance $s' = s'_{F'}$ behind the second surface of the lens [Fig. 5(a)], and at the distance $s' = s'_{F'} + \mathrm{d}s$ [Fig. 5(b)], where $\mathrm{d}s$ is optimized according to the minimization of a gyration radius r_g calculated from the equation (for a circular exit pupil) [26]

$$r_g^2 = \frac{2}{H^2}\int_0^H (\delta y')^2 h \mathrm{d}h, \tag{25}$$

where h denotes the radial coordinate in the plane of the exit pupil of the lens, H is its maximal value, and $\delta y'$ is the transverse spherical aberration in the image plane. In Fig. 5, y_0 denotes the impinging height of rays on the first surface of the lens. In the aforementioned figures, the calculations were done for the wavelength $\lambda = 587$ nm and for an object at infinity.

One can see from Fig. 5 that the mentioned double curvature membrane lens has very good imaging quality for imaging of an axial point at infinity even if it is a simple lens. The gyration radius $r_g = 1.29\mathrm{e}-01$ mm. Therefore, the presented formulas are designed from initial paraxial equations, and the procedure can be used as a very good starting point for further optimization processes.

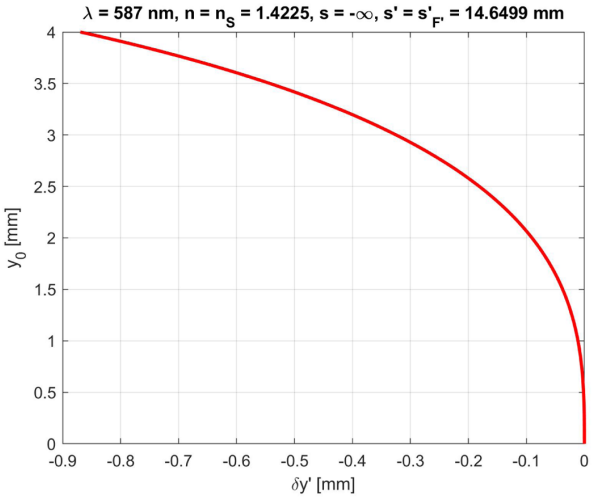


Fig. 5. Transverse spherical aberration of the double curvature membrane lens. (a) Transverse spherical aberration $\delta y'$ in an image plane of the lens in an axial distances $s' = s'_{F'}$ from the second surface of the lens; y_0 denotes impinging height of the ray on the first surface of the lens. (b) Transverse aberration $\delta y'$ in an image plane of the lens in an axial distance s' from the second surface of the lens, while s' is optimized for minimal gyration radius r_g; y_0 denotes impinging height of the ray on the first surface of the lens.

5. CONCLUSION

The paper presented a theoretical analysis of the double curvature membrane lens with different axial thicknesses of the membranes whose optical properties vary with changes in inner uniform pressure. The paper presented new formulas for calculation of membrane deflections, radii of curvatures of the membranes, and axial geometry, which offer to minimize the third-order spherical aberration of the lens for an object at infinity. The presented theory was examined on specific examples.

Funding. České Vysoké učení Technické v Praze (SGS20/093/OHK1/2T/11).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. J. D. Campbell, "On the theory of initially tensioned circular membranes subjected to uniform pressure," Quart. J. Mech. Appl. Math. **9**, 84–93 (1956).
2. A. Miks, J. Novak, and P. Novak, "Generalized refractive tunable-focus lens and its imaging characteristics," Opt. Express **18**, 9034–9047 (2010).
3. L. Li, Q.-H. Wang, and W. Jiang, "Liquid lens with double tunable surfaces for large power tunability and improved optical performance," J. Opt. **13**, 115503 (2011).
4. G. Li and T. Mauger, "Adaptive lenses for vision correction and optical imaging," in *Frontiers in Optics 2012/Laser Science XXVIII* (OSA, 2012).
5. H. Ren and S.-T. Wu, *Introduction to Adaptive Lenses* (Wiley, 2012).
6. Y.-K. Fuh, M.-X. Lin, and S. Lee, "Characterizing aberration of a pressure-actuated tunable biconvex microlens with a simple spherically-corrected design," Opt. Lasers Eng. **50**, 1677–1682 (2012).
7. P. Zhao, Ç. Ataman, and H. Zappe, "Spherical aberration free liquid-filled tunable lens with variable thickness membrane," Opt. Express **23**, 7567–7591 (2015).
8. P. Zhao, Ç. Ataman, and H. Zappe, "Gravity-immune liquid-filled tunable lens with reduced spherical aberration," Appl. Opt. **55**, 21264–21278 (2016).
9. P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek, "Calculation of nonlinearly deformed membrane shape of liquid lens caused by uniform pressure," Appl. Opt. **56**, 5939–5947 (2017).
10. P. Pokorný, F. Šmejkal, P. Kulmon, P. Novák, J. Novák, A. Mikš, M. Horák, and M. Jirásek, "Deformation of a prestressed liquid lens membrane," Appl. Opt. **56**, 9368–9376 (2017).
11. A. Mikš and F. Šmejkal, "Dependence of the imaging properties of the liquid lens with variable focal length on membrane thickness," Appl. Opt. **57**, 6439–6445 (2018).
12. H. Zhou, X. Zhang, Z. Xu, P. Wu, and H. Yu, "Universal membrane-based tunable liquid lens design for dynamically correcting spherical aberration over user-defined focal length range," Opt. Express **27**, 37667–37679 (2019).
13. "Varioptic," http://www.varioptic.com.
14. "Optotune," http://www.optotune.com.
15. "TAG Optics," https://www.tag-optics.com.
16. M. Jirásek, F. Šmejkal, and M. Horák, "Pressurized axisymmetric membrane deforming into a prescribed shape," Int. J. Solids Struct. **198**, 1–16 (2020).
17. B. Audoly and Y. Pomeau, *Elasticity and Geometry: From Hair Curls to the Non-linear Response of Shells* (OUP Oxford, 2010).
18. S. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells* (Textbook Publishers, 2003).
19. A. Volmir, Flexible Plates and Shells, Technical Report (Air Force Flight Dynamics Laboratory, Research and Technology Division, Air Force Systems Command, 1967).
20. L. Scales, *Introduction to Non-Linear Optimization*, Computer Science Series (Macmillan Education UK, 1985).
21. M. Aoki, *Introduction to Optimization Techniques: Fundamentals and Applications of Nonlinear Programming*, Macmillan series in applied computer sciences (Macmillan, 1971).
22. "Sylgard 184 silicone elastomer kit," http://www.dowcorning.com/applications/search/products/Details.aspx?prod=01064291.
23. A. Mikš, *Applied Optics* (CTU in Prague, 2009).
24. W. Welford, *Aberrations of Optical Systems* (CRC Press, 2017).
25. A. Mikš and J. Novák, "Third-order aberration coefficients of a thick lens," Appl. Opt. **51**, 7883–7886 (2012).
26. A. Mikš and P. Pokorný, "Spherical aberration of an optical system and its influence on depth of focus," Appl. Opt. **56**, 5009–5015 (2017).

applied optics

3D optical two-mirror scanner with focus-tunable lens

PETR POKORNY* AND ANTONIN MIKS

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 19 May 2015; revised 2 July 2015; accepted 11 July 2015; posted 13 July 2015 (Doc. ID 241320); published 31 July 2015

The paper presents formulas for a ray tracing in the optical system of two-mirror optical scanner with a focus-tunable lens. Furthermore, equations for the calculation of focal length which ensure focusing of a beam in the desired point in a detection plane are derived. The uncertainty description of such focal length follows as well. The chosen vector approach is general; therefore, the application of formulas in various configurations of the optical systems is possible. In the example situation, the authors derived formulas for mirrors' rotations and the focal length depending on the position of the point in the detection plane.

OCIS codes: (080.2740) Geometric optical design; (080.4035) Mirror system design; (110.1085) Adaptive imaging; (220.1080) Active or adaptive optics.

http://dx.doi.org/10.1364/AO.54.006955

1. INTRODUCTION

In the industrial application it is very often important to determine the spatial parameters of different kinds of objects (e.g., a surface topography of the object under study), or to perform specific modification in three-dimensional space (e.g., cutting or welding). Several principles exist for such purposes. One of the possibilities is the usage of 3D optical scanners for their speed and accuracy [1–4]. The beam-steering device can be constructed in many ways. From the chosen type of construction the principles of usage are derived. As an example, let one name one-mirror scanners, two-mirror scanners, polygonal, acousto-optical or electro-optical systems, etc. [1]. In the last decades the adaptive optical elements development has been a focus of interest and especially the focus-tunable lenses have been studied [5–11].

From the general point of view one can divide the optical scanners into three categories: (a) measurement devices for revision and testing of current objects, (b) projection technology, and (c) high-power technology in industry.

The first of the above-mentioned categories covers the application in civil engineering and architecture, such as documentation of interiors, mining, building foundation, volume measurement, underground network measurement, or similar. Also mechanical engineering is part of this category, as well as reverse engineering, quality control, automobile industry, documentation of traffic networks, digital terrain modelling, historical and art documentation, security purposes, crack detection, water engineering, archaeology, or medicine. The second characteristic covers the entertainment industry, light effects, video mapping or theatre lightning. And the last category, the application of high-power sources in industry, describes manufacturing technologies for cutting, welding, engraving, or coating.

In the last couple of years many papers have been written that were especially focused on the one-dimensional or two-dimensional description of mirror scanners. Flat mirrors were analyzed by Pegis and Rao [12]. Two-dimensional free-of-distortion scanners were presented by Sabban *et al.* [13]. Image shape analysis as an intersection with the detection plane was presented as well [14–17]. Stability of optical systems was analyzed by Friedman and Schweitzer [18]. Profiling based on beam reflection is shown by Shinozaki *et al.* [19]. Three-dimensional analysis of the position and accuracy of the beam's spot in the detection plane was presented by Pokorny [3]. The adaptive optical components were studied in detail by Miks *et al.* [5–11].

In this paper, the authors present a general vector approach for a ray tracing in the optical beam-steering device of a two-mirror scanner with a plano–convex focus-tunable lens. The position of the beam's spot is calculated in the detection plane as a function of the mirrors' angles of rotation. Moreover, the calculation of the focal length of the focus-tunable lens is presented based on the condition of beam focusation in the detection plane. This condition finds utilization in many practical applications. At the end of the paper the results are presented in examples. The calculation of the mirrors' rotations for one type of scanner is carried out as well. To the authors' best knowledge, derived formulas of such type have never been published yet.

1559-128X/15/226955-06$15/0$15.00

2. GENERAL VECTOR MODEL FOR RAY-TRACING ANALYSIS

Based on the calculation of the beam spot position in the detection plane presented in the previous authors' articles [3,4], the modification for the system with a plano–convex focus-tunable lens is going to be presented.

Suppose the situation in Fig. 1. System K transforms the incoming light to a beam of parallel rays. The position of such system's last surface is given by vector $\mathbf{s}_0$. The vertex V_{L2} of a rear surface of a focus-tunable lens defines its position by a vector $\mathbf{l}_{V2}$ in such a way that points S_0, V_{L2}, and O_1 define a line with a unit directional vector $\mathbf{o}$. Let this line be an optical axis of the lens. Unit vector $\mathbf{a}_0$ with point S, which is given by position vector $\mathbf{s}$, determines a parallel ray emerging from the system K. Unit vector $\mathbf{a}_1$ with point L, given by position vector $\mathbf{l}$, then defines a ray behind the lens.

The position of the mirrors in the system is defined by vectors $\mathbf{o}_1$ and $\mathbf{o}_2$, and by their unit normal vectors $\mathbf{n}_1$ and $\mathbf{n}_2$. Position vectors $\mathbf{r}_1$ and $\mathbf{r}_2$ denote places where the ray impinges and reflects from the mirrors. Unit vectors $\mathbf{a}_2$ and $\mathbf{a}_3$ then denote directions of the ray after the reflection on the first and the second mirror. Planar detector, detection plane ξ, is supposed to be given by a normal vector $\mathbf{q}$ and it is distanced by a value of d from the origin of the coordinate system. The intersection of the ray with such plane ξ is determined by position vector $\mathbf{p}$ of the point P.

For the next analysis let one suppose the rotationally symmetric plano–convex focus-tunable lens, which is very often used in practice. The considered situation is depicted in Fig. 2.

In Fig. 2, point S and unit directional vector $\mathbf{a}_0$ describe an off-axis ray. The point of intersection L_0 lays in height h, which can be calculated as a distance between the incoming ray and the optical axis of the lens. From the depicted situation it is obvious that $\mathbf{a}_0$ equals $\mathbf{o}$. According to the law of refraction the ray is refracted in the point L_0 and the unit directional vector $\mathbf{a}_L$ defines its direction in the lens with index of refraction n_L. In the point L, the ray is refracted again and emerges the lens in the direction of unit vector $\mathbf{a}_1$. Without the loss of generality for practical purposes, where the geometrical aberrations caused by the optical components are usually smaller than diffraction effects, this model of off-axis ray can be converted to the paraxial one with the use of formulas of geometrical optics [20,21]. Therefore, the following simplifications hold: $h \to 0$, $\mathbf{a}_0 = \mathbf{o}$, $(\mathbf{s} - \mathbf{s}_0) \to \mathbf{0}$, $(\mathbf{l} - \mathbf{l}_{V2}) \to \mathbf{0}$, $(\mathbf{a}_1 - \mathbf{o}) \to \mathbf{0}$, where $\mathbf{0}$ denotes null vector. The authors show a practical example and a proof of such simplification at the end of this paper.

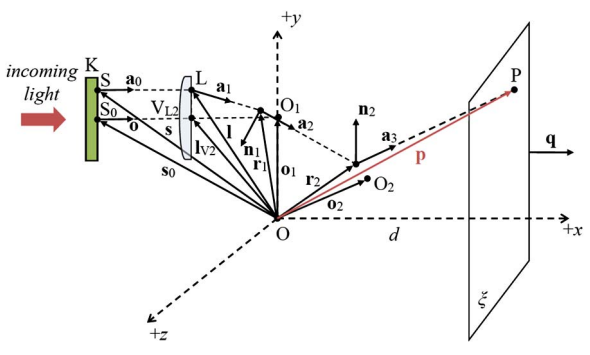


Fig. 1. Two-mirror scanner.

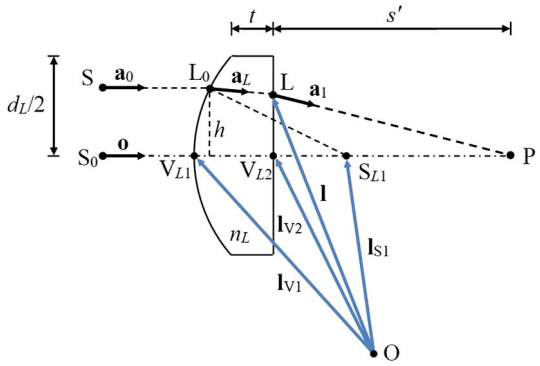


Fig. 2. Ray-tracing of a beam in a plano–convex lens.

For the geometrical characteristics, let the position of the lens be given by stable center V_{L2} of its planar surface by its positional vector $\mathbf{l}_{V2}$. As the focal length of the lens changes, the position of front vertex V_{L1} changes. The center of curvature moves its position as well. As is obvious in Figs. 1 and 2, one can characterize the position of point V_{L1} as follows:

$$\begin{aligned}\mathbf{l}_{V1} &= \mathbf{l}_{V2} - t_L\mathbf{o} \\ &= \mathbf{l}_{V2} - \left[t + R - \frac{1}{2}\sqrt{4R^2 - d_L^2}\right]\mathbf{o}. \end{aligned} \tag{1}$$

In Eq. (1) t_L denotes variable axial thickness of the lens; t is the fixed thickness of the lens measured at the edge; d_L is the diameter of the lens; R denotes the lens first surface's radius of curvature, which can be enumerated from the formula $R = f'(n_L - 1)$, whereas f' is the back focal length of the lens and n_L is the index of refraction of its material. One can derive the following formula for variable position of the center of curvature, it holds:

$$\begin{aligned}\mathbf{l}_{S1} &= \mathbf{l}_{V1} + R\mathbf{o} \\ &= \mathbf{l}_{V2} - \left[t - \frac{1}{2}\sqrt{4R^2 - d_L^2}\right]\mathbf{o}. \end{aligned} \tag{2}$$

Let one start with ray tracing through the mirror system with the description of unit normal vector n_i (φ_i) of the i-th mirror, rotated by angle φ_i around the axis with unit directional vector $\mathbf{c}_i$. From the analytical geometry it can be shown that the following formula represents the aforementioned situation [3,4]:

$$\begin{aligned}\mathbf{n}_i(\varphi_i) = \mathbf{n}_i(0)\cos\varphi_i + \mathbf{c}_i[\mathbf{c}_i \cdot \mathbf{n}_i(0)](1 - \cos\varphi_i) \\ + [\mathbf{c}_i \times \mathbf{n}_i(0)]\sin\varphi_i. \end{aligned} \tag{3}$$

As the paraxial model is supposed, the position vector $\mathbf{r}_1$ equals $\mathbf{o}_1$. Afterwards, one can write for the unit directional vector of the ray reflected on the first mirror as follows [3,4]:

$$\mathbf{a}_2(\varphi_1) = \mathbf{a}_1 - 2\mathbf{n}_1(\varphi_1)[\mathbf{a}_1 \cdot \mathbf{n}_1(\varphi_1)]. \tag{4}$$

With the use of analytical geometry [22,23] the formula for the point $\mathbf{r}_2$ of the ray's intersection with the second mirror can be derived, it holds:

$$\mathbf{r}_2 = \mathbf{o}_1 + \frac{[\mathbf{o}_2 - \mathbf{o}_1]\cdot\mathbf{n}_2(\varphi_2)}{\mathbf{a}_2(\varphi_1)\cdot\mathbf{n}_2(\varphi_2)}\mathbf{a}_2(\varphi_1). \tag{5}$$

Similar to Eq. (4), the unit directional vector of the ray after the reflection on the second mirror can be written as follows [3,4]:

$$\mathbf{a}_3(\varphi_2) = \mathbf{a}_2(\varphi_1) - 2\mathbf{n}_2(\varphi_2)[\mathbf{a}_2(\varphi_1)\cdot\mathbf{n}_2(\varphi_2)]. \tag{6}$$

The ray's intersection with the detection plane, i.e., the point P, is afterwards given by positional vector $\mathbf{p}$. It can be shown that the following formula holds [3,4]:

$$\mathbf{p} = \mathbf{r}_2 + \frac{d - \mathbf{r}_2\cdot\mathbf{q}}{\mathbf{a}_3(\varphi_2)\cdot\mathbf{q}}\mathbf{a}_3(\varphi_2). \tag{7}$$

The beam spot's position calculation is therefore completely described.

The focus-tunable lens can be manipulated in such a way that the beam will be focused in the point P in the detection plane. Let one derive the formula for the calculation of the focal length.

The optical power of the aforementioned plano–convex lens will be given as follows [20,21]:

$$\varphi = \frac{1}{f'} = \frac{n_L - 1}{R}. \tag{8}$$

In Eq. (8), f' denotes the variable focal length, n_L is the index of refraction, R is the radius of curvature of the first lens' spherical surface. As the optical power changes, the vertex V_{L1} of the spherical surface shifts on the optical axis.

With the use of formulas of geometrical optics the distance s' between the back lens' surface and the image plane (distance between points V_{L2} and P) of such a lens can be written, it holds [20,21]:

$$s' = f'\left(1 - \varphi\frac{t_L}{n_L}\right). \tag{9}$$

The meaning of the symbols in Eq. (9) is obvious from previous statements.

If one substitutes Eqs. (1) and (8) in Eq. (9), then the following formula comes:

$$f' = s' + \frac{t_L}{n_L} = s' + \frac{1}{n_L}\left[t + R - \frac{1}{2}\sqrt{4R^2 - d_L^2}\right]. \tag{10}$$

After some elementary rearrangements the quadratic formula for the focal length can be written as follows:

$$\frac{f'^2}{n_L^2}[(n_L - 1)^2 - 1] + 2\left(s' + \frac{t}{n_L}\right)\frac{f'}{n_L} - \frac{d_L^2}{4n_L^2} - \left(s' + \frac{t}{n_L}\right)^2 = 0. \tag{11}$$

The suitable solution for our purposes follows:

$$f' = -\frac{1}{n_L(n_L - 2)}\left(t + n_L s' - \frac{1}{2}\sqrt{D_f}\right). \tag{12}$$

The value of D_f can be calculated from the following formula:

$$\begin{aligned} D_f = {} & 4n_L^4 s'^2 - 8n_L^3 s'(s' - t) \\ & + n_L^2[d_L^2 + 4(s'^2 + t^2) - 16s't] \\ & + n_L[8t(s' - t) - 2d_L^2] + 4t^2. \end{aligned} \tag{13}$$

As obvious from Eqs. (12) and (13), the value of ideal focal length for focusing the beam in the point P in the detection plane can be calculated if the value of s' is known. From the aforementioned formulas, the value of s' can be calculated from the relation as follows:

$$s' = |\mathbf{o}_1 - \mathbf{l}_{V2}| + |\mathbf{r}_2 - \mathbf{o}_1| + |\mathbf{p} - \mathbf{r}_2|. \tag{14}$$

Let one determine the uncertainty $u(f')$ of the calculated focal length given by Eq. (12). Suppose the dependency of such uncertainty on uncertainties of side thickness of the lens $u(t)$, the index of refraction $u(n_L)$, the value of distance between points V_{L2} and P $u(s')$, and the lens diameter $u(d_L)$. Without any loss of generality one can consider the aforementioned uncertainties very small compared to the absolute values of their mutually independent parent variables, i.e., the values of uncertainties are expected to be comparable to the total differentials of the given model. Suppose next the even probability distribution for such uncertainties. The conditions describe the situation of random errors only; therefore, one considers systematic errors to be suppressed. With these assumptions the law of variance propagation can be used [24,25] and the uncertainty of the focal length can be expressed as follows:

$$u(f') = \sqrt{U_1^2u^2(t) + U_2^2u^2(n_L) + U_3^2u^2(s') + U_4^2u^2(d_L)}, \tag{15}$$

where coefficients U_i $(i = 1, \ldots, 4)$ are given with following formulas, it holds:

$$U_1 = N(DV_1 - 1),$$

$$U_2 = N^2\left[(n_L - 2)n_L DV_2 + (n_L - 1)\left(2t - \sqrt{D_f}\right) + n_L^2 s'\right],$$

$$U_3 = N(DV_3 - n_L),$$

$$U_4 = NDV_4,$$

$$V_1 = 8[s'n_L^3 + (t - 2s')n_L^2 + (s' - 2t)n_L + t],$$

$$V_2 = 16s'^2n_L^3 - 24s'(s' - t)n_L^2 + 2[d_L^2 + 4(s'^2 + t^2) - 16s't]n_L + 8t(s' - t) - 2d_L^2,$$

$$V_3 = 8[s'n_L^4 - n_L^3(2s' - t) + s'n_L^2 - 2tn_L^2 + tn_L],$$

$$V_4 = 2d_L(n_L^2 - 2n_L),$$

$$D = \frac{\sqrt{D_f}}{4D_f}, \qquad N = \frac{1}{n_L(n_L - 2)}.$$

Suppose now the situation when one calculates the smallest possible beam's spot in the detection plane and the position of such an ideal image plane as well. As known from the theory of optical imaging, the unique plane where the beam spot has the smallest diameter exists for each optical system and for a given field of view.

The transverse ray aberration $\delta y'$ of the plano–convex lens (see Fig. 2), and for an object at the infinity can be calculated with formulas from [9,20,21] as follows:

$$\delta y' = -\frac{f'}{2}\left(\frac{h}{f'}\right)^3\left[\frac{n_L^3 - 2n_L^2 + 2}{n_L(n_L-1)^2}\right], \tag{16}$$

where h is the incidence height of the aperture ray and n_L is the index of refraction. This formula is valid for thin plano–convex lenses, having the first radius spherical and the second radius infinite. Formula (16) gives sufficiently accurate results also for thick lenses with moderate f-number. For the f-number F [20,21] of the lens it holds: $F = f'/d_L$. Substituting the formula for the f-number of the optical system in Eq. (16) results in the following equation for maximal value of transverse ray aberration $\delta y'_{\max}$ in the paraxial image plane, it holds ($h = d_L/2$):

$$\delta y'_{\max} = -\frac{f'}{16F^3}\left[\frac{n_L^3 - 2n_L^2 + 2}{n_L(n_L-1)^2}\right]. \tag{17}$$

The position D_{opt} of an optimal image plane, where the diameter of the circle of confusion $d_{\min}$ is minimal, according to the paraxial image plane can be calculated both with the value of minimal circle of confusion's diameter $d_{\min}$ with formulas stated in Miks and Novak's paper [26] as follows:

$$D_{\text{opt}} = 4F\delta y'_{\max}/3, \qquad d_{\min} = |\delta y'_{\max}|/3. \tag{18}$$

Let one set the minimal diameter of the circle of confusion from Eq. (18) equal to the diameter of the Airy disc ($d_A = 2.44\lambda F$). Afterwards, the minimal f-number for the lens can be calculated using Eqs. (17) and (18), one gets

$$F_{\min} = \frac{1}{3.29}\sqrt[4]{\frac{f'}{\lambda}\left[\frac{n_L^3 - 2n_L^2 + 2}{n_L(n_L-1)^2}\right]}. \tag{19}$$

Using Eq. (19) one can calculate such f-number of the plano–convex lens that the lens can be considered as a physically ideal optical system.

Commercially available focus-tunable lenses [27] are usually placed between two plano-parallel plates. Therefore, an aberration induced by such plates has to be considered for precise analysis as well. Aberrations of a plano-parallel plate are described in detail in [28]. Transversal spherical aberration $\delta y'_p$ induced by thickness d_p of a plate from material with index of refraction n_p is given with formula as follows [28]:

$$\delta y'_p = \frac{d_p(n_p^2 - 1)}{2n_p^3}\left(\frac{h}{f'}\right)^3. \tag{20}$$

As obvious from Eqs. (16) and (20), the aberration induced by the plate's thickness is much smaller than the lens' aberration ($|\delta y'_p| \ll |\delta y'_L|$). Therefore, such aberration given by Eq. (20) can be neglected. The plano-parallel plate behind the lens does not affect the lens' focal length. However, the paraxial image plane will be shifted about the value $\Delta = d_p(n_p - 1)/n_p$ [28].

3. EXAMPLES

A. Example 1

Suppose that the plano–convex focus-tunable lens Optotune EL-10-30 [27] with parameters $d_L = 11$ mm, $t = 2.45$ mm, $\lambda = 587$ nm, and $n_L = 1.559$ is used in the optical system of a two-mirror scanner with the following parameters:

$$\mathbf{s}_0 = -50\mathbf{e}_1 + 30\mathbf{e}_2, \qquad \mathbf{l}_{V2} = -30\mathbf{e}_1 + 30\mathbf{e}_2,$$
$$\mathbf{o}_1 = 30\mathbf{e}_2, \qquad \mathbf{o}_2 = \mathbf{0}, \qquad \mathbf{c}_1 = \mathbf{e}_1,$$
$$\mathbf{c}_2 = \mathbf{e}_3, \mathbf{n}_1(0) = -(\mathbf{e}_1 + \mathbf{e}_2)/\sqrt{(2)},$$
$$\mathbf{n}_2(0) = (\mathbf{e}_1 + \mathbf{e}_2)/\sqrt{(2)}, \qquad \mathbf{q} = \mathbf{e}_1, \qquad \mathrm{d} = 50 \text{ mm},$$

where $\mathbf{e}_1$, $\mathbf{e}_2$, and $\mathbf{e}_3$ denote the basis of the coordinate system.

For the zero rotation of the mirrors the focal length equals to 111.727 mm and the axial thickness of the lens is $t_L = 2.693$ mm. The maximal transverse ray aberration $\delta y'_{\max}$ of the plano–convex lens (see Fig. 2), and for object at the infinity, can be calculated from Eq. (16), and one gets ($h = d_L/2$)

$$\delta y'_{L\,\max} = -\frac{f'}{16}\left(\frac{d_L}{f'}\right)^3\left[\frac{n_L^3 - 2n_L^2 + 2}{n_L(n_L-1)^2}\right] = -0.01270 \text{ mm}. \tag{21}$$

If one considers the axial thickness equal to 2.693 mm, OSLO accurate calculation for a thick lens gives the value $\delta y'_{L(\text{OSLO})} = -0.01271$ mm. Such a result can be obtained with the aforementioned Eq. (7) for general ray tracing in a given optical system as well. Therefore, the approximation used in Eq. (16) gives enough accurate results with an error less than 0.08% for a given situation.

For the aberration induced by a plano-parallel plate with thickness $d_p = 0.5$ mm, which is made by glass BK7 ($n_p = 1.516$), one gets with the use of Eq. (20) ($h = d_L/2$) the following value:

$$\delta y'_{p\,\max} = \frac{d_p(n_p^2 - 1)}{16n_p^3}\left(\frac{d_L}{f'}\right)^3 = 0.000011 \text{ mm}. \tag{22}$$

As obvious, the value of $\delta y'_{p\,\max}$ is much smaller than $\delta y'_{L\,\max}$; therefore, it can be neglected.

If one considers the wavelength equal for example to 587 nm, then the diffraction effect of the system expressed by the radius of the Airy disc [20,21] can be calculated from the following formula: $r_A = 1.22\lambda F = 0.0073$ mm, where λ is the wavelength. As is obvious, the assumption of the paraxial model was justified because the transverse ray aberration is comparable with the radius of the Airy disc and the diameter of the circle of confusion $d_{\min} = 0.0042$ mm, which is smaller than the diameter of the Airy disc ($d_A = 2r_A = 0.0146$ mm).

The beam spots in the detection plane for different angles of rotation φ_1 and φ_2 are shown in Fig. 3. The calculated focal lengths for the beams' focusation in the detection plane respecting the aforementioned situation are shown in Fig. 4.

B. Example 2

Consider now the situation of known position of the point P in the detection plane. The question is, how shall one rotate the mirrors and how the focal length of the focus-tunable lens in the system has to be to fulfil the condition of the focused beam in the point P? Suppose the same parameter of the scanner as in Example 1.

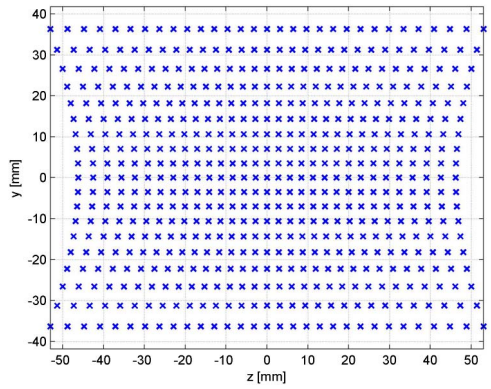


Fig. 3. Beam spots in the detection plane for different angles of rotation $\varphi_1 = \{-30°, -28°, \ldots, +28°, +30°\}$, $\varphi_2 = \{-18°, -16°, \ldots, +16°, +18°\}$.

Substituting the given scanner's parameters into the aforementioned formulas, one gets the following expression of the coordinates of the point P, i.e., the position vector $\mathbf{p} = (p_x, p_y, p_z)$

$$\begin{aligned} p_x &= d, \\ p_y &= d \tan 2\varphi_2, \\ p_z &= -\tan \varphi_1 \left(o_{1y} + \frac{d}{\cos 2\varphi_2} \right), \end{aligned} \tag{23}$$

where o_{1y} is the y-direction component of the vector $\mathbf{o}_1 = (o_{1x}, o_{1y}, o_{1z})$. Afterwards, the angles of rotation can be expressed from the Eqs. (23) as follows:

$$\begin{aligned} \varphi_1 &= -a\tan\left(\frac{p_z}{o_{1y} + \sqrt{p_y^2 + d^2}} \right), \\ \varphi_2 &= \frac{1}{2}\, a\tan\left(\frac{p_y}{d} \right). \end{aligned} \tag{24}$$

The distance between points V_{L2} and P can be then calculated from the formula

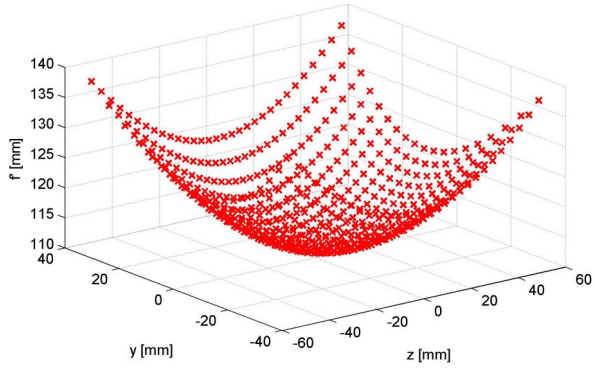


Fig. 4. Focal lengths for the beams' focusation in the detection plane for different angles of rotation $\varphi_1 = \{-30°, -28°, \ldots, +28°, +30°\}$, $\varphi_2 = \{-18°, -16°, \ldots, +16°, +18°\}$.

Table 1. Coordinates of Points in the Detection Plane, Corresponding Angles of Mirrors' Rotations, and Focal Lengths for the Scanner from Example 1

point	p_z [mm]	p_y [mm]	φ_1 [°]	φ_2 [°]	f' [mm]
1	0	0	0.0000	0.0000	+111.727
2	−20	+30	+12.7609	+15.4819	+122.260
3	−20	0	+14.0362	0.0000	+114.186
4	−20	−30	+12.7609	−15.4819	+122.260
5	+20	+30	−12.7609	+15.4819	+122.260
6	+20	0	−14.0362	0.0000	+114.186
7	+20	−30	−12.7609	−15.4819	+122.260

$$s' = -l_{V2x} + \sqrt{d^2 + p_y^2 + \frac{(d^2 + p_y^2)p_z^2}{\left(o_{1y} + \sqrt{p_y^2 + d^2}\right)^2}} + \sqrt{o_{1y}^2 + \frac{l_{V2y}^2 p_z^2}{\left(o_{1y} + \sqrt{p_y^2 + d^2}\right)^2}}, \tag{25}$$

where l_{V2x} and l_{V2y} are components of the vector $\mathbf{l}_{V2} = (l_{V2x}, l_{V2y}, l_{V2z})$.

Equation (12) can be then used for the calculation of the focus-tunable lens' focal length.

In Table 1, there are shown given coordinates of points in the detection plane and corresponding angles on the mirrors' rotations and focal lengths for focusation of the lens in the system. The points in the detection plane are shown in Fig. 5.

C. Example 3

Suppose the uncertainties of the focus-tunable lens' parameters in three possibilities as follows:

(a) $u(t) = u(dL) = 0.1$ mm, $u(nL) = 1e-6$,
(b) $u(t) = u(dL) = 0.3$ mm, $u(nL) = 1e-5$,
(c) $u(t) = u(dL) = 0.5$ mm, $u(nL) = 1e-4$.

For each of the aforementioned cases let the uncertainty $u(s') = 0.01 \cdot s'$, i.e., one percent from the value s'. In Fig. 6, the dependency of uncertainties of focal length on the distance between points V_{L2} and P is shown for the scanner from Example 1 and its focal length focused in the detection plane.

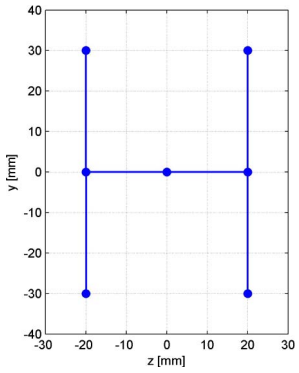


Fig. 5. Given points in the detection plane from Example 1 with parameters of the scanner from Example 1.

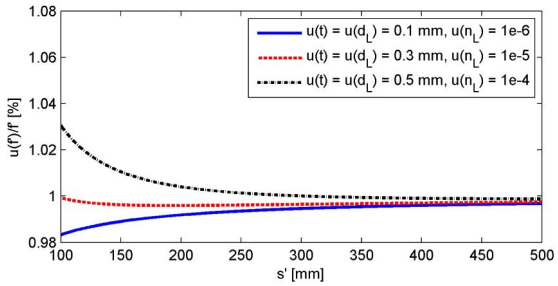


Fig. 6. Dependency of uncertainties of focal length $u(f')$ on the distance between points V_{L2} and P for different cases of uncertainties of the focus-tunable lens' parameters.

4. CONCLUSION

The ray-tracing analysis of the optical system of a two-mirror scanner with a focus-tunable plano–convex lens was presented. The general vector approach was used, which ensures the applicability for different optical systems. The exact formulas for the calculation of the beam spot in the detection plane was presented as a dependency on the angles of the mirrors' rotation. Afterwards, the calculation of focal length of the lens was derived according the condition of focusation in the detection plane. The uncertainty analysis of such focal length is presented as well. In the end of the paper, the authors show the practical examples and usage of derived formulas. For the most used type of two-mirror scanner the formulas for exact calculation of the angles of the mirrors' rotation was presented for given positions of points in the detection plane. The presented formulas can find utilization in many practical applications of optical scanners.

Funding. Czech Technical University in Prague (SGS15/125/OHK1/2T/11).

REFERENCES

1. G. F. Marshall and G. E. Stutz, *Handbook of Optical and Laser Scanning* (CRC/Taylor & Francis, 2011), Vol. **147**.
2. G. Vosselman and H.-G. Maas, *Airborne and Terrestrial Laser Scanning* (Whittles Publishing, 2010).
3. P. Pokorny, "One-mirror and two-mirror three-dimensional optical scanners—position and accuracy of laser beam spot," Appl. Opt. **53**, 2730–2740 (2014).
4. A. Mikš and P. Pokorný, "3D optical scanners," Fine Mech. Opt. **56**, 137–141 (2012).
5. A. Mikš and J. Novák, "Paraxial imaging properties of double conjugate zoom lens system composed of three tunable-focus lenses," Opt. Lasers Eng. **53**, 86–89 (2014).
6. A. Mikš, J. Novák, and P. Novák, "Algebraic and numerical analysis of imaging properties of thin tunable-focus fluidic membrane lenses with parabolic surfaces," Appl. Opt. **52**, 2136–2144 (2013).
7. A. Mikš and J. Novák, "Three-component double conjugate zoom lens system from tunable focus lenses," Appl. Opt. **52**, 862–865 (2013).
8. A. Mikš and J. Novák, "Analysis of three-element zoom lens based on refractive variable-focus lenses," Opt. Express **19**, 23989–23996 (2011).
9. A. Miks, J. Novak, and P. Novak, "Generalized refractive tunable-focus lens and its imaging characteristics," Opt. Express **18**, 9034–9047 (2010).
10. A. Miks and J. Novak, "Analysis of two-element zoom systems based on variable power lenses," Opt. Express **18**, 6797–6810 (2010).
11. A. Mikš and J. Novák, "Third-order aberrations of the thin refractive tunable-focus lens," Opt. Lett. **35**, 1031–1033 (2010).
12. R. J. Pegis and M. Rao, "Analysis and design of plane-mirror systems," Appl. Opt. **2**, 1271–1274 (1963).
13. J. Cohen-Sabban, Y. Cohen-Sabban, and A. Roussel, "Distortion-free 2-D space and surface scanners using light deflectors," Appl. Opt. **22**, 3935–3942 (1983).
14. Y. J. Li and J. Katz, "Laser beam scanning by rotary mirrors. I. Modeling mirror-scanning devices," Appl. Opt. **34**, 6403–6416 (1995).
15. Y. J. Li, "Laser beam scanning by rotary mirrors. II. Conic-section scan patterns," Appl. Opt. **34**, 6417–6430 (1995).
16. Y. J. Li, "Single-mirror beam steering system: analysis and synthesis of high-order conic-section scan patterns," Appl. Opt. **47**, 386–398 (2008).
17. Y. J. Li, "Beam deflection and scanning by two-mirror and two-axis systems of different architectures: a unified approach," Appl. Opt. **47**, 5976–5985 (2008).
18. Y. Friedman and N. Schweitzer, "Classification of stable configurations of plane mirrors," Appl. Opt. **37**, 7229–7234 (1998).
19. R. Shinozaki, O. Sasaki, and T. Suzuki, "Fast scanning method for one-dimensional surface profile measurement by detecting angular deflection of a laser beam," Appl. Opt. **43**, 4157–4163 (2004).
20. A. Mikš, *Applied Optics* (Czech Technical University, 2009).
21. M. Born, E. Wolf, and A. B. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 1999).
22. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review* (Dover, 2000).
23. K. Rektorys, *Survey of Applicable Mathematics* (MIT, 1969).
24. K.-R. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models* (Springer, 1999).
25. L. Mervart and Z. Lukes, *Adjustment Calculus* (Czech Technical University, 2007).
26. A. Miks and J. Novak, "Dependence of camera lens induced radial distortion and circle of confusion on object positron," Opt. Lasers Technol. **44**, 1043–1049 (2012).
27. "Optotune," 2014, www.optotune.com.
28. A. Miks and P. Pokorny, "Analytical expressions for the circle of confusion induced by plane-parallel plate," Opt. Lasers Eng. **50**, 1517–1521 (2012).

applied optics

Diffraction by perfect and imperfect amplitude grating

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
*Corresponding author: petr.pokorny@fsv.cvut.cz

Received 30 June 2020; revised 28 August 2020; accepted 4 September 2020; posted 9 September 2020 (Doc. ID 401606); published 14 October 2020

The paper presents theoretical formulas for calculation of diffraction by perfect infinite and finite amplitude gratings with Fresnel and Fraunhofer approximations. Further, general formulas for diffraction by an imperfect diffraction grating are derived where edges of the grating are described with general harmonic functions. Such a formalism provides enough power to accurately characterize imperfections of diffraction gratings, and it serves as a simple tool for a solution to a diffraction problem.

https://doi.org/10.1364/AO.401606

1. INTRODUCTION

Amplitude diffraction gratings [1–8] are fundamental optical components that are used in many practical applications, for example, during construction or measurement with spectrometers, monochromators, and similar. As the demand for accuracy is continuously growing in industry, diffraction gratings have to be manufactured with very high precision. While imperfections or defects are present, those have to be identified, and one has to know their effect on measurement.

The paper presents theoretical formulas for calculation of diffraction by perfect infinite and finite amplitude gratings with Fresnel and Fraunhofer approximations [1,9–13]. Further, general formulas for diffraction by an imperfect diffraction grating are derived where edges of the grating are described with general harmonic functions. Such a formalism provides enough power to accurately characterize imperfections of diffraction gratings, and it serves as a simple tool for a solution to a diffraction problem.

Many studies of diffraction by gratings have been presented [2–8] using various approaches. Similarly, several authors published different approaches to solve diffraction problems by imperfect circular apertures, slits, etc. [14–23]. The diffraction by gratings with rough edges was studied in [24–29]. Nevertheless, the approach presented in this paper is unique, and it offers the very simple possibility to analyze diffraction by imperfect gratings without any expensive software tools.

2. DIFFRACTION BY APERTURE

It is known from the diffraction theory that the amplitude of the field in arbitrary point P of the space behind an aperture (or a series of apertures) can be calculated with the formula [1,9–13]

$$U(P) = -\frac{i}{\lambda}\iint_A U(M)\frac{\exp(ikr_{PM})}{r_{PM}}\cos(\mathbf{n}, \mathbf{r}_{PM})\mathrm{d}A, \quad (1)$$

where the integration is performed across the surface A of the aperture. Equation (1) is the so-called Sommerfeld solution of diffraction (Sommerfeld diffraction integral or Rayleigh–Sommerfeld solution of the second kind). This formula enables us to calculate the field properties $U(P)$ in an arbitrary point P of an area bounded by surface S, if one knows the field $U(M)$ on such a surface, while r_{PM} is the distance between points P and M, and $\cos(\mathbf{n}, \mathbf{r}_{PM})$ denotes the cosine of the angle between inner normal vector $\mathbf{n} = (0, 0, 1)$ to the surface S with a direction of the vector $\mathbf{r}_{PM}$ (see Fig. 1). Intensity of the field $I(P)$ in point P is then given as follows:

$$I(P) = |U(P)|^2. \quad (2)$$

It is usually possible to use several simplifications during calculations of practical examples, which makes numerical evaluation easier. Distance r_{PM} between points P and M is given with the formula (according Fig. 1, $z_M = 0$)

$$r_{PM} = \sqrt{(x_P - x_M)^2 + (y_P - y_M)^2 + z_P^2}$$
$$\approx z_P + \frac{(x_P - x_M)^2 + (y_P - y_M)^2}{2z_P}, \quad (3)$$

where the Taylor series neglecting higher orders was used. Moreover, if one supposes only small diffraction angles, i.e., $\cos(\mathbf{n}, \mathbf{r}_{PM}) \approx 1$, then Eq. (1) can be approximated with satisfactory accuracy as follows:

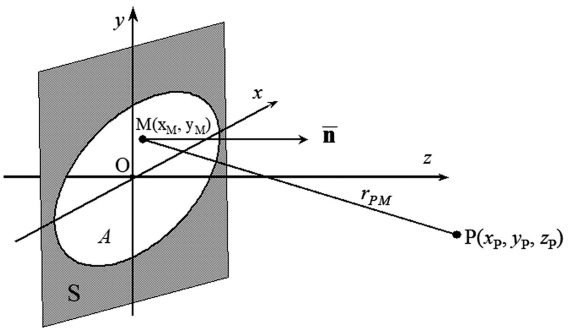


Fig. 1. Diffraction by aperture.

$$U(P) = C \iint_A U(M) \times \exp\left\{\frac{ik}{2z_P}\left[(x_P - x_M)^2 + (y_P - y_M)^2\right]\right\} \mathrm{d}x_M \mathrm{d}y_M,$$

$$C = -\frac{i}{\lambda}\frac{\exp(ikz_P)}{z_P}, \tag{4}$$

where r_{PM} is substituted by Eq. (3) in the exponent of Eq. (1), and the denominator is equal to $r_{PM} = z_P$. Equation (4) represents the so-called Fresnel approximation of the Sommerfeld diffraction integral.

In the case of diffraction by an aperture of characteristic dimensions much smaller than distance z_P, one can suppose

$$\exp\left[\frac{ik}{2z_P}(x_M^2 + y_M^2)\right] \approx 1. \tag{5}$$

The same approximation would hold for coordinates x_P and y_P, which states that the diffraction is studied near the optical axis.

In the case of diffraction of a convergent spherical (or approximately spherical) wave with the center in point P or in its close distance, it holds:

$$U(M) = T(M)\exp\left[-\frac{ik}{2z_P}(x_M^2 + y_M^2)\right], \tag{6}$$

where function $T(M)$ characterizes properties of these waves in the plane of the aperture. Equations (5) and (6) give the so-called Franhofer approximation of diffraction integral. If one denotes

$$u = x_P/z_P \quad \text{and} \quad v = y_P/z_P, \tag{7}$$

the field in point P can be described by formulas

$$U(P) = C \iint_A U(M)\exp[-ik(ux_M + vy_M)]\mathrm{d}x_M \mathrm{d}y_M, \tag{8}$$

$$U(P) = C \iint_A T(M)\exp[-ik(ux_M + vy_M)]\mathrm{d}x_M \mathrm{d}y_M. \tag{9}$$

One can see that considering the Franuhofer approximation of the diffraction integral, the field $U(P)$ is proportional to Fourier transformation of the field in the plane of the aperture.

3. DIFFRACTION BY AMPLITUDE GRATING

A. Diffraction by Perfect Amplitude Grating

Consider now an amplitude diffraction grating [9–13] (e.g., Ronchi grating), whose scheme is shown in Fig. 2, where p denotes the period and b the width of the transparent part of the grating. Afterwards, such a grating can be characterized with the formula ($i = \sqrt{-1}$)

$$f(\xi) = A\sum_{n=-\infty}^{\infty}\mathrm{rect}\left(\frac{\xi - np}{b}\right) = \sum_{n=-\infty}^{\infty} c_n \exp\left(i\frac{2\pi}{p}n\xi\right), \tag{10}$$

where [11]

$$\mathrm{rect}(x) = \begin{cases} 1, & |x| < 1/2, \\ 1/2, & |x| = 1/2, \\ 0, & |x| > 1/2, \end{cases}$$

A is the amplitude transitivity of the grating, and Fourier coefficients c_n are given as follows [9–12]:

$$c_n = A\frac{b}{p}\mathrm{sinc}\left(\frac{nb}{p}\right), \tag{11}$$

where $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. Further, consider that the grating is opaque outside the region (i.e., the field outside the mentioned region does not affect the diffraction).

Without loss of generality, one can suppose $A = 1$ for the next purposes. The goal is to calculate the state of the field $U(x, y, z)$ in distance z behind the grating. Using the Fresnel approximation of the Sommerfeld diffraction integral [1,9,13] for a complex amplitude of wavefield in the point of Cartesian coordinates (x, y, z), one gets

$$U(x, y, z) = C\int_{-d}^{d}\int_{-c}^{c} U(x_M, y_M, 0) \times \sum_{n=-\infty}^{\infty} c_n \exp\left(i\frac{2\pi}{p}nx_M\right) \times \exp\left[ik\frac{(x - x_M)^2 + (y - y_M)^2}{2z}\right]\mathrm{d}x_M \mathrm{d}y_M, \tag{12}$$

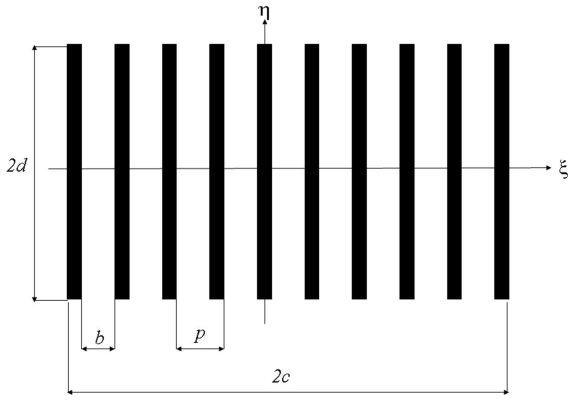


Fig. 2. Amplitude grating.

where $U(x_M, y_M, 0)$ denotes the complex amplitude of the field impinging on the amplitude grating, $k = 2\pi/\lambda$ is the wavenumber, λ is the wavelength, $2d$ is the height, $2c$ is the width of the grating, and (x_M, y_M) are coordinates in the plane of the grating. Suppose now that the impinging field is a plane wave. Therefore, one can set without loss of generality $U(x_M, y_M, 0) = 1$. The derivation of the field state in the plane behind the amplitude grating is presented in a previous paper [30]. It can be shown that the following formula holds [30]:

$$U(x, y, z) = -\frac{i}{\pi}\exp(ikz)\sum_{n=-\infty}^{\infty} U_n \exp\left[i\left(\gamma - \frac{\beta^2}{4\alpha}\right)\right], \tag{13}$$

where

$$U_n = c_n\left[F(\tau_1) - F(\tau_2)\right]\left[F(\theta_1) - F(\theta_2)\right], \tag{14}$$

while

$$\tau_1 = (x-c)\sqrt{\alpha} + \frac{\beta}{2\sqrt{\alpha}}, \quad \tau_2 = (x+c)\sqrt{\alpha} + \frac{\beta}{2\sqrt{\alpha}},$$
$$\theta_1 = (y-d)\sqrt{\alpha}, \quad \theta_2 = (y+d)\sqrt{\alpha},$$
$$\alpha = \pi/(\lambda z), \quad \beta = -2\pi n/p, \quad \gamma = 2\pi nx/p, \tag{15}$$

and $F(\tau)$ denotes

$$F(\tau) = \int_0^{\tau} \exp(i\xi^2)\mathrm{d}\xi = \sqrt{\frac{\pi}{2}}\left[C(\tau\sqrt{2/\pi}) + iS(\tau\sqrt{2/\pi})\right], \tag{16}$$

where $C(\tau\sqrt{2/\pi})$ and $S(\tau\sqrt{2/\pi})$ are Fresnel integrals [10–13] defined by formulas

$$C(x) = \int_0^x \cos\left(\frac{\pi}{2}t^2\right)\mathrm{d}t, \quad S(x) = \int_0^x \sin\left(\frac{\pi}{2}t^2\right)\mathrm{d}t. \tag{17}$$

The field intensity is then given according to Eq. (2):

$$I(x, y, z) = \frac{1}{\pi^2}\left|\sum_{n=-\infty}^{\infty} U_n \exp\left[i\left(\gamma - \frac{\beta^2}{4\alpha}\right)\right]\right|^2. \tag{18}$$

The addition in Eq. (18) demonstrates that the image of intensity will be identical to its template for distances $z = Nz_T = N(2p^2/\lambda)$, where N is a constant. The distance $z_T = 2p^2/\lambda$ is the so-called Talbot distance [30–33].

For an infinite diffraction grating ($c \to \infty$ and $d \to \infty$), one can apply the known properties of limit cases of Fresnel integrals [10–13]. Afterwards, the complex amplitude and the intensity can be calculated with formulas [30]

$$U(x, y, z) = -i\exp\left[i\left(kz + \frac{\pi}{2}\right)\right]\sum_{n=-\infty}^{\infty} c_n \exp\left[i\left(\gamma - \frac{\beta^2}{4\alpha}\right)\right], \tag{19}$$

$$I(x, y, z) = \left|\sum_{n=-\infty}^{\infty} c_n \exp\left[i\left(\gamma - \frac{\beta^2}{4\alpha}\right)\right]\right|^2. \tag{20}$$

Consider now a diffraction problem by a finite grating. One can calculate the state of the field $U(x, y, z)$ in distance z behind the plane of the grating, which is described by Eq. (10), with Eq. (8) (Fraunhofer approximation, $U(M) = 1$), and it holds:

$$U(x, y, z) = C\int_{-d}^{d}\int_{-c}^{c}\sum_{n=-\infty}^{\infty} c_n \exp\left(i\frac{2\pi}{p}nx_M\right) \times \exp[-ik(ux_M + vy_M)]\mathrm{d}x_M\mathrm{d}y_M. \tag{21}$$

Equation (21) describes diffraction by the grating that is impinged on by a convergent spherical wave with radius z. Considering further the formula $\int_{-c}^{c}\exp(i\psi\xi)\mathrm{d}\xi = 2\sin\psi c/\psi$, one can modify Eq. (21) to express the state of the field $U(x, y, z)$ after simplification, and it holds:

$$U(x, y, z) = -4i\exp(ikz)\frac{\sin(kvd)}{\lambda kvz}\sum_{n=-\infty}^{\infty} c_n S_n,$$
$$S_n = \sin(\psi c)/\psi, \quad \psi = \frac{2\pi}{p}n - ku. \tag{22}$$

And it follows for the intensity on the plane in the center of the convergent spherical wave (which is distanced by z from the grating):

$$I(x, y, z) = 16\left[\frac{\sin(kvd)}{\lambda kvz}\sum_{n=-\infty}^{\infty} c_n S_n\right]^2. \tag{23}$$

The limit case for $u = v = 0$ then is represented by the formula

$$I(0, 0, z) = 4\left(\frac{pd}{\lambda z\pi}\right)^2\left[\sum_{n=-\infty}^{\infty} c_n \frac{\sin(2\pi nc/p)}{n}\right]^2. \tag{24}$$

B. Diffraction by Imperfect Amplitude Grating

Suppose now the case of an imperfect amplitude grating. Let any individual slit of the grating be described as an individual diffraction by the aperture. The field behind the imperfect grating then will be characterized as a superposition of contributions of those individual sub-apertures. Generally, one can consider the following formula to describe the grating, and it holds:

$$f(\xi, \eta) = A\sum_{n=-\infty}^{\infty}\mathrm{rect}\left[\frac{\xi - \xi_{n,0}(\eta)}{b_n(\eta)}\right], \tag{25}$$

where A is the constant, $\xi_{n,0}(\eta)$ denotes the function of the center line of the nth slit (sub-aperture), and $b_n(\eta)$ is the function of its width. Figure 3 distinctly shows a scheme of the considered grating.

Further, one can suppose

$$\xi_{n,0}(\eta) = \frac{1}{2}[\xi_{n,d}(\eta) + \xi_{n,b}(\eta)],$$
$$b_n(\eta) = \xi_{n,b}(\eta) - \xi_{n,d}(\eta), \tag{26}$$

where $\xi_{n,d}(\eta)$ and $\xi_{n,b}(\eta)$ are functions of individual edges of the nth grating's sub-aperture (slit). The edge-function can be described by various types of prescriptions (approximations),

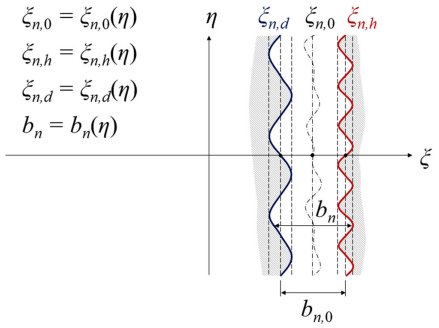


Fig. 3. Scheme of n the slit (sub-aperture) of imperfect amplitude grating.

e.g., power series, Legendre series, Fourier trigonometric series, or similar [34].

The aforementioned general formalism covers various types of gratings, which can have every slit of a different shape [described by individual edge functions $\xi_{n,d}(\eta)$ and $\xi_{n,h}(\eta)$]. In general, central lines of slits [Eq. (26)] of the grating do not have to be parallel, as one can vary individual shapes of the slits' edges.

For completeness, let an existence condition be defined. The grating can be physically possible (manufactured) if the following condition holds: $\max[\xi_{n-1,h}(\eta)] < \min[\xi_{n,d}(\eta)]$ for every η.

Suppose now that the grating has N slits; therefore, $n \in [1, N]$. Afterwards, the complex amplitude $U(x, y, z)$ and the intensity $I(x, y, z)$ in plane z behind the grating can be expressed with the superposition principle as follows:

$$U(x, y, z) = \sum_{n=1}^{N} U_{S,n}, \qquad I(x, y, z) = \left| \sum_{n=1}^{N} U_{S,n} \right|^2, \tag{27}$$

where $U_{S,n}$ is the contribution of the nth sub-aperture of the grating. In the case of the Fraunhofer approximation of the Sommerfeld diffraction integral, Eq. (8), one can write the following formula for the grating of height $2d$:

$$U_{S,n} = CK \int_{-d}^{d} \int_{\xi_{n,d}(y_M)}^{\xi_{n,h}(y_M)} \exp\left[-ik(ux_M + vy_M)\right] \mathrm{d}x_M \mathrm{d}y_M, \tag{28}$$

where $K = \text{const.}$, $u = x/z$, $v = y/z$, $C = -\frac{i}{\lambda}\frac{\exp(ikz)}{z}$, and $k = \frac{2\pi}{\lambda}$.

Consider now a specific case where the grating's edges are described by harmonic functions as follows:

$$\xi_{n,d}(\eta) = \xi_{n,0}(0) - \frac{b_{n,0}}{2} - A_{n,1} \sin(\Omega_{n,1}\eta + \varphi_{n,1}),$$

$$\xi_{n,h}(\eta) = \xi_{n,0}(0) + \frac{b_{n,0}}{2} + A_{n,2} \sin(\Omega_{n,2}\eta + \varphi_{n,2}), \tag{29}$$

where $\xi_{n,0}(0)$ and $b_{n,0}$ denote, respectively, the position of the center line and width of the sub-aperture for coordinate $\eta = 0$; $A_{n,1}$, $A_{n,2}$ are amplitudes of edge-functions; $\Omega_{n,1}$, $\Omega_{n,2}$ are angular frequencies of edge-functions; and $\varphi_{n,1}$, $\varphi_{n,2}$ are their initial phase shifts. Setting next $U_{S,n} = CKU_{S,n,xy}$, one then gets the following formula after appropriate simplification:

$$\begin{aligned} U_{S,n,xy} = i D_{S,n,2} \int_{-d}^{d} &\left\{\exp(-ikvy_M) \right. \\ &\left. \times \exp[-i\beta_{n,2} \sin(\Omega_{n,2} y_M + \varphi_{n,2})]\right\} \mathrm{d}y_M \\ -i D_{S,n,1} \int_{-d}^{d} &\left\{\exp(-ikvy_M) \right. \\ &\left. \times \exp[i\beta_{n,1} \sin(\Omega_{n,1} y_M + \varphi_{n,1})]\right\} \mathrm{d}y_M, \end{aligned} \tag{30}$$

where

$$\beta_{n,2} = A_{n,2}ku, \quad \beta_{n,1} = A_{n,1}ku,$$

$$D_{S,n,1} = \frac{1}{ku} \exp\left[-iku\left(\xi_{n,0}(0) - \frac{b_{n,0}}{2}\right)\right],$$

$$D_{S,n,2} = \frac{1}{ku} \exp\left[-iku\left(\xi_{n,0}(0) + \frac{b_{n,0}}{2}\right)\right]. \tag{31}$$

It holds for the limit case $u = 0$:

$$\begin{aligned} U_{S,n,xy}\big|_{u=0} = \int_{-d}^{d} &\exp(-ikvy_M)[b_0 \\ &+ A_{n,1} \sin(\Omega_{n,1} y_M + \varphi_{n,1}) \\ &+ A_{n,2} \sin(\Omega_{n,2} y_M + \varphi_{n,2})]\mathrm{d}y_M. \end{aligned} \tag{32}$$

For $v = 0$, one gets

$$\begin{aligned} U_{S,n,xy}\big|_{v=0} = &\, i D_{S,n,2} \int_{-d}^{d} \exp[-i\beta_{n,2} \sin(\Omega_{n,2} y_M + \varphi_{n,2})]\mathrm{d}y_M \\ &-i D_{S,n,1} \int_{-d}^{d} \exp[i\beta_{n,1} \sin(\Omega_{n,1} y_M + \varphi_{n,1})]\, \mathrm{d}y_M, \end{aligned} \tag{33}$$

and for $u = v = 0$,

$$\begin{aligned} U_{S,n,xy}\big|_{u=v=0} = &\, 2b_{n,0}d + \frac{2A_{n,2}}{\Omega_{n,2}} \sin(\Omega_{n,2} d) \sin\varphi_{n,2} \\ &+ \frac{2A_{n,1}}{\Omega_{n,1}} \sin(\Omega_{n,1} d) \sin\varphi_{n,1}. \end{aligned} \tag{34}$$

Using Eqs. (28)–(34) in (27), one can express the state of the field in the plane behind the grating. Afterwards, one obtains for intensity

$$I(x, y, z) = \frac{K^2}{\lambda^2 z^2} \left| \sum_{n=1}^{N} U_{S,n,xy} \right|^2. \tag{35}$$

In the limit case $I(0, 0, z)$, the following formula holds:

$$\begin{aligned} I(0, 0, z) = 4\frac{K^2}{\lambda^2 z^2} \left| \sum_{n=1}^{N} \left[b_{n,0}d + \frac{A_{n,2}}{\Omega_{n,2}} \sin(\Omega_{n,2} d) \sin\varphi_{n,2} \right.\right. \\ \left.\left. + \frac{A_{n,1}}{\Omega_{n,1}} \sin(\Omega_{n,1} d) \sin\varphi_{n,1} \right] \right|^2. \end{aligned} \tag{36}$$

If one considers small amplitudes $A_{n,1}$ and $A_{n,2}$, then the function $U_{S,n,xy}$ in Eq. (30) can be simplified as follows:

$$\begin{aligned} U_{S,n,xy} \approx\, & i D_{S,n,2} \int_{-d}^{d} \exp(-ikvy_M) \\ & \times [1 - i\beta_{n,2} \sin(\Omega_{n,2} y_M + \varphi_{n,2})]\, dy_M \\ & - i D_{S,n,1} \int_{-d}^{d} \exp(-ikvy_M) \\ & \times [1 + i\beta_{n,1} \sin(\Omega_{n,1} y_M + \varphi_{n,1})] dy_M \\ = & \, i(D_{S,n,2} - D_{S,n,1}) \int_{-d}^{d} \exp(-ikvy_M)\, dy_M \\ & + \beta_{n,2} D_{S,n,2} \int_{-d}^{d} \exp(-ikvy_M) \sin(\Omega_{n,2} y_M + \varphi_{n,2})\, dy_M \\ & + \beta_{n,1} D_{S,n,1} \int_{-d}^{d} \exp(-ikvy_M) \sin(\Omega_{n,1} y_M + \varphi_{n,1})\, dy_M. \end{aligned} \tag{37}$$

Further, one gets after modification

$$\begin{aligned} U_{S,n,xy} \approx\, & i(D_{S,n,2} - D_{S,n,1})\, Q \\ & + \beta_{n,1} D_{S,n,1}\, Q_{n,1} \\ & + \beta_{n,2} D_{S,n,2}\, Q_{n,2}, \end{aligned} \tag{38}$$

where

$$\begin{aligned} i(D_{S,n,2} - D_{S,n,1})\, Q = & \frac{4}{k^2 uv} \sin(kvd) \sin\left(ku \frac{b_{n,0}}{2}\right) \\ & \times [\cos(ku\xi_{n,0}(0)) - i \sin(ku\xi_{n,0}(0))], \end{aligned} \tag{39}$$

$$\begin{aligned} Q_{n,1} = & \frac{2}{\Omega_{n,1}^2 - k^2 v^2} \left[\Omega_{n,1} \cos(kvd) \sin(d\Omega_{n,1}) \right. \\ & \left. - kv \sin(kvd) \cos(d\Omega_{n,1})\right] \sin\varphi_{n,1} \\ & + i \frac{2}{\Omega_{n,1}^2 - k^2 v^2} \left[\Omega_{n,1} \sin(kvd) \cos(d\Omega_{n,1}) \right. \\ & \left. - kv \cos(kvd) \sin(d\Omega_{n,1})\right] \cos\varphi_{n,1}, \end{aligned} \tag{40}$$

$$\begin{aligned} Q_{n,2} = & \frac{2}{\Omega_{n,2}^2 - k^2 v^2} \left[\Omega_{n,2} \cos(kvd) \sin(d\Omega_{n,2}) \right. \\ & \left. - kv \sin(kvd) \cos(d\Omega_{n,2})\right] \sin\varphi_{n,2} \\ & + i \frac{2}{\Omega_{n,2}^2 - k^2 v^2} \left[\Omega_{n,2} \sin(kvd) \cos(d\Omega_{n,2}) \right. \\ & \left. - kv \cos(kvd) \sin(d\Omega_{n,2})\right] \cos\varphi_{n,2}. \end{aligned} \tag{41}$$

Substituting Eqs. (39)–(41) into Eq. (38) can be used for approximate calculation of the state of the field if the amplitudes of the grating's sub-apertures are small.

While deriving Eq. (37), the first two orders of Taylor series $\exp(\pm iX) \approx 1 \pm iX - X^2/2 + \ldots$ were used. Maximal relative error $\varepsilon_{rel.}$ of such an approximation of integrands in Eq. (30) can be estimated from properties of orthogonal series expansion, and it holds: $\varepsilon_{rel.} \approx -X^2/[2(1 \pm iX)]$. Applying in Eq. (30) and Eq. (37), one gets the following formula for the estimation of the maximal relative error $\delta U_{S,n,xy}$:

$$\begin{aligned} \delta U_{S,n,xy} \approx & -\frac{1}{2U_{S,n,xy}} \left[i D_{S,n,2} \int_{-d}^{d} \exp(-ikvy_M) \right. \\ & \times \beta_{n,2}^2 \sin^2(\Omega_{n,2} y_M + \varphi_{n,2})\, dy_M \\ & + i D_{S,n,1} \int_{-d}^{d} \exp(-ikvy_M) \\ & \left. \times \beta_{n,1}^2 \sin^2(\Omega_{n,1} y_M + \varphi_{n,1})\, dy_M \right]. \end{aligned} \tag{42}$$

Equation (42) can simply be used for an analysis of the accuracy of field-state calculation with Eq. (38).

Let one consider a specific case of the imperfect grating that has the same shape of all the slits' edges, i.e., integrals in Eq. (30) will be constants. Suppose now Eq. (30) for Fraunhofer diffraction by a grating in the following form:

$$U_{S,n,xy} = i(D_{S,n,2} G_2 - D_{S,n,1} G_1). \tag{43}$$

Further, let one assume

$$\begin{aligned} D_{S,n,1} &= B_1 \exp[-iku\xi_{n,0}(0)], \\ D_{S,n,2} &= B_2 \exp[-iku\xi_{n,0}(0)], \end{aligned} \tag{44}$$

where

$$B_1 = \frac{1}{ku} \exp\left(i \frac{kub_{n,0}}{2}\right),\ B_2 = \frac{1}{ku} \exp\left(-i \frac{kub_{n,0}}{2}\right). \tag{45}$$

Afterwards, one gets for Eq. (30)

$$U_{S,n,xy} = i(B_2 G_2 - B_1 G_1) \exp[-iku\xi_{n,0}(0)]. \tag{46}$$

If p denotes the grating's period, then it holds:

$$\xi_{n+1,0}(0) = \xi_{n,0}(0) + p. \tag{47}$$

After summing contributions from all slits of the grating, one gets

$$\begin{aligned} \sum_{n=1}^{N} U_{S,n,xy} &= i(B_2 G_2 - B_1 G_1) \sum_{n=1}^{N} \exp[-iku\xi_{n,0}(0)] \\ &= i(B_2 G_2 - B_1 G_1) \exp[ig(N-1)] \frac{\sin(Ng)}{\sin g}, \end{aligned} \tag{48}$$

where $g = kup/2$.

In a specific case of the same shape for both slits' edges, i.e., $A_{n,1} = A_{n,2}$, $\Omega_{n,1} = \Omega_{n,2}$, $\varphi_{n,1} = \varphi_{n,2}$, and $b_{n,0} = b_0$ is constant, one gets

$$U_S = \sum_{n=1}^{N} U_{S,n,xy}$$

$$= b_0 \exp[ig(N-1)]\, G_1 \left[\frac{\sin(h)}{h}\right]\left[\frac{\sin(Ng)}{\sin g}\right], \quad (49)$$

where $h = kub_0/2$.

Further, it holds for a perfect grating ($A_{n,1} = A_{n,2} = 0$): $G_1 = G_0$, where

$$G_0 = 2d\left[\frac{\sin(dkv)}{dkv}\right]. \quad (50)$$

Therefore, one gets

$$U_{S,0} = b_0 \exp[ig(N-1)]\, G_0 \left[\frac{\sin(h)}{h}\right]\left[\frac{\sin(Ng)}{\sin g}\right]. \quad (51)$$

Afterwards, the following formula for the difference between amplitudes of imperfect and perfect gratings holds:

$$U_S - U_{S,0} = (G_1 - G_0) b_{n,0}$$

$$\times \exp[ig(N-1)]\left[\frac{\sin(h)}{h}\right]\left[\frac{\sin(Ng)}{\sin g}\right]. \quad (52)$$

The difference in intensities between perfect and imperfect gratings with same-shaped edges is then given with the formula ($g = kup/2, h = kub_0/2$)

$$\Delta I = I_0 - I = [G_0^2 - |G_1|^2] b_0^2 \left[\frac{\sin(h)}{h}\right]^2 \left[\frac{\sin(Ng)}{\sin g}\right]^2. \quad (53)$$

Finally, the relative difference in intensities can be expressed as follows:

$$\frac{\Delta I}{I_0} = \frac{G_0^2 - |G_1|^2}{G_0^2} = 1 - \left(\frac{|G_1|}{G_0}\right)^2. \quad (54)$$

4. EXAMPLES

A. Example 1–Diffraction by Perfect Amplitude Grating

Figure 4 shows the intersection of intensity of a diffraction pattern calculated with Eqs. (20) and (18) (Fresnel approximation) for infinite and finite perfect diffraction gratings with parameters $p = 0.5$ mm, $b = 0.25$ mm, $\lambda = 633$ nm, $c = d = 5$ mm, $A = 1$, $y = 0$ mm, and $N_{\max} = 500$ in Talbot distance [30–32] $z = z_T = 2p^2/\lambda = 789.8894$ mm. Figure 5 then presents an intensity diffraction pattern where the convergent spherical wave with radius of curvature z impinges on the grating, calculated by Eq. (23) (Fraunhofer diffraction). The horizontal axis in Fig. 5 is expressed in angular coordinate ω, and it holds: $\tan\omega = \frac{x}{z}$.

B. Example 2–Diffraction by Imperfect Amplitude Grating with Irregular Grating Period

Suppose now the situation of an imperfect amplitude grating, described by Eq. (25) and the following parameters:

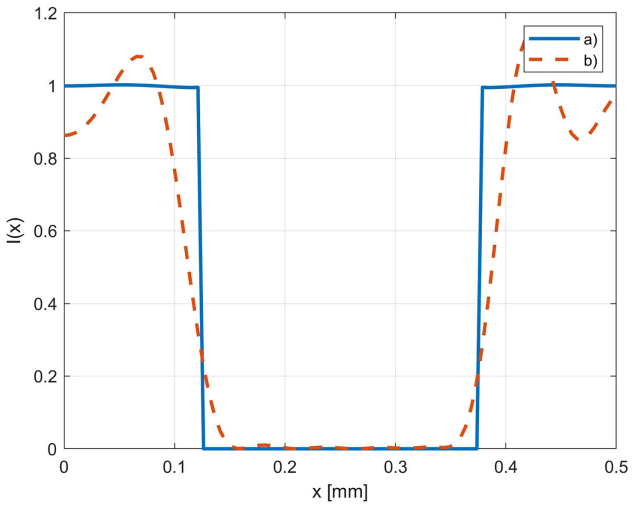


Fig. 4. Intersection of diffraction pattern intensity formed by perfect a) infinite and b) finite amplitude gratings with parameters: $p = 0.5$ mm, $b = 0.25$ mm, $\lambda = 633$ nm, $c = d = 5$ mm, $A = 1$, $y = 0$ mm, $z = z_T = 789.8894$ mm, $N_{\max} = 500$.

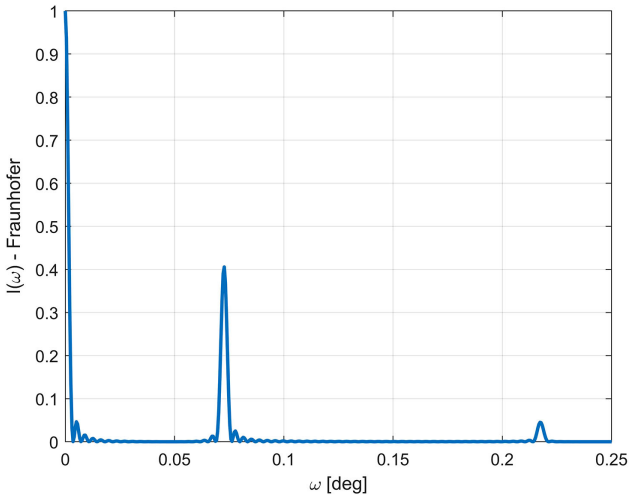


Fig. 5. Intersection of diffraction pattern intensity (Fraunhofer diffraction) formed by perfect finite amplitude grating with parameters: $p = 0.5$ mm, $b = 0.25$ mm, $\lambda = 633$ nm, $c = d = 5$ mm, $A = 1$, $y = 0$ mm, $z = z_T = 789.8894$ mm, $N_{\max} = 500$.

$$\xi_{n,0}(0) = n_p(1 + A_p)p_0, \quad n_p \in [-N_p, N_p],$$

$$b_{n,0} = b_0, \quad A_{n,1} = A_1, \quad A_{n,2} = A_2,$$

$$\Omega_{n,1} = \Omega_1, \quad \Omega_{n,2} = \Omega_2, \quad \varphi_{n,1} = \varphi_1, \quad \varphi_{n,2} = \varphi_2, \quad (55)$$

where $A_p = 0.02$, $p_0 = 0.5$ mm, $N_p = 9$, $b_0 = 0.25$ mm, and $A_1 = A_2 = \Omega_1 = \Omega_2 = \varphi_1 = \varphi_2 = 0$. Consider dimensions of the grating to be $c = d = 5$ mm, and the parameters of the impinging field can be set without loss of generality as follows: $A = K = 1$, $\lambda = 633$ nm, $z = 150$ mm. Therefore, this grating has sharp linear edges and a variable period, which linearly changes with the distance from its center.

