

Systematic search for new solutions in lens design

Hou, Z.

DOI

[10.4233/uuid:485031ef-4fcf-4c2f-9a6b-41b037e88afe](https://doi.org/10.4233/uuid:485031ef-4fcf-4c2f-9a6b-41b037e88afe)

Publication date

2023

Document Version

Final published version

Citation (APA)

Hou, Z. (2023). *Systematic search for new solutions in lens design*. [Dissertation (TU Delft), Delft University of Technology]. <https://doi.org/10.4233/uuid:485031ef-4fcf-4c2f-9a6b-41b037e88afe>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.



SYSTEMATIC SEARCH FOR NEW SOLUTIONS IN LENS DESIGN

Zhe Hou | 侯哲

SYSTEMATIC SEARCH FOR NEW SOLUTIONS IN LENS DESIGN

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus prof. dr. ir. T. H. J. van der Hagen,
chair of the Board for Doctorates,
to be defended publicly on
Tuesday 11 April 2023 at 12:30 o'clock

by

Zhe HOU

侯哲

Master of Science in Applied Physics
Delft University of Technology, The Netherlands
born in Hohhot, Nei Mongol, P. R. China.

This dissertation has been approved by the promotor.

Composition of the doctoral committee:

Rector Magnificus	chairperson
Prof. dr. H. P. Urbach	Delft University of Technology, promotor
Dr. F. Bociort	Delft University of Technology, copromotor

Independent members:

Prof. dr. T. H. W. Bäck	Leiden University
Prof. dr. B. H. W. Hendriks	Delft University of Technology
Prof. dr. ir. W. L. Ijzerman	Eindhoven University of Technology
Prof. dr. rer. nat. habil. S. Sinzinger	Technische Universität Ilmenau, Germany
Dr. B. Kruizinga	TNO (Toegepast Natuurwetenschappelijk Onderzoek)
Prof. dr. S. Stallinga	Delft University of Technology, reserve member



This research has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/under REA grant agreement number 608082.

Keywords: Optical Design, Lens Design Method, Optimization, Saddle Point

Printed by: Ipskamp Printing

Cover art: Designed by Hui LIN. It contains the elements of lenses, saddles, with simple components forming complex patterns.

Copyright © 2023 by Zhe Hou

ISBN 978-94-6384-431-4

An electronic version of this dissertation is available at
<http://repository.tudelft.nl/>.

*To my dearest parents
I love you*

CONTENTS

Summary	1
Samenvatting	3
1 Introduction	7
References	10
2 Optical Lens Design	11
2.1 Designing a Magnifying Glass	11
2.2 Modern Optical Design	15
2.2.1 Optical System Optimization	17
2.2.2 Getting the Starting Point	19
2.2.3 Local Optimization	19
2.2.4 Global Optimization	22
2.2.5 Saddle Point Method	26
2.3 Conclusion	29
References	29
3 The Saddle Point Construction (SPC) Method	31
3.1 Saddle Point and Design Landscape	31
3.2 The SPC Method	32
3.2.1 The General Version of SPC	32
3.2.2 The Special Version of SPC	36
3.3 Recommendations for Applying SPC	37
3.3.1 Adding Lens Elements	37
3.3.2 Switching to Alternative Minima	39
3.4 Practical flow of implementing SPC in lens design	41
3.5 Conclusion	44
References	45
4 Lens Design Landscape Exploration with Simple Systems	47
4.1 Wide-angle Pinhole Lens	48
4.2 Switching Between Local Minima	49
4.3 Designing a Pinhole System Starting from a Single Lens	51
4.4 Decomposing a High-dimensional Search for New Minima in a Succession of One-dimensional Searches	52
4.5 Effect of Changing the Field of View on the Landscape	57
4.6 Discussion and Conclusion	62
References	63

5	Applying SPC in Complex Systems	65
5.1	Introduction	65
5.2	Investigation of a Wide-angle Lens with Six Elements	66
5.2.1	Design a Wide-angle Lens with SPC	67
5.2.2	Stepwise Analysis of SPC for a Wide-angle System	69
5.3	Combining Designers' Knowledge and SPC to Improve Design Efficiency.	78
5.3.1	Ultraviolet (UV) Microscope Dry Objective	78
5.3.2	Constraints in Optical Design Software	80
5.3.3	Applying SPC to Induce Structural Changes in Lens	83
5.3.4	Further Consideration when Applying SPC to Complex System	91
5.4	Applying SPC to a Photo-Lithographic Objective	95
5.5	Conclusion	99
	References	100
6	Simultaneous Multiple Surface (SMS) 2D Designs as Starting Points for Lens Optimization	103
6.1	Introduction	103
6.2	Starting Point with the SMS Method	105
6.3	Comparing Results of Different Design Approaches.	107
6.3.1	System 1	110
6.3.2	System 2	112
6.4	Design Landscape for Aspheric Systems	114
6.5	Conclusions.	116
	References	117
7	Conclusion and Outlook	119
7.1	Conclusion	119
7.2	Outlook	121
	References	122
A	Appendices	123
A.1	Surface Specifications of the Wide-angle Pinhole Lenses with Four Elements	123
A.2	Surface Specifications of the Wide-angle Pinhole Lenses with Six Elements.	124
A.3	Surface Specifications of the SMS Constructed Systems.	125
A.4	Definition of the Qcon Polynomials.	127
	References	127
	Curriculum Vitæ	129
	List of Publications	131
	Acknowledgement	133

SUMMARY

In this thesis, we explore how lens design and optimization techniques can adapt to the design (optimization) space in order to increase the lens design efficiency. We extensively discuss the Saddle Point Construction (SPC), a method that can systematically search for new solutions, as well as replace high-dimensional searches with a discrete number of one-dimensional searches to increase efficiency.

The optical lens design space is known to be very complicated. This is a high-dimensional space defined by the degrees of freedom and the quantified performance of a parameterized optical system. This high-dimensional space is also nonlinear which means there are multiple local minima (one can think of the bottom of a crater in the geological landscape) with different values (deep and shallow craters). In modern optical lens design, the task of obtaining a useful design solution can be described as searching for the global minimal value in the high-dimensional space.

For sophisticated lens designs, the number of the degrees of freedom increases in order to gain more control of the design. However, for the search of solutions, it brings two major challenges. First, the size of the high dimensional space increases according to the power law and the number of points to be evaluated increases tremendously. For example, if 100 points need to be evaluated in each one-dimensional direction, a three-dimensional space requires 100^3 points to be evaluated. It is 100 times of the two-dimensional case, and 10, 000 times of the one-dimensional case. For a complicated optics systems which usually contains more than twenty variables, it is practically impossible (given the current computing power) to evaluate every points in the high-dimensional space in order to find the global minimum. Second, the number of local minima increases with the number of the degrees of freedom used. It indicates the number of undesired local minima becomes larger. Hence, the difficulty of finding best minima also increases.

We demonstrate in this thesis that, considering an actual lens design process, the design space is not only statically complicated, but also dynamically changes through the design process. As a consequence of the dynamic aspect, the number of local minima varies and the effectiveness of the optimization techniques can be impacted. In Chapter 6, a study using Simultaneous Multiple Surface (SMS) method as one of the design strategies, compares the effectiveness of different strategies under static and dynamic design landscape. Our recommendation is to structurally work with the static "snapshot" of the design landscape to keep the complexity under control. In Chapter 5, we explain how to use SPC to achieve this by improving a deep ultra-violet photo-lithographic objective.

Over the past decades, different design and optimization techniques have been examined on their effectiveness of obtaining a global optimal solution. Some of the techniques have been integrated into commercial lens design software. However, these algorithms are based almost exclusively on generally applicable mathematical models and use little or no specific knowledge about the optical system (and its design landscape).

As a result, little information is available on how these techniques should be used in a practical design task instead of implementing them as a lottery draw.

Saddle Point Construction is a design method which can rapidly construct saddle points with Morse Index 1 in a design space. These saddle points serve as agents to guide the local optimization to obtain new minima. Different from other optimization techniques, it also reveals a special structure in the lens design landscape: certain saddle points existing in the landscape are reducible to minima of simpler systems plus one additional lens element. By exploiting this property, searching for minima in a high-dimensional lens design landscape can be simplified to several one-dimensional scans combining with standard local optimization routines. The effort for obtaining a satisfactory lens design solution is therefore reduced.

Previous research shows potentials of SPC as a systematic lens design technique. In this thesis, we further examine the practicality of SPC as a global optimization or semi-global optimization (to generate a small pool of solutions) technique. The dynamic aspect of the design space is particularly of interest since it is relevant to an actual design practice.

Different design examples are investigated in this thesis. They range from relative simple triplet design to complex photo-lithographic objective consisting of more than twenty lens elements. We demonstrate in Chapter 4 that for a simple triplet system, the method of SPC can obtain all the minima found by alternative methods. We also discover that some of the saddle points cannot be constructed using SPC. Nevertheless, the redundancy in the solution network makes it possible to still obtain the minima without identifying each saddle points.

When the configuration of the system changes (e.g. from polychromatic to monochromatic), we observe that not all the minima found by alternative methods can be easily obtained using SPC. We also observe that these minima, which are not found by SPC, are solutions having high merit function values (worse performance) compared to the best system. In the designs that we investigated, using SPC can obtain the best performing system(s). The systems with high value of merit functions are usually not stable. It means that they are more likely to change (occur ray failure or merge with an existing solution) when the system configuration changes. These phenomenons are explained in details in the latter part of Chapter 4. We also argue that in practice these minima with high merit function values are not likely to be considered, therefore, they are less important.

Given the demonstrated effectiveness of SPC in relative simple systems, Chapter 5 provides study cases with more complexity and practical relevancy. We argue that at a certain stage of the lens design process, having a whole-landscape search is not preferred. Instead, a better strategy would be searching in a confined landscape guided by designer's experience. We demonstrate that SPC can be particularly effective for this purpose. For complicated systems consisting many lens elements, SPC can rapidly provide a small pool of solution candidates, while keeping system configuration not drastically changed. In the same chapter, we have provided some guidelines on how to use SPC in practical design tasks. Given the demonstrated effectiveness, we believe SPC can certainly be a useful tool in the lens designer's toolbox.

SAMENVATTING

In dit proefschrift onderzoeken wij hoe lensontwerp en optimalisatietechnieken kunnen worden aangepast aan de ontwerp(optimalisatie)ruimte om de efficiëntie van het lensontwerp te vergroten. Wij bespreken uitgebreid de zadelpuntconstructie (Saddle Point Construction in het Engels, afgekort als SPC), een methode die systematisch naar nieuwe oplossingen kan zoeken, maar ook hoog-dimensionale zoekopdrachten kan vervangen door een discreet aantal eendimensionale zoekopdrachten om de efficiëntie te verhogen.

Het is bekend dat de ontwerpruimte van de optische lens erg gecompliceerd is. Dit is een hoogdimensionale ruimte die wordt gedefinieerd door de vrijheidsgraden en de gekwantificeerde prestaties van een geparametriseerd optisch systeem. Deze hoogdimensionale ruimte is ook niet-lineair, wat betekent dat er meerdere lokale minima zijn (denk aan de bodem van een krater in het geologische landschap) met verschillende waarden (diepe en ondiepe kraters). Bij het ontwerpen van moderne optische lenzen kan de taak van het verkrijgen van een bruikbare ontwerp oplossing worden omschreven als het zoeken naar de globale minimale waarde in de hoogdimensionale ruimte.

Voor geavanceerde lensontwerpen neemt het aantal vrijheidsgraden toe om meer controle over het ontwerp te krijgen. Het zoeken naar oplossingen brengt het echter twee grote uitdagingen met zich mee. Ten eerste neemt de grootte van de hoogdimensionale ruimte toe volgens de machtswet en neemt het aantal te evalueren punten enorm toe. Als er bijvoorbeeld 100 punten moeten worden geëvalueerd in elke eendimensionale richting, heeft een driedimensionale ruimte 100^3 punten nodig om te worden geëvalueerd. Het is 100 keer het tweedimensionale geval en 10.000 keer het eendimensionale geval. Voor een gecompliceerd optisch systeem dat gewoonlijk meer dan twintig variabelen bevat, is het praktisch onmogelijk (gezien de huidige rekenkracht) om elk punt in de hoog-dimensionale ruimte te evalueren om het globale minimum te vinden. Ten tweede neemt het aantal lokale minima toe met het aantal gebruikte vrijheidsgraden. Het geeft aan dat het aantal ongewenste lokale minima groter wordt. Daarom neemt ook de moeilijkheid om de beste minima te vinden toe.

In dit proefschrift laten wij zien dat, uitgaande van een echt lensontwerpproces, de ontwerpruimte niet alleen statisch gecompliceerd is, maar ook dynamisch verandert tijdens het ontwerpproces. Als gevolg van het dynamische aspect varieert het aantal lokale minima en kan de effectiviteit van de optimalisatietechnieken worden beïnvloed. In Hoofdstuk 6, een studie waarin de Gelijktijdig meerdere oppervlakken (Simultaneous Multiple Surface in het Engels, afgekort als SMS) methode als een van de ontwerpstrategieën wordt gebruikt, wordt de effectiviteit van verschillende strategieën in een statisch en dynamisch ontwerplandschap vergeleken. Onze aanbeveling is om structureel te werken met de statische "momentopname" van het ontwerplandschap om de complexiteit onder controle te houden. In Hoofdstuk 5 leggen wij uit hoe SPC te gebruiken om dit te bereiken door een diep ultraviolet fotolithografisch objectief.

In de afgelopen decennia zijn verschillende ontwerp- en optimalisatietechnieken onderzocht op hun effectiviteit voor het verkrijgen van een globaal optimale oplossing. Sommige technieken zijn geïntegreerd in commerciële lensontwerpssoftware. Deze algoritmen zijn echter bijna uitsluitend gebaseerd op algemeen toepasbare wiskundige modellen en gebruiken weinig of geen specifieke kennis over het optische systeem (en zijn ontwerplandschap). Als gevolg hiervan is er weinig informatie beschikbaar over hoe deze technieken zouden moeten worden gebruikt in een praktische ontwerptaak in plaats van ze te implementeren als een loterijtrekking.