Figure 6 shows an intersection of normalized intensities in plane $y = 0$ for Fraunhofer diffraction by a perfect grating

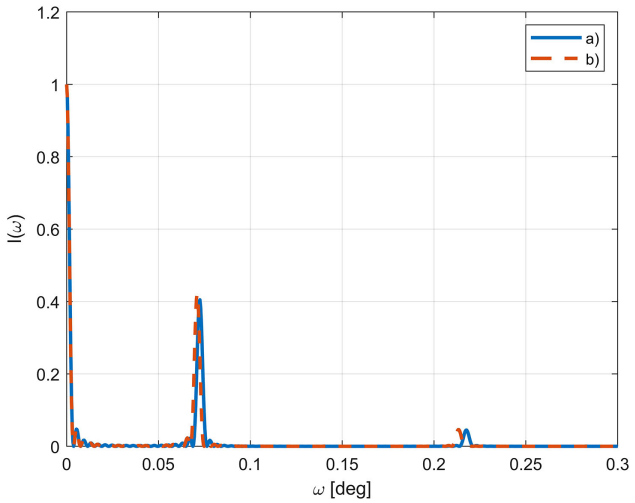


Fig. 6. Intersection of normalized intensities in plane $y = 0$ for (a) perfect and (b) imperfect gratings with irregular periods.

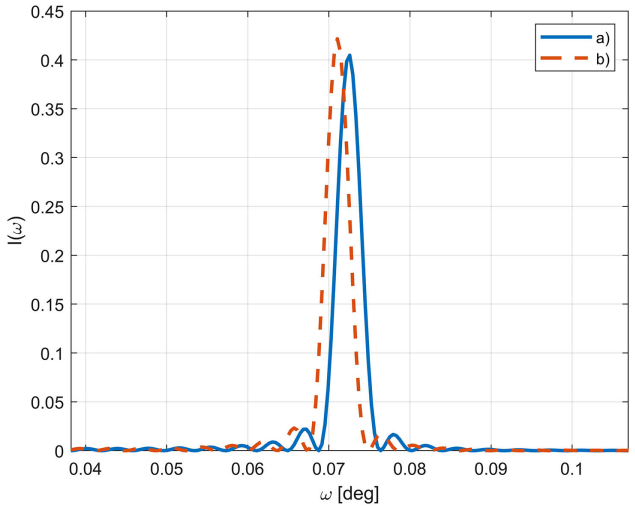


Fig. 7. Detail of intersection of normalized intensities in plane $y = 0$ for (a) perfect and (b) imperfect gratings with irregular periods for the peak of the first diffraction order.

that has constant period p_0, and for the imperfect grating with the aforementioned parameters. Figure 7 presents the same situation in detail around the peak of the first diffraction order. Figure 8 then presents the difference in those intensities. One can see that irregularity of the grating affects the positions of diffraction maxima. Further, the change in intensities in individual peaks is visible as well.

C. Example 3–Diffraction by Imperfect Amplitude Grating with the Same Shape of Slits

Suppose now the imperfect amplitude grating that has slits of all the same shape, and the edges of those slits are identical. Let the parameters of the grating be

$$\xi_{n,0}(0) = n_p p, \quad n_p \in [-N_p, N_p], \quad b_{n,0} = b,$$

$$A_{n,1} = A_{n,2} = A_0, \quad \Omega_{n,1} = \Omega_{n,2} = \Omega, \quad \varphi_{n,1} = \varphi_{n,2} = \varphi, \tag{56}$$

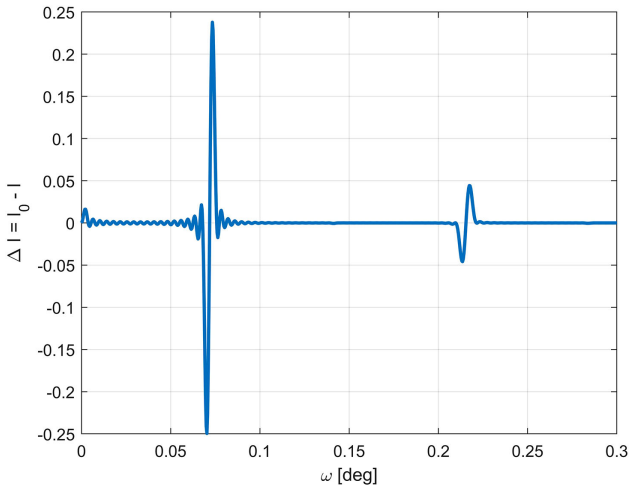


Fig. 8. Difference in intensities of perfect (I_0) and imperfect (I) gratings with irregular periods in plane $y = 0$.

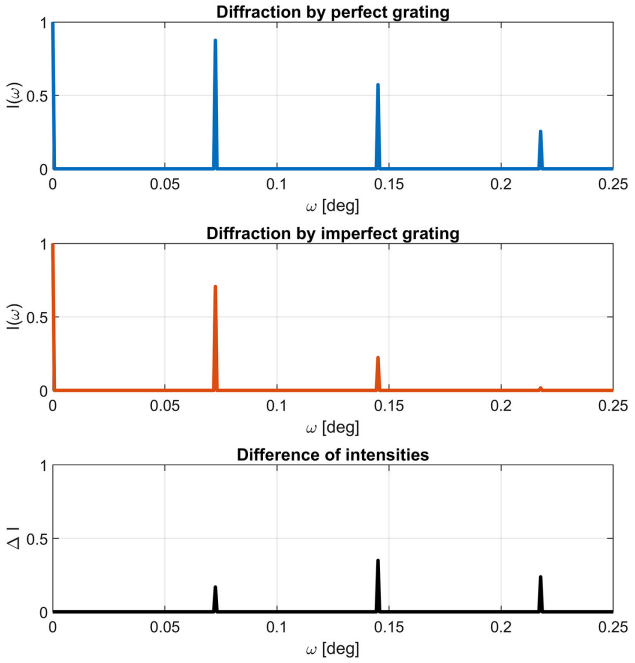


Fig. 9. Diffraction by perfect and imperfect gratings, which have slits of all the same shape, and the edges of those slits are identical.

where $p = 0.5$ mm, $N_p = 10$, $b = 0.1$ mm, $A_0 = 0.05$ mm, and $\Omega = 1$, $\varphi = 0$. Consider dimensions of the grating to be $c = d = 5$ mm, and the parameters of the impinging field can be set without loss of generality as follows: $A = K = 1$, $\lambda = 633$ nm, $z = 789.8894$ mm.

Figure 9 shows a comparison of intensity profiles for diffraction by a perfect grating (with zero amplitude of edges) and the aforementioned imperfect grating. Results were calculated using Eqs. (49), (51), and (53). Identical results can be calculated with Eq. (35). For such an imperfect grating, the intensity loss in diffraction orders is significant compared to the perfect grating.

5. CONCLUSION

The paper presented theoretical formulas for calculation of diffraction by perfect infinite and finite gratings with Fresnel and Fraunhoffer approximations of diffraction problems. Further, the unique general formulas for an imperfect grating were derived as well, where the shape edge-functions of individual slits (sub-apertures) were described with general harmonics. This approach offers precise characterization of imperfections of amplitude gratings. Moreover, it gives a simple solution to a diffraction problem. Therefore, the results presented in the paper contribute to possibilities of analysis of imaging (diffraction) by diffraction gratings, and presents outputs that can find wide practical applications.

Funding. Czech Technical University in Prague (SGS20/093/OHK1/2T/11).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 1999).
2. H. A. Rowland, "Imperfections in diffraction grating spectra," Philos. Mag. Ser. **35**, 397–419 (1893).
3. C. M. Sparrow, "Theory of imperfect gratings," Astrophys. J. **49**, 65–95 (1919).
4. R. F. Stamm and J. J. Whalen, "Energy distribution of diffraction gratings as a function of groove form (calculations by an equation of Henry A Rowland)," J. Opt. Soc. Am. **36**, 2–12 (1946).
5. U. Fano, "On the theory of imperfect diffraction gratings," J. Opt. Soc. Am. **38**, 921–929 (1948).
6. W. Friedl and B. Hartenstein, "Energy distribution of diffraction gratings as a function of groove form," J. Opt. Soc. Am. **45**, 398–399 (1955).
7. A. Hessel and A. A. Oliner, "A new theory of wood's anomalies on optical gratings," Appl. Opt. **4**, 1275–1297 (1965).
8. S. Fujiwara and Y. Iguchi, "Parallel wood's anomalies of optical gratings," J. Opt. Soc. Am. **58**, 361–367 (1968).
9. J. Gaskill and J. W. Sons, *Linear Systems, Fourier Transforms, and Optics*, Wiley Series in Pure and Applied Optics (Wiley, 1978).
10. A. Papoulis, *Systems and Transforms with Applications in Optics*, McGraw-Hill series in systems science (R.E. Krieger, 1981).
11. R. Bracewell and R. Bracewell, *The Fourier Transform and Its Applications*, Electrical engineering series (McGraw Hill, 2000).
12. J. Goodman, *Introduction to Fourier Optics*, McGraw-Hill physical and quantum electronics series (W. H. Freeman, 2005).
13. A. Mikš, *Applied Optics* (CTU in Prague, 2009).
14. N. George and G. M. Morris, "Diffraction by serrated apertures," J. Opt. Soc. Am. **70**, 6–17 (1980).
15. D. L. Jaggard and Y. Kim, "Diffraction by band-limited fractal screens," J. Opt. Soc. Am. A **4**, 1055–1062 (1987).
16. Y. Kim, D. L. Jaggard, and H. Grebel, "Diffraction by fractally serrated apertures," J. Opt. Soc. Am. A **8**, 20–26 (1991).
17. J. M. Auerbach and V. P. Karpenko, "Serrated-aperture apodizers for high-energy laser systems," Appl. Opt. **33**, 3179–3183 (1994).
18. M. Gu and X. S. Gan, "Fresnel diffraction by circular and serrated apertures illuminated with an ultrashort pulsed-laser beam," J. Opt. Soc. Am. A **13**, 771–778 (1996).
19. S. Mehrabkhani and T. Schneider, "Is the Rayleigh-Sommerfeld diffraction always an exact reference for high speed diffraction algorithms?" Opt. Express **25**, 30229–30240 (2017).
20. S. Chatterjee, V. C. Vani, R. R. Haghighi, and R. K. Banyal, "Diffraction by ruled gratings with variable spacing: fundamental method of intensity calculation," Appl. Opt. **58**, 8638–8647 (2019).
21. C. Buitrago-Duque and J. Garcia-Sucerquia, "Non-approximated Rayleigh–Sommerfeld diffraction integral: advantages and disadvantages in the propagation of complex wave fields," Appl. Opt. **58**, G11–G18 (2019).
22. I. Sizova, T. Moskalev, and L. Mikheev, "Laser beam shaping with circular serrated apertures. I Spatial filtering," Appl. Opt. **58**, 4905–4909 (2019).
23. I. Sizova, T. Moskalev, and L. Mikheev, "Laser beam shaping with circular serrated apertures. II Theory of the beam profile formation," Appl. Opt. **58**, 4910–4917 (2019).
24. F. J. Torcal-Milla, L. M. Sanchez-Brea, and E. Bernabeu, "Diffraction of gratings with rough edges," Opt. Express **16**, 19757–19769 (2008).
25. F. J. Torcal-Milla, L. M. Sanchez-Brea, and E. Bernabeu, "Self-imaging of gratings with rough strips," J. Opt. Soc. Am. A **25**, 2390–2394 (2008).
26. L. M. Sanchez-Brea, F. J. Torcal-Milla, and E. Bernabeu, "Far field of gratings with rough strips," J. Opt. Soc. Am. A **25**, 828–833 (2008).
27. S. Teng, Y. Cui, and Z. Li, "Talbot effect of grating with fractal rough edges," J. Opt. **18**, 015601 (2015).
28. F. J. Torcal-Milla and L. M. Sanchez-Brea, "Diffraction by random Ronchi gratings," Appl. Opt. **55**, 5855–5859 (2016).
29. F. J. Torcal-Milla and L. M. Sanchez-Brea, "Diffraction by gratings with random fill factor," Appl. Opt. **56**, 5253–5257 (2017).
30. A. Mikš and P. Pokorný, "Edge spread function of Talbot phenomenon," Optik **127**, 8065–8069 (2016).
31. H. Talbot, "LXXVI. Facts relating to optical science. No. IV," The London, Edinburgh, Dublin Philos. Mag. J. Sci. **9**, 401–407 (1836).
32. K. Patorski, "I the self-imaging phenomenon and its applications," in *Progress in Optics* (Elsevier, 1989), pp. 1–108.
33. J. Wen, Y. Zhang, and M. Xiao, "The Talbot effect: recent advances in classical optics, nonlinear optics, and quantum optics," Adv. Opt. Photon. **5**, 83–130 (2013).
34. K. Rektorys, *Survey of Applicable Mathematics*, Mathematics and Its Applications (Springer, 2013).

Optik 127 (2016) 8065–8069

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Edge spread function of Talbot phenomenon

Antonín Mikš, Petr Pokorný *

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thakurova 6, 166 29 Prague, Czech Republic

ARTICLE INFO

Article history:
Received 7 April 2016
Accepted 1 June 2016

Keywords:
Diffraction theory
Talbot phenomenon
Self-imaging
Edge spread function

ABSTRACT

The paper presents and analyses the edge spread function of Talbot phenomenon—self-imaging of a periodical structure in given distances behind a template, if the template is illuminated by a plane wave. This phenomenon has many applications in optical metrology and in industry. Detailed mathematical analysis with an amplitude grating illuminated by a monochromatic plane wave is performed, and a simple relation for the calculation of the edge spread function width of Talbot imaging of Ronchi grating is shown which allows a simple analysis of the edge spread function for different grating sizes without any difficult calculations. The results bring a complete insight to the problematics and it can serve as a great theoretical background in many applications which utilize Talbot phenomenon.

1. Introduction

English scientist William Henry Fox Talbot (1800–1877) discovered a phenomenon in 1836 [1,2] and he described the following properties of a propagating light field. When the wave field (e.g., a plane wave) impinges any periodical structure (for example an amplitude grating or a field of holes in an opaque iris) the unique distances from such the structure exist where the structure is self-imaged. Such a phenomenon – Talbot phenomenon – has many applications in optical metrology, interferometry, or microelectronics [2–11].

This paper presents a detailed mathematical description of such a phenomenon for an amplitude grating illuminated by a plane wave. A simple formula for a calculation of the edge spread function width, which affects the quality of Talbot image, is shown. Such a study of the edge spread function of Talbot phenomenon on the amplitude grating has not been published yet. The results of the analysis bring a complete insight to the problematics and it can serve as a great theoretical background in many applications where Talbot phenomenon is used.

2. Mathematical description of Talbot phenomenon

Let one describes a mathematical theory of the phenomenon on the example of an amplitude grating. Suppose the amplitude diffraction grating (Ronchi grating) shown in Fig. 1 where p denotes period and b is width of a transparent part of the grating.

* Corresponding author.
E-mail address: petr.pokorny@fsv.cvut.cz (P. Pokorný).

http://dx.doi.org/10.1016/j.ijleo.2016.06.003
0030-4026/

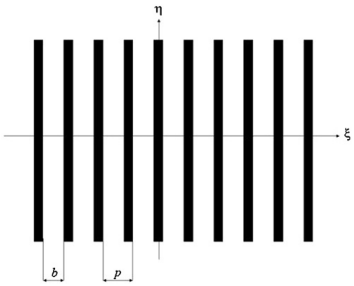


Fig. 1. Amplitude grating.

The grating shown in Fig. 1 can be described with the following formula:

$$f(\xi) = A \sum_{n=-\infty}^{\infty} \mathrm{rect}\left(\frac{\xi - np}{b}\right) = \sum_{n=-\infty}^{\infty} c_n \exp\left(\mathrm{i}\frac{2\pi}{p} n\xi\right). \tag{1}$$

In Eq. (1) i denotes imaginary part ($\mathrm{i} = \sqrt{(-1)}$), A is amplitude transparency of the grating and Fourier coefficients c_n are given as follows [12,13]:

$$c_n = A\frac{b}{p}\mathrm{sinc}\left(\frac{nb}{p}\right). \tag{2}$$

Without any loss of generality, one can suppose $A = 1$. Using Fresnel approximation in Sommerfeld diffraction formula [12–17], the complex amplitude $U(x, y, z)$ of a wave field in a point given by rectangular coodrinates $[x, y, z]$ behind the grating can be described as follows:

$$\begin{aligned} U(x,y,z) = C \int_{-d}^{d}\int_{-c}^{c} U(\xi,\eta,0) \sum_{n=-\infty}^{\infty} c_n \exp\left(\mathrm{i}\frac{2\pi}{p} n\xi\right) \times \\ \times \exp\left[\mathrm{i}k\frac{(x-\xi)^2 + (y-\eta)^2}{2z}\right] \mathrm{d}\xi\,\mathrm{d}\eta, \\ C = -\frac{\mathrm{i}}{\lambda}\frac{\exp(\mathrm{i}kz)}{z}, \end{aligned} \tag{3}$$

where $U(\xi, \eta, 0)$ denotes the wave field impinging the grating, $k = 2\pi/\lambda$ is wave number, λ is used wavelength, $2d$ is height and $2c$ is width of the grating, and (ξ, η) are coordinates in the grating's plane. Let one suppose a plane wave impinging the grating; afterwards, the generality of the study will not be lost with denoting $U(\xi, \eta, 0) = 1$. With the use of substitution:

$$\begin{aligned} &t = x - \xi,\ u = y - \eta,\ \alpha = \pi/\lambda z, \\ &\beta = -2\pi n/p,\ \gamma = -\beta x, \end{aligned} \tag{4}$$

the following formula can be used of the calculation of integrals in Eq. (3), it holds:

$$\int_{\varepsilon_1}^{\varepsilon_2} \exp[\mathrm{i}(\alpha t^2 + \beta t + \gamma)]\,\mathrm{d}t = \frac{1}{\sqrt{\alpha}}\exp\left[\mathrm{i}\left(\gamma - \frac{\beta^2}{4\alpha}\right)\right] \int_{\varepsilon_1\sqrt{\alpha}+\frac{\beta}{2\sqrt{\alpha}}}^{\varepsilon_2\sqrt{\alpha}+\frac{\beta}{2\sqrt{\alpha}}} \exp(\mathrm{i}v^2)\,\mathrm{d}v. \tag{5}$$

Afterwards, Eq. (3) can be rewritten using Eq. (5) as follows:

$$U(x,y,z) = -\frac{\mathrm{i}}{\pi}\exp(\mathrm{i}kz) \sum_{n=-\infty}^{\infty} U_n \exp\left(-\mathrm{i}\beta x\right) \exp\left(-\mathrm{i}\frac{\beta^2}{4\alpha}\right), \tag{6}$$

where

$$U_n = c_n \left[F\left(\theta_1\right) - F\left(\theta_2\right)\right] \left[F(\tau_1) - F(\tau_2)\right], \tag{7}$$

whereby

$$\begin{aligned}\tau_1 &= (x-c)\sqrt{\alpha}+\frac{\beta}{2\sqrt{\alpha}},\\ \tau_2 &= (x+c)\sqrt{\alpha}+\frac{\beta}{2\sqrt{\alpha}},\\ \theta_1 &= (y-d)\sqrt{\alpha},\\ \theta_2 &= (y+d)\sqrt{\alpha}.\end{aligned} \tag{8}$$

$F(\tau)$ in Eq. (7) is given with the following formula:

$$\begin{aligned}F(\tau) &= \int_0^{\tau} \exp(\mathrm{i}\zeta^2)\,\mathrm{d}\zeta\\ &= \sqrt{\frac{\pi}{2}}\left[C(\tau\sqrt{2/\pi})+\mathrm{i}\,S(\tau\sqrt{2/\pi})\right].\end{aligned} \tag{9}$$

In Eq. (9), C and S denote Fresnel integrals [15,16,18]. Intensity of the field in given point $[x, y, z]$ is than calculated with the formula:

$$\begin{aligned}I(x,y,z) &= |U(x,y,z)|^2\\ &= \frac{1}{\pi^2}\Big|\sum_{n=-\infty}^{\infty} U_n \exp\left(-\mathrm{i}\,\beta x\right)\ \exp\left(-\mathrm{i}\frac{\beta^2}{4\alpha}\right)\Big|^2.\end{aligned} \tag{10}$$

Suppose now that $z = Nz_T$, where $z_T = 2p^2/\lambda$ and $N = 1, 2, 3, \ldots$. Afterwards, the third component in Eq. (10) is equal to one. The distance z_T is so called Talbot distance [2,3]. The intensity of the field in Talbot distance is than given as follows:

$$I(x,y,Nz_T) = \frac{1}{\pi^2}\Big|\sum_{n=-\infty}^{\infty} A_n \exp\left(\mathrm{i}\,\frac{2\pi n}{p}x\right)\Big|^2, \tag{11}$$

where

$$A_n = c_n\left[F\left(\theta_{1n}\right)-F\left(\theta_{2n}\right)\right]\left[F(\tau_{1n})-F(\tau_{2n})\right], \tag{12}$$

$$\begin{aligned}\theta_{1n} &= \frac{y-d}{p}\sqrt{\frac{\pi}{2N}},\ \theta_{2n} = \frac{y+d}{p}\sqrt{\frac{\pi}{2N}},\\ \tau_{1n} &= \sqrt{\frac{\pi}{2N}}\left(\frac{x-c}{p}-2Nn\right),\ \tau_{2n} = \sqrt{\frac{\pi}{2N}}\left(\frac{x+c}{p}-2Nn\right).\end{aligned} \tag{13}$$

In the case of the grating with infinite size, i.e. $c \to \infty$, $d \to \infty$, it holds: $A_n \to c_n$. Therefore, Talbot image of the grating is its own copy. For a finite sized grating, it is obvious from Eq. (11) that Talbot image has the same period as the grating. However, while the grating is defined with Fourier coefficients c_n, its Talbot image has different Fourier coefficients A_n. Moreover, one can see from Eqs. (11) and (13) that Talbot image of Ronchi grating does not depend on the used wavelength.

3. Edge spread function of Talbot phenomenon

Let one study the effect of finite size of the grating on its Talbot image now. Because the size of the grating is limited the edge of the image will not be steep as in the case of infinite grating. It will be gradual and the transition will have given width. Function describing such a transition is called the edge spread function. The main focus of the following study will be given on the effect of finite size of the grating on the edge spread function.

Let one define the width Δ of the edge spread function (similarly as in the theory of optical imaging) as a distance between x-coordinates of the points where the tangent to the edge spread function intersects the top and the bottom of the pulse; therefore, it holds: $\Delta = -A/I'(x_i)$, where A is amplitude of the pulse and $I'(x_i)$ is derivative of Eq. (11) in the inflection point on the side of the transition with x-coordinate x_i. In Fig. 2 the situation for the grating with once amplitude ($A = 1$) and size 10 mm × 10 mm is shown. For the edge spread function width Δ it holds: $\Delta = 1.54$ μm. Summation in Eq. (11) is approximated for n from $-N_{\max}$ to $+N_{\max}$. One can see Gibbs phenomenon as well, as a result of approximation of discontinuous functions (grating) with Fourier series [12,19].

One of the criteria for an image quality in the theory of optical imaging is an area under the edge spread function. The smaller the area the better image quality of the optical system from the point of edge sharpness view. Such a situation for the aforementioned Talbot imaging is shown in Fig. 2 with yellow colour.

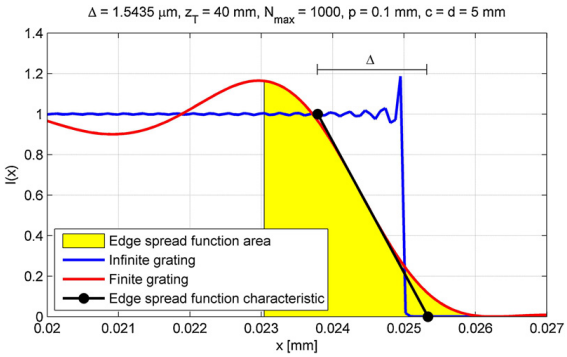


Fig. 2. Edge spread function for grating with size 10 mm × 10 mm. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

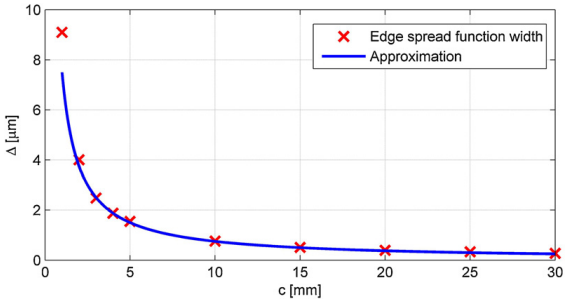


Fig. 3. Dependency of edge spread function width Δ on grating size for grating period $p = 0.1$ mm.

Fig. 3 shows a dependency of the edge spread function on the grating size. Let one suppose the grating with size $2c \times 2c$ (i.e., $c = d$) and period p. Analysis of the aforementioned formulas gives the following approximate expression for the edge spread function width Δ, it holds:

$$\Delta \approx 0.75\, p^2/c. \tag{14}$$

One can use Eq. (14) for an easy calculation of the edge spread function width of Talbot imaging; the formula gives results with an error less than 10% for gratings where $p/c > 0.1$.

4. Effect of used wavelength on quality of Talbot image

Let one study the effect of a spectral width dλ of a used light on the quality of Talbot imaging with a grating. As is obvious from the aforementioned formulas, Talbot distance $z_T = 2p^2/\lambda$ is wavelength dependent. For the case of small changes of the wavelength, i.e. $d\lambda << \lambda$, the change of Talbot distance can be calculated as follows:

$$dz_T = -2\left(\frac{p}{\lambda}\right)^2 d\lambda = -z_T\left(\frac{d\lambda}{\lambda}\right). \tag{15}$$

As is obvious from Eq. (15), the change in Talbot distance is dependent on the ratio p/λ squared for a given spectral width dλ. For usual cases in practice, where period p is much larger than used average wavelength λ, the change in Talbot distance will be very significant. For an example, the grating with period $p = 0.1$ mm, used average wavelength $\lambda = 500$ nm, and spectral width $d\lambda = 20$ nm cause change in Talbot distance $dz_T = -1.6$ mm which is a significant value compared to Talbot distance $z_T = 40$ mm. Usage of Talbot phenomenon in practical technical applications requires sources of light with very tight spectral width (lasers, laser diodes, etc.). Fig. 4 shows the edge spread functions dependent on used wavelengths for a grating illumination. The calculation was done for a plane distanced from the grating by Talbot distance for wavelength $\lambda = 500$ nm ($z_T = 40$ mm). As is obvious, the small change in wavelength ($d\lambda = 1$ nm) cause significant changes in intensity profiles. Therefore, the grating image will be blurred in such an image plane.

5. Conclusion

Talbot phenomenon—a self-imaging of a periodical structure in given distances behind its template, when the structure is illuminated by a plane or a spherical wave, has many practical applications in optical metrology and in industry. The paper presented an analysis of the effect of a finite size of the grating on the edge spread function of Talbot image. The effect of a spectral width on Talbot distance was studied as well. It was shown that Talbot image of Ronchi finite grating has the same period as in the case of infinite grating; however, the amplitudes are different and they are not dependent on the used

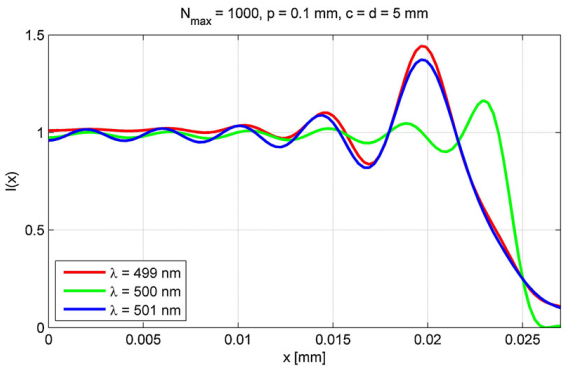


Fig. 4. The effect of wavelength changes on edge spread function for grating with size 10 mm × 10 mm and period $p = 0.1$ mm.

wavelength which is used for an illumination. Afterwards, it was shown that a small change of wavelength affects the edge spread function significantly. The simple relation for the calculation of the edge spread function width of Talbot imaging of Ronchi grating is shown. It allows the simple analysis of the edge spread function for different grating sizes without any difficult calculations.

Acknowledgement

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS15/125/OHK1/2T/11

References

[1] W.H.F. Talbot, Facts relating to optical science No. IV, Philos. Mag. 9 (1836) 401–407.
[2] K. Patorski, The self-imaging phenomenon and its applications, Prog. Opt. 27 (1989) 3–108.
[3] J. Wen, Y. Zhang, M. Xiao, The Talbot effect: recent advances in classical optics, nonlinear optics, and quantum optics, Adv. Opt. Photon. 5 (2013) 83–130.
[4] D. Malacara, Optical Shop Testing, John Wiley & Sons, 2007.
[5] J.C. Bhattacharya, Measurement of the refractive index using the Talbot effect and a moire technique, Appl. Opt. 28 (1989) 2600–2604.
[6] Y. Nakano, K. Murata, Measurements of phase objects using the Talbot effect and moiré techniques, Appl. Opt. 23 (1984) 2296–2299.
[7] G.S. Spagnolo, D. Ambrosini, D. Paoletti, Displacement measurement using the Talbot effect with a Ronchi grating, J. Opt. A: Pure Appl. Opt. 4 (2002) S376–80.
[8] L. Wang, D. Yun, Talbot effect and its application in measurement of in-plane strain, Opt. Laser Eng. 18 (1993) 109–114.
[9] X. Jin, J. Zhang, J. Bai, C. Hou, X. Hou, Calibration method for high-accuracy measurement of long focal length with Talbot interferometry, Appl. Opt. 51 (2012) 2407–2413.
[10] S. Lee, Talbot interferometry for measuring the focal length of a lens without moiré fringes, J. Opt. Soc. Korea 19 (2015) 165–168.
[11] A. Isoyan, F. Jiang, Y.C. Cheng, F. Cerrina, P. Wachulak, L. Urbanski, et al., Talbot lithography: self-imaging of complex structures, J. Vac. Sci. Technol. B 27 (6) (2009) 2931–2937.
[12] R.N. Bracewell, The Fourier Transform and Its Aplications, McGraw-Hill, 2000.
[13] J.D. Gaskill, Linear Systems, Fourier Transforms and Optics, John Wiley & Sons, 1978.
[14] A. Miks, Applied Optics, CTU in Prague, Prague, 2009.
[15] J. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, 1968.
[16] A. Papoulis, Systems and Transforms with Applications in Optics, McGraw-Hill, New York, 1968.
[17] O.K. Ersoy, Diffraction, Fourier Optics and Imaging, John Wiley & Sons, 2007.
[18] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, Dover Publications, 2007.
[19] G. Arfken, Mathematical Methods for Physics, Academic Press, NewYork, 2005.

Optik - International Journal for Light and Electron Optics 239 (2021) 166885

Contents lists available at ScienceDirect

Optik - International Journal for Light and Electron Optics

journal homepage: www.elsevier.com/locate/ijleo

Original research article

Explicit calculation of Point Spread Function of optical system

Antonín Mikš, Petr Pokorný *

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic

ARTICLE INFO

Keywords:
Point spread function
Pupil function
Explicit calculation
Sonin's integral

ABSTRACT

The paper presents novel formulas for an explicit calculation of the Point Spread Function for an optical system with a circular pupil and the case of axial-point imaging, which were derived with the use of Sonin's integral. In such a case, it is not needed to perform integration in the pupil, as it is usual with different methods. The approach is presented and verified with examples. The presented apparatus wide-spreads the theory of calculation of the Point Spread Function, and one can use it as an alternative solution of the mentioned task.

1. Introduction

One of generally well-known facts of the theory of optical imaging is that an image of a point, which is created by an optical system, can be described as specific energy distribution (diffraction) characterised by intensity of the wave-field — so called PSF which stands for the Point Spread Function [1–5]. If the optical system is rotationally symmetric and without any aberrations, the maximal intensity of the wave-field will lay in a plane which corresponds to a geometric-optical image of the point at the object.

Intensity of such the field then depends on a shape of the pupil (circle, annulus, rectangle, slit, etc.), on a transmissivity, and on a wavelength of used light. Further, the intensity distribution varies with distance where it is studied (calculated, measured). In the case of a circular pupil, the problem is well described in Refs. [1–16], and several studies have been presented on the topic of so called apodisation — the effect of pupil shape and transmissivity on PSF, see, for example, Refs. [17–19].

This paper covers the area of explicit calculation of the Point Spread Function for an optical system with a circular pupil and the case of axial-point imaging. As will be presented, the Sonin's integral [20,21] is used and novel explicit formulas are derived. With the proposed method, it is not needed to perform integration in the pupil, as it is usual with different methods of PSF calculation. One can explicitly enumerate the wave-field state in the image plane just with the use of proper series. The approach is then presented and verified with examples.

The presented apparatus wide-spreads the theory of calculation of the Point Spread Function, and one can use it as an alternative solution of the mentioned task, where the process of integration in the exit pupil of the optical system can be eliminated.

2. Imaging of axial point by optical system with circular pupil

At the beginning, suppose an optical system without optical aberrations and let the object and image space to be air. For the following analysis, only the scalar wave-field will be considered, as it is a very accurate approximation for optical system with numerical apertures less than 0.7 [8,9]. This condition is fulfilled in most of practical situations of optical systems (an exception would be, for example, photographic objectives, microscope objectives with small or medium value of numerical aperture, telescope lenses, etc.). Polarisation properties of the field can be neglected in this situation as well.

* Corresponding author.
E-mail address: petr.pokorny@fsv.cvut.cz (P. Pokorný).

https://doi.org/10.1016/j.ijleo.2021.166885
Received 27 February 2021; Accepted 31 March 2021
Available online 7 April 2021

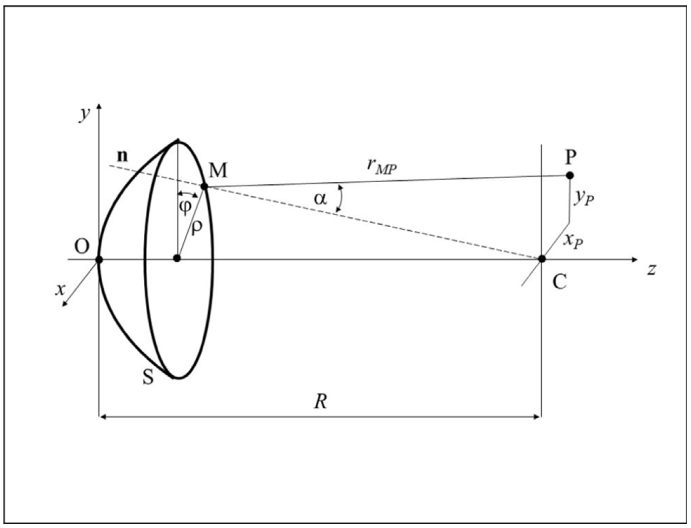


Fig. 1. Calculation of diffraction of convergent spherical wave.

Suppose now the situation in Fig. 1. Let the scalar wave-field $U(M)$ at the wave-front S in image space of the optical system is given. It is known from theory of optical imaging that the amplitude $U(P)$ of the wave-field at point P in image space of the optical system can be expressed as follows [1–5]:

$$U(P) = -\frac{\mathrm{i}}{\lambda}\iint_S U(\mathrm{M})\frac{\exp(\mathrm{i}kr_{MP})}{r_{MP}}\cos\alpha\ \mathrm{d}S\,, \tag{1}$$

where M is the point at the wave-front S, r_{MP} is the distance between points P and M, λ is the wavelength of light, $k = 2\pi/\lambda$ is the wave-number, and α is the angle between inner normal $\mathbf{n}$ of the wave-front S and the vector $\mathbf{r}_{MP}$, i denotes the imaginary unit. Suppose next that the wave-field $U(M)$ is a convergent spherical wave with radius R having its origin in point C, i.e. $U(M) = A(M)\exp(-ikR)/R$, where $A(M)$ denotes the complex amplitude of such wave at point M.

Further, let the optical system is rotationally symmetric with circular pupil, and let values of Δ, x_P, and y_P are much less than radius R of the wave-front (i.e., the angle α is small as well). Then one can approximate $\cos\alpha \approx 1$, and $\exp(ikr_{MP})/r_{MP} \approx \exp(ikr_{MP})/R$. With such assumptions, one can simply conclude the following approximate formula for the amplitude $U(P)$ at the point P (for $A(M) \approx A(\rho)$ and circular pupil), it holds [1–6]:

$$U(P) = -\mathrm{i}\frac{2\pi a^2}{\lambda R}\left[\frac{\exp(\mathrm{i}k\bar{R})}{R}\right]\int_0^1 A(\rho)J_0(\tau\rho)\ \rho\,\mathrm{d}\rho\,, \tag{2}$$

where

$$\tau = \frac{2\pi}{\lambda}\left(\frac{a}{R}\right)\sqrt{x_P^2 + y_P^2} = \frac{\pi\sqrt{x_P^2 + y_P^2}}{\lambda c} = \frac{\pi t}{\lambda c}\,, \tag{3}$$

$t = \sqrt{x_P^2 + y_P^2}$, $\rho = r/a$, $\bar{R} = r_{OP} - R$, $a = r_{\max}$ is the maximal value of r, J_0 is the Bessel function of the first kind, and $c = R/(2a)$ is the f-number of the optical system in image space.

In the presence of aberrations of the optical system, one can modify and express Eq. (2) as follows [1–5]:

$$U(P) = L\int_0^1 P(\rho)J_0(\tau\rho)\ \rho\,\mathrm{d}\rho\,, \tag{4}$$

where

$$L = -\mathrm{i}\frac{2\pi a^2}{\lambda R}\left[\frac{\exp(\mathrm{i}k\bar{R})}{R}\right] = -\mathrm{i}\left(\frac{\pi}{2\lambda c^2}\right)\exp(\mathrm{i}k\bar{R})\,, \tag{5}$$

$$P(\rho) = T(\rho)\exp\left[\mathrm{i}kW(\rho)\right] = T(\rho)\{\cos[kW(\rho)] + i\sin[kW(\rho)]\}\,, \tag{6}$$

where function $T(\rho)$ characterises the amplitude pupil transmissivity and function $W(\rho)$ is the wave aberration of the optical system [1–5].

Finally, intensity $I(P)$ in the case of imaging of axial point can be generally written using Eq. (4) as follows:

$$I(\mathrm{P}) = |U(\mathrm{P})|^2 = K\left|\int_0^1 P(\rho)J_0(\tau\rho)\ \rho\,\mathrm{d}\rho\right|^2, \tag{7}$$

where

$$K = |L|^2 = \left(\frac{\pi}{2\lambda c^2}\right)^2. \tag{8}$$

3. Defocusation and spherical aberration of the third and fifth order

Consider now the optical system which images an axial point with presence of optical aberrations and defocusation. Such a case covers various practical situations of optical systems, where an image of the axial point with as small aberrations as possible is desired.

The wave aberration for imaging of axial point can be written, in the case of aberrations of the third order (Seidel aberrations), as follows [1–6]:

$$W(\rho) = W_0 + W_2\rho^2 + W_4\rho^4\,, \tag{9}$$

where

$$W_2 = -\frac{\Delta}{8c^2}\,, \quad W_4 = \frac{\delta s'_K}{16c^2}\,, \tag{10}$$

Δ in the longitudinal defocusation, $\delta s'_K$ is the longitudinal spherical aberration for an aperture ray passing the edge of the entrance pupil, W_0 is a constant, and c is the f-number of the optical system in the image space.

In the case of aberrations of the fifth order one can write [1–6]:

$$W(\rho) = W_0 + W_2\rho^2 + W_4\rho^4 + W_6\rho^6\,. \tag{11}$$

If one assumes the image laying in the optimal image plane, where the Strehl ratio has maximal value, then it holds [22]:

$$W_2 = -W_4 - \frac{9}{10}W_6\,, \quad \Delta = -8c^2W_2\,, \tag{12}$$

where Δ denotes the distance of optimal image plane from the paraxial image plane of the optical system. Considering $\delta s'_{ext}$ to be the extreme value of the longitudinal spherical aberration and r_0 the radial range of the zone from the optical axis where the longitudinal spherical aberration has zero value (i.e., it is corrected), one can write following equations for coefficients W_4 and W_6, it holds:

$$W_6 = -\frac{\delta s'_{ext}}{6c^2r_0^4}\,, \quad W_4 = -\frac{3}{2}W_6r_0^2\,. \tag{13}$$

4. Explicit calculation of point spread function

Enumeration of integral in Eq. (4) or (7) for the intensity in the image plane of the axial point, i.e., the calculation of the Point Spread Function (PSF), can be performed with the use of Sonin's integral [20,21], where it holds:

$$J_{m+n+1}(x) = \frac{x^{n+1}}{2^n\Gamma(n+1)}\int_0^{\pi/2} J_m(x\sin t)\sin^{m+1}t\cos^{2n+1}t\,\mathrm{d}t\,. \tag{14}$$

If $m = 0$, $x = \tau$, and $\sin t = \rho$ in Eq. (14), one can write Sonin's integral as follows:

$$\int_0^1 J_0(\tau\rho)(1-\rho^2)^n\rho\,\mathrm{d}\rho = 2^n\Gamma(n+1)\frac{J_{n+1}(\tau)}{\tau^{n+1}}\,, \tag{15}$$

where $\Gamma(n+1)$ is Gamma function [23]. Further, if one expands function $T(\rho)\sin[kW(\rho)]$ and $T(\rho)\cos[kW(\rho)]$ as series:

$$T(\rho)\sin[kW(\rho)] = \sum_s p_s(1-\rho^2)^s\,, \tag{16}$$

$$T(\rho)\cos[kW(\rho)] = \sum_s q_s(1-\rho^2)^s\,,$$

then the amplitude $U(P)$ can be calculated with formula:

$$U(\mathrm{P}) = L\int_0^1 P(\rho)J_0(\tau\rho)\;\rho\,\mathrm{d}\rho = L\sum_s 2^s(q_s + ip_s)\Gamma(s+1)\frac{J_{s+1}(\tau)}{\tau^{s+1}}\,. \tag{17}$$

With Eq. (17), one can explicitly calculate the amplitude of the wave-field for the case of imaging of the axial point of the optical system, if the series expansion of the function $P(\rho)$ with basis $\{(1-\rho^2)^s\}$, where $s = 0, 1, 2, \ldots$, is known. Afterwards, one simply gets PSF of the optical system with Eq. (7), i.e. $I(\mathrm{P}) = |U(\mathrm{P})|^2$. The normalised intensity distribution (assuming $I_n(\tau = 0) = 1$) then can be calculated with formula:

$$I_n(\mathrm{P}) = K_n\left|\sum_s 2^s(q_s + ip_s)\Gamma(s+1)\frac{J_{s+1}(\tau)}{\tau^{s+1}}\right|^2\,, \tag{18}$$

where

$$K_n = 4\left|\sum_s q_s\right|^{-2}\,. \tag{19}$$

Axial Point Spread Function (APSF) can be calculated as a limit case for $\tau = 0$, it holds:

$$U(\tau = 0) = L \sum_s \frac{(q_s + ip_s)}{2(s+1)}\,, \tag{20}$$

and the normalised intensity of APSF is then given as follows:

$$\mathrm{APSF}_n = I_n(\tau = 0) = K_n \left| \sum_s \frac{(q_s + ip_s)}{2(s+1)} \right|^2 . \tag{21}$$

5. Expansion of function $P(\rho)$ into a series of $\{(1-\rho^2)^s\}$

As was shown in the aforementioned parts of the paper, explicit calculation of PSF or APSF can be done if one is able to find the series expansion (6) of $P(\rho)$ with functions $\{(1-\rho^2)^s\}$, where $s = 0, 1, \ldots$, i.e. one has to calculate coefficients p_s and q_s of Eqs. (16). Exact (explicit) calculation of those coefficients is difficult and it is appropriate to calculate then numerically. The following part will present several methods for such purpose.

5.1. Numerical calculation of coefficients p_s and q_s with least squares method

One of approaches to calculate coefficients p_s and q_s of Eqs. (16) is numerical enumeration where the function $P(\rho)$ is approximated on a discrete set of values ρ_m.

Suppose that one knows values of $P(\rho_m) = F_c(\rho_m) + i\, F_s(\rho_m)$ at points ρ_m, where

$$F_c(\rho_m) = T(\rho_m)\cos[kW(\rho_m)]\,, \quad F_s(\rho_m) = T(\rho_m)\sin[kW(\rho_m)]\,, \tag{22}$$

$m = 0, 1, \ldots M$, $\rho_0 = 0$, $\rho_M = 1$. Further, let S denotes the degree of expansion $\{(1-\rho^2)^s\}$, i.e. $s = 0, 1, \ldots, S$, $M \geq S$, then one can calculate coefficients p_s and q_s by solving the system of equations, for example with least squares method [23].

System's matrix can be symbolically written as follows (with indexing from 0): $\mathbf{A} = \left(a_{m+1,s+1}\right) = (1-\rho_m{}^2)^s$. Further, let column matrices $\mathbf{F}_c$ and $\mathbf{F}_s$ are defined: $\mathbf{F}_c = \left(c_{m+1,1}\right) = T(\rho_m)\cos[f(\rho_m)]$, $\mathbf{F}_s = \left(s_{m+1,1}\right) = T(\rho_m)\sin[f(\rho_m)]$. Afterwards, coefficient matrices $\mathbf{q}_s = \left(q_{s+1,1}\right) = q_s$ and $\mathbf{p}_s = \left(p_{s+1,1}\right) = p_s$ can be calculated as follows:

$$\mathbf{q}_s = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{F}_c\,, \quad \mathbf{p}_s = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{F}_s\,. \tag{23}$$

5.2. Calculation of coefficients p_s and q_s with Taylor series

Another approach of calculation of coefficients p_s and q_s in Eqs. (16) is with Taylor series [23]. Without any loss of generality, suppose that the pupil amplitude transmissivity is constant, i.e. $T(\rho) = T_0 = 1$.

Taylor series for a function $f(\rho) = \sin[kW(\rho)]$, where $W(\rho) = \sum_{i=1}^{N} W_{2i}\rho^{2i}$ is even function, which has small values around the point $\rho = 0$ (Maclaurin series), can be written as follows:

$$\sin[kW(\rho)] = \sum_{s=0}^{S} F_{2s}\rho^{2s}\,, \tag{24}$$

where

$$F_{2s} = \left.\frac{f^{(2s)}(\rho)}{(2s)!}\right|_{\rho=0}\,. \tag{25}$$

Series expansion of

$$\sin[kW(\rho)] = \sum_{s=0}^{S} p_s(1-\rho^2)^s \tag{26}$$

is determined with coefficients p_s. With the use of binomial theorem then one gets:

$$\sum_{s=0}^{S} p_s(1-\rho^2)^s = \sum_{s=0}^{S}\sum_{k=0}^{s} p_s \binom{s}{k}(-1)^k\rho^{2k} = \sum_{s=0}^{S} F_{2s}\rho^{2s}\,, \tag{27}$$

where $\binom{s}{k} = \frac{s!}{(s-k)!k!}$.

With comparison of members of corresponding powers of ρ and after modification, considering the symmetry of given problem, one can express coefficients p_s as follows:

$$p_s = (-1)^s \sum_{i=s}^{S} \binom{i}{s} F_{2i}\,. \tag{28}$$

Coefficients F_{2i} can be calculated with Eq. (25), where one can enumerate high-order derivatives with general Faà di Bruno's theorem [24–26] (firstly presented by Arbogast in 1800 [24] and named by Faà di Bruno in 1855 [26]). It holds:

$$f^{(n)}(u(\rho)) = \sum \frac{n!}{m_1!m_2!\cdots m_n!} \cdot f^{(m_1+m_2+\cdots m_n)}(u(\rho)) \cdot \prod_{j=1}^{n} \left(\frac{u^{(j)}(\rho)}{j!} \right)^{m_j}, \tag{29}$$

where summing is done with all tuples of positive numbers $(m_1,\ m_2,\ \ldots,\ m_n)$ which satisfies condition $\sum_{j=1}^{n} j \cdot m_j = 1 \cdot m_1 + 2 \cdot m_2 + \cdots + n \cdot m_n = n$.

In the case of aberrations of the fifth order one set $N = 3$. Without any loss of generality one can suppose $W_0 = 0$ as well. Therefore, it holds: $W(\rho) = W_2\rho^2 + W_4\rho^4 + W_6\rho^6$. As an example, consider $S = 10$. In such a case, one can express the first 11 coefficients p_s as follows:

$$\begin{aligned}
p_0 &= F_0 + F_2 + F_4 + F_6 + F_8 + F_{10} + F_{12} + F_{14} + F_{16} + F_{18} + F_{20}\,, \\
p_1 &= F_2 + 2\,F_4 + 3\,F_6 + 4\,F_8 + 5\,F_{10} + 6\,F_{12} + 7\,F_{14} + 8\,F_{16} + 9\,F_{18} + 10\,F_{20}\,, \\
p_2 &= F_4 + 3\,F_6 + 6\,F_8 + 10\,F_{10} + 15\,F_{12} + 21F_{14} + 28\,F_{16} + 36\,F_{18} + 45\,F_{20}\,, \\
p_3 &= F_6 + 4\,F_8 + 10\,F_{10} + 20\,F_{12} + 35\,F_{14} + 56\,F_{16} + 84\,F_{18} + 120\,F_{20}\,, \\
p_4 &= F_8 + 5\,F_{10} + 15\,F_{12} + 35\,F_{14} + 70\,F_{16} + 126\,F_{18} + 210\,F_{20}\,, \\
p_5 &= F_{10} + 6\,F_{12} + 21\,F_{14} + 56\,F_{16} + 126\,F_{18} + 252\,F_{20}\,, \\
p_6 &= F_{12} + 7\,F_{14} + 28\,F_{16} + 84\,F_{18} + 210\,F_{20}\,, \\
p_7 &= F_{14} + 8\,F_{16} + 36\,F_{18} + 120\,F_{20}\,, \\
p_8 &= F_{16} + 9\,F_{18} + 45\,F_{20}\,, \\
p_9 &= F_{18} + 10\,F_{20}\,, \\
p_{10} &= F_{20}\,.
\end{aligned} \tag{30}$$

Afterwards, one gets coefficients F_{2s} for function $f(\rho) = \sin[kW(\rho)]$ with Eq. (25) as follows:

$$\begin{aligned}
F_0 &= 0\,, \\
F_2 &= W_2\,k\,, \\
F_4 &= W_4\,k\,, \\
F_6 &= \left(-\frac{{W_2}^3}{6}\right) k^3 + W_6\,k\,, \\
F_8 &= \left(-\frac{{W_2}^2\,W_4}{2}\right) k^3\,, \\
F_{10} &= \frac{{W_2}^5}{120}\,k^5 + \left(-\frac{W_6\,{W_2}^2}{2} - \frac{W_2\,{W_4}^2}{2}\right) k^3\,, \\
F_{12} &= \frac{{W_2}^4\,W_4}{24}\,k^5 + \left(-\frac{{W_4}^3}{6} - W_2\,W_6\,W_4\right) k^3\,, \\
F_{14} &= \left(-\frac{{W_2}^7}{5040}\right) k^7 + \left(\frac{W_6\,{W_2}^4}{24} + \frac{{W_2}^3\,{W_4}^2}{12}\right) k^5 + \left(-\frac{{W_4}^2\,W_6}{2} - \frac{W_2\,{W_6}^2}{2}\right) k^3\,, \\
F_{16} &= \left(-\frac{{W_2}^6\,W_4}{720}\right) k^7 + \left(\frac{W_6\,{W_2}^3\,W_4}{6} + \frac{{W_2}^2\,{W_4}^3}{12}\right) k^5 + \left(-\frac{W_4\,{W_6}^2}{2}\right) k^3\,, \\
F_{18} &= \frac{{W_2}^9}{362880}\,k^9 + \left(-\frac{W_6\,{W_2}^6}{720} - \frac{{W_2}^5\,{W_4}^2}{240}\right) k^7 \\
&\quad + \left(\frac{{W_2}^3\,{W_6}^2}{12} + \frac{{W_2}^2\,{W_4}^2\,W_6}{4} + \frac{W_2\,{W_4}^4}{24}\right) k^5 + \left(-\frac{{W_6}^3}{6}\right) k^3\,, \\
F_{20} &= \frac{{W_2}^8\,W_4}{40320}\,k^9 + \left(-\frac{W_6\,{W_2}^5\,W_4}{120} - \frac{{W_2}^4\,{W_4}^3}{144}\right) k^7 \\
&\quad + \left(\frac{{W_2}^2\,W_4\,{W_6}^2}{4} + \frac{W_2\,{W_4}^3\,W_6}{6} + \frac{{W_4}^5}{120}\right) k^5\,.
\end{aligned} \tag{31}$$

To derive the aforementioned formulas, one can use properties of even function $W(\rho) = \sum_{i=1}^{N} W_{2i}\rho^{2i}$. Therefore, derivatives at point $\rho = 0$ can be expressed as follows:

$$W^{(2s)}(\rho)\Big|_{\rho=0} = (2s)!W_{2s}\,, \quad W^{(2s+1)}(\rho)\Big|_{\rho=0} = 0\,, \tag{32}$$

and further it holds:

$$f^{(m)}(kW(\rho)) = \sin\left[kW(\rho) + m\frac{\pi}{2}\right]\Big|_{\rho=0} = \sin\left(m\frac{\pi}{2}\right). \tag{33}$$

Similarly to Eq. (28), one gets coefficients q_s of $g(\rho) = \cos[kW(\rho)] = \sum_{s=0}^{S} q_s(1-\rho^2)^s$ as follows:

$$q_s = (-1)^s \sum_{i=s}^{S} \binom{i}{s} G_{2i}\,, \tag{34}$$

where G_{2i} denotes coefficients of Taylor series calculated with formula: $G_{2i} = g^{(2i)}(\rho)/(2i)!\Big|_{\rho=0}$. Further, it holds:

$$g^{(m)}(kW(\rho)) = \cos\left[kW(\rho) + m\frac{\pi}{2}\right]\Big|_{\rho=0} = \cos\left(m\frac{\pi}{2}\right), \tag{35}$$

and one gets coefficients G_{2i} after modification:

$$\begin{aligned}
&G_0 = 1\,, \quad G_2 = 0\,, \quad G_4 = \left(-\frac{{W_2}^2}{2}\right)k^2\,, \\
&G_6 = \left(-W_2\,W_4\right)\,k^2\,, \\
&G_8 = \frac{{W_2}^4}{24}\,k^4 + \left(-\frac{{W_4}^2}{2} - W_2\,W_6\right)k^2\,, \\
&G_{10} = \frac{{W_2}^3\,W_4}{6}\,k^4 + \left(-W_4\,W_6\right)\,k^2\,, \\
&G_{12} = \left(-\frac{{W_2}^6}{720}\right)k^6 + \left(\frac{W_6\,{W_2}^3}{6} + \frac{{W_2}^2\,{W_4}^2}{4}\right)k^4 + \left(-\frac{{W_6}^2}{2}\right)k^2\,, \\
&G_{14} = \left(-\frac{{W_2}^5\,W_4}{120}\right)k^6 + \left(\frac{W_6\,{W_2}^2\,W_4}{2} + \frac{W_2\,{W_4}^3}{6}\right)k^4\,, \\
&G_{16} = \frac{{W_2}^8}{40320}\,k^8 + \left(-\frac{W_6\,{W_2}^5}{120} - \frac{{W_2}^4\,{W_4}^2}{48}\right)k^6 \\
&\quad + \left(\frac{{W_2}^2\,{W_6}^2}{4} + \frac{W_2\,{W_4}^2\,W_6}{2} + \frac{{W_4}^4}{24}\right)k^4\,, \\
&G_{18} = \frac{{W_2}^7\,W_4}{5040}\,k^8 + \left(-\frac{W_6\,{W_2}^4\,W_4}{24} - \frac{{W_2}^3\,{W_4}^3}{36}\right)k^6 \\
&\quad + \left(\frac{{W_4}^3\,W_6}{6} + \frac{W_2\,W_4\,{W_6}^2}{2}\right)k^4\,, \\
&G_{20} = \left(-\frac{{W_2}^{10}}{3628800}\right)k^{10} + \left(\frac{W_6\,{W_2}^7}{5040} + \frac{{W_2}^6\,{W_4}^2}{1440}\right)k^8 \\
&\quad + \left(-\frac{{W_2}^4\,{W_6}^2}{48} - \frac{{W_2}^3\,{W_4}^2\,W_6}{12} - \frac{{W_2}^2\,{W_4}^4}{48}\right)k^6 \\
&\quad + \left(\frac{{W_4}^2\,{W_6}^2}{4} + \frac{W_2\,{W_6}^3}{6}\right)k^4\,.
\end{aligned} \tag{36}$$

6. Examples

In the following examples, the optical system affected by residual spherical aberration of the fifth order will be considered, and coefficients p_s and q_s of Eqs. (16) will be calculated. Suppose that the optical system has corrected spherical aberration for zone $r_0 = 1$ (pupil edge) and let the pupil amplitude transmissivity is constant, i.e. $T(\rho) = T_0 = 1$. Extreme value of longitudinal spherical aberration $\delta s'_{ext}$ will then occurs for zone $r = \sqrt{0.5} = 0.707$. Let the f-number of the optical system be $c = 4$ and $\delta s'_{ext} = 24\lambda c^2$. Afterwards, one gets aberration coefficients for optimal image plane as:

$$\begin{aligned}
&W_6 = \delta s'_{ext}/6c^2 = 4\lambda\,, \quad W_4 = -3W_6/2 = -\delta s'_{ext}/4c^2 = -6\lambda\,, \\
&W_2 = -W_4 - 9W_6/10 = 2.4\lambda\,, \quad \Delta = -8c^2W_2 = 0.8\delta s'_{ext} = 19.2\lambda c^2\,.
\end{aligned} \tag{37}$$

Finally, let the wavelength is $\lambda = 587\,\text{nm}$.

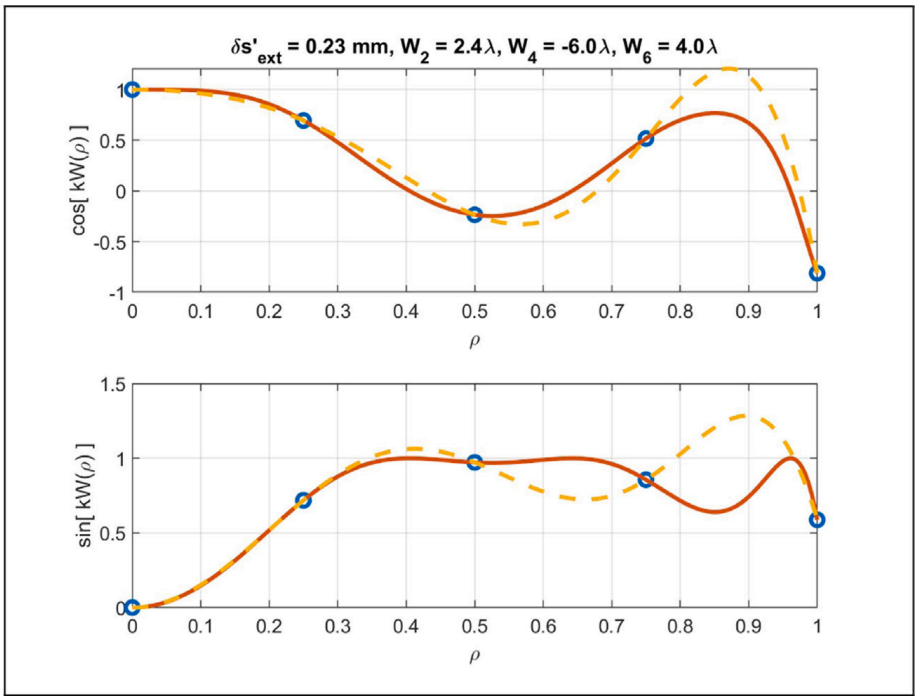


Fig. 2. Approximation of functions $\cos kW(\rho)$ and $\sin kW(\rho)$ for $S = M = 5$ (solid line — exact functional values, dashed line — approximation, blue points — uniform points distribution on the interval [0,1].

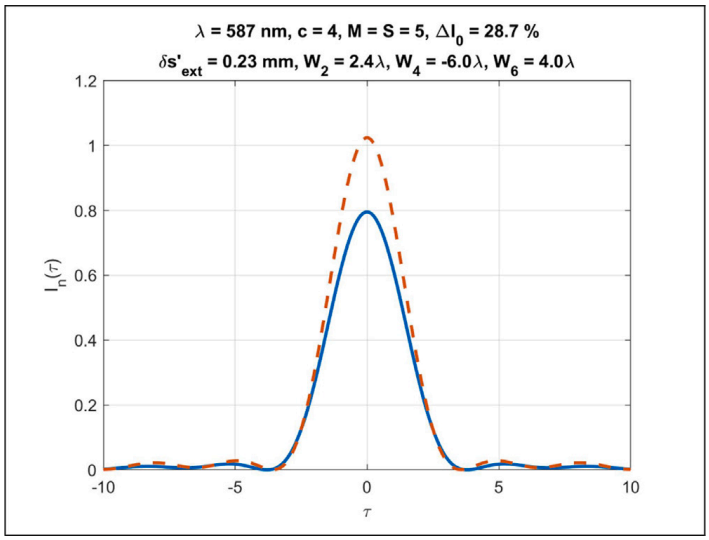


Fig. 3. Comparison between the calculation of normalised intensity distribution with Eq. (18) (dashed line) and numerical integration with Eq. (7) (solid line), $S = M = 5$.

6.1. Numerical calculation of coefficients with least squares method

As the first example, let the coefficients p_s and q_s of Eqs. (16) are calculated with least squares method. As was mentioned, properties of the wave aberration are known, and one can simple compose system matrix for selected points ρ_m, where $m = 0, 1, \ldots, M$, $\rho_0 = 0$, $\rho_M = 1$. Coefficients then can be calculated for selected S with Eqs. (23).

In the case that $P(\rho_m)$ is not known at point ρ_0, it is appropriate to fill the matrix of system by appropriate boundary conditions. For the case of uniform amplitude it holds: $\sum_s q_s = T_0 = 1$ and $\sum_s p_s = 0$. Therefore, new row of values $\left(a_{M+1,s+1}\right) = 1$ should be added to matrix $\mathbf{A}$, and matrices $\mathbf{F}_c$ and $\mathbf{F}_s$ should be filled by values $\left(c_{M+1,1}\right) = T_0 = 1$, and $\left(s_{M+1,1}\right) = 0$.

Fig. 2 shows comparison of approximations of function $\cos[kW(\rho)]$ and $\sin[kW(\rho)]$ for the case $S = M = 5$ and ρ_m uniformly distributed at interval $[0, 1]$. Fig. 3 then compares normalised intensities calculated with Eq. (18) and numerical integration by Eq. (7). One can see that error ΔI_0 in the calculation of normalised intensities for $\tau = 0$ is equal 28.7% in this case. Figs. 4 and 5 then presents similar results for $S = M = 18$, the order which were found iteratively to fulfil the condition $\Delta I_0 < 1\%$.

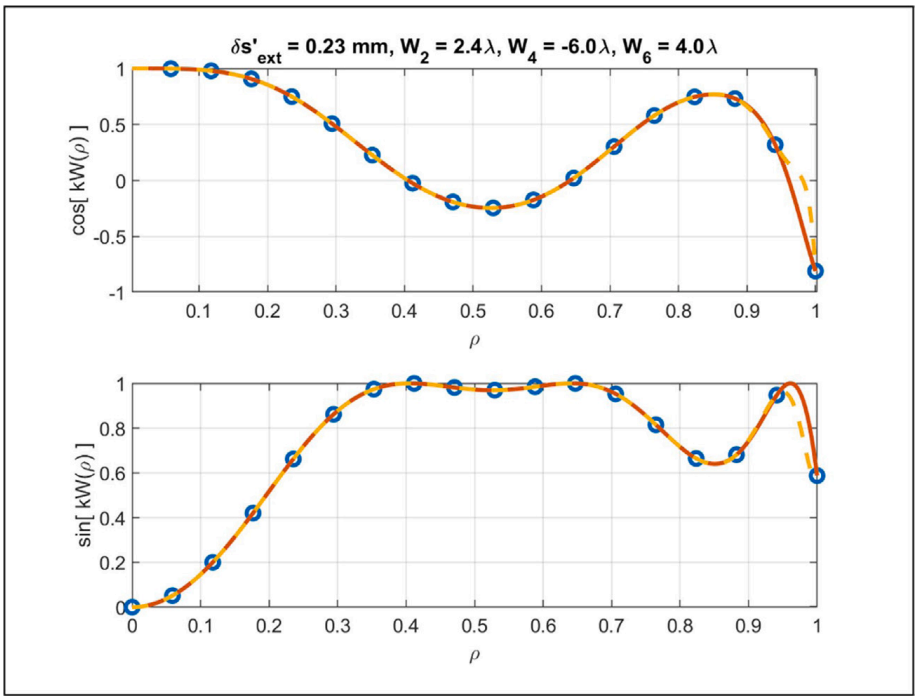


Fig. 4. Approximation of functions $\cos kW(\rho)$ and $\sin kW(\rho)$ for optimised $S = M = 18$ such that the error in calculation of normalised intensity ΔI_0 for $\tau = 0$ is less than 1% (solid line — exact functional values, dashed line — approximation, blue points — uniform points distribution on the interval [0,1].

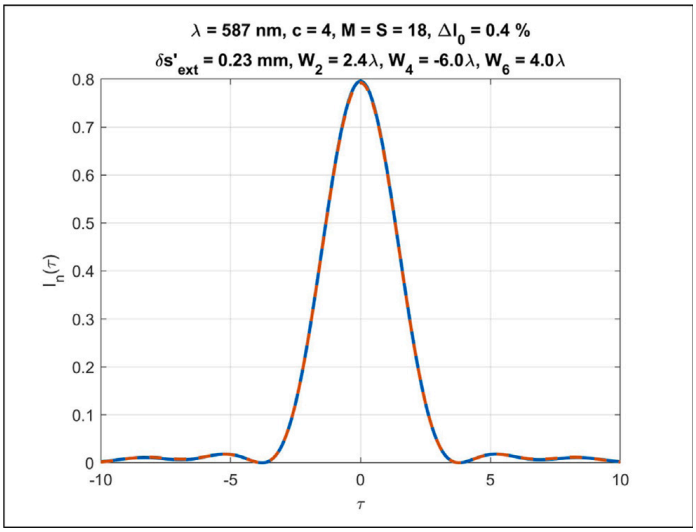


Fig. 5. Comparison between the calculation of normalised intensity distribution with Eq. (18) (dashed line) and numerical integration with Eq. (7) (solid line) for optimised $S = M = 18$ such that the error in calculation of normalised intensity ΔI_0 for $\tau = 0$ is less than 1%.

6.2. Numerical calculation of coefficients with Taylor series

The second example considers defocusation $W_2 = 0.5\lambda$ and $W_2 = W_6 = 0$. The goal is to find number of members (order) of the series for coefficients p_s and q_s to reach the accuracy of $\Delta I_0 < 1\%$ at $\tau = 0$, both for series in Eqs. (16) calculated analytically with Eqs. (28) and (34) and for numerical calculation by least squares method.

Fig. 6 shows result of iterative calculation with analytical formulas (28) and (34), where the desired condition is fulfilled for $M = S = 9$, i.e. for ten coefficients. Fig. 7 then presents similar results for numerical calculation by least squares method. In such a case, $M = S = 6$ only, i.e. one needs seven members of the series.

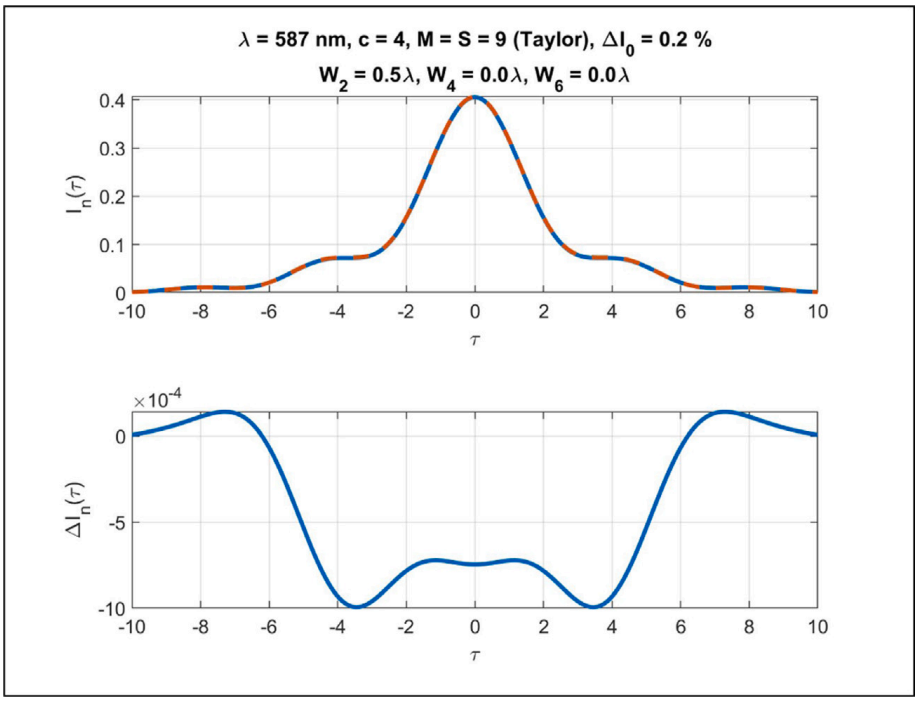


Fig. 6. Comparison between the calculation of normalised intensity distribution with Eq. (18) (dashed line in the top figure) and numerical integration with Eq. (7) (solid line in the top figure) if the coefficients p_s and q_s of series (16) were calculated with analytic formulas (28) and (34) (Taylor series) for optimised $S = M = 9$ such that the error in calculation of normalised intensity ΔI_0 for $\tau = 0$ is less than 1%; the bottom figure shows the difference of normalised intensities between explicit and numerical calculation.

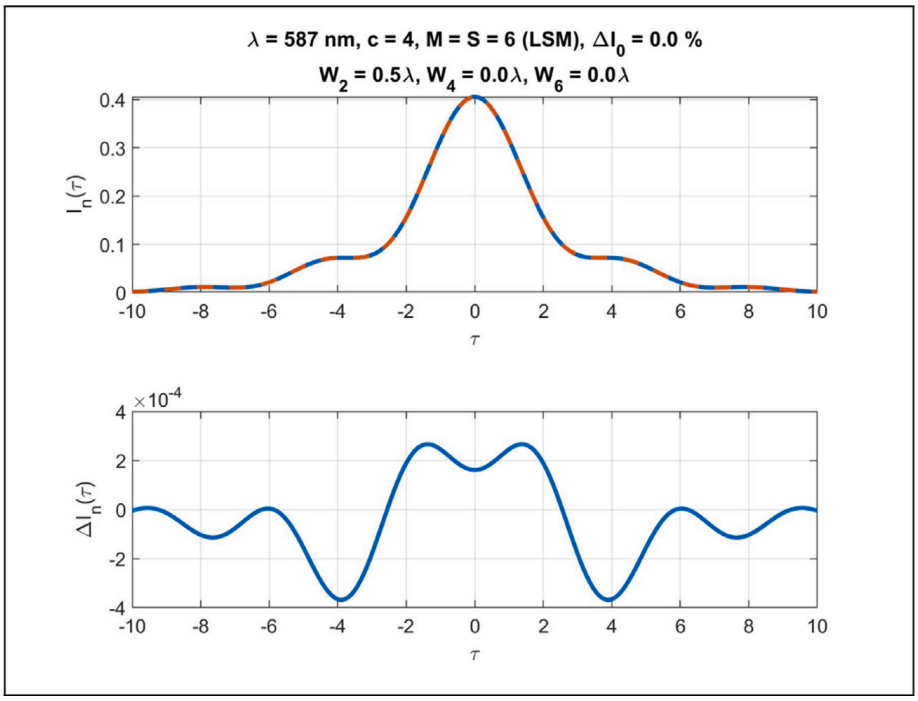


Fig. 7. Comparison between the calculation of normalised intensity distribution with Eq. (18) (dashed line in the top figure) and numerical integration with Eq. (7) (solid line in the top figure) if the coefficients p_s and q_s of series (16) were calculated with the least squares approach for optimised $S = M = 6$ such that the error in calculation of normalised intensity ΔI_0 for $\tau = 0$ is less than 1%; the bottom figure shows the difference of normalised intensities between explicit and numerical calculation.

7. Conclusion

The paper presented new formulas for an explicit calculation of the Point Spread Function for axial point imaging by the optical system with circular pupil, where Sonin's integral was used during derivation. With presented method, one can explicitly enumerate the intensity distribution in the image plane without any integration. Further, the derived formulas were presented and verified with examples.

The results fulfil the field of theory of the Point Spread Function calculation, and the proposed procedure can be used in various cases where the reader wants to avoid the integration in the plane of exit pupil of the optical systems. The derived formulas can be further used for analysis of influence of individual aberrations coefficients on distribution of the Points Spread Function.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by the Grant Agency of the Czech Technical University in Prague, Czech Republic, grant No. SGS21/092/OHK1/2T/11.

References

[1] M. Born, E. Wolf, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, Cambridge University Press, 1999.
[2] A. Maréchal, M. Françon, Diffraction, Structure Des Images: Influece de la Coherence de la Lumiere, Vol. 1, Éditions de la Revue d'optique théorique et instrumentale, 1960.
[3] E. O'Neill, Introduction To Statistical Optics, Dover Books on Physics, Dover Publications, 2003.
[4] J. Goodman, Introduction To Fourier Optics, W. H. Freeman, 2017.
[5] A. Mikš, Applied optics, CTU in Prague, 2009.
[6] W. Welford, Aberrations of Optical Systems, CRC Press, 2017.
[7] J. Braat, P. Dirksen, A.J.E.M. Janssen, Assessment of an extended nijboer–zernike approach for the computation of optical point-spread functions, J. Opt. Soc. Amer. A 19 (5) (2002) 858, http://dx.doi.org/10.1364/josaa.19.000858.
[8] A. Miks, J. Novak, P. Novak, Calculation of point-spread function for optical systems with finite value of numerical aperture, Optik 118 (11). http://dx.doi.org/10.1016/j.ijleo.2006.05.002.
[9] A.S. Marathay, G.B. Parrent, Use of scalar theory in optics, Journal of the Optical Society of America 60 (2). http://dx.doi.org/10.1364/josa.60.000243.
[10] D.G. Flagello, A.E. Rosenbluth, Lithographic tolerances based on vector diffraction theory, Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena 10 (6). http://dx.doi.org/10.1116/1.585959.
[11] J.J. Stamnes, H. Heier, Scalar and electromagnetic diffraction point-spread functions, Appl. Opt. 37 (17). http://dx.doi.org/10.1364/AO.37.003612.
[12] C.W. McCutchen, Generalized aperture and the three-dimensional diffraction image, J. Opt. Soc. Am. 54 (2). http://dx.doi.org/10.1364/josa.54.000240.
[13] J.C. Heurtley, Scalar Rayleigh-Sommerfeld and Kirchhoff diffraction integrals: A comparison of exact evaluations for axial points, J. Opt. Soc. Am. 63 (8). http://dx.doi.org/10.1364/josa.63.001003.
[14] H. Osterberg, L.W. Smith, Closed solutions of Rayleigh's diffraction integral for axial points, J. Opt. Soc. Am. 51 (10). http://dx.doi.org/10.1364/josa.51.001050.
[15] Y. Li, Predictions of Rayleigh's diffraction theory for the effect of focal shift in high-aperture systems, J. Opt. Soc. Am. A 25 (7). http://dx.doi.org/10.1364/josaa.25.001835.
[16] C.J.R. Sheppard, P. Török, Focal shift and the axial optical coordinate for high-aperture systems of finite fresnel number, J. Opt. Soc. Am. A 20 (11). http://dx.doi.org/10.1364/josaa.20.002156.
[17] P. Jacquinot, B. Roizen-Dossier, Apodisation, in: Progress in Optics, Elsevier, 1964, http://dx.doi.org/10.1016/s0079-6638(08)70570-5.
[18] A.N.K. Reddy, M. Hashemi, Apodization pupils: design and performance (review), J. Phys.: Conf. Ser. 1096. http://dx.doi.org/10.1088/1742-6596/1096/1/012140.
[19] J. Ojeda-Castañeda, C.M. Gómez-Sarabia, Tuning field depth at high resolution by pupil engineering, Adv. Opt. Photonics 7 (4). http://dx.doi.org/10.1364/aop.7.000814.
[20] Sonin integral, https://encyclopediaofmath.org/wiki/Sonin_integral.
[21] G. Watson, A Treatise on the Theory of Bessel Functions, Cambridge Mathematical Library, Cambridge University Press, 1995.
[22] A. Miks, J. Novak, P. Novak, Dependence of Strehl ratio on f-number of optical system, Appl. Opt. 51 (17) (2012) 3804, http://dx.doi.org/10.1364/ao.51.003804.
[23] K. Rektorys, Survey of Applicable Mathematics, Mathematics and Its Applications, Springer Netherlands, 2013.
[24] L. Arbogast, Du calcul des derivations, 1800.
[25] B. Tortolini, Annali di scienze matematiche e fisiche, no. sv. 6, Tip. delle bella arti, 1855.
[26] Faa di Bruno's formula, https://en.wikipedia.org/wiki/Fa%C3%A0_di_Bruno%27s_formula.