Zadelpuntconstructie is een ontwerpmethode waarmee snel zadelpunten met Morse Index 1 in een ontwerpruimte kunnen worden geconstrueerd. Deze zadelpunten dienen als agenten om de lokale optimalisatie te begeleiden om nieuwe minima te verkrijgen. Anders dan bij andere optimalisatietechnieken, onthult het ook een speciale structuur in het lensontwerplandschap: bepaalde in het landschap bestaande zadelpunten zijn herleidbaar tot minima van eenvoudigere systemen plus één extra lens-element. Door gebruik te maken van deze eigenschap kan het zoeken naar minima in een hoog-dimensionaal lensontwerplandschap worden vereenvoudigd tot verschillende eendimensionale scans in combinatie met standaard lokale optimalisatieroutines. De inspanning voor het verkrijgen van een bevredigende lensontwerp oplossing wordt daardoor vermindert.

Eerder onderzoek toont de mogelijkheden van SPC als een systematische lensontwerptechniek. In dit proefschrift onderzoeken wij verder de bruikbaarheid van SPC als een techniek voor globale optimalisatie of semi-globale optimalisatie (om een kleine pool van oplossingen te genereren). Vooral het dynamische aspect van de ontwerpruimte is interessant, omdat het relevant is voor een daadwerkelijke ontwerp werkwijze.

In dit proefschrift worden verschillende ontwerpvoorbeelden onderzocht. Ze variëren van een relatief eenvoudig tripletontwerp tot een complex fotolithografisch objectief bestaande uit meer dan twintig lenselementen. Wij demonstreren in Hoofdstuk 4 dat voor een eenvoudig triplet-systeem de methode SPC alle minima kan verkrijgen die gevonden zijn met alternatieve methoden. Wij ontdekken ook dat sommige zadelpunten niet met SPC kunnen worden geconstrueerd. Desalniettemin maakt de redundantie in het oplossingsnetwerk het mogelijk om nog steeds de minima te verkrijgen zonder elk zadelpunt te identificeren.

Wanneer de configuratie van het systeem verandert (bijvoorbeeld van polychromatisch naar monochromatisch), zien wij dat niet alle minima gevonden door alternatieve methoden gemakkelijk kunnen worden verkregen met behulp van SPC. Wij zien ook dat deze minima, die niet door SPC worden gevonden, oplossingen zijn met hoge waardefunctiewaarden (slechtere prestaties) in vergelijking met het beste systeem. In de ontwerpen die wij hebben onderzocht, kan het gebruik van SPC de best presterende systemen opleveren. De systemen met een hoge waarde van merit-functies zijn meestal niet stabiel. Het betekent dat de kans groter is dat ze veranderen (straalstoring optreden of samensmelten met een bestaande oplossing) wanneer de systeemconfiguratie verandert. Deze fenomenen worden in detail uitgelegd in het laatste deel van hoofdstuk 4. Wij betogen ook dat deze minima met hoge merit-functiewaarden in de praktijk waarschijnlijk niet in overweging worden genomen en daarom minder belangrijk zijn.

Gezien de aangetoonde effectiviteit van SPC in relatief eenvoudige systemen, biedt

Hoofdstuk 5 studiegevallen met meer complexiteit en praktische relevantie. Wij stellen dat in een bepaalde fase van het lensontwerpproces het niet de voorkeur heeft om het hele landschap te doorzoeken. In plaats daarvan zou het een betere strategie zijn om te zoeken in een beperkt landschap, geleid door de ervaring van de ontwerper. Wij laten zien dat SPC hiervoor bijzonder effectief kan zijn. Voor gecompliceerde systemen die uit veel lenselementen bestaan, kan SPC snel een kleine pool van oplossingskandidaten bieden, terwijl de systeemconfiguratie niet drastisch verandert. In hetzelfde hoofdstuk hebben wij enkele richtlijnen gegeven voor het gebruik van SPC in praktische ontwerptaken. Gezien de aangetoonde effectiviteit, geloven wij dat SPC zeker een nuttig hulpmiddel kan zijn in de gereedschapskist van de lensontwerper.

1

INTRODUCTION

Optical design, generally speaking, can refer to the activities that manipulate light to achieve a certain purpose. This can be designing a telescope to observe remote celestial bodies, designing a concentrator to maximally harvest the solar energy, or designing optical fibre to maximize efficiency for information transmission. In this work we will focus on imaging system design. That means using optical lenses and mirrors as means to achieve high quality images. The focus of this thesis will also be the design of optical lenses. However, the insights summarized in this thesis can also be extended to the design with other optical manipulators (e.g. mirrors).

In lens design, the light is usually modeled with its geometrical model (or ray model). The light propagation is described in terms of light rays. When the wavelength of the light is very small compared to the size of the structures with which light interacts, it is an excellent approximation. That also indicates the wave property of light such as interference and diffraction, is not captured by the geometrical model. The utilization of the simpler geometric model is mostly driven by the practicality of optical design: with geometrical optics where light is simplified as rays pointing towards the traveling direction of radiation energy, billion of rays can be traced in seconds for a standard modern personal computer. The performance of the system can be quickly determined for subsequent analysis or optimization. On the other side, with the more rigorous wave model, designing and analyzing a common optical system is computationally expensive, hence very time consuming and inconvenient to rapidly compute and optimize the system performance.

In some occasions, a wave model is necessary to be applied to understand the system performance. For instance, when an imaging system is only limited by diffraction, analyzing the intensity distribution of a focused beam at the focal plane requires a wave model. A geometrical model will only predict an extreme small spot without the information about the intensity distribution. As a result, depending on the performance requirements, practice for optical system design may involve both geometrical and wave models. A geometrical model is usually used to rapidly obtain a system solution that is good enough in terms of its performance predicted by the geometrical model. When

approximation with the geometrical model is no longer accurate, the wave model helps to provide further analysis which gives feedback to fine-tune the system performance. The later part is usually involved in systems with a demanding performance requirement such as microscopic objectives or photo-lithographic projection systems. In this thesis, the discussion is focused on designing optical systems based on the principles of geometrical optics. The wave model related analysis is not in scope.

In geometrical optics, one of most important laws would be the law of refraction (also know as the Snell's law). It describes how a light ray travels at the interface of two different media:

$$n \sin i = n' \sin r, \quad (1.1)$$

where n and n' are the refractive indices of the media before and after refraction of a ray; i and r are the angles of the incidence and refraction at the interface between the two media. According to the law of refraction, lenses are designed to manipulate the light.

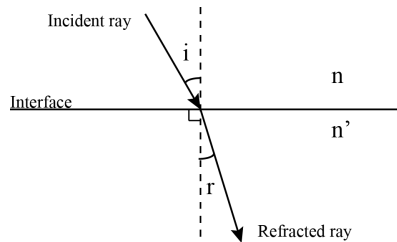


Figure 1.1: Snell's law: the light ray is refracted at the interface between the two media. The relation between the incident ray and the refracted ray is given by Equation 1.1.

The earliest known lenses were made by the ancient Egyptians and Mesopotamians [1]. They were made from polished crystal, often quartz, and are dated as early as 750BC for Assyrian lenses such as the Nimrud/Layard lens [2]. The exact function of these lenses is not clear, with some authors suggesting that they were used as an optical lens (magnifying glasses) and others suggesting a decorative function. These lenses were mainly made by trial and error where actual understanding of the light was very limited. When people had more knowledge about geometrical optics, optical design started to pay attention to how light rays travel. With the increasing understanding of the nature of light, the design of optics also evolved from simple lenses to complicated systems which can achieve resolutions down to nanometers. This can not be achieved without the development of modern computer-aided optical design with which the performance of the optical system can be rapidly estimated and numerical optimization can be applied to further improve the system performance.

In modern optical lens design, an optical system is parameterized in design software. The designer studies the design requirements and determines what the initial configuration of the optical lens should be. Figure 1.2 provides a flowchart of a modern optical design process. The optical designer then needs to start modifying the system based on his/her knowledge and observation with a purpose of meeting the design requirements. Numerical optimization, which is widely used in the computer-aid design, is also heavily applied in this step for lens design. With the help of the optimizer, the designer "guides" the system towards a desired configuration.

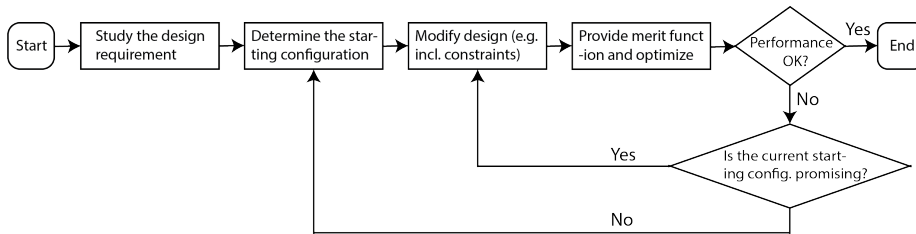


Figure 1.2: Flowchart of the modern optical design process.

The function describing the imaging quality (i.e. the merit function used in optimization) is a highly nonlinear function of the design parameters. As will be seen in Chapters 4 and 5, this nonlinearity leads to multiple local minima in the design landscape. Their number increases with the complexity (e.g. using more lenses) of the design problem. When running standard local optimization with such a merit function, the result is sensitive to the choice of the starting point. The optimization can either provide a satisfactory solution or be trapped in a poor-quality local minimum. This presents one of the major challenges of today’s optical lens design: to efficiently obtain the global optimal solution among many inferior ones. As the achievable minimum is sensitive to its starting point of the design, this challenge can also be reformulated as: to determine a starting configuration that would land at a global optimal. For complicated systems, finding such optimal starting configurations mostly depends on an optical designer’s knowledge and accumulated experience [3][4].

The exploration of various global optimization techniques (elaborated in the next chapter) has been given much attention over the last few decades. Given ever-increasing computational power, the hope is that these global optimization tools can lessen the burden on the optical designer by efficiently providing the solution candidates such that the designer can focus on other important tasks instead of trial and error.

One fact is that these global optimization algorithms rely almost exclusively on generally applicable mathematical models and use little or no specific knowledge about the optical system (and its design landscape). As a result, setting up the proper configurations (similar to the step-size of a local optimizer) of these global optimizer becomes another task of trial and error.

In this thesis, we particularly focus on the Saddle Point Construction (SPC) method which uses a special property in the lens design landscape: certain saddle points existing in the landscape are reducible to minima of simpler systems plus one additional element. By investigating how SPC behaves in different design scenarios, we provide insights for systematically design optical systems using SPC.

OUTLINE OF THE DISSERTATION

In Chapter 2, we provide some background regarding modern lens design. To better explain a design process, a simple example of designing a magnifying glass is used. We also discuss some of the major optimization techniques available either in commercial design software or publications.

Chapter 3 is dedicated to SPC where we give the mathematical proof for the general

version of SPC and explain in detail how SPC could be used in practical lens design. A flowchart on how to apply the method step by step is also provided.

In Chapter 4, a wide-angle pinhole lens design network containing saddle points and minima is given where we demonstrate in one scenario that using SPC can obtain all the minima via the constructed saddle points. However, we further discuss that it is not always valid. When lens design specification changes (e.g. field of view), the design landscape changes correspondingly which alters the number of minima and saddle points that can be retrieved via SPC.

Lenses with more complexities are used as study examples in Chapter 5. For a complex lens system with more number of elements, analyzing the solutions network becomes difficult. In practical design tasks, it is not always necessary to run a global search to list all the solutions in the design space. A small pool of solution candidates without drastically changing the system configurations, could help the design to progress in a more controlled way. We demonstrate in the chapter how SPC can be beneficial in such design scenarios as well as providing some practical insights about the method.

Chapters 4 and 5 describe the dynamic aspect of the lens design landscape - instead of remaining static, the design landscape changes continuously through the design process via modifying system specifications, applying constraints, etc. This adds the uncertainty of successfully obtaining a good solution. In Chapter 6, we deviate from SPC and compare different design strategies for simple aspheric lenses. It is observed that if a good starting configuration including all the constraints and necessary degrees of design freedom can be determined, reaching a global optimal solution can be very efficient. We use Simultaneous Multiple Surface (SMS) method in this chapter to construct such good starting configurations.

At last, conclusions of the thesis and some future outlooks are given in Chapter 7.

REFERENCES

- [1] Wikipedia contributors, *History of optics — Wikipedia, the free encyclopedia*, (2019), [Online; accessed December-2019].
- [2] Wikipedia contributors, *Nimrud lens — Wikipedia, the free encyclopedia*, (2019), [Online; accessed December-2019].
- [3] I. Livshits and V. Vasilyev, *Q and A tutorial on optical design*, *Advanced Optical Technologies* **2**, 31 (2013).
- [4] D. Shafer, *Doing more with less*, *Proc. SPIE* **2537**, 11 (1995).

2

OPTICAL LENS DESIGN

Successful design will require the application of individual creativity as well as artful manipulation and thorough comprehension of the numeric tools available in lens design programs.[1]

R. R. Shannon

In this chapter, we start with a simple example of how a magnifying glass can be designed on paper (instead of on a computer). This section is added since nowadays the design is mostly done with professional software and the story of how a lens can be designed from scratch based on theory, is less clear. After that, we explain the modern optical approach with an emphasis on optimization techniques. The SPC method is briefly mentioned at the end of the chapter in comparison with other design/optimization techniques and it is elaborated on in the next chapter.

2.1. DESIGNING A MAGNIFYING GLASS

We use a simple magnifying glass as an example to demonstrate the basic approach of lens design. As the name suggests, a magnifying glass magnifies objects and is commonly used to assist reading when the letters are too small to read. Assume the magnification requirement of the magnifier is $6\times$. We know from textbooks [2] that such a magnifier can be made using a convex lens. The magnified image is a virtual image that is then captured by the human-eye. The magnification of a magnifier is defined as the angular magnification which is the ratio between the subtended angles to the eye of the object and the image. This is illustrated in Figure 2.1-a. A larger angle means a larger image is formed on the retina of the eye. The object and the image are normally located at a distance of 250mm from the eye. The distance represents the near-point, which is the closest distance that the eye can view comfortably.