Optik - International Journal for Light and Electron Optics 230 (2021) 166317

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.com/locate/ijleo

Original research article

Influence of circular and annular pupil function on Axial Point Spread Function of optical system

Antonín Mikš, Petr Pokorný *

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic

ARTICLE INFO

Keywords:
Axial Point Spread Function
Point spread function
Pupil function

ABSTRACT

The paper presents an analysis of influence of shape and transmissivity of a pupil of an optical system on an axial energy distribution while the optical system images axial point. Therefore, a depth of focus is studied. Specifically, the analysis is done for uniform illumination of a circular pupil, and for the circular and annular pupil in the case of Gaussian amplitude distribution. Analytical formulas for Point Spread Function and Axial Point Spread Function are described, and characteristics of properties of the depth of focus of the optical system are derived. The novel analytic formulas presented in the paper can be used for determination of the value of the depth of focus of the optical system for given threshold limit of normalized axial intensity.

1. Introduction

It is well-known from the theory of optical imaging that an image of a point, created by an optical system, is a diffraction energy distribution characterised by an intensity of wave-field (PSF – Point Spread Function). In the case of a rotational symmetric optical system without aberrations, the maximal intensity of the wave-field will lay in a plane corresponding to a geometric-optical image of a point of an object. Intensity of the field will be dependent on a shape of the pupil (circle, annulus, rectangle, slit, etc.), on a transmissivity, and on a wavelength of used light. Further, the intensity distribution will be influenced by a distance where it is studied (calculated, measured). In the case of a circular pupil, the problem is well described in Refs. [1–11], and several studies have been presented on the topic of apodization – the effect of pupil shape and transmissivity on PSF, see, for example, Refs. [12–14].

The goal of this paper is to analyse the influence of shape and transmissivity of a pupil of an optical system on an axial energy distribution while the optical system images axial point. Therefore, a depth of focus is studied. Specifically, the analysis is done for uniform illumination of a circular pupil, and for the circular and annular pupil in the case of Gaussian amplitude distribution.

The analysis and modelling of the axial PSF (APSF) has been carried out by many authors. In the paper [15], the APSF is modelled and analysed for applications in wide-field microscopy. Authors of paper [16] model the transverse and axial PSF with the Martinez-Corral filter (the Martinez-Corral filter consists of a transparent annulus and central clear circular aperture of area less than the area of the annulus) and with new filter of a definite number of black and white annuli of a certain number of circles, where the center is a clear circular disc. Afterwards, the authors discuss several examples. The application confocal microscopy is shown in [17], where the authors propose pupil filter, which are composed by a number of concentric annular zones with constant real transmittance. The number of zones and their widths can be adapted according to the shape requirements. The presented method is then applied to design filters that produce axial super-resolution in scanning systems. As another example, the paper [18] can be stated, where the

* Corresponding author.
E-mail address: petr.pokorny@fsv.cvut.cz (P. Pokorný).

https://doi.org/10.1016/j.ijleo.2021.166317
Received 25 November 2020; Accepted 14 January 2021
Available online 23 January 2021

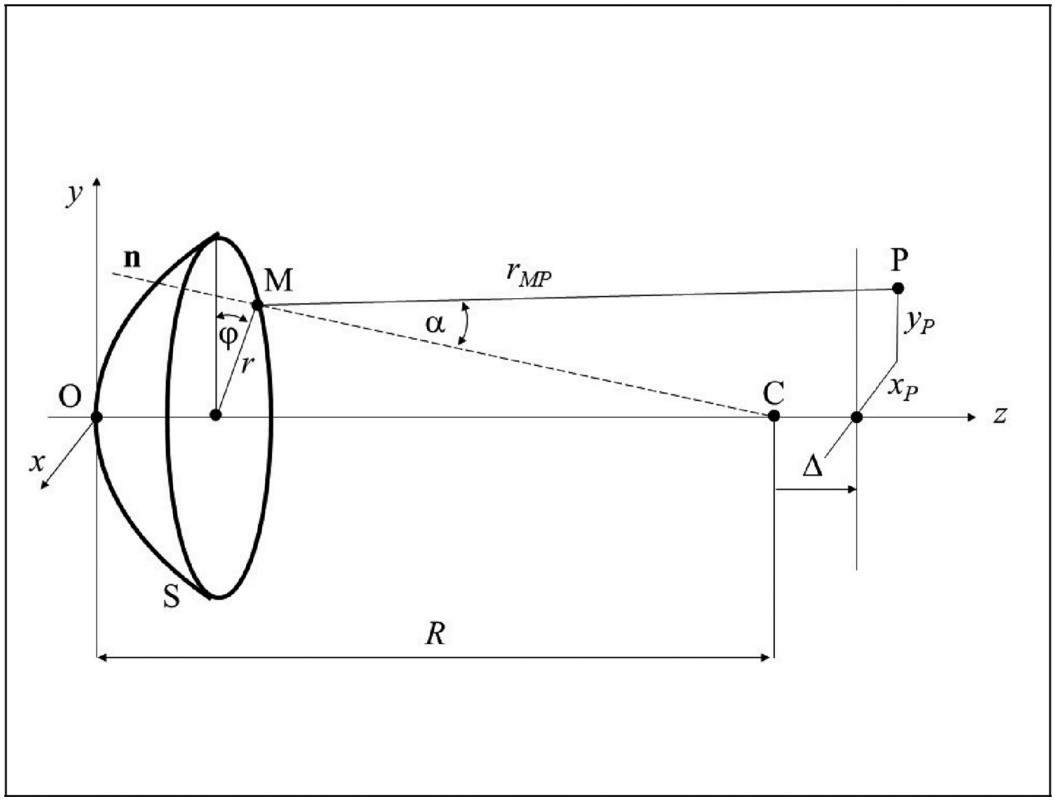


Fig. 1. Calculation of diffraction of convergent spherical wave.

authors present a new simple iterative method which is used to binarize pupil filters designed to control the intensity distribution in the focal volume of an optical system.

In comparison to the aforementioned papers, the presented work describes the PSF and APSF analytically. Specifically, the analysis is done for the case of a circular and annular pupil with uniform or Gaussian amplitude distribution in the pupil plane. Further, analytical formulas are derived which can be used for characterisation of the depth of focus properties, which have not been presented before. Those novel formulas are a useful completion of theory and analysis of PSF and APSF.

2. Imaging of axial point by optical system with circular pupil

To introduce the reader into the analysis, consider now a physically perfect optical system, i.e. the optical system without aberrations, and let an object and image space are an air. In the analysis, consider the scalar wave-field only, which gives enough accurate results for optical systems with numerical apertures less than 0.7 [3–6]. This condition can be fulfilled in most of practical situations of optical systems (an exception can be microscope objectives with high magnification). Therefore, polarization properties of the field does not have to be considered.

Suppose a scalar wave-field $U(M)$ on a wave-front S in an image space of the optical system to be known. Further, it is known from the theory of optical imaging that the amplitude $U(P)$ in a point P in an image space of the optical system can be expressed by the following formula, it holds [1,2]:

$$U(P) = -\frac{\mathrm{i}}{\lambda}\iint_S U(M)\frac{exp(\mathrm{i}kr_{\mathrm{MP}})}{r_{\mathrm{MP}}}cos\alpha \;\; \mathrm{d}S\,, \tag{1}$$

where M is the point on the wave-front S, r_{MP} is the distance between points P and M, λ is the wavelength of light, $k = 2\pi/\lambda$ is the wave-number, α is the angle between the inner normal n of the wave-front S in the point M and the vector $\mathrm{r_{MP}}$, and i is the imaginary unit. Suppose the field $U(M)$ to be a convergent spherical wave, i.e. $U(M) = A(M)exp(-\mathrm{ikR})/R$, with the radius R having its centre in the point C. Further, let the point P is in a plane perpendicular to an axis $\overline{\mathrm{OC}}$ which is distanced by Δ from the point C. Further, coordinates of the point P with respect to the point O are $(x_P, y_P, R+\Delta)$, and coordinates of the point M with respect to the point O are $(rsin\varphi, rcos\varphi, z_M)$, where

$$z_M(2R - z_M) = r^2. \tag{2}$$

Fig. 1 presents the considered situation.

According the Fig. 1, the following formula for r_{MP} holds:

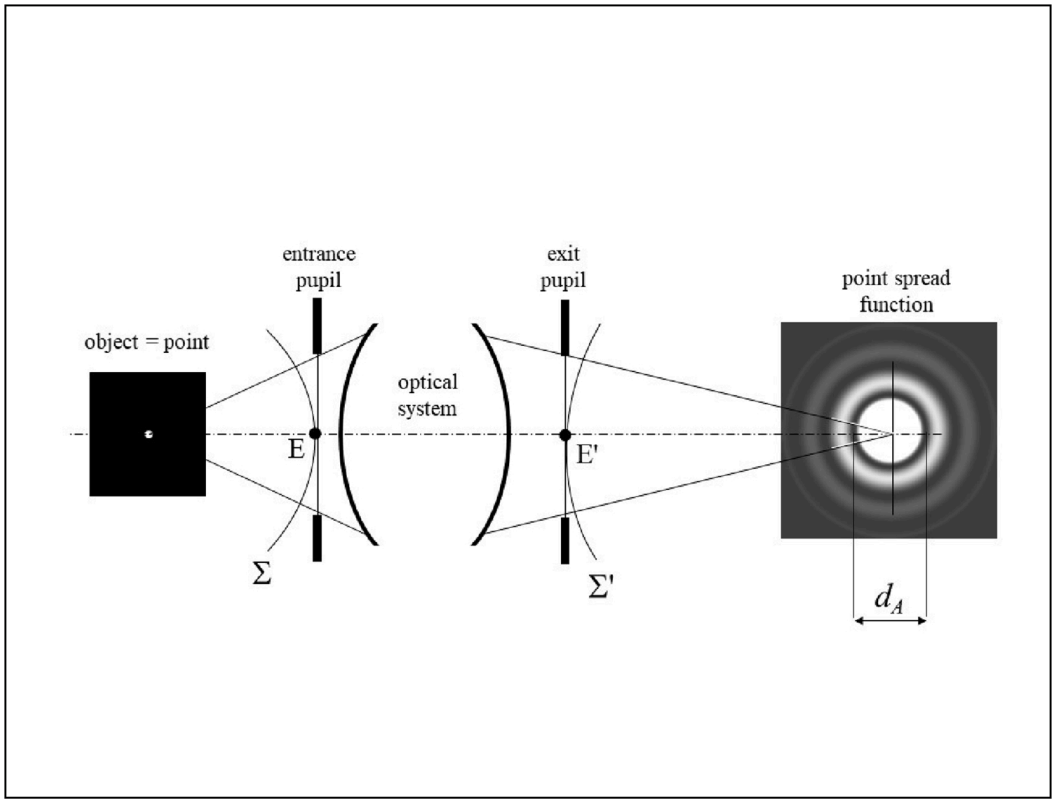


Fig. 2. Imaging of axial point with aberration-free optical system ($\Delta = 0$).

$$r_{\mathrm{MP}} = r_{\mathrm{OP}} - \Delta\frac{r^2}{2R^2} - \frac{r}{R}(x_P sin\phi + y_P cos\phi)\,, \tag{3}$$

where r_{OP} is the distance between points O and P.

Consider now a rotational symmetric optical system with a circular pupil. Further, let the values Δ, x_P, and y_P are much less than the radius R of the wave-front, i.e., the angle α is very small. In such a case, one can set: $cos\alpha \approx 1$, $exp(\mathrm{ikr_{MP}})/r_{\mathrm{MP}} \approx exp(\mathrm{ikr_{MP}})/R$, and the amplitude $U(P)$ of the field in the point P, for $U(A) = A(\rho)$ and the circular pupil, can be calculated with approximate formula [1,2]:

$$U(P) = -\mathrm{i}\frac{2\pi a^2}{\lambda R}\left[\frac{exp(\mathrm{ik}\overline{R})}{R}\right] \tag{4}$$

$$\times \int_0^1 A(\rho) exp\left(-\mathrm{ik}\frac{a^2\Delta}{2R^2}\rho^2\right) J_0(\tau\rho)\ \ \rho\,\mathrm{d}\rho\,,$$

where $\rho = r/a$, $a = r_{max}$ is the maximal value of r, and J_0 is the Bessel function of the first kind [19]. Further, one gets:

$$\overline{R} = r_{\mathrm{OP}} - R \approx \Delta + \frac{\Delta^2 + x_P^2 + y_P^2}{2R} \approx \Delta\,. \tag{5}$$

The value τ can be calculated with formula:

$$\tau = \frac{2\pi}{\lambda}\left(\frac{a}{R}\right)\sqrt{x_P^2 + y_P^2} = \frac{\pi\sqrt{x_P^2 + y_P^2}}{\lambda c} = \frac{\pi t}{\lambda c}\,, \tag{6}$$

where $c = R/(2a)$ is the f-number of the optical system in the image space, $t = \sqrt{x_P^2 + y_P^2}$, and $g = \pi t/c$.

Afterwards, in the case of $A(\rho) = A_0$ where A_0 denotes constant, the amplitude of the wave-field in the plane which is perpendicular to the axis $\overline{\mathrm{OC}}$ and lays in the point C ($\overline{R} \approx \Delta = 0$), can be calculated with Eq. (4) as follows:

$$U_0(P) = -\mathrm{i}\left(\frac{2A_0\,\pi\,a^2}{\lambda R^2}\right)\ \frac{2J_1(\tau)}{\tau}\,. \tag{7}$$

The intensity in such a plane then can be calculated with the formula:

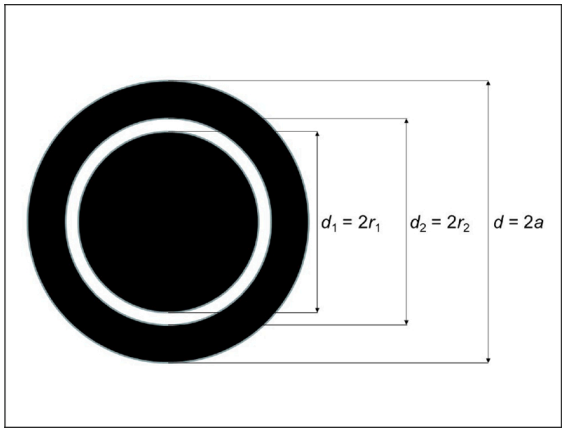


Fig. 3. Scheme of annular pupil of a diameter $d = 2a$ which is transparent in region created by two circles of diameters $d_1 = 2r_1$ and $d_2 = 2r_2$.

$$I_0(\tau) = I_0(P) = |U_0(P)|^2 = \left(\frac{\pi A_0}{2\lambda c^2}\right)^2 \left[\frac{2J_1(\tau)}{\tau}\right]^2. \tag{8}$$

One can see that the image of the point, in the aforementioned situation, is a diffraction distribution called Point Spread Function (PSF), and the image is characterised with a specific energy distribution. It is known [1,2] that the diameter d_A of the central part of the diffraction image (so called Airy disc) is given with the formula: $d_A = 2.4\lambda c$. As an illustration, Fig. 2 shows optical scheme of imaging of a lighting point by an optical system without aberrations ($\Delta = 0$).

In a general case, one gets for the intensity $I(P)$ with Eq. (4):

$$I(\tau, \Delta) = I(P) = |U(P)|^2 \tag{9}$$

$$= K \left| \int_0^1 A(\rho) exp\left(-\mathrm{i}k\frac{\Delta}{8c^2}\rho^2\right) J_0(\tau\rho)\ \rho\,\mathrm{d}\rho \right|^2,$$

where

$$K = \left(\frac{\pi}{2\lambda c^2}\right)^2. \tag{10}$$

For a normalized intensity $I_n(0, \Delta)$ in the centre of the diffraction pattern ($\tau = 0$) one gets the following formula, $A(\rho) = A_0$, $I_n(0,0) = 1$:

$$I_n(0, \Delta) = 4\left[\frac{sin(k\beta/2)}{k\beta}\right]^2, \tag{11}$$

where

$$\beta = \frac{\Delta}{8c^2}. \tag{12}$$

One can simply derive from Eq. (11) that minimal values of intensity $I_n(0, \Delta)$ are located in places with $\Delta = 8n\lambda c^2$, where $n = \pm 1, \pm 2, \ldots$.

3. Imaging of axial point by optical system with annular pupil and uniform amplitude distribution

Suppose now a situation of imaging with an optical system which has an annular pupil of a diameter $d = 2a$, and that this pupil is transparent in annular region created by two circles of diameters $d_1 = 2r_1$ and $d_2 = 2r_2$. The considered pupil is schematically shown in Fig. 3.

Afterwards, instead of Eq. (9), one can write the following formula for the intensity of light $I(\tau, \Delta) = I(P)$:

$$I(\tau, \Delta) = K \left| \int_{\rho_1}^{\rho_2} A(\rho) exp\left(-\mathrm{i}k\frac{\Delta}{8c^2}\rho^2\right) J_0(\tau\rho)\ \rho\,\mathrm{d}\rho \right|^2, \tag{13}$$

where $\rho_1 = r_1/a$ and $\rho_2 = r_2/a$. Then, the normalized intensity distribution $I_n(0, \Delta)$ in the centre of the diffraction pattern can be expressed from Eq. (13) as follows ($\tau = 0, J_0(0) = 1, A(\rho) = A_0$, and $I_n(0,0) = 1$):

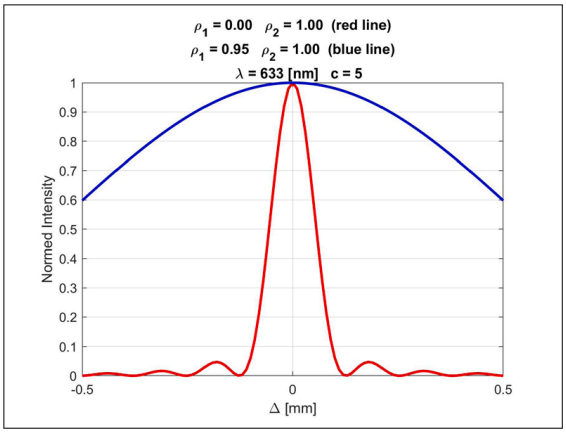


Fig. 4. Normalized intensity distribution $I_n(0,\Delta)$ in the centre of diffraction pattern.

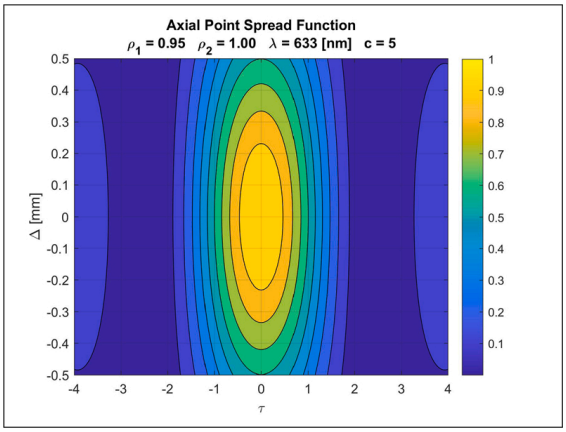


Fig. 5. Normalized intensity distribution $I_n(\tau,\Delta)$ for $\rho_1 = 0.95$ and $\rho_2 = 1.00$.

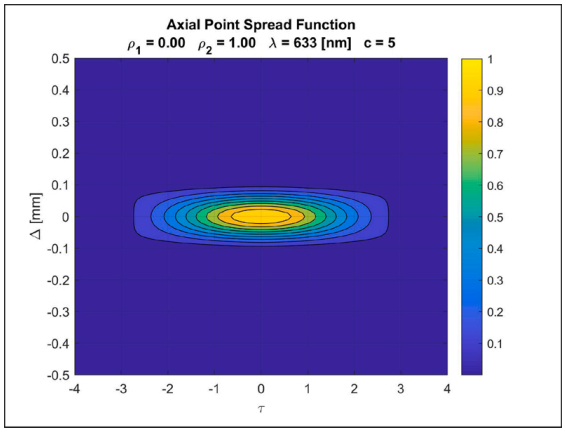


Fig. 6. Normalized intensity distribution $I_n(\tau,\Delta)$ for $\rho_1 = 0$ and $\rho_2 = 1$.

$$I_n(0,\Delta) = \frac{4}{(\rho_2^2-\rho_1^2)^2}\left[\frac{sin[k\beta\,(\rho_2^2-\rho_1^2)/2]}{k\beta}\right]^2. \tag{14}$$

In Fig. 4, there is the normalized intensity $I_n(0,\Delta)$ in the centre of the diffraction pattern shown for the case of $\rho_1 = 0$, $\rho_2 = 1$, and $\rho_1 = 0.95$, $\rho_2 = 1$. Fig. 5 then shows the normalized intensity $I_n(\tau,\Delta)$ for $\rho_1 = 0.95$ and $\rho_2 = 1$, and Fig. 6 presents the normalized intensity $I_n(\tau,\Delta)$ for $\rho_1 = 0$ and $\rho_2 = 1$. In all the figures, the wavelength $\lambda = 633$ nm and the f-number of the optical system $c = 5$.

Suppose now that the argument in the sinus in Eq. (14) is small. Further, one can write an approximate formula for the normalized intensity in the centre of the diffraction pattern (suppose that $sin\alpha \approx \alpha - \alpha^3/6$ for small α), it holds:

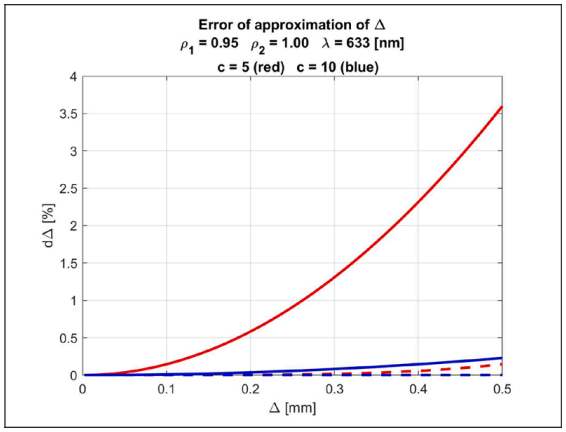


Fig. 7. Relative error dΔ of Δ_m calculated with Eq. (16) (solid lines) and Eq. (18) (dashed lines) from its exact value Δ.

$$I_n(0,\Delta) \approx \left(1 - \frac{\Delta_\rho^2 k^2 \beta^2}{24}\right)^2, \tag{15}$$

where $\Delta_\rho = \rho_2^2 - \rho_1^2$.

If one demands to calculate the minimal value of Δ_m on the axis where the normalized intensity reaches $I_m = I_n(0,\Delta_m)$, then, after substitution from Eq. (12) into Eq. (15), one gets after simplification the following approximate formula:

$$\Delta_m(I_m) = \pm\frac{16c^2}{k\Delta_\rho}\sqrt{6\left(1 - \sqrt{I_m}\right)}. \tag{16}$$

A more accurate results can be obtained considering the following approximation in Eq. (14): $sin\alpha \approx \alpha - \alpha^3/6 + \alpha^5/120$. The normalized intensity in the centre of the diffraction pattern then can be calculated approximately as follows:

$$I_n(0,\Delta) \approx \left(1 - \frac{\Delta_\rho^2 k^2 \beta^2}{24} + \frac{\Delta_\rho^4 k^4 \beta^4}{1920}\right)^2. \tag{17}$$

And substituting from Eq. (12) into Eq. (15) results in:

$$\Delta_m(I_m) = \pm\frac{16c^2}{k\Delta_\rho}\sqrt{10 - 2\sqrt{30\sqrt{I_m} - 5}}, \tag{18}$$

while one gets a real solution only for minimal usable value $I_m = (5/30)^2 \approx 0.028$.

One can see from Eqs. (16) and (18) that the depth of focus on the axis of the optical system, which will be characterised by the value of Δ_m, is inversely proportional to $\Delta_\rho = \rho_2^2 - \rho_1^2$. Considering that the area of the annulus is $\pi(\rho_2^2 - \rho_1^2) = \pi\Delta_\rho$, one can see that the depth of focus on the optical axis will be the same for different optical systems with different annular pupils but of the same areas; and the depth will be, for constant values c and k, inversely proportional to this area of the annular pupil.

Fig. 7 presents relative errors (in percents) of approximate formulas (16) and (18) for calculation of Δ_m from its exact value Δ, for two f-numbers of the optical system $c = 5$ and $c = 10$. One can see that the error of approximation with Eq. (16) is small for larger values of the f-number c, i.e., for optical systems with not too large numerical aperture NA ($c = 1/(2\ \mathrm{NA})$).

Fig. (8) presents normalized intensities of light $I_n(\tau,\Delta)$ for different starting values of ρ_1. The values of ρ_2 are chosen to obtain the same area of the annular pupil as in the case of the example presented in Fig. 5. Further, in Fig. 8, the red lines characterise the depth of focus for the axial region where $I_n(0,\Delta) > 0.75$. One can see that the depth of focus is the same regardless the change in ρ_1. However, the transverse profile of the intensity changes, and transverse minimums of the intensity are closer for larger values of ρ_1.

Consider now the problem of calculation of the first transverse minimum of the PSF in the image plane. Suppose the situation for $\Delta = 0$ and $A(\rho) = A_0$. Afterwards, the intensity described by Eq. (13) in transverse direction can be approximated by a series expansion. One gets after simplification:

$$I(\tau,0) \approx A_0 K \left| \sum_{k=0}^{\infty} \frac{(-1)^k \tau^{2k}}{4^k (2k+2)(k!)^2} \left(\rho_2^{2k+2} - \rho_1^{2k+2}\right) \right|^2 \tag{19}$$

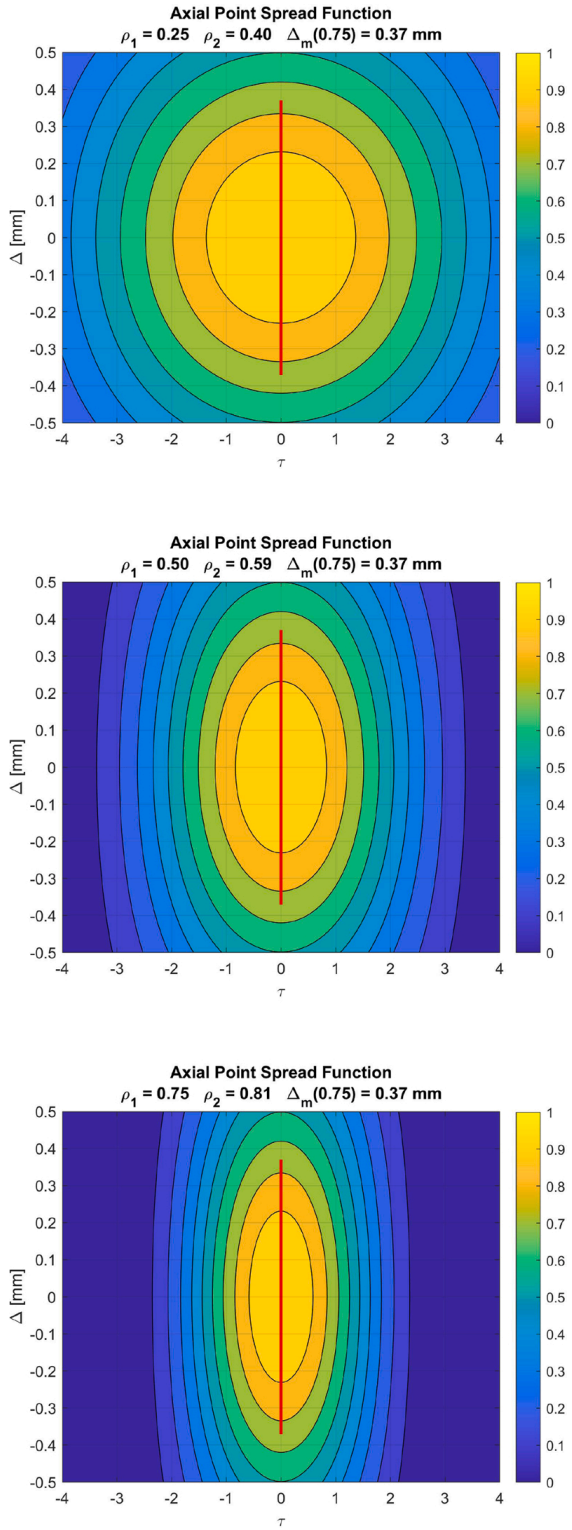


Fig. 8. Normalized intensity distribution $I_n(\tau, \Delta)$ for different values ρ_1 and ρ_2 selected to obtain the same area of the annulus as in Fig. 5 (red lines – axial depth of field for $I_n(0, \Delta) > 0.75$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$= A_0 K \left| \frac{1}{2}\left(\rho_2{}^2 - \rho_1{}^2\right) - \frac{\tau^2}{16}\left(\rho_2{}^4 - \rho_1{}^4\right) + \frac{\tau^4}{384}\left(\rho_2{}^6 - \rho_1{}^6\right) + \ldots \right|^2$$

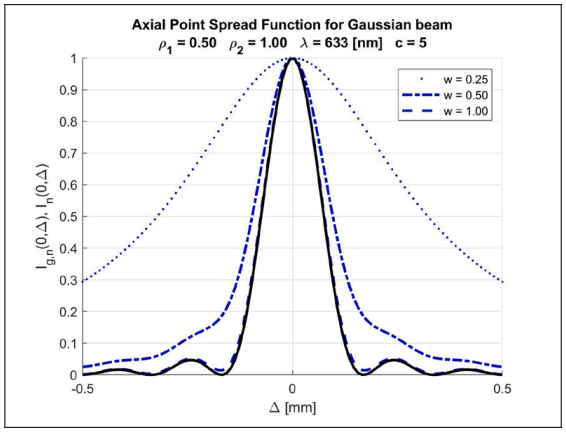


Fig. 9. Effect of the value w on the profile of normalized intensity distribution $I_{g,n}(0,\Delta)$ and comparison with $I_n(0,\Delta)$.

where the series expansion of the Bessel function of the first kind J_0 was used. Considering the first members only, than the transverse value τ_0 of the location of the minimum of the intensity $I(\tau,0)$ has to fulfil the following equation:

$$T_6\,{\tau_0}^6 + T_4{\tau_0}^4 + T_2{\tau_0}^2 + T_0 \;\; = 0 \tag{20}$$

where

$$\begin{aligned} T_6 &= {\rho_1}^6 + {\rho_1}^4{\rho_2}^2 + {\rho_1}^2{\rho_2}^4 + {\rho_2}^6 \\ T_4 &= -48({\rho_1}^4 + {\rho_1}^2{\rho_2}^2 + {\rho_2}^4) \\ T_2 &= 1152({\rho_1}^2 + {\rho_2}^2) \\ T_0 &= -9216. \end{aligned}$$

In the case of circular pupil, i.e., $\rho_1 = 0$, $\rho_1 = 1$, one gets $\tau_0 \approx 3.66$. Compared to a value $\tau_A = (\pi r_A)/(\lambda c) = 1.2\pi \approx 3.77$, where $r_A = d_A/2$ denotes the radius of the Airy disc, it is obvious that the approximate calculation with Eq. (20) gives enough accurate results for the guess of the first minimum of the diffraction pattern (PSF) in transverse direction. For values of ρ_1 and ρ_2 in Fig. 8 one gets: $\rho_1 = 0.25$, $\rho_2 = 0.40$, $\tau_0 \approx 7.27$; $\rho_1 = 0.50$, $\rho_2 = 0.59$, $\tau_0 \approx 4.38$; $\rho_1 = 0.75$, $\rho_2 = 0.81$, $\tau_0 \approx 3.07$. It is obvious that choosing the values of ρ_1 and ρ_2 affects and define the depth of focus of the optical system, and one gets better resolution of the system with larger radii of inner circle of the annular pupil.

4. Imaging of axial point by optical system with annular pupil for Gaussian beams

Consider now a situation when the field in the plane of the pupil of the optical system corresponds to Gaussian beams. In that case, one can write for $A(\rho)$ the following formula [1,2]:

$$A(\rho) = exp\left(-\frac{\rho^2}{w^2}\right), \tag{21}$$

where w is constant. After substitution into Eq. (13) for the intensity generated by the annular pupil, one gets for the Gaussian beam on the optical axis ($\tau = 0$):

$$I_g(0,\Delta) = K\left|\int_{\rho_1}^{\rho_2} exp\left(-\frac{\rho^2}{w^2}\right) exp\left(-\mathrm{i}k\frac{\Delta}{8c^2}\rho^2\right)\;\rho\,\mathrm{d}\rho\right|^2. \tag{22}$$

Further, it holds for an axial distribution of the intensity for $\tau = 0$ after simplification:

$$I_g(0,\Delta) = \frac{K\,w^4}{4\left(\beta^2\,k^2\,w^4+1\right)}\left[G_1^2 + G_2^2 - 2G_1G_2cos(k\beta\Delta_\rho)\right], \tag{23}$$

where

$$\begin{aligned} G_1 &= exp\left(-\frac{{\rho_1}^2}{w^2}\right) \quad G_2 = exp\left(-\frac{{\rho_2}^2}{w^2}\right) \\ \beta &= \frac{\Delta}{8c^2} \quad \Delta_\rho = \rho_2^2 - \rho_1^2. \end{aligned}$$

Afterwards, one gets the following formula for the normalized intensity $I_{g,n}(0,\Delta)$ ($I_{g,n}(0,0) = 1$):

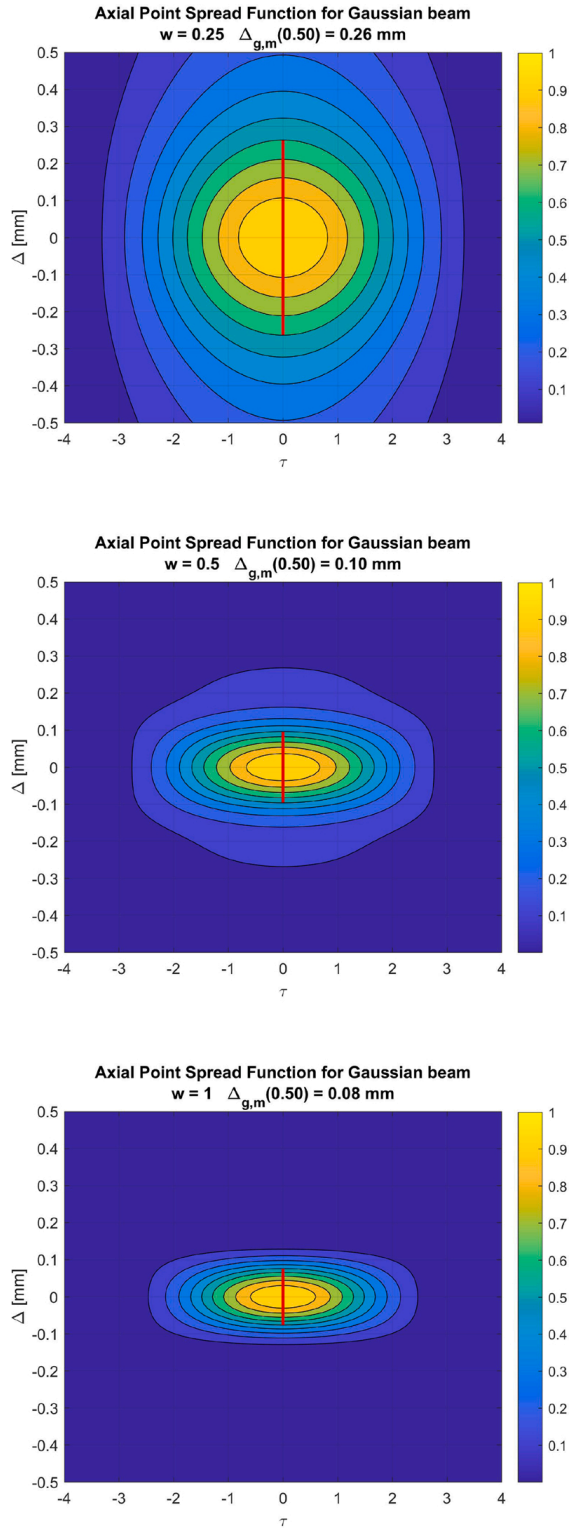


Fig. 10. Normalized intensity distribution $I_{g,n}(\tau,\Delta)$ for Gaussian beam for parameters of the annular pupil $\rho_1 = 0.50$ and $\rho_2 = 1.00$, and $\lambda = 633$ nm, $c = 5$, for different values of w (red lines – axial depth of field for $I_n(0,\Delta) > 0.50$). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$I_{g,n}(0,\Delta) = \frac{G_1^2 + G_2^2 - 2G_1G_2cos(k\beta\Delta_\rho)}{(G_1 - G_2)^2\left(\beta^2 k^2 w^4 + 1\right)}. \tag{24}$$

It can be simply proven that the limit case for $w \to \infty$ gives the formula for $I_n(0,\Delta)$ in Eq. (14). Fig. 9 shows the influence of value w on the profile of the intensity $I_{g,n}(0,\Delta)$ compared to the intensity $I_{g,n}(0,\Delta)$, for the situation of $\rho_1 = 0.50$, $\rho_2 = 1.00$, $\lambda = 633$ nm, and $c = 5$. It is obvious that the lower values of w generate larger depth of focus compared to the uniform amplitude distribution in the

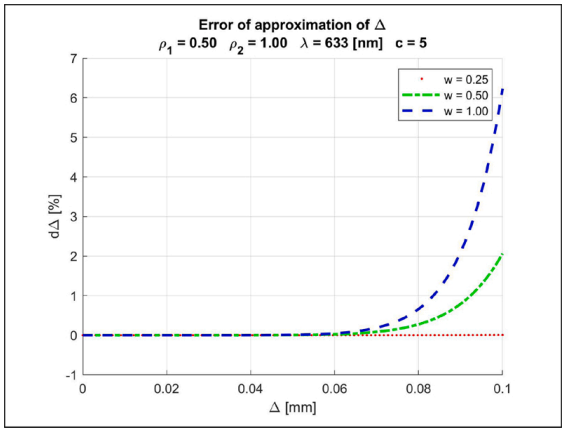


Fig. 11. Relative error dΔ of $\Delta_{g,m}$ calculated with Eq. (26) from its exact value Δ for parameters $\rho_1 = 0.50$, $\rho_2 = 1.00$, $\lambda = 633$ nm, and $c = 5$ for different values of w.

aperture.

An approximate expression of the intensity $I_{g,n}(0,\Delta)$ can be obtained using the Taylor series of the 7-th order in Eq. (24) with respect to the variable β. With such a procedure, and substituting for β, one gets the following approximate formula of the 6-th order:

$$I_{g,n}(0,\Delta) \approx 1 + D_2\Delta^2 + D_4\Delta^4 + D_6\Delta^6\,, \tag{25}$$

where

$$\begin{aligned} D_2 &= -\left(\frac{k}{8c^2}\right)^2 \left[w^4 - \frac{{\Delta_\rho}^2 G_1 G_2}{(G_1 - G_2)^2}\right], \\ D_4 &= \left(\frac{k}{8c^2}\right)^4 \left[w^8 - \frac{{\Delta_\rho}^2 G_1 G_2 \left({\Delta_\rho}^2 + 12\,w^4\right)}{12\,(G_1 - G_2)^2}\right], \\ D_6 &= -\left(\frac{k}{8c^2}\right)^6 \left[w^{12} - \frac{{\Delta_\rho}^2 G_1 G_2 \left({\Delta_\rho}^4 + 30\,{\Delta_\rho}^2 w^4 + 360\,w^8\right)}{360\,(G_1 - G_2)^2}\right]. \end{aligned}$$

Afterwards, to calculate an approximate value of $\Delta_{g,m}$, where the axial intensity distribution reaches the value of $I_{g,m} = I_{g,n}(0,\Delta_{g,m})$, $\Delta_{g,m}$ has to fulfil the following equation:

$$(1 - I_{g,m}) + D_2{\Delta_{g,m}}^2 + D_4{\Delta_{g,m}}^4 + D_6{\Delta_{g,m}}^6 = 0\,. \tag{26}$$

Fig. 10 presents the axial PSFs of the optical system for specific cases of the Gaussian beam in the plane of pupil, together with the depth of focus for $I_{g,n}(0,\Delta) > 0.50$. Afterwards, Fig. 11 shows relative errors of $\Delta_{g,m}$ from its accurate values Δ, which are calculated with Eq. (26) for $\rho_1 = 0.50$, $\rho_2 = 1.00$, $\lambda = 633$ nm, $c = 5$, and different values of w. One can see that for larger values of w, the approximation is appropriate for limited values of Δ only. Fig. 10 shows that as the profile of the intensity is closer to the one of the corresponding uniform amplitude distribution in the pupil, the approximation with the 6-th order (Eq. (25)) is accurate only for a close central region of the APSF, i.e., for Δ close to zero. The more accurate results for $\Delta_{g,m}$ for a non-central region of the APSF can be calculated numerically. An estimation of the relative error $\delta I_{g,n}$ of the approximation with Eq. (25) can be calculated with next member of the Taylor series of Eq. (24), one gets:

$$\delta I_{g,n} \approx \frac{1}{I_{g,n}}\left(\frac{k}{8c^2}\right)^8 \left[w^{16} - \frac{{\Delta_\rho}^2 G_1 G_2\,\left({\Delta_\rho}^6 + 56\,{\Delta_\rho}^4 w^4 + 1680\,{\Delta_\rho}^2 w^8 + 20160\,w^{12}\right)}{20160\,(G_1 - G_2)^2}\right]. \tag{27}$$

5. Conclusion

The paper presented an analysis of the 3D PSF (Point Spread Function) and APSF (Axial Point Spread Function) for the case of a circular pupil and the case of an annular pupil. The situation of a uniform amplitude distribution and of a Gaussian beams was studied, and new analytic formulas (Eqs. (16), (18), (20), and (26)) were derived which characterise properties of the depth of focus of the optical system. Those novel formulas have not been presented in previous works, and they are a useful completion of theory and analysis of the APSF.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

This work was supported by the Grant Agency of the Czech Technical University in Prague, Grant No. SGS20/093/OHK1/2T/11.

References

[1] M. Born, E. Wolf, Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, Cambridge University Press, 1999. https://books.google.cz/books?id=nUHGpfNsGyUC.
[2] A. Mikš, Applied Optics, CTU in Prague, 2009.
[3] A. Miks, J. Novak, P. Novak, Calculation of point-spread function for optical systems with finite value of numerical aperture, Optik 118 (11) (2006), https://doi.org/10.1016/j.ijleo.2006.05.002.
[4] A.S. Marathay, G.B. Parrent, Use of scalar theory in optics, J. Opt. Soc. Am. 60 (2) (1970), https://doi.org/10.1364/josa.60.000243.
[5] D.G. Flagello, A.E. Rosenbluth, Lithographic tolerances based on vector diffraction theory, J. Vacuum Sci. Technol. B 10 (6) (1992), https://doi.org/10.1116/1.585959.
[6] J.J. Stamnes, H. Heier, Scalar and electromagnetic diffraction point-spread functions, Appl. Opt. 37 (17) (1998), https://doi.org/10.1364/AO.37.003612.
[7] C.W. McCutchen, Generalized aperture and the three-dimensional diffraction image, J. Opt. Soc. Am. (2) (1964), https://doi.org/10.1364/josa.54.000240.
[8] J.C. Heurtley, Scalar Rayleigh-Sommerfeld and Kirchhoff diffraction integrals: a comparison of exact evaluations for axial points, J. Opt. Soc. Am. 63 (8) (1973), https://doi.org/10.1364/josa.63.001003.
[9] H. Osterberg, L.W. Smith, Closed solutions of Rayleigh's diffraction integral for axial points, J. Opt. Soc. Am. 51 (10) (1961), https://doi.org/10.1364/josa.51.001050.
[10] Y. Li, Predictions of Rayleigh's diffraction theory for the effect of focal shift in high-aperture systems, J. Opt. Soc. Am. A 25 (7) (2008), https://doi.org/10.1364/josaa.25.001835.
[11] C.J.R. Sheppard, P. Török, Focal shift and the axial optical coordinate for high-aperture systems of finite Fresnel number, J. Opt. Soc. Am. A 20 (11) (2003), https://doi.org/10.1364/josaa.20.002156.
[12] P. Jacquinot, B. Roizen-Dossier, Apodisation. Progress in Optics, Elsevier, 1964, https://doi.org/10.1016/s0079-6638(08)70570-5.
[13] A.N.K. Reddy, M. Hashemi, Apodization pupils: design and performance (review), J. Phys. Conf. Ser. 1096 (2018), https://doi.org/10.1088/1742-6596/1096/1/012140.
[14] J. Ojeda-Castañeda, C.M. Gómez-Sarabia, Tuning field depth at high resolution by pupil engineering, Adv. Opt. Photon. 7 (4) (2015), https://doi.org/10.1364/aop.7.000814.
[15] L. Lalaoui, M. Bouafia, D. Issaad, A. Medjahed, Axial point spread function modeling by application of effective medium approximations: Widefield microscopy, Optik 212 (2020), https://doi.org/10.1016/j.ijleo.2020.164646.
[16] A.M. Hamed, Computation of the lateral and axial point spread functions in confocal imaging systems using binary amplitude mask, Pramana 66 (6) (2006), https://doi.org/10.1007/BF02708457.
[17] M. Martinez-Corral, M. Caballero, E.H.K. Stelzer, J. Swoger, Tailoring the axial shape of the point spread function using the Toraldo concept, Opt. Express 10 (1) (2002) 98, https://doi.org/10.1364/oe.10.000098.
[18] M. Martínez-Corral, L. Muñoz-Escrivá, M. Kowalczyk, T. Cichocki, One-dimensional iterative algorithm for three-dimensional point-spread function engineering, Opt. Lett. 26 (23) (2001) 1861, https://doi.org/10.1364/ol.26.001861.
[19] G. Watson, A Treatise on the Theory of Bessel Functions, Cambridge Mathematical Library, Cambridge University Press, 1995.

Optik - International Journal for Light and Electron Optics 201 (2020) 163483

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.com/locate/ijleo

Original research article

Analytic formulas for calculation of axial defocus in microscopy

Antonín Mikš, Petr Pokorný*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic

ARTICLE INFO

Keywords:
Microscopy
Imaging
Defocus

ABSTRACT

The paper analyses a possibility of determination of axial defocus of lighting point, which is imaged by an optical system without aberrations, from a knowledge of intensity distribution or an amount of light (amount of energy – fractional total energy) detected in a region around a point's image centre. Analytic formulas for calculation of the defocus are derived, and they are verified by numerical examples, simulations and experiment. Outputs show that the method can evaluate the axial depth of the particle up to several micro-metres with accuracy in a range of fraction of percents.

1. Introduction

In microscopy [1–10], there are many cases where one studies an object which composes of practically point micro-particles, for example bacteria, coloured segments of structures with fluorochromes in the fluorescence microscopy, or similar. These micro-particles are located in positions in an object space which are given by their spatial coordinates (x, y, z). It is quite simple to determine transverse coordinates (x, y). This problematic is studied by localization microscopy, and one can found detailed information in Birk's book [8] for example. However, accurate determination of axial z-coordinate, or axial defocusation, is still challenging.

The goal of this paper is to analyse a possibility of determination of the axial defocus of lighting point, which is imaged by an optical system without aberrations, from a knowledge of intensity distribution or an amount of light (amount of energy – fractional total energy) detected in a region around a point's image centre. To the authors best knowledge, this approach has not been analytically studied and presented in literature yet.

Formulas for calculation of amount of light detected in a circle are known, see for example Martin's [1] or Born and Wolf's book [2]. However, there is no solution for the axial longitudinal defocus of the lighting point presented. Franke et al. [11] present a solution of this issue only experimentally, the authors used area of circular ring between two radii around an image of point particle. Fuchs et al. [12] show theoretical and experimental image formation study in the presence of astigmatic aberrations, and macroscopic location scheme of micrometer-sized particles for the single camera astigmatism particle tracking velocimetry (APTV) technique.

The first part of this paper is focused on imaging with a physically ideal optical system without aberrations and analysis of a diffraction of convergent spherical wave-front. The optical system without aberrations is justly selected as all microscope objectives are constructed with such an image quality that one can neglect effects of aberrations on image. Formulas for calculation of an intensity distribution in a plane located in the centre of convergent spherical wave-front are presented both for circular and rectangular aperture (detector). Relations for the intensity distribution in an axially displaced (defocused) plane are then shown as well. As will be derived, analysis of an intensity and illumination distribution can give information about the distance between mentioned

* Corresponding author.
E-mail address: petr.pokorny@fsv.cvut.cz (P. Pokorný).

https://doi.org/10.1016/j.ijleo.2019.163483
Received 20 June 2019; Accepted 24 September 2019

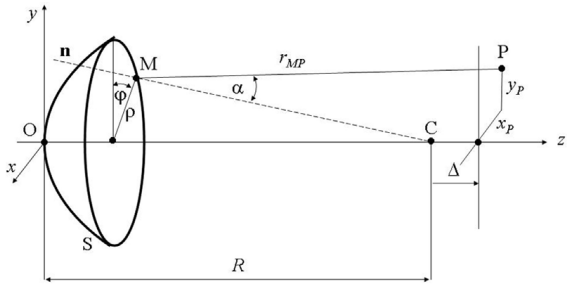


Fig. 1. Scheme for a calculation of spherical wave diffraction.

planes – the defocus. Energy invariance than helps with practical application in microscopy for example, where the intensity distribution is registered on a plane of a detector.

The second part of the paper presents numerical examples and verify derived formulas. The formulas are simply applicable in practice which has not been possible in previous published works. The final part than shows simple experimental verification of the presented approach.

2. Optical imaging of axial points and calculation of axial defocus

Suppose now a physically ideal optical system, i.e. optical system without aberrations. It can be justly selected as all microscope objectives are constructed with such an image quality that the image is not distorted at all and only diffraction effects come into consideration.

Let a scalar wave-filed amplitude $U(\mathrm{M})$ on a spherical wave-front S in an image space of the optical system is known (see Fig. 1). It is known from the theory of optical imaging [1,2] that the amplitude $U(\mathrm{P})$ of a wave-field in a point P in the image space of the optical system is given with formula:

$$U(\mathrm{P}) = -\frac{\mathrm{i}}{\lambda}\iint_S U(\mathrm{M})\frac{\exp(\mathrm{ikr_{MP}})}{r_{\mathrm{MP}}}\cos\alpha\,\mathrm{d}S\,, \tag{1}$$

where M is the point on the wave-front S, r_{MP} is the distance between points P and M, λ is the wave-length, $k = 2\pi/\lambda$ is the wave-number, and α is the angle between the inner unit normal vector $\mathbf{n}$ of the wave-front S in the point M and the vector $\mathbf{r}$ (between points M and P) and i denotes the imaginary unit.

Suppose next that the amplitude $U(\mathrm{M})$ is a convergent spherical wave of the radius R having its centre in the point C and having the amplitude in a unit distance from C equal to A, therefore $U(\mathrm{M}) = A\exp(-\mathrm{ikR})/R$.

Let the point P is located in a plane perpendicular to the axis $\bar{\mathrm{OC}}$,which is distanced by a Δ from the point C. Therefore, coordinates of the point P with respect to the point O are $[x_P,\ y_P,\ R+\Delta]$. Coordinates of the point M with respect to the point O are $[\rho\sin\varphi,\ \rho\cos\varphi,\ z_M]$, where

$$z_M(2R - z_M) = \rho^2, \quad \rho = \sqrt{x_M^2 + y_M^2}\,, \quad \cos\varphi = y_M/\rho, \quad \sin\varphi = x_M/\rho. \tag{2}$$

The distance r_{MP} can be calculated as follows, according Fig. 1 ($r_{\mathrm{OP}} = \bar{\mathrm{OP}}$),

$$\begin{aligned} r_{\mathrm{MP}} &= r_{\mathrm{OP}} - \Delta\frac{\rho^2}{2R^2} - \frac{\rho}{R}(x_P\sin\varphi + y_P\cos\varphi) \\ &= r_{\mathrm{OP}} - \Delta\frac{x_M^2 + y_M^2}{2R^2} - \frac{1}{R}(x_P x_M + y_P y_M). \end{aligned} \tag{3}$$

Let the values of Δ, x_P and y_P are significantly smaller than the wave-front's radius R. Afterwards, one can consider the angle α very small and it holds: $\cos\alpha \approx 1$, and $\exp(\mathrm{ikr_{MP}})/r_{\mathrm{MP}} \approx \exp(\mathrm{ikr_{MP}})/R$. Moreover, for areas close to the optical axis one can write the following formula:

$$r_{\mathrm{OP}} - R \approx \Delta + \frac{\Delta^2 + x_P^2 + y_P^2}{2R} \approx \Delta\,. \tag{4}$$

Substituting the afore-mentioned assumptions into Eq. (1), one can express approximate analytic formulas for the complex amplitude $U(\mathrm{P})$ in point P. Afterwards, an intensity distribution $I(\mathrm{P})$ for points P and an illumination $L(\tau_P)$ (time-averaged mean amount of energy) for a region τ_P will be given with formulas:

$$I(\mathrm{P}) = U(\mathrm{P})U^*(\mathrm{P})\,, \quad L(\tau) = \iint_{\tau_P} I(\mathrm{P})\,\mathrm{d}\tau_P\,, \tag{5}$$

where * denotes complex conjugate. The following part of the paper discusses situations of circular and rectangular aperture.

2.1. Calculation of axial defocus for circular aperture

Suppose now that the optical system is rotational symmetric with a circular pupil. One can then write a modified formula for the amplitude $U(\mathrm{P})$ in point P. With the assumptions mentioned in Eq. (4) one can modify the formula for the complex amplitude in the case of circular aperture as follows [1,2]:

$$U(\mathrm{P}) = -\mathrm{i}\frac{a^2}{R^2}\mathrm{kA}\exp(\mathrm{i}k\Delta)\int_0^1 \exp\left(-\mathrm{i}k\frac{a^2\Delta}{2R^2}r^2\right)J_0(\tau r)\,r\mathrm{d}r\,, \tag{6}$$

where $a = \rho_{\max}$ is the maximal value of ρ (see Fig. 1), and J_0 is the Bessel function of the first kind. If one defines new variable:

$$\tau = \frac{2\pi}{\lambda}\left(\frac{a}{R}\right)\sqrt{x_P^2 + y_P^2} = \frac{\pi\sqrt{x_P^2 + y_P^2}}{\lambda c} = \frac{\pi t}{\lambda c} = \frac{g}{\lambda}\,, \tag{7}$$

where $c = R/(2a)$ is the f-number of the optical system in the image space, $t = \sqrt{x_P^2 + y_P^2}$ and $g = \pi t/c$. The amplitude of the field in the plane perpendicular to the axis $\bar{\mathrm{OC}}$ crossing the centre C ($\Delta \to 0$) then can be calculated from Eq. (6), one gets:

$$U_{\Delta\to 0}(\tau) = -\mathrm{i}\left(\frac{a^2}{R^2}\mathrm{kA}\right)\frac{2J_1(\tau)}{\tau}\,, \tag{8}$$

where the position of point P is expressed with variable τ and one supposes rotational symmetry of the image. Corresponding intensity distribution of the field is then given by formula:

$$I_{\Delta\to 0}(\tau) = U_{\Delta\to 0}(\tau)U^*_{\Delta\to 0}(\tau) = \left(\frac{a^2}{R^2}\mathrm{kA}\right)^2\left[\frac{2J_1(\tau)}{\tau}\right]^2. \tag{9}$$

One can see from Eq. (9) that image of the point is not a point, but it is a diffraction energy distribution — Point Spread Function. It is known from literature [1,2] that the diameter d_A of the central part of the diffraction image is so called Airy disc: $d_A = 2.4\lambda c$. In general case, one can use Eq. (6) and the intensity $I(\mathrm{P})$ in the point P can be calculated as follows:

$$I(\tau,\,\Delta) = \left(\frac{a^2}{R^2}\mathrm{kA}\right)^2\left|\int_0^1 \exp(-\mathrm{i}k\beta r^2)J_0(\tau r)r\mathrm{d}r\right|^2, \tag{10}$$

where

$$\beta = \frac{\Delta}{8c^2}\,. \tag{11}$$

An amount of light $\mathrm{d}L$ in a circular ring of the radius τ and width $\mathrm{d}\tau$ is given by formula:

$$\mathrm{d}L = 2\pi\tau\mathrm{d}\tau I(\tau,\,\Delta)\,. \tag{12}$$

Afterwards, an illumination $L = L(\tau_M,\,\Delta)$ in a circle of the radius τ_m can be calculated as follows:

$$L(\tau_M,\,\Delta) = 2\pi\int_0^{\tau_m} I(\tau,\,\Delta)\tau\,\mathrm{d}\tau\,. \tag{13}$$

Suppose now that the defocus Δ is small, i.e. the value of β is small as well. Afterwards, one can approximately write:

$$\exp(-\mathrm{i}k\beta r^2) \approx 1 - \mathrm{i}k\beta r^2 - \frac{1}{2}k^2\beta^2 r^4\,. \tag{14}$$

Afterwards, the complex amplitude in Eq. (6) can be re-written as follows:

$$U(\tau,\,\Delta) \approx -\mathrm{i}\frac{a^2}{R^2}\mathrm{kA}\exp(\mathrm{i}k\Delta)\int_0^1\left(1 - \mathrm{i}k\beta r^2 - \frac{1}{2}k^2\beta^2 r^4\right)J_0(\tau r)\,r\mathrm{d}r\,. \tag{15}$$

It is generally known that the following formulas hold:

$$I_1(\tau) = \int_0^1 rJ_0(\tau r)\,\mathrm{d}r = \frac{J_1(\tau)}{\tau}\,, \tag{16}$$

$$I_3(\tau) = \int_0^1 r^3J_0(\tau r)\,\mathrm{d}r = \frac{2\tau J_0(\tau) + (\tau^2 - 4)J_1(\tau)}{\tau^3}\,, \tag{17}$$

$$I_5(\tau) = \int_0^1 r^5J_0(\tau r)\,\mathrm{d}r = \frac{(\tau^2 - 8)[4\tau J_0(\tau) + (\tau^2 - 8)J_1(\tau)]}{\tau^5}\,. \tag{18}$$

Therefore, one can modify Eq. (15) to the form:

$$U(\tau,\,\Delta) \approx -\mathrm{i}\frac{a^2}{R^2}\mathrm{kA}\exp(\mathrm{i}k\Delta)\left[I_1(\tau) - \mathrm{i}k\beta I_3(\tau) - \frac{1}{2}(k\beta)^2 I_5(\tau)\right]. \tag{19}$$

An approximate formula for the intensity distribution then follows:

$$I(\tau, \Delta) \approx \left(\frac{a^2}{R^2}\mathrm{kA}\right)^2 [I_1^2(\tau) + (k\beta)^2(I_3^2(\tau) - I_1(\tau)I_5(\tau))]\,. \tag{20}$$

Suppose next that the normalized intensity is equal to one for a limit case $\tau \to 0$ and $\Delta \to 0$, i.e.:

$$\lim_{\tau\to 0,\Delta\to 0} I(\tau, \Delta) = \frac{1}{4}\left(\frac{a^2}{R^2}\mathrm{kA}\right)^2 = 1\,. \tag{21}$$

Afterwards, the following formula for the normalized intensity $I_n(\tau, \Delta)$ holds:

$$I_n(\tau, \Delta) \approx 4[I_1^2(\tau) + (k\beta)^2(I_3^2(\tau) - I_1(\tau)I_5(\tau))]\,. \tag{22}$$

Normalization of the illumination L_n in the circle of the radius τ_m can be calculated for a condition $L_n(\infty, 0) = 1$. It holds:

$$L_n(\tau_m, \Delta) = \frac{1}{2}\int_0^{\tau_m} I_n(\tau, \Delta)\tau\,\mathrm{d}\tau\,. \tag{23}$$

Substituting Eq. (22) in Eq. (23), one gets the normalized illumination L_n as an approximation:

$$L_n(\tau_m, \Delta) \approx 2(k\beta)^2\left[\frac{8J_1^2(\tau_m)}{\tau_m^4} - \frac{2(J_0^2(\tau_m) + J_1^2(\tau_m))}{\tau_m^2}\right] - [J_0^2(\tau_m) + J_1^2(\tau_m) - 1]\,. \tag{24}$$

If the normalized intensity distribution $I_n(\tau_m, \Delta)$ is known, one derives the defocus $\Delta = \Delta(I_n(\tau, \Delta))$ from Eqs. (22) and (11), it holds:

$$\Delta = \frac{4c^2}{k}\sqrt{\frac{I_n(\tau, \Delta) - 4I_1^2(\tau)}{I_3^2(\tau) - I_1(\tau)I_5(\tau)}}\,. \tag{25}$$

After modification of Eq. (24), one derives the formula for the defocus $\Delta = \Delta(L_n(\tau_m, \Delta))$ as a function of normalized illumination $L_n(\tau_m, \Delta)$, it holds:

$$\Delta = \frac{4c^2\tau_m^2}{k}\sqrt{\frac{1 - J_0^2(\tau_m) - J_1^2(\tau_m) - L_n(\tau_m, \Delta)}{\tau_m^2[J_0^2(\tau_m) + J_1^2(\tau_m)] - 4J_1^2(\tau_m)}}\,. \tag{26}$$

Eqs. (25) and (26) are approximate solutions of the problem for small values of Δ. In the case of large defocus, there is no simple analytic solution. Nevertheless, there are other options. One can pre-calculate nominal values of L_n for series of values Δ and τ_m with exact solutions in Eqs. (10) and (13) and approximate function $L_n(\tau_m, \Delta)$ with appropriate polynomials. Different approach can be based on optimization algorithms where a goal function implements Eqs. (10) or (13).

2.2. Calculation of axial defocus for rectangular aperture

Suppose that the aperture contributing to imaging into point P has a rectangular shape with width $2a$ and height $2b$. With assumptions shown in Eqs. (2)–(4), one can re-write, after several simple modifications, Eq. (1) as follows:

$$\begin{aligned} U(\mathrm{P}) &\approx -\mathrm{i}\frac{A}{\lambda R^2}\int_{-b}^{b}\int_{-a}^{a}\exp\left\{\mathrm{i}k\left[\Delta\left(1 - \frac{x_M^2 + y_M^2}{2R^2}\right) - \frac{1}{R}(x_P x_M + y_P y_M)\right]\right\}\mathrm{d}x_M\,\mathrm{d}y_M \\ &= -\mathrm{i}\frac{A}{\lambda R^2}\exp(\mathrm{i}k\Delta)\,U_x\,U_y\,, \end{aligned} \tag{27}$$

where

$$U_x = \int_{-a}^{a}\exp[\mathrm{i}(\alpha x_M^2 + \beta_x x_M)]\,\mathrm{d}x_M\,, \qquad U_y = \int_{-b}^{b}\exp[\mathrm{i}(\alpha y_M^2 + \beta_y y_M)]\,\mathrm{d}y_M\,,$$
$$\alpha = -k\frac{\Delta}{2R^2}\,, \qquad \beta_x = -k\frac{x_P}{R}\,, \quad \beta_y = -k\frac{y_P}{R}\,.$$

Solution of integrals in Eqs. (27) can be found using Fresnel integrals [13–15]. The following formulas can be easily derived or they are known from differential and integral calculus:

$$\begin{aligned} &\int_{\varepsilon_1}^{\varepsilon_1}\exp[\mathrm{i}(\alpha t^2 + \beta t)]\,\mathrm{d}t = \frac{1}{\sqrt{\alpha}}\exp\left(-\mathrm{i}\frac{\beta^2}{4\alpha}\right)\int_{\varepsilon_1\sqrt{\alpha}+\frac{\beta}{2\alpha}}^{\varepsilon_2\sqrt{\alpha}+\frac{\beta}{2\alpha}}\exp(\mathrm{i}\nu^2)\,\mathrm{d}\nu, \\ &\int_{E_1}^{E_2}\exp(\mathrm{i}\nu^2)\,\mathrm{d}\nu = \int_0^{E_2}\exp(\mathrm{i}\nu^2)\,\mathrm{d}\nu - \int_0^{E_1}\exp(\mathrm{i}\nu^2)\,\mathrm{d}\nu = F(E_2) - F(E_1)\,, \\ &F(\tau) = \int_0^{\tau}\exp(\mathrm{i}\xi^2)\,\mathrm{d}\xi = \sqrt{\frac{\pi}{2}}\left[C\left(\tau\sqrt{\frac{2}{\pi}}\right) + \mathrm{i}S\left(\tau\sqrt{\frac{2}{\pi}}\right)\right], \end{aligned} \tag{28}$$

where C and S denote Fresnel integrals and α, β, E_1 and E_2 are constants. Afterwards, one can write:

$$U_x = \frac{1}{\sqrt{\alpha}} \exp\left(-\mathrm{i}\frac{\beta_x^2}{4\alpha}\right)[F(E_{2x}) - F(E_{1x})]\,, \tag{29}$$

$$U_y = \frac{1}{\sqrt{\alpha}} \exp\left(-\mathrm{i}\frac{\beta_y^2}{4\alpha}\right)[F(E_{2y}) - F(E_{1y})]\,,$$

$$E_{1x} = -a\sqrt{\alpha} + \frac{\beta_x}{2\alpha} = -a\sqrt{\alpha} + \frac{R}{\Delta}\frac{\lambda c}{\pi}\tau_x\,, \quad E_{2x} = a\sqrt{\alpha} + \frac{\beta_x}{2\alpha} = a\sqrt{\alpha} + \frac{R}{\Delta}\frac{\lambda c}{\pi}\tau_x\,,$$

$$E_{1y} = -b\sqrt{\alpha} + \frac{\beta_y}{2\alpha} = -b\sqrt{\alpha} + \frac{R}{\Delta}\frac{\lambda c}{\pi}\tau_y\,, \quad E_{2y} = b\sqrt{\alpha} + \frac{\beta_y}{2\alpha} = b\sqrt{\alpha} + \frac{R}{\Delta}\frac{\lambda c}{\pi}\tau_y\,,$$

$$\tau_x = \frac{\pi}{\lambda c}x_P\,, \quad \tau_y = \frac{\pi}{\lambda c}y_P\,,$$

where $c = R/(2\sqrt{a^2 + b^2})$ is the f-number of the optical system in the image space. Intensity in point P than can be calculated as follows:

$$I(\mathrm{P}, \Delta) = U(\mathrm{P}, \Delta)U^*(\mathrm{P}, \Delta) = \left(\frac{A}{\lambda R^2}\right)^2 \frac{1}{\alpha^2}[F(E_{2x}) - F(E_{1x})]^2[F(E_{2y}) - F(E_{1y})]^2 = \frac{A^2}{\pi^2\Delta^2}[F(E_{2x}) - F(E_{1x})]^2[F(E_{2y}) - F(E_{1y})]^2. \tag{30}$$

Similarly as in the case of circular aperture, one can approximate the afore-mentioned formulas for small defocuses, i.e. $\alpha \to 0$, and $\exp(\mathrm{i}\alpha t^2) \approx 1 + \mathrm{i}\alpha t^2 - \frac{1}{2}\alpha^2 t^4$. Therefore, U_x and U_y from Eqs. (27) can be approximated as follows:

$$U_x \approx \int_{-a}^{a}\left(1 + \mathrm{i}\alpha x_M^2 - \frac{1}{2}\alpha^2 x_M^4\right)\exp(\mathrm{i}\beta_x x_M)\,\mathrm{d}x_M$$

$$= \sin(a\beta_x)\left[\frac{1}{\beta_x}(2 - a^4\alpha^2 + 2\mathrm{i}a^2\alpha) + \frac{4}{\beta_x^3}(3a^2\alpha^2 - \mathrm{i}\alpha) - \frac{24}{\beta_x^5}\alpha^2\right] + \cos(a\beta_x)\left[\frac{4}{\beta_x^2}(-a^3\alpha^2 + \mathrm{i}a\alpha) + \frac{24}{\beta_x^4}a\alpha^2\right],$$

$$U_y \approx \int_{-b}^{b}\left(1 + \mathrm{i}\alpha y_M^2 - \frac{1}{2}\alpha^2 y_M^4\right)\exp(\mathrm{i}\beta_y y_M)\,\mathrm{d}y_M$$

$$= \sin(b\beta_y)\left[\frac{1}{\beta_y}(2 - b^4\alpha^2 + 2\mathrm{i}b^2\alpha) + \frac{4}{\beta_y^3}(3b^2\alpha^2 - \mathrm{i}\alpha) - \frac{24}{\beta_y^5}\alpha^2\right] + \cos(b\beta_y)\left[\frac{4}{\beta_y^2}(-b^3\alpha^2 + \mathrm{i}b\alpha) + \frac{24}{\beta_y^4}b\alpha^2\right]. \tag{31}$$

The intensity in point P than will be given by the formula:

$$I(\mathrm{P}, \Delta) = U(\mathrm{P}, \Delta)U^*(\mathrm{P}, \Delta) = \left(\frac{A}{\lambda R^2}\right)^2 U_x U_x^* U_y U_y^* \approx \left(\frac{A}{\lambda R^2}\right)^2 (I_{x,4}\Delta^4 + I_{x,2}\Delta^2 + I_{x,0})(I_{y,4}\Delta^4 + I_{y,2}\Delta^2 + I_{y,0}), \tag{32}$$

where

$$I_{x,4} = \left(\frac{k}{2R^2}\right)^4 \frac{1}{\beta_x^{10}}[(a^4\beta_x^4 - 12a^2\beta_x^2 + 24)\sin(a\beta_x) + 4(a^3\beta_x^3 - 6a\beta_x)\cos(a\beta_x)]^2\,,$$

$$I_{x,2} = \left(\frac{k}{2R^2}\right)^2 \frac{1}{\beta_x^6}[(40 - 8a^2\beta_x^2)\cos(2a\beta_x) + 32a\beta_x\sin(2a\beta_x) + 24a^2\beta_x^2 - 40]\,,$$

$$I_{x,0} = \frac{4}{\beta_x^2}\sin^2(a\beta_x)\,,$$

$$I_{y,4} = \left(\frac{k}{2R^2}\right)^4 \frac{1}{\beta_y^{10}}[(b^4\beta_y^4 - 12b^2\beta_y^2 + 24)\sin(b\beta_y) + 4(b^3\beta_y^3 - 6b\beta_y)\cos(b\beta_y)]^2\,,$$

$$I_{y,2} = \left(\frac{k}{2R^2}\right)^2 \frac{1}{\beta_y^6}[(40 - 8b^2\beta_y^2)\cos(2b\beta_y) + 32b\beta_y\sin(2b\beta_y) + 24b^2\beta_y^2 - 40]\,,$$

$$I_{y,0} = \frac{4}{\beta_y^2}\sin^2(b\beta_y)\,.$$

Next, let the normalized intensity is equal to one for a limit case $\beta_x \to 0$, $\beta_y \to 0$, and $\Delta \to 0$. Afterwards, one can derive the following formulas from Eq. (32):

$$
\begin{aligned}
&I_{\Delta\to 0}(\mathrm{P},\,\Delta) = \lim_{\Delta\to 0} I(\mathrm{P},\,\Delta)\\
&= \left(\frac{A}{\lambda R^2}\right)^2 I_{x,0} I_{y,0} = \left(\frac{A}{\lambda R^2}\right)^2 \frac{16}{\beta_x^2\beta_y^2}\sin^2(a\beta_x)\sin^2(b\beta_y),\\
&I_0 = \lim_{\beta_x\to 0,\beta_y\to 0} I_{\Delta\to 0}(\mathrm{P},\,\Delta) = 16a^2b^2\left(\frac{A}{\lambda R^2}\right)^2.
\end{aligned}
\tag{33}
$$

Eqs. (33) presents generally known formulas of diffraction theory [2]. Therefore, the afore-mentioned derivations leads to correct values and the procedure is proved. The normalized intensity $I_n(\mathrm{P},\,\Delta)$ than can be described as follows:

$$
I_n(\mathrm{P},\,\Delta) \approx \frac{1}{16a^2b^2}(I_{x,4}\Delta^4 + I_{x,2}\Delta^2 + I_{x,0})(I_{y,4}\Delta^4 + I_{y,2}\Delta^2 + I_{y,0})\,. \tag{34}
$$

At this moment, one can find solution for the defocus Δ by calculating roots of the polynomial of the 8-th order which is given by modification of Eq. (34):

$$
\begin{aligned}
&16a^2b^2 I_n(\mathrm{P},\,\Delta) = I_{x,4}I_{y,4}\Delta^8 + (I_{x,4}I_{y,2} + I_{x,2}I_{y,4})\Delta^6\\
&+ (I_{x,4}I_{y,0} + I_{x,2}I_{y,2} + I_{x,0}I_{y,4})\Delta^4 + (I_{x,2}I_{y,0} + I_{x,0}I_{y,2})\Delta^2 + I_{x,0}I_{y,0}\,.
\end{aligned}
\tag{35}
$$

Simplified analytic formulas for calculation of the defocus can be derived supposing the approximation $\exp(\mathrm{i}\alpha t^2) \approx 1 + \mathrm{i}\alpha t^2$. The normalized intensity in point P than can be derived with similar procedure, and the following formula holds:

$$
I_n(\mathrm{P},\,\Delta) \approx \frac{1}{16a^2b^2}(G_4\Delta^4 + G_2\Delta^2 + G_0)\,, \tag{36}
$$

where

$$
\begin{aligned}
G_4 &= \left(\frac{k}{2R^2}\right)^4 \frac{16}{\beta_x^6\beta_y^6}[(a^2\beta_x^2 - 2)\sin(a\beta_x) + 2a\beta_x\cos(a\beta_x)]^2\\
&\quad [(b^2\beta_y^2 - 2)\sin(b\beta_y) + 2b\beta_y\cos(b\beta_y)]^2\,,\\
G_2 &= \left(\frac{k}{2R^2}\right)^2 \frac{16}{\beta_x^6\beta_y^2}[(a^2\beta_x^2 - 2)\sin(a\beta_x) + 2a\beta_x\cos(a\beta_x)]^2\sin^2(b\beta_y)\\
&+ \left(\frac{k}{2R^2}\right)^2 \frac{16}{\beta_x^2\beta_y^6}[(b^2\beta_y^2 - 2)\sin(b\beta_y) + 2b\beta_y\cos(b\beta_y)]^2\sin^2(a\beta_x)\,,\\
G_0 &= \frac{16}{\beta_x^2\beta_y^2}\sin^2(a\beta_x)\sin^2(b\beta_y)\,.
\end{aligned}
$$

Afterwards, one can find an approximate solution for the defocus $\Delta = \Delta(I_n(\mathrm{P},\,\Delta))$ from the formula:

$$
\Delta = \pm\sqrt{\frac{-G_2 \pm \sqrt{G_2^2 - 4G_4[G_0 - 16a^2b^2 I_n(\mathrm{P},\,\Delta)]}}{2G_4}}\,. \tag{37}
$$

In the afore-mentioned situation of the rectangular aperture, there is no analytic solution for the normalized illumination. It can be seen from the assumption of calculation in a region which has its width $2\tau_{x,m}$ and height $2\tau_{y,m}$ (note that the variables τ is used). Afterwards, the normalized illumination $L_n(\tau_{x,m},\,\tau_{y,m},\,\Delta)$ would be given by formula:

$$
L_n(\tau_{x,m},\,\tau_{y,m},\,\Delta) = K\int_{-\tau_{y,m}}^{\tau_{y,m}}\int_{-\tau_{x,m}}^{\tau_{x,m}} I_n(\tau_x,\,\tau_y,\,\Delta)\,\mathrm{d}\tau_x\mathrm{d}\tau_y\,, \tag{38}
$$

where K denotes the normalization constant. Substituting Eq. (34) in Eq. (38) leads to integration of a type which has no analytic solution. Therefore, there is no simple analytic formula for calculating the defocus Δ from the rectangular aperture.