Figure 2.1-b shows how the virtual image is formed considering an ideal thin lens scenario. Certain rules are used to derive such a plot: 1) A light ray, which starts from the

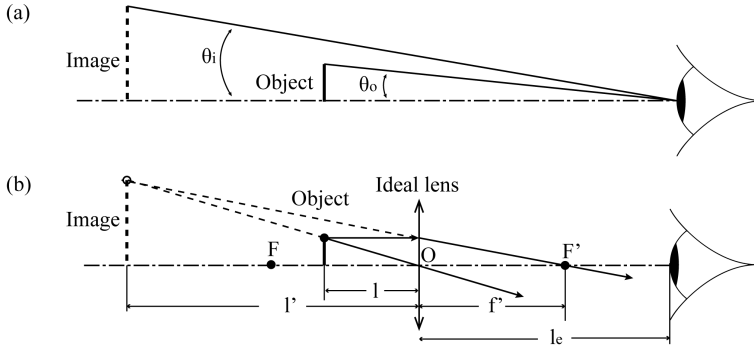


Figure 2.1: (a) The magnification of a magnifying glass is defined as the ratio between the image subtended angle θ_i and the object subtended angle θ_o . (b) Illustration of a virtual, enlarged image of the object through a magnifying glass. Equations 2.3 - 2.5 explain the relation of the parameters.

object plane and is parallel to the optical axis, will pass the focal point at image side; 2) A light ray, which starts from the object plane and passes the nodal point (the intersection between the ideal lens and the optical axis), will remain its in original direction.

In this thesis, we use the following convention for the geometrical optics formulas: points (object, image or focal point) sitting on the left side of the lens have a negative distance to the lens. Points sitting on the right side of the lens have a positive distance. The radius of the optical surfaces is positive when the center of curvature is to the right of the surface vertex, and vice versa.

The subtended angle θ_o of the object to the eye writes

$$\theta_o = \frac{h_o}{250 \text{ mm}}, \quad (2.1)$$

where h_o stands for the height of the object. The subtended angle θ_i of the magnified image writes

$$\theta_i = \frac{mh_i}{l_e - l'}, \quad (2.2)$$

where $m(= \frac{l'}{l})$ is the lateral magnification and l_e is the distance from the eye to the lens. Since the image is virtual ($l' < 0$), the expression $l_e - l'$ indicates the distance between the image and the eye. Together with the thin lens equation

$$\frac{1}{l'} - \frac{1}{l} = \frac{1}{f'}, \quad (2.3)$$

we can write Equation 2.2 into

$$\theta_i = \left(1 - \frac{l'}{f'}\right) \cdot \frac{h_o}{l_e - l'}. \quad (2.4)$$

With Equation 2.1 and Equation 2.4, the magnification m_{mag} of the magnifier is

$$m_{mag} = \frac{\theta_i}{\theta_o} = \left(1 - \frac{l'}{f'}\right) \cdot \frac{250 \text{ mm}}{l_e - l'}. \quad (2.5)$$

Often, we want to set the distance between the image and the eye, $l_e - l'$, as the near-point distance (250 mm). In this case, to maximize the magnification m_{mag} , we need to maximize the numerator of Equation 2.5 and set $l' = -250$ mm, $l_e \approx 0$. It means that the observer holds the magnifier close to the eye. This magnification is then given by

$$m_{mag} = 1 + \frac{250 \text{ mm}}{f'} \quad (2.6)$$

Now we substitute m_{mag} with 6, then we have $f' = 50$ mm. As a result, we need a magnifying glass with a focal length of 50 mm to achieve a magnification of $6\times$. Now we know the focal length $f' = 50$ mm and the image distance $l' = -250$ mm. Given Equation 2.3, the object distance is calculated to be 41.67 mm.

For a spherical lens, the radii of the two spherical surfaces of the lens can be calculated using Lensmaker's equation

$$(n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right) = \frac{1}{f'}, \quad (2.7)$$

where n stands for the refractive index of the lens material. R_1 and R_2 are the radii of the two spherical surfaces. The focal length f' is 50 mm. Let us use a common optical material such as BK7 glass. It has a refractive index (n) of approximately 1.5 in the visible spectrum. For simplicity, we use 1.5 as our refractive index value.

Now in Equation 2.7, the two radii can still vary and we do not have a unique set of values to satisfy the equation. We can choose to have a double-convex lens by forcing the absolute radii to be identical, where $R_1 = 50$ mm, and $R_2 = -50$ mm.

Equation 2.7 assumes the thickness of the lens is very small compared to the radius of the two surfaces (a thin lens scenario). There is also the equation when the thickness d of the lens is considered:

$$(n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} + \frac{(n - 1)d}{nR_1R_2} \right) = \frac{1}{f'}. \quad (2.8)$$

With a thickness d of 3 mm, the radii can be now calculated using Equation 2.8 to be $R_1 = 49.50$ mm and $R_2 = -49.50$ mm. These values are very close to the case when using thin lens approximation (50 mm and -50 mm). At this phase, the above derivation shows that based on first order (Gaussian) optics, the following specification can achieve a magnification factor of $6\times$, given that the object locates 41.67 mm from the lens:

Table 2.1: Magnifying Lens Specification

Magnification	$6\times$
Number of lens	1
Diameter (mm)	20
Refractive index	1.5
Thickness (mm)	3
Radii (mm)	49.50, -49.50
Effective Focal Length (EFL, mm)	50

For the aforementioned example, we use the most basic Lensmaker's formula to design a $6\times$ magnifier with a double-convex lens. The condition of double-convex is chosen by us since according to Equation 2.8, there are infinite choices of radii that satisfy

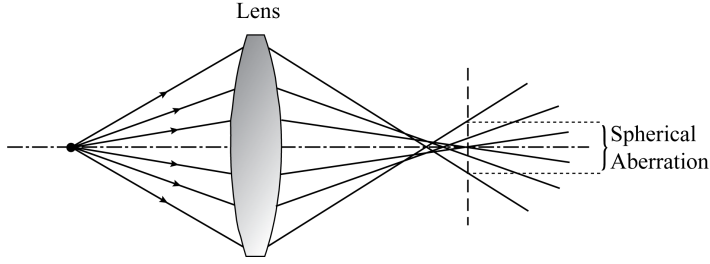


Figure 2.2: Illustration of the spherical aberration in the optical system. The light rays emitting from one on-axis object point is not converging perfectly into one on-axis point on the image plane. The deviation from the expected on-axis point is called spherical aberration.

the equation. Note that the Lensmaker's formula is based on first-order optics, where the angle between the rays and the optical axis is assumed to be very small. This assumption no longer holds true when optical aberration needs to be treated. In the following part, we show an example of how to improve the magnifier design considering minimizing the spherical aberration. An illustration of spherical aberration is given in Figure 2.2. Light rays emit from an on-axis point and their deviation from the expected refracted¹ on-axis point is called spherical aberration.

The primary monochromatic aberrations are the spherical aberration, coma, astigmatism, field curvature and distortion. There are axial and lateral chromatic aberrations when chromatic application is considered. Given the thin lens model, the coefficient of each aberration term can be explicitly expressed. The term of the spherical aberration can be expressed as ([3], 29.4.7):

$$S_I = \frac{h^4}{4f'^3} \left(\frac{3n+2}{n} M^2 - \frac{4(n+1)}{(n-1)n} XM + \frac{n+2}{(n-1)^2 n} X^2 + \frac{n^2}{(n-1)^2} \right). \quad (2.9)$$

where h indicates the ray height on the lens surface. X stands for the bending factor of the thin lens ([3], 10.1.1):

$$X = \frac{R_1 + R_2}{R_2 - R_1}, \quad (2.10)$$

and M stands for the position parameter ([3], 10.1.2):

$$M = \frac{1+m}{1-m}, \quad (2.11)$$

m is the magnification factor and it is 6 as defined previously. In this case, $M = -\frac{7}{5}$. In Equation 2.9, for a chosen ray, parameter f' , M , and n are all given. Therefore, S_I becomes a quadratic function for the bending factor X . There is a value of X that minimize the value of S_I . In this case, when $X = -1$, the minimum value of S_I is $8.16(\frac{h^4}{4(f')^3})$. The spherical aberration of each ray is proportional to its ray height given the fixed lens configuration.

¹For a mirror, it is reflected.

Now if we consider both the Lensmaker's Equation 2.7 and the spherical term in Equation 2.9, the two radii of the lens, which minimize the spherical aberration, can be solved as $R_1 = \infty, R_2 = -25$ mm. These two values are very different from the values (50 mm and -50 mm) calculated based on first-order optics. It is worth mentioning that Equation 2.9 assumes a thin lens, hence the comparison of the radii is done between thin lens models. It is also observed that in our magnifier example, considering a 3 mm lens thickness in Lensmaker's formula does not alter the radii significantly.

When more aberration terms are considered, the number of variables in this example is not enough to analytically solve the values for the minimal values of each aberration term. Normally, new degrees of freedom will be introduced to the system (e.g. adding new lenses). However, more numbers of variables also implies an increasing complexity of the formulae. Modern optical design handles it by using computer-aided design techniques. Numerical optimization and rapid ray tracing are done by computers, which improves the efficiency of optical design.

2.2. MODERN OPTICAL DESIGN

Prior to the development of digital computers, lens optimization was a hand-calculation task using trigonometric and logarithmic tables to plot Two-Dimensional (2D) cuts through the multi-dimensional space. The development of the digital computer in the 20th century has drastically shaped the way optical design is being practiced. The burden of calculation is tremendously reduced by using computers. Instead of calculating each ray path on paper, optical systems are parameterized into different variables representing the physical construction of an optical system (a physical model). For example, a simple two-element air-spaced spherical lens can be represented with nine variables (four radii of curvature, two thickness one airspace, and two glass types). With increasing complexity (number of elements, type of surfaces) of the system, the number of variable can reach hundreds. Using modern computers allows rapid ray tracing calculation with which the performance of the lens system can be quickly modelled and evaluated.

Helped by the technique of numerical optimization, the vast high-dimensional design space can be searched to look for the set of variables providing the optimal performance of the system. Lens optimization was studied as early as the 1940s [4]. The earliest attempt for optimizing a doublet using computers dates back to the 1950s, when James G. Baker used the *Harvard Card Programmed Calculators* [5]. A good description of the start of computer-aided optical design is given in a 1963 paper by Feder [6]. In order to optimize an optical system using the computer, it is necessary to construct a merit function. The value of this merit function predicts the performance of the system. Normally, the smaller the value of the merit function, the better the performance should be. For an imaging system, such a merit function can be the sum of the aberrations, the size of the focused spot, etc. The merit function in optical design is commonly defined as [7]:

$$MF(\mathbf{x}) = \sum_{j=1}^m w_j [\tilde{f}_j(\mathbf{x}) - \tilde{f}_{tar,j}]^2, \quad (2.12)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_N)$ is a vector describing a point in the N -dimensional variable

space. $\tilde{f}_j(\mathbf{x})$ are the operands with target values $\tilde{f}_{tar,j}$. $\tilde{f}_j(\mathbf{x}) - \tilde{f}_{tar,j}$ could be defined as elementary aberration such as ray or wavefront deviations for selected rays. w_j are the positive weighting factors, therefore MF is a single number giving the residual of various operands and their target values. The weights and targets can be absorbed in the definition of the operands, therefore we have

$$MF(\mathbf{x}) = \mathbf{f}^T(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}), \quad (2.13)$$

where \mathbf{f} is a vector having the operands as components.

Given the merit function, a high dimensional space is determined. Each sets of value of the variables represents a point in this high dimensional space and the value of the merit function is determined at that point. Ideally, the value of the merit function is directly associated with the performance of the optical system being designed. The minimal value of the merit function is desired because it represents the best system performance.

Once the variables are chosen from the system, the next step is to apply numerical optimization techniques to minimize the value of the merit function defined in Equation 2.12:

$$\begin{aligned} &\text{minimize } MF(\mathbf{x}), \\ &\text{subject to some constraints,} \end{aligned} \quad (2.14)$$

with optimization, a new set of values of the variables \mathbf{x}_{opt} is obtained such that the value of the merit function is a minimal.

As shown in the example of the magnifier, the requirements of working distance and magnification have to be fulfilled. At the same time, the spherical aberration for a thin lens model has to be minimized. To satisfy these conditions, a set of radii of the optical surfaces has been calculated. However, in practice, spherical aberration is not the only aberration contributor that limits the image performance. At the same time, some thickness of the lens needs to be added for practical reasons. Given the previously calculated radii and other parameters of the magnifier, a merit function value can be determined after the lens is modeled in an optical design software. This merit function value can usually be further reduced for better system performance with optimization techniques.

We applied optimization for the previous magnifier example in optical design software CODE V. Three different cases are parameterized in CODE V as starting points for the optimization. They are calculated using the Lensmaker's formula (Equation 2.7), Lensmaker's formula considering a thick lens (Equation 2.8), and the spherical aberration formula for a thin lens (Equation 2.9). The variables are the two radii of the surfaces. The thickness is also used as a constrained variable to optimize for a thick lens configuration. The radii and merit function values in each step of the optimization of the three cases are listed in Table 2.2.

In CODE V, the object distance is set as 41.67 mm and the image distance is at 250 mm. An EFL of 50 mm is used as a constraint of the system. Only one wavelength is used and the material refractive index is set as 1.5. An entrance pupil diameter of 20 mm is used and the fields are defined with three object heights: 0, 3, and 5 mm. The merit function used here is the default transverse-ray function in CODE V. It is a composite value, scaled so that it is the mean square of the weighted image radius.

When optimized directly with the input setting, the starting points of both thin lens approximation using paraxial optics and the one optimized for spherical aberration converge to the same local minimum (radii are given in the "Optimized thin" column in Table 2.2). We also notice that the starting point for the thin lens which is optimized for spherical aberration already has very similar radii compared to the CODE V optimized system. It indicates that for such a configuration, the image quality is dominated by the spherical aberration and the theoretical calculation predicts the image quality well. However, it is also noticed that the merit function value of these systems is very big ($9.06E5 \mu\text{m}^2$). Here we use this example only to demonstrate how a simple lens design and optimization flow works. In practical design, the merit function value needs to be further reduced and additional degrees of freedoms (such as new lenses) are required.