3. Examples

3.1. Example 1 - Calculating the defocus from intensity measurements

Suppose the following parameters of an optical system and imaging: the f-number of the optical system $c = 2$, the pixel size of the detector $p = 1.25\,\mu\mathrm{m}$, and the wave-length $\lambda = 633$ nm. Fig. 2 shows simulated normalized intensity distribution calculated with presented formulas for the circular detector with dimension $\rho_{\max} = a = 100p$ and nominal values of defocuses $\Delta = [1,\,3,\,5,\,10]\,\mu\mathrm{m}$. Fig. 3 shows simulated normalized intensity distribution calculated for the rectangular detector with dimensions $a = b = 100p$. Calculated defocuses from intensity distributions by the approximate solution from Eq. (25), by finding roots of the polynomial from Eq. (35), and the approximate solution by Eq. (37) are presented in Table 1. Afterwards, the normalized illuminations for the circular aperture were calculated according Eq. (23) and one get values of reconstructed defocuses with Eq. (26) which are shown in Table 1 as well.

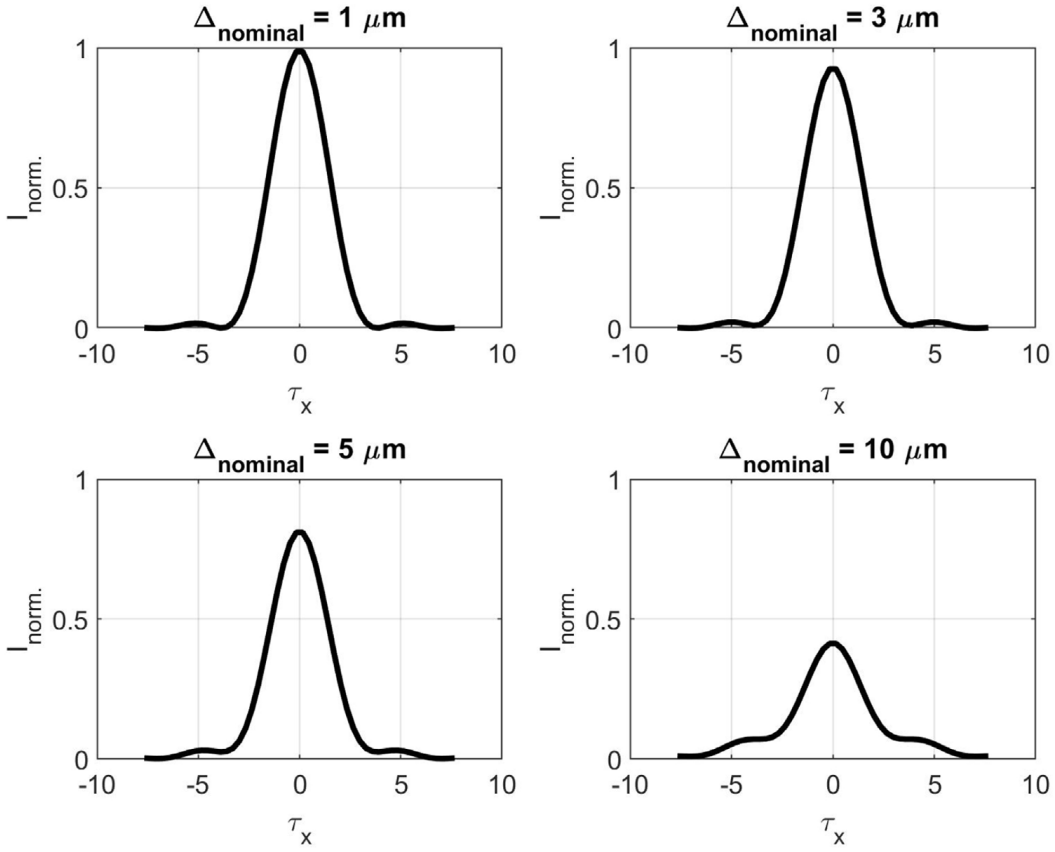


Fig. 2. Simulated normalized intensity distribution for circular aperture and parameters: $c = 2$, $p = 1.25\,\mu\text{m}$, $\lambda = 633$ nm, $\rho_{\max} = a = 100p$.

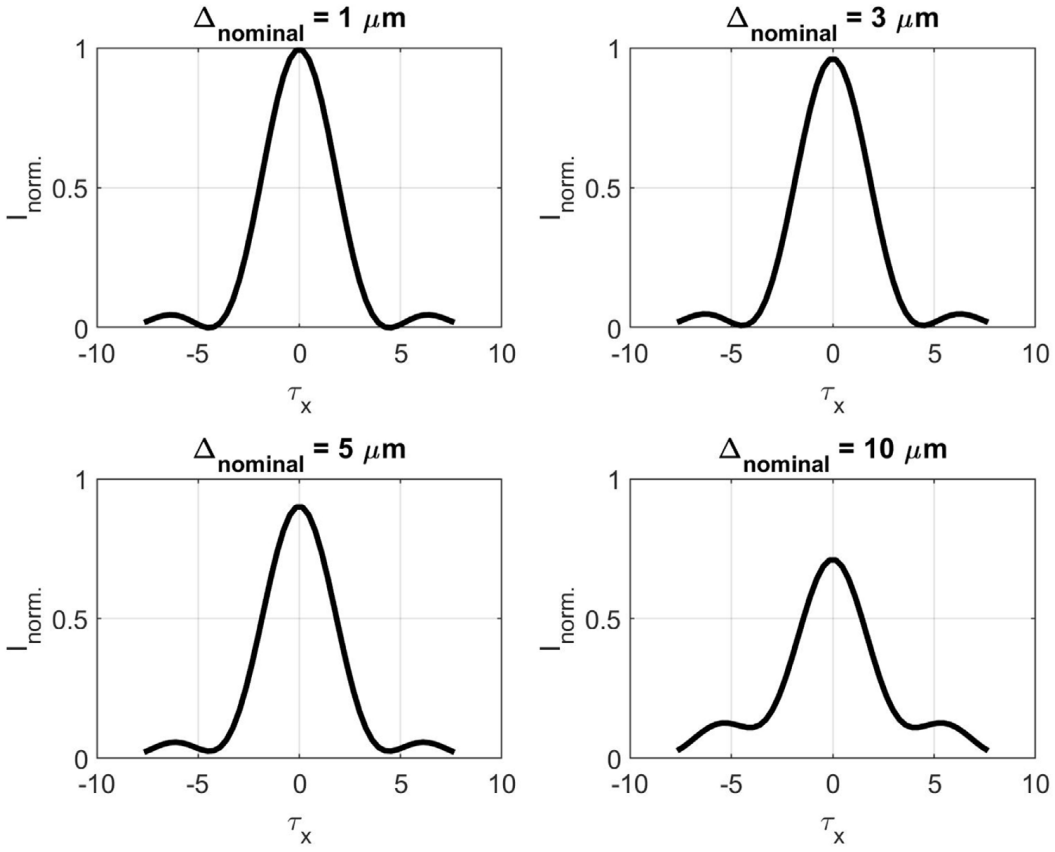


Fig. 3. Simulated normalized intensity distribution for rectangular aperture and parameters: $c = 2$, $p = 1.25\,\mu\text{m}$, $\lambda = 633$ nm, $a = b = 100p$.

Table 1
Calculated defocuses from intensity distributions for circular detector by solving Eq. (25), for rectangular detector by solving Eqs. (35) and (37), and defocuses calculated from normalized illumination for circular aperture by solving Eq. (26)

Nominal defocus [μm]	Eq. (25) [μm]	Eq. (35) [μm]	Eq. (37) [μm]	Eq. (26) [μm]
1.00	0.99	1.00	0.80	1.02
3.00	2.98	3.00	2.43	3.06
5.00	4.92	4.97	4.12	5.25
10.00	9.17	10.00	8.90	11.63

3.2. Example 2 - Calculating the defocus for its large values

Consider now a situation when the value of defocus Δ has larger value than in the Example 3.1. Suppose the following parameters of the optical system with a circular aperture: $c = 2$ and $c = 1$, $\lambda = 633$ nm. In Fig. 4, there is a plot of the function $L_n(\tau_m,\ \Delta)$ for values Δ in interval $[0,\ 0.05]$ mm, $c = 2$ for $\tau_1 = 3.832$, $\tau_2 = 2\tau_1$ and $\tau_3 = 3\tau_1$. If one knows value of the function $L_n(\tau_m,\ \Delta)$ it is possible to get value Δ. As shown in Fig. 4, it is necessary to use larger values of τ_m for larger Δ for accurate calculation. Function $L_n(\tau_m,\ \Delta)$ for Δ in interval $[0,\ 0.01]$ mm and parameters of the optical system $c = 1$, $\tau_1 = 3.832$, $\tau_2 = 2\tau_1$ and $\tau_3 = 3\tau_1$ is shown in Fig. 5. In Fig. 6, there is a dependence of $L_n(\tau_{x,m},\ \tau_{y,m},\ \Delta)$ on Δ for a square aperture with dimensions $2\tau_m \times 2\tau_m$ presented, which was calculated by Eq. (38) for $c = 2$.

As shown in Fig. 5, one can detect depth differences of the magnitude of 10^{-2} mm if $L_n(\tau_m,\ \Delta)$ is known (measured) for the optical system with $c = 1$ (numerical aperture = 0.5). Comparison between Fig. 4 (circular aperture) and Fig. 6 (square aperture) shows similar curves in both cases. Therefore, it is not important which of the aperture shape is used. Values of $L_n(\tau_m,\ \Delta)$ for the square aperture are slightly higher than in the case of circular one as the square area is $4\tau_m^2$ which is more than the circular $3.14\tau_m^2$.

4. Experimental verification and application in microscopy

4.1. Design of experimental setup

This part of the paper presents experimental verification of the aforementioned theoretical analysis. Fig. 7 shows a principal optical scheme of the measurement unit. The source of light S illuminates through the condenser C a shade with very small pin-hole which has a diameter smaller that the diameter of the Airy disc (in the plane of the pin-hole) of the optical system OS (microscope objective). The pin-hole is imaged on the CCD sensor. The signal from the CCD sensor is transferred to the computer where the calculation is performed.

Fig. 8 shows a scheme of utilisation in practice. The proposed method can determine depth of sample with measuring the intensity or illumination distribution. The source of light S illuminates trough the condenser C a very small particle, i.e. a microsphere [16] or a pin-hole [17], which can be considered as a point source (if its diameter will be smaller than the Airy disc's diameter $d_A = 2.4\lambda c$ in the object space of the optical system OS). Such a particle is imaged by the optical system OS on the sensor (CCD) which registers the intensity distribution. The signal is transferred to the computer which processes it. If the particle is shifted from the position of the best focus of the optical system by the value of Δ, the intensity distribution can be recorded and the defocus can be evaluated using the presented formulas in the previous part of the paper.

Table 2 shows diameters of Airy disc in the object space d_A and in the image space for several microscope objectives (magnification/numerical aperture) and wave-length. If the pin-hole or the observed particle has smaller diameter than d_A, it can be

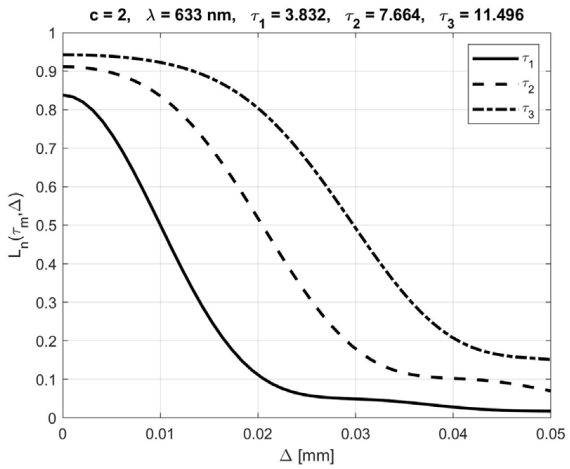


Fig. 4. Dependence of $L_n(\tau_m,\ \Delta)$ on Δ and τ_m for circular aperture of radius τ_m and $c = 2$.

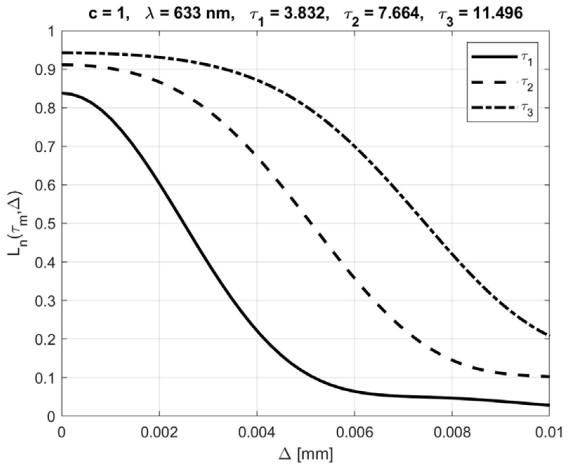


Fig. 5. Dependence of $L_n(\tau_m,\ \Delta)$ on Δ and τ_m for circular aperture of radius τ_m and $c = 1$.

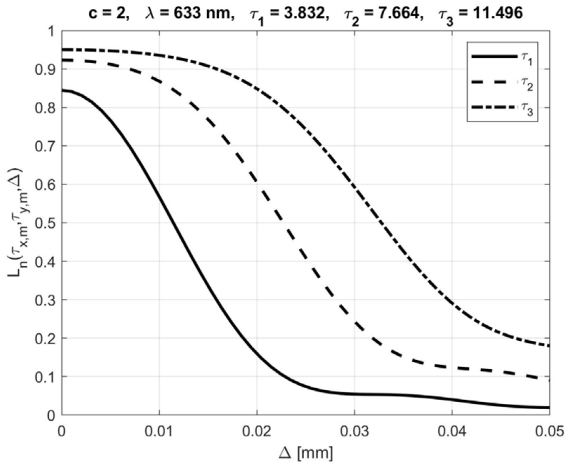


Fig. 6. Dependence of $L_n(\tau_{x,m},\ \tau_{y,m},\ \Delta)$ on Δ for $2\tau_m \times 2\tau_m$ square aperture and $c = 2$.

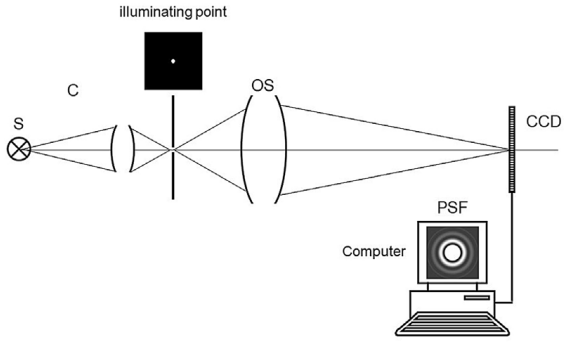


Fig. 7. Principal optical scheme of the measurement unit.

considered as a point source.

4.2. Experimental verification

An experimental setup for a verification of the afore-mentioned situation was realized. Mono-chromatic light of the wave-length $\lambda = 633$ nm was carried by the optical fibre and focused approximately to the object plane of the microscope objective by the lens of the condenser. A pin-hole (Edmund optics 1 μm Aperture Diameter Precision Pinhole [17]) was placed in the object plane of the microscope objective (Meopta Achromat 10x0.30). The pin-hole fulfilled the condition on a smaller diameter than the Airy disc's diameter, and it could be supposed as a point source of light (lighting point particle). An image of the pin-hole was registered by a common CCD sensor (machine vision Edmund optics camera EO-18112 Color USB 3.0 [17]), which is commercially easily available.

Images of point source were registered for different values of defocuses and the evaluation was processed in the area of Airy disc

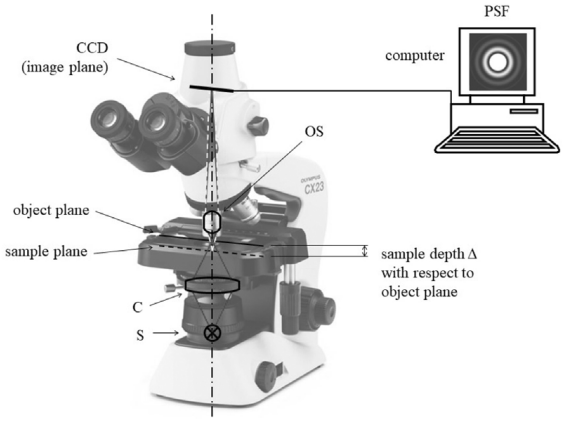


Fig. 8. Principal scheme of sample's depth determination in microscopy.

Table 2
Diameters of Airy disc in the object space (d_A) and in the image space (d'_A) of microscope objectives

objective	$c = 1/(2NA)$	d_A [mm]	d'_A [mm]
10x/0.25	2.00	0.0026	0.0262
20x/0.45	1.11	0.0015	0.0291
40x/0.65	0.77	0.0010	0.0403
40x/0.95	0.53	0.0007	0.0276
60x/0.85	0.59	0.0008	0.0462
100x/1.4	0.36	0.0005	0.0468

Table 3
Results of experimental verification for the microscope objective and imaging parameters $m = 11.63$ (calibrated with respect to experimental setup), NA = 0.30, $\lambda = 633$ nm, $p = 1.25\,\mu$m, $d = 1\,\mu$m

Nominal defocus [μm]	2.00	4.00	6.00
Eq. (25) [μm]	2.46	3.57	5.05
Eq. (26) [μm]	1.52	4.12	5.42

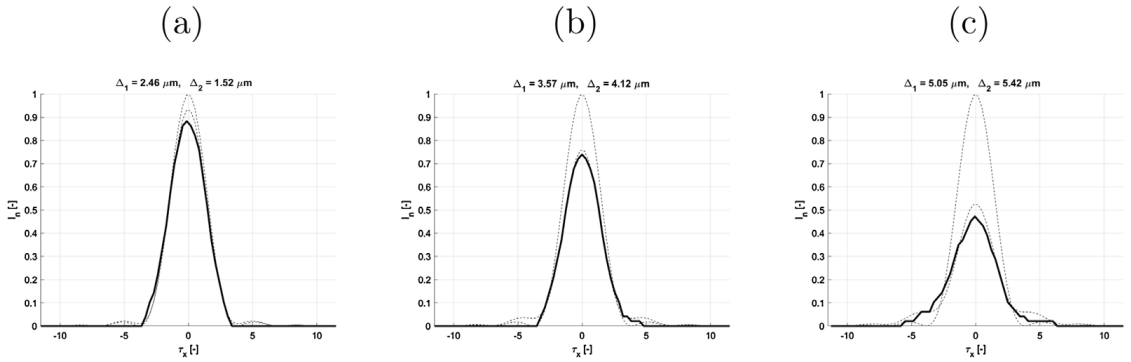


Fig. 9. Results of experimental verification for the microscope objective and imaging parameters $m = 11.63$ (calibrated with respect to experimental setup), NA = 0.30, $\lambda = 633$ nm, $p = 1.25$ μm, $d = 1\,\mu$m; normalised images for nominal defocus: a) 2 μm, b) 4 μm, c) 6 μm (solid bold line – measured data, dashed lines – theoretical values for zero and nominal defocus; Δ_1 - solution with Eq. (25), Δ_2 - solution with Eq. (26)).

($\tau_m = 3.832$). After a background variation suppression, all images were normalized with respect to the nominal image without the defocus. Table 3 and Fig. 9 presents results of the evaluation.

One can see very good agreement between nominal and reconstructed values. Differences from nominal values can be caused by noise or inaccuracies in registered intensities with a common CCD sensor. One should be able to retrieve much more better results with professional camera for metrologic applications. Other possibility in the source of differences is an uncertainty in setting the nominal value of the defocus on the microscope table which was 1 μm for this case.

5. Conclusion

The paper analysed a possibility of determination of axial defocus of lighting point, which is imaged by an optical system without aberrations, from a knowledge of intensity distribution or an amount of light (amount of energy – fractional total energy) detected in a region around a point's image centre. Analytic formulas for calculation of the defocus were derived for rectangular and circular aperture of the optical system, and they were verified by numerical examples, simulations and experiment. Outputs show that the method can evaluate the axial depth of the particle up to several micro-metres with accuracy in a range of fraction of percents.

Funding

This work was supported by the Grant Agency of the Czech Technical University in Prague, grant No. SGS18/105/OHK1/2T/11.

References

[1] L.C. Martin, The theory of the microscope, Proc. Phys. Soc. 43 (2) (1931) 186 http://stacks.iop.org/0959-5309/43/i=2/a=308.
[2] M. Born, L. Born, E. Wolf, M. Born, A. Bhatia, P. Clemmow, F. Clemmow, D. Gabor, A. Stokes, A. Taylor, et al., Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light, Cambridge University Press, 1999, https://books.google.cz/books?id=nUHGpfNsGyUC.
[3] T. Tkaczyk, Field Guide to Microscopy, Field Guide Series, SPIE, 2010, https://books.google.cz/books?id=BSJFAQAAIAAJ.
[4] D. Murphy, M. Davidson, Fundamentals of Light Microscopy and Electronic Imaging, Wiley, 2012, https://books.google.cz/books?id=f8qhtqRjxQMC.
[5] P. Török, F. Kao, Optical Imaging and Microscopy: Techniques and Advanced Systems, Springer Series in Optical Sciences, Springer, 2013, https://books.google.cz/books?id=sRnzCAAAQBAJ.
[6] U. Kubitscheck, Fluorescence Microscopy: From Principles to Biological Applications, Wiley, 2017, https://books.google.cz/books?id=kqNtDgAAQBAJ.
[7] J. Pawley, Handbook of Biological Confocal Microscopy, Springer, 2012, https://books.google.cz/books?id=qtB9BwAAQBAJ.
[8] U. Birk, Super-Resolution Microscopy: A Practical Guide, Wiley, 2017, https://books.google.cz/books?id=u44tDwAAQBAJ.
[9] J. Heath, Super-Resolution Microscopy, Wiley, 2018.
[10] H. Erfle, Super-Resolution Microscopy: Methods and Protocols, Methods in Molecular Biology, Springer, 2017, https://books.google.cz/books?id=wynmswEACAAJ.
[11] C. Franke, M. Sauer, S. van de Linde, Photometry unlocks 3d information from 2d localization microscopy data, Nat. Methods 14 (2016), https://doi.org/10.1038/nmeth.4073 41 EP.
[12] T. Fuchs, R. Hain, C.J. Kähler, Three-dimensional location of micrometer-sized particles in macroscopic domains using astigmatic aberrations, Opt. Lett. 39 (5) (2014) 1298–1301, https://doi.org/10.1364/OL.39.001298.
[13] J. Goodman, Introduction to Fourier Optics, McGraw-Hill physical and quantum electronics series, W.H. Freeman, 2005, https://books.google.cz/books?id=ow5xs_Rtt9AC.
[14] A. Papoulis, Systems and Transforms With Applications in Optics, McGraw-Hill series in systems science, Malabar, Robert Krieger Publishing Company, Florida, 1968https://books.google.cz/books?id=gqoeAQAAIAAJ.
[15] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables, Applied mathematics series, Dover Publications, 1965, https://books.google.cz/books?id=MtU8uP7XMvoC.
[16] Cospheric, http://www.cospheric.com/.
[17] Edmund Optics, https://www.edmundoptics.com/.

Journal of the Optical Society of America A
OPTICS, IMAGE SCIENCE, AND VISION

Calculation of a lens system with one or two aspherical surfaces having corrected spherical aberration

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 9 June 2020; revised 17 July 2020; accepted 21 July 2020; posted 21 July 2020 (Doc. ID 399361); published 7 August 2020

The paper presents a detailed theoretical analysis of characteristics of a rotationally symmetric lens system with one or two aspherical surfaces having corrected spherical aberration and reduced coma aberration for a given position of the object and the image. Formulas for surface shape optimization are derived, and the procedure for calculating the aspherical system is shown. The presented formulas are verified with examples of ray tracing.

https://doi.org/10.1364/JOSAA.399361

1. INTRODUCTION

Designing aspherical surfaces has been a challenging task for decades, and it has been studied in many books [1–7] and papers [8–23]. There is a variety of computational tools available that help with optimization of lenses' forms, e.g., Zemax [24] or Oslo [25]. For example, reduction of a spherical aberration of a classic lens can be done with one aspherical surface for a given object position.

One can find many papers [8–10,13–19,23] on topics of minimization of the spherical aberration, which solve the problem by various approaches. The goal of this paper is to present a simple method of designing a lens system with one or two aspherical surfaces that has, for a given object position, corrected spherical aberration. To the authors' best knowledge, the presented simple procedure has not been published before. The formulas are valid for ray tracing through several optical surfaces, while one or two of them are aspherical.

As a specific case, a situation of a simple lens in air is specifically discussed. However, it can be easily generalized for optical systems of many surfaces.

In the paper, formulas for a polynomial approximation of an aspherical surface are derived, and tools for a general meridional ray analysis are presented, as well as formulas for ray tracing of off-axis rays. Moreover, the paper studies a coma aberration minimization for the aspherical lens. The final part of the paper presents examples that demonstrate the proposed approaches of aspherical lens design and verify derived equations. The method can serve as an alternative possibility to optical design in commercially available software.

2. DESIGN OF ASPHERICAL SURFACE OF OPTICAL SYSTEM

A. General Situation of Ray Refraction on Optical Surfaces

Let one study a ray refraction by a system of optical surfaces now. Figure 1 shows a scheme of meridional aperture ray refraction on an i-th and $(i+1)$-st optical surface (spherical or aspherical). Symbols in the figure represent: n_i, index of refraction of i-th space; σ_i, angle between a ray impinging on the i-th surface with an optical axis (axis z) of the optical system; σ_{i+1}, angle between a ray impinging on the $(i+1)$-st surface and the optical axis; s_i, axial distance between point A_i and vertex V_i of the i-th surface; y_i, transverse distance of point B_i from the optical axis; d_i, distance between vertex V_{i+1} of $(i+1)$-st surface of the system and vertex V_i of the i-th surface; z_i, z coordinate of point B_i measured from the vertex V_i of the i-th surface; t_i, distance between points A_i and B_i. Meanings of other symbols are obvious in Fig. 1. To clarify a sign convention used in the paper, the distance s_i is negative if the point A_i is located on the left side of the vertex V_i, and positive if opposite. Angles σ are measured from the optical axis and are positive if this direction is clockwise and negative if opposite.

Suppose now that the optical system consists of m optical surfaces and images point $A \equiv A_1$ into the point $A' \equiv A_{m+1}$. Let $[AA']$ denote an optical path distance (a product of geometrical distance and index of refraction) of a general meridional ray, and let $[AA']_0$ symbolize an optical path distance of a ray passing along the optical axis of the system. The difference of optical path distances of those rays is then given as follows:

$$\delta = [AA'] - [AA']_0. \tag{1}$$

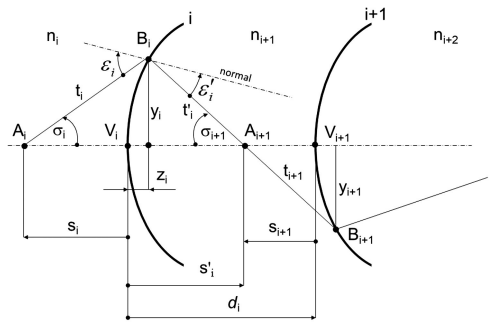


Fig. 1. Ray refraction on optical surfaces.

According to Fig. 1, one then gets the following formula for rotationally symmetric spherical or aspherical surfaces:

$$\delta = \sum_{i=1}^{m} \delta_i = \sum_{i=1}^{m} \left[n_{i+1} \left(t'_i - s'_i \right) - n_i (t_i - s_i) \right]$$
$$= \sum_{i=1}^{m} \left[y_i \left(n_{i+1} \tan \frac{\sigma_{i+1}}{2} - n_i \tan \frac{\sigma_i}{2} \right) - (n_{i+1} - n_i)\, z_i \right]. \tag{2}$$

It is known that for a stigmatically imaged point A into point A′, an equation $\delta = 0$ has to be fulfilled. In general, the surface (surfaces) shape will affect the optical path difference, and it can be used for the designing process described below. An optimal optical design can be reached in a situation if the shape of an optical surface (surfaces) will generate zero optical path difference δ.

B. Situation of a Simple Lens in Air

Suppose a specific situation of a simple lens in air is now ($m = 2$, $n_1 = 1$, $n_2 = n$, $n_3 = 1$), where the shape of the first surface of the lens is known, e.g., spherical or aspherical. Therefore, one can easily calculate parameters of rays coming through this surface.

In general, ray tracing of meridional rays through an aspherical surface [5], whose surface is described by the formula $z = f(y)$, can be calculated using the following formulas ($i = 1, 2, 3, \ldots, m$):

$$y_i = (s_i - z_i) \tan \sigma_i, \quad \tan \omega_i = \left(\frac{\mathrm{d}z}{\mathrm{d}y} \right)_i,$$
$$\varepsilon_i = \omega_i - \sigma_i, \quad \sin \varepsilon'_i = \frac{n_i}{n_{i+1}} \sin \varepsilon_i,$$
$$\sigma_{i+1} = \sigma_i + \varepsilon_i - \varepsilon'_i, \quad s'_i = \frac{y_i}{\tan \sigma_{i+1}} + z_i, \tag{3}$$

where ω_i is the angle between the normal in point (y_i, z_i) and axis z, and ε_i and ε'_i denote, respectively, angles of the impinging and refracted rays with respect to the normal of the i-th surface.

As the first simple case, suppose a spherical shape of the first surface with the vertex radius of curvature r_1. Further, let given (selected) parameters of the lens be: the axial distance s_1 between the object and the vertex of the first surface, the impinging height y_1 on the first surface of the lens, lens thickness d, index of refraction n of a lens material, and paraxial axial distance s'_2 behind the second surface of the lens. Then one calculates

$$z_1 = \frac{y_1^2}{r_1 \left[1 + \sqrt{1 - (y_1/r_1)^2} \right]}, \quad \tan \sigma_1 = \frac{y_1}{s_1 - z_1},$$
$$\sin \varepsilon_1 = \left(\frac{s_1}{r_1} - 1 \right) \sin \sigma_1, \quad \sin \varepsilon'_1 = \frac{\sin \varepsilon_1}{n},$$
$$\sigma_2 = \sigma_1 + \varepsilon_1 - \varepsilon'_1, \quad s'_1 = \frac{y_1}{\tan \sigma_2} + z_1, \quad s_2 = s'_1 - d. \tag{4}$$

Afterwards, the second surface of the lens is calculated in such a way that the lens images an axial object point as a stigmatic image—without spherical aberration.

Further, it holds for a simple lens in air, according to Fig. 1, that

$$\tan \sigma_2 = \frac{y_2}{s_2 - z_2}, \quad \tan \sigma_3 = \frac{y_2}{s'_2 - z_2}. \tag{5}$$

The solution to Eq. (5) then leads to

$$y_2 = \frac{\tan \sigma_2 \tan \sigma_3 (s'_2 - s_2)}{\tan \sigma_2 - \tan \sigma_3}, \quad z_2 = \frac{s_2 \tan \sigma_2 - s'_2 \tan \sigma_3}{\tan \sigma_2 - \tan \sigma_3}. \tag{6}$$

With Eq. (2), for a case of stigmatically imaged point A into point A′, i.e., $\delta = \delta_1 + \delta_2 = 0$, one gets the following formula:

$$\delta_1 + y_2 \left(\tan \frac{\sigma_3}{2} - n \tan \frac{\sigma_2}{2} \right) - (1 - n) z_2 = 0, \tag{7}$$

where

$$\delta_1 = y_1 \left(n \tan \frac{\sigma_2}{2} - \tan \frac{\sigma_1}{2} \right) - (n - 1) z_1 \tag{8}$$

is an optical path difference introduced by the first surface of the lens. Substituting Eq. (6) into Eq. (7) then gives the following formula for a calculation of the angle σ_3:

$$\alpha \sin \sigma_3 + \beta \cos \sigma_3 + \gamma = 0, \tag{9}$$

where

$$\alpha = \cos \sigma_2 (\delta_1 - s'_2 + n s_2) - n (s_2 - s'_2),$$
$$\beta = -\sin \sigma_2 (\delta_1 - s'_2 + n s_2),$$
$$\gamma = \sin \sigma_2 (s_2 - s'_2). \tag{10}$$

Afterwards, the solution to Eq. (9) leads to

$$\sin \sigma_3 = -\frac{\alpha \gamma + \beta \sqrt{\alpha^2 + \beta^2 - \gamma^2}}{\alpha^2 + \beta^2}. \tag{11}$$

Substituting Eq. (11) into Eq. (6) and simple calculation then give parametric coordinates of the point on the second surface of the lens (an aspherical surface).

In the case of the first aspherical surface of the lens (bi-aspherical lens), one uses the same procedure only with application of general Eq. (3) instead of Eq. (4). Therefore, the given problem is solved. This procedure leads to simpler calculations than, for example, [18].

Using the aforementioned formulas in Eqs. (10), (11), and (6), one can calculate coordinates (y_2, z_2) of the point on

the aspherical (second) surface. The profile of the lens can be obtained after a selection of several values y_1 and repeating the aforementioned procedure for different impinging heights.

Suppose now that one wants to ray trace a ray through a rotational symmetric aspherical lens for a different position of an object than the lens was designed for. To be able to use ray-tracing procedures, one has to know a function $z_2 = f(y_2)$, i.e., an analytical description of the surface shape.

Suppose that the rotational symmetric aspherical surface is described with the following formula (power series):

$$z_2 = \sum_{i=1}^{N} a_i y_2^{2i}, \tag{12}$$

where $a_1 = 1/(2r_2)$, and r_2 is the vertex radius of the last (second) surface of the lens. One has to determine coefficients a_i at the moment. Therefore, a series of coordinates $(z_2, y_2)_k$ for different values $(y_1)_k$ can be calculated, where $k = 1, 2, 3, \ldots, K$, $K > N$. Afterwards, coefficients a_i can be calculated with the least-squares method [26] as follows. Two matrices **A** and **B** can be arranged with elements

$$A_{ki} = (y_2^{2i})_k, \quad B_k = (z_2)_k, \tag{13}$$

where k denotes the index of the point $(z_2, y_2)_k$ and number of the row in matrices **A** and **B**, and i is the column number in matrix **A**. Unknown coefficients $\mathbf{a} = [a_1, a_2, \ldots, a_N]^T$ then can be calculated with the known formula of the least-squares method [26]:

$$\mathbf{a} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{B}, \tag{14}$$

where T denotes matrix transposition. Therefore, Eq. (12) is an approximate description of the aspherical surface with known coefficients a_i. Accuracy of calculated coordinate z_2 of the aspherical surface depends on a number of series elements N. The usual value is $N = 5$ in practice. At this point, it has to be noted that matrix $\mathbf{A}^T\mathbf{A}$ can be badly conditioned in some cases for large values of y_2, and results calculated with Eq. (14) can be inaccurate if one uses a low number of valid digits. Therefore, for such situations, it is more appropriate to use a normalized variable $\bar{y}_2 = y_2/(y_1)_{\max}$ instead of y_2, where $(y_1)_{\max}$ denotes the maximal value of height y_1 on the first surface of the lens.

3. RAY TRACING OF MERIDIONAL RAYS THROUGH AN ASPHERICAL LENS

Suppose now the aspherical lens in air with the first spherical surface and the second aspherical surface, which is determined by Eq. (12), and coefficients a_i are known. Afterwards, an analysis of a general meridional ray passing through the lens can be processed with two possibilities.

In the first approach, the following parameters are known: $r_1, d, n, s_1, y_1, a_1, a_2, a_3, a_4, a_5$ ($N = 5$). Afterwards, Eq. (4) is used to calculate refraction on the first surface, and values of s_2 and σ_2 are obtained. Further, one can generally substitute the formula $y_2 = (s_2 - z_2)\tan\sigma_2$ into Eq. (12), and the following condition for the value of z_2 holds after simplification:

$$\sum_{m=0}^{10} b_m z_2^m = 0, \tag{15}$$

where

$$\begin{aligned}
b_{10} &= a_5 \tan^{10}\sigma_2,\\
b_9 &= -10 b_{10} s_2,\\
b_8 &= 45 b_{10} s_2^2 + C,\\
b_7 &= -120 b_{10} s_2^3 - 8 C s_2,\\
b_6 &= 210 b_{10} s_2^4 + 28 C s_2^2 + D,\\
b_5 &= -252 b_{10} s_2^5 - 56 C s_2^3 - 6 D s_2,\\
b_4 &= 210 b_{10} s_2^6 + 70 C s_2^4 + 15 D s_2^2 + E,\\
b_3 &= -120 b_{10} s_2^7 - 56 C s_2^5 - 20 D s_2^3 - 4 E s_2,\\
b_2 &= 45 b_{10} s_2^8 + 28 C s_2^6 + 15 D s_2^4 + 6 E s_2^2 + F,\\
b_1 &= -10 b_{10} s_2^9 - 8 C s_2^7 - 6 D s_2^5 - 4 E s_2^3 - 2 F s_2 - 1,\\
b_0 &= b_{10} s_2^{10} + C s_2^8 + D s_2^6 + E s_2^4 + F s_2^2,
\end{aligned} \tag{16}$$

and where

$$C = a_4\tan^8\sigma_2,\ D = a_3\tan^6\sigma_2,\ E = a_2\tan^4\sigma_2,\ F = a_1\tan^2\sigma_2. \tag{17}$$

The solution to Eq. (15) (finding the root of the polynomial on the left-hand side) then gives the desired value of the coordinate z_2 of the ray intersection with the aspherical surface. Substituting the value z_2 into the formula $y_2 = (s_2 - z_2)\tan\sigma_2$ and simple calculation then give the corresponding value of coordinate y_2. The angle ω_2 between a normal to the aspherical surface in the point (y_2, z_2) and the optical axis can be calculated according to Eq. (3) as follows:

$$\tan\omega_2 = 2\sum_{i=1}^{N} i a_i y_2^{2i-1}. \tag{18}$$

Afterwards, one can calculate the ray refraction by the aspherical surface with Eq. (3).

The second possibility of calculation of coordinates (y_2, z_2) of the ray intersection with the aspherical surface is based on an iterative procedure.

1. In the first step, the following initial guess is stated: $(y_2)^0 = s_2\tan\sigma_2$, $(z_2)^0 = 0$.
2. In the second step, the following formulas hold:

$$(z_2)^p = \sum_{i=1}^{N} a_i (y_2^{2i})^{p-1}, \quad (y_2)^p = \left[s_2 - (z_2)^p\right]\tan\sigma_2. \tag{19}$$

The iterative procedure [Eq. (19)] is repeated for given $p = 1, 2, 3, \ldots$ (p denotes the iteration number) as long as the following condition is fulfilled: $|(z_2)^p - (z_2)^{p-1}| < \Delta$, where Δ denotes stated tolerance (e.g., $\Delta = 10^{-6}$ mm). Therefore, the coordinates of the ray intersection (y_2, z_2) are calculated, and

next, ray tracing can be done with the same procedure as in the aforementioned situation.

The problem of calculation (ray tracing) of the general meridional ray is therefore solved. Generally, if the index 2 in Eqs. (15) and (19) is replaced by i, for example, then one gets the system of equations for a system of i rotational symmetric optical surfaces of an arbitrary shape.

4. RAY TRACING OF OFF-AXIAL RAYS THROUGH AN ASPHERICAL SURFACE

Analysis of a general off-axial ray can be calculated according to formulas presented in [3,12]. Another approach can be as follows. Suppose the situation shown in Fig. 2. The formula for the aspherical surface in Eq. (12) can be rewritten as follows:

$$F(x, y, z) = z - \sum_{i=1}^{N} a_i \left(x^2 + y^2\right)^i = 0. \tag{20}$$

Suppose next the following formula for the impinging ray on the aspherical surface:

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{p}_1, \tag{21}$$

where $\mathbf{r} = (x, y, z)$ is the positional vector of the ray intersection with the aspherical surface, $\mathbf{r}_0 = (x_0, y_0, z_0)$ is the positional vector of the ray intersection with a tangential plane τ that touches the aspherical surface in the vertex V, $\mathbf{p}_1 = (p_x, p_y, p_z)$ is the unit directional vector of the ray, and t denotes the parameter, which depends on the point of the ray intersection with the surface, to be calculated. Next, let the symbols in Fig. 2 denote: $\mathbf{n}$, unit normal vector of the aspherical surface in a direction of the impinging ray; $\mathbf{s}_2$, unit directional vector of the refracted ray; n_1, index of refraction of the space in front of the aspherical surface; n_2, index of refraction of the space behind the aspherical surface.

One can obtain the formula of the $2N$-th order for the parameter t after substitution of components of vector $\mathbf{r}$ from Eq. (21) into Eq. (20). The solution to the formula leads to the desired value of the parameter t. Substituting the value of the parameter back into Eq. (21) then gives coordinates of the intersection of the off-axial ray with the aspherical surface. Afterwards, the normal vector of the aspherical surface can be calculated, and using the law of refraction then gives a direction of the refracted ray.

Therefore, in the aforementioned situation, the parameter t has to fulfill the condition

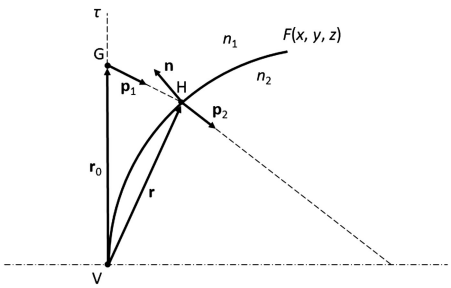


Fig. 2. Ray tracing of off-axial rays through an aspherical surface.

$$0 = z_0 + tp_z - \sum_{i=1}^{N} a_i \left(Kt^2 + Lt + M\right)^i, \tag{22}$$

where

$$K = p_x^2 + p_y^2, \quad L = 2(x_0 p_x + y_0 p_y), \quad M = x_0^2 + y_0^2.$$

The condition ($2N$-th order polynomial) in Eq. (22) was obtained by substituting the components of vector $\mathbf{r}$ from Eq. (21) into Eq. (20) and by the rearrangement.

For $N = 5$ and after simple modification, one gets the formula of the 10th order for the parameter t; it holds that

$$\sum_{i=0}^{10} c_i t^i = 0, \tag{23}$$

where

$$c_{10} = K^5 a_5,\ c_9 = 5K^4 L a_5,\ c_8 = 5K^3(2L^2 + KM)a_5 + K^4 a_4,$$

$$c_7 = 10K^2 L(L^2 + 2KM)a_5 + 4K^3 L a_4,$$

$$c_6 = 5K(2K^2M^2 + 6KL^2M + L^4)a_5 + 2K^2(3L^2 + 2KM)a_4 + K^3 a_3,$$

$$c_5 = L(30K^2M^2 + 20KL^2M + L^4)a_5 + 4KL(L^2 + 3KM)a_4 + 3K^2 L a_3,$$

$$c_4 = 5M(2K^2M^2 + 6KL^2M + L^4)a_5 + (6K^2M^2 + 12KL^2M + L^4)a_4 + 3K(L^2 + KM)a_3 + K^2 a_2,$$

$$c_3 = 10LM^2(L^2 + 2KM)a_5 + 4LM(L^2 + 3KM)a_4 + L(L^2 + 6KM)a_3 + 2KLa_2,$$

$$c_2 = 5M^3(2L^2 + KM)a_5 + 2M^2(3L^2 + 2KM)a_4 + 3M(L^2 + KM)a_3 + (L^2 + 2KM)a_2 + Ka_1,$$

$$c_1 = 5LM^4 a_5 + 4LM^3 a_4 + 3LM^2 a_3 + 2LMa_2 + La_1 - s_z,$$

$$c_0 = M^5 a_5 + M^4 a_4 + M^3 a_3 + M^2 a_2 + Ma_1 - z_0.$$

Therefore, the problem of determination of the parameter t is solved by finding the root of Eq. (23).

The unit normal vector $\mathbf{n} = (n_x, n_y, n_z)$ of the aspherical surface (see Fig. 2) can be calculated from the formula $\mathbf{n} = -\nabla F(x, y, z)/|\nabla F(x, y, z)|$. It holds after modification that

$$n_x = 2\left[\sum_{i=1}^{N} a_i i x (x^2+y^2)^{i-1}\right] / |\nabla F(x,y,z)|,$$

$$n_y = 2\left[\sum_{i=1}^{N} a_i i y (x^2+y^2)^{i-1}\right] / |\nabla F(x,y,z)|,$$

$$n_z = -|\nabla F(x,y,z)|^{-1},$$

$$|\nabla F(x,y,z)|^2 = 1 + 4\left[\sum_{i=1}^{N} a_i i x (x^2+y^2)^{i-1}\right]^2 + 4\left[\sum_{i=1}^{N} a_i i y (x^2+y^2)^{i-1}\right]^2. \tag{24}$$

The unit directional vector $\mathbf{s}_2$ of the refracted ray can be calculated with the law of refraction in the vector form; it holds that

$$\mathbf{p}_2 = \frac{n_1}{n_2}\mathbf{p}_1 - \frac{1}{n_2}\mathbf{n}\left[\sqrt{n_2^2 - n_1^2 + n_1^2(\mathbf{p}_1\cdot\mathbf{n})^2} + n_1(\mathbf{p}_1\cdot\mathbf{n})\right]. \tag{25}$$

As in the previous section (the case of the meridional ray), the intersection $\mathbf{r} = (x, y, z)$ of the general off-axial ray with the aspherical surface can be calculated with the iterative procedure. If in the first step the following initial guess is stated: $(x)^0 = x_0$, $(y)^0 = y_0$, then the following formulas for an iterative procedure similar to Eq. (19) hold:

$$(z)^p = \sum_{i=1}^{N} a_i\left[(x^2+y^2)^i\right]^{p-1},$$

$$(x)^p = x_0 + \frac{(z)^p - 2z_0}{p_z}p_x,$$

$$(y)^p = y_0 + \frac{(z)^p - 2z_0}{p_z}p_y. \tag{26}$$

This procedure calculates the ray intersection with a plane perpendicular to the optical axis with point $(z)^p$ in it. If the intersection (coordinates) $\mathbf{r} = (x, y, z)$ is calculated, then one continues with a similar procedure according to Eqs. (24) and (25).

5. EFFECT OF LENS SHAPE ON COMA ABERRATION

It has been stated above that the spherical aberration of the simple lens can be eliminated if one of its surfaces is aspherical. During imaging of off-axial points, images will no longer be points, but so-called spot diagrams. The sizes of the diagrams depend on the shape of the lens for given object plane position, focal length, and entrance numerical aperture, and the used wavelength is determined by a dispersion formula of the lens material. In other words, the effect of the lens shape on coma aberration depends on a ratio of vertex radii of curvatures of the lens. Therefore, one can affect coma aberration during designing the lens with eliminated spherical aberration as well. It is known from geometrical optics [1,3,5,7] that meridional (tangential) coma $\delta y'_m$ can be expressed with the formula

$$\delta y'_m = 3y'_0\left(\frac{\delta m}{m_0} - \frac{\delta s'}{p'}\right) = 3y'_0 C_m, \tag{27}$$

where y'_0 is the image size, $\delta m = m - m_0$ is the deviation from the sine condition, $m = \sin\sigma/\sin\sigma'$ is the transverse magnification of the lens in air, σ is the entrance aperture angle, σ' is the exit aperture angle, m_0 is the paraxial magnification, $\delta s'$ is the lateral spherical aberration, p' is the distance between the image plane and the exit pupil of the optical system, and C_m is the coma coefficient. In the case of the aspherical lens with corrected spherical aberration, it holds that $\delta s' = 0$. In the case of minimized coma of the aspherical lens, the radii of curvatures r_1 and r_2 have to fulfill the following approximate formula:

$$\frac{r_2 + r_1}{r_2 - r_1} \approx \left(\frac{2n^2}{n+1} - 1\right)\left(1 + \frac{2f'}{s}\right), \tag{28}$$

where n is the index of refraction of the lens, f' is its focal length, and s is the object distance from the lens. Eq. (28) was derived from the theory of aberrations of the third order; see, for example, [5,7]. It is valid for a lens with both spherical surfaces. However, one can use the mentioned formula as a very good approximation for an aspherical lens, as will be shown in Section 6.B.

6. EXAMPLES

A. Example 1—Plan-Hyperbolic Lens

It is known that the rotational symmetric plan-hyperbolic lens has no spherical aberration for a parallel beam of rays impinging on the lens from a side of its planar surface. And similarly, the plan-hyperbolic lens transforms the ray coming from its focal point impinging on the hyperbolic surface to a parallel beam.

The general formula for a curve of the second order (conic section) has the following form, as known from analytic geometry [26]:

$$y^2 = 2R_0 z + (\varepsilon^2 - 1)z^2, \tag{29}$$

where R_0 is the radius of curvature of the curve in its vertex, and ε is its numerical eccentricity. For the case of a plan-hyperbolic lens in air, it holds that (as can be easily proven with the Fermat principle)

$$y^2 = 2R_0 x + (n^2-1)z^2 = -2f'(n-1)z + (n^2-1)z^2, \tag{30}$$

where f' is the focal length, and n is the index of refraction of the lens material.

Consider now the following parameters for a plan-hyperbolic lens: $s_1 = -\infty$, $r_1 = \infty$, $s'_2 = f' = 100$ mm, where f' is the image focal length of the lens; $n_1 = 1$, $n_2 = n = 1.5$, $n_3 = 1$; the axial thickness of the lens is $d = 10$ mm; and the impinging height is $y_1 = 25$ mm.

Is such a situation, this limit case has the following results:

$$s_2 = \infty, \quad \sin\sigma_2 = \frac{y_2}{t_2} = \frac{y_2}{\sqrt{y_2^2 + (z_2 - s_2)^2}} = 0,$$

$$\cos\sigma_2 = 1, \quad \delta_1 = 0,$$

$$\alpha = (n-1)s_2', \quad \beta = -ny_2, \quad \gamma = y_2,$$

and one gets after substitution into Eq. (11)

$$\sin\sigma_3 = -\frac{f'y_2(n-1) - ny_2\sqrt{n^2y_2^2 + f'^2(n-1)^2 - y_2^2}}{n^2y_2^2 + f'^2(n-1)^2}.$$

Further, with the use of Eq. (6), one gets after simplification

$$y^2 = -2f'(n-1)z + (n^2-1)z^2,$$

which is the second formula of Eq. (30) for hyperbola. Therefore, it is analytically shown that the presented procedure in the paper is valid.

Numerically, the results of this example are: $\alpha = 50$ mm, $\beta = -37.5$ mm, $\gamma = 25$ mm, $\sigma_3 = 13.2917$ deg, $y_2 = 25$ mm, and $z_2 = -5.8258$ mm.

B. Example 2

As the second example, a general rotational symmetric aspherical lens in air is designed. Suppose the following input parameters: $(y_1)_{\max} = 40$ mm, $r_1 = 63$ mm, $d = 20$ mm, $n_1 = 1$, $n_2 = n = 1.516$, $n_3 = 1$, $s_1 = -4000$ mm, $s_2' = 100$ mm, and $N = 5$. The calculation is processed for heights (coordinates) y_1 selected from an interval $y_1 \in [0, (y_1)_{\max}]$. It is appropriate to select non-uniform distribution for large numerical apertures of the lens with smaller spacing for larger values of y_1. Afterwards, one calculates a set of coordinates $(y_2, z_2)_k$ of points on the aspherical surface for a given height $(y_1)_k$ with Eqs. (3), (4), (6), (11), and, e.g., $k = 1, 2, \ldots, 20$. With Eqs. (13) and (14), one then gets the following coefficients of the aspherical surface described with Eq. (12): $a_5 = -9.8135\text{e}-19$ mm^{-9}, $a_4 = 9.4575\text{e}-15$ mm^{-7}, $a_3 = -5.3593\text{e}-11$ mm^{-5}, $a_2 = 5.8261\text{e}-07$ mm^{-3}, $a_1 = -1.0953\text{e}-03$ mm^{-1}. The vertex radius of curvature of the aspherical surface then equals $r_2 = 1/(2a_1) = -456.497$ mm. Afterwards, one can ray trace rays through the aspherical lens with Eqs. (3), (4), and (15)–(18). Figure 3(a) shows a residual transverse spherical aberration, and Fig. 3(b) shows a dependency of the coefficient C_m of the meridional coma aberration on the height y_1 (impinging height on the first surface of the lens). The radius of curvature of the first surface ($r_1 = 63$ mm) was chosen to satisfy the condition in Eq. (28) as much as possible. Table 1 presents numerically a dependency of the maximal value of the coefficient C_m on the first radius of curvature r_1 and on the value

$$\Delta X = \frac{r_2 + r_1}{r_2 - r_1} - \left(\frac{2n^2}{n+1} - 1\right)\left(1 + \frac{2f'}{s}\right). \tag{31}$$

C. Example 3

Suppose now a design of an aspherical lens with a given first surface different from the spherical shape. Consider a parabolic surface described with the formula $z_1 = k_1 y_1^2$. The known parameters of the design are: $r_1 = 63$ mm, $k_1 = 1/(2r_1) = 0.008$ mm^{-1}, $(y_1)_{\max} = 40$ mm, $d = 20$ mm, $n = 1.516$, $s_1 = -4000$ mm, $s_2' = 100$ mm, and $N = 5$. The calculation was processed for values of $y_1 \in [0, (y_1)_{\max}]$ with non-uniform distribution similar to the previous example. However, the ray tracing through the first surface was calculated with Eq. (3), where $\tan\omega_1 = \mathrm{d}z_1/\mathrm{d}y_1 = 2k_1y_1$. With the presented design procedure, one gets the following coefficients of the second aspherical surface of the lens: $a_5 = 2.3812\text{e}-18$ mm^{-9},

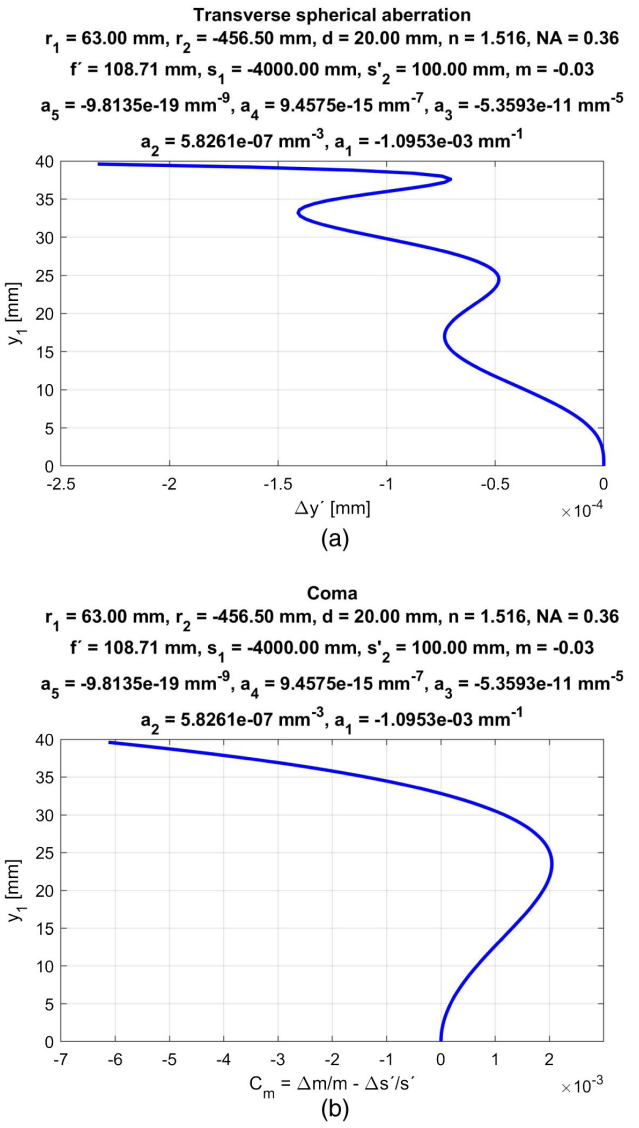


Fig. 3. Results of aspherical lens design for example 2. (a) Transverse spherical aberration and (b) coma aberration.

Table 1. Dependency of the Maximal Value of the Coefficient C_m on the First Radius of Curvature r_1 and on the Value ΔX

r_1	∞	200	100	63	50	45
$(C_m)_{\max}$	0.20	0.14	0.07	−0.01	−0.07	−0.12
ΔX	−1.79	−1.28	−0.74	0.02	0.59	0.98

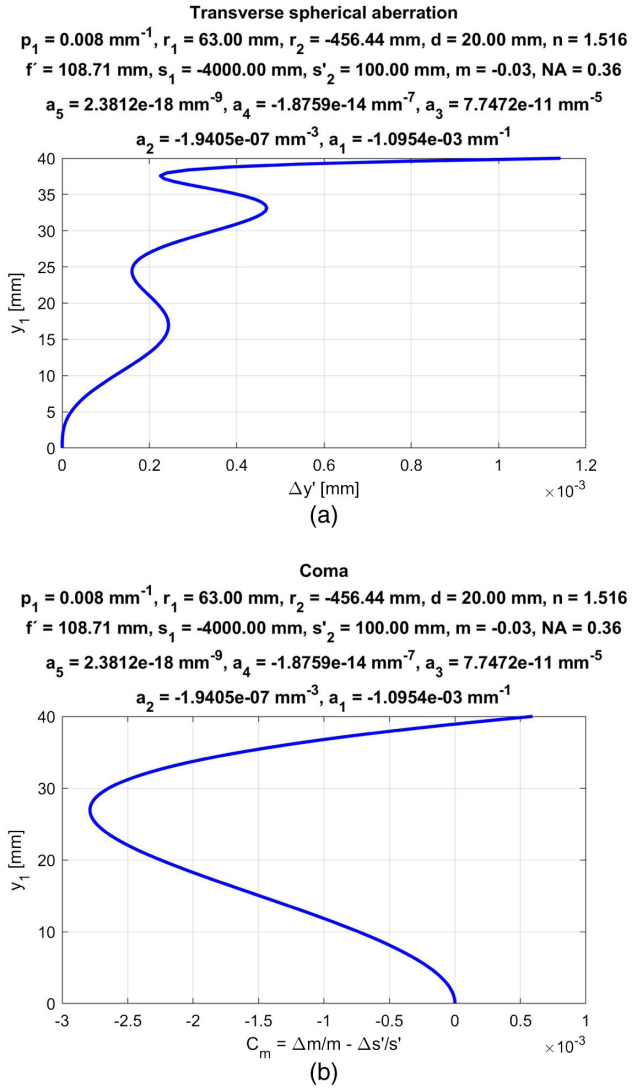


Fig. 4. Results of aspherical lens design for example 3. (a) Transverse spherical aberration and (b) coma aberration.

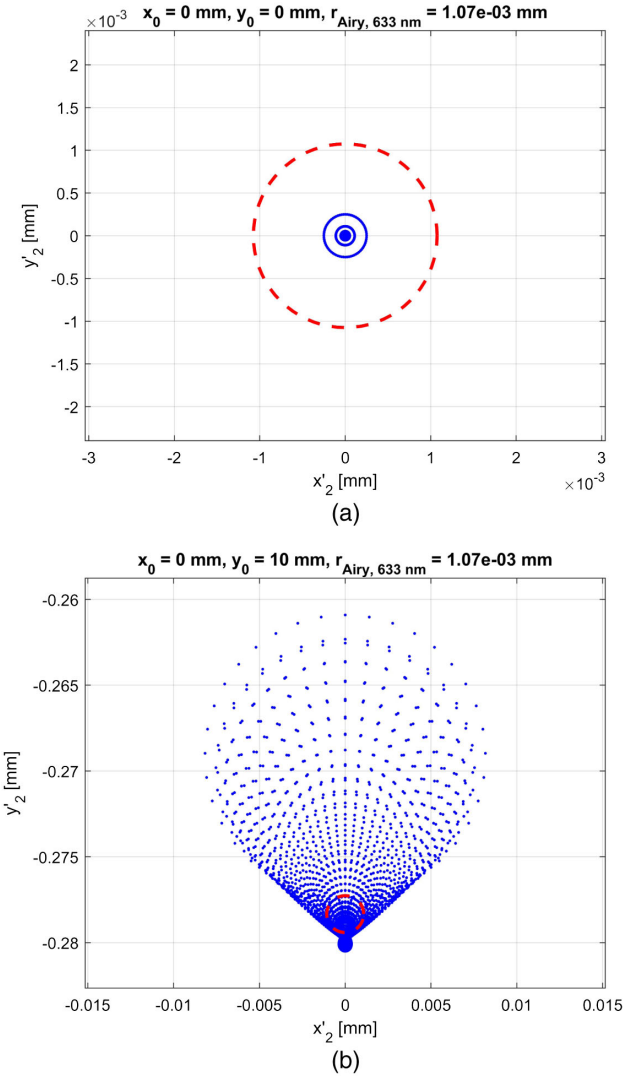


Fig. 5. Results of off-axial ray tracing for example 4. (a) Spot diagram for central beam: blue, spot diagram; red circle, Airy disc for $\lambda = 633$ nm. (b) Spot diagram for excentric beam: blue, spot diagram; red circle, Airy disc for $\lambda = 633$ nm.

$a_4 = -1.8759\text{e}-14\ \text{mm}^{-7}$, $a_3 = 7.7472\text{e}-11\ \text{mm}^{-5}$, $a_2 = -1.9405\text{e}-07\ \text{mm}^{-3}$, $a_1 = -1.0954\text{e}-03\ \text{mm}^{-1}$. The vertex radius of curvature of the aspherical surface then equals $r_2 = 1/(2a_1) = -456.44$ mm. Figure 4(a) shows results of the design with calculated transverse spherical aberration, and Fig. 4(b) presents the coma aberration of the lens.

D. Example 4

The last example shows a ray tracing of off-axial rays with the aspherical lens with geometrical and material parameters as in the second example.

Figures 5(a) and 5(b) show calculated spot diagrams in a plane (x'_2, y'_2) in the axial distance s'_2 from the vertex of the second (aspherical) surface of the lens for a point object in a plane (x_0, y_0), which is in the axial distance s_1 in front of the vertex of the first (spherical) surface of the lens. The beam of rays is limited by a value $(y_1)_{\max}$ in a plane tangential to the first surface of the lens. Refraction on the second surface was calculated with Eqs. (20)–(25). It is possible to get the same results with the iterative procedure in Eq. (26).

7. CONCLUSION

The paper presented formulas for designing a lens system with one or two aspherical surfaces. A specific case of an aspherical lens with optimized parameters was shown; the lens has the first surface planar, spherical or aspherical, and the second surface of the aspherical shape. The lens has corrected spherical aberration for point objects on the optical axis. This example can be easily generalized into complex optical systems with one or two aspheric surfaces. Design of the optical system with one or two aspherical surfaces using the presented formulas is, to the authors' knowledge, the simplest way that has been published.

New original formulas [Eqs. (15)–(17)] for ray tracing of a general meridional ray through the aspherical lens were derived, and formulas for approximation of the aspherical surface with

power series as well. Moreover, new procedures [Eqs. (23)–(26)] for ray tracing of a general off-axial ray were presented. The calculation can be done for an arbitrary position of a point object. Further, an approximate formula in Eq. (28) for minimization of the coma aberration of the aspherical lens was presented.

The paper brings a valuable contribution and deeper insight into the theory of aspherical surfaces with an analytic study of many challenges. Not every reader has a possibility to use modern optical programs such as Zemax, Oslo, and Code V, which are very expensive; the presented formulas can be very easily coded, and a design of aspherical lenses can be done without large financial costs.

Funding. Czech Technical University in Prague (SGS20/093/OHK1/2T/11).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. M. M. Rusinov, *Aspherical Surfaces in Optics* (Nedra, 1973).
2. B. Jurek, *Optical Surfaces* (Akademia, 1976).
3. M. Herzberger, *Modern Geometrical Optics, Pure and Applied Mathematics* (Interscience, 1958).
4. M. Born, E. Wolf, and A. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 2000).
5. B. Havelka, *Geometrical Optics I, II* (NČSAV, 1956).
6. W. Smith, *Modern Optical Engineering*, 4th ed. (McGraw-Hill Education, 2007).
7. W. Welford, *Aberrations of Optical Systems*, Series in Optics and Optoelectronics (Taylor & Francis, 1986).
8. M. Herzberger and H. O. Hoadley, "The calculation of aspherical correcting surfaces," J. Opt. Soc. Am. **36**, 334–340 (1946).
9. G. D. Wassermann and E. Wolf, "On the theory of aplanatic aspheric systems," Proc. Phys. Soc. London Sect. B **62**, 2–8 (1949).
10. K. Miyamoto, "On the design of optical systems with an aspheric surface," J. Opt. Soc. Am. **51**, 21–22 (1961).
11. J. L. Rayces and X. Cheng, "Numerical integration of the profile of aspheric surfaces," in *International Optical Design* (Optical Society of America, 2006), paper ThB3.
12. D. P. Feder, "Differentiation of ray-tracing equations with respect to construction parameters of rotationally symmetric optics," J. Opt. Soc. Am. **58**, 1494–1505 (1968).
13. C. E. Gutiérrez, "Aspherical lens design," J. Opt. Soc. Am. A **30**, 1719–1726 (2013).
14. G. Castillo-Santiago, M. Avendaño-Alejo, R. Díaz-Uribe, and L. C. Neda, "Analytic aspheric coefficients to reduce the spherical aberration of lens elements used in collimated light," Appl. Opt. **53**, 4939–4946 (2014).
15. J. C. Valencia-Estrada, R. B. Flores-Hernández, and D. Malacara-Hernández, "Singlet lenses free of all orders of spherical aberration," Proc. R. Soc. London Ser. A **471**, 20140608 (2015).
16. N. del C. Lozano-Rincón and J. C. Valencia-Estrada, "Paraboloid–aspheric lenses free of spherical aberration," J. Mod. Opt. **64**, 1146–1157 (2017).
17. M. Avendaño-Alejo, E. Román-Hernández, L. C. Neda, and V. I. Moreno-Oliva, "Analytic conic constants to reduce the spherical aberration of a single lens used in collimated light," Appl. Opt. **56**, 6244–6254 (2017).
18. R. G. González-Acuna and H. A. Chaparro-Romo, "General formula for bi-aspheric singlet lens design free of spherical aberration," Appl. Opt. **57**, 9341–9345 (2018).
19. R. G. González-Acuña, H. A. Chaparro-Romo, and J. C. Gutiérrez-Vega, "General formula to design a freeform singlet free of spherical aberration and astigmatism," Appl. Opt. **58**, 1010–1015 (2019).
20. P. D. Lin and C.-Y. Tsai, "Determination of unit normal vectors of aspherical surfaces given unit directional vectors of incoming and outgoing rays," J. Opt. Soc. Am. A **29**, 174–178 (2012).
21. A. Mikš and P. Novák, "Determination of unit normal vectors of aspherical surfaces given unit directional vectors of incoming and outgoing rays: comment," J. Opt. Soc. Am. A **29**, 1356–1357 (2012).
22. J. C. Valencia-Estrada and J. García-Márquez, "Freeform geometrical optics I: principles," Appl. Opt. **58**, 9455–9464 (2019).
23. J. C. Valencia-Estrada and J. García-Márquez, "General formula to design a freeform singlet free of spherical aberration and astigmatism: comment," Appl. Opt. **59**, 3422–3424 (2020).
24. https://www.zemax.com/products/opticstudio.
25. http://www.lambdares.com.
26. K. Rektorys, *Survey of Applicable Mathematics* (MIT, 1969).

applied optics

Spherical aberration of an optical system and its influence on depth of focus

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Applied Optics Group, Thákurova 7, 166 29 Prague, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 29 March 2017; revised 22 May 2017; accepted 22 May 2017; posted 23 May 2017 (Doc. ID 291653); published 9 June 2017

This paper analyzes the influence of spherical aberration on the depth of focus of symmetrical optical systems for imaging of axial points. A calculation of a beam's caustics is discussed using ray equations in the image plane and considering longitudinal spherical aberration as well. Concurrently, the influence of aberration coefficients on extremes of such a curve is presented. Afterwards, conditions for aberration coefficients are derived if the Strehl definition should be the same in two symmetrically placed planes with respect to the paraxial image plane. Such conditions for optical systems with large aberrations are derived with the use of geometric-optical approximation where the gyration diameter of the beam in given planes of the optical system is evaluated. Therefore, one can calculate aberration coefficients in such a way that the optical system generates a beam of rays that has the gyration radius in a given interval smaller than the defined limit value. Moreover, one can calculate the maximal depth of focus of the optical system respecting the aforementioned conditions.

OCIS codes: (080.1010) Aberrations (global); (080.2720) Mathematical methods (general); (080.2740) Geometric optical design; (350.4600) Optical engineering.

https://doi.org/10.1364/AO.56.005099

1. INTRODUCTION

The depth of focus of an optical system is a very important imaging characteristic, and it describes an area in image space where a diameter of a circle of confusion is approximately the same or less than a given threshold value. The diameter of the circle of confusion characterizes an area where a point is imaged into and which we accept as an image point. As it is known from optical theory, the image of a point is not a point due to aberrations of the optical system and diffraction effects. Rather, it is an energy distribution, so-called a point spread function [1–5]. In the case of an optical system with a circular aperture and without any aberrations, the diameter of the central part of the point spread function is called the Airy disk, and it is given by formula $d_A = 2.4\lambda F$, where λ is the wavelength and F is the f-number of the optical system; one considers imaging the axial points of a rotationally symmetric optical system, and the image is analyzed in the paraxial image plane [1–10]. If the optical system is affected by aberrations [1–10], then the diameter of the circle of confusion depends on the amount of aberrations. In the case of axial points of the rotationally symmetric optical system, the diameter of the circle of confusion depends on spherical aberration. Therefore, one can modify the diameter of the circle of confusion with an appropriate spherical aberration, which depends on the distance from the paraxial image plane. If one accepts some threshold value of the circle of confusion for practical applications, then there exists an area where the diameter of the circle of confusion is less than or equal to the considered value. Such an area is called the depth of focus of the optical system. Therefore, one can influence the depth of focus of the optical system with an appropriate modification of spherical aberration, in the case of axial points.

The goal of this paper is to present the influence of spherical aberration on the depth of focus of an optical system for the case of imaging of axial points of a rotationally symmetric optical system. The other way to achieve such an effect is with the usage of appropriate optical elements, which will change the amplitude, phase, and polarization of a passing wave field [11–20]. Interesting reviews of modifying the depth of focus are presented in [11,12]. This paper continues the topic that was presented in the previous work [21], where only the geometric-optical approach was applied. Readers can find other references discussing the topic in [21] as well. This presented paper shows a more general analysis considering the diffraction point of view together with the geometric-optical description. The practical usability is presented as well.

In the first part of this paper, analytical formulas for the calculation of caustics of a ray beam are presented with the use of ray equations in the image plane and with longitudinal

spherical aberration. Afterwards, the influence of aberration coefficients on extremes of such a curve is discussed. This part fills the theory of caustics, e.g., studies on evaluating and using caustics in symmetric optical systems as was published by Shealy and Burkhard [22–26], which were also used by Andersen for automatically calculating caustics from high-order aberration coefficients [27]. The next part focuses on the derivation of the requirements for aberration coefficients of an optical system if the Strehl definition should remain the same for two symmetrically placed planes with respect to the paraxial image plane. For optical systems with large aberrations, such conditions are derived with the use of the geometric-optical approach where the gyration diameter of the system is evaluated in given planes. Therefore, aberration coefficients of spherical aberration can be calculated in such a way that the optical system generates a beam of rays that has less or equal gyration diameter than the given threshold value. Moreover, one can calculate the maximum possible depth of focus of the optical system that fulfills the aforementioned conditions. In the end of the paper, the derived formulas are presented with practical examples. To the best of the authors' knowledge, such a general analysis of a similar topic has not been published yet. Thus, this paper gives important answers for many practical applications of the discussed topic. Results of the presented analysis have a significant impact, especially for optical scanning systems, which work generally with axial beams. Therefore, spherical aberration is fundamental.

2. RAY EQUATIONS AND CAUSTICS IN AN IMAGE PLANE

Consider the situation depicted in Fig. 1, where the axial ray tracing (aperture ray) in the image space of the rotationally symmetric optical system is presented. The image space is homogenous, and isotropic rays become straight lines. Let O be the origin of the defined coordinate system (e.g., the center of the exit pupil of the optical system), line $\overline{BC}$ denote an outgoing ray from the optical system, ξ be the paraxial image plane, and A_0 be the paraxial image of the optical system. Due to rotational symmetry, one can analyze only a meridional section. From mathematics, a line can be described by the following formula [28]:

$$y = kx + h, \tag{1}$$

where $k = k(h)$ is the tangent of the angle between the ray and the positive direction of the x axis. It holds for

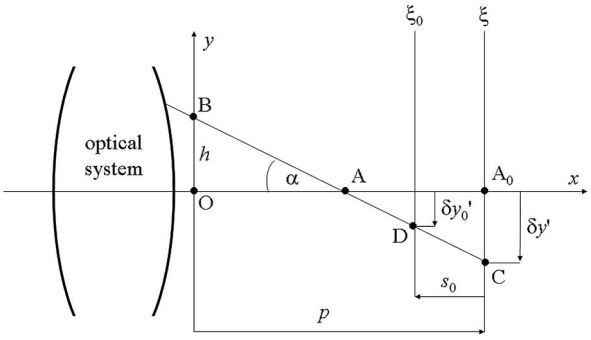


Fig. 1. Transverse spherical aberration.

$k = \tan(180° - \alpha) = -\tan(\alpha)$, and $h = \overline{OB}$. According to Fig. 1, it holds for

$$k = -(h - \delta y')/p, \tag{2}$$

where $\delta y' = \delta y'(h)$ is the transverse ray aberration of the optical system (transverse spherical aberration) [1–10]. The meaning of the other values is obvious from Fig. 1. The transverse ray aberration of the aperture ray in the paraxial image plane ξ can be, in the case of a rotationally symmetric optical system, written in the form ($\delta y'$ is an odd function of h)

$$\delta y' = \sum_{m=2}^{M} a_{2m-1} h^{2m-1}, \tag{3}$$

where a_{2m-1} denotes aberration coefficients. Substituting Eq. (3) into Eq. (2), one gets

$$k = \frac{-h + \sum_{m=2}^{M} a_{2m-1} h^{2m-1}}{p}. \tag{4}$$

As it is obvious from Eqs. (1) and (2) that the equation of a ray (line) depends only on one parameter h for a constant value of p (see Fig. 1). Therefore, those rays define a family of lines, and one can calculate an envelope of this family (i.e., a curve that is tangent to each member of the family of lines [28,29]). From mathematics (from the definition of an envelope), one can calculate the envelope of lines [given by Eq. (1)] from the following system of equations [6,7,28,29]:

$$\phi(x, y, h) = 0, \partial\phi(x, y, h)/\partial h = 0, \tag{5}$$

where

$$\phi(x, y, h) = y - kx - h = 0. \tag{6}$$

The lines' envelope is called a caustic in optic [4,6,7,29]. It can be shown that caustics are curves where the wavefronts' centers of curvature are positioned [4,6,7,29]. The caustics formula can be derived excluding the parameter h from Eq. (5). As it is obvious from Eqs. (4)–(6), such a problem is generally unsolvable. For this reason, the coordinates of caustics are expressed in the parametrical form. Using Eqs. (2)–(6), one gets formulas for caustics' coordinates $x_c = x_c(h)$ and $y_c = y_c(h)$ in the following form:

$$x_c = \frac{p}{g(h)}, \qquad y_c = h - \frac{f(h)}{g(h)}, \tag{7}$$

where

$$f(h) = h - \sum_{m=2}^{M} a_{2m-1} h^{2m-1},$$

$$g(h) = 1 - \sum_{m=2}^{M} (2m-1) a_{2m-1} h^{2(m-1)}. \tag{8}$$

In the case of an optical system without any aberrations, i.e., $a_{2m-1} = 0$ for $m = 2, 3, \ldots, M$, one gets $x_c = p$, $y_c = 0$, and therefore caustics become a point.