We also optimized the system with a lens thickness constrained at 3 mm. Three starting points all converge to the same local minimum. It indicates that there is no difference in which starting configuration we choose. We will show later in the thesis that it is not the case for more complicated optical systems (e.g. with more lenses).

Table 2.2: Optimization results of the magnifying glass in CODE V

	Starting point			Optimized thin			Optimized thick		
	Radii (mm)	THK (mm)	MF (μm^2)	Radii (mm)	THK (mm)	MF (μm^2)	Radii (mm)	THK (mm)	MF (μm^2)
Thin lens	50, -50	0	3.56E6	2247, -25	0	9.06E5	108, -32	3	4.98E6
Thick lens	49.50, -49.50	3	6.45E6	/	/	/			
Thin lens optimized for SA	$\infty, -25$	0	9.07E5	2247, -25	0	9.06E5			

THK: thickness. SA: spherical aberration.

In summary, the process of the modern optical lens design can be simplified into two steps: 1) determine parameterized configurations of the system as starting points; 2) use optimization tools to optimize the system performance. In practice, the two steps can be iterated many times to finally obtain a satisfactory system. Mastering both steps requires a good understanding of the optical systems as well as in-depth knowledge of optimization. We will explain these two parts in the following sections, with an emphasis on the optimization techniques.

2.2.1. OPTICAL SYSTEM OPTIMIZATION

When the optical system and its merit function are treated as an optimization problem, there are two kinds of situations: 1) the merit function is a *convex* function² and there is only one global minimum; 2) the merit function is a *non-convex* function and there are multiple local minima.

Simple cases such as the Magnifier example, the spherical aberration is a quadratic function of the bending factor. The given number of equations equals the number of

²Geometrically, a function is *convex* if a line segment drawn from any point $(\mathbf{x}, f(\mathbf{x}))$ to another point $(\mathbf{y}, f(\mathbf{y}))$ – called the *chord* from \mathbf{x} to \mathbf{y} – lies on or above the graph of f .

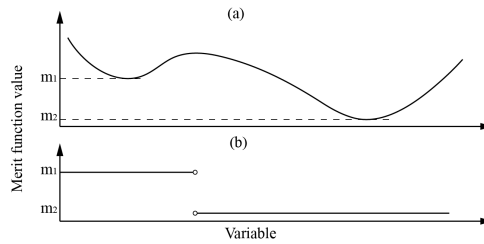


Figure 2.3: Illustration of basins of attraction in a one-dimensional optimization space. (a) The merit function with two local minima having value m_1 and m_2 . (b) The corresponding basins of attraction in the plotted merit function region. At the boundary of the two basins, it is unclear where the optimization leads to, therefore marked as circles.

variables and a unique solution of the parameterized system can be obtained. However, in optical design, the common situation is that the number of equations (resulting either from aberration coefficients or ray aberrations) is larger than the number of variables³. In this case, an approximation of the solution can be given. This is achieved by local optimization, where one minimum can be acquired.

It is also observed that when the number of the variables (e.g. the number of lenses) increases, the number of the solutions also increases⁴. This is because the merit function becomes *non-convex* as the variable increases. The results of the local optimization is then sensitive to the initial point where optimization starts.

BASINS OF ATTRACTION

In an optimization problem, the basin of attraction defines a set of points in the optimization space, where the optimization leads to the same local minimum. An example of a one-dimensional merit function and its corresponding basins of attraction is given in Figure. 2.3.

The illustration of the basin of attraction is a direct way of visualizing optimization results in the optimization space, especially in a non-convex problem⁵. The result of a local optimizer is determined by its starting point. Since the goal of this problem is to find the global minimum among the many, there are two consequent strategies: 1) starting from the basin of attraction that corresponds to the global minimum. It requires choosing the starting point based on some prior knowledge. 2) searching for all (most of) the basins of attraction, list all the minima and obtain the global minimum.

In optical design, the former strategy is largely associated with the conventional optical design strategies. The starting point can be determined from the literature and patent data, where the system is believed to have a larger chance to lead to a satisfactory min-

³For imaging systems, each surface is described by a finite (often small) number of parameters. If one thinks in terms of aberration series, the number of aberration coefficients is in principle infinite, while the number of variables of imaging systems is finite. In practice, where ray tracing is used, the number of rays required for an accurate modelling of the image quality is typically larger than the number of design variables.

⁴A good example is illustrated in Page 61 of Reference [8]. The contour plot of the merit function of a doublet with respect to two variables shows four local minima.

⁵In Chapter 4, basins of attraction on the two-dimensional hyper-plane are used for demonstration purposes.

imum. Another approach is to compute the system parameters based on an analytical model. The assumption is that the starting point provided by the analytical model is in the basin of attraction of the global minimum. However, when the optical system becomes increasingly complicated, e.g. adding lenses, making surfaces aspheric, etc., it is very difficult to fully analyze the system and provide a starting point. The strategy of finding multiple minima and selecting the best among them becomes preferred in this situation.

The most straightforward method to search for the global minimum is to systematically evaluate the merit function value on the multi-dimensional grid. The problem of this method is that the computational time increases exponentially with the number of variables. By applying a local optimizer, the amount of merit function evaluations can be reduced when starting from a basin of attraction. The task of finding the global minimum is then transformed to finding the basin of attraction containing the global minimum.

2.2.2. GETTING THE STARTING POINT

When multiple minima are presented in the optical design space, the result of the local optimization is sensitive to its choice of the starting point. In the conventional optical design practice, large effort is spent on obtaining a good starting point.

As demonstrated in the beginning of this chapter, simple systems such as the Magnifier can be computed using first-order optics and a third-order aberration model to obtain a relative good starting point. For more complicated systems which contain more optical elements and difficult requirement, there are design strategies summarized by lens designers based on their in-depth knowledge and extensive experience [9][10]. However, mastering these design methods can require years of practice.

A common practice is that optical designers scan through the existing catalog of the available designs. This can be via patent search or by looking up examples in optical design books [11] [12]. The one with the closest specification with respect to the design requirement is chosen to serve as the starting point of the design. The assumption is that this chosen system is already in the basin of attraction where good system performance is expected. By doing some fine adjustment, the preferred system should be achievable.

Methods, such as the Simultaneous Multiple Surface (SMS) method [13] which obtains a starting point based on iterative ray tracing process, have also been getting much attention recently. The SMS method can be very effective for obtaining starting points consisting of aspherical lenses. In Chapter 6, we will look at the SMS method and the corresponding design landscape.

2.2.3. LOCAL OPTIMIZATION

In this section, we briefly describe the basic concepts and methods used for local optimization. We will elaborate a bit more on the damped least-square method as it is commonly applied for lens design optimization.

Given the starting point and the merit function, the local optimizer aims to find a local minimum (minimization problem) within a basin of attraction. Numerically, this is achieved by iterative steps. In the N -dimensional space, the local optimization process

consists of moving from the starting point towards the minimum in several steps. The value of the merit function reduces after each step, until a minimum is reached. When close enough to a minimum, further iteration will not produce any significant changes in the system parameters and the process is called convergent.