Now we will show another way to calculate caustics of a beam of rays. This approach is based on the usage of longitudinal spherical aberration $\delta s'$. Consider the situation in

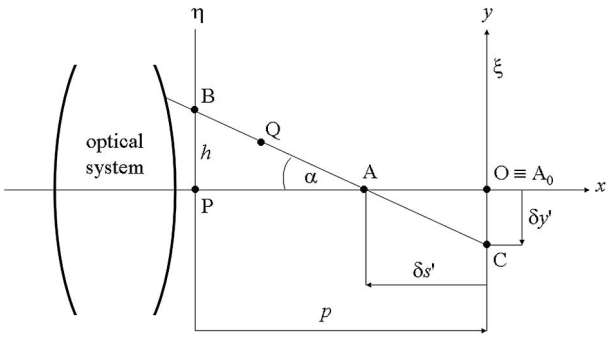


Fig. 2. Calculation of caustics of a beam of rays.

Fig. 2. Let the origin of the coordinate system O be in the paraxial image plane ξ, which is positioned at a distance p from the plane of the exit pupil η. As is obvious from Fig. 2, point O is identical with paraxial image point A_0. The ray equation of $\overline{BC}$ can be expressed in parametrical form. It holds for coordinates $x = x(\alpha, t)$, $y = y(\alpha, t)$ of an arbitrary point Q on the ray:

$$x = \delta s' + t \cos \alpha, \qquad y = t \sin \alpha, \tag{9}$$

where t denotes the distance between points Q and A. Longitudinal spherical aberration $\delta s'$ can be written in the form as follows:

$$\delta s' = \sum_{n=1}^{N} a_{2n} (\tan \alpha)^{2n}, \tag{10}$$

where s_{2n} denotes aberration coefficients.

As it is known [14,15], one can calculate an envelope of lines given by Eq. (9) (generally curves) if Jacobian J of the transformation from ray coordinates (α, t) to Cartesian (x, y) is equal to zero:

$$J = \begin{vmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial x}{\partial t} \\ \frac{\partial y}{\partial \alpha} & \frac{\partial y}{\partial t} \end{vmatrix} = 0. \tag{11}$$

Using Eqs. (9)–(11) then gives the following formula for parameter t:

$$t = \frac{2}{\cos \alpha} \sum_{n=1}^{N} n s_{2n} (\tan \alpha)^{2n}. \tag{12}$$

Substituting Eq. (12) into Eq. (9) gives the analytical formulas for the coordinates of caustics (x_c, y_c):

$$x_c = \sum_{n=1}^{N} (2n+1) s_{2n} (\tan \alpha)^{2n},$$
$$y_c = 2 \sum_{n=1}^{N} n s_{2n} (\tan \alpha)^{(2n+1)}. \tag{13}$$

Now we will study the extreme values of the caustic's coordinate y_c depending on parameter h. The necessary condition for the extreme, $\partial y_c / \partial h = 0$, and the combination with Eqs. (7) and (8) gives the following formula:

$$f(h) e(h) = 0, \tag{14}$$

where

$$e(h) = \sum_{m=2}^{m=M} 2(2m-1)(m-1) a_{2m-1} h^{2m-3}. \tag{15}$$

Solving Eq. (14) results in value h where the coordinate y_c will be external.

Let one solve Eq. (14) for the case of an optical system with aberrations of third and fifth orders, i.e., for $M = 3$. After rearrangement, Eq. (14) has the form

$$10 a_5^2 h^8 + 13 a_3 a_5 h^6 + (3 a_3^2 - 10 a_5) h^4 - 3 a_3 h^2 = 0. \tag{16}$$

The solution of Eq. (16) then gives the following values of h for extremes:

$$\begin{aligned} h_1 &= 0, & h_2 &= 0, \\ h_3 &= \frac{\sqrt{30}}{10} \frac{\sqrt{-a_3 a_5}}{a_5}, & h_4 &= -h_3, \\ h_5 &= \sqrt{-\frac{a_3 + \sqrt{a_3^2 + 4a_5}}{2a_5}}, & h_6 &= -h_5, \\ h_7 &= \sqrt{-\frac{a_3 - \sqrt{a_3^2 + 4a_5}}{2a_5}}, & h_8 &= -h_7. \end{aligned} \tag{17}$$

In the case of an optical system with aberrations of third, fifth, and seventh orders, the transverse ray aberration (in the paraxial image plane) can be written with Eq. (3) for $M = 4$ as follows:

$$\delta y' = a_3 h^3 + a_5 h^5 + a_7 h^7. \tag{18}$$

With Eq. (18), one can easily find the solution for a value of h_0 where $\delta y' = 0$:

$$h_{01} = \sqrt{-\frac{a_5 + \sqrt{a_5^2 - 4a_3 a_7}}{2a_7}},$$
$$h_{02} = \sqrt{-\frac{a_5 - \sqrt{a_5^2 - 4a_3 a_7}}{2a_7}}. \tag{19}$$

For the case of an optical system with only third- and fifth-order aberrations ($M = 3$), a similar analysis gives

$$h_0 = \pm \sqrt{-\frac{a_3}{a_5}}. \tag{20}$$

3. DEPTH OF FOCUS FROM THE POINT OF WAVE OPTICS

We will now look at the aforementioned problem from the point of wave optics. Let the optical system have only third- and fifth-order aberrations (the most important in practice). The wave aberration W of the optical system can be expressed as series of Seidel polynomials [1,3–5,9] as follows:

$$W = W_{20} r^2 + W_{40} r^4 + W_{60} r^6, \tag{21}$$

where r is the normalized ($r_{\max} = 1$) radial polar coordinate of the point in the plane of the exit pupil or the reference sphere of the optical system, W_{20} characterizes longitudinal defocus, W_{40} spherical aberration of the third order, and W_{60} is spherical aberration of the fifth order.

The Strehl ratio (Strehl definition) is defined as a ratio between the maximum value of the point spread function of a real

optical system with aberrations and the maximum value of the point spread function of an ideal optical system (only diffraction limited) without aberrations ($W = 0$). The Strehl definition *S.D.* of the optical system with small aberrations ($W < 0.5\lambda$) can be expressed with the following formula [1,3–5,9]:

$$S.D. \approx 1 - (2\pi/\lambda)^2 E_0, \tag{22}$$

where

$$E_0 = \overline{W^2} - \overline{W}^2, \overline{W} = 2\int_0^1 W(r) r\mathrm{d}r, \overline{W^2} = 2\int_0^1 W^2(r) r\mathrm{d}r. \tag{23}$$

In the aforementioned situation, one gets for the variance of wave aberration E_0 as follows [1,3,9]:

$$\begin{aligned} E_0 &= \frac{W_{20}^2}{12} + W_{20}\left(\frac{W_{40}}{6} + \frac{3W_{60}}{20}\right) \\ &\quad + \left(\frac{4W_{40}^2}{45} + \frac{W_{40}W_{60}}{6} + \frac{9W_{60}^2}{112}\right) \\ &= e_2 W_{20}^2 + e_1 W_{20} + e_0, \end{aligned} \tag{24}$$

where one can calculate aberration coefficients with the following formulas [1,3,9]:

$$\begin{aligned} W_{20} &= -\frac{s_0}{8F^2}, W_{40} = -\frac{3}{2} q_0 W_{60}, \\ W_{60} &= \frac{\Delta s'_k}{24F^2(1-q_0)} = -\frac{\Delta s'_{\mathrm{ext}}}{6F^2 q_0^2}, \end{aligned} \tag{25}$$

where F is the f-number of the optical system [1–4,9], s_0 is the defocus, $\Delta s'_k$ is the longitudinal spherical aberration at the edge of the exit pupil ($r = 1$), $q_0 = r_0^2$ is the correction zone where the longitudinal spherical aberration is corrected (i.e., it equals zero for $r = r_0$), and $\Delta s'_{\mathrm{ext}}$ is the extreme value of longitudinal spherical aberration [1,3,4].

Suppose now that the Strehl definition has to be the same in two planes symmetrically placed with respect to the paraxial image plane by the value of $\pm s_0$. If such a condition shall be fulfilled, it has to be $e_1 = 0$ according Eq. (24). The solution of this equation thus gives

$$W_{40} = -\frac{9}{10} W_{60}. \tag{26}$$

Comparing Eqs. (25) and (26) results in

$$W_{40} = -\frac{3}{2} q_0 W_{60} = -\frac{9}{10} W_{60}. \tag{27}$$

Therefore,

$$q_0 = r_0^2 = 3/5. \tag{28}$$

One can summarize the aforementioned analysis as follows. If the Strehl definition of an optical system with small aberrations for axial points should have the same value in two planes symmetrically placed with respect to the paraxial image plane, then the spherical aberration of the optical system should be corrected for zone $r_0 = \sqrt{3/5} = 0.7746$.

If $(E_0)_p$ is the demanded value of the variance of wave aberration in planes placed at distances $\pm(s_0)_p$ from the paraxial image plane, then one gets, using Eqs. (24) and (26), the following:

$$W_{40} = \pm 5.351\sqrt{12(E_0)_p - (W_{20})_p^2}, \tag{29}$$

where

$$(W_{20})_p = -\frac{(s_0)_p}{8F^2}. \tag{30}$$

According to Eq. (26),

$$W_{60} = -\frac{10}{9} W_{40} = \mp 5.946\sqrt{12(E_0)_p - (W_{20})_p^2}. \tag{31}$$

With Eqs. (29) and (31), one can calculate coefficients W_{40} and W_{60} of wave aberration of the optical system in the case of spherical aberration of the third and fifth order, which ensures that the maximum value of the Strehl definition will be the same in planes at distances $\pm(s_0)_p$ from the paraxial image plane of the optical system. For a real solution, it has to be fulfilled, according to Eqs. (29) and (30), as follows:

$$|(s_0)_p| \le 16F^2\sqrt{3(E_0)_p}. \tag{32}$$

Equation (32) can be used for the initial calculation of the maximum depth of focus with given variance $(E_0)_p$.

Note that between coefficients of wave aberration W_{40} and W_{60} and coefficients a_3 and a_5 of transverse ray aberration the following equations hold:

$$W_{40} = \frac{H^4}{4R} a_3, \qquad W_{60} = \frac{H^6}{6R} a_5, \tag{33}$$

where H is maximum (edge) height h (see Fig. 1) and R is the radius of the reference sphere. If point O (Fig. 1) is identical with the center of the exit pupil of the optical system, then $R = p$. The value r in Eq. (21) is then given as $r = h/H$.

4. CALCULATION OF ABERRATION COEFFICIENTS WITHIN THE GEOMETRIC-OPTICAL APPROXIMATION

In the case of large aberrations ($W > \lambda$, with respect to the optimal image point), it is not possible to use the Strehl definition for a relevant description of the problem of optical image quality. As presented by Miyamoto [8,30–32], the effect of diffraction of light can be neglected if the wave aberration fulfills the condition $W > 2\lambda$, and optical imaging can be described using geometric-optical approximation, based on ray aberrations of the optical system. Such results are practically identical with the ones obtained with the diffraction theory of optical imaging. Consider now a situation and the way of calculation of the depth of focus in the case of spherical aberration of the third and fifth order, where transverse spherical aberration in the paraxial image plane is given with the formula

$$\delta y' = a_3 h^3 + a_5 h^5. \tag{34}$$

The mean value of the radius r_g of geometric-optical circle of confusion (gyration radius [10,32]) can be calculated with the following formula (valid for the circular exit pupil of optical system):

$$r_g^2 = \frac{1}{\pi H^2}\int_0^{2\pi}\int_0^H (\delta y_0')^2 h\mathrm{d}h\mathrm{d}\varphi = \frac{2}{H^2}\int_0^H (\delta y_0')^2 h\mathrm{d}h$$
$$= \frac{2}{H^2}\int_0^H (\delta y' - s_0 \tan\alpha)^2 h\mathrm{d}h, \tag{35}$$

where integration is performed over the plane of the exit pupil of the optical system, (h, φ) are polar coordinates in the plane of the exit pupil, H is the maximum value of h, s_0 is the distance between the paraxial image plane and the plane where the size of the geometric-optical circle of confusion is calculated, α is the aperture angle (always positive) between the ray and the optical axis (see Fig. 1), and $\delta y_0'$ is the transverse spherical aberration in the plane ξ_0, which is located at distance s_0 from the paraxial image plane. As shown by Miyamoto [32], the geometric-optical intensity I_g of the imaged point, in the first approximation, is in straight relation with the gyration radius [32]:

$$I_g = \beta_1 + \beta_2(r_g^2), \tag{36}$$

where β_1 and β_2 are constants.

After integration of Eq. (35) with substituted Eq. (34), one obtains the following equation for gyration radius:

$$r_g^2 = g_2 s_0^2 + g_1 s_0 + g_0, \tag{37}$$

where

$$g_2 = H^2/(2p^2),$$
$$g_1 = -H^4(3H^2 a_5 + 4a_3)/(6p),$$
$$g_0 = H^6(10H^4 a_5^2 + 24H^2 a_3 a_5 + 15a_3^2)/60. \tag{38}$$

Equation (38) can be written with Eq. (33) in the following form, too, $(R = p)$:

$$g_2 = H^2/(2p^2),$$
$$g_1 = -8W_{40}/3 - 3W_{60},$$
$$g_0 = 2p^2(10W_{40}^2 + 24W_{40}W_{60} + 15W_{60}^2)/(5H^2). \tag{39}$$

Let one demand an equal gyration radius r_g in two planes, which are symmetrically placed with respect to the paraxial image plane by value $\pm s_0$. To fulfill such a condition, it holds according to Eq. (37) that $g_1 = 0$. Finding the solution of this equation, one obtains

$$a_3 = -\frac{3}{4}H^2 a_5. \tag{40}$$

Let $(r_g)_p$ be the gyration radius in planes distanced from the paraxial image plane by the value $\pm(s_0)_p$. Then the aberration coefficient a_5 can be calculated from Eqs. (37) and (40) as follows:

$$a_5 = \pm\frac{4\sqrt{210}\sqrt{2p^2(r_g^2)_p - H^2(s_0)_p^2}}{7H^5 p}. \tag{41}$$

Equations (40) and (41) can be used for the calculation of aberration coefficients a_3 and a_5 of the transverse ray aberration of the optical system, which requires the gyration radius $(r_g)_p$ in planes that are shifted by the value $\pm(s_0)_p$ from the paraxial image plane.

To obtain the real solution of our problem, one can derive the maximum possible depth of focus $2(s_0)_{\lim}$ from Eq. (41) as

$$|(s_0)_{\lim}| \le \sqrt{2}\frac{p}{H}(r_g)_p. \tag{42}$$

One can calculate the maximum possible depth of focus of the optical system in advance with Eq. (42), and the optical system will have the aforementioned properties. If the desired value of the depth of focus $2(s_0)_p$ is larger than $2(s_0)_{\lim}$, one has to change some of parameters H or p in such a way that Eq. (42) will be fulfilled.

The next value of interest is the minimum gyration radius $(r_g)_{\min}$. From the necessary condition for the extreme and with Eq. (29) one gets

$$(r_g)_{\min} = \sqrt{g_0}. \tag{43}$$

Afterwards, the minimum value of the gyration diameter $(D_g)_{\min}$ is

$$(D_g)_{\min} = 2(r_g)_{\min} = 2\sqrt{g_0}. \tag{44}$$

The correction zone can be calculated from Eqs. (25), (33), and (40) as follows:

$$q_0 = r_0^2 = 3/4. \tag{45}$$

One can summarize the aforementioned analysis as follows. If the value of gyration diameter of the optical system for axial points should have the same values in two planes symmetrically placed with respect to the paraxial image plane, then the spherical aberration of the optical system should be corrected for zone $r_0 = \sqrt{3/4} = 0.866$.

Demanding the same value of gyration radius r_g for $s_0 = s_{01}$ and $s_0 = s_{02}$, one gets the following from Eq. (37):

$$a_5 = -\frac{9H(s_{01} + s_{02}) \pm \sqrt{96D}}{7H^2 p},$$
$$a_3 = -\frac{3}{4}H^2 a_5 + \frac{3(s_{01} + s_{02})}{4H^2 p}, \tag{46}$$

where

$$D = -9H^2(s_{01}^2 + s_{02}^2) + 17H^2 s_{01}s_{02} + 70r_g^2 p^2. \tag{47}$$

For a solution in a real domain, $D \ge 0$. Therefore, it is not possible to set the values of s_{01} and s_{02} independently.

5. EXAMPLE

Let one calculate aberration coefficients using the aforementioned procedure. Suppose the following initial parameters: $(s_0)_p = \pm 2$ mm (depth of focus = 4 mm), $(D_g)_p = 1$ mm, $p = 50$ mm, and $H = 12.5$ mm.

One obtains the following with the use of Eqs. (39) and (40): $a_3 = 1.5899\mathrm{e}-3$ mm^{-2} and $a_5 = -1.3567\mathrm{e}-5$ mm^{-4}. Next, for the minimum gyration diameter it holds that $(D_g)_{\min} = 0.707$ mm. Equation (20) can be used for the calculation of height h_0 where the spherical aberration is corrected, $h_0 = 10.825$ mm.

Figure 3 shows a plot of transverse spherical aberration $\delta y'$ of the optical system, which will fulfill the aforementioned

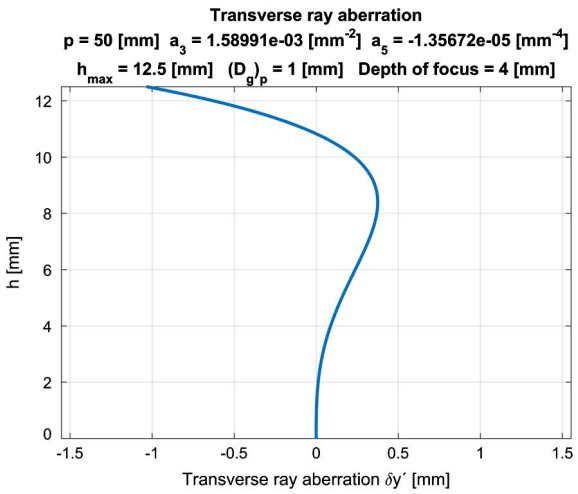


Fig. 3. Transverse spherical aberration.

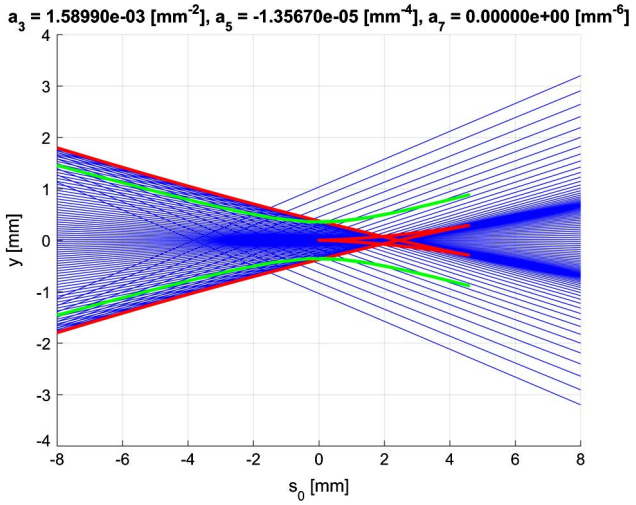


Fig. 5. Dependency of caustics (red line) of a beam of rays (blue lines) with respect to the gyration radius (green line) on defocus s_0 for a constant step in ray height h in the aperture plane.

conditions. In Fig. 4, a dependency of the gyration diameter D_g on defocus s_0 is shown. Figure 5 then shows the beam of rays and its caustics.

Now let us study the energy properties of light that are transformed by the given optical system into an area given by the gyration diameter. By splitting the aperture into subapertures of equal areas, one can easily model such a phenomenon with rays coming through the centroid of such subapertures.

Let a circular sector be defined by angle $\Delta\varphi$. Let such a sector be divided into N parts of the same area in a radial direction. Afterwards, the centroid of ith part of the annulus will be given by radius r_{Ti}, and it holds that

$$r_{\mathrm{Ti}} = \frac{4}{3\Delta\varphi}\frac{(r_i^3 - r_{i-1}^3)}{(r_i^2 - r_{i-1}^2)}\sin\frac{\Delta\varphi}{2}, \tag{48}$$

where r_i and r_{i-1} denote the edge radii of the section of the ith annulus, which can be calculated by the following recurrent formula:

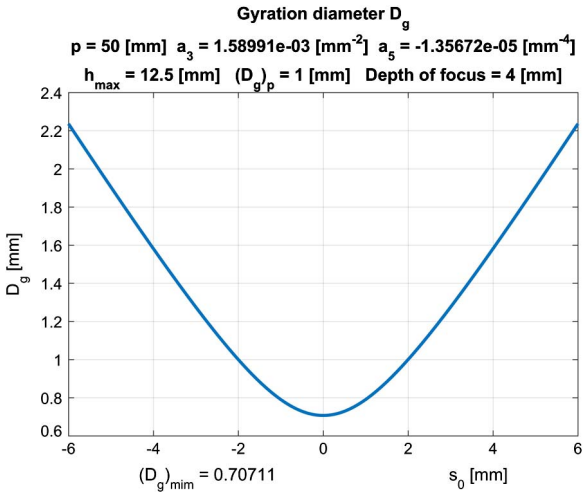


Fig. 4. Dependency of gyration diameter D_g on defocus s_0.

$$r_{i+1} = \sqrt{\frac{R^2}{N} + \Delta\phi \cdot r_i^2}, \qquad i = 0, \ldots, N, r_0 = 0. \tag{49}$$

For a limit case where $\Delta\varphi \to 0$, one gets the following from Eq. (48):

$$r_{\mathrm{Ti0}} = \frac{2}{3}\frac{(r_i^3 - r_{i-1}^3)}{(r_i^2 - r_{i-1}^2)}. \tag{50}$$

Figure 6 shows an energetic study for aperture division into equal areas. In the plane given by defocus $-s_0$ with respect to the paraxial image plane in an area surrounded by the gyration radius, there is focused 54% of energy in the plane $+s_0$ 78% of energy.

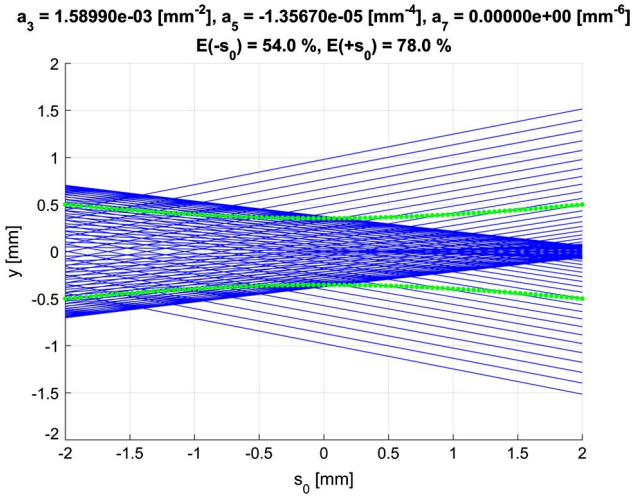


Fig. 6. Energetic study of a beam with division of aperture into equal areas (blue, rays coming from centroids of the subapertures; green, gyration radius).

6. CONCLUSION

This paper presented the problem of depth of focus of an optical system both from the point of view of geometrical optics and from the point of wave optics. The formulas for the calculation of aberration coefficients of longitudinal spherical aberration were derived as well as formulas for wave aberration. Both were studied under the condition that the optical system with such aberration coefficients will generate a beam of rays with a gyration diameter or a Strehl definition equal to or less than a chosen threshold value in a given interval of image space. Moreover, formulas for the calculation of initial maximum depth of focus were derived for an optical system fulfilling the aforementioned conditions.

Funding. Czech Technical University in Prague (CVUT) (SGS17/004/OHK1/1T/11).

REFERENCES

1. A. Mikš, *Applied Optics* (CTU, 2009).
2. S. F. Ray, *Applied Photographic Optics* (Focal, 2002).
3. E. L. O'Neill, *Introduction to Statistical Optics* (Addison-Wesley, 1963).
4. M. Born and E. Wolf, *Principles of Optics* (Oxford University, 1964).
5. W. T. Welford, *Aberrations of Optical Systems* (Adam Hilger, 1986).
6. M. Herzberger, *Modern Geometrical Optics* (Interscience, 1958).
7. O. N. Stavroudis, *The Optics of Rays, Wavefronts and Caustics* (Academic, 1972).
8. K. Miyamoto, "Wave optics and geometrical optics in optical design," Prog. Opt. **1**, 31–66 (1961).
9. A. Miks, J. Novak, and P. Novak, "Dependence of Strehl ratio on f-number of optical system," Appl. Opt. **51**, 3804–3810 (2012).
10. A. Miks and J. Novak, "Dependence of camera lens induced radial distortion and circle of confusion on object position," Opt. Laser Technol. **44**, 1043–1049 (2012).
11. Z. Zalevsky, "Extended depth of focus imaging: a review," SPIE Rev. **1**, 018001 (2010).
12. J. Ojeda-Castañeda and C. M. Gómez-Sarabia, "Tuning field depth at high resolution by pupil engineering," Adv. Opt. Photon. **7**, 814–880 (2015).
13. A. Vázquez-Villa, J. A. Delgado-Atencio, S. Vázquez-Montiel, J. Castro-Ramos, and M. Cunill-Rodríguez, "Aspheric lens to increase the depth of focus," Opt. Lett. **40**, 2842–2845 (2015).
14. S. Liu and H. Hua, "Extended depth-of-field microscopic imaging with a variable focus microscope objective," Opt. Express **19**, 353–362 (2011).
15. T. Colomb, N. Pavillon, J. Kühn, E. Cuche, C. Depeursinge, and Y. Emery, "Extended depth-of-focus by digital holographic microscopy," Opt. Lett. **35**, 1840–1842 (2010).
16. D. Elkind, Z. Zalevsky, U. Levy, and D. Mendlovic, "Optical transfer function shaping and depth of focus by using a phase only filter," Appl. Opt. **42**, 1925–1931 (2003).
17. H. Wang and F. Gan, "High focal depth with a pure-phase apodizer," Appl. Opt. **40**, 5658–5662 (2001).
18. F. Diaz, F. Goudail, B. Loiseaux, and J. P. Huignard, "Design of a complex filter for depth of focus extension," Opt. Lett. **34**, 1171–1173 (2009).
19. J. Ojeda-Castañeda and L. R. B. Valdós, "Arbitrarily high focal depth with finite apertures," Opt. Lett. **13**, 183–185 (1988).
20. C. Rivolta, "Depth of focus of optical systems with a small amount of spherical aberration," Appl. Opt. **29**, 3249–3254 (1990).
21. A. Mikš and J. Novák, "Dependence of depth of focus on spherical aberration of optical systems," Appl. Opt. **55**, 5931–5935 (2016).
22. D. L. Shealy and D. G. Burkhard, "Flux density for ray propagation in discrete index media expressed in terms of the intrinsic geometry of the deflecting surface," Opt. Acta **20**, 287–301 (1973).
23. D. G. Burkhard and D. L. Shealy, "Flux density for ray propagation in geometrical optics," J. Opt. Soc. Am. **63**, 299–304 (1973).
24. D. L. Shealy and D. G. Burkhard, "Caustic surfaces and irradiance for reflection and refraction from an ellipsoid, elliptic paraboloid, and elliptic cone," Appl. Opt. **12**, 2955–2959 (1973).
25. D. L. Shealy and D. G. Burkhard, "Analytical illuminance calculation in a multi-interface optical system," Opt. Acta **22**, 485–501 (1975).
26. D. G. Burkhard and D. L. Shealy, "Simplified formula for the illuminance in an optical system," Appl. Opt. **20**, 897–909 (1981).
27. T. B. Andersen, "Optical aberration functions: computation of caustic surfaces and illuminance in symmetrical systems," Appl. Opt. **20**, 3723–3728 (1981).
28. K. Rektorys, *Survey of Applicable Mathematics* (MIT, 1969).
29. Y. A. Kravtsov and Y. I. Orlov, *Caustics, Catastrophes and Wave Fields* (Springer, 1999).
30. K. Miyamoto, "On a comparison between wave optics and geometrical optics by using Fourier analysis. I. General theory," J. Opt. Soc. Am. **48**, 57–63 (1958).
31. K. Miyamoto, "On a comparison between wave optics and geometrical optics by using Fourier analysis. II. Astigmatism, coma, spherical aberration," J. Opt. Soc. Am. **48**, 567–575 (1958).
32. K. Miyamoto, "On a comparison between wave optics and geometrical optics by using Fourier analysis. III. Image evaluation by spot diagram," J. Opt. Soc. Am. **49**, 35–40 (1959).

1998 Vol. 59, No. 7 / 1 March 2020 / Applied Optics Research Article

applied optics

Fundamental design parameters of two-component optical systems: theoretical analysis

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 22 November 2019; revised 14 January 2020; accepted 14 January 2020; posted 14 January 2020 (Doc. ID 383985); published 25 February 2020

The paper presents a detailed theoretical analysis of two-component optical systems of Petzval objective, tele-objectives, reverse tele-objectives, and objectives of anallactic type. This type of optical system is popular in practice, especially in the field of photographic technologies and surveying devices (theodolites, levelling devices, etc.), where anallactic telescopes with inner focusing are used. The paper presents methods of designing of fundamental parameters of the objective, i.e., focal distances of the objective's components and their mutual distance, and radii of curvatures of individual surfaces if the components are cemented doublets. Further, a detailed analysis of aberration properties of those optical systems is presented.

https://doi.org/10.1364/AO.383985

1. INTRODUCTION

Optical systems composed of two optical components find wide utilization in practice [1–9]. In a field of photographic technologies, tele-objectives are the most used. The optical system of those objectives is usually made by two components; from those, the first has positive optical power, while the second has negative power. This construction has an effect of shorter length of the objective than its focal length [4–9]. Such tele-objectives are commercially offered by different companies. The next type of two-component optical system, also widely used, is the so-called Petzval objective (both components have positive optical power), reverse tele-objectives (the first element has negative power and the second has positive optical power), and objectives of observation telescopes. In the field of measurement, usually the most used are so-called anallactic telescopes, which find application in surveying devices. In such systems, the first optical component has positive optical power, while the second (the inner component) has negative. A measurement reticle is placed in the image plane of the system. Focusing on differently distanced objects is then done by moving the second component, while the distance of the tube remains constant, i.e., so-called inner focusing. For example, in geodesy, the current offer of commercial companies includes many optical devices whose telescope has an objective composed by two optical components with mentioned positive and negative optical power. The first component is usually made by two or three optical elements; the second usually has one or two lenses. The eyepiece of the telescope is a positive optical system usually made by three to five lenses. The anallactic construction of the telescope promises usage in other fields than geodesy; further, one can name possible applications in endoscopes or other micro-optical systems. As is known from the theory [10–20], a so-called anallactic point of such anallactic telescopes (a vertex of focusing cone) has no fixed position. Its position depends on a distance of a target from an entrance pupil. This movement causes an error in measurement distance [12–20]. While changing the position of the target, optical aberrations of the optical system vary as well, which causes other errors during measurement [1,21]. The inventor of the anallactic telescope is a known designer of geodetic devices, i.e., Wild [10,11]. Before his invention of anallactic telescope, telescopes with a Keppler type of construction were used. With this system, the eyepiece is moving with respect of the objective while focusing, so-called outer focusing, and a distance of the objective from a reticle, too.

This paper presents a detailed theoretical analysis of imaging properties of a general two-element optical system. Further, an aberration theory of the third order is applied for a derivation of formulas, which can be used for designing such objectives. With the presented derived formulas, one can calculate starting parameters for further optimization of the optical system of anallactic telescope, which can be processed in software OSLO, ZEMAX, or similar. Further, two examples show a detailed procedure for calculating the telephoto lens and anallactic telescope objectives.

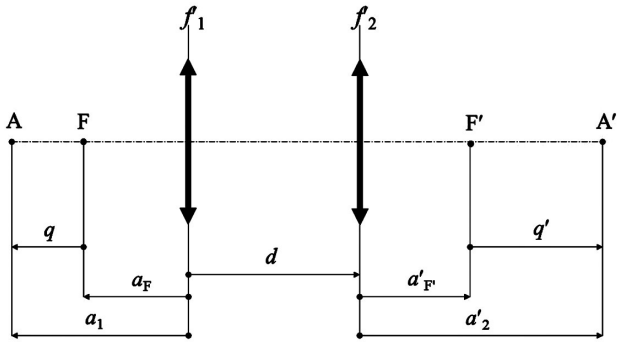


Fig. 1. Optical scheme of a two-component objective.

2. IMAGING PROPERTIES OF TWO-COMPONENT OPTICAL SYSTEMS

Suppose an optical system is composed by two optical components. The optical component means an optical system of one or more simple lenses (doublet, triplet, etc.).

In Fig. 1, there is an optical scheme of such a system shown. The meaning of individual symbols is as follows: f_1' and f_2' are focal lengths of the first and the second optical component of the objective, d is their mutual distance, a_1 is the distance between the first object axial point A and the object principal plane of the first optical component, a_2' is the distance between the image A′ of the axial object point A and the image principal plane of the second optical component, a_F is the distance between the object focal point F of the objective and the principal plane of the first component, $a'_{F'}$ is the distance between the image focal point F' and the image principal plane of the second component, q is distance between point A and the object focal point F, and q' is the distance between point A′ and the image focal point F′.

One can show that the following formulas are valid, using the fundamental imaging equations, it holds [1–3]:

$$qq' = -f'^2, \quad a_F = -f'\left(1 - d/f_2'\right),$$
$$a'_{F'} = f'\left(1 - d/f_1'\right), \quad f' = \frac{f_1' f_2'}{f_1' + f_2' - d}, \tag{1}$$

where f' is the focal length of the optical system. A transverse magnification m of the optical system than can be calculated with formula

$$m = \frac{y'}{y} = \frac{f'}{q} = -\frac{q'}{f'} = \frac{a_1' a_2'}{a_1 a_2}, \tag{2}$$

where y is the object's height and y' is height of the image. Equations (1) and (2) represent imaging properties of the two-component optical system in air.

3. FUNDAMENTALS OF ANALLACTIC OBJECTIVE CONSTRUCTION

The anallactic telescope's objective is composed of two optical components. The first has positive optical power, while the second (the inner component) has negative. A scheme of the objective is shown in Fig. 2. Without any loss of generality, there are individual components in a form of thin lenses. In the image plane, there is a reticle on a glass plate. On the plate, there

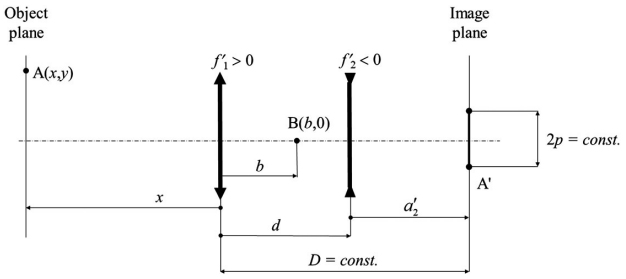


Fig. 2. Optical scheme of an anallactic objective of the telescope.

are two lines of mutual distance $2p$. Focusing on differently distanced objects is done by moving the second component, while the distance D remains constant, so-called inner focusing. The stability of distance D is appropriate in many applications, e.g., geodesy [10–25], where a mechanical prolongation of the telescope is not practical.

One can derive the following formulas with imaging equations

$$a_1' = \frac{a_1 f_1'}{a_1 + f_1'}, \quad a_2 = a_1' - d, \quad a_2' = \frac{a_2 f_2'}{a_2 + f_2'},$$
$$m = \frac{a_1' a_2'}{a_1 a_2} = \frac{p}{y}, \tag{3}$$

where a_i and a_i' $(i = 1, 2)$ denote image and object axial distances and m is the transverse magnification of the objective. Let one denote $a_1 = x$. With the use of condition $D = d + a_2' = \text{const.}$ and the formula for the transverse magnification, then, using Eq. (3), the following formulas hold:

$$d - D + \frac{f_2'\left(\frac{f_1' x}{f_1' + x} - d\right)}{f_2' + \left(\frac{f_1' x}{f_1' + x} - d\right)} = 0,$$
$$\frac{f_1' f_2' y}{\left(f_1' + x\right)\left(f_2' + \frac{f_1' x}{f_1' + x} - d\right)} - p = 0. \tag{4}$$

After elimination of distance d from Eq. (4), one obtains

$$F(x, y) = Ax^2 + Bxy + Cy^2 + Ex + Fy + G, \tag{5}$$

where

$$A = p^2 f_2', \quad B = p f_1'\left(D - f_1' - 2f_2'\right),$$
$$C = f_1'^2 f_2', \quad E = 2p^2 f_1' f_2',$$
$$F = f_1'^2 p\left(D - 2f_2'\right), \quad G = p^2 f_1'^2 f_2'. \tag{6}$$

Equation (5) describes the curve on which the point A(x, y) lies while changing the distance x. It can be easily shown that Eq. (5) represents a hyperbola with asymptotes:

$$y_{a1} = k_1 x + q_1,$$
$$y_{a2} = k_2 x + q_2, \tag{7}$$

where, using the procedure clearly described in [26] (vol. I, section 9.6), one obtains after modification and simplification

the following formulas for asymptotes parameters k and q, it holds that

$$k_{1,2} = \frac{p\left[f_1' + 2f_2' - D \mp \sqrt{(f_1' - D)(f_1' + 4f_2' - D)}\right]}{2f_1' f_2'},$$

$$q_i = \frac{p(2f_2' p - 2f_1' f_2' k_i + f_1' k_i D)}{p(f_1' + 2f_2' - D) - 2f_1' f_2' k_i}, \quad i = 1, 2.$$

Angle β between asymptotes can be calculated as follows [26]:

$$\tan\beta = \left|\frac{k_2 - k_1}{1 + k_1 k_2}\right|. \tag{8}$$

Further, let one suppose that the focal length of the objective f_0' (when the system is focused at infinity), the length D of the telescope, position of the anallactic point $b = -q_2/k_2$, mutual distance of reticle lines $2p$, and parameter k_2 (usually $k_2 = -0.01$) as known values. Afterward, focal lengths f_1' and f_2' and the distance d_0 (for and object in infinity) can be calculated with formulas:

$$G = b^2 p^2 + (2k_2 b^2 p + 6bp^2)D + (p - bk_2)^2 D^2,$$

$$f_1' = -\frac{bp - pD - bk_2 D + \sqrt{G}}{4p + 2bk_2 - 2k_2 D},$$

$$f_2' = \frac{f_1'^2 k_2 p - f_1' k_2 p D}{f_1'^2 k_2^2 - 2f_1' k_2 p + p^2},$$

$$d_0 = f_1' + f_2' - f_1' f_2'/f_0'. \tag{9}$$

While focusing the object in the finite distance x from the first optical component of the objective, distance d_x between the objective's components can be calculated as follows:

$$d_x = \frac{H + f_1' x + D(f_1' + x)}{2(f_1' + x)},$$

$$H = \sqrt{[f_1' x - D(f_1' + x)][4f_1' f_2' + (f_1' + 4f_2')x - D(f_1' + x)]}. \tag{10}$$

Let one examine the difference between the asymptotes and the curve of point A(x, y). The y coordinate of the point A(x, y) can be calculated with the following formula:

$$y = \frac{p}{2f_1' f_2'}\left[(f_1' + x)(2f_2' - D) + f_1' x - H\right]. \tag{11}$$

Afterward, the image height y' can be calculated as follows:

$$y' = \frac{f_1' f_2'}{f_1' x - (f_1' + x)(d_x - f_2')} y. \tag{12}$$

If point A(x, y) lies on the second asymptote, its image height y_a' can be calculated with formula

$$y_a' = U - a'\left(\frac{U}{f_2'} - k_2 + \frac{q_2}{f_1'}\right), \quad U = q_2 + d_x\left(k_2 - \frac{q_2}{f_1'}\right). \tag{13}$$

Afterward, the difference of image heights $\Delta y' = y' - y_a'$ can be easily calculated with Eqs. (12) and (13). Distance x_B between point A(x, y) and anallactic point B, then, can be calculated with the simple formula $x_B = -y_a/k_2$. It can be shown that, for longer distances x_B, the difference of image heights $\Delta y'$ is small. Thus, measuring a value of y, the distance x can be easily calculated knowing the asymptote's parameters. However, for measurement of short distances, a correction from deviation of asymptote from its parent curve has to be considered.

5. SEIDEL ABERRATION COEFFICIENTS

This part of the paper is focused on aberration properties of the optical system. Suppose now that the optical system is composed by thin lenses. Seidel aberration coefficients (Seidel sums) S_{I}, S_{II}, S_{III}, S_{IV}, and S_{V} can be calculated as follows [18,27,28]:

$$S_{\mathrm{I}} = \sum_{i=1}^{K} h_i^4 M_i, \quad S_{\mathrm{II}} = \sum_{i=1}^{K} h_i^3 \bar{h}_i M_i + \sum_{i=1}^{K} h_i^2 N_i,$$

$$S_{\mathrm{III}} = \sum_{i=1}^{K} h_i^2 \bar{h}_i^2 M_i + 2\sum_{i=1}^{K} h_i \bar{h}_i N_i + \sum_{i=1}^{K} \varphi_i, \; S_{\mathrm{IV}} = \sum_{i=1}^{K} p_i \varphi_i$$

$$S_{\mathrm{V}} = \sum_{i=1}^{K} h_i \bar{h}_i^3 M_i + 3\sum_{i=1}^{K} \bar{h}_i^2 N_i + \sum_{i=1}^{K} \frac{\bar{h}_i}{h_i}(3 + p_i)\varphi_i, \tag{14}$$

where S_{I} denotes the Seidel sum for the spherical aberration, S_{II} is the Seidel sum for the coma, S_{III} is the Seidel sum for the astigmatism, S_{IV} is the Seidel sum for the field curvature (the Petzval sum), and S_{V} is the Seidel sum for the distortion. In Eq. (14), the variables have the following meaning: h_i is the impinging height of the paraxial aperture ray on the ith lens, $\bar{h}_i$ is the impinging height of the paraxial chief ray on the ith lens, $p_i = 1/n_i$, n_i is the so-called equivalent index of refraction of the ith component of the optical system (a thin component can be composed by many thin lenses in contact), $\varphi_i = 1/f_i'$ is the optical power of the ith optical component, and f_i' is its focal distance. Parameters M_i and N_i characterize the spherical aberration and the coma aberration of the ith component [18,27,28].

If the optical system is made by K thin lenses, the chromatic sums of the optical system are given as follows [1–3,18]:

$$C_{\mathrm{I}} = \sum_{i=1}^{K} h_i^2 \frac{\varphi_i}{\nu_i} P_{i\lambda}, \quad C_{\mathrm{II}} = \sum_{i=1}^{K} h_i \bar{h}_i \frac{\varphi_i}{\nu_i} P_{i\lambda},$$

where C_{I} is the chromatic sum for the lateral chromatic aberration, C_{II} is the chromatic sum for the transversal chromatic aberration, ν_i denotes the Abbe number of the ith lens, and $P_{i\lambda}$ is the relative partial dispersion of the ith lens. The optical system will be corrected from chromatic aberrations if $C_{\mathrm{I}} = 0$ and $C_{\mathrm{II}} = 0$.

Consider now a general two-component optical system, which fulfils the following conditions: $\varphi = 1$, $S_{\mathrm{I}} = S_{\mathrm{II}} = S_{\mathrm{III}} = S_{\mathrm{V}} = 0$, and $S_{\mathrm{IV}} \neq 0$ or $S_{\mathrm{IV}} = 0$. The solution of Eq. (14) for the object position $s_1 = \infty$ and any arbitrary value of entrance pupil position $\bar{s}_1$ gives the following formulas:

$$M_1 = -\frac{M_F}{d^3}(\varphi_1 d - 1)^2, \quad M_2 = \frac{M_F}{d^3(\varphi_1 d - 1)^2},$$
$$N_1 = \frac{(\varphi_1 d - 1)[(\varphi_2 - 2\varphi_1)d + N_F]}{d^2},$$
$$N_2 = \frac{(2\varphi_2 - \varphi_1)d + N_F}{d^2(\varphi_1 d - 1)}, \tag{15}$$

where

$$M_F = 2p(\varphi_1^2 d - \varphi_1 - \varphi_2 + \varphi_1\varphi_2 d)\bar{s}_1 + 3(\varphi_2 - \varphi_1)d + 2p\varphi_2 d,$$
$$N_F = p(\varphi_1^2 d - \varphi_1 - \varphi_2 + \varphi_1\varphi_2 d)\bar{s}_1 + p\varphi_2 d.$$

Afterward, substituting Eq. (15) into formulas

$$\bar{M}_i = f'^3 M_i - 2f_i'^2 N_i Y_i + 1.56 Y_i^2,$$
$$\bar{N}_i = f_i'^2 N_i - 1.31 Y_i, \quad i = 1, 2,$$

one obtains the values of $\bar{M}_1$, $\bar{M}_2$, $\bar{N}_1$, and $\bar{N}_2$:

$$Y_i = \frac{s_i' + s_i}{s_i' - s_i} = \frac{m_i + 1}{m_i - 1} = -1 - \frac{2}{s_i\varphi_i},$$
$$Y_{i+1} = \frac{h_i\varphi_i}{h_{i+1}\varphi_{i+1}}(Y_i - 1) - 1, \tag{16}$$

where s_i, s_i' denote axial object and image distance of the ith component. Parameters $\bar{M}_i$ and $\bar{N}_i$ characterize the spherical aberration and the coma aberration of the ith component, which has unit focal distance and unit transversal magnification ($m = -1$) [27,28].

Let one demand the two-component optical system ($K = 2$) having corrected all aberrations of the third order. Afterward, the solution of the system in Eqs. (14) or (15) for $\varphi = 1$, $S_{\mathrm{I}} = S_{\mathrm{II}} = S_{\mathrm{III}} = S_{\mathrm{IV}} = S_{\mathrm{V}} = 0$, and $p_1 = p_2 = p$ results in following formulas:

$$\varphi_1 = -\varphi_2, \quad d = 1/\varphi_1^2,$$
$$M_1 = 2\varphi_1^3(\varphi_1 - 1)^2(p + 3), \quad N_1 = \varphi_1^2(\varphi_1 - 1)(p + 3),$$
$$M_2 = -2\varphi_1^7(p + 3)/(\varphi_1 - 1)^2, \; N_2 = \varphi_1^4(p + 3)/(\varphi_1 - 1). \tag{17}$$

It is known from the aberration theory of the third order [2,3] that, if the aberration coefficient S_q equals zero, then the aberration coefficient S_{q+1} does not depend on a position of the entrance pupil ($q = \mathrm{I}$, II, III, IV, V). In the presented case, the aberration coefficients equal zero ($S_{\mathrm{I}} = S_{\mathrm{II}} = S_{\mathrm{III}} = S_{\mathrm{IV}} = S_{\mathrm{V}} = 0$); therefore, Eq. (17) does not depend on the entrance pupil position. Without any loss of generality, one can then consider $\bar{h}_1 = 0$ or $\bar{h}_2 = 0$, etc. In the case of $S_{\mathrm{I}} = S_{\mathrm{II}} = S_{\mathrm{III}} = S_{\mathrm{IV}} = S_{\mathrm{V}} = 0$, the parameters $\bar{M}_i$ and $\bar{N}_i$ of the individual optical component of the system are calculated as follows:

$$\bar{M}_1 = 7.24\varphi_1(\varphi_1 - 1) + 1.56, \; \bar{N}_1 = 3.62\varphi_1 - 2.31,$$
$$\bar{M}_2 = \frac{7.24\varphi_1^3(\varphi_1 - 1) - 5.68\varphi_1^2 + 3.12\varphi_1 + 1.56}{(\varphi_1 - 1)^2},$$
$$\bar{N}_2 = \frac{3.62\varphi_1^2 - 1.31(\varphi_1 + 1)}{(\varphi_1 - 1)}. \tag{18}$$

Using the formulas in Section 5, one can calculate the values of radii of curvatures of individual lenses. These values are starting points for the next optimization process with optical design software, e.g., ZEMAX [29], OSLO [30], etc.

6. CALCULATION OF CEMENTED DOUBLET'S PARAMETERS

This part of the paper presents a calculation of radii of curvatures of a cemented doublet for focal length equal $f' = 1$ mm. In the case of different focal lengths, the calculated radii of curvature have to be multiplied by the final focal length's value.

The solution in Eq. (18) gives values of parameters $\bar{M}$ and $\bar{N}$ for individual components of the optical system. If the condition $1 < \bar{M} - 0.86\bar{N}^2 < 2.76$ is satisfied, then the optical component can be realized as a simple lens [27,28]. If it is not the case, the component can be manufactured usually as a cemented doublet or two cemented doublets.

The radii of curvatures r_1, r_2, and r_3 of the cemented doublet can be calculated, knowing the values of $\bar{M}$ and $\bar{N}$; thus, with the solution of equations presented in [27,28], it holds

$$\Phi_1 = \frac{\nu_1(1 + C_1\nu_2)}{\nu_1 - \nu_2}, \quad \Phi_2 = 1 - \Phi_1,$$
$$K_1 = -\Phi_2/2, \quad K_2 = \Phi_1/2, \tag{19}$$
$$A_i = \Phi_i\frac{n_i + 2}{n_i}, \quad B_i = \Phi_i K_i,$$
$$C_i = \Phi_i^3\left[\frac{n_i}{2(n_i - 1)}\right]^2, \quad E_i = \Phi_i\frac{n_i + 1}{n_i}, \tag{20}$$
$$T = \frac{1}{2}\left(\frac{\Phi_1}{n_1 - 1} + \frac{\Phi_2}{n_2 - 1} + 1\right), \tag{21}$$
$$R_1 = \frac{E_2 T + \bar{N}}{E_1 + E_2}, \quad R_2 = R_1 - T,$$
$$Q_i = A_i R_i^2 - 2B_i R_i + C_i, \quad i = 1, 2, \tag{22}$$
$$\rho_1 = R_1 + K_1 + \frac{\Phi_1}{2(n_1 - 1)}, \quad \rho_2 = R_1 + K_1 - \frac{\Phi_1}{2(n_1 - 1)},$$
$$\rho_3 = R_2 + K_2 - \frac{\Phi_2}{2(n_2 - 1)}, \tag{23}$$
$$r_j = 1/\rho_j, \quad j = 1, 2, 3, \tag{24}$$
$$\left|Q_1 + Q_2 - \bar{M}\right| \le \delta\bar{M}. \tag{25}$$

In Eqs. (19)–(25), Φ_i is the optical powers of lenses, n_i is the indices of refraction of lenses materials, ν_i is the Abbe numbers of lenses materials of the doublet, and C_{I} is the coefficient of the chromatic position aberration of the doublet.

Calculation is processed in a way that one searches such a pair of lens materials to satisfy Eq. (25) with demanded accuracy. A value of $\delta\bar{M}$ can be selected $\delta\bar{M} \approx 0.02$, for example. Equations (19)–(25) can be easily coded in a computing software, e.g., MATLAB [31]. The calculation for a materials catalogue of optical glasses can then be processed in a fraction of second.

7. EXAMPLES

A. Example 1: Calculation of Two-Component Optical System with Two Doublets with Optical Powers $\varphi_1 = -\varphi_2$

For example, if one selects $\varphi_1 = 1.5\ \mathrm{mm}^{-1}$, the following values of paraxial parameters of the objective hold: $f' = 1$ mm, $f_1' = 0.6667$ mm, $f_2' = -0.6667$ mm, $s_1 = \infty$, $h_1 = 1$ mm, $h_2 = 0.3333$ mm, $\bar{h}_1 = 0$ mm, $\bar{h}_2 = 0.4444$ mm, $d = 0.4444$ mm, and $D = 0.7778$ mm. Substituting $\varphi_1 = 1.5\ \mathrm{mm}^{-1}$ in Eq. (18), one obtains $\bar{M}_1 = 6.99$, $\bar{N}_1 = 3.12$, $\bar{M}_2 = 22.71$, and $\bar{N}_2 = 9.74$. These values can be used to determine a structure of individual objective components [27,28]. It is obvious from parameters $\bar{M}$ and $\bar{N}$ that the components can be simple cemented doublets. With formulas presented in Section 5, the parameters of cemented doublets presented in Tables 1 and 2 are calculated (length units of radii of curvatures are in mm). In the tables, n denotes the index of refraction of the glass, ν the Abbe number, and r represents radii of the curvature. Doublets are achromatic (for wavelengths λ_F and λ_C), i.e., the coefficient of chromatic aberration $C_{\mathrm{I}} = 0$. Monochromatic aberrations are calculated for wavelength λ_d.

Figure 3 presents results of MTF curves calculated in OSLO software for the first possibility of the objective (N-BAK4/F2) – (N-BAK1/N-LAF2) for the focal length of the objective $f' = 100$ mm, the f-number $F = 6.2$, and the field-of-view angle $2w = 12°$. Radii of curvatures of lenses of the first and the second component of the objective, presented in Tables 1 and 2, are obtained by multiplying the tabled values by 66.666 for the first component and by −66.666 for the second component.

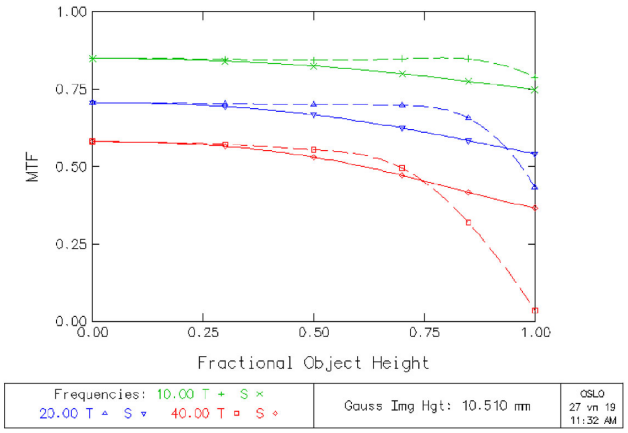


Fig. 3. Polychromatic MTF. An image plane shifted by $s_0 = 0.107$ mm with respect to the paraxial image plane.

The distance d is multiplied by 100. Therefore, the parameters of the system with the objective's focal length $f' = 100$ mm are calculated. Indices of refraction are given for wavelength λ_d. In Fig. 3, it is shown that the imaging quality is good even if the objective's parameters are calculated with the third-order aberrations only. Therefore, it is a good starting point for further optimization calculations with optical design software ZEMAX, OSLO, etc.

B. Example 2: Calculation of Anallactic Two-Component Optical System with Two Doublets

Now let one suppose a situation of the anallactic objective with the following parameters: the focal length of the objective for an object point in infinity $f_0' = 200$ mm, the length of the telescope $D = 150$ mm, parameter $k_2 = -0.01$, $p = -2$ mm, and position of the anallactic point $b = 75$ mm. Then, with the use of Eq. (9), one obtains $f_1' = 132.8724$ mm, $f_2' = -101.0087$ mm, $s_1 = \infty$, $d = 98.9701$ mm. Let the components be cemented doublets. Then, for a calculation of their parameters, one should use modified parameters corresponding to $f' = 1$ mm. In such a case, the optical powers and distance between the components would be $\varphi_1 = 1.5052\ \mathrm{mm}^{-1}$, $\varphi_2 = -1.9800\ \mathrm{mm}^{-1}$, $d = 0.4949$ mm,

Table 1. Parameters of the First Component

First Component: $\bar{M}_1 = 6.99$, $\bar{N}_1 = 3.12$, $C_{\mathrm{I}} = 0$, $f' = 1$ mm.

Glass 1	n_1	ν_1	Glass 2	n_2	ν_2	r_1 [mm]	r_2 [mm]	r_3 [mm]
N-BAK4	1.57125	55.70	F2	1.62408	36.11	0.38448	−0.42081	1.73245
N-BAF52	1.61173	46.30	N-SF14	1.76859	26.32	0.39630	−0.79064	2.22643
N-K5	1.52458	59.22	SF5	1.67764	31.97	0.38239	−0.65460	4.91016

Table 2. Parameters of the Second Component

Second Component: $\bar{M}_2 = 22.71$, $\bar{N}_2 = 9.74$, $C_{\mathrm{I}} = 0$, $f' = 1$ mm.

Glass 1	n_1	ν_1	Glass 2	n_2	ν_2	r_1 [mm]	r_2 [mm]	r_3 [mm]
N-BAK1	1.57487	57.27	N-LAF2	1.74791	44.57	0.20761	−0.33031	0.60065
LLF1	1.55099	45.47	N-LAF7	1.75459	34.56	0.20953	−0.35822	0.71106
N-BAF4	1.60897	43.43	N-LAF7	1.75459	34.56	0.20805	−0.30925	0.51819

Table 3. Parameters of the First and Second Components of the Anallactic Telescope

First Component: $\bar{M}_1 = 5.39$, $\bar{N}_1 = 2.72$, $C_I = 0$, $f' = 1$ mm.								
Glass 1	n_1	v_1	**Glass 2**	n_2	v_2	r_1 [mm]	r_2 [mm]	r_3 [mm]
N − K5	1.52458	59.22	N-SF8	1.69413	31.06	0.41714	−0.62050	−44.26320
Second Component: $\bar{M}_1 = 8.60$, $\bar{N}_1 = 7.03$, $C_I = 0$, $f' = 1$ mm.								
Glass 1	n_1	v_1	**Glass 2**	n_2	v_2	r_1 [mm]	r_2 [mm]	r_3 [mm]
SF5	1.67764	31.97	SF11	1.79190	25.55	0.26398	−0.28086	0.68256

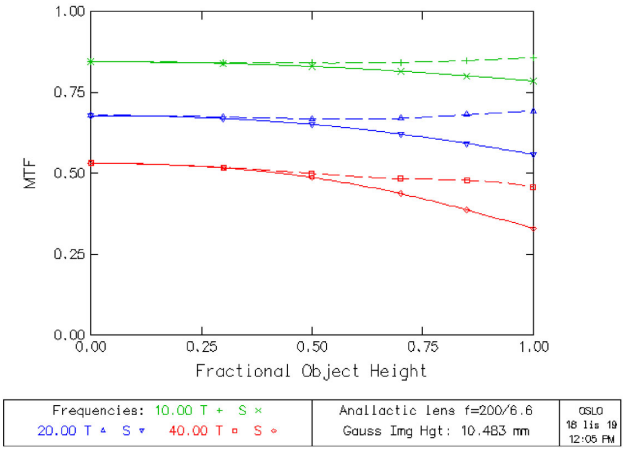


Fig. 4. Polychromatic MTF for imaging with the anallactic lens with parameters from Example 3.

$p_1 = p_2 = p = 0.62$ mm, $\bar{s}_1 = 0$. Afterward, one obtains $\bar{M}_1 = 5.39$, $\bar{N}_1 = 2.72$, $\bar{M}_2 = 8.60$, and $\bar{N}_2 = 7.03$. With the formulas presented in Section 5, the parameters of cemented doublets for each component presented in Table 3 are calculated (length units of radii of curvatures are in mm). In the tables, n denotes the index of refraction of the glass, v the Abbe number, and r is the radii of curvature. To obtain values of radii of curvature of an individual doublet's surface, one has to multiply the tabled value by the corresponding focal length. Therefore, one obtains for the first component $r_1 = 55.4264$ mm, $r_2 = -82.4473$ mm, and $r_3 = -5881.3576$ mm, and, for the second component, $r_1 = -26.6643$ mm, $r_2 = 28.3693$ mm, and $r_3 = -68.9445$ mm.

Figure 4 shows MTF curves for the presented objective calculated in OSLO. It is shown that the imaging quality is good even if the objective's parameters are calculated with the third-order aberrations only. Therefore, it is a good starting point for further optimization calculations with optical design software ZEMAX, OSLO, etc.

8. CONCLUSION

The paper presented a detailed analysis of fundamental imaging properties of a two-component optical systems (Petzval lens, telephoto and inverse telephoto lens, and anallactic objective lens). Detailed theoretical analysis of those optical systems was presented considering both the paraxial and aberrations properties. The aberration properties were analysed with the third-order theory (Seidel aberrations). General Eqs. (15), (17), and (18) were derived, which can be used for calculation of fundamental optical parameters of those optical systems (focal lengths f_1' and f_2' of components of the optical system, its mutual distance d, radii of curvature of individual components, etc.). The calculated parameters can be used as a good starting point for further optimization calculations in optical design software, e.g., ZEMAX or OSLO.

Funding. České Vysoké Učení Technické v Praze, The Czech Technical University in Prague (SGS18/105/OHK1/2T/11).

REFERENCES

1. A. Mikš, *Applied Optics* (CTU, 2009).
2. H. Chretien, *Calcul des Combinaisons Optiques* (Masson, 1980).
3. W. T. Welford, *Aberrations of the Symmetrical Optical Systems* (Academic, 1974).
4. S. F. Ray, *Applied Photographic Optics* (Focal, 2002).
5. R. Kingslage and R. B. Johnson, *Lens Design Fundamentals* (Elsevier, 2010).
6. H. Gross, F. Blechinger, and B. Achtner, *Handbook of Optical Systems, Volume IV: Survey of optical instruments* (Wiley, 2008).
7. W. Smith, *Modern Optical Engineering*, 4th ed. (McGraw-Hill, 2007).
8. M. Laikin, *Lens Design*, 4th ed. (CRC Press, 2006).
9. T. Kazamaki and F. Kondo, "New series of distortionless telephoto lenses," J. Opt. Soc. Am. **46**, 22–31 (1956).
10. H. Wild, "Neue Nivellierinstrumente," Z. Instrumentenkd. **29**, 329–344 (1909).
11. H. Wild, "Der neue Theodolit," Schweiz. Z. Vermess. Kulturtech. **23**, 103–105 (1925).
12. O. Eggert, "Ein Beitrag zur Theorie des Fernrohrs mit Fokussierlinse," Z. Vermess. **23**, 833–841 (1929).
13. H. Schulz, "Der anallaktische Punkt beim Fernrohr mit innerer Einstellinse," Z. Istrumentenkd. **56**, 357–360 (1936).
14. R. Roelofs, "Fadendistanzmesser mit Innenfokussierung," Z. Istrumentenkd. **61**, 137–147 (1941).
15. S. V. Jelisejev, *Geodetic Instruments and Equipment* (Nedra, 1973).
16. B. Havelka, "Distance-measuring telescope with inner focusing," Phys. Technol. **1**, 257–265 (1946).
17. E. Keprt, *Construction of Surveying Devices* (1951).
18. B. Havelka, *Geometrical Optics I, II* (NČSAV, 1956).
19. F. Deumlich, *Instrumentenkunde der vermessungstechnik* (VEB Verlag für Bauwesen, 1967).
20. W. Uhink, "Betrachtungen über Fernrohre mit Entfernungmeßfäden," Z. Instrumentenkd. **52**, 435–442 (1932).
21. A. König and H. Köehler, *Die Fernrohre und Entfernungsmesser* (Springer, 1959).
22. E. O. Henrici, "Use of telescopes with internal focusing for stadia surveying," Trans. Opt. Soc. **22**, 20 (1920).
23. J. E. Jackson, "Telescopes anallactic or otherwise," Surv. Rev. **23**, 51–58 (1975).
24. S. P. Roy, "The design of anallactic surveying telescopes with a positive focusing lens," Surv. Rev. **25**, 130–142 (1979).
25. S. K. Duggal, *Surveying I* (McGraw-Hill, 2011).
26. K. Rektorys, *Survey of Applicable Mathematics I, II* (Kluwer, 1991).
27. A. Mikš, J. Novák, and P. Novák, "Method of zoom lens design," Appl. Opt. **47**, 6088–6098 (2008).
28. A. Mikš, "Modification of the formulas for third-order aberration coefficients," J. Opt. Soc. Am. A **19**, 1867–1871 (2002).
29. https://www.zemax.com/.
30. https://www.lambdares.com/.
31. https://www.mathworks.com/.

Optics and Lasers in Engineering 50 (2012) 1517–1521

Contents lists available at SciVerse ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

Analytical expressions for the circle of confusion induced by plane-parallel plate

Antonin Miks *, Petr Pokorny

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thakurova 7, 166 29 Prague 6, Czech Republic

ARTICLE INFO

Article history:
Received 21 April 2012
Received in revised form
15 June 2012
Accepted 19 June 2012
Available online 6 July 2012

Keywords:
Geometrical optics
Aberrations
Optical imaging

ABSTRACT

This article provides a detailed theoretical analysis of the imaging properties of a plane-parallel plate, including exact formulas derived for the calculation of the ray aberrations and the wave aberrations induced by the plane-parallel plate and formulas for the 3rd and 5th order aberration. Formulas for a diameter of the circle of confusion are derived for both the 3rd and the 5th order aberration, as well as relations for optimal location of the image point and shape of the optimal image surface.

1. Introduction

The plane-parallel plate [1–9] is widely used in numerous optical systems. Many optical prisms [4] are generally equivalent to the plane-parallel plate. It is therefore important to know their influence on imaging properties for the optical system, which they are part of. Detailed calculation of the wave aberration induced by the plane-parallel plate is explained in the paper [2] by Braat. The objective of this paper is to perform a general analysis of the geometric-optical properties of the plane-parallel plate and its influence on image quality. The ray aberration and the circle of confusion (RMS spot size) were selected as criteria for image quality. The advantage of the circle of confusion over the wave aberration is that it indicates the size of the area, where almost all energy of the beam of rays is concentrated in. Thus, point is not imaged as a point, but as certain area with a specific intensity profile and size. The circle of confusion affects the image sharpness. Using the diameter of the circle of confusion one can very roughly determine (estimate) the *MTF* (Modulation Transfer Function) of the optical system for low spatial frequencies and vice versa. Robb's paper [10] provides very complex formulas to calculate the RMS image size for an annular entrance pupil. Robb's work does not apply the explicit equations for the plane-parallel plate. Using formulas from Robb's work to apply them to the plane-parallel plate would result in very complex and lengthy calculations including the equations from Buchdahl's paper [11]. This is very impractical and difficult to use in practice.

This paper provides exact formulas for calculating the transverse ray aberrations induced by the plane-parallel plate. Furthermore, the paper presents a detailed analysis of ray aberrations, the wave aberration, the circle of confusion, the optimal position of the image point and the shape of the optimal image surface for the 3rd and the 5th order aberration. The formulas used are very clear and simple, hence very useful and easy to use in practice (for example, for the lens design and initial lens design etc.). To our best knowledge, there is neither the paper nor the book with publication of the same relations.

2. Ray aberrations induced by plane-parallel plate

The beam of rays is considered to pass through the plane-parallel plate with a thickness d made of a material (for example, optical glass) with a refractive index n. The above explained situation is shown in Fig. 1.

The surrounding medium in front of and behind the plate is assumed to be air. Further on the plane-parallel plate a homocentric beam of rays strikes having a peak at a point B, lying in the plane ξ_0. Next two rays of this homocentric beam of rays are considered. The first ray intersect the plate at a point O (0,0,0), which has been chosen as the origin of the coordinate system. Let $\mathbf{s}_0$ be a unit direction vector of this ray. For the ray refraction by the law of refraction [1,3] the following equation holds:

$$\mathbf{s}_0' = \mathbf{s}_0/n + (\cos\varepsilon_0' - \cos\varepsilon_0/n)\mathbf{k} \quad (1)$$

where $\mathbf{s}_0$' is a unit direction vector of the refracted ray, $\mathbf{k}=(0,0,1)$ is a normal vector to ξ_1, which is the first surface of the plate, ε_0 and ε_0' denote the angle of incidence and the refraction angle of

* Corresponding author. Tel.: +420 2 24354948.
E-mail address: miks@fsv.cvut.cz (A. Miks).

0143-8166/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.optlaseng.2012.06.010

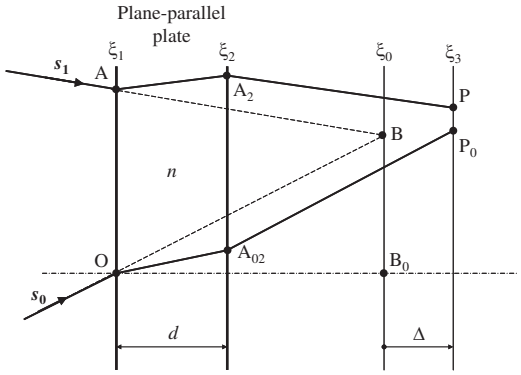


Fig. 1. Refraction of beams by the plane-parallel plate.

the beam on the first surface of the plate. After the refraction at the surface ξ_2 the ray will continue in direction of the unit vector $\mathbf{s}_0''=\mathbf{s}_0$ and intersect the plane ξ_3 at a point P_0, which is given by the vector

$$\mathbf{r}_{03}=\left(\frac{d}{n\cos\varepsilon_0'}+\frac{\mathbf{bk}+\Delta-d}{\cos\varepsilon_0}\right)s_0+d\left(1-\frac{\cos\varepsilon_0}{n\cos\varepsilon_0'}\right)\mathbf{k} \tag{2}$$

where $\mathbf{b}$ is a position vector of the point B and $\Delta=d(n-1)/n$ is the paraxial displacement caused by the plate [4]. The ray defined by the unit direction vector $\mathbf{s}_1$ strikes the plane ξ_1 at a point A, which is given by the vector $\mathbf{a}$. After the refraction on the plane ξ_2 the ray will continue in the direction of the unit vector $\mathbf{s}_1''=\mathbf{s}_1$ and intersect the plane at a point P, which is given by the vector

$$\mathbf{r}_{13}=\mathbf{a}+\left(\frac{d}{n\cos\varepsilon_1'}+\frac{\mathbf{bk}+\Delta-d}{\cos\varepsilon_1}\right)\mathbf{s}_1+d\left(1-\frac{\cos\varepsilon_1}{n\cos\varepsilon_1'}\right)\mathbf{k} \tag{3}$$

For the transverse ray aberration $\boldsymbol{\delta r}'=(\delta x',\delta y',\delta z')$ in the plane ξ_3, with respect to the ideal image of point B (i.e., the point which is determined by the vector $\mathbf{b}+\Delta\mathbf{k}$), the following exact relation is derived.

$$\begin{aligned}\boldsymbol{\delta r}'=\mathbf{r}_{13}-(\mathbf{b}+\Delta\mathbf{k})&=\mathbf{a}+\left(\frac{d}{n\cos\varepsilon_1'}+\frac{\mathbf{bk}-d/n}{\cos\varepsilon_1}\right)\mathbf{s}_1\\&\quad+d\left(1-\frac{\cos\varepsilon_1}{n\cos\varepsilon_1'}\right)\mathbf{k}-(\mathbf{b}+\Delta\mathbf{k}).\end{aligned} \tag{4}$$

Eq. (4) allows to perform an exact (accurate) calculation of the ray aberrations of the plane-parallel plate in general.

Assuming that the angle of incidence ε of the rays on the plate is not too large, $\varepsilon\le 30^\circ$ (area of the 3rd and 5th order), $\cos\varepsilon$ and $\cos\varepsilon'$ can be simplified

$$\begin{aligned}&\cos\varepsilon\approx 1-\frac{\sin^2\varepsilon}{2}-\frac{\sin^4\varepsilon}{8},\quad \cos\varepsilon'\approx 1-\frac{\sin^2\varepsilon}{2n^2}-\frac{\sin^4\varepsilon}{8n^4}\\&\frac{1}{\cos\varepsilon'}\approx 1+\frac{\sin^2\varepsilon}{2n^2}+\frac{3\sin^4\varepsilon}{8n^4},\quad \frac{1}{\cos\varepsilon}\approx 1+\frac{\sin^2\varepsilon}{2}+\frac{3\sin^4\varepsilon}{8}\end{aligned} \tag{5}$$

Let vector $\mathbf{b}=(0,b_y,b_z)$. If w is denoted as the half-angle of the field of view, $\tan w=b_y/b_z$ is valid. According to Fig. 1 the following relations can be easily derived ($\mathbf{s}_1=(s_{1x},s_{1y},s_{1z})$):

$$\begin{aligned}&\cos\varepsilon_1=\mathbf{s}_1\mathbf{k}=s_{1z}=\frac{1}{\sqrt{1+A_x^2+(\tan w-A_y)^2}},\\&s_{1x}=-A_x\cos\varepsilon_1,\quad s_{1y}=(\tan w-A_y)\cos\varepsilon_1\end{aligned} \tag{6}$$

where $A_x=a_x/b_z$ and $A_y=a_y/b_z$ are the numerical apertures for the axial point B_0.

Substituting the terms in Eq. (4) with the relations described in (5) and (6), one gets the following formulas (after a long calculation) for the transverse ray aberration of the 5th order:

$$\begin{aligned}\delta x'&=-LA_x(X_0-X_1\tan w+X_2\tan^2 w-X_3\tan^3 w+X_4\tan^4 w)\\\delta y'&=-L(Y_0-Y_1\tan w+Y_2\tan^2 w-Y_3\tan^3 w+Y_4\tan^4 w-Y_5\tan^5 w)\end{aligned} \tag{7}$$

where

$$\begin{aligned}&L=\frac{3d(n^2-1)^2}{8n^5},\quad N=\frac{4n^2}{3(n^2-1)},\\&X_0=(A_x^2+A_y^2)^2-N(A_x^2+A_y^2),\\&X_1=4A_x^2A_y+4A_y^3-2A_yN,\\&X_2=2A_x^2+6A_y^2-N,\\&X_3=4A_y,\quad X_4=1,\\&Y_0=A_y(A_x^2+A_y^2)(A_x^2+A_y^2-N),\\&Y_1=A_x^4+5A_y^4+6A_x^2A_y^2-A_x^2N-3A_y^2N,\\&Y_2=6A_x^2A_y+10A_y^3-3A_yN,\\&Y_3=2A_x^2+10A_y^2-N,\\&Y_4=5A_y,\quad Y_5=1\end{aligned}$$

Knowing the transverse ray aberrations $\delta x'$a $\delta y'$, the wave aberration W can be calculated from relation [6,9]

$$W=W_Q-W_P=\int_P^Q(\delta x'dA_x+\delta y'dA_y) \tag{8}$$

where the integration is performed on a reference sphere from the point P to the point Q, as is shown in Fig. 2 for the case of an elliptical pupil. At first one integrates from the point P to B and then from B to the point Q.

Using the Eq. (7) in (8), one obtains after integration the following relation for the 5th order wave aberration

$$\begin{aligned}W=&\left[\frac{K}{4}-\frac{L}{6}(A_x^2+A_y^2)\right](A_x^2+A_y^2)^2-\left[K-L(A_x^2+A_y^2)\right]A_y(A_x^2+A_y^2)\tan w\\&+\frac{1}{2}\left[K(A_x^2+3A_y^2)-L(A_x^2+A_y^2)(A_x^2+5A_y^2)\right]\tan^2 w\\&-A_y\left[K-\frac{2}{3}L(3A_x^2+5A_y^2)\right]\tan^3 w\\&-\frac{1}{2}L(A_x^2+5A_y^2)\tan^4 w+LA_y\tan^5 w\\=&\,W_3+W_5\end{aligned} \tag{9}$$

where

$$K=d(n^2-1)/2n^3,\quad L=3d(n^2-1)^2/8n^5$$

$$W_3=K\left[\frac{1}{4}(A_x^2+A_y^2)^2-(A_x^2+A_y^2)A_y\tan w+\frac{1}{2}(A_x^2+3A_y^2)\tan^2 w-A_y\tan^3 w\right]$$

$$W_5=L\left[\begin{array}{l}\frac{1}{6}(A_x^2+A_y^2)^3+(A_x^2+A_y^2)^2A_y\tan w-\frac{1}{2}(A_x^2+A_y^2)(A_x^2+5A_y^2)\tan^2 w\\+\frac{2}{3}(3A_x^2+5A_y^2)A_y\tan^3 w-\frac{1}{2}(A_x^2+5A_y^2)\tan^4 w+A_y\tan^5 w\end{array}\right]$$

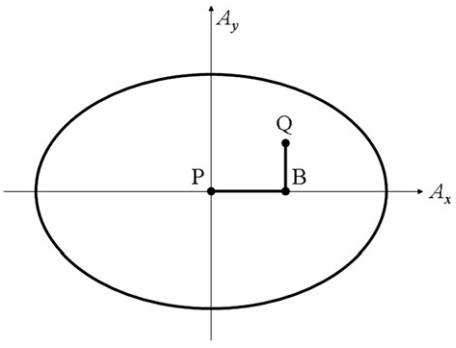


Fig. 2. Integration path for the calculation of wave aberration W.

Neglecting in relation (9) the term W_5 which stands for small quantities, one obtains the relation valid for the 3rd order wave aberration.

3. Calculation of the circle of confusion

To calculate the center of the spot diagram the substitution $A_x=A\cos\varphi$, $A_y=A\sin\varphi$ can be used, where A is a numerical aperture of the incident beam on the plane-parallel plate, φ is the polar angle in the plane ξ_1.

For the coordinates of the center of the spot diagram (energetic center of the beam of rays spot) one obtains for the 5th order aberration

$$\langle\delta x'\rangle=0,\langle\delta y'\rangle=\frac{1}{\pi A_M^2}\int_0^{2\pi}\int_0^{A_M}\delta y' A dA d\varphi = L[(A_M^4-A_M^2N)\tan w+(3A_M^2-N)\tan^3 w+\tan^5 w] \quad (10)$$

where A_M is the maximal numerical aperture. Diameter d_c of the circle of confusion (RMS spot size) is determined by the following formula:

$$d_c=2\rho_c=2\sqrt{\langle\rho_c^2\rangle} \quad (11)$$

where

$$\langle\rho_c^2\rangle=\frac{1}{\pi A_M^2}\int_0^{2\pi}\int_0^{A_M}[(\delta x'-\langle\delta x'\rangle)^2+(\delta y'-\langle\delta y'\rangle)^2]AdAd\varphi \quad (12)$$

Using the relations (7), (10), (11) and (12) to calculate the circle of confusion's diameter d_c in the 5th aberration order one gets

$$d_c=\frac{2}{3}LA_M\sqrt{\sum_{i=0,2,4,6,8}D_i\tan^i w}, \quad (13)$$

where

$$D_0=\frac{3}{2}A_M^8-\frac{18}{5}A_M^6N+\frac{9}{4}A_M^4N^2$$
$$D_2=36A_M^6-54A_M^4N+18A_M^2N^2$$
$$D_4=171A_M^4-144A_M^2N+\frac{45}{2}N^2$$
$$D_6=201A_M^2-72N$$
$$D_8=117/2$$

The diameter of the circle of confusion indicates the size of the "spot" in which most of the energy of the beam of rays spread caused by the aberrations is concentrated. The point therefore does not appear as a point, but as a spot. As a result of this phenomenon the resolution of the optical system will be reduced and the number of lines R per unit of length is bound to the circle of confusion's diameter by the approximate relation $R\approx 1/d_c$. In the case of a physically perfect optical system the central part of the diffraction pattern is called Airy disk, whose diameter is in the case of a circular and uniformly illuminated pupil given by $d_A=1.22\lambda/A_M$, where λ is the wavelength of the light. If the diameter of the circle of confusion is less then the diameter of the Airy disk, i.e. $d_c<d_A$, then the aberrations of the optical system do not affect the image quality and the optical system can be practically considered as an optical system with a good image quality.

For a plane-parallel plate thickness d, which does not cause visible degradation of the image quality, one gets the expression

$$d<\frac{4.9\lambda n^5}{A_M^2(n^2-1)^2\sqrt{\sum_{i=0,2,4,6,8}D_i\tan^i w}} \quad (14)$$

In the next step the position of the "image point", where the circle of confusion has the smallest diameter will be determined. Through shifting the image plane by a value of D with respect to the paraxial image plane of the optical system the ray aberrations changes in the following way:

$$\delta x'_D=\delta x'-DA\cos\varphi,\delta y'_D=\delta y'-DA\sin\varphi \quad (15)$$

Using relations (7), (10), (11), (12) and (15) one obtains for the diameter of the circle of confusion, in a plane shifted by the value of D with respect to the paraxial image plane, the following relation:

$$(d_c)_D=2A_M\sqrt{\sum_{i=0,2,4,6,8}d_i\tan^i w} \quad (16)$$

where

$$d_0=\frac{1}{6}A_M^8L^2-\frac{2}{5}A_M^6L^2N+\frac{1}{2}A_M^4LD+\frac{1}{4}A_M^4L^2N^2-\frac{2}{3}A_M^2LND+\frac{1}{2}D^2$$
$$d_2=4A_M^6L^2-6A_M^4L^2N+2A_M^2L^2N^2+4A_M^2LD-2DLN$$
$$d_4=19A_M^4L^2-16A_M^2L^2N+\frac{5}{2}L^2N^2+3LD$$
$$d_6=\frac{67}{3}A_M^2L^2-8L^2N$$
$$d_8=\frac{13}{2}L^2$$

From the necessary condition for the extreme $(\partial(d_c)_D/\partial D=0)$ one obtains for the optimal position of the image point the following relation:

$$D_{opt}=L\left[A_M^2\left(\frac{2}{3}N-\frac{1}{2}A_M^2\right)+2(N-2A_M^2)\tan^2 w-3\tan^4 w\right] \quad (17)$$

From relation (17) it is obvious that the optimal image point lies on the rotational surface (with respect to the optical axis of the system—the axis z) whose z coordinate, with respect to the paraxial image plane, is equal to D_{opt} (optimal image surface). Substituting relation (17) into relation (16) one obtains for the diameter of the circle of confusion in the optimal image point the following relation:

$$(d_c)_{\min}=2A_ML\sqrt{\sum_{i=0,2,4,6,8}d_{im}\tan^i w} \quad (18)$$

where

$$d_{0m}=\frac{1}{24}A_M^8-\frac{1}{15}A_M^6N+\frac{1}{36}A_M^4N^2$$
$$d_{2m}=2A_M^6-\frac{7}{3}A_M^4N+\frac{2}{3}A_M^2N^2$$
$$d_{4m}=\frac{19}{2}A_M^4-6A_M^2N+\frac{1}{2}N^2$$
$$d_{6m}=\frac{31}{3}A_M^2-2N$$
$$d_{8m}=2 \quad (19)$$

For the ratio of the circles of confusion in the optimal image point $(d_c)_{\min}$ and the paraxial image plane d_c in the 5th order aberration one obtains from the relations (13) and (18) the following relation assuming a point lying on the optical axis of the system $(\tan w=0)$

$$\frac{(d_c)_{\min}}{d_c}=\frac{1}{3}\sqrt{\frac{3A_M^4/2-12A_M^2N/5+N^2}{2A_M^4/3-8A_M^2N/5+N^2}}\approx\frac{1}{3} \quad (20)$$

This relation shows that the diameter of the circle of confusion in the optimal image plane is approximately three times smaller than in the paraxial plane. Assuming, for example, an aperture as large as $A_M=0.8$ and a refractive index of the plate $n=1.5$ ($N=2.4$ and $N^2=5.8$), then from the Eq. (20) one obtains: $(d_c)_{\min}/d_c=0.35$ showing that $1/3=0.33$ is quite good estimation.

In the case of 3rd order aberration (Seidel aberrations), one obtains from the previous relations, that

$$K = d(n^2-1)/2n^3$$
$$\delta x'_{III} = K[A^3 \cos\phi - A^2 \sin 2\varphi \tan w + A\cos\varphi \tan^2 w]$$
$$\delta y'_{III} = K[A^3 \sin\varphi - A^2(1+2\sin^2\varphi)\tan w + 3A\sin\varphi \tan^2 w - \tan^3 w] \quad (21)$$

$$W = K\left[\frac{1}{4}A^4 - A^3 \sin\varphi \tan w + \frac{1}{2}A^2(1+2\sin^2\phi)\tan^2 w - A\sin\varphi \tan^3 w\right] \quad (22)$$

$$d_c = KA_M\sqrt{A_M^4 + 8A_M^2 \tan^2 w + 10\tan^4 w} \quad (23)$$

$$(d_c)_D = A_M\sqrt{K^2(A_M^4 + 8A_M^2 \tan^2 w + 10\tan^4 w) + 2D^2 - 8DK(A_M^2/3 + \tan^2 w)} \quad (24)$$

$$(d_c)_{\min} = \frac{KA_M}{3}\sqrt{A_M^4 + 24A_M^2 \tan^2 w + 18\tan^4 w} \quad (25)$$

For the ratio of the circles of confusion in the optimal image point $(d_c)_{\min}$ and the paraxial image plane d_c in the 3rd order aberration one obtains from the relations (23) and (25) the following relation assuming a point lying on the optical axis of the system ($\tan w = 0$):

$$\frac{(d_c)_{\min}}{d_c} = \frac{1}{3}$$

This relation shows that the diameter of the circle of confusion in the optimal image plane is three times smaller than in the paraxial plane.

If the diameter of the circle of confusion is less then the diameter of the Airy disk, i.e. $(d_c)_{\min} < d_A$, then the aberrations of the optical system do not affect the image quality so much and the optical system can be practically considered as an optical system with a good image quality. From relations (22) and (25) assuming for example for $\tan w = 0$ (axial point) one obtains for the spherical aberration coefficient of the 3rd order relation: $W_{40} = (3/8F)(d_c)_{\min}$. As is explained in [6], it holds that if the Strehl ratio is greater than 0.8, than $W_{40} < 0.95\lambda$ must hold. Substituting into the previous relation one gets: $(d_c)_{\min} = (8F/3)\,0.95\lambda = 2.53\lambda F \approx d_A$. This simple calculation provides evidence that the claim $(d_c)_{\min} < d_A$ provides a relatively good estimation for practice.

Using relation (18) one gets the following expression for the thickness d of the plane-parallel plate:

$$d < \frac{1.63\lambda n^5}{A_M^2(n^2-1)^2\sqrt{\sum_{i=0,2,4,6,8} d_{im}\tan^i w}} \quad (26)$$

For a point lying on the optical axis ($\tan w = 0$) one obtains from (26) the following relation for the thickness d:

$$d < \frac{7.34\lambda n^3}{A_M^4(n^2-1)} \quad (27)$$

In the case of optical systems (for example, photographic lenses) with large values for the wave aberration (a few wavelengths), the wave aberration does not provide any useful information while the circle of confusion provides precise information on the quality of imaging. Using the diameter d_c of the circle confusion (RMS spot size), one can very roughly determine (estimate) the *MTF* (Modulation Transfer Function) of the optical system for low spatial frequencies and vice versa, as is indicated by the simple relation from [9].

$$MTF(R) \approx 1 - (4F\lambda/\pi)R - (\pi^2/2)R^2 d_c^2 \approx 1 - 1.273F\lambda R - 4.935R^2 d_c^2, \quad (28)$$

where F denotes the f-number of the optical system, λ denotes the wavelength, R denotes the spatial frequency i.e. number of lines/mm for the diameter d_c of the circle of confusion measured in mm. Although being a rough estimate, this information is very useful in practice.

4. Example

In Fig. 3 the behavior of the transverse ray aberration in the case of 3rd and 5th order aberration for $d = 10$ mm, $n = 1.5$, $a_x = 0$, $a_y = 20$, $a_z = 0$, $b_x = 0$, $b_y \in \langle 0, 46.6\rangle$, $b_z = 100$ and $A_M = 0.2$ is shown. As one can see from Fig. 3, the deviations in the case of the 3rd order aberration (Eq. (21)) are lower than 3% and in the case of the 5th order aberration the deviations of our Eq. (7) from the exact results (exact ray tracing—Eq. (4)) are negligible. In this case (Fig. 3), the error is smaller than 0.2%. For the most optical systems (occurring in practice) our Eqs. (7) and (9) are sufficiently accurate for calculation of the transverse ray aberrations and the wave aberration. Eq. (4) allows to perform accurate (exact) calculation of the ray aberrations of the plane-parallel plate in general.

5. Conclusion

The objective of the article was to provide a detailed geometric-optical analysis of a plane-parallel plates's imaging characteristics. A number of novel analytical relations were derived helping to extend the theoretical knowledge in the field of geometrical optics on the new relations in the area of aberration induced by the plane-parallel plate or its optical prism equivalent. In particular, relation (4) allows the exact calculation of the transverse ray aberration induced by the plane-parallel plate. Furthermore, relations (7) allow to calculate the 5th order transverse ray aberration and relations (13), (14), (16), (18), (20) and (26) allow to calculate the diameter of the circle of confusion in both paraxial and optimal image plane and determine the optimal point position and shape of the optimal image surface.

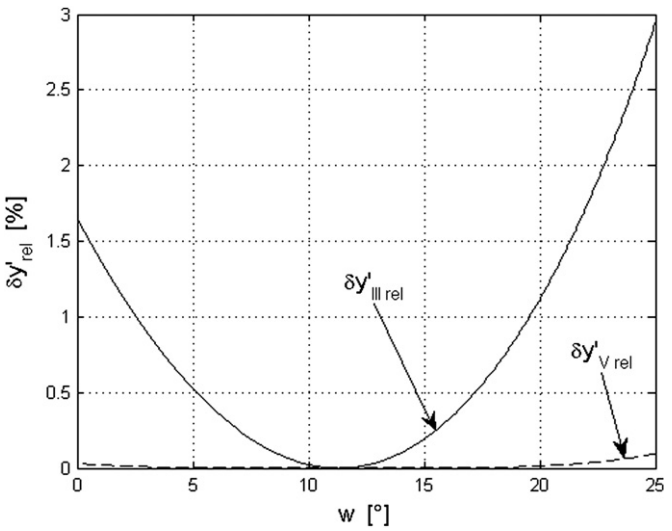


Fig. 3. Dependence of relative deviations of transverse ray aberration $\delta y'_{rel} = 100|(\delta y'_{exact} - \delta y')/\delta y'_{exact}|$ of the 3rd and 5th order aberration on the half-angle of the field of view w for maximal numerical aperture $A_M = 0.2$.

Relation (9) presents the wave aberration in form of the apparent influence of the 3rd and 5th order aberrations on the final value of the wave aberration. In addition, the above stated relations allow to calculate the thickness of the plane-parallel plate without visible deterioration of the image quality, the influence of a prism inserted into the optical system on its aberration, the effect of a cover glass and water layers on the quality of imaging when observing biological structures in confocal microscopy, etc. The relations derived can also be applied to the area of primary optical systems design and to the tolerance analysis of optical systems containing optical plane-parallel plates and prisms. These relations have a simple form and can be easily programmed for example in the MATLAB computing environment, etc.

Acknowledgments

This work has been supported by the Grant P102/10/2377 from the Czech Science Foundation.

References

[1] Herzberger M. Modern geometrical optics. New York: Interscience; 1958.
[2] Braat J. Analytical expressions for the wave-front aberration coefficients of a tilted plane-parallel plate,. Appl Opt 1997;36:8459–67.
[3] Miks A. Applied optics. Prague: Czech Technical University Press; 2009.
[4] Smith W. Modern optical engineering. 3rd ed. New York: McGraw-Hill; 2000.
[5] Eerenbeemd J, Stallinga S. Compact system description for systems comprising a tilted plane parallel plate,. Appl Opt 2007;46:319–26.
[6] Welford WT. Aberrations of the symmetrical optical systems. London: Academic Press; 1974.
[7] Murty MVRK. The use of a single plane parallel plate as a lateral shearing interferometer with a visible gas laser source,. Appl Opt 1964;3:531–4.
[8] Malacara D. Optical shop testing. New York: John Wiley and Sons; 2007.
[9] Sliusarev GG. Aberration and optical design theory. Bristol: Adam Hilger; 1984.
[10] Robb PN. Analytic merit function based on Buchdahl's aberration coefficients,. J Opt Soc Am 1976;66:1037–41.
[11] Buchdahl HA. Optical aberration coefficients. III. The computation of the tertiary coefficients,. J Opt Soc Am 1958;48:747–56.

applied optics

Fundamental characterization of *a priori* measurement accuracy of terrestrial laser scanning

JINDŘICH BRZOBOHATÝ, FILIP ŠMEJKAL, AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 18 September 2020; revised 19 October 2020; accepted 20 October 2020; posted 21 October 2020 (Doc. ID 410571); published 11 November 2020

This paper summarizes, presents, and derives fundamental formulas and principles of characterization of *a priori* measurement accuracy of terrestrial laser scanning. The first part is focused on a general mathematical description of geometrical properties of the scanner, and useful general formulas for uncertainty analysis of all types of scanner constructions are presented. Afterwards, a description of physical aspects affecting terrestrial laser scanning follows. The final part of the paper summarizes the presented analysis on examples.

https://doi.org/10.1364/AO.410571

1. INTRODUCTION

Terrestrial laser scanning has become a very popular technique of quantitative characterization of surface geometry and properties of large areas of interest in a very short time in the last decades [1–6]. Laser scanners are able to measure millions of points in a time scale of fractions of seconds. The fundamental idea is to measure spherical coordinates (horizontal and vertical angles and distance) of points of interest. Accuracy of the acquired point-clouds depends on mechanical construction and uncertainties of the scanner's components, which measure the mentioned angles, and on a quality of distance measurement. There are many commercial companies offering terrestrial laser scanners for practical applications [7–11].

As for the accuracy of scanning outputs (geometric uncertainty of individual points of measured point-clouds), the first component of the *a priori* analysis of the quality is based on a description of uncertainties of geometrical parameters of the scanner. It can be formulated using theoretical foundations presented in previous works of the authors [12] or [13], which will be generalized in this paper. The second component of the accuracy analysis is based on uncertainty estimation of distance measurement. The process of distance measurement during terrestrial laser scanning is the so-called reflector-less electronic distance measurement (EDM) [14], as there are no artificial reflectors placed on the targets. In general, there exist many principles of EDM used in practice. The fundamental and the simplest one is the time-of-flight method (TOF), which accurately measures time of a light-pulse traveling between a source and a target. As the second category, one can name methods based on measuring a phase change—amplitude-modulated continuous wave method or phase-modulated continuous wave method [6]. Those methods offer better accuracy. However, their implementation is limited by the speed of measurement. Therefore, the terrestrial laser scanning uses generally the TOF method as a basic principle of distance measurement, and it will be analyzed in this paper.

In general, the main goal of the paper is to present and derive formulas and principles based on fundamental mathematical and physical foundations, which can be used for characterization of *a priori* accuracy of terrestrial laser scanning. The first part is focused on a mathematical description of geometrical properties of the scanner. Useful general formulas for uncertainty analysis of all types of scanner constructions are presented. Afterwards, physical aspects affecting terrestrial laser scanning are described. Specifically, a reflector-less measurement of distances with the TOF method, reflectivity of natural surfaces, the influence of spot size of laser beam on inclined surface and non-planar geometry of a target object, and the influence of an inhomogeneous refractive index on a measurement are discussed. The final part of the paper summarizes the presented analysis on examples. To the authors' best knowledge, such a general summary of effects that bring uncertainty in terrestrial laser scanning has not been published. There are many studies covering in principle experimentally several separate topics [15–18]. The formulas and principles described in this paper, therefore, offer a useful tool that can be used for *a priori* analysis of many situations of terrestrial laser scanning.

2. GENERAL MATHEMATICAL DESCRIPTION OF ACCURACY OF TERRESTRIAL LASER SCANNING

A scheme of a Cartesian clockwise coordinate system for mathematical description of point location determination measured by a terrestrial laser scanner is presented in Fig. 1. A global coordinate system (X, Y, Z) is determined by its origin G. The system contains a local coordinate system of the scanner (x, y, z), whose starting point O is given by a position vector $\mathbf{X}_O$ in the global system of coordinates, and the axes of the system of the scanner are rotated by angles roll, pitch, and yaw. The system of the scanner further contains a steering unit, whose internal coordinate system (x', y', z') is given by an origin O' determined by a position vector $\mathbf{x}_{O'}$ within the system of the scanner, and its axes are rotated by angles α, β, γ. A measured point R is determined by position vectors $\mathbf{X}_R$ in the global system, $\mathbf{x}_R$ in the scanner system and $\mathbf{r}_R$ in the system of the steering unit.

Transformation relationships between the individual systems can be characterized by the following equations, which holds that

$$\mathbf{X}_R = \mathbf{X}_O + \mathbf{S}\mathbf{x}_R, \quad \mathbf{x}_R = \mathbf{x}_{O'} + \mathbf{R}\mathbf{r}_R, \tag{1}$$

where

$$\mathbf{X}_R = [X_R, Y_R, Z_R]^T, \quad \mathbf{X}_O = [X_O, Y_O, Z_O]^T,$$
$$\mathbf{x}_R = [x_R, y_R, z_R]^T, \quad \mathbf{x}_{O'} = [x_{O'}, y_{O'}, z_{O'}]^T,$$
$$\mathbf{r}_R = [x'_R, y'_R, z'_R]^T,$$
$$\mathbf{S} = \mathbf{R}_z(\text{Yaw})\mathbf{R}_y(\text{Pitch})\mathbf{R}_x(\text{Roll}),$$
$$\mathbf{R} = \mathbf{R}_z(\gamma)\mathbf{R}_y(\beta)\mathbf{R}_x(\alpha),$$

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

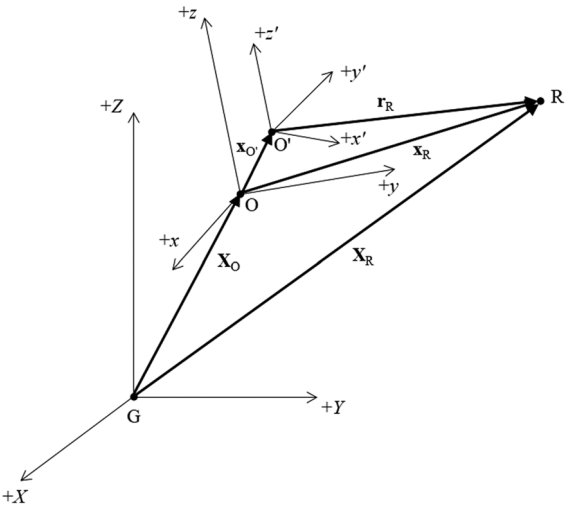


Fig. 1. A scheme of coordinate systems of a terrestrial laser scanner.

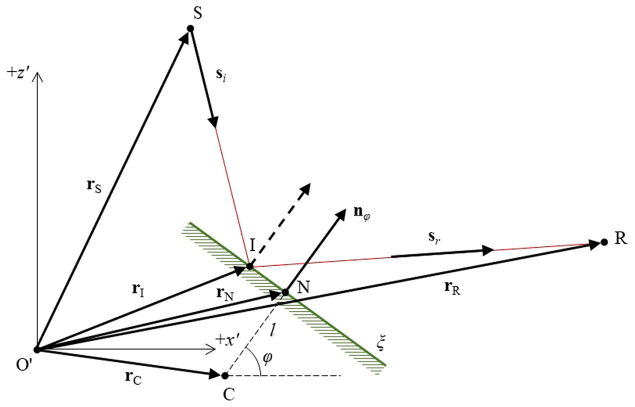


Fig. 2. A scheme of steering device of a terrestrial laser scanner.

In terrestrial systems, the steering unit is generally made by one rotating reflective surface (a mirror, a monogon, an optical, polygonal, or pyramid prism), which steers the ray in vertical direction (relative to the body of the scanner). Further, the whole unit of the scanner is rotated to achieve steering in the horizontal direction. There are other methods of realization of steering units (two-mirror [1,12], etc.), but they will not be considered further in this paper. We consider the case of one reflective surface of the steering system, which realizes direction of a line of sight in a vertical plane (relative to the body of the scanner). A scheme for the mathematical description is presented in Fig. 2.

Suppose that a center of rotation C of the reflective surface ξ is given by a position vector, $\mathbf{r}_C$, and let the reflective surface further rotate around a line parallel to the axis y' of the coordinate system of the unit by an angle φ. A distance between the reflective surface ξ and the point C is denoted, l. An equation of the plane of the reflective surface can be characterized by a location of a point N given by a position vector $\mathbf{r}_N$ and by a unit normal vector $\mathbf{n}$. Let us further assume that the laser ray landing on the reflective surface is given by a point S determined by a position vector $\mathbf{r}_S$ and a unit direction vector $\mathbf{s}_i$. The ray lands on the reflective surface in a point I determined by a position vector $\mathbf{r}_I$, and it is reflected according to the law of reflection and continues in a direction given by a unit direction vector $\mathbf{s}_r$.

We can use the law of reflection to calculate the vector $\mathbf{s}_r$. It holds in the vector form [19,20] that

$$\mathbf{s}_r = \mathbf{s}_i - 2\mathbf{n}(\mathbf{s}_i \cdot \mathbf{n}), \tag{2}$$

where $\mathbf{s}_i$ denotes a unit vector in a direction of the incident ray, $\mathbf{n}$ is a unit normal vector of interface at the point of intersection in the direction of the incident ray, $\mathbf{s}_r$ is a unit vector in a direction of the reflected ray, and the middle dot denotes a scalar product of vectors. Relation for rotation of a unit vector $\mathbf{n}_0$ around an axis given by a unit direction vector $\mathbf{c}$ by an angle φ, which intersects an origin of an auxiliary coordinate system, is given by the following formula [12]:

$$\mathbf{n}_\varphi = \mathbf{n}_0 \cos\varphi + \mathbf{c}(\mathbf{c} \cdot \mathbf{n}_0)(1 - \cos\varphi) + (\mathbf{c} \times \mathbf{n}_0)\sin\varphi. \tag{3}$$

Then, we can obtain a formula for position of the point R within the steering unit after modifying elementary equations of analytical geometry. It holds that

$$\mathbf{r}_R = r_S + \frac{(\mathbf{r}_C - \mathbf{r}_S + l\mathbf{n}_\varphi)\cdot\mathbf{n}_\varphi}{\mathbf{s}_i\cdot\mathbf{n}_\varphi}\mathbf{s}_i + p_r[\mathbf{s}_i - 2\mathbf{n}_\varphi(\mathbf{s}_i\cdot\mathbf{n}_\varphi)], \quad (4)$$

where p_r is a parameter numerically equal to spatial distance of the point I from the point R. For theoretical purposes (e.g., a beam spot position accuracy analysis, or study of a position of the beam spot for different rotations and configurations of the sweeping unit), we can determine the value of the parameter p_r from a nominal definition of plane location of a target that the beam impinges. If the plane of the target ρ, where the beam impinges and from which it is reflected, is given by a point D, a position vector $\mathbf{r}_D$, and a normal vector $\mathbf{n}_D$, then after, it holds for the parameter p_r that

$$p_r = \frac{\left[\mathbf{r}_D - \mathbf{r}_S - \frac{(\mathbf{r}_C - \mathbf{r}_S + l\mathbf{n}_\varphi)\cdot\mathbf{n}_\varphi}{\mathbf{s}_i\cdot\mathbf{n}_\varphi}\mathbf{s}_i\right]\cdot\mathbf{n}_D}{[\mathbf{s}_i - 2\mathbf{n}_\varphi(\mathbf{s}_i\cdot\mathbf{n}_\varphi)]\cdot\mathbf{n}_D}. \quad (5)$$

In practical applications, a location of a target is most often determined using a measurement of transit time (TOF method). If a speed of light in an environment is $v = c/n$ ($c = 299792458$ m/s [19,20] is the speed of light in vacuum, and n is a refractive index of the environment), then a distance d of a measured point on the target from a reference point (the source of the line of sight) is given by $d = v\frac{\Delta\tau}{2}$, where $\Delta\tau$ denotes a transit time between sending and receiving a signal. In our case, the distance between the source and the reflection point is given by $d = p_i + p_r$. After modification, we can express the parameter p_r using the measurement of transit time $\Delta\tau$ as follows:

$$p_r = v\frac{\Delta\tau}{2} - \frac{(\mathbf{r}_C - \mathbf{r}_S + l\mathbf{n}_\varphi)\cdot\mathbf{n}_\varphi}{\mathbf{s}_i\cdot\mathbf{n}_\varphi}. \quad (6)$$

Using the aforementioned relationships, it is possible to analyze the influence of uncertainties of the scanner's construction components, or scanner's position, on the resulting global position of point R. For these analyses, it is possible to utilize simulation methods in which the individual uncertainties defined by relevant probability distributions are added to modeled nominal values, and consequent *a posteriori* analysis can characterize the uncertainty of point R. Alternatively, in particular situations (case studies), it is possible to analytically express the relationships of the location of point R and estimate the uncertainties through the law of variance propagation [12,21–24].

Let us consider a model given by a vector function $\mathbf{f}$ of n variables x_i. If the function $\mathbf{f}$ has continuous partial derivatives, real errors of variables x_i are small relative to function values that the vector function $\mathbf{f}$ takes on, and they have even probability distribution with zero mean value, then the covariance matrix of function $\mathbf{f}$ will be given by a known relationship [12,21–24],

$$\Sigma_{\mathbf{f}} = \mathbf{J}\Sigma_x\mathbf{J}^T, \quad (7)$$

where Σ_x is a covariance matrix of parameters x_i and $\mathbf{J}$ is a Jacobi matrix of first derivatives of the function $\mathbf{f}$. Variance of individual function values then belongs to a diagonal of the matrix $\Sigma_{\mathbf{f}}$.

If we consider the analysis of spatial location of the point R and we know its covariance matrix Σ_R, then position accuracy of the point can be unambiguously described by an ellipsoid of errors, whose sizes of half-axes are given by square roots of eigenvalues of the covariance matrix, and the directions of half-axes are given by the corresponding eigenvectors.

Table 1. Commercially Available Timing Measurement Units from Companies AMS [26] and Texas Instruments[a,b] [27]

Company	Sensor	Accuracy [ps]	Std (Type) [ps]	Range
AMS	TDC-GP30	11	1,2 (RMS)	10000 ps–4,1 ms
TI	THS788	13	8 (accuracy)	0 ps–7 s
AMS	TDC-GP22	22/90	35/45 (std)	700 ps–4 ms/3,5 ps–2500 ps
TI	TDC7201	55	35 (std)	8 ps–250 ps/12 ps–2000 ms

[a]Texas Instruments is abbreviated as TI.
[b]TDC stands for time to digital converter.

3. PHYSICAL ASPECTS AFFECTING ACCURACY OF TERRESTRIAL LASER SCANNING

A. Reflector-less Measurement of Distance with the Method Time of Flight

The method of TOF [1,2,6,14] is based on a very precise duration measurement of time of electromagnetic wave pulse propagation through an environment between a source and a target of the pulse. The pulse is generated by a pulse diode. Afterwards, it is divided in a device (e.g., a semi-transparent mirror), and one of its parts switches on a timing measurement unit. Another part of the pulse is sent into the environment, and after reflection from the target, it returns to the device, where it switches off the timing measurement unit. The timing measurement unit is an electrical device that counts the sum of periods generated by an oscillator. The count of periods is converted by techniques of signal distribution to a transit time of a wave [1,2,6,14,25]. Then, for a measured distance D, it holds that

$$2D = v\tau, \quad (8)$$

where $v = c/n$ is a speed of pulse propagation through an environment, c is the speed of light in vacuum, n is a refractive index of the environment, and τ is a measured transit time.

If millimeter accuracy of the measured distance is required, the transit time τ needs to be measured with picosecond accuracy. An approximate estimate can be obtained by a trivial reasoning—differentiation of the relationship [Eq. (8)], in which we assume $n = 1$ for sake of simplicity (the approximate value for air). After substitution, we obtain a first-order approximation of the required accuracy of measured transit time $\delta\tau = 2\delta D/c$. Substituting for the speed of light c and an assumed distance measurement error $\delta D = 1$ mm, we obtain $\delta t = 6.67$ ps.

Table 1 presents an overview of parameters of currently available timing measurement units (where TDC stands for time to digital converter) offered by companies AMS [26] and Texas Instruments [27].

If we denote the transit time τ and a registered value of amplitude or a count of pulses as $I(\tau)$, then we can determine an estimate of measured transit time $\bar{\tau}$, which we consider a target

value during distance measurement, for example, as a center of gravity (first moment) according to the following formula [1,21]:

$$\bar{\tau} = \frac{1}{P}\int_{-\infty}^{+\infty} \tau I(\tau)\mathrm{d}\tau, \quad P = \int_{-\infty}^{+\infty} I(\tau)\mathrm{d}\tau, \tag{9}$$

where the integration is calculated numerically on the examined range of transit times. An uncertainty σ_τ can then be determined as a square root of variance σ_τ^2, for which it holds that

$$\sigma_\tau^2 = \left(\frac{1}{P}\int_{-\infty}^{+\infty} \tau^2 I(\tau)\mathrm{d}\tau\right) - \bar{\tau}^2. \tag{10}$$

A distance $D_{\bar{\tau}}$ and its uncertainty σ_D corresponding to a calculated transit time $\bar{\tau}$ and its uncertainty σ_τ is then given as follows:

$$D_{\bar{\tau}} = \frac{v}{2}\bar{\tau}, \quad \sigma_D = \frac{v}{2}\sigma_\tau, \tag{11}$$

where v denotes a pulse speed through the environment.

Let us further assume that the switching of the timing measurement unit itself is influenced by a random error. Without further knowledge of the unit, we can assume the switching error to be uniformly distributed with an interval width of $\Delta\tau$. Then, an uncertainty $u_{\Delta\tau}$ can be estimated as a standard deviation of the given uniform distribution, and the resulting uncertainty u_τ of measured transit time can be estimated as a cubic sum of the above-mentioned uncertainties; therefore, it holds that

$$u_\tau = \sqrt{\sigma_\tau^2 + u_{\Delta\tau}^2}, \quad u_{\Delta\tau} = \frac{\Delta\tau}{\sqrt{12}}. \tag{12}$$

For an uncertainty u_D of the measured distance D, we then obtain the estimate,

$$u_D = \frac{v}{2}\sqrt{\sigma_\tau^2 + u_{\Delta\tau}^2}. \tag{13}$$

The aforementioned uncertainty estimations of the transit time or the distance can be further implemented into mathematical tools presented in Section 2 as components of the covariance matrix used in the law of covariance propagation [Eq. (7)], and one can then estimate the overall effect of uncertainties on measurement results.

B. Reflectivity of Natural Surfaces

Reflectivity is an optical property of surfaces that is expressed by a ratio of reflected intensity to the amount of intensity of incident radiation [20]. A coefficient of reflectivity R, which quantitatively characterizes this property, depends on the material (chemical composition, structure, temperature, roughness, color) on which the reflection occurs, and on the type of incident radiation (wavelength, light polarization, etc.).

Let us denote $I_0(\lambda)$ an intensity of incident radiation on a reflective surface, $I_r(\lambda)$ as an intensity of a reflected radiation, and λ the wavelength of light; then, a spectral coefficient of reflectivity R_λ for one wavelength will be given by the following formula:

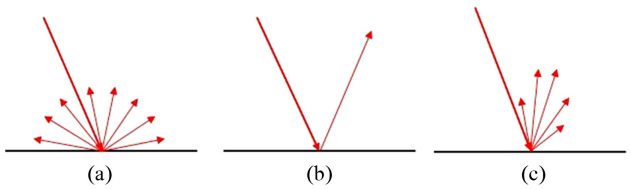


Fig. 3. A scheme of electromagnetic radiation propagation on (a) Lambertian surface, (b) mirror surface, and (c) glossy surface [1,2,6].

$$R_\lambda = R(\lambda) = I_r(\lambda)/I_0(\lambda). \tag{14}$$

A coefficient of reflectivity R for given widths of wavelengths λ_1 and λ_2 will then be given by

$$R = \int_{\lambda_1}^{\lambda_2} I_r(\lambda)\mathrm{d}\lambda \Big/ \int_{\lambda_1}^{\lambda_2} I_0(\lambda)\mathrm{d}\lambda. \tag{15}$$

Reflectivity of a surface is a very important property for measurements by contact-less methods. If properties of the surface are such that the radiation reflects in precise direction as the radiation impinges the surface, then the detector can register the reflected intensity. In the opposite case, this is not possible, and contact-less methods fail. Reflectivity can be divided into three basic categories based on type of surface on which the reflection occurs [6,20]:

- Lambertian surface (ideally matte, ideally diffuse surface) reflects light energy to all directions [19,20,28]. Such surface does not exist in practical cases, but walls of brick buildings (plasters) or school blackboards have properties that resemble it.
- Mirror surfaces that have the feature of perfect reflection [28] governed by the law of reflection.
- Special category of mirror surfaces is glossy surfaces. The ray shatters, but a dominant direction of the following propagation is the direction of perfect reflection. An illustrative example is presented in Fig. 3. These properties of materials negatively influence distance measurement by contact-less methods. In most situations, it is, therefore, not appropriate to measure in wet conditions, namely when the reflective surfaces are inclined compared to the line of view direction.

Similarly to reflectivity, registered values of measurement are also influenced by transmissivity of atmosphere, which expresses a fraction of intensity after it has traveled a certain distance to intensity in the beginning of the distance. Transmissivity is quantitatively characterized, correspondingly to reflectivity, by a coefficient of transmissivity T, which also depends on atmospheric conditions and on properties of the radiation passing through. The coefficient of transmissivity T can never be determined precisely (atmospheric conditions are changing in time and space); hence, idealized models need to suffice.

Reflectivity of materials and transmissivity of atmosphere have high influence on contact-less measurements. For a diffusely reflective target, such as a building or a rock, it is possible to idealize the reflected radiation as hemispherically dispersed (see Fig. 3). Intensity decreases rapidly in each direction. Power of the radiation that is reflected from the target and returns

Table 2. Reflectivity of Selected Materials[a]

Material	Reflectivity (%)
Dimension lumber (pine, clean, dry)	94
Snow	80–90
White masonry	85
Limestone, clay	≤75
Deciduous trees	≈60
Coniferous trees	≈30
Carbonate sand (dry)	57
Carbonate sand (wet)	41
Beach sands, bare areas in dessert	≈50
Rough wood pallet (clean)	25
Concrete, smooth	24
Asphalt with pebbles	17
Lava	8
Black neoprene (synthetic rubber)	5

[a]From [8].

to the detector will be, in fact, only a fraction of the originally transmitted signal. If we denote the power of the output signal P_v and P_p as the power of the received signal, then according to [4], assuming reflection on a Lambertian surface, it holds that

$$P_p = R\frac{AT^2}{2\pi D^2}P_v, \tag{16}$$

where R is the coefficient of reflectivity, A is the size surface of a detector, T is the coefficient of transmissivity in the atmosphere, and D is the distance. Therefore, during the distance measurement, the detector needs to sufficiently amplify the received signal such that it is possible to correctly evaluate the measurement.

Table 2 depicts reflectivity of selected materials in a way that is presented on a website of the company Riegl [8]. In other literature, e.g., [4,5], the authors habitually refer to those values.

C. Gaussian Beams

Gaussian beams are of great importance in laser technology because they describe properties of radiation generated by lasers [1,19,20,28,29]. Let us consider transverse unlimited Gaussian beams. These are beams whose intensity in the transverse direction decreases according to a function (if we denote the direction of wave propagation z), $\exp[-(x^2+y^2)/a]$, where a is a constant. It is known that for complex amplitude of a transverse unlimited Gaussian beam in a point having coordinates $[x, y, z]$ in distance z from the beam waist (the narrowest part of the beam), it holds that [1,19,20,28,29]

$$U(x,y,z) = \frac{w_0}{w(z)}\exp\left[-\frac{x^2+y^2}{w^2(z)}\right] \times \exp\left[i\left(\varphi(z) - k\frac{x^2+y^2}{2R(z)} - kz\right)\right], \tag{17}$$

where

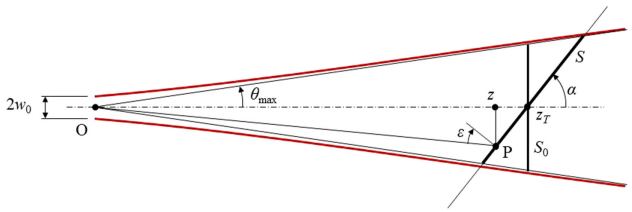


Fig. 4. Intersection of a Gaussian beam spot with an inclined plane.

$$w(z) = \frac{w_0}{z_0}\sqrt{z_0^2+z^2}, \quad z_0 = k\frac{w_0^2}{2}, \quad \varphi(z) = \arctan\frac{z}{z_0},$$
$$R(z) = z\left(1+\frac{z_0^2}{z^2}\right),$$

$k = 2\pi/\lambda$ is the wavenumber, λ is the wavelength, and w_0 denotes the beam waist. Field intensity in the point $[x, y, z]$ can then be described by following formula [19,20,29]:

$$I(x,y,z) = K|U(x,y,z)|^2, \tag{18}$$

where K is a constant. Substituting Eq. (17) into Eq. (18), we obtain for the intensity of transverse unlimited Gaussian beam in the point $[x, y, z]$,

$$I(x,y,z) = K\left[\frac{w_0}{w(z)}\right]^2\exp\left[-\frac{2(x^2+y^2)}{w^2(z)}\right]. \tag{19}$$

Analysis of Gaussian beam asymptotes can show that for the divergence angle θ, it holds that [19,20,29]

$$\tan\theta = \lambda/(\pi w_0), \quad w_0\theta \approx \lambda/\pi, \quad w_0\theta = M^2\lambda/\pi, \tag{20}$$

where the second formula is an approximation for small angles, and the third formula implements the so-called M^2 factor [30], which is in practice used to characterize a quality of laser beam focusing. Generally, it holds that $M^2 \geq 1$ for lasers, and for Gaussian beams, one supposes $M^2 = 1$.

D. Influence of Spot Size of Laser Beam on Inclined Surface on Transit Time Measurement

Let one study a case of a Gaussian beam impinging an inclined plane. Consider the situation depicted in Fig. 4. A transverse unlimited Gaussian beam having a radius of the beam waist w_0 propagates along the axis z, and it impinges a plane inclined by an angle α, while the plane intersects the propagation axis at a distance z_T. The intensity image at this plane will then cover a surface S, which will be elliptical, and a distribution of intensity can be determined using Eq. (18). Let us further assume that there is a point P at the main axis of the elliptical trace having coordinates $[0, y, z]$. Considering geometry in Fig. 4, we can write the following formula for coordinates of the intersection:

$$y = z_s\sin\alpha, \quad z = z_T + z_s\cos\alpha, \quad z_s = z_T\frac{\sin\theta}{\sin(\theta+\alpha)}. \tag{21}$$

If we neglect small curvature of the beam element that goes through the point, then we can write for a transit time τ of a

duration of pulse traveling from a middle point of the beam waist O to a point P,

$$\tau = \frac{2\sqrt{y^2 + z^2}}{v}, \tag{22}$$

where v denotes the speed of pulse propagation through the environment. In ideal case in which all points on the beam spot on the slope become elementary sources of radiation, which returns to the measuring device, we can express a transit time $\tau_{\max}$ for a maximal angle $\theta_{\max}$ after substituting Eq. (21) in Eq. (22) as follows:

$$\tau_{\max} = \frac{2z_T}{v}\sqrt{\left[1 + \frac{\cos\alpha\sin\alpha}{\sin(\alpha + \theta_{\max})}\right]^2 + \frac{\sin^4\alpha}{\sin^2(\alpha + \theta_{\max})}}. \tag{23}$$

For small angles of dispersion, considering the first approximation, and after further rearrangement, we obtain

$$\tau_{\max} \doteq \frac{2z_T}{v(\sin\alpha + \theta_{\max}\cos\alpha)}\sqrt{\sin^2\alpha + 2\theta_{\max}\sin 2\alpha}. \tag{24}$$

A difference $\Delta\tau = \tau_{\max} - \tau_{\min}$, where $\tau_{\min}$ is obtained using the same approach but with $\theta_{\min} = -\theta_{\max}$, it is then possible to write

$$\Delta\tau = \frac{2z_T}{v}\left(\frac{\sqrt{\sin^2\alpha + 2\theta_{\max}\sin 2\alpha}}{\sin\alpha + \theta_{\max}\cos\alpha} - \frac{\sqrt{\sin^2\alpha - 2\theta_{\max}\sin 2\alpha}}{\sin\alpha - \theta_{\max}\cos\alpha}\right). \tag{25}$$

Through rearrangement and simplification for small angles, it holds that

$$\Delta\tau \doteq \frac{2z_T}{v}\left(\sqrt{\sin^2\alpha + \theta_{\max}\sin 2\alpha} - \sqrt{\sin^2\alpha - \theta_{\max}\sin 2\alpha}\right). \tag{26}$$

The value $\Delta\tau$ defines the minimal resolution with which the pulse counter must work depending on the slope of the surface, since when the distance measurement by the TOF will not be influenced. Respectively, if the sampling of the timing measurement unit will be less than at least half of $\Delta\tau$, then the width of the beam impinging the slope cannot influence distance measurement results of the TOF method. In case that the time measurement unit is able to sample measured intensities with frequency that sufficiently covers the interval $\Delta\tau$, it is possible to advance to characterization of accuracy of transit time measurement τ.

In practice, especially when a very precise measurement using the method described above is considered, it is possible to suppress the influence of beam divergence by using a so-called field aperture, which prevents registration of intensity from parts of the beam that are remote from its central part. However, utilization of the field aperture results in larger dimensions of the optical system of measurement devices. Therefore, field aperture is not used, e.g., in laser scanning, or in cheaper electronic range-finders for regular applications.

For illustration, let us consider the following parameters: $\lambda = 900$ nm, $M^2 = 1$, $v = 3 \cdot 10^8$ m/s, $\theta_{\max} = 0.0100$ deg, $w_0 = 1.6370$ mm (parameters of the range-finder of laser scanner Riegl VZ-400i [8]), and $z_T = 100$ m. Then, for the value $\alpha = 30$ deg, the corresponding $\Delta\tau \approx 200$ ps, and for $\alpha = 65$ deg, the corresponding $\Delta\tau \approx 100$ ps. Comparing these values to the above-stated resolutions of timing measurement units of transit time, it is apparent that the slope of the reflective surface can realistically influence measured distances in the presently available devices.

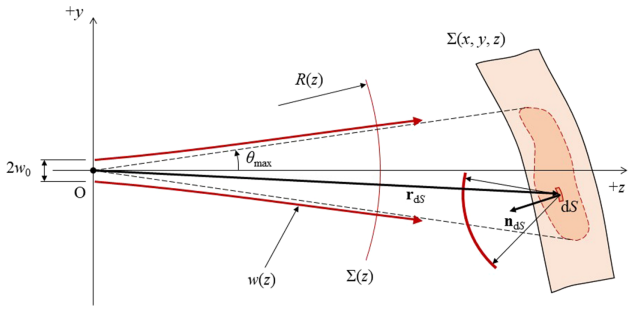


Fig. 5. A reflection of Gaussian beam from a non-planar surface $\Sigma(x, y, z)$.

E. Influence of Non-planar Geometry of a Target Object on a Registered Intensity of a Reflected Pulse

Non-planar geometry is a further source influencing distance measurement by the method of transit time. It has been shown that a spot of Gaussian beam is not a point, but a flat of certain size. Since the spot becomes wider with increasing distance, registered distribution of intensity can be distorted during measuring geometrically non-planar surfaces (e.g., an inside corner between two walls). As in the previous section, it holds that it is possible to suppress the influence of divergence by using a field aperture, but the approach is not often used in practice due to dimension requirements on sensors.

Let us consider the situation depicted in Fig. 5. A transverse unlimited Gaussian beam with a location of the beam waist in the origin of the coordinate system, whose divergence is determined by parameters w_0 and $\theta_{\max}$, propagates through an environment along an axis z and impinges a non-planar topography described by a function $\Sigma(x, y, z)$. A spot of the Gaussian beam is projected to the surface with a shape that depends precisely on geometry of the reflective surface topography. Complex amplitude and intensity within individual points of the trace can be described by the Eqs. (17) and (18). Inside the spot, an elementary flat dS whose location is given by a position vector $\mathbf{r}_{dS}$ and whose slope is given by a unit normal vector $\mathbf{n}_{dS}$ becomes a source of elementary spherical waves (according to Huygens–Fresnel principle [20]), and the waves propagates back to the environment. Let us assume further that the surface from which the radiation is reflected is Lambertian [19,20,28,29], i.e., ideally matte, ideally diffuse, and reflecting light energy evenly in all directions. The detector of the reflected radiation is considered to be in the origin of coordinate system for the sake of simplicity. A registered field in the location of the detector is then given by a superposition of elementary reflected waves (according to the principle of superposition [19,20,29].

For a complex amplitude dU, given by a contribution of the elementary flat dS, in the location of the detector, we can write [19,20,29]

$$dU = -\frac{i}{\lambda} R_{dS} U(\mathbf{r}_{dS}) \frac{\exp(ikr_{dS})}{r_{dS}} \cos(-\mathbf{r}_{dS}, \mathbf{n}_{dS}) dS, \quad (27)$$

where $i = \sqrt{-1}$ denotes the complex unit, λ is the wavelength of the radiation, R_{dS} is the coefficient of reflectivity, $U(\mathbf{r}_{dS})$ is the complex amplitude in the position $\mathbf{r}_{dS}$, $k = 2\pi/\lambda$ is the wavenumber, r_{dS} is vector size, and $\cos(-\mathbf{r}_{dS}, \mathbf{n}_{dS})$ denotes the cosine of an angle between the vector $-\mathbf{r}_{dS}$ and the unit normal vector. $\mathbf{n}_{dS}$.

Let us now denote $S_{\Delta t}$ the sum of elementary flats $S_{\Delta t}$ such that the Gaussian beams impinges them, and reflects from them in a time interval Δt. For a transit time τ registered in the detector, it holds that $\tau = 2t$, and therefore for an interval $\Delta\tau$, it holds that $\Delta\tau = 2\Delta t$. A complex amplitude $U(\Delta\tau)$ registered in the detector in a certain interval of transit time $\Delta\tau$ (assuming a coherence time [19,20,29] of a given device longer than the interval) $\Delta\tau$ is given by a superposition of contributions of elementary flats on which the radiation impinged in the time interval $\Delta t = \Delta\tau/2$. We can, therefore, write

$$U(\Delta\tau) = -\frac{i}{\lambda} \int_{S_{\Delta t}} R_{dS} U(\mathbf{r}_{dS}) \frac{\exp(ikr_{dS})}{r_{dS}} \cos(-\mathbf{r}_{dS}, \mathbf{n}_{dS}) \mathrm{d}S, \quad (28)$$

where integration is performed over the area $S_{\Delta t}$. A registered intensity of radiation in the given interval $\Delta\tau$ will then be given as

$$I(\Delta\tau) = K|U(\Delta\tau)|^2, \quad (29)$$

where K is a constant.

Therefore, we can analyze the influence of object surface distribution on a registered intensity with the following algorithm:

1. Define parameters of a Gaussian beam and a geometry of an object $\Sigma(x, y, z)$ on which we analyze the case.
2. For individual points of the object $\Sigma(x, y, z)$, we calculate the complex amplitude $U(x, y, z)$ when the beam impinges the object using the Eq. (17).
3. For individual points we calculate the contributions dU for the propagation of reflected radiation according to Eq. (27).
4. For individual points of the object $\Sigma(x, y, z)$, we calculate the transit times τ corresponding to the speed of pulse propagation in a given environment. For the sake of simplicity, it is possible to choose rectilinear propagation of beams from the origin of the coordinate system (in which we consider the beam waist) to the given points of the object, neglecting small curvature of the rays of the Gaussian beam.
5. We integrate (sum) contributions of the complex amplitudes dU for selected intervals $\Delta\tau_i$, in which we select the time interval considering the resolution of the timing measurement unit.
6. For individual beginnings of intervals $\Delta\tau_i$, we compute registered intensity according to the relationship [Eq. (29)].

In this way, it is possible to obtain a data set for an analysis of influence of geometric distribution of an object on a registered intensity in dependence on a transit time τ. Using Eqs. (9)–(13), we can consequently determine an estimate of the actual value of measured transit time $\bar{\tau}$, an estimate of the corresponding distance $D_{\bar{\tau}}$, and its uncertainty u_D. These estimations can be used as components of the covariance matrix in the law of covariance propagation presented in Section 2 [see Eq. (7)], and one can estimate the overall effect of uncertainties on measurement results.

F. Influence of an Inhomogeneous Refractive Index on a Measurement of a Point Location

Quality of position determination of a point with a direction and TOF distance measurement of line of sight is influenced also by variation of a refractive index of a surrounding environment. In practical applications, we can consider terrestrial atmosphere to be an isotropic non-homogeneous environment, since the refractive index depends on a number of physical factors—notably temperature, pressure, and humidity [19,20,28,29]. These can be variable within a measurement of one sight line, their variation cannot be usually precisely determined (turbulent atmosphere), and they need to be modeled for simplified situations. The changes in distribution of the refractive index of the environment curve trajectory of a ray by virtue of validity of the Fermat's principle [19,20,28,29], and the timing measurement unit measures a different value of transit time than in a case of pulse propagation through an isotropic homogeneous environment (as a straight line). The direction of the ray that is registered then also does not correspond to reality, and it will be distorted by the curvature of the ray. For very precise measurements, it is, therefore, necessary to introduce so-called atmospheric corrections, which might be complicated for long distances as it is not possible to measure atmospheric parameters in real time. However, approximate equations are often implemented in internal software of measurement devices.

Let us now assume that a function of refractive index within a time interval is dependent on location only, i.e.,

$$n = n(\mathbf{r}) = n(x, y, z), \quad (30)$$

where $\mathbf{r} = [x, y, z]$ denotes a position vector of an examined location. In practice, it is very difficult to fully describe such a functional dependence. Nevertheless, it is possible to obtain a set of discrete data for selected positions and to approximate the model by a suitable expansion into a series of base functions [21], or to use spatial interpolation methods from discrete data sets (e.g., linear or spline interpolation).

It is known [19,31–33] that for one specific place it is possible to express the dependence of refraction index on wavelength λ, pressure p, temperature t, and relative humidity ν using the so-called Edlén equation [33], which is valid in order of 10^{-8} in the range of normal pressures and for wavelengths of visible light. In practice, it turns out that for wavelength $\lambda = 633$ nm (He–Ne laser, the most frequently used one), it is possible to use a simplified formula [19],

$$n = 1 + 7.86 \cdot 10^{-7} \frac{p}{273 + t} - 1.5 \cdot 10^{-11} \nu (t^2 + 160). \quad (31)$$

Let us consider the situation depicted in Fig. 6. Let us assume a reference center of a range-finder in the origin of a coordinate system O, and let a point T be in a geometric distance s_0 from the origin of the coordinate system. Evidently, it holds that

$$s_0 = \sqrt{x_T^2 + y_T^2 + z_T^2}, \quad (32)$$

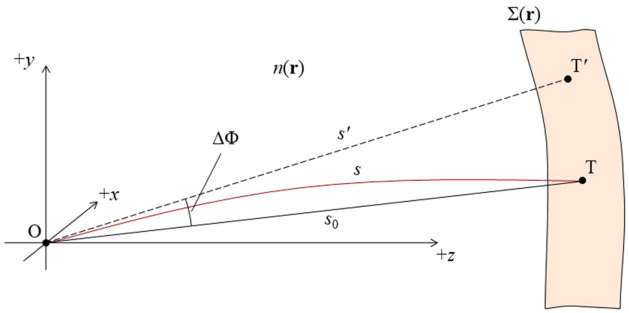


Fig. 6. The influence of inhomogeneous refractive index on a measurement of angles and distances.

where x_T, y_T, and z_T are Cartesian coordinates of point T. Let us also assume that there is an inhomogeneous isotropic environment characterized by a refractive index $n = n(\mathbf{r}) = n(x, y, z)$, a function of coordinates, that is potentially variable in time. In our case, we do not further study the dependence on time, and we only consider the spatial change of the refractive index.

It is known from geometrical optics that an equation of ray is given by the following formula [19,20,28,29]:

$$\nabla n(\mathbf{r}) = \frac{d}{ds}\left[n(\mathbf{r})\frac{d\mathbf{r}}{ds}\right], \tag{33}$$

where ds is a curve arc element in place $\mathbf{r}$ and ∇ is the Hamilton's operator [21]. Since the refractive index is not generally constant, a curve of sight between points O and T is not a straight line, as is apparent from the ray Eq. (33). A line of sight will, therefore, be targeted in a wrong direction (on point T', see Fig. 6, in which the point is deviated from by the angle) $\Delta\Phi$, and also, a wrong distance s corresponding to the length of the ray will be measured instead of the correct distance s_0. An error of the measurement will then be given by the relationship $\Delta s = s - s_0$.

If we know the spatial distribution of the refractive index and the initial conditions (location of the source and direction of the ray, i.e., we solve a Cauchy boundary problem), it is possible to calculate the true trajectory of the ray and the pulse path length through solving Eq. (33), a partial differential equation of second order. An analytical solution of the equation only exists for very specific cases. In practice, it is, therefore, necessary to solve the equation numerically. Methods of solution can be found by the reader in the literature [19,21,34–43].

For a curve arc element in location $\mathbf{r}$, we can generally write

$$ds = v(\mathbf{r})dt = \frac{c}{n(\mathbf{r})}dt, \tag{34}$$

where $v(\mathbf{r})$ is a beam propagation speed in a given location, c is the speed of light in vacuum, and dt stands for a time element. Time T necessary to travel the distance s between the initial and the final point along the curve will be given as

$$T = \frac{1}{c}\int n(\mathbf{r})\mathrm{d}s\,, \tag{35}$$

where we integrate by the curve of the beam. If a discrete set of distribution values of the refraction index and path of the beam is known, the most convenient method of numerical solution is given by the following relationship:

$$T = \frac{1}{c}\sum_{i=0}^{K-1} n(\mathbf{r}_i)\sqrt{\| \mathbf{r}_{i+1} - \mathbf{r}_i \|^2}, \tag{36}$$

where $\mathbf{r}_j$ denotes a vector of jth point of the curve such that the curve has been divided into K sections, the initial vector of the beam is $\mathbf{r}_0$, and the location vector of the beam is $\mathbf{r}_K$.

Through the calculation of a spatial propagation of a beam in inhomogeneous environment and of a transit time required to overcome a path along a curve thus determined, we can analyze in a simple manner the influence of a refractive index distribution on a transit time, and therefore on a measurement of spatial location using the TOF method, which is dependent on this time. Geometric analysis can be conducted afterwards to analyze a directional deviation $\Delta\Phi$ from rectilinear propagation. Afterwards, one can use those deviations to estimate the overall *a priori* uncertainty of the scanning results with the law of variance propagation presented in Section 2 [see Eq. (7)].

4. EXAMPLES

A. Example 1

Let us now consider an analysis of the influence of construction errors of the steering and rotation unit of a scanner, of the error of estimation of the transit time, and of the error in determining tilt of the scanner [see Eq. (1)].

For now, let us omit the influence of a coordinate placement of the scanner in a global system, i.e., let us assume zero uncertainties for the vector $\mathbf{X}_O$ (the influence of uncertainty in placement of the scanner would proportionately increase the uncertainty of the resulting coordinates). Without loss of generality, we can further assume the parameters in Eq. (1) as $\mathbf{X}_O = \mathbf{0}$, $\mathbf{x}_{O'} = \mathbf{0}$, $\mathbf{R} = \mathbf{R}_z(\gamma)$, where γ denotes a rotation angle of a body of the scanner around its vertical axis. Then, for the coordinate of point R in the global system, we can write

$$\mathbf{X}_R = \mathbf{R}_z(\mathrm{Yaw})\mathbf{R}_y(\mathrm{Pitch})\mathbf{R}_x(\mathrm{Roll})\mathbf{R}_z(\gamma)\mathbf{r}_R(\varphi, \tau), \tag{37}$$

where roll, pitch, and yaw determine tilt and orientation of the scanner, φ is an angle of steering unit reflective surface tilt, and $\Delta\tau$ is a transit time. To define geometric parameters of the steering unit and for parameters of the measurement and its uncertainties, let us further assume in this illustrative example that $l = 2$ cm, $\mathbf{r}_S = [0, 0, 5]^T$ cm, $\mathbf{r}_C = [-1, 0, 0]^T$ cm, $\mathbf{s}_i = [0, 0, -1]^T$, $\mathbf{c} = [0, 1, 0]^T$, $\mathbf{n}_0 = [1/\sqrt{2}, 0, 1/\sqrt{2}]^T$, $\Delta\tau = 0.2$ μs, $u(\Delta\tau) = 20$ ps, $u(\varphi) = 0.0001$ deg, $u(\gamma) = 0.005$ deg, Roll = Pitch = Yaw = 0 deg, $u(\mathrm{Roll}) = u(\mathrm{Pitch}) = 0.001$ deg, $u(\mathrm{Yaw}) = 0.01$ deg. Figure 7 depicts location of the points and their corresponding ellipsoids of errors determined through Eqs. (1)–(7) at a scale of 500:1. The Jacobi matrix was constructed in a corresponding way for measures for which we consider uncertainty, and a covariance matrix of input parameters can be in this case constructed as a diagonal matrix, whose diagonal contains squares of uncertainties of these parameters (estimates of covariances); therefore, we assume their mutual independence. Using the above-stated tools, it is possible to analyze numerous specific

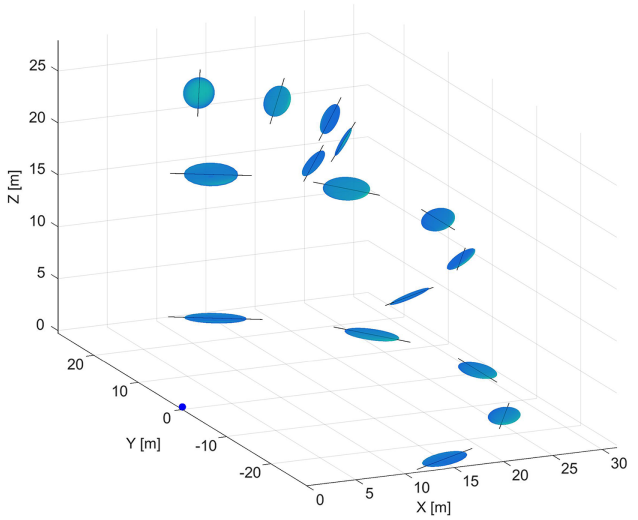


Fig. 7. The location of points in a global system and ellipsoids of errors at a scale of 500:1.

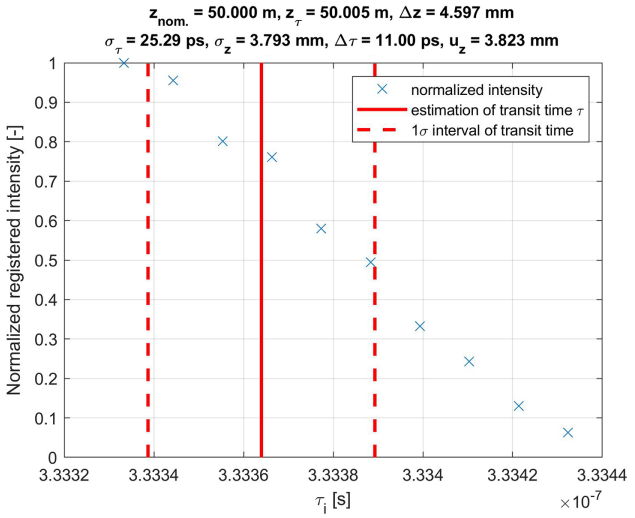


Fig. 8. Normalized registered intensity for a corner target and reconstruction of transit time.

situations and, therefore, evaluate the *a priori* estimates of uncertainty of measurements.

B. Example 2

The introduced algorithm for analysis and reconstruction of measured distance by the method of transit time can be applied to various analyzed geometries of the reflective surface. In this case, we will consider a simple situation in which the reflective surface is an outer corner between two walls. The axis of sight of the range-finder is simultaneously the axis of symmetry of the walls. Nominal distance of the corner is selected as $z_{\text{nom.}} = 50$ m. Figure 8 shows the transit time reconstruction, assuming that the beam propagates in the form of unlimited Gaussian beam and its parameters correspond to a terrestrial laser scanner Riegl VZ-400i [8], in particular, $\theta_{\max} = 0.0100$ deg, $w_0 = 1.6370$ mm, $\lambda = 900$ nm, assuming $M^2 = 1$, $v = 3 \cdot 10^8$ m/s, and resolution of the timing measurement unit $\Delta\tau = 11$ ps. Further, a constant coefficient of reflectivity is assumed for individual points of the reflective surface. In Fig. 8, blue crosses symbolize normalized intensities registered by the unit at times τ_i, the red line shows estimation of the transit time τ, and red dashed lines show 1σ interval of estimated uncertainty of the transit time τ. It is apparent from the results that divergence of rays has crucial influence on the evaluated distance and its uncertainty.

C. Example 3

Let us now illustrate a case of ray tracing under so-called Dirichlet boundary condition, i.e., if we know the positions of start and end point and the goal is to determine the path of a line of sight through the environment between those points, while the refractive index of the environment is a function of position. Without loss of generality, let us consider a plane case in which distribution of temperature is given by the following formula:

$$t(x, y) = 25 + 10 \exp\left[-\frac{(x-50)^2}{50^2}\right] \exp(-0.3y) \quad [^\circ\text{C}]. \tag{38}$$

Let us further consider a wavelength of the radiation $\lambda = 633$ nm, an atmospheric pressure $p = 101325$ Pa, and a relative humidity $v = 50\%$. The dependence of refractive index can then be expressed by Eq. (31). The beam was traced using an optimization method for searching an initial direction that allows the beam to intersect the target point in minimal distance. Step of the calculation was selected as $ds = 0.01$ m. Figure 9 shows the deviation of the calculated beam from rectilinear propagation, when the start point was chosen as the origin of the coordinate system and the end point has coordinates $\mathbf{r}_B = [100, 5]$ m. Further, a deviation from geometric length of the beam s from rectilinear distance s_0 is calculated, i.e., $\Delta s = s - s_0 = 5.6 \cdot 10^{-5}$ mm, and the angle deviation of the initial direction of rectilinear propagation $\Delta\Phi = -0.0035$ deg. It is obvious that the distribution of refractive index of an environment can have significant influence, especially on directions of measured rays.

5. CONCLUSION

The paper summarized a general theoretical description of selected mathematical and physical aspects influencing the accuracy of measurement by terrestrial laser scanners, together with an apparatus that can be easily employed in specific situations and analyses.

The first part was focused on a mathematical description of geometrical properties of the scanner, and useful general formulas for uncertainty analysis of all types of scanner constructions were presented. Afterwards, physical aspects affecting the scanning were described. Specifically, a reflector-less measurement of distances with the TOF method, reflectivity of natural surfaces, the influence of spot size of laser beam on inclined surface and non-planar geometry of a target object, and the influence of

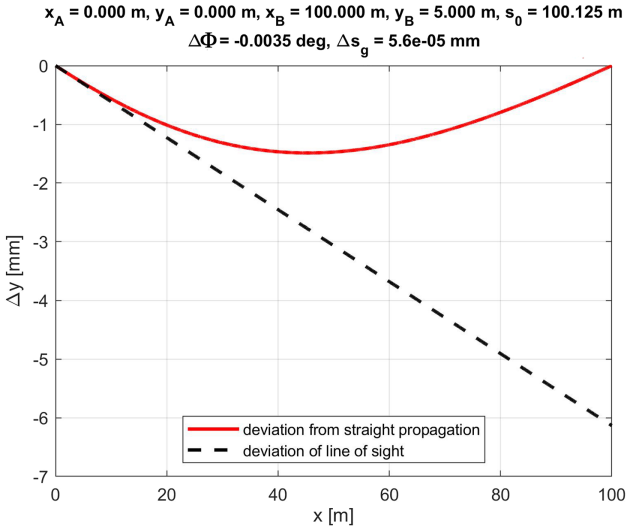


Fig. 9. Deviation of ray propagation from rectilinear case.

an inhomogeneous refractive index on a measurement were discussed.

Compared to the currently published, mainly experimental studies focused on specific measurement devices and situations, the work is a valuable summary addition and completion of the laser scanning area, with which it is possible to predict *a priori* uncertainties of results.

Funding. Czech Technical University in Prague (SGS20/093/OHK1/2T/11).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. G. F. Marshall and G. E. Stutz, *Handbook of Optical and Laser Scanning*, Optical Science and Engineering (CRC Press, 2004).
2. G. Vosselman and H. G. Maas, *Airborne and Terrestrial Laser Scanning* (Whittles Publishing, 2010).
3. G. Marshall, *Laser Beam Scanning: Opto-Mechanical Devices, Systems, and Data Storage Optics*, Optical Science and Engineering (Taylor & Francis, 1985).
4. J. Shan and C. Toth, *Topographic Laser Ranging and Scanning: Principles and Processing* (CRC Press, 2017).
5. A. Wehr and U. Lohr, "Airborne laser scanning—an introduction and overview," ISPRS J. Photogramm. Remote Sens. **54**, 68–82 (1999).
6. T. Schulz, "Calibration of a terrestrial laser scanner for engineering geodesy," Ph.D. dissertation (ETH Zurich, 2008).
7. "Leica geosystems," http://hds.leica-geosystems.com/en/index.htm.
8. "RIEGL Laser Measurement Systems," http://www.riegl.com/.
9. "Surphaser," http://www.surphaser.com/.
10. "FARO," http://www.faro.com/.
11. "Topcon," https://www.topconpositioning.com/mass-data-and-volume-collection/laser-scanners/gls-2000.
12. P. Pokorny, "One-mirror and two-mirror three-dimensional optical scanners—position and accuracy of laser beam spot," Appl. Opt. **53**, 2730–2740 (2014).
13. P. Pokorny and A. Miks, "3D optical two-mirror scanner with focus-tunable lens," Appl. Opt. **54**, 6955–6960 (2015).
14. J. M. Rueger, *Electronic Distance Measurement* (Springer Berlin Heidelberg, 1996).
15. E. Lambrou and G. Pantazis, "Evaluation of the credibility of reflectorless distance measurement," J. Surv. Eng. **136**, 165–171 (2010).
16. E. Lambrou, "Modeling the deviations of the reflectorless distance measurement due to the laser beam's incident angle," Int. J. Appl. Sci. Technol. **8**, 1–13 (2018).
17. K. Kowalczyk and J. Rapinski, "Investigating the error sources in reflectorless EDM," J. Surv. Eng. **140**, 06014002 (2014).
18. M. Zámečnková, A. Wieser, H. Woschitz, and C. Ressl, "Influence of surface reflectivity on reflectorless electronic distance measurement and terrestrial laser scanning," J. Appl. Geod. **8**, 311–326 (2014).
19. A. Mikš, *Applied Optics* (CTU in Prague, 2009).
20. M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University Press, 1999).
21. K. Rektorys, *Survey of Applicable Mathematics*, Mathematics and Its Applications (Springer Netherlands, 2013).
22. K. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models* (Springer, 1999).
23. A. Fridman, A. Sabak, and P. Makinen, *The Quality of Measurements: A Metrological Reference* (Springer New York, 2011).
24. M. Grabe, *Measurement Uncertainties in Science and Technology* (Springer International Publishing, 2016).
25. J. Uren and W. F. Price, *Surveying for Engineers* (Palgrave Macmillan, 2010).
26. "AMS," https://ams.com.
27. "Texas Instruments," http://www.ti.com.
28. U. Teubner and H. Brückner, *Optical Imaging and Photography: Introduction to Science and Technology of Optics, Sensors and Systems*, De Gruyter STEM (De Gruyter, 2019).
29. J. Braat and P. Török, *Imaging Optics* (Cambridge University Press, 2019).
30. ISO Central Secretary: "Lasers and laser-related equipment–test methods for laser beam widths, divergence angles and beam propagation ratios–Part 1: Stigmatic and simple astigmatic beams," Standard ISO 11146-1:2005, International Organization for Standardisation, Geneva, CH (2005).
31. B. Edlén, "The refractive index of air," Metrologia **2**, 71–80 (1966).
32. K. P. Birch and M. J. Downs, "An updated Edlén equation for the refractive index of air," Metrologia **30**, 155–162 (1993).
33. K. P. Birch and M. J. Downs, "Correction to the updated Edlén equation for the refractive index of air," Metrologia **31**, 315–316 (1994).
34. L. Montagnino, "Ray tracing in inhomogeneous media," J. Opt. Soc. Am. **58**, 1667–1668 (1968).
35. E. W. Marchand, "Ray tracing in gradient-index media," J. Opt. Soc. Am. **60**, 1–7 (1970).
36. A. Sharma, D. V. Kumar, and A. K. Ghatak, "Tracing rays through graded-index media: a new method," Appl. Opt. **21**, 984–987 (1982).
37. A. Sharma, "Computing optical path length in gradient-index media: a fast and accurate method," Appl. Opt. **24**, 4367–4370 (1985).
38. J. Puchalski, "Numerical determination of ray tracing: a new method," Appl. Opt. **31**, 6789–6799 (1992).
39. J. Puchalski, "Numerical determination of continuous ray tracing: the four-component method," Appl. Opt. **33**, 1900–1906 (1994).
40. B. D. Stone and G. W. Forbes, "Optimal interpolants for Runge–Kutta ray tracing in inhomogeneous media," J. Opt. Soc. Am. A **7**, 248–254 (1990).
41. B. D. Stone and G. W. Forbes, "Differential ray tracing in inhomogeneous media," J. Opt. Soc. Am. A **14**, 2824–2836 (1997).
42. B. Richerzhagen, "Finite element ray tracing: a new method for ray tracing in gradient-index media," Appl. Opt. **35**, 6186–6189 (1996).
43. Y. Nishidate, T. Nagata, S. Ya Morita, and Y. Yamagata, "Ray-tracing method for isotropic inhomogeneous refractive-index media from arbitrary discrete input," Appl. Opt. **50**, 5192–5199 (2011).

One-mirror and two-mirror three-dimensional optical scanners—position and accuracy of laser beam spot

Petr Pokorny

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thakurova 7, 166 29 Prague 6, Czech Republic (petr.pokorny@fsv.cvut.cz)

Received 6 February 2014; revised 21 March 2014; accepted 21 March 2014; posted 24 March 2014 (Doc. ID 206101); published 18 April 2014

This article presents several fundamental formulas for ray tracing in optical systems used in 3D optical scanners. A procedure for numerical modeling of one-mirror and two-mirror optical systems is presented, and the calculation of positioning and accuracy of the laser beam spot in a detection plane is carried out. Finally, a point position and accuracy depending on a transit time is evaluated. © 2014 Optical Society of America

OCIS codes: (080.0080) Geometric optics; (110.0110) Imaging systems; (120.0120) Instrumentation, measurement, and metrology; (220.0220) Optical design and fabrication; (280.0280) Remote sensing and sensors.

http://dx.doi.org/10.1364/AO.53.002730

1. Introduction

Three-dimensional (3D) optical scanners [1–11] are devices that are able to perform noncontact, very fast, and accurate measurements of 3D physical objects. Areas of application of 3D optical scanners are in civil engineering, architecture, interior modeling, earthwork (building foundations, mining, etc.), components and modules in engineering, automobile body shaping, measuring the condition of road surfaces, railway lines, detecting and documenting cracks, security and traffic control (laser gates), archaeology, and so forth [3]. Furthermore, these systems are widely used in safety technology such as the protection of objects and detection of human presence in safety zones of production systems (e.g., machining centers and foundries). Another area of 3D scanner application is in laser technology (cutting, welding, engraving, surface finishing of materials, etc.), medicine, or the entertainment industry (laser effects, theaters). There are many companies [4–9] that deal with this issue and commercially offer 3D scanners for the above-mentioned applications.

There are several studies of 1D and 2D optical scanners that have been developed and presented in many papers. Fundamental analysis of systems of plane mirrors by matrix multiplication is carried out in [12]. Other works are focused on design of distortion-free two-mirror scanners [13], analysis of various types of scanning patterns in the observation plane [14–17], stability of plane-mirror systems for as many as three mirrors [18], or 1D surface profile measurement by detection of the angles of deflection [19].

In the above-mentioned papers the attention is focused on 1D or 2D problems. In this paper, besides 2D analysis, 3D positioning and accuracy determination is carried out as well. This approach is required in practical situations. First, a procedure for ray tracing through a scanner optical system will be presented, and the formulas for positioning and accuracy of a laser beam spot in a detection plane will be derived. Subsequently, a method for a 3D position and accuracy estimation based on a time-of-flight method is carried out. A vector approach is used

1559-128X/14/122730-11$15.00/0

for its simplicity and versatility and can be applied in many practical cases.

A 3D optical scanner comprises a radiation source, an optical or opto-mechanical directional device for laser beams, a radiation detector, and an evaluation system. Light (beam) coming from the source of radiation is transformed by means of an optical system, deflected by a directional device to an accurately determined direction, and impacts to the measured object. After the reflection from the object, a part of the diffused light returns through the optical system and impacts on the radiation detector. The evaluation system then determines the spatial coordinates of the object point. The distance between the object point and scanner is mostly determined by modulating the light or by measuring the elapsed time between the transmitted and received signal (time of flight method). There are other ways to determine the distance of the object from the scanner, for example, the triangulation method [11], but these are rarely used and therefore they are not discussed.

There are usually one or two rotating mirrors used in the design process. One-mirror scanners are used where a larger angular range (field of view) of the swept beams is needed. This type of scanner is usually formed by a laser module with a single rotating mirror, which rotates around a horizontal axis (H), while the laser beam sweeps in a plane perpendicular to the axis of mirror rotation (vertical plane). The laser module is then rotated around a vertical axis (V) perpendicular to the axis of the mirror rotation, and thereby it sweeps the laser beam in the horizontal position (horizontal) plane. An achievable field of view can be, for example, 360° × 320° (H × V). Two-mirror scanners are used primarily in the field of laser technology in engineering and elsewhere where a large angular field of view is not required. The optical system of this scanner consists of two mirrors that revolve around two different axes and thus lead to sweep the laser beam. There are several companies that provide ready-made modules for two-mirror scanners [11]. An achievable field of view may be, for example, 80° × 80° (H × V).

2. Fundamental Formulas for Laser Beam Tracing through the Scanner Optical System

A. Law of Refraction and Reflection in the Vector Form

To derive formulas for laser beam tracing one can employ the basic equations of classical electrodynamics—Maxwell's equations [20–26]. However, the complete procedure is beyond the scope of this work and can be found in [20,22–26]. Assume that a unit normal vector $\mathbf{N}$ of the interface is directed to the incident ray of light and the angle of incidence and refraction is measured out from the normal, as is shown in Fig. 1, where $\mathbf{A}_I$ and $\mathbf{A}_T$ are unit vectors along the incident and refracted rays. All above-stated vectors lie in a plane, which is

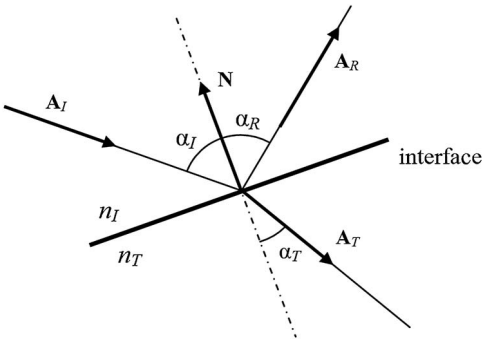


Fig. 1. Refraction and reflection on the plane.

called the plane of incidence, as known from fundamentals of geometrical optics.

Angles of incidence α_I and refraction α_T are connected with the well-known Snell's law of refraction at an interface between media of refractive indices n_I and n_T [20–22,25–28]:

$$n_I \sin \alpha_I = n_T \sin \alpha_T, \tag{1}$$

where subscript I and T denote rays of incidence and refraction.

Consider the interface shape of any surface described by function $F(\mathbf{r}) = 0$, where $\mathbf{r}$ is a position vector of any point of the interface. As known from differential geometry of surfaces, a normal unit vector of the interface is obtained from the formula [20,21,29]

$$\mathbf{N} = \frac{\operatorname{grad} F(\mathbf{r})}{|\operatorname{grad} F(\mathbf{r})|}. \tag{2}$$

With the use of vector calculus, one can obtain the law of refraction in the form [20]

$$\mathbf{A}_T = \frac{n_I}{n_T}\mathbf{A}_I - \frac{1}{n_T}\mathbf{N}\left(\sqrt{n_T^2 - n_I^2 \sin^2 \alpha_I} - n_I \cos \alpha_I\right). \tag{3}$$

A similar situation is in the case of the law of reflection. As can be proved, for example, in [20–22,25–28], the angle of incidence is equal to the angle of reflection. Therefore, the law of reflection can be retained in the vector form as follows [20,26]:

$$\begin{aligned} \mathbf{A}_R &= \mathbf{A}_I + 2\mathbf{N} \cos \alpha_I, \\ \mathbf{A}_R &= \mathbf{A}_I - 2\mathbf{N}(\mathbf{N} \cdot \mathbf{A}_I). \end{aligned} \tag{4}$$

B. Rotation of Vectors around an Axis

Consider unit vector $\mathbf{N}$, which rotates around an axis defined by unit directional vector $\mathbf{C}$. As is well known in mathematics [30], one gets the formula

$$\mathbf{N}(\varphi) = \mathbf{N}\cos\varphi + \mathbf{C}(\mathbf{C}\cdot\mathbf{N})(1-\cos\varphi) + (\mathbf{C}\times\mathbf{N})\sin\varphi, \tag{5}$$

where φ denotes rotation angle. For differentially small rotations $d\varphi$ ($\cos d\varphi \approx 1, \sin d\varphi \approx d\varphi$) one gets

$$\mathbf{N}(d\varphi) = \mathbf{N} + (\mathbf{C}\times\mathbf{N})d\varphi. \tag{6}$$

During the rotation of the scanner's mirror the unit directional normal vector $\mathbf{N}$ of the mirror is transformed with respect to Eq. (5). The effect of rotation axis errors can be determined with Eq. (6).

C. Reflection from the Mirror and the System of Mirrors

Let a beam, which is defined by unit directional vector $\mathbf{A}_1$, intersect planar mirror Z with unit normal vector $\mathbf{N}$ directed to the incident beam, as is shown in Fig. 2.

Unit directional vector $\mathbf{A}_2$ of the reflected beam is then determined with the use of Eq. (4) as

$$\mathbf{A}_2 = \mathbf{A}_1 - 2\mathbf{N}(\mathbf{A}_1\cdot\mathbf{N}) = \mathbf{M}\mathbf{A}_1. \tag{7}$$

Matrix $\mathbf{M}$ and vectors $\mathbf{A}_1$, $\mathbf{A}_2$, and $\mathbf{N}$ are given as

$$\mathbf{M} = \begin{pmatrix} 1-2N_x^2 & -2N_xN_y & -2N_xN_z \\ -2N_xN_y & 1-2N_y^2 & -2N_yN_z \\ -2N_xN_z & -2N_yN_z & 1-2N_z^2 \end{pmatrix},$$

$$\mathbf{A}_1 = \begin{pmatrix} A_{1x} \\ A_{1y} \\ A_{1z} \end{pmatrix}, \qquad \mathbf{A}_2 = \begin{pmatrix} A_{2x} \\ A_{2y} \\ A_{2z} \end{pmatrix},$$

$$\mathbf{N} = \begin{pmatrix} N_x \\ N_y \\ N_z \end{pmatrix}. \tag{8}$$

For the reflection on k mirrors with repeated use of Eq. (7) one gets the determinant

$$\mathbf{A}_{k+1} = \begin{vmatrix} \mathbf{A}_1 & 2(\mathbf{A}_1\cdot\mathbf{N}_1) & 2(\mathbf{A}_1\cdot\mathbf{N}_2) & 2(\mathbf{A}_1\cdot\mathbf{N}_3) & \dots & 2(\mathbf{A}_1\cdot\mathbf{N}_k) \\ \mathbf{N}_1 & 1 & 2(\mathbf{N}_1\cdot\mathbf{N}_2) & 2(\mathbf{N}_1\cdot\mathbf{N}_3) & \dots & 2(\mathbf{N}_1\cdot\mathbf{N}_k) \\ \mathbf{N}_2 & 0 & 1 & 2(\mathbf{N}_2\cdot\mathbf{N}_3) & \dots & 2(\mathbf{N}_2\cdot\mathbf{N}_k) \\ \mathbf{N}_3 & 0 & 0 & 1 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{N}_k & 0 & 0 & 0 & \dots & 1 \end{vmatrix}, \tag{9}$$

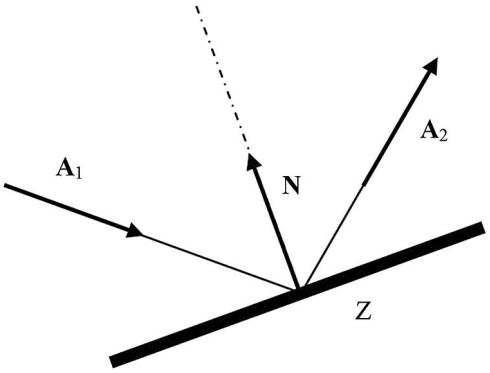


Fig. 2. Reflection on the planar mirror.

where $\mathbf{A}_{k+1}$ is the unit directional vector of the reflected beam from the system of k mirrors. For the system of two mirrors, the most practically used case, $\mathbf{A}_3$, is derived according to Eq. (9) as follows:

$$\mathbf{A}_3 = \begin{vmatrix} \mathbf{A}_1 & 2(\mathbf{A}_1\cdot\mathbf{N}_1) & 2(\mathbf{A}_1\cdot\mathbf{N}_2) \\ \mathbf{N}_1 & 1 & 2(\mathbf{N}_1\cdot\mathbf{N}_2) \\ \mathbf{N}_2 & 0 & 1 \end{vmatrix}$$
$$= \mathbf{A}_1 - 2\mathbf{N}_1(\mathbf{A}_1\cdot\mathbf{N}_1) - 2\mathbf{N}_2(\mathbf{A}_1\cdot\mathbf{N}_2) + 4\mathbf{N}_2(\mathbf{N}_1\cdot\mathbf{N}_2)(\mathbf{A}_1\cdot\mathbf{N}_1). \tag{10}$$

D. Beam Reflection from a Rotating Mirror and a System of Rotating Mirrors

If one rotates a mirror around any axis, the direction of the reflected beam will be changed depending on the angle of the mirror rotation and the direction of the axis around which the mirror rotates. Thus the direction of the mirror's unit normal directional vector will be changed with respect to Eq. (5). Consider the mirror's rotation around vector $\mathbf{C}_1$ by angle φ_1. Then with Eq. (5) one gets for the rotation of the mirror's normal vector

$$\mathbf{N}(\varphi_1) = \mathbf{N}(0)\cos\varphi_1 + \mathbf{C}_1(\mathbf{C}_1\cdot\mathbf{N}(0))(1-\cos\varphi_1) + (\mathbf{C}_1\times\mathbf{N}(0))\sin\varphi_1, \tag{11}$$

where $\mathbf{N}(0)$ denotes the initial position of the unit directional vector. For unit vector $\mathbf{A}_2$ of the reflected beam from the rotated mirror according to Eq. (7) one gets

$$\mathbf{A}_2 = \mathbf{A}_1 - 2\mathbf{N}(\varphi_1)(\mathbf{A}_1 \cdot \mathbf{N}(\varphi_1)), \tag{12}$$

where $\mathbf{A}_1$ is the unit directional vector of the incident beam. For the reflection from the k rotating mirrors one gets, with Eqs. (5) and (7), formulas

$$\begin{aligned}\mathbf{N}_i(\varphi_i) &= \mathbf{N}_i(0)\cos\varphi_i + \mathbf{C}_i(\mathbf{C}_i \cdot \mathbf{N}_i(0))(1-\cos\varphi_i) \\ &\quad + [\mathbf{C}_i \times \mathbf{N}_i(0)]\sin\varphi_i, \\ \mathbf{A}_{i+1}(\varphi_i) &= \mathbf{A}_i(\varphi_i) - 2\mathbf{N}_i(\varphi_i)(\mathbf{A}_i(\varphi_i) \cdot \mathbf{N}_i(\varphi_i)),\end{aligned} \tag{13}$$

where $i = 1, 2, \ldots, k$ and $\mathbf{A}_1(\varphi_1) = \mathbf{A}_1(0)$ is the directional vector of the beam, which is the incident to mirror Z_1.

3. Calculation of Optical Beam Spot in Detection Plane and Its Accuracy

To calculate beam pattern position in the detection plane, it is necessary to model the trajectory of the beam through an optical system. First, the procedure for one-mirror scanners will be shown, then for two-mirror scanners.

A. One-Mirror Scanners

Consider the situation shown in Fig. 3. The ray of light, which comes from the source S and is given by unit directional vector $\mathbf{A}_1$, goes through the timer T_S and impacts the mirror, which is rotated around axis $\mathbf{C}_1$ by angle φ_1. Furthermore, the entire scanner rotates around the axis of the unit directional vector $\mathbf{C}_2$ by angle φ_2. Let the mirror axis intersect the mirror at its center and consider the mirror center in the beginning of the coordinate system. Denote d_S as a distance from the timer T_S to the center of the mirror. The reflected beam continues and intersects detection plane ξ at point P. Subsequently, light of a beam is reflected from the point P and goes back in the reverse direction, impacts the mirror, and after reflection continues to the timer T_D. Consider the distance from the center of the mirror to the timer T_D to be d_D.

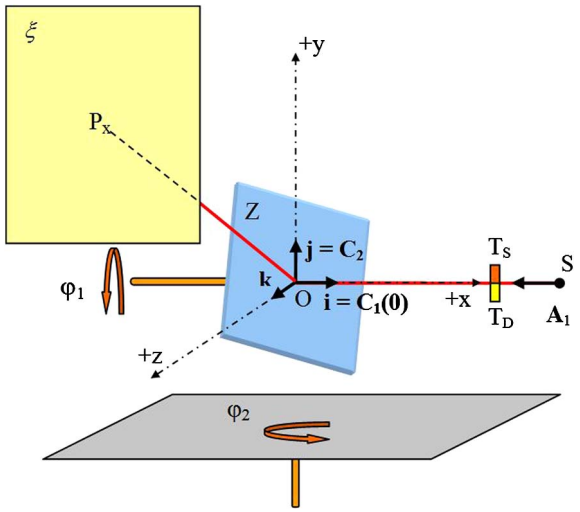


Fig. 3. One-mirror scanner.

The reflection from mirror Z is given by Eqs. (11) and (12). Due to complete scanner rotation by angle φ_2, the final unit vector $\mathbf{A}_3$ of the beam can be estimated as follows:

$$\begin{aligned}\mathbf{A}_3 &= \mathbf{A}_2\cos\varphi_2 + \mathbf{C}_2(\mathbf{C}_2 \cdot \mathbf{A}_2)(1-\cos\varphi_2) \\ &\quad + (\mathbf{C}_2 \times \mathbf{A}_2)\sin\varphi_2\end{aligned} \tag{14}$$

or

$$\mathbf{A}_3^T = \mathbf{R}_y(\varphi_2)\mathbf{A}_2^T, \tag{15}$$

where

$$\mathbf{R}_y(\varphi_2) = \begin{pmatrix} \cos\varphi_2 & 0 & \sin\varphi_2 \\ 0 & 1 & 0 \\ -\sin\varphi_2 & 0 & \cos\varphi_2 \end{pmatrix} \tag{16}$$

is the rotation matrix around y axis by angle φ_2 [29–31].

To calculate vector $\mathbf{r}_P$—the intersection of the beam with detection plane ξ, which is given by formula [29–31]

$$\xi \equiv ((\mathbf{r} - b\mathbf{q}) \cdot \mathbf{q}) = 0, \tag{17}$$

where $\mathbf{q}$ is the unit directional normal vector of the plane ξ, $\mathbf{r}$ is the position vector of an arbitrary point in the plane ξ, and b is the distance of the plane from the origin—then one gets, due to the condition of the reflection in the origin of coordinate system (x, y, z),

$$\mathbf{r}_P = p\mathbf{A}_3. \tag{18}$$

Subsequently, substitution to Eq. (17) leads to determination of parameter p, and the resulting formula follows:

$$\mathbf{r}_P = \frac{b}{\mathbf{A}_3 \cdot \mathbf{q}}\mathbf{A}_3. \tag{19}$$

For practical purposes, the requirement for vector $\mathbf{r}_P$ as a function of angles of rotation φ_1 and φ_2 and a transit time t is made. One can write the following formula:

$$\mathbf{r}_P = d\mathbf{A}_3, \tag{20}$$

where d denotes a distance from the center of the mirror to point P. If one considers c to be a velocity of used light in a medium, the formula for a measured distance d_m is obtained as

$$d_m = ct = 2d + d_S + d_D, \tag{21}$$

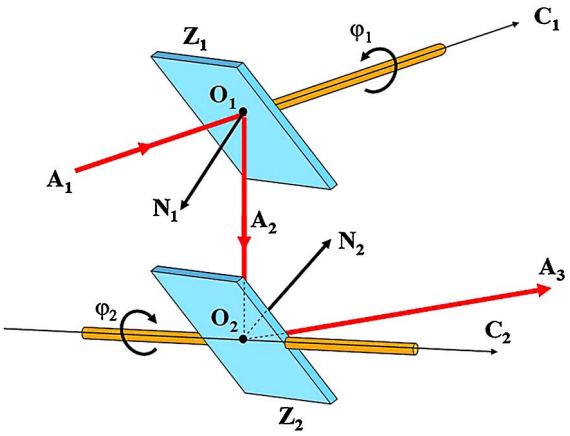


Fig. 4. System of two mirrors.

where transit time t is given as a difference between detected time in the timer T_D and T_S ($t = T_D - T_S$).

Rearranging Eq. (21) and substituting to Eq. (20) leads to the following formula:

$$\mathbf{r}_P = \frac{1}{2}(ct - d_S - d_D)\mathbf{A}_3. \tag{22}$$

From equations stated above one can see that components of the unit directional vector $\mathbf{A}_3$ are functions of angles of rotation φ_1 and φ_2. This vector is multiplied by a function of measured transit time t and parameters of used optical system.

B. Two-Mirror Scanners

Consider a system of two planar mirrors Z_1 and Z_2 as shown in Fig. 4. Suppose that mirror Z_1 passes through point O_1 and rotates around an axis having unit directional vector $\mathbf{C}_1$, and mirror Z_2 passes through point O_2 and rotates around an axis having unit directional vector $\mathbf{C}_2$. Further, let $\mathbf{N}_1$ and $\mathbf{N}_2$ be unit vectors of normals to mirrors Z_1 and Z_2, and $\mathbf{A}_1$ is the unit directional vector of the beam incident on mirror Z_1, $\mathbf{A}_2$ is the unit directional vector of the reflected beam from mirror Z_1 and the incident on mirror Z_2, and $\mathbf{A}_3$ is the unit directional vector of the reflected beam from mirror Z_2.

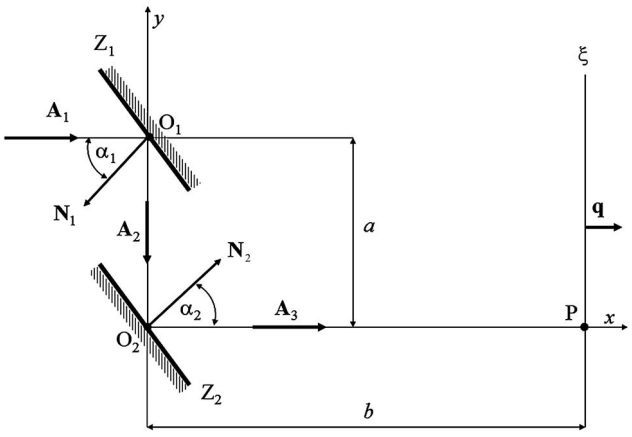


Fig. 5. System of two mirrors in the plane (x, y).

Rotation of mirror Z_1 by angle φ_1 around an axis that is defined by unit directional vector $\mathbf{C}_1$ and mirror Z_2 by angle φ_2 around the axis that is defined by unit directional vector $\mathbf{C}_2$ with respect to initial positions can be easily described with Eq. (13).

Now calculate the coordinates of the intersection of the beam reflected from the system of mirrors Z_1 and Z_2 in detection plane ξ (Fig. 5). Denote unit vectors in the directions of the coordinate axes (x, y, z) as (i, j, k). Axes x and y lie in the plane of Fig. 5 and the z axis is perpendicular to this plane. We say that the mirror system is in initial position, if vectors $\mathbf{A}_1$ and $\mathbf{A}_3$ are lying in mutually parallel planes. Vectors belonging to the initial position are denoted $\mathbf{A}_1(0)$, $\mathbf{A}_2(0)$, $\mathbf{A}_3(0)$, $\mathbf{N}_1(0)$, and $\mathbf{N}_2(0)$. The beam coming from the mirror system in the initial position then intersects plane ξ, which is located at distance b from point O_2, at point P. Points O_1 and O_2 are separated by the value of a and point O_2 is placed in the origin of the coordinate system.

Vector $\mathbf{r}_P$ of the beam intersection with detection plane ξ is obtained by a similar procedure as in the case of a one-mirror scanner, but considering the reflections on two mirrors the resulting formula is

$$\mathbf{r}_P = \mathbf{r}_2 + p_3\mathbf{A}_3(\varphi_2), \tag{23}$$

where

$$\mathbf{r}_2 = a\mathbf{j} - a\frac{\mathbf{j}\cdot\mathbf{N}_2(\varphi_2)}{\mathbf{A}_2(\varphi_1)\cdot\mathbf{N}_2(\varphi_2)}\mathbf{A}_2(\varphi_1), \tag{24}$$

$$p_3 = \frac{b + \mathbf{r}_2\cdot\mathbf{q}}{\mathbf{A}_3(\varphi_2)\cdot\mathbf{q}}, \tag{25}$$

$$\mathbf{A}_3(\varphi_2) = \mathbf{A}_2(\varphi_1) - 2\mathbf{N}_2(\varphi_2)(\mathbf{A}_2(\varphi_1)\cdot\mathbf{N}_2(\varphi_2)), \tag{26}$$

$$\mathbf{A}_2(\varphi_1) = \mathbf{A}_1(0) - 2\mathbf{N}_1(\varphi_1)(\mathbf{A}_1(0)\cdot\mathbf{N}_1(\varphi_1)), \tag{27}$$

where $\mathbf{r}_2$ is the position vector of the point where the beam intersects mirror Z_2. $\mathbf{N}_1(\varphi_1)$ and $\mathbf{N}_2(\varphi_2)$ are given by Eq. (13).

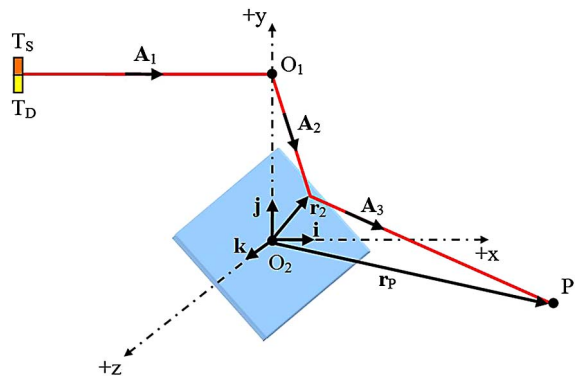


Fig. 6. Reflection on second mirror.

To calculate vector $\mathbf{r}_P$ as a function of angles of rotation φ_1 and φ_2 and a transit time t, consider the situation depicted in Fig. 6. Measured transit time t can be expressed as a sum of several parts. First, the beam goes from the timer T_S to point O_1 and passes a distance d_S. Then it is reflected from the first mirror and passes a distance a' to a point given by vector $\mathbf{r}_2$. Subsequently, the beam is reflected from the second mirror and through a distance d to point P. Next, light is reflected and a part of it goes reversely by the distance d and a'. At last, from point O_1 to the timer T_D it goes through a distance d_D. If c is considered to be a speed of used light, then the formula for measured distance d_m follows:

$$d_m = ct = 2(d + a') + d_S + d_D. \tag{28}$$

Suppose next the above-mentioned situation where centers of mirrors are separated by the value of a, point O_2 is placed in the origin of the coordinate system, and point O_1 lies on axis y. One can see that position vector $\mathbf{r}_P$ given by the formula

$$\mathbf{r}_P = \mathbf{r}_2 + d\mathbf{A}_3(\varphi_2), \tag{29}$$

where d is given with Eq. (28) as follows:

$$d = \frac{1}{2}(ct - d_S - d_D) - a'. \tag{30}$$

The last determined value is a'. With respect to Fig. 6 and to Eq. (24) one obtains

$$a' = |\mathbf{r}_2 - a\mathbf{j}| = -a\frac{\mathbf{j}\cdot\mathbf{N}_2(\varphi_2)}{\mathbf{A}_2(\varphi_1)\cdot\mathbf{N}_2(\varphi_2)}. \tag{31}$$

With respect to the above-stated results, the final formula for position vector $\mathbf{r}_P$, which is given as a function of angles of rotation φ_1 and φ_2 and transit time t, can be expressed as follows:

$$\mathbf{r}_P = a\mathbf{j} + a\frac{\mathbf{j}\cdot\mathbf{N}_2(\varphi_2)}{\mathbf{A}_2(\varphi_1)\cdot\mathbf{N}_2(\varphi_2)}[\mathbf{A}_3(\varphi_2) - \mathbf{A}_2(\varphi_1)] + \frac{1}{2}(ct - d_S - d_D)\mathbf{A}_3(\varphi_2). \tag{32}$$

Vectors $\mathbf{A}_2(\varphi_1)$ and $\mathbf{A}_3(\varphi_2)$ are given by Eqs. (26) and (27).

C. Estimation of Accuracy

Positional accuracy of the laser scanner's beam spot in the detection plane is evaluated by the coordinate standard deviation and by the sample coordinate standard deviation, which are defined as

$$\sigma_{xy} = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{2}}, \qquad s_{xy} = \sqrt{\frac{s_x^2 + s_y^2}{2}}, \tag{33}$$

where σ_i and s_i denote the standard deviation and the sample standard deviation of the ith coordinate. The sample standard deviation of one coordinate is given as

$$s_i = \sqrt{\frac{\sum_j \delta_{ij}^2}{n'}}, \tag{34}$$

where $\delta_{ij} = i'_j - i_j$ is the difference between the average value of the jth point of the ith coordinate i'_j and the measured values of the jth point i_j and n' is the number of redundant measurements.

To calculate the values of the standard deviations σ_x and σ_y, one can use the well-known law of variance propagation [32,33]. Consider a model given by a vector function $\mathbf{f} = (f_1, f_2, \ldots, f_m)^T$ of n variables x_i $(i = 1, \ldots, n)$. If function $\mathbf{f}$ has continuous partial derivatives, actual errors of variables x_i are small compared to the functional values and have an even distribution with zero expected value. The relationship for the variance–covariance matrix of functional values can be derived [32,33] as

$$\Sigma_f = \mathbf{J}\Sigma_x\mathbf{J}^T. \tag{35}$$

Variances of functional values then will be situated on the diagonal of the matrix Σ_f. In Eq. (35) $\mathbf{J}$ denotes the Jacobi matrix of partial derivatives of the model with respect to all variables [32,33] and Σ_x is the variance–covariance matrix of the variables [32,33]. If variables are mutually independent, thus $\mathrm{Cov}(x_i, x_j) = 0$, $i \neq j$, the variance–covariance matrix is then diagonal and one can rewrite Eq. (35) to the form

$$\sigma_{fi}^2 = \sum_{j=1}^{n} (\partial f_i/\partial x_j)^2 \sigma_{xj}^2. \tag{36}$$

For our purposes, one can consider that one-mirror and two-mirror optical scanners fulfill the above-stated conditions; therefore, the law of variance propagation can be used. Relations between standard deviations and uncertainties are clearly described in [34].

In 3D problems, accuracy of a measured point is completely characterized by the error ellipsoid. As known, its semiaxes are determined as a square root of eigenvalues [29–31] of the variance–covariance matrix [Eq. (35)] and directions of semiaxes are represented by eigenvectors [29–31].

Another easy method to identify the characteristics of accuracy is numerical modeling. If we know the standard deviations of all variables and functional relationships, we can, for example, simulate actual measurements, using the MATLAB pseudo-random number generator. With a sufficient number of repetitions, we can *a posteriori* calculate the characteristics of precision—the sample standard deviations [Eq. (34)].

4. Examples

A. Example 1—One-Mirror Scanner

Consider a situation of a one-mirror scanner with parameters (Fig. 3) $\mathbf{A}_1 = -\mathbf{i}$, $\mathbf{C}_1 = \mathbf{i}$, $\mathbf{C}_2 = \mathbf{j}$, $\mathbf{N}(0) = (\mathbf{i} + \mathbf{k})/\sqrt{(2)}$, $\mathbf{q} = \mathbf{k}$, $b = 100$ m, $d_S = 0.05$ m, and $d_D = 0.05$ m.

With the use of Eqs. (11), (12), and (14)–(18) one gets

$$\mathbf{A}_2 = -\mathbf{j} \sin \varphi_1 + \mathbf{k} \cos \varphi_1,$$

$$\mathbf{A}_3 = \mathbf{i} \cos \varphi_1 \sin \varphi_2 - \mathbf{j} \sin \varphi_1 + \mathbf{k} \cos \varphi_1 \cos \varphi_2,$$

$$\mathbf{r}_P = \mathbf{i}b \tan \varphi_2 - \mathbf{j}b \frac{\tan \varphi_1}{\cos \varphi_2} + \mathbf{k}b. \tag{37}$$

Coordinates of point P on the detection plane are then given by

$$x_P = b \tan \varphi_2, \qquad y_P = -b \frac{\tan \varphi_1}{\cos \varphi_2}, \qquad z_P = b. \tag{38}$$

The above-stated situation is shown in Fig. 7.

The first and second formulas in Eq. (38) are parametrical equations of a curve in the detection plane. After a simple treatment one gets the equation of the curve in explicit form,

$$y_P = -\frac{\tan \varphi_1}{\sin \varphi_2} x_P. \tag{39}$$

Using the Taylor series [29–31] for coordinates x_P and y_P for small angles, considering just the power of 3 for angles of rotation, one gets

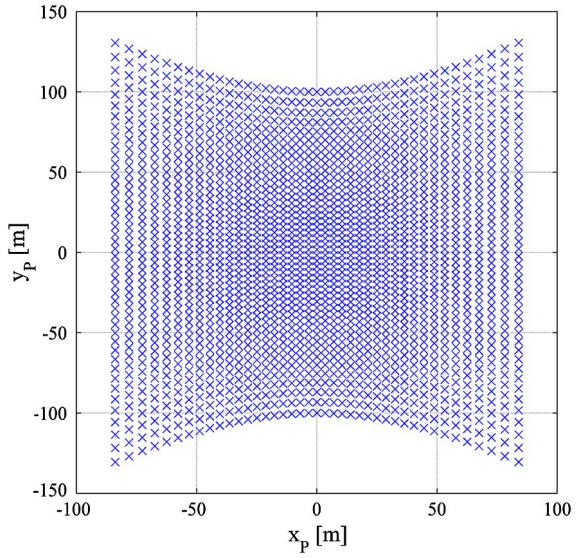


Fig. 7. Point P in the detection plane for $\varphi_1 = \{-45°, -43°, \ldots, 43°, 45°\}$ and $\varphi_2 = \{-40°, -38°, \ldots, 38°, 40°\}$ with parameters from example 1.

$$x_P \approx b\varphi_2 + \frac{b}{3}\varphi_2^3, \qquad y_P \approx -b\varphi_1 - \frac{b}{2}\varphi_1\varphi_2^2 - \frac{b}{3}\varphi_1^3, \tag{40}$$

where angles are substituted in radians. If we limit the series [Eq. (40)] only to the first components, we see that the coordinates of point P in the detection plane are approximately proportional to the angles of rotation. For relative error of the linear approximation of relations [Eq. (40)] one gets

$$\frac{\delta x_P}{x_P} = \frac{1}{3}\varphi_2^2, \qquad \frac{\delta y_P}{y_P} = \frac{1}{2}\varphi_2^2. \tag{41}$$

For example, for angle $\varphi_2 = 10° = 0.17$ rad is the relative error of linear approximation $\delta x_P/x_P \approx 1.0\%$ and $\delta y_P/y_P \approx 1.4\%$.

Consider mutually independent angles of the mirror rotation and the angle of the entire scanner rotation with standard deviations $\sigma_{\varphi 1}$ and $\sigma_{\varphi 2}$, then for exact analysis of the positional accuracy of the laser beam spot in the detection plane one uses Eq. (36). The coordinate standard deviation is then given by

$$\sigma_{xy} = \sqrt{\frac{\sigma_x^2 + \sigma_y^2}{2}}, \tag{42}$$

where

$$\sigma_x^2 = \left(\frac{\partial x_P}{\partial \varphi_1}\right)^2 \sigma_{\varphi 1}^2 + \left(\frac{\partial x_P}{\partial \varphi_2}\right)^2 \sigma_{\varphi 2}^2 = \left(\frac{b}{\cos^2 \varphi_2}\right)^2 \sigma_{\varphi 2}^2, \tag{43}$$

$$\sigma_y^2 = \left(\frac{\partial y_P}{\partial \varphi_1}\right)^2 \sigma_{\varphi 1}^2 + \left(\frac{\partial y_P}{\partial \varphi_2}\right)^2 \sigma_{\varphi 2}^2 = \left(\frac{-b}{\cos^2 \varphi_1 \cos \varphi_2}\right)^2 \sigma_{\varphi 1}^2 + \left(\frac{-b \tan \varphi_1 \sin \varphi_2}{\cos^2 \varphi_2}\right)^2 \sigma_{\varphi 2}^2. \tag{44}$$

Figure 8 presents the coordinate standard deviation in the detection plane.

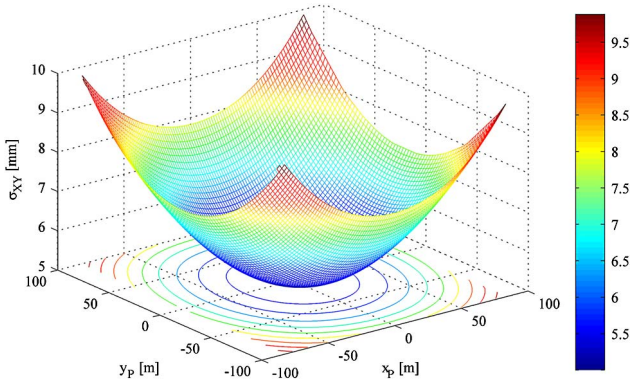


Fig. 8. Coordinate standard deviation in the detection plane—example 1.

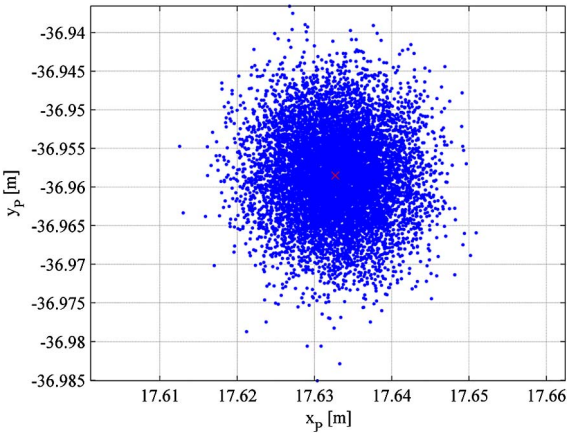


Fig. 9. Simulation of the laser beam spot in the detection plane for the expected values of the angles $\varphi_1 = 20°$ and $\varphi_2 = 10°$ and 10,000 repetitions.

For the simulation and *a posteriori* determination of the characteristics of accuracy assume a normal distribution with zero expected value of mirror rotation errors and standard deviations $\sigma_{\varphi 1} = \sigma_{\varphi 2} = 50$ μrad. Figure 9 shows the simulation of the laser beam spot for the expected values of the angles $\varphi_1 = 20°$ and $\varphi_2 = 10°$ for 10,000 repetitions. Figure 10 presents the progression of the sample coordinate standard deviation with an increase in the number of repetitions.

Position vector $\mathbf{r}_P$, as a function of angles of rotation φ_1 and φ_2 and transit time t, is determined with respect to Eq. (22). For its components, with consideration of the above-stated formulas, it holds that

$$\begin{aligned} x_P &= \frac{1}{2}(ct - d_S - d_D)\cos\varphi_1 \sin\varphi_2, \\ y_P &= -\frac{1}{2}(ct - d_S - d_D)\sin\varphi_1, \\ z_P &= \frac{1}{2}(ct - d_S - d_D)\cos\varphi_1 \cos\varphi_2. \end{aligned} \tag{45}$$

Diagonal elements of the variance–covariance matrix, calculated using law of variance propagation, represent coordinate accuracy of the measured point. Variables φ_1, φ_2, and t are mutually independent; the variance–covariance matrix will be diagonal. Product $ct = d_m$ is more suitable for a numerical

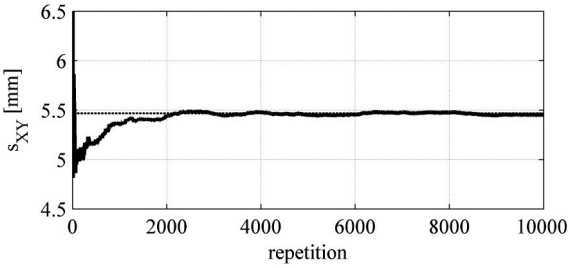


Fig. 10. Progression of sample coordinate standard deviation with an increase in the number of repetitions for angles $\varphi_1 = 20°$ and $\varphi_2 = 10°$.

estimation of accuracy instead of t, which is disproportionately small in comparison with other variables. One can simply derive the following formula for standard deviation σ_{ct} using law of variance propagation:

$$\sigma_{ct} = c\sigma_t, \tag{46}$$

where σ_t represents accuracy of measured transit time. Equation (46) is suitable for expression of requirements for technical parts of the scanner optical system. For accuracy in measured pseudodistance d_m equal to 0.001 m, with speed of used light approximately equal to $3 \cdot 10^8$ m, it holds that $\sigma_t \approx 3.3 \cdot 10^{-12}$ s $= 3.3$ ps; the timer has to accomplish this condition.

Applying law of variance propagation [Eq. (36)] on the situation described with Eq. (45), one obtains

$$\begin{aligned} \sigma_x^2 = &\left[\frac{1}{2}(d_S + d_D - ct)\sin\varphi_1 \sin\varphi_2\right]^2 \sigma_{\varphi 1}^2 \\ &+ \left[\frac{1}{2}(ct - d_S - d_D)\cos\varphi_1 \cos\varphi_2\right]^2 \sigma_{\varphi 2}^2 \\ &+ \left[\frac{1}{2}\cos\varphi_1 \sin\varphi_2\right]^2 \sigma_{ct}^2, \end{aligned} \tag{47}$$

$$\sigma_y^2 = \left[\frac{1}{2}(d_S + d_D - ct)\cos\varphi_1\right]^2 \sigma_{\varphi 1}^2 + \left[-\frac{1}{2}\sin\varphi_1\right]^2 \sigma_{ct}^2, \tag{48}$$

$$\begin{aligned} \sigma_z^2 = &\left[\frac{1}{2}(d_S + d_D - ct)\sin\varphi_1 \cos\varphi_2\right]^2 \sigma_{\varphi 1}^2 \\ &+ \left[\frac{1}{2}(d_S + d_D - ct)\cos\varphi_1 \sin\varphi_2\right]^2 \sigma_{\varphi 2}^2 \\ &+ \left[\frac{1}{2}\cos\varphi_1 \cos\varphi_2\right]^2 \sigma_{ct}^2. \end{aligned} \tag{49}$$

Eigenvalues and eigenvectors of the diagonal variance–covariance matrix with elements given by Eqs. (47)–(49) can be used for determination of error ellipsoids in 3D situations. In Fig. 11, ellipsoids in plane (xz) are shown for different values of transit time t. As is obvious, influence of timer accuracy is more significant in closer parts of the scanned field. Remote parts are more affected with errors of angles of rotation.

B. Example 2—Two-Mirror Scanner

Consider a two-mirror scanner with parameters (Fig. 12)

$$\begin{aligned} &\mathbf{A}_1(0) = \mathbf{i}, \quad \mathbf{C}_1 = \mathbf{i}, \quad \mathbf{C}_2 = \mathbf{k}, \quad \mathbf{q} = \mathbf{i}, \\ &\alpha_1 = \alpha_2 = 45°, \\ &\mathbf{N}_1(0) = -(\mathbf{i} + \mathbf{j})/\sqrt{2}, \quad \mathbf{N}_2(0) = (\mathbf{i} + \mathbf{j})/\sqrt{2}, \end{aligned} \tag{50}$$

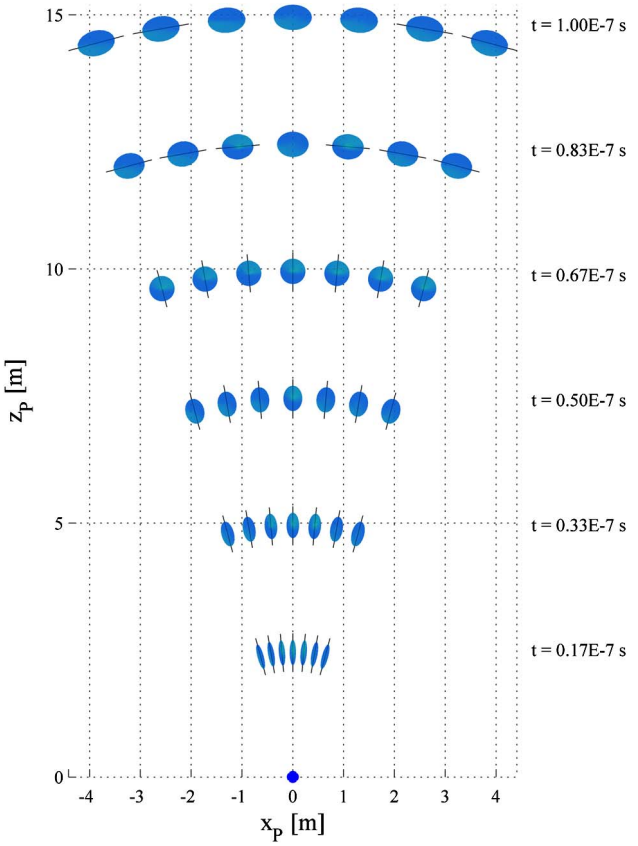


Fig. 11. Error ellipsoids (scaled 500:1) for different transit time t, $\varphi_1 = 0°$ and $\varphi_2 = \{-15°, -10°, \ldots, 10°, 15°\}$ with parameters from example 1.

$a = 0.05$ m, $b = 100$ m, $d_S = 0.05$ m, and $d_D = 0.05$ m.

With the use of Eqs. (13), (26), and (27) one gets

$$\mathbf{N}_1(\varphi_1) = -\frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j}\cos\varphi_1 + \mathbf{k}\sin\varphi_1), \tag{51}$$

$$\mathbf{A}_2(\varphi_1) = -(\mathbf{j}\cos\varphi_1 + \mathbf{k}\sin\varphi_1), \tag{52}$$

$$\mathbf{N}_2(\varphi_2) = \frac{\mathbf{i}}{\sqrt{2}}(\cos\varphi_2 - \sin\varphi_2) + \frac{\mathbf{j}}{\sqrt{2}}(\cos\varphi_2 + \sin\varphi_2), \tag{53}$$

$$\mathbf{A}_3(\varphi_2) = \mathbf{i}\cos\varphi_1(1-2\sin^2\varphi_2) + \mathbf{j}\cos\varphi_1\sin 2\varphi_2 - \mathbf{k}\sin\varphi_1. \tag{54}$$

Substitution of the estimated vectors in (24) and (26) leads to

$$\mathbf{r}_2 = -a\mathbf{k}\tan\varphi_1, \tag{55}$$

$$\mathbf{r}_P = \mathbf{i}b + \mathbf{j}b\tan 2\varphi_2 - \mathbf{k}\tan\varphi_1\left(a + \frac{b}{\cos 2\varphi_2}\right). \tag{56}$$

Coordinates of point P in the detection plane are then given by the following formulas:

$$x_P = b, \qquad y_P = b\tan 2\varphi_2, \qquad z_P = -\tan\varphi_1\left(a + \frac{b}{\cos 2\varphi_2}\right). \tag{57}$$

The second and third formulas in (57) are parametrical equations of the curve in the detection plane. After a simple treatment, one gets the equation of the curve in explicit form:

$$z_P = -\left(\frac{\tan\varphi_1}{\sin 2\varphi_2}\right)y_P - a\tan\varphi_1. \tag{58}$$

For small mirror rotations, with the use of the Taylor series, one gets from Eq. (57):

$$y_P \approx 2b\varphi_2 + \frac{8}{3}b\varphi_2^3, \qquad z_P \approx -\varphi_1(a+b) - 2b\varphi_1\varphi_2^2, \tag{59}$$

where just the powers of 3 are considered for the rotation angles. Angles φ_1 and φ_2 are substituted in radians. If we limit the series [Eq. (57)] only to the first members, we see that the coordinates of point P in the detection plane are approximately proportional to the angles of rotation. For relative error of the linear approximation of relations [Eq. (57)] one gets

$$\frac{\delta y_P}{y_P} = \frac{4}{3}\varphi_2^2, \qquad \frac{\delta z_P}{z_P} \approx \frac{2}{1 + a/b}\varphi_2^2 \approx 2\varphi_2^2, \tag{60}$$

where we supposed $a/b \ll 1$. For example, for angle $\varphi_2 = 10° = 0.17$ rad is the relative error of linear approximation $\delta y_P/y_P \approx 4\%$ and $\delta z_P/z_P \approx 6\%$. With the use of the same procedure as in example 1, one gets for the coordinate standard deviation formula

$$\sigma_{yz} = \sqrt{\frac{\sigma_y^2 + \sigma_z^2}{2}}, \tag{61}$$

where

$$\sigma_y^2 = \left(\frac{\partial y_P}{\partial\varphi_1}\right)^2\sigma_{\varphi 1}^2 + \left(\frac{\partial y_P}{\partial\varphi_2}\right)^2\sigma_{\varphi 2}^2 = \left(\frac{2b}{\cos^2 2\varphi_2}\right)^2\sigma_{\varphi 2}^2, \tag{62}$$

$$\sigma_z^2 = \left(\frac{\partial z_P}{\partial\varphi_1}\right)^2\sigma_{\varphi 1}^2 + \left(\frac{\partial z_P}{\partial\varphi_2}\right)^2\sigma_{\varphi 2}^2 = \left(\frac{b + a\cos 2\varphi_2}{\cos^2\varphi_1\cos 2\varphi_2}\right)^2\sigma_{\varphi 1}^2 + \left(\frac{2b\tan\varphi_1\tan 2\varphi_2}{\cos 2\varphi_2}\right)^2\sigma_{\varphi 2}^2. \tag{63}$$

Figure 13 presents the coordinate standard deviation in the detection plane.

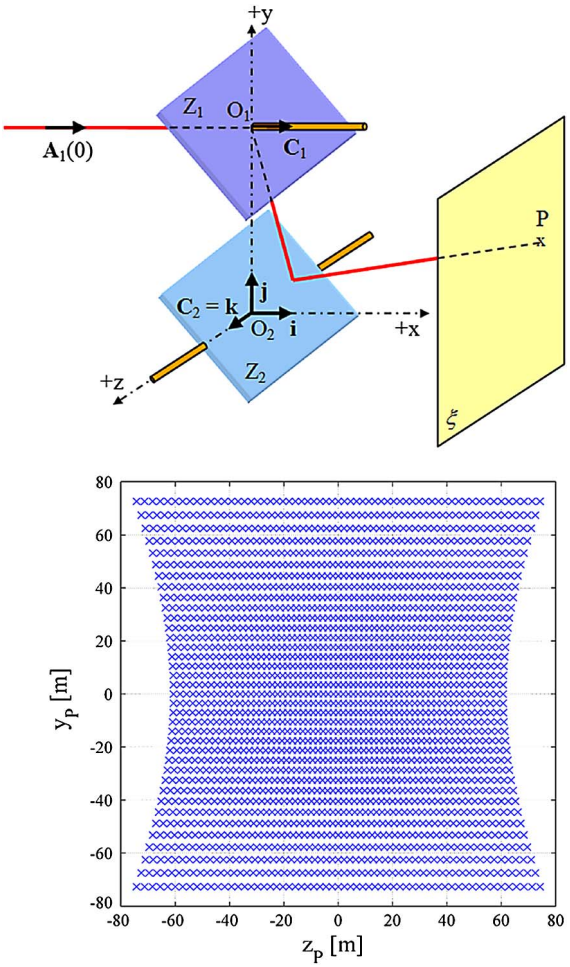


Fig. 12. Scheme of two-mirror scanner and point P in the detection plane for $\varphi_1 = \{-30°, -29°, \ldots, 29°, 30°\}$, $\varphi_2 = \{-18°, -17°, \ldots, 17°, 18°\}$ with parameters from example 2.

For the simulation and *a posteriori* determination of the accuracy characteristics one can assume a normal distribution with zero expected value of mirror rotation errors and with supposed standard deviations and obtain similar results as in example 1.

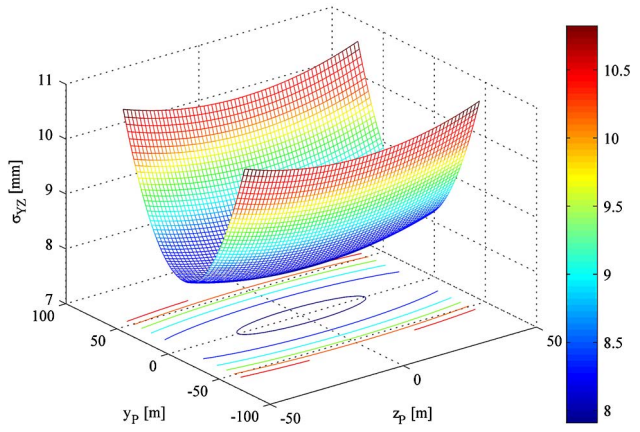


Fig. 13. Coordinate standard deviation in the detection plane—example 2.

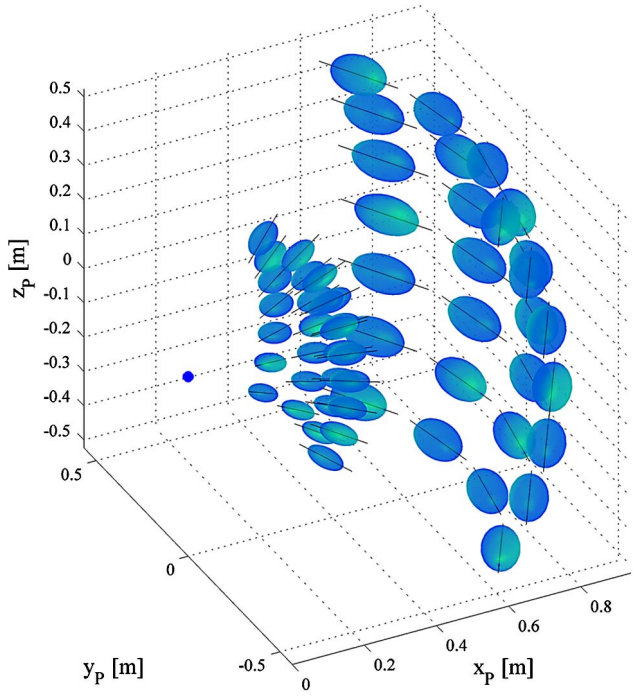


Fig. 14. Error ellipsoids (scaled 1000:1) for transit time $t = \{0.33 \cdot 10^{-8}\ \text{s}, 0.67 \cdot 10^{-8}\ \text{s}\}$, $\varphi_1 = \{-30°, -20°, \ldots, 20°, 30°\}$ and $\varphi_2 = \{-15°, -5°, 5°, 15°\}$ with parameters from example 2.

Position vector $\mathbf{r}_P$, as a function of angles of rotation φ_1 and φ_2 and transit time t, is determined with respect to Eq. (32). For its components, with a consideration of the above-stated formulas, it holds that

$$x_P = \left[\frac{1}{2}(ct - d_S - d_D) - \frac{a}{\cos\varphi_1}\right]\cos\varphi_1 \cos 2\varphi_2,$$
$$y_P = \left[\frac{1}{2}(ct - d_S - d_D) - \frac{a}{\cos\varphi_1}\right]\cos\varphi_1 \sin 2\varphi_2,$$
$$z_P = \frac{1}{2}[(d_S + d_D - ct)\sin\varphi_1]. \tag{64}$$

With the use of same procedure as in example 1 one can calculate diagonal elements of the variance–covariance matrix. Subsequently, error ellipsoids are determined with eigenvalues and eigenvectors. In Fig. 14, one can see scaled ellipsoids in the model situation from example 2. Standard deviation for a timer was chosen as $\sigma_t = 3.3 \cdot 10^{-13}$ s. It is equivalent with accuracy 0.0001 m in the measured pseudodistance d_m. Standard deviations of mirror rotations were chosen as $\sigma_{\varphi 1} = \sigma_{\varphi 2} = 50$ μrad.

As is obvious, timer accuracy more affects the closer parts. Influence of angles of rotation is more significant in the remote parts.

5. Conclusions

The objective of this article was to present tracing of laser beams through the system of mirrors that is used for the construction of 3D laser scanners. The procedures for modeling of one-mirror and

two-mirror systems, which are used most often in practice, were derived. Furthermore, an analysis of the position of the laser beam spot in the detection plane was done. Subsequently, formulas for point position as a function of angles of rotation and a measured transit time were derived; application in practical 3D situations then can be done.

Position accuracy in a detection plane was determined with the use of coordinate standard deviation, calculated with the law of variance propagation, and with the use of the sample coordinate standard deviation, calculated *a posteriori* from the numerical modeling. In practical applications it is important to estimate accuracy in three dimensions. The most suitable approach is to calculate error ellipsoids from the variance–covariance matrix. One can make this procedure with the use of eigenvalues and eigenvectors; computations and results have been analyzed.

Formulas and procedures shown in the paper are a simple solution for modeling of the optical systems for laser scanners and for practical use in production. Exact analysis has shown that the accuracy of the presented optical systems is less then 10 mm for 100 m distance, if we consider the expected value of mirror rotation errors 50 μrad. Numerical modeling endorsed the results of the exact solution; moreover, the size and the direction of the laser beam spot scattering in the detection plane shown is actually appropriate for construction processes. As analysis in three dimensions shows, influence of timer accuracy is more significant in closer parts of the scanned field. Remote parts are more affected with errors of angles of rotation.

This work was supported by the grant 13-31765S from the Czech Science Foundation and SGS14/110/OHK1/2T/11 from Czech Technical University in Prague.

References

1. G. F. Marshall and G. E. Stutz, *Handbook of Optical and Laser Scanning*, Vol. **147** (CRC, 2011).
2. G. Vosselman and H.-G. Maas, *Airborne and Terrestrial Laser Scanning* (Whittles, 2010).
3. "Control System—Laser Scanning—Surveying Works," 2013, http://www.controlsystem.cz/en.
4. "Surphaser 3D Scanners," 2013, http://www.surphaser.com.
5. "Leica Geosystems—HDS," 2013, http://hds.leica-geosystems.com/en/index.htm.
6. "TOPCON Global Gateway," 2013, http://global.topcon.com.
7. "RIEGL Laser Measurement Systems," 2013, http://www.riegl.com.
8. "FARO Laser Scanner Focus3D," 2013, http://www.faro.com/focus/uk.
9. "MDL," 2013, http://www.mdl.co.uk/en/14735.aspx.
10. "Galvos—Scanning Mirrors—Optical Scanners," 2013, http://www.cambridgetechnology.com.
11. "Laser Sensors—IR Temperature Sensors—High Precision Displacement, and Position Measurement—Micro-Epsilon Measurement," 2013, http://www.micro-epsilon.com/index.html.
12. R. J. Pegis and M. Rao, "Analysis and design of plane-mirror systems," Appl. Opt. **2**, 1271–1274 (1963).
13. J. Cohen-Sabban, Y. Cohen-Sabban, and A. Roussel, "Distortion-free 2-D space and surface scanners using light deflectors," Appl. Opt. **22**, 3935–3942 (1983).
14. Y. J. Li and J. Katz, "Laser beam scanning by rotary mirrors. I. Modeling mirror-scanning devices," Appl. Opt. **34**, 6403–6416 (1995).
15. Y. J. Li, "Laser beam scanning by rotary mirrors. II. Conic-section scan patterns," Appl. Opt. **34**, 6417–6430 (1995).
16. Y. J. Li, "Single-mirror beam steering system: analysis and synthesis of high-order conic-section scan patterns," Appl. Opt. **47**, 386–398 (2008).
17. Y. J. Li, "Beam deflection and scanning by two-mirror and two-axis systems of different architectures: a unified approach," Appl. Opt. **47**, 5976–5985 (2008).
18. Y. Friedman and N. Schweitzer, "Classification of stable configurations of plane mirrors," Appl. Opt. **37**, 7229–7234 (1998).
19. R. Shinozaki, O. Sasaki, and T. Suzuki, "Fast scanning method for one-dimensional surface profile measurement by detecting angular deflection of a laser beam," Appl. Opt. **43**, 4157–4163 (2004).
20. P. Pokorny, "Theoretical foundations of one-mirror and two-mirror optical scanners," bachelor's thesis (Czech Technical University in Prague, Faculty of Civil Engineering, Prague, 2012).
21. A. Miks, *Applied Optics* (Czech Technical University, 2009), p. 230.
22. M. Born, E. Wolf, and A. B. Bhatia, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University, 1999).
23. A. Miks, *Physics 2: Electromagnetic Field* (Czech Technical University, 2005), p. 162.
24. A. Yariv, *Quantum Electronics* (Wiley, 1967).
25. B. E. A. Saleh and M. C. Teich, *Fundamentals of Photonics* (Wiley, 2007).
26. J. A. Stratton, *Electromagnetic Theory* (Adams, 2008).
27. A. Lipson, S. G. Lipson, and H. Lipson, *Optical Physics*, 4th ed. (Cambridge University, 2010), p. 572.
28. F. Trager, *Springer Handbook of Lasers and Optics* (Springer, 2007).
29. G. A. Korn and T. M. Korn, *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review* (Dover, 2000).
30. E. Madelung, *Die mathematischen Hilfsmittel des Physikers* (Springer-Verlag, 1964).
31. K. Rektorys, *Survey of Applicable Mathematics* (SNTL, 1969).
32. K.-R. Koch, *Parameter Estimation and Hypothesis Testing in Linear Models* (Springer, 1999).
33. L. Mervart and Z. Lukes, *Adjustment Calculus* (Czech Technical University, 2007).
34. S. V. Gupta, *Measurement Uncertainties: Physical Parameters and Calibration of Instruments* (Springer-Verlag, 2012).

applied optics

Deformation of a spherical lens caused by its own weight

Petr Pokorný,* Filip Šmejkal, and Antonín Mikš

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague, Czech Republic
*Corresponding author: petr.pokorny@fsv.cvut.cz

Received 29 August 2017; revised 2 November 2017; accepted 19 November 2017; posted 21 November 2017 (Doc. ID 305910); published 14 December 2017

This paper presents a method for calculation of a surface deformation of a spherical lens or a plane-parallel plate caused by its own weight. Formulas for the case of support at the edge or for the inner radius support are presented. Since it is not possible to find an analytical solution with appropriate boundary conditions, this paper presents numerical methods enabling us to calculate an approximate solution. The mentioned deflection has enormous impact in the field of optical metrology of large lenses where it has to be taken into account during precise measurements.

OCIS codes: (080.3630) Lenses; (120.3940) Metrology; (120.4630) Optical inspection; (220.4840) Testing.

https://doi.org/10.1364/AO.56.009984

1. INTRODUCTION

It is very important to measure geometrical and optical parameters very accurately during manufacturing of optical components. Surface topography is the one of the parameters that has a significant impact on imaging quality. Many approaches exist of precise measuring of the surface in the case of spherical or aspherical lenses [1–11], and interferometrical measurements are the most precise of them [12]. Many companies offer interferometers for very accurate measuring of optical surface topographies [7–11]. Deflection from a nominal shape of a surface under test is caused by manufacturing errors and by deformation caused by its own weight that can be present during the measuring process. Moreover, there are some limits for the accuracy that can be achieved during the surface shape measurement.

Many researchers presented works dealing with theory of deformation of plates over the past years. The deflection in individual points of supported planes was studied by Nadai [13,14]. Similarly, determination of plainness and bending of optical flats were presented by Emerson [15]. Elastic deflections of a thick circular mirror horizontally placed on a ring support was studied by Selke [16]. A paper on point supports of a telescope mirror and its deflections was shown by Nelson *et al.* [17]. A study of the symmetrical bending of an elastic circular plate supported at numerous internal points was published by Nong and Bao-lian [18]. Deflections of a thin or thick annular mirror and application to active mirror support optimization were studied by Arnold [19]. Mikš and Novák [20] presented in detail the effect of self-deflection of a plane-parallel plate caused by its own weight on interferometrical measurements.

The main goal of this paper is to present an analysis of a deformation of a spherical lens positioned horizontally together with several approaches of a numerical approximate solution of derived governing equations. Vertically positioned lenses are not investigated in this paper as the influence of self-deflection is not practically measurable in this case for typical lenses. Moreover, in practical industrial applications there are almost no situations of vertical positioning because of complications with fastening of lenses that could lead to unwanted tensions and deformations.

First, fundamental equations for the calculation of the deformation are presented. Afterward, different numerical approximate solutions are compared. The importance of the presented topic for very accurate measurements is shown in an example. To the best of the authors' knowledge, such a study for a spherical lens has not been published yet. Therefore, the reader can find the solution and a useful tool for practical modeling of the aforementioned issue in this paper.

2. GOVERNING EQUATIONS FOR A CALCULATION OF DEFORMATION OF A SPHERICAL LENS CAUSED BY ITS OWN WEIGHT

A lens can be considered as a circular plate of variable thickness from the point of view of a theory of elasticity. And a deflection caused by a lens's own weight will be much smaller than its own thickness.

First, consider a lens to be supported at the edge. As is known from the theory of elasticity, a deflection w in a vertical

direction, if one supposes a symmetrical loading, can be described by the equations [21–24]

$$\varphi = -\frac{dw}{dr}, \qquad M = D(r)\left[\frac{d\varphi}{dr} + \nu\frac{\varphi}{r}\right], \tag{1}$$

$$\frac{d^2\varphi}{dr^2} + \left[\frac{1}{r} + \frac{1}{D(r)}\frac{dD(r)}{dr}\right]\frac{d\varphi}{dr} + \left[\frac{\nu}{D(r)}\frac{dD(r)}{dr} - \frac{1}{r}\right]\frac{\varphi}{r} = \frac{Q(r)}{D(r)}, \tag{2}$$

$$D(r) = \frac{Ed^3(r)}{12(1-\nu^2)}, \qquad q(r) = d(r)\rho g, \tag{3}$$

where $w = w(r)$ is a vertical deflection of the plate (lens), r is a radial distance from the center, ν is the Poisson number, E is the Young modulus of elasticity, ρ is a volume density, g is the gravitational acceleration, $d(r)$ is a plate's thickness in the position r from its center, $\varphi = \varphi(r)$ is a slope of a tangent to the centerline, $M = M(r)$ is a radial inner moment on a length unit, $D(r)$ is a bending stiffness, $q(r)$ is a flat load, and $Q(r)$ is a pushing force per a unit distance. Note that the aforementioned formulas are valid only for thin circular plates where a ratio of its thickness and diameter is less than or equal to 1/8. This condition covers usual optical components that come for testing in output control processes during construction of optical systems.

Suppose now that the lens is spherical (it has both surfaces spherical), and therefore its thickness can be described as

$$\begin{aligned} d(r) &= d_0 - R_1\left(1 - \sqrt{1-(r/R_1)^2}\right) + R_2\left(1 - \sqrt{1-(r/R_2)^2}\right), \\ &\approx d_0 - a_2 r^2 - a_4 r^4 - a_6 r^6, \end{aligned} \tag{4}$$

where

$$a_2 = \frac{1}{2}\left(\frac{1}{R_1} - \frac{1}{R_2}\right), \quad a_4 = \frac{1}{8}\left(\frac{1}{R_1^3} - \frac{1}{R_2^3}\right), \quad a_6 = \frac{1}{16}\left(\frac{1}{R_1^5} - \frac{1}{R_2^5}\right),$$

where R_1 a R_2 are radii of the curvature of the lens and d_0 is its axial thickness. Using Eqs. (3) and (4) one gets

$$\begin{aligned} \frac{1}{D(r)}\frac{dD(r)}{dr} &= \frac{3r}{d(r)}\left[\frac{1}{R_2\sqrt{1-(r/R_2)^2}} - \frac{1}{R_1\sqrt{1-(r/R_1)^2}}\right] \\ &\approx -\frac{6(a_2 r + 2a_4 r^3 + 3a_6 r^5)}{d(r)}. \end{aligned} \tag{5}$$

With the use of Eq. (5) and appropriate boundary conditions one can numerically solve Eq. (2). Boundary conditions characterize a way of supporting a lens. For example, in the case of support at the edge (as is shown in Fig. 1) boundary conditions have the form

$$w(r)|_{r=a} = 0, \qquad M(r)|_{r=a} = 0, \tag{6}$$

where a denotes the radius of the lens, and the pushing force $Q(r)$ can be calculated as

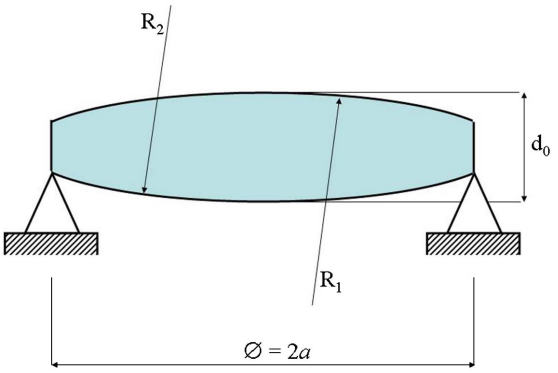


Fig. 1. Scheme of an edge support of a spherical lens.

$$\begin{aligned} Q(r) &= \frac{1}{r}\int_0^r q(r)r\mathrm{d}r = \rho g[I(r) - I(0)] \\ &\approx \rho g(d_0 r/2 - a_2 r^3/4 - a_4 r^5/6 - a_6 r^7/8), \end{aligned} \tag{7}$$

where

$$\begin{aligned} I(r) &= \int \mathrm{d}(r)r\mathrm{d}r = \frac{r^2}{2}(d_0 - R_1 + R_2) \\ &\quad - \frac{R_1\sqrt{1-(r/R_1)^2}}{3}(R_1^2 - r^2) \\ &\quad + \frac{R_2\sqrt{1-(r/R_2)^2}}{3}(R_2^2 - r^2), \\ I(0) &= \frac{R_2^3 - R_1^3}{3}. \end{aligned} \tag{8}$$

In the case of an inner support of the lens on a chosen radius R_2 (see Fig. 2), there will be a discontinuity at the point $r = r_s$, and one has to find a solution in two cases as a function $w_1(r)$ in interval $r \in [0, r_s]$ and as a function $w_2(r)$ in interval $r \in [r_s, a]$. The boundary conditions then become

$$w_1(r)|_{r=r_s} = w_2(r)|_{r=r_s} = 0, \qquad M_2(r)|_{r=a} = 0. \tag{9}$$

Further, conditions of continuity have to be prescribed at the point $r = r_s$; they hold

$$\varphi_1(r)|_{r=r_s} = \varphi_2(r)|_{r=r_s}, \qquad M|_{r=r_s} = M|_{r=r_s}. \tag{10}$$

The pushing force $Q(r)$ will be the same for the interval $r \in [0, r_s]$ as in the case of the edge-supported lens. At the point $r = r_2$, there will be a change as a result of reaction of the support, and therefore the following formulas become valid:

$$\begin{aligned} Q_1(r) &= \frac{1}{r}\int_0^r q(r)r\mathrm{d}r = \frac{\rho g}{r}[I(r) - I(0)], \\ Q_2(r) &= -\frac{1}{r}\int_r^a q(r)r\mathrm{d}r = -\frac{\rho g}{r}[I(a) - I(r)]. \end{aligned} \tag{11}$$

A difference in forces $Q_1(r_s)$ and $Q_2(r_s)$ is equal to the support reaction.

In the following part of this paper, a numerical solution of the problem described by Eq. (2) will be presented.

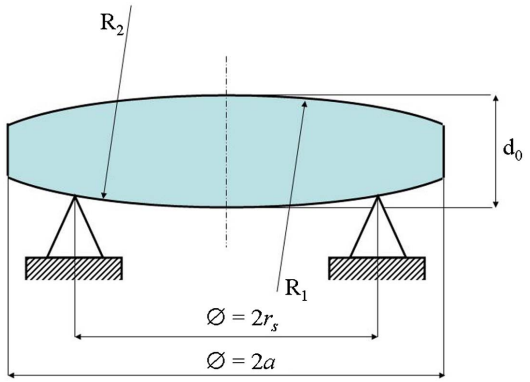


Fig. 2. Scheme of an inner support of a spherical lens.

3. NUMERICAL SOLUTION

This part is going to study several solutions of a calculation of the lens deformation. First, a power series solution is shown, presenting a universal method with transforming the problem to a nonlinear optimization issue. It is followed by the Runge–Kutta method and the finite differences method.

A. Power Series Solution Using Optimization Algorithms

Suppose that a lens is supported at its edge. If a deflection is prescribed as a power series of order $2N$,

$$w = \sum_{i=0}^{N} b_{2i} r^{2i}, \tag{12}$$

where b_{2i} are coefficients of the supposed power series and r is the radius. Then it holds

$$\varphi = -\frac{dw}{dr} = -\sum_{i=0}^{N} (2i) b_{2i} r^{2i-1}, \tag{13}$$

and therefore

$$\frac{d\varphi}{dr} = -\sum_{i=0}^{N} (2i)(2i-1) b_{2i} r^{2(i-1)},$$
$$\frac{d^2\varphi}{dr^2} = -\sum_{i=0}^{N} (4i)(i-1)(2i-1) b_{2i} r^{2(i-3)}. \tag{14}$$

Boundary conditions of Eq. (9) then have the following form:

$$w(r)|_{r=a} = 0 = \sum_{i=0}^{N} b_{2i} a^{2i}, \tag{15}$$

$$M(r)|_{r=a} = 0 = D(a)\left[-\sum_{i=0}^{N} (2i)(2i-1) b_{2i} a^{2(i-1)} - \nu \sum_{i=0}^{N} (2i) b_{2i} a^{2(i-1)}\right]. \tag{16}$$

It is obvious that the following condition has to be valid to fulfill Eq. (16):

$$\sum_{i=0}^{N} (2i)(2i-1+\nu) b_{2i} a^{2(i-1)} = 0. \tag{17}$$

An approximate numerical solution of a problem defined by Eq. (2) can be found with Eqs. (5)–(8), e.g., within an optimization algorithm [25,26].

Suppose now an inner support of the lens. Let a deflection $w_1(r)$ be an even series of order $2N_1$ (considering function $w_1(r)$ symmetrical about axis $x = 0$), and let a deflection $w_2(r)$ be a general power series of order N_2 (e.g., function $w_2(r)$ cannot be approximated with even powers of r only). Then,

$$w_1 = \sum_{i=0}^{N_1} b_{2i} r^{2i}, \qquad w_2 = \sum_{i=0}^{N_2} c_i r^i, \tag{18}$$

where b_{2i} and c_i are coefficients of the supposed power series and r is the radius. Therefore,

$$\varphi_1 = -\frac{dw_1}{dr} = -\sum_{i=0}^{N_1} (2i) b_{2i} r^{2i-1},$$
$$\varphi_2 = -\frac{dw_2}{dr} = -\sum_{i=0}^{N_2} i c_i r^{i-1}, \tag{19}$$

and similarly

$$\frac{d\varphi_1}{dr} = -\sum_{i=0}^{N_1} (2i)(2i-1) b_{2i} r^{2(i-1)},$$
$$\frac{d\varphi_2}{dr} = -\sum_{i=0}^{N_2} i(i-1) c_i r^{i-2}, \tag{20}$$

$$\frac{\mathrm{d}^2\varphi_1}{\mathrm{d}r^2} = -\sum_{i=0}^{N_1} (4i)(i-1)(2i-1) b_{2i} r^{2i-3},$$
$$\frac{d^2\varphi_2}{dr^2} = -\sum_{i=0}^{N_2} i(i-1)(i-2) c_i r^{i-3}. \tag{21}$$

Boundary conditions for the aforementioned situation will have the form

$$w_1(r)|_{r=r_s} = 0 = \sum_{i=0}^{N_1} b_{2i} r_s^{2i} = w_2(r)|_{r=r_s} = \sum_{i=0}^{N_2} c_i r_s^i, \tag{22}$$

$$M(r)|_{r=a} = 0 = -D(a) \sum_{i=0}^{N_2} i(i-1+\nu) c_i a^{i-2}, \tag{23}$$

and conditions of continuity

$$\varphi_1(r)|_{r=r_s} = \varphi_2(r)|_{r=r_s} = -\sum_{i=0}^{N_1} (2i) b_{2i} r_s^{2i-1} = -\sum_{i=0}^{N_2} i c_i r_s^{i-1}, \tag{24}$$

$$M_1(r)|_{r=r_s} = M_2(r)|_{r=r_s}$$
$$= -D(r_s) \sum_{i=0}^{N_1} (2i)(2i-1+\nu) b_{2i} r_s^{2(i-1)}$$
$$= -D(r_s) \sum_{i=0}^{N_2} i(i-1+\nu) c_i r_s^{i-2}. \tag{25}$$

Optimization algorithms [25,26] are a very powerful tool for solving many problems. The main goal of those methods is to minimize a user-defined function of variable unknowns (find function minimum). In the presented case of the solution of differential Eq. (2) the coefficients of the power series of Eq. (12) or Eq. (18) become the unknowns. A goal function to be minimized will be shown below.

Substituting Eqs. (12)–(14) into Eq. (2), one gets

$$\sum_{i=0}^{N} E_i(r) b_{2i} = C(r), \tag{26}$$

where

$$\begin{aligned}
E_i(r) &= e_i(r) + A(r)c_i(r) + B(r)a_i(r),\\
e_i(r) &= -(4i)(i-1)(2i-1)r^{2i-3},\\
c_i(r) &= -(2i)(2i-1)r^{2(i-1)},\\
a_i(r) &= -(2i)r^{2i-1},\\
A(r) &= \left[\frac{1}{r} + \frac{1}{D(r)}\frac{dD(r)}{dr}\right],\\
B(r) &= \frac{1}{r}\left[\frac{\nu}{D(r)}\frac{dD(r)}{dr} - \frac{1}{r}\right], \qquad C(r) = \frac{Q(r)}{D(r)}.
\end{aligned}$$

If coefficients b_{2i} in Eq. (26) are converging to a correct solution, then the formula

$$\phi_1(r_m) = \left[\sum_{i=0}^{N} E_i(r_m) b_{2i}\right] - C(r_m) \tag{27}$$

is converging to zero for given point r_m. Similarly, for boundary conditions Eqs. (15) and (16) one can write

$$\phi_2 = \sum_{i=0}^{N} b_{2i} a^{2i}, \qquad \phi_3 = \sum_{i=0}^{N} (2i)(2i-1+\nu) b_{2i} a^{2(i-1)}. \tag{28}$$

The goal function ϕ for $M+1$ points $r_m (m = 0, \ldots, M)$ in the interval $[0, a] (r_0 = 0, r_M = a)$ defined in a least-squares sense can be formulated as

$$\phi = \left[\sum_{m=0}^{M} W_1 \phi_1^2(r_m)\right] + W_2\phi_2^2 + W_3\phi_3^2, \tag{29}$$

where W_1, W_2, and W_3 are weight coefficients. One can choose the goal function in a sense of a sum of absolute values, and it holds as follows:

$$\phi = \left[\sum_{m=0}^{M} W_1 |\phi_1(r_m)|\right] + W_2|\phi_2| + W_3|\phi_3|. \tag{30}$$

If the algorithm finds such coefficients b_{2i}, which will minimize the goal function of Eq. (29) or Eq. (30), then such coefficients define the approximate solution of Eq. (2) with the boundary conditions Eq. (9).

For the lens with the inner support, the goal function can be defined in a similar approach,

$$\phi = \sum_{m=0}^{M_s} W_1 \phi_1^2(r_m) + \sum_{m=M_s+1}^{M} W_2 \phi_2^2(r_m) + \sum_{j=3}^{8} W_j \phi_j^2, \tag{31}$$

where $m = M_s$ denotes the position of the support in the interval $[0, a]$, $W_1, \ldots, W_8$ are weights, and

$$\begin{aligned}
\phi_1(r_m) &= \left[\sum_{i=0}^{N_1} E_i(r_m) b_{2i}\right] - \frac{Q_1(r_m)}{D(r_m)},\\
\phi_2(r_m) &= \left[\sum_{i=0}^{N_2} F_i(r_m) c_i\right] - \frac{Q_2(r_m)}{D(r_m)},\\
F_i(r_m) &= p_i(r_m) + A(r_m)h_i + B(r_m)q_i,\\
p_i(r_m) &= -i(i-1)(i-2)r^{i-3},\\
h_i(r_m) &= -i(i-1)r^{i-2}, \qquad q_i(r_m) = -ir^{i-1},\\
\phi_3 &= \sum_{i=0}^{N_1} b_{2i} r_s^{2i}, \qquad \phi_4 = \sum_{i=0}^{N_2} c_i r_s^i,\\
\phi_5 &= \sum_{i=0}^{N_2} i(i-1+\nu) c_i a^{i-2},\\
\phi_6 &= \sum_{i=0}^{N_1} (2i) b_{2i} r_s^{2i-1} - \sum_{i=0}^{N_2} i c_i r_s^{i-1},\\
\phi_7 &= \sum_{i=0}^{N_1} (2i)(2i-1+\nu) b_{2i} r_s^{2(i-1)},\\
\phi_8 &= \sum_{i=0}^{N_1} i(i-1+\nu) c_i r_s^{i-2}.
\end{aligned} \tag{32}$$

The goal function as a sum of absolute values can be constructed similarly, as in Eq. (30).

The aforementioned formulas present an approximate solution for function φ. Deflection w then can be easily calculated with, for example, numerical integration [27,28], with respect to Eq. (1),

$$w(r) = -\int_0^r \phi(\xi)\mathrm{d}\xi + c, \tag{33}$$

where c denotes an integration constant, which can be calculated from boundary conditions.

B. Calculation of Deflection with Runge–Kutta Method

One can modify Eq. (2) as follows:

$$\frac{d^2\varphi}{dr^2} = \frac{Q(r)}{D(r)} - \left[\frac{1}{r} + \frac{1}{D(r)}\frac{dD(r)}{dr}\right]\frac{d\varphi}{dr} - \left[\frac{\nu}{D(r)}\frac{dD(r)}{dr} - \frac{1}{r}\right]\frac{\varphi}{r}. \tag{34}$$

Denoting Eq. (34) as $\varphi'' = f(r, \varphi, \varphi')$, one can find an approximate solution for the function φ with the Runge–Kutta Method of the fourth order [28]. Recursion formulas then hold

$$\begin{aligned}
\varphi_{n+1} &= \varphi_n + h\varphi_n' + \frac{h}{6}(k_1 + k_2 + k_3),\\
\varphi_{n+1}' &= \varphi_n' + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4),\\
k_1 &= hf(r_n, \varphi_n, \varphi_n'), \qquad k_2 = hf\left(r_n + \frac{h}{2}, \varphi_n + \frac{h}{2}\varphi_n', \varphi_n' + \frac{k_1}{2}\right),\\
k_3 &= hf\left(r_n + \frac{h}{2}, \varphi_n + \frac{h}{2}\varphi_n' + \frac{h}{4}k_1, \varphi_n' + \frac{k_2}{2}\right),\\
k_4 &= hf\left(r_n + h, \varphi_n + h\varphi_n' + \frac{h}{2}k_2, \varphi_n' + k_3\right),
\end{aligned} \tag{35}$$

where h is a calculation step. The resulting solution has to satisfy the boundary conditions of Eq. (9). It is obvious that one should make an initial guess of starting values in a way to represent the form of expected shape as close a possible.

To find a solution for the internally supported lens at the radius r_s, the aforementioned Eq. (35) can be used without any changes assuming Eqs. (1)–(5) and (7)–(8) for interval $[0, r_s]$, and Eqs. (1)–(5) and the second formula of Eq. (11) for interval $[r_s, a]$ with the appropriate solution of continuity in the point r_s.

C. Calculation of Deflection with Finite Differences Method

The finite differences method [27] is based on expressing derivations of a given function in a given point as a linear combination of function values in surrounding points. In our case, derivations of the function φ can be expressed as

$$\frac{d\varphi}{dr} \approx \frac{\varphi_{i+1} - \varphi_{i-1}}{2h},$$
$$\frac{d^2\varphi}{dr^2} \approx \frac{\varphi_{i-1} - 2\varphi_i + \varphi_{i+1}}{h^2}, \tag{36}$$

where φ_i is the value of the function to be found in given point r_i ($i = 0, \ldots, N$, therefore interval $[0, a]$ is divided into $N + 1$ points), $\varphi_{i\pm1}$ is the function value in the point $r_{i\pm1}$ (see Fig. 3), and $h = r_i - r_{r-1} = r_{i+1} - r_i$ is the distance between calculation points.

Substituting Eq. (36) in Eq. (2) one gets the governing equation for a solution of the aforementioned problem with the use of the finite differences method; it holds

$$\frac{\varphi_{i-1} - 2\varphi_i + \varphi_{i+1}}{h^2} + A(r_i)\frac{\varphi_{i+1} - \varphi_{i-1}}{2h} + B(r_i)\varphi_i = C(r_i), \qquad i = 1, 2, 3\ldots, N, \tag{37}$$

where

$$A(r) = \left[\frac{1}{r} + \frac{1}{D(r)}\frac{dD(r)}{dr}\right],$$
$$B(r) = \frac{1}{r}\left[\frac{\nu}{D(r)}\frac{dD(r)}{dr} - \frac{1}{r}\right],$$
$$C(r) = \left[\frac{Q(r)}{D(r)}\right] \tag{38}$$

could be calculated with Eqs. (5)–(7).

Considering an assumption of a symmetric problem the boundary conditions of Eq. (9) can be extended with

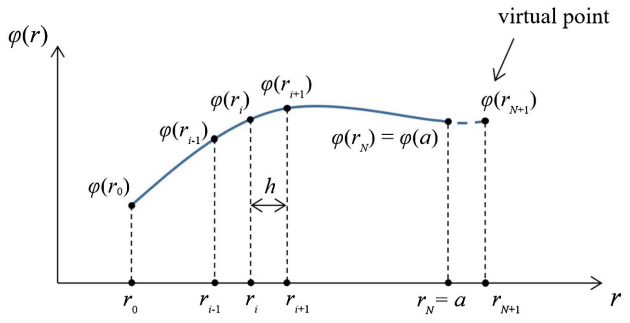


Fig. 3. Method of finite differences.

$$\varphi(r)|_{r=0} = \varphi_0 = 0. \tag{39}$$

Therefore, the value of function $\varphi(r)$ in point $r = 0$ is known and does not have to be assumed as an unknown to be calculated [note $i \neq 0$ in Eq. (37)]. Using Eq. (1) the moment the boundary condition of Eq. (9) has the following form:

$$M(r)|_{r=a} = D(r_N)\left[\frac{\varphi_{N+1} - \varphi_{N-1}}{2h} + \frac{\nu}{r}\varphi_N\right] = 0. \tag{40}$$

Besides the value of φ_{N+1} is not an unknown value in the considered interval, this value states in the last formula, where $i = N$, and therefore it is necessary to calculate it from the boundary condition (40) as

$$\varphi_{N+1} = -\frac{2h\nu}{r_N}\varphi_N + \varphi_{N-1}, \tag{41}$$

and, consequently, it has to be substituted into the last formula.

For an illustration let one formulate the system of Eq. (37) using matrix notation. The values of function φ in discrete points of the interval $(0, a]$ are calculated from the system of equations

$$\mathbf{Af} = \mathbf{C}, \tag{42}$$

where $\mathbf{A} \in \mathbb{R}^{N\times N}$ is a tridiagonal matrix of the system whose elements δ_i, ($i = 1, 2, \ldots, N$) on the main diagonal can be expressed as follows:

$$\delta_i|_{i<N} = -\frac{2}{h^2} + B(r_i),$$
$$\delta_i|_{i=N} = -\frac{2}{h^2} - \frac{2\nu}{r_N h} - A(r_N)\frac{\nu}{r_N} + B(r_N). \tag{43}$$

The upper collateral diagonal elements $\alpha_i (i = 1, 2, \ldots, N - 1)$ are then given with formula

$$\alpha_i = \frac{1}{h^2} + \frac{A(r_i)}{2h}, \tag{44}$$

and the lower collateral diagonal elements $\beta_i (i = 2, \ldots N)$ as

$$\beta_i|_{i<N} = \frac{1}{h^2} - \frac{A(r_i)}{2h}, \qquad \beta_i|_{i=N} = \frac{2}{h^2}. \tag{45}$$

Next, $\mathbf{f} \in \mathbb{R}^{N\times1}$ in Eq. (42) denotes a column matrix of values $\varphi_i (i = 1, \ldots, N)$ to be found, and $\mathbf{C} \in \mathbb{R}^{N\times1}$ is a matrix with elements $C(r_i) (i = 1, \ldots, N)$.

Similarly to the previous approaches, the aforementioned formulas present an approximate solution for function φ. Deflection w then can be easily calculated with, for example, numerical integration [27,28], with respect to Eq. (1).

4. EXAMPLE

In this section, first, all the aforementioned methods are compared within each other and to the finite elements method used in a commercially available software ANSYS [29], too. Afterward, a dependency of the deflection on material parameters of the lens is presented in the second subsection.

A. Comparison of Presented Numerical Solutions

To compare the presented numerical approaches, suppose a lens with the following parameters: $R_1 = 600$ mm, $R_2 = -600$ mm, $d_0 = 5$ mm, $a = 50$ mm, $E = 82$ GPa, $\rho = 2510$ kg/m^3, $\nu = 0.206$. Results of the approximate solution

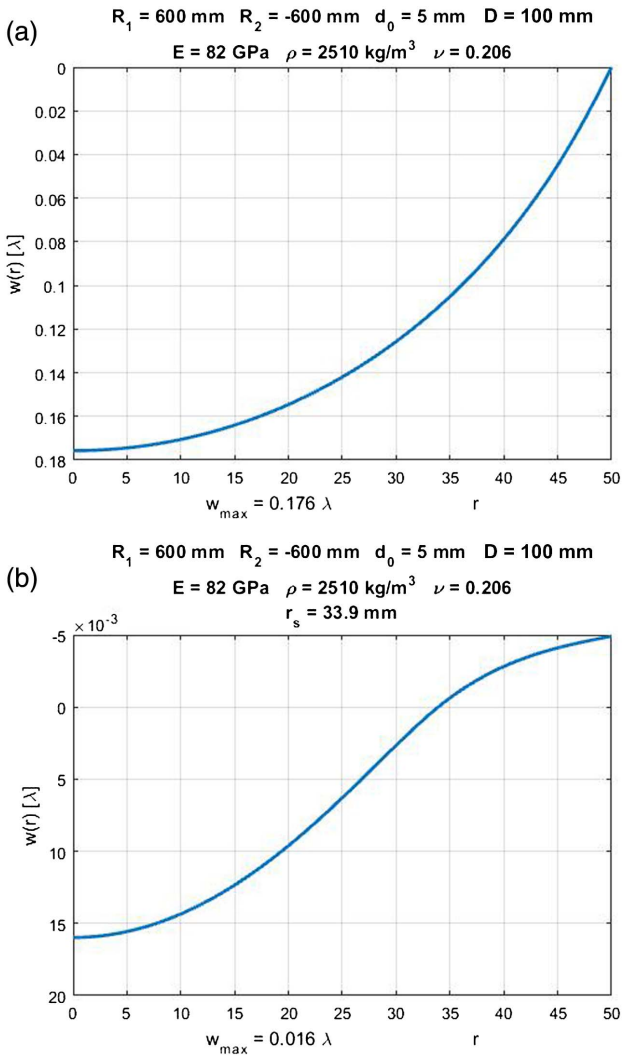


Fig. 4. Example of calculation of lens deflection.

presented in the previous part of the paper are shown in Fig. 4. Figure 4(a) shows the result for a lens supported at the edge, and Fig. 4(b) shows the result for a support in radial distance $r_s = 33.9$ mm. Please note that the scale of the second figure is in 10^{-3}. It demonstrates a large difference between the deflection for edge and inner support. The figure shows only one result for all presented methods, because differences between chosen numerical approaches are negligible. Besides the presented method the finite elements method was used with the software ANSYS giving the same results. Note that the interval h of the division in the radial direction was chosen, $h = 0.05$ mm.

B. Dependency of Deflections on a Material of a Lens

Let one study the effect of a material of a lens on deflections. First, suppose a positive (biconvex) lens with the following geometrical parameters: $R_1 = 300$ mm, $R_2 = -300$ mm, $d_0 = 10$ mm, $a = 50$ mm. Material parameters chosen for a calculation are shown in Table 1. In the table, E denotes the Young modulus, ρ is the volume density, and ν is the Poisson ratio. The deflections of the lens for support at the edge are shown in Fig. 5(a), and results for the inner support at the radius $r_s = 33.9$ mm $= 0.678a$ are shown in Fig. 5(b). Note that the curves for materials N-BK7 and N-LAF21 coincide in this example despite the fact that materials are different. As is obvious, the maximal deflection for the edge support is $w_{\max} = \lambda/7$ (where $\lambda = 633$ nm) for the material SF57, and the maximal deflection for the inner support is $w_{\max} = \lambda/78$.

Table 1. Material Parameters of a Lens

Glass Type	E [GPa]	ρ [g/cm^3]	ν
SF57	54	5.51	0.248
N-BK7	82	2.51	0.206
N-LAF21	124	4.28	0.295

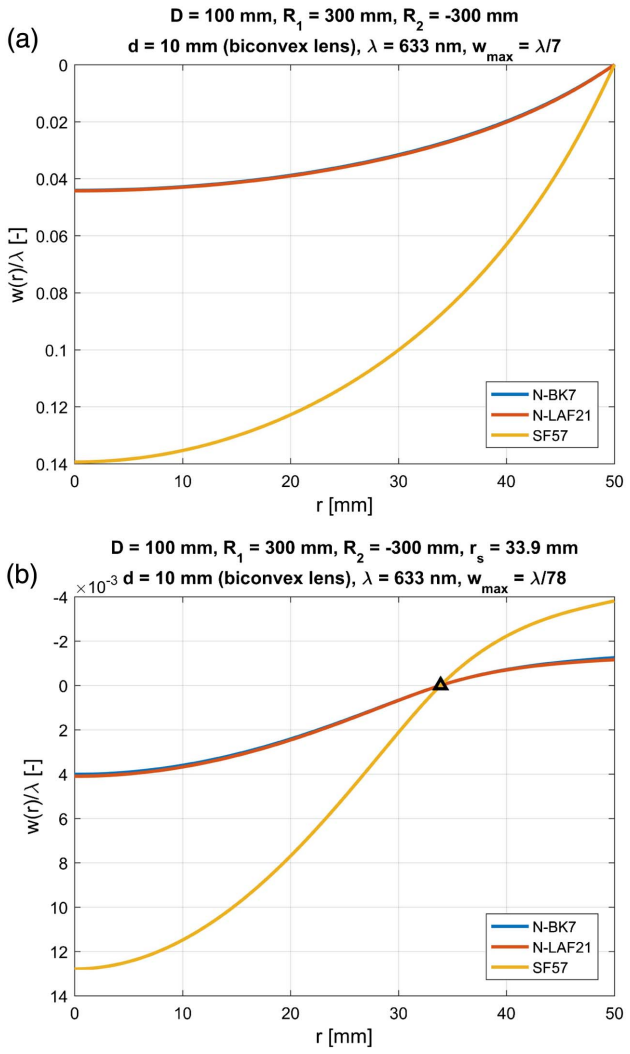


Fig. 5. (a) Comparison of deflections of a positive (biconvex) lens for different materials and support at the edge (related to the wavelength $\lambda = 633$ nm), and (b) comparison of deflections of a positive (biconvex) inner-supported lens for different materials (related to the wavelength $\lambda = 633$ nm). Note that the curves for materials N-BK7 and N-LAF21 coincide in this example despite the fact that materials are different.

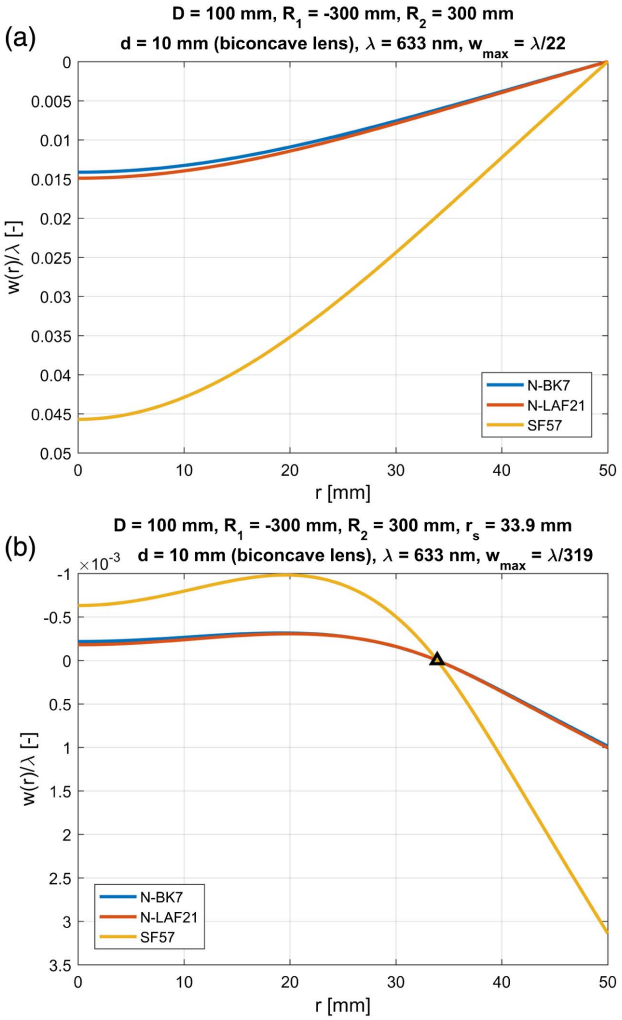


Fig. 6. (a) Comparison of deflections of a negative (biconcave) lens for different materials and support at the edge (related to the wavelength $\lambda = 633$ nm), and (b) comparison of deflections of a negative (biconcave) inner-supported lens for different materials (related to the wavelength $\lambda = 633$ nm).

One can see that the self-deflection could result in non-negligible errors in combination with inappropriate support of the lens under test.

Second, a similar analysis is done for a negative (biconcave) lens with the following geometrical parameters: $R_1 = -300$ mm, $R_2 = 300$ mm, $d_0 = 10$ mm, $a = 50$ mm. The lens is again made by three different materials, which are the same as in the previous example (see Table 1). Results for the edge support of the lens are shown in Fig. 6(a) and for the inner support in Fig. 6(b).

As is obvious from the presented results of numerical examples, the self-deflection caused by the lens's weight cannot be neglected for the case of measurements with high accuracy, e.g., interferometry. Especially, with the use of inappropriate material the deflection can enormously affect measured results.

5. CONCLUSION

This paper presented an analysis of a numerical solution of a differential equation characterizing a deflection of a spherical lens caused by its own weight, based on a chosen support and lens material. Several calculation methods were shown and formulas for an easy implementation were presented.

Deflection of lenses depends on geometrical and material parameters of the tested component. One of the examples presented in this paper shows a case study for the positive lens with diameter of 100 mm, edge thickness of 10 mm, and radii of curvatures equal to 300 mm. The deflection for this case, considering a material SF57, was approximately $\lambda/7$ for the edge support. Such a value is easily measurable during interferometrical testing. Therefore, if one does not include such a phenomenon into the correction of measurement results, it can lead to a wrong evaluation of optical quality of the final product. In the presented case, if a production tolerance would be $\lambda/10$, for example, the self-deflection effect could lead to an elimination of a component that should remain in the production process.

The reader can use the derived formulas and calculation procedures presented in this paper for solving specific problems in practice that can vary for different geometries of tested components. Especially during measurements with high accuracy, e.g., interferometrical testing of large lenses, it is very important to consider self-deflection of tested parts.

Generally, there are other effects that affect the testing procedure, e.g., variation in temperature, which can affect atmospheric condition as well as elongation of component itself. Usually, conditions for testing of a high accuracy are of such quality that these phenomena are eliminated (temperature stabilization, etc.). Nevertheless, the self-deflection cannot be practically removed; therefore, one has to use inner support for the tested component or numerically correct final results.

Funding. Czech Technical University in Prague (CTU). (SGS17/004/OHK1/1T/11).

REFERENCES

1. D. Malacara, *Optical Shop Testing*, Wiley Series in Pure and Applied Optics (Wiley, 2007).
2. B. Dörband and H. Gross, *Handbook of Optical Systems, Metrology of Optical Components and Systems*, Gross/Optical Systems V1-V6 special prices until 6 V ST published (Wiley, 2012).
3. R. Leach, *Optical Measurement of Surface Topography* (Springer Berlin Heidelberg, 2011).
4. T. Yoshizawa, *Handbook of Optical Metrology: Principles and Applications*, 2nd ed. (CRC Press, 2015).
5. A. Mikš, *Applied Optics* (CTU Publishing, 2009).
6. P. Hariharan, *Optical Interferometry*, Electronics & Electrical (Academic, 2003).
7. http://www.zygo.com.
8. http://www.trioptic.com.
9. http://www.4dtechnology.com.
10. http://www.optotech.de.
11. http://www.schneider-om.com.
12. A. Mikš, "Interferometrical methods for evaluation spherical surfaces in optics," Fine Mech. Opt. **46**, 29–35 (2001).
13. A. Nadai, "Die verbiegungen in einzelnen punkten unterstiitzter kreisfiirmiger platten," Physik **23**, 366–376 (1922).
14. A. Nadai, "Theorie der plattenbiegung und ihre experimentelle best tigung," Z. Angew. Math. Mech. **2**, 381–398 (1922).
15. W. B. Emerson, "Determination of planeness and bending of optical flats," J. Res. Natl. Bur. Stand. **49**, 241–247 (1952).
16. L. A. Selke, "Theoretical elastic deflections of a thick horizontal circular mirror on a ring support," Appl. Opt. **9**, 149–153 (1970).

17. J. E. Nelson, J. Lubliner, and T. S. Mast, "Telescope mirror supports: plate deflections on point supports," Proc. SPIE **322**, 212–228 (1982).
18. L. Nong and F. Bao-lian, "The symmetrical bending of an elastic circular plate supported at K internal points," Appl. Math. Mech. **12**, 1091–1096 (1991).
19. L. Arnold, "Uniform-load and actuator influence functions of a thin or thick annular mirror: application to active mirror support optimization," Appl. Opt. **35**, 1095–1106 (1996).
20. A. Mikš and J. Novák, "Influence of the gravity induced deflection of a reference flat of the Fizeau interferometer on the measurement accuracy of testing flat surfaces in optics," Fine Mech. Opt. **59**, 228–232 (2014).
21. S. Timoshenko and S. Woinowsky-Krieger, *Theory of Plates and Shells*, Engineering Societies Monographs (McGraw-Hill, 1959).
22. A. Volmir, *Flexible Plates and Shells* (Defense Technical Information Center, 1967).
23. E. Ventsel and T. Krauthammer, *Thin Plates and Shells: Theory, Analysis, and Applications* (CRC Press, 2001).
24. R. Szilard, *Theories and Applications of Plate Analysis: Classical Numerical and Engineering Methods* (Wiley, 2004).
25. L. E. Scales, *Introduction to Non-Linear Optimization*, Macmillan Computer Science Series (Springer, 1985).
26. M. Aoki, *Introduction to Optimization Techniques: Fundamentals and Applications of Nonlinear Programming*, Macmillan Series in Applied Computer Science (Macmillan, 1971).
27. J. van Kan, A. Segal, and F. Vermolen, *Numerical Methods in Scientific Computing* (Delft Academic, 2014).
28. G. Korn and T. Korn, *Mathematical Handbook for Scientists and Engineers: Definitions, Theorems, and Formulas for Reference and Review*, Dover Civil and Mechanical Engineering (Dover, 2013).
29. http://www.ansys.com.

applied optics

Simple method for determination of parameters of cemented doublet

ANTONÍN MIKŠ AND PETR POKORNÝ*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thákurova 7, 166 29 Prague 6, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 20 May 2016; revised 12 June 2016; accepted 13 June 2016; posted 14 June 2016 (Doc. ID 266558); published 8 July 2016

This paper proposes a simple noninvasive method that makes it possible to calculate the inner design parameters of the cemented doublet using measurements of its chosen paraxial optical and geometrical parameters without any damage to the system under testing (e.g., dismantling). Derived formulas are based on the knowledge of measured values of the lenses thicknesses, the radii of curvatures of the first and the last doublet's surfaces, the paraxial focal length, and positions of the object and the image focal point. Practical usefulness of the proposed method is demonstrated on the real measurement of a known doublet.

OCIS codes: (080.3630) Lenses; (120.3940) Metrology; (120.4630) Optical inspection; (220.4840) Testing.

http://dx.doi.org/10.1364/AO.55.005456

1. INTRODUCTION

In practical industrial applications the issue of determining the internal parameters of the optical system (e.g., photographic objective, etc.) without any damage to its components is in focus very often. The optical systems are usually composed by individual simple lenses or cemented doublets. A determination of parameters of simple lenses is very well known and it is an easily solvable problem, because one is able to measure the radii of curvatures of the lens' surfaces and the central thickness. Afterwards, the index of refraction can be calculated from the known formula for the optical power of the lens. A much more complicated issue comes to role with testing the cemented doublet if one wants to determine its internal parameters without any damage to the component, which is advantageous for practical purposes.

A cemented doublet belongs to frequently used optical systems in practice [1–6]. The doublet is composed of two spherical lenses, where the second radius of the first lens is identical to the first radius of the second lens, and both lenses are cemented together. In practice, it is possible to measure external parameters and aberrations relatively easily by various measurement techniques [7–14]. However, the internal parameters cannot be measured directly.

The authors' previous work [15] describes the possibility of determination of internal parameters of a cemented doublet, which is based on the measurement of paraxial parameters and wave aberration for a point on the optical axis of the doublet (spherical aberration). With regard to the fact that it is not possible to obtain a simple analytical solution for a determination of an unknown doublet's internal parameters, the optimization technique has to be used [16,17].

The aim of this work is to propose a novel and simple noninvasive method that makes it possible to obtain the internal design parameters of the cemented doublet using measurements of some paraxial optical and geometrical parameters of the doublet. The simplification of the previous procedure [15] is based on the measurement of internal thicknesses of the individual lenses, which can be by available commercial measurement devices, e.g., OptiCentric by the company Trioptics [13]. Such an approach provides a possibility to derive simple and explicit analytical formulas for the calculation of the remaining internal parameters only with the knowledge of the focal length, position of the object focal point, and position of the image focal point. Therefore, it is not needed to measure the wave aberration and calculate the internal parameters with complex optimization techniques, as it was necessary in the authors' previous paper [15].

2. DETERMINATION OF PARAMETERS OF CEMENTED DOUBLET

A scheme of the cemented doublet is shown in Fig. 1. Such an optical system has seven design parameters, namely three radii of curvature (r_1, r_2, r_3), two values of central thickness (d_1, d_2), and two values of refractive index (n_1, n_2) of individual lenses, from which the doublet is composed by. In Fig. 1, F and F' denote object and image focal point of the doublet, n_0 and n_3 are refractive indices of the object and image media, V_1 and V_2 are vertices of the doublet laying on the optical axis, s_F and $s'_{F'}$ are distances of the object and the image focal points from the vertices.

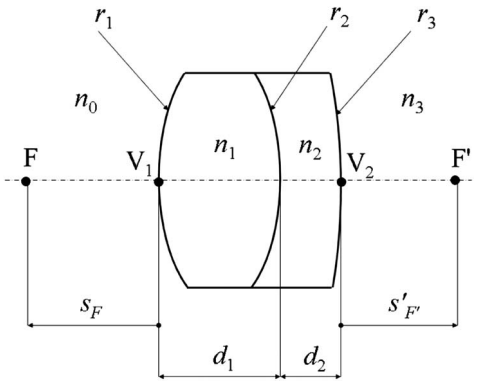


Fig. 1. Optical scheme of cemented doublet.

Using equations for paraxial ray tracing [1–6] one can derive the following formulas for optical power φ, position of the object focal point s_F, and position of the image focal point $s'_{F'}$ of the doublet. It holds:

$$\varphi = \varphi_1 + \varphi_2 + \varphi_3 - D_1(\varphi_1\varphi_2 + \varphi_1\varphi_3) - D_2(\varphi_1\varphi_3 + \varphi_2\varphi_3) + D_1D_2\varphi_1\varphi_2\varphi_3, \quad (1)$$

$$s_F = n_0(D_1\varphi_2 + D_1\varphi_3 + D_2\varphi_3 - D_1D_2\varphi_2\varphi_3 - 1)/\varphi, \quad (2)$$

$$s'_{F'} = -n_3(D_1\varphi_1 + D_2\varphi_1 + D_2\varphi_2 - D_1D_2\varphi_1\varphi_2 - 1)/\varphi, \quad (3)$$

$$s' = \frac{n_0n_3(D_1 + D_2 - D_1D_2\varphi_2) - ss'_{F'}\varphi}{\varphi(s_F - s)},$$

where s is the distance of the object from the first surface of the doublet, s' is the distance of the image from the last surface of the doublet. Optical powers $\varphi_1, \varphi_2, \varphi_3$ of individual refractive surfaces are given by the following formulas:

$$\varphi_1 = (n_1 - n_0)/r_1, \qquad \varphi_2 = (n_2 - n_1)/r_2,$$
$$\varphi_3 = (n_3 - n_2)/r_3,$$

and it holds for the reduced values of thickness D_1 and D_2:

$$D_1 = d_1/n_1, \qquad D_2 = d_2/n_2.$$

As one can see from Fig. 1, it is possible to measure relatively easily the radii of curvature r_1 and r_3, overall central thickness of the doublet $d = d_1 + d_2$, positions of the object and image focal points s_F and $s'_{F'}$, and focal length $f' = n_3/\varphi$ between the second principal plane and the image focal point F' of the doublet using the methods given, for example, in [7–15]. Next, it is possible to measure thicknesses d_1 and d_2 of the individual lenses using available commercial devices, e.g., Trioptics' OptiCentric [13].

Suppose now that the doublet is in the air, i.e., it holds for the object and image refractive indices $n_0 = 1$, $n_3 = 1$. Such a situation is the most common in practical situations. As was already stated in the introduction, the aim of this work is to determine the internal parameters n_1, n_2, and r_2 of the cemented doublet if one is able to measure radius r_1 of its first surface, radius r_3 of its last surface, thickness d_1 of the first lens and d_2 of the second lens, the focal length $f' = 1/\varphi$ (where φ is the optical power of the doublet when the surrounding media is air), position s_F of the object focal point, and position $s'_{F'}$ of the image focal point. The remaining parameters (n_1, n_2, and r_2) then can be calculated as follows.

Solution of the system of Eqs. (1)–(3) gives the following formula for the index of refraction n_2 of the second lens:

$$n_2 = \frac{d_2 + \varphi^2 d_2(s_Fs'_{F'} + r_1r_3 - r_3s_F - r_1s'_{F'})}{d_2 + r_3 - \varphi r_3(d_1 - r_1) + \varphi^2 s'_{F'}(d_2s_F + r_3s_F - d_2r_1 - r_1r_3)}. \quad (4)$$

Afterwards, the index of refraction n_1 of the first lens can be calculated as follows:

$$n_1 = \frac{\varphi d_1 n_2 r_3(r_1 - s_F)}{r_1(d_2n_2 - d_2 + n_2r_3) - \varphi n_2r_3s_F(d_1 - r_1)}. \quad (5)$$

Finally, one can find the value of the second radius r_2 of the doublet with one of the following formulas:

$$r_2 = -\frac{d_1\left[\frac{d_2(n_2-1)}{n_2r_3} + 1\right](n_1 - n_2)}{n_1\left[\varphi s_F + \frac{d_1(n_2-1)}{n_1r_3} + \frac{d_2(n_2-1)}{n_2r_3} + 1\right]}, \quad (6)$$

$$r_2 = -\frac{d_2\left[\frac{d_1(n_1-1)}{n_1r_1} - 1\right](n_1 - n_2)}{n_2\left[\varphi s'_{F'} + \frac{d_1(n_1-1)}{n_1r_1} + \frac{d_2(n_1-1)}{n_2r_1} - 1\right]}. \quad (7)$$

In summary, one should follow the steps below for the calculation of inner parameters of the doublet:

1. The index of refraction n_2 of the second lens is calculated with Eq. (4).
2. Such a value is used in Eq. (5) and the index of refraction n_1 of the first lens can be calculated.
3. Using the values of indices of refraction in Eq. (6) or Eq. (7) gives the value of radius r_2 of the doublet.

The issue of inner parameters determination is therefore solved.

3. EXAMPLE

In this example the authors will present the proposed novel method of the calculation of parameters of the doublet in the case when the refractive indices n_1, n_2 and radius r_2 of the doublet are not known. As a comparison with the authors' previous method [15], the same cemented doublet will be used.

Consider a cemented doublet having the following nominal parameters: $r_1 = 57.008$ mm, $r_2 = -40.738$ mm, $r_3 = -173.786$ mm, $d_1 = 8$ mm, $d_2 = 4$ mm, $n_1 = 1.51874$ (Schott N-BK7), $n_2 = 1.62409$ (Schott F2) for the wavelength $\lambda = 546$ nm. Further, one obtains $f' = 100.029$ mm, $s_F = -98.603$ mm, $s'_{F'} = 93.599$ mm.

Measurements of individual parameters of the doublets were carried out in laboratories of Meopta-optika company [18]. Measurements were performed using several measuring instruments in order to obtain a higher reliability. Measurements of parameters f', s_F, and $s'_{F'}$ were carried out using OTS 200 from OEG-Messtechnik, OptiCentric MOT 2R from Trioptics, and the interferometer OWI 150 XT from Optotech. Measurements of the radii of curvatures were performed using the interferometer OWI 150 XT from

Table 1. Comparison of Nominal and Calculated Values of Inner Parameters of Cemented Doublet

	n_1	n_2	r_2[mm]
Nominal Values	1.51874	1.62409	–40.738
Calculated Values	1.51856	1.62373	–40.743

Optotech and the interferometer Zygo Verifire ATZ from ZYGO. The following values for individual parameters were obtained: $r_{1m} = 57.036$ mm, $r_{3m} = -174.068$ mm, $s'_{F'm} = 93.696$ mm, $s_{Fm} = -98.704$ mm, $f'_m = 100.128$ mm, $d_{1m} = 8.001$ mm, and $d_{2m} = 3.998$ mm. The accuracy of the measured parameters was as follows. The accuracy of the measurements of parameters f', s_F, and $s'_{F'}$ was $\pm 0.03\%$, the accuracy of the measurement of the thickness (d_1, d_2) was ± 0.001 mm with the use of the device Trioptics OptiCentric, and the accuracy of the radii of curvature measurement was ± 0.003 mm. Using the measured values and Eqs. (4)–(7) one can obtain the unknown parameters (n_1, n_2, r_2).

The nominal and resulting calculated values of the internal doublet parameters are given in Table 1. As one can see, the differences of the calculated parameters from the nominal parameters are very small, and the proposed method is efficient for the described problem of the determination of the inner parameters of the doublet in practice.

4. CONCLUSION

The paper presented novel and simple explicit formulas for the calculation of the inner parameters of a cemented doublet (n_1, n_2, r_2) based only on the knowledge of the first and back radii of curvature (r_1, r_3), central thickness (d_1, d_2), and the doublet's paraxial parameters, i.e., focal lengths f', position of the object focal point s_F, and position of the image focal point $s'_{F'}$. It was shown with the example of a real doublet that the proposed method is very efficient for the described issue of the determination of the inner doublet's parameters and that the method can find practical usage.

Funding. Grant Agency of the Czech Technical University in Prague (SGS15/125/OHK1/2T/11).

Acknowledgment. The authors would like to thank Dr. Ševčík from Meopta-optika company for the performed measurements of doublets.

REFERENCES

1. M. Herzberger, *Modern Geometrical Optics* (Interscience, 1958).
2. A. Mikš, *Applied Optics* (Czech Technical University, 2009).
3. M. Berek, *Grundlagen der Praktischen Optik* (Walter de Gruyter & Co., 1970).
4. M. Born and E. Wolf, *Principles of Optics* (Oxford University, 1964).
5. W. T. Welford, *Aberrations of the Symmetrical Optical Systems* (Academic, 1974).
6. A. Miks, J. Novak, and P. Novak, "Generalized refractive tunable-focus lens and its imaging characteristics," Opt. Express **18**, 9034–9047 (2010).
7. D. Malacara, *Optical Shop Testing* (Wiley, 2007).
8. G. V. Kreopalova, N. L. Lazareva, and D. T. Puriajev, *Optical Measurements* (Maschinostroenie, 1987).
9. J. Picht, *Meß- und Prüfmethoden der optischen Fertigung* (Akademie-Verlag, 1953).
10. J. Flügge, *Einführung in die Messung der optischen Grundgrössen* (Verlag Braun, 1954).
11. B. Dorband, H. Miller, and H. Gross, *Handbook of Optical Systems*, Vol. **5** of Metrology of Optical Components and Systems (Wiley, 2012).
12. http://www.zygo.com/.
13. http://www.trioptics.com/.
14. http://www.oeg-messtechnik.de/.
15. A. Miks and J. Novak, "Experimental method of determination of parameters of cemented doublet," Appl. Opt. **54**, 7940–7943 (2015).
16. L. E. Scales, *Introduction to Non-linear Optimization* (Springer, 1985).
17. E. M. T. Hendrix and B. G. Toth, *Introduction to Nonlinear and Global Optimization* (Springer, 2010).
18. http://www.meopta.com/en/.

applied optics

Use of diffraction grating for measuring the focal length and distortion of optical systems

ANTONIN MIKS AND PETR POKORNY*

Czech Technical University in Prague, Faculty of Civil Engineering, Department of Physics, Thakurova 7, 166 29 Prague 6, Czech Republic
**Corresponding author: petr.pokorny@fsv.cvut.cz*

Received 27 August 2015; revised 4 November 2015; accepted 4 November 2015; posted 4 November 2015 (Doc. ID 248855); published 30 November 2015

The paper presents an experimentally simple, accurate, and inexpensive method for measuring the focal length and distortion of optical systems using a diffraction grating, where both of the properties are determined from the transversal distances of diffraction maximums in one measurement. The proposed approach does not require any special components or any expensive equipment. A detailed theoretical analysis is performed, and the estimation of uncertainties is studied as well. Afterward, the method is demonstrated with a computer simulation and experimental measurement, and compared with commercially available measurement devices. It is shown that the method provides sufficiently accurate results for many practical applications; therefore, it is appropriate for laboratory testing and for industrial applications.

OCIS codes: (050.1950) Diffraction gratings; (120.3940) Metrology; (120.4630) Optical inspection; (120.4820) Optical systems; (220.4840) Testing.

http://dx.doi.org/10.1364/AO.54.010200

1. INTRODUCTION

Focal length [1–3] is a fundamental characteristic of optical systems, and it is very important to be able to measure its value for a given optical system. The current state of the art is described in many papers [4–31]. The methods differ in the complexity of their experimental equipment and the achievable precision of the measured focal length. The error in these measurements ranges from a fraction of a percent to a few percent [4–31].

Classical methods of measuring the focal length using conjugate distance equations are given in [5–14]. Usage of two diffraction gratings for the measurement of the focal length was presented in [15–22]. Such double gratings behind the lens under test generate a moiré effect, and the focal length is calculated from the evaluation of the moiré pattern. The methods' accuracy is in the range from 0.15% to 0.05% for long focal lengths [22]. The method proposed in this paper requires only one grating, and the process of evaluation is much simpler, as will be presented in the following parts.

The imaging quality of the optical system depends on the residual aberrations of the system [1–3]. The fidelity of the geometrical parameters of the image created by the optical system is characterized by an optical distortion [32]. For a chosen object point the image is not a point; it is formed by an energy distribution—a point spread function [1,3]. Afterward, the distortion is defined as a distance from the maximum of the point spread function to a paraxial image of the point [32]. The distortion of the optical system can be measured by many methods [6–11], which are characterized by varying degrees of accuracy and experimental demands.

In the usual practice (e.g., in photography or cinematography), the focal length does not have to be known with a high accuracy. Therefore, it is useful to use some simple measurement method for its determination.

The aim of this work is to present a method for a measurement of the focal length and the distortion of the optical system with the use of a diffraction grating. To the authors' best knowledge, this simple and accurate method has not been published yet; therefore, it can expand the range of existing methods, and it can find utilization in many laboratory and industrial applications.

2. METHOD DESCRIPTION

In Fig. 1 the principal scheme of the measurement method is shown. A point source (A) is placed at the focal plane of a collimating lens (CL). The collimated beam is then directed toward a linear transmissive amplitude grating (G) where diffraction occurs. Such a phenomenon is very well known, and it can be described with the following formula [1]:

$$\sin \alpha_m = m\lambda/a = m\lambda N, \qquad m = 0, \pm 1, \pm 2, \ldots \quad (1)$$

In Eq. (1) a denotes the grating's period, m is the diffraction order, α_m is the diffraction angle of the mth diffraction order (the angle between the diffracted beam and the normal of the grating), λ is the wavelength of the used monochromatic light,

1559-128X/15/3410200-07$15/0$15.00

and N denotes the spatial frequency of the grating—that is, the number of grooves per unit of length. For simplicity and readability, only the zero (0), the plus first (+1), and the minus first (−1) diffraction orders are shown in Fig. 1.

The measured optical system is placed behind the diffraction grating, and the outgoing diffracted beams are focused in points $A^{(0)}$, $A^{(+1)}$, and $A^{(-1)}$. As was already mentioned in Section 1, the image of the point A will not be a point, but the energy distribution—the point spread function [1,3]—will be formed. In the sense of geometrical optics one can characterize the energy distribution with a circle of confusion—a spot diagram [2,32].

Let one suppose that the parallel beam of rays impinges the optical system under an angle of incidence α. It is known from the theory of geometrical optics [1–3] that for the optical system without any aberration such a beam will be focused on a point which is at a distance y from the optical axis, and the following formula holds for such a distance and the optical system in air [1–3]:

$$y = -f' \tan \alpha. \tag{2}$$

In Eq. (2) f' is the focal length of the optical system. If one measures the distance y and the angle α is known, the focal length f' of the optical system can be easily calculated with the use of Eq. (2).

In a general situation, the beam does not have to impinge on the diffraction grating perpendicularly. Suppose that the parallel beam is impinging on the diffraction grating G under an angle β measured from the grating's normal. Consequently, one gets [1]

$$a(\sin \alpha_m - \sin \beta) = m\lambda, \qquad m = 0, \pm 1, \pm 2, \ldots \tag{3}$$

Afterward, a diffraction angle α_m of the mth diffraction order can be calculated from Eq. (3) as follows:

$$\sin \alpha_m = \sin \beta + m\lambda/a. \tag{4}$$

Let one denotes an angle between the diffracted beam and the z axis as α_{mz}. With the use of Eq. (4) one gets the following formula for such an angle; it holds

$$\alpha_{mz} = \arcsin(\sin \beta + m\lambda/a) - \beta. \tag{5}$$

The maximal angle of the grating's rotation $\beta^m_{\max}$ is given according to Eq. (5) as follows:

$$\beta^m_{\max} = \arcsin(1 - m\lambda/a). \tag{6}$$

Figure 2 shows the dependency between the angles α_{mz} and β for the diffraction grating with the spatial frequency $N = 200$ lines/mm; the diffraction orders $m = 1$, $m = 2$, $m = 3$, and $m = 4$; and the wavelength $\lambda = 587$ nm. It is obvious that with the rotation of the diffraction grating by the angle β around the x axis one can continuously change the angle α_{mz} of the field of view of the optical system.

It is gradually arising from the aforementioned statements that if one measures the distance $y = y^{(+1)} = \overline{A^{(0)}A^{(+1)}}$ (or $y = y^{(-1)} = \overline{A^{(0)}A^{(-1)}}$) (see Fig. 1) for $\beta = 0$ and $m = 1$, the focal length f' of the optical system can be calculated from the following formula:

$$f' = -y^{(+1)}/\tan \alpha_{1z} = -y^{(+1)}\sqrt{a^2 - \lambda^2}/\lambda. \tag{7}$$

It is obvious from Eqs. (5) and (6) and Fig. 2 that the change in the angle β changes the distance $y^{(m)}_\beta$ of the energy centrum of the point spread function of the optical system. Therefore, if one measures the distances $y^{(m)}_\beta$ for different angles β of the grating's rotation, for example, for $m = 1$ and $m = 2$, both the distortion and the focal length of the optical system can be simply calculated. Let one denote $y^{(m)}_{0\beta} = -f' \tan \alpha_{mz}$ as a "paraxial image distance" of the point A from the point $A^{(0)}$. Afterward, the distortion δy of the optical system for the given paraxial image distance can be calculated as follows:

$$\delta y = y^{(m)}_\beta - y^{(m)}_{0\beta}. \tag{8}$$

The given issue of the measurement of the focal length and the distortion of the optical system is therefore simply solved.

Note that the same effect can be reached without the rotation of the diffraction grating by using several diffraction gratings with different periods a (with different spatial frequency N). The mechanical construction of such a measurement device with several gratings will be simpler; the angle of rotation does not have to be measured at all. During the last decades gratings with a variable frequency [33–38] and liquid-crystals-on-silicon based spatial light modulators [39,40] have been

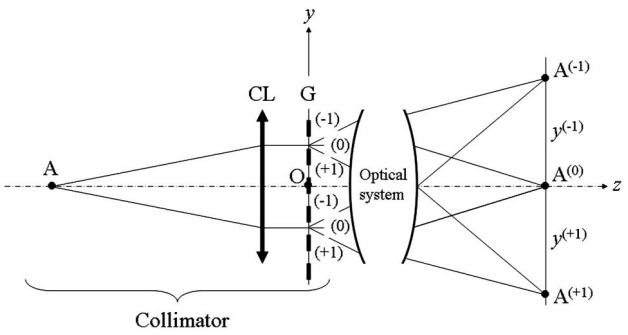


Fig. 1. Principal scheme of measurement device (A, point source; CL, collimating lens; G, diffraction grating; $A^{(m)}$, image of mth diffraction order). The focal length and distortion of an optical system can be calculated from transversal distances between diffraction maximums and the zero-order maximum.

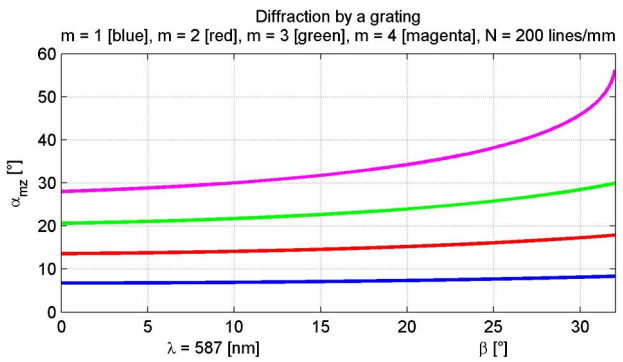


Fig. 2. Dependency between the angle α_{mz} of a field of view of the optical system and the rotation angle β of the diffraction grating for $N = 200$ lines/mm, $m = 1$ (blue line), $m = 2$ (red line), $m = 3$ (green line), and $m = 4$ (magenta line).

developed. However, the quality and temporary stability of such commercially available devices or generated structures do not fulfil the demands for the presented application. Another disadvantage of the usage of such devices is the additional electronics for generation of variable structures. From this point of view the classical diffraction gratings are still the best option and the most suitable solution for practical applications.

For the elimination of the aberrations of the optical system on the measured values, the F-number of the tested system should be set to a high value, for example, $F = 10$.

The measurement of the focal length of the optical system will proceed as follows:

1. The collimator with the diffraction grating without rotation ($\beta = 0$) is placed in front of the tested optical system, and the collimator's optical axis is aligned to be parallel with the optical axis of the tested system.
2. The position of the first diffraction order ($m = +1$) $y^{(+1)} = \overline{A^{(0)}A^{(+1)}}$ ($y^{(+1)} < 0$ for $f' > 0$) is measured, and the focal length of the tested optical system is calculated with Eq. (7); it holds $f'^{(+1)} = -y^{(+1)}(a^2 - \lambda^2)^{1/2}/\lambda$. Afterward, the distance $y^{(-1)} = \overline{A^{(0)}A^{(-1)}}$ is measured, and the focal length for the minus first diffraction order (m=−1) is calculated; it holds $f'^{(-1)} = y^{(-1)}(a^2 - \lambda^2)^{1/2}/\lambda$. The resulting focal length of the optical system is then calculated as follows: $f' = (f'^{(+)} + f'^{(-)})/2$.

The focal length f' is therefore calculated. This result is used for the calculation of the distortion of the optical system. However, this value is not explicitly equal to the paraxial focal length because the measured values $y^{(+1)}$ and $y^{(-1)}$ differ from the paraxial ones by the distortion of the optical system. Nevertheless, if the values $y^{(+1)}$ and $y^{(-1)}$ are smaller than the quarter of the semidiameter of the field of view of the tested optical system (it can be adjusted by the diffraction grating with appropriate spatial frequency N), the effect of the distortion can be practically neglected. Afterward, the measured focal length can be considered as the paraxial focal length of the tested optical system.

The measurement of the distortion of the optical system proceeds as follows:

3. The positions of maximums of the higher diffraction orders than $m = 1$ are measured with $\beta = 0$, that is, values $y^{(+m)} = \overline{A^{(0)}A^{(+m)}}$ and $y^{(-m)} = \overline{A^{(0)}A^{(-m)}}$. Afterward, the value $\bar{y}^{(+m)} = (y^{(+m)} - y^{(-m)})/2$ is calculated.
4. The value of $\tan \alpha_{mz} = m\lambda/[a^2 - (m\lambda)^2]^{1/2}$ is calculated with the "paraxial image height" $y_{0\beta}^{(m)} = -f' \tan \alpha_{mz}$, where the focal length f' is known from paragraph 2.
5. The distortion belonging to such paraxial image height $y_{0\beta}^{(+m)}$ is given with Eq. (8); it holds $\delta y = \bar{y}^{(+m)} - y_{0\beta}^{(+m)}$.

The distortion for the paraxial image height $y_{0\beta}^{(+m)}$ is therefore calculated for the diffraction grating perpendicular to the optical axis ($\beta = 0$). The distortion of different paraxial image heights can be calculated with the following procedure (similarly as in the previous paragraphs).

6. The diffraction grating is rotated by the angle β (or it is changed with a different one with higher spatial frequency and $\beta = 0$); the angle α_{mz} is calculated from the formula $\alpha_{mz} = \arcsin(\sin\beta + m\lambda/a) - \beta$, and the paraxial image height from the formula $y_{0\beta}^{(m)} = -f' \tan \alpha_{mz}$.
7. The positions of the first and the second diffraction orders are measured, namely, $y_\beta^{(\pm m)} = \overline{A^{(0)}A^{(\pm m)}}$. Afterward, the value $\bar{y}_\beta^{(m)} = (y_\beta^{(+m)} - y_\beta^{(-m)})/2$ is calculated.
8. The distortion belonging the paraxial image height $y_{0\beta}^{(m)}$ is afterward calculated as follows: $\delta y_\beta^{(m)} = \bar{y}^{(m)} - y_{0\beta}^{(m)}$.

The issue of the determination of the distortion of the optical system is therefore solved.

3. MEASUREMENT ERRORS AND UNCERTAINTY ANALYSIS

It was already stated above that one of the sources of error is the uncertainty of determination of the paraxial focal length caused by the optical system's distortion. As is known, the distortion of the optical system for small angles of the field of view is described enough accurately (in the sense of geometrical optics) by the theory of aberrations of the third order, and one can express the distortion $\delta y'_D$ by the following approximate formula [32]:

$$\delta y'_D = -\frac{1}{2}(S_{II}A'^2_M tgw' + S_V tg^3 w'). \quad (9)$$

In Eq. (9) S_{II} is the coefficient of coma, S_V is the coefficient of distortion, $2w'$ is the angle of the field of view in the image space, and A'_M is the maximal numerical aperture of the optical system in the image space. It can be shown that the distortion of almost every optical system is practically zero for small angles of the field of view (until approx. 5°) [41–44]. Therefore, the value of the paraxial focal length of the tested optical system can be considered unaffected by the distortion.

In Fig. 3 an example of aberrations of the objective Tessar (AKLIN USP 2,165,328 [42]) with the focal length 99.767 mm is shown for the half-angle of the field of view

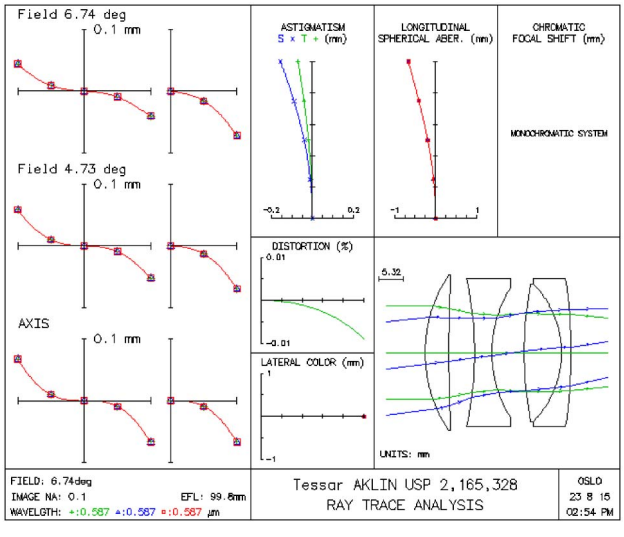


Fig. 3. Aberrations of objective Tessar $f' = 100$ mm for the half-angle of the field of view $w = 6.74°$ and for the wavelength $\lambda = 578$ nm.

$w = 6.742°$ (in object space) and for the wavelength 578 nm. Such an angle of the field of view corresponds to a diffraction angle of the first diffraction order ($m = 1$) of a grating with spatial frequency $N = 200$ lines/mm. The coefficients of coma and distortion of the objective are $S_{II} = 1.63$, $S_V = -0.705$. As is obvious from Fig. 3, the distortion is less than 0.01% for the image height $y = 11.794$ mm ($w' = 7.385°$, $A'_M = 0.1$). Therefore, the effect of the distortion on the paraxial focal length is negligible. The calculation was performed in the OSLO software [45].

Let us study the uncertainty of evaluated values of the measurement now. Imperfections of the used diffraction grating, an unstable source of light, misalignment of components of the measurement setup, and the quality of the used sensor can affect the results of the measurement procedure.

First, one can analyze the uncertainty in the image height, $u(y_\beta^{(m)})$, for a given image height, $y_\beta^{(m)}$. For the rotated diffraction grating about the angle β one can use Eqs. (2) and (5), that is, $y_\beta^{(m)} = -f' \tan \alpha_{mz}$, where $\alpha_{mz} = \arcsin(\sin \beta + m\lambda/a) - \beta$. Suppose now that systematic errors of the measurement are suppressed and random errors are very small in comparison to absolute values of the mutually independent parent variables, that is, the effect of absolute errors can be compared with the value of a total differential [46] of the functional value of the image height $y_\beta^{(m)}$. Moreover, suppose that the errors have an even probability distribution (one can use such an assumption because the errors of the aforementioned values can be positive or negative as well, and small errors are more probable than the large ones). Without the loss of generality such conditions can be accepted for the studied situation. Therefore, the effect of uncertainties can be described with the law of variance propagation [46–48], and the estimation of the uncertainty $u(y_\beta^{(m)})$ of the image height $y_\beta^{(m)}$ is given with the following formula:

$$u\left(y_\beta^{(m)}\right) = \pm\sqrt{\tan^2 \alpha_{mz} u^2(f') + \left(U_G^{(m)}\right)^2 + u^2(r)}, \quad \textbf{(10)}$$

where

$$U_G^{(m)} = \frac{f'}{\cos^2 \alpha_{mz}}\left[(U_\beta^{(m)})^2 u^2(\beta) + (U_a^{(m)})^2 u^2(a) + (U_\lambda^{(m)})^2 u^2(\lambda)\right]^{1/2},$$

$$U_\beta^{(m)} = 1 - \frac{\cos \beta}{\cos \alpha_m}, \qquad U_a^{(m)} = \frac{m\lambda}{a^2 \cos \alpha_m},$$

$$U_\lambda^{(m)} = \frac{-m}{a \cos \alpha_m}, \qquad \alpha_m = \arcsin(\sin \beta + m\lambda/a),$$

$u(f')$ is the uncertainty in the focal length, $u(\beta)$ is the uncertainty in the angle of the diffraction grating's rotation, $u(a)$ is the uncertainty in the grating's period, $u(\lambda)$ is the uncertainty in the wavelength, and $u(r)$ denotes the uncertainty of the measured image height, which has to be considered according the accuracy of the used reading device (e.g., an objective microscope, the resolution of digital sensor).

The estimation of the uncertainty of the focal length calculated with the use of paragraph 2 of the aforementioned measurement procedure can be expressed with a similar approach. With the use of the law of variance propagation [46–48] and after an elementary rearrangement, one gets

$$u(f') = \pm\sqrt{\frac{1}{2}[(yK)^2 u^2(a) + (ayK/\lambda)^2 u^2(\lambda) + (f'/y)^2 u^2(y)]}, \quad \textbf{(11)}$$

where $K = a/[\lambda(a^2 - \lambda^2)^{1/2}]$ and $u(y)$ is calculated with Eq. (10) for $\beta = 0$ and $m = 1$ (for the angle $\beta = 0$ the following equality holds: $u(y_{\beta=0}^{(+1)}) = u(y_{\beta=0}^{(-1)})$).

The estimation of uncertainty of the distortion of the optical system can be derived with the use of Eq. (10), because the

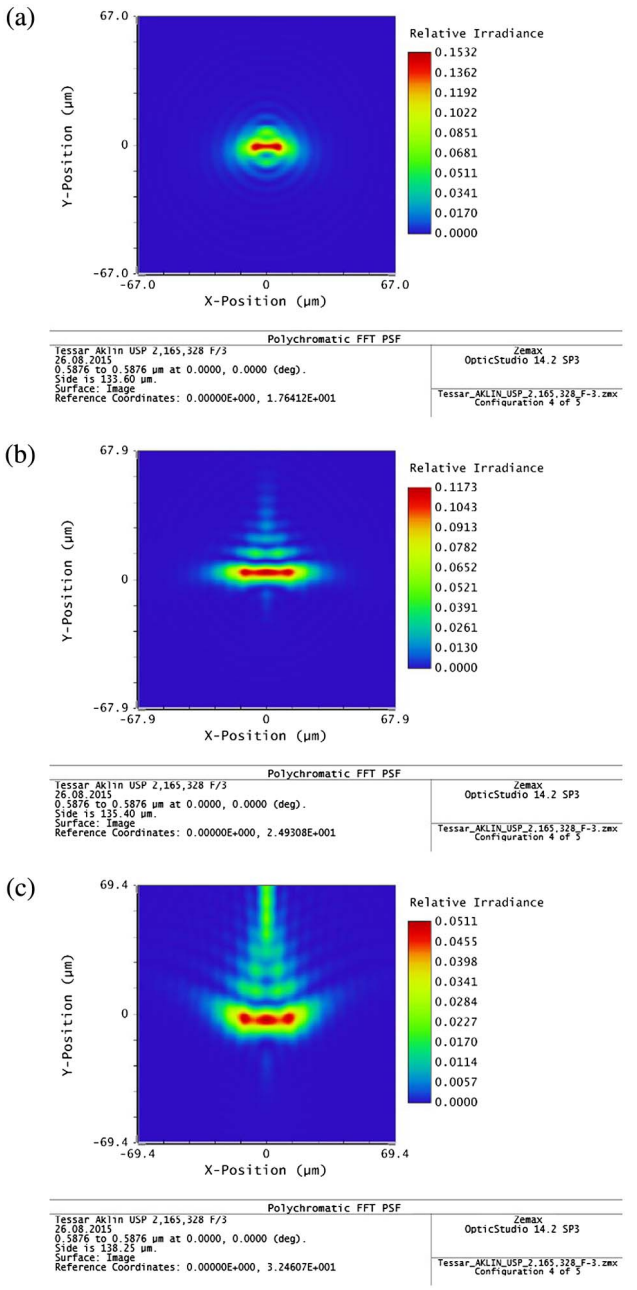


Fig. 4. Point spread functions in the paraxial image focal plane for the objective Tessar $f' = 100$ mm for the angle of the field of view α_{mz} equal (a) 10°, (b) 14°, and (c) 18° and for the diameter of the input beam $D = 10$ mm.

value of the paraxial image height is present in Eq. (8) for the calculation of the distortion. Next, the uncertainty in determination of the position of the diffraction orders' maximums will affect the result. Applying the law of variance propagation [46–48] gives the following formula for the uncertainty $u(\delta y_\beta^{(m)})$ of the distortion of the optical system; it holds

$$u(\delta y_\beta^{(m)}) = \pm\frac{\sqrt{2}}{2}\sqrt{u^2(r) + 2u^2(y_{0\beta}^{(m)})}, \quad (12)$$

where $u(r)$ denotes the uncertainty in determination of the position of diffraction orders' maximums, and $u(y_{0\beta}^{(m)})$ is calculated with Eq. (10) for $u(r) = 0$.

The last unknown information for the uncertainty analysis is the expression of uncertainties $u(x)$ of input variables x. The manufacturers usually provide estimations of absolute errors ($\Delta(x)$), peak-to-valley values (PV(x)), or root-mean-square values (RMS(x)). If one assumes a normal probability distribution of errors with 95% confidence interval, the estimation of the uncertainties of variable x can be calculated as follows: $u(x) \approx \Delta(x)/1.96 \approx \mathrm{PV}(x)/3 \approx \mathrm{RMS}(x)$.

The given issue of estimation of the uncertainty of the focal length and the uncertainty of the distortion of the tested optical system for paraxial image heights is therefore solved.

4. SIMULATION OF MEASUREMENT METHOD

The measurement of the focal length and the distortion of an objective Tessar (AKLIN USP 2,165,328 [42]) was simulated in the software Zemax OpticStudio [49]. A fictive diffraction grating was placed in front of the objective perpendicularly ($\beta = 0$) to an optical axis. The spatial frequency of the grating was changed according the chosen angle of the field of view α_{mz} for the second diffraction order; the frequency is given with the following formula: $N = \sin\alpha_{mz}/(m\lambda)$. The wavelength used was 587 nm. In Fig. 4 the point spread functions in the paraxial image focal plane are shown for the angles of the field of view 10°, 14°, and 18° and for a diameter of an input beam $D = 10$ mm.

Similarly, the point spread functions were simulated for different angles of the field of view, and the positions of the diffraction maximums (energy centres) were registered. The evaluation of the distortion from the second diffraction orders with the procedure described in Section 2 (paragraphs 3–8) are shown in Table 1. The calculated distortion corresponds with a direct calculation with Zemax OpticStudio [49]. The following values were chosen for the estimation of uncertainty: $u(\lambda) = \pm 2.3\mathrm{e}-2$ nm (corresponds to a temperature stability $\mathrm{d}\lambda/\mathrm{d}T = 0.045$ nm/K), $u(a) = \pm 1.3\mathrm{e}-7$ mm (corresponds to an error 0.01 lines/mm), $u(r) = \pm 2.6\mathrm{e}-3$ mm, $u(\beta) = \pm 0.001°$. The focal length was calculated according the procedure described in paragraphs 1 and 2 in Section 2 from the position of the first diffraction maximums; it holds: $f' = 99.975$ mm, $u(f') = \pm 0.016$ mm.

5. EXPERIMENTAL VERIFICATION OF THE METHOD ON THE MEASUREMENT OF FOCAL LENGTH

In Fig. 5(a) the laboratory setup for the experimental verification of the method for the measurement of the focal length of the optical system with the use of a diffraction grating is shown. Figure 5(b) shows the registered image of the diffraction pattern. A point source is realized with a distributed Bragg reflector laser diode ($\lambda = 631.6$ nm) coupled to a fiber. The uncertainty $u(\lambda) = \pm$ 2.3e-2 nm corresponds to such a source. A diverging beam of rays is collimated with a doublet (Linos, $D = 25.4$ mm, $f' = 100$ mm), and a diffraction pattern formed by a grating [$a = 0.127$ mm, $\beta = 0$, $u(a) = \pm 0.01\%$, $u(\beta) = \pm 1'$] is registered on the CMOS camera (Thorlabs DCC1645C, pixel size 3.6 μm) after passage through a tested objective (COSMICAR 1:1.4, $f' = 75$ mm), where the F-number $F = 16$ was set by an iris aperture.

In Fig. 6 the evaluation of the registered image is shown. The evaluation algorithm determined the positions of the diffraction maximums according the zero diffraction order with the uncertainty $u(r) = \pm$ (pixel size)/8.

As is obvious from the positions of the diffraction maximums that the evaluated area lies approximately in a paraxial image space. Therefore, the distortion of the tested objective can be considered zero for such a measured range. The results

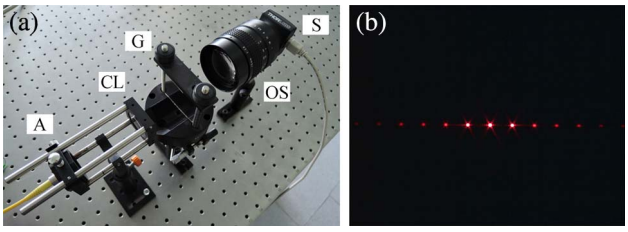


Fig. 5. Experimental verification of the method of measurement of the focal length with the use of diffraction grating: (a) laboratory setup (A, point source; CL, collimating lens; G, diffraction grating; OS, optical system under test; S, sensor), (b) registered diffraction pattern.

Table 1. Measurement of Distortion of the Objective Tessar $f' = 100$ mm for Different Angles of the Field of View

Angle of Field of View α_{mz} [°]	Frequency of Diffraction Grating N [lines/mm]	Measured Position of Second Diffraction Maximum $y^{(+2)}$ [mm]	Calculated Paraxial Image Height of Second Diffraction Maximum $y_{0\beta}^{(+2)}$ [mm]	Distortion δy [mm]	Uncertainty $u(\delta y)$ [mm]
10	147.91	−17.6412	−17.6468	0.0056 (−0.03%)	± 0.0034
12	177.10	−21.2605	−21.2730	0.0125 (−0.06%)	± 0.0040
14	206.07	−24.9308	−24.9535	0.0227 (−0.09%)	± 0.0045
16	234.78	−28.6613	−28.6990	0.0377 (−0.13%)	± 0.0052
18	263.22	−32.4607	−32.5204	0.0597 (−0.18%)	± 0.0058
20	291.33	−36.3358	−36.4299	0.0941 (−0.26%)	± 0.0065

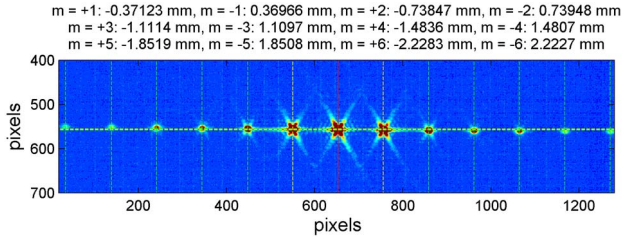


Fig. 6. Evaluation of registered diffraction pattern and calculated distances of diffraction maximums from the maximum of the zero order.

Table 2. Evaluated Focal Length of the Objective COSMICAR $f' = 75$ mm Calculated from Six Diffraction Maximums

Diffraction Order m	Focal Length f' [mm]
1	74.487
2	74.492
3	74.427
4	74.490
5	74.430
6	74.549

of the evaluated focal length are shown in Table 2. The values for diffraction orders higher than one were calculated with the formula stated in paragraph 4 of Section 2, namely, $f' = -y_{0\beta}^{(m)} / \tan \alpha_{mz}$, where $\tan \alpha_{mz} = m\lambda/[a^2 - (m\lambda)^2]^{1/2}$. The uncertainty of the focal length calculated from the first diffraction maximum is ± 0.066 mm ($\pm 0.09\%$).

The accuracy of the proposed method was verified by comparing the results with those of a commercially measurement device. The measurements of the focal length of the aforementioned objective were carried out using the OptiCentric MOT 2R from the company Trioptic [12] and with the OTS 200 from the company OEG-Messtechnik [50]. The computed focal length was $f' = 74.52 \pm 0.08$ mm. As is obvious, the results of the tested focal length are statistically the same because of the intersection of the confidence intervals.

6. CONCLUSION

The paper presented an experimentally simple, rather accurate and inexpensive method for a measurement of the focal length and the distortion of optical systems with the use of a diffraction grating. In a comparison with different methods [4–31], the proposed approach does not require any special components or any complicated, and therefore expensive, equipment. The detailed theoretical analysis was performed, and the estimation of uncertainties was studied as well. Afterward, the method was demonstrated via both computer simulations and experimental measurements. As is obvious from the results of the simulation and the experimental measurement, the method provides sufficiently accurate results for many practical applications. Moreover, there is no commercially available device measuring the focal length and the distortion of the optical system in one step. Therefore, the proposed method is appropriate for laboratory testing and for industrial applications as well.

Funding. Czech Technical University in Prague (SGS15/125/OHK1/2T/11).

REFERENCES

1. M. Born and E. Wolf, *Principles of Optics* (Oxford University, 1964).
2. M. Herzberger, *Modern Geometrical Optics* (Interscience, 1958).
3. A. Miks, *Applied Optics* (Czech Technical University, 2009).
4. A. Miks and J. Novak, "Estimation of accuracy of optical measuring systems with respect to object distance," Opt. Express **19**, 14300–14314 (2011).
5. J. Nemec, "Measurement of focal length," Fine Mech. Opt. **16**, 12–25 (1971).
6. D. Malacara, *Optical Shop Testing* (Wiley, 2007).
7. V. A. Afanasjev, *Optical Measurements* (High School, 1981).
8. G. V. Kreopalova, N. L. Lazareva, and D. T. Puriajev, *Optical Measurements* (Maschinostroenie, 1987).
9. J. Picht, *Mess-und Prüfmethoden der optischen Fertigung* (Akademie Verlag, 1953).
10. J. Flügge, *Einführung in die Messung der optischen Grundgrössen* (Verlag Braun, 1954).
11. B. Dorband, H. Miller, and H. Gross, *Handbook of Optical Systems: Metrology of Optical Components and Systems* (Willey, 2012), Vol. 5.
12. Trioptics, 2015, http://www.trioptics.com/.
13. Möeller-Wedel Optical, 2015, http://www.moeller-wedel-optical.com/.
14. L. Liao, F. C. Bráulio de Albuquerque, R. E. Parks, and J. M. Sasian, "Precision focal-length measurement using imaging conjugates," Opt. Eng. **51**, 113604 (2012).
15. I. Glatt and O. Kafri, "Determination of the focal length of nonparaxial lenses by moiré deflectometry," Appl. Opt. **26**, 2507–2508 (1987).
16. E. Keren, K. M. Kreske, and O. Kafri, "Universal method for determining the focal length of optical systems by moiré deflectometry," Appl. Opt. **27**, 1383–1385 (1988).
17. S. Trivedi, J. Dhanotia, and S. Prakash, "Measurement of focal length using phase shifted moiré interferometry," Opt. Lasers Eng. **51**, 776–782 (2013).
18. Y. Nakano and K. Murata, "Measurements of phase objects using the Talbot effect and moiré techniques," Appl. Opt. **23**, 2296–2299 (1984).
19. Y. Nakano and K. Murata, "Talbot interferometry for measuring the focal length of a lens," Appl. Opt. **24**, 3162–3166 (1985).
20. K. V. Sriram, M. P. Kothiyal, and R. S. Sirohi, "Direct determination of focal length by using Talbot interferometry," Appl. Opt. **31**, 5984–5987 (1992).
21. X. Jin, J. Zhang, J. Bai, Ch. Hou, and X. Hou, "Calibration method for high-accuracy measurement of long focal length with Talbot interferometry," Appl. Opt. **51**, 2407–2413 (2012).
22. J. Luo, J. Bai, J. Zhang, Ch. Hou, K. Wang, and X. Hou, "Long focal-length measurement using divergent beam and two gratings of different periods," Opt. Express **22**, 27921–27931 (2014).
23. Y. P. Kumar and S. Chatterjee, "Technique for the focal-length measurement of positive lenses using Fizeau interferometry," Appl. Opt. **48**, 730–736 (2009).
24. Z. Yang, Z. Gao, J. Dou, and X. Wang, "Focal length measurement based on the wavefront difference method by a Fizeau interferometer," Appl. Opt. **53**, 5598–5605 (2014).
25. M. de Angelis, S. De Nicola, P. Ferraro, A. Finizio, and G. Pierattini, "A new approach to high accuracy measurement of the focal lengths of lenses using a digital Fourier transform," Opt. Commun. **136**, 370–374 (1997).
26. J. Dhanotia and S. Prakash, "Focal length and radius of curvature measurement using coherent gradient sensing and Fourier fringe analysis," Optik **124**, 2115–2120 (2013).
27. D. R. Neal, J. Copland, D. A. Neal, D. M. Topa, and P. Riera, "Measurement of lens focal length using multi-curvature analysis of

Shack-Hartmann wavefront data," Proc. SPIE **5523**, 243–255 (2004).
28. J. Wu, J. Chen, A. Xu, X. Gao, and S. Zhuang, "Focal length measurement based on Hartmann-Shack principle," Optik **123**, 485–488 (2012).
29. L. Motka, "Measurement of focal length with SHS," Fine Mech. Opt. **59**, 43–45 (2014).
30. C. Mafusire and A. Forbes, "Mean focal length of an aberrated lens," J. Opt. Soc. Am. A **28**, 1403–1409 (2011).
31. V. I. Meshcheryakov, M. I. Sinelnikov, and O. K. Filippov, "Measuring the focal lengths of long-focus optical systems," J. Opt. Technol. **66**, 458–459 (1999).
32. A. Miks and J. Novak, "Dependence of camera lens induced radial distortion and circle of confusion on object position," Opt. Laser Technol. **44**, 1043–1049 (2012).
33. A. R. Tanguay, Jr., C. S. Wu, P. Chavel, T. C. Strand, and A. A. Sawchuk, "Physical characterization of the variable grating mode liquid crystal device," Opt. Eng. **22**, 687–694 (1983).
34. B. H. Soffer, J. D. Margerum, A. M. Lackner, D. Boswell, and A. R. Tanguay, Jr.., T. C. Strand, A. A. Sawchuk, and P. Chavel, "Variable grating mode liquid crystal device for optical processing computing," Mol. Cryst. Liq. Cryst. **70**, 145–161 (1981).
35. M. Aschwanden and A. Stemmer, "Polymeric, electrically tunable diffraction grating based on artificial muscles," Opt. Lett. **31**, 2610–2612 (2006).
36. J. Kim, J. H. Suh, B. Y. Lee, S. U. Kim, and S.-D. Lee, "Optically switchable grating based on dye-doped ferroelectric liquid crystal with high efficiency," Opt. Express **23**, 12619–12627 (2015).
37. J. Sun, A. K. Srivastava, L. Wang, V. G. Chigrinov, and H. S. Kwok, "Optically tunable and rewritable diffraction grating with photoaligned liquid crystals," Opt. Lett. **38**, 2342–2344 (2013).
38. J. Kim, J. H. Na, and S. D. Lee, "Fully continuous liquid crystal diffraction grating with alternating semi-circular alignment by imprinting," Opt. Express **20**, 3034–3042 (2012).
39. R. H. Chen, *Liquid Crystal Display: Fundamental Physics and Technology* (Wiley, 2011).
40. Holoeye, 2015, http://www.holoeye.com.
41. S. F. Ray, *Applied Photographic Optics* (Focal, 2002).
42. W. Smith, *Modern Optical Engineering*, 3rd ed. (McGraw-Hill, 2000).
43. A. Cox, *A System of Optical Design* (Focal, 1964).
44. W. Merte, R. Richter, and M. V. Rohr, *Das Photographische Objektiv* (Springer, 1932).
45. OSLO, 2015, http://www.lambdares.com/oslo.
46. K. Rektorys, *Survey of Applicable Mathematics* (MIT, 1969).
47. M. Grabe, *Measurement Uncertainties in Science and Technology* (Springer, 2014).
48. S. Gupta, *Measurement Uncertainties: Physical Parameters and Calibration of Instruments* (Springer, 2012).
49. Zemax, 2015, http://www.zemax.com.
50. OEG-Messtechnik, 2015, http://www.oeg-messtechnik.de.