GENERAL PROBLEM STATEMENT

Given the j -th component, $f_j(\mathbf{x})$, of the operands vector \mathbf{f} can be written as the following formula with a Taylor expansion

$$f_j(\mathbf{x}_0 + \Delta\mathbf{x}) = f_j(\mathbf{x}_0) + \Delta\mathbf{x}^T \cdot \nabla f_j(\mathbf{x}_0) + \frac{1}{2} \Delta\mathbf{x}^T \cdot H_j \cdot \Delta\mathbf{x}, \quad (2.15)$$

where $\Delta\mathbf{x}$ describes the size and direction at a certain optimization iteration. H_j is the Hessian matrix of f_j at $\mathbf{x} = \mathbf{x}_0$. Its elements are

$$(H_j^N) = \begin{bmatrix} \frac{\partial^2 f_j}{\partial x_{0,1}^2} & \cdots & \frac{\partial^2 f_j}{\partial x_{0,1} \partial x_{0,N}} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f_j}{\partial x_{0,N} \partial x_{0,1}} & \cdots & \frac{\partial^2 f_j}{\partial x_{0,N}^2} \end{bmatrix}.$$

The merit function in Equation 2.13 can be written at \mathbf{x} in the vicinity of \mathbf{x}_0 using Taylor expansion :

$$MF(\mathbf{x}) = \mathbf{f}_0^T \cdot \mathbf{f}_0 + 2\Delta\mathbf{x}^T \cdot \mathbf{J}^T \cdot \mathbf{f}_0 + \frac{1}{2} \Delta\mathbf{x}^T \cdot \mathbf{H} \cdot \Delta\mathbf{x}, \quad (2.16)$$

where $\mathbf{f}_0 = \mathbf{f}(\mathbf{x}_0)$, $\mathbf{J} = \nabla \mathbf{f}(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0}$, which is the Jacobian matrix of \mathbf{f} at $\mathbf{x} = \mathbf{x}_0$ (with elements: $J_{ij} = \frac{\partial f_i}{\partial x_j}|_{x_j=x_{j,0}}$), $\mathbf{H} = 2(\mathbf{J}^T \cdot \mathbf{J} + \sum_{j=1}^m f_j(\mathbf{x}_0) \cdot H_j)$, which is the Hessian matrix of MF at $\mathbf{x} = \mathbf{x}_0$. To find the minimum of $MF(\mathbf{x})$, numerical iteration steps are executed. The gradient of MF with respect to the optimization variables vanishes

$$\nabla MF(\mathbf{x}) = 2\mathbf{J}^T \cdot \mathbf{f}_0 + \mathbf{H} \cdot \Delta\mathbf{x} = 0. \quad (2.17)$$

Different methods are used to solve these equations. We mention here the most common ones.

NEWTONIAN METHOD

The Newtonian methods solve Equation 2.17. The computation of the Hessian matrix \mathbf{H} is expensive. As a result, approximation of the Hessian matrix is usually used. The Gauss-Newton method is a common method, where the H is approximated by $2\mathbf{J}^T \cdot \mathbf{J}$. The Equation 2.17 can then be written as

$$\begin{aligned} 2\mathbf{J}^T \cdot \mathbf{f}_0 + 2\mathbf{J}^T \cdot \mathbf{J} \cdot \Delta\mathbf{x} &= 0, \text{ then} \\ \Delta\mathbf{x} &= -(\mathbf{J}^T \cdot \mathbf{J})^{-1} \cdot \mathbf{J}^T \cdot \mathbf{f}_0, \end{aligned} \quad (2.18)$$

where Δx is the steps used for iteration calculation. The advantage of the Newtonian method is that it converges fast when the operands are linear (in the vicinity of a local minimum). However, when starting far from the local minimum, the operands are usually very nonlinear. In such cases, it may be difficult for the algorithm to find the vicinity of a local minimum.

STEEPEST DESCENT

The method of steepest descent ignores the second order term and only uses the first order term for Equation 2.16. The step for the numerical optimization is defined with a negative direction towards the gradient of the function. If we start the optimization at \mathbf{x}_0 , the step can be written as

$$\Delta x = -t \cdot \nabla MF(\mathbf{x}_0), \quad \mathbf{x} = \mathbf{x}_0 + \Delta x,$$

where $t > 0$, represents the size of the step. We define a function of t as

$$\phi(t) = MF(\mathbf{x}_0 - t \cdot \nabla MF(\mathbf{x}_0)). \quad (2.19)$$

The derivative of $\phi(t)$ at $t = 0$ is

$$\phi'(0) = -\|\nabla MF(\mathbf{x}_0)\| < 0. \quad (2.20)$$

Given $MF(\mathbf{x})$ is continuous differentiable, for $t > 0$, we have $\phi(t) < \phi(0)$. Hence, we have

$$MF(\mathbf{x}_0) > MF(\mathbf{x}_0 + \Delta x). \quad (2.21)$$

That is, the method of steepest descent is guaranteed to make at least some progress toward a minimized function value during each iteration. However, when in the vicinity of the minima, the steepest descent converges slower than the Newtonian method.

METHOD OF DAMPED LEAST-SQUARE

The method of damped least-square is the one commonly used as a local optimizer in optical design. It has a behavior between Gauss-Newton and steepest descent method. Instead of solving Equation 2.18, it solves for

$$\begin{aligned} 2\mathbf{J}^T \cdot \mathbf{f}_0 + 2(\mathbf{J}^T \cdot \mathbf{J} + \lambda \mathbf{I}) \cdot \Delta x &= 0, \\ \Delta x &= -(\mathbf{J}^T \cdot \mathbf{J} + \lambda \mathbf{I})^{-1} \cdot \mathbf{J}^T \cdot \mathbf{f}_0, \end{aligned} \quad (2.22)$$

where λ is non-negative damping factor, \mathbf{I} is identity matrix. If λ is regarded as an independent variable, the angle between Δx and $\nabla MF(\mathbf{x}_0)$ (equals $-2 \cdot \mathbf{J}^T \cdot \mathbf{f}_0$) is a monotonically decreasing function of λ . With the λ goes to infinity, the angle goes to zero. By adjusting the value of λ , the method of damped least-square has a behaviour between the steepest descent method and the Gauss-Newton method: when $\lambda \rightarrow \infty$, the method acts like the steepest descent method. It converges slowly to the local minimum even if

starts far from it. When $\lambda = 0$, the method acts like the Gauss-Newton method. It converges rapidly in the vicinity of the local minimum.

In lens design programs, the identity matrix I in Equation 2.22 is replaced with a diagonal matrix Q , with its diagonal elements scaling different variables (e.g. curvatures and refractive index). As a result, the convergence of the optimization become faster [14]. The implementation of Equation 2.22 is usually further refined by using the method of singular value decomposition, where the matrix J that must be inverted, is written as a product of three matrices. One of them is the matrix of singular values. Removing the contribution of certain singular values ensures that the algorithm remains stable and that parameter changes, which are too large, can be avoided [7].

All major commercial lens design programs (CODE V, Zemax, OSLO, SYNOPSIS) use as their main local optimization engine some version of the damped least-squares method. Equation 2.22 gives the basic idea, but the implementation details vary greatly between programs and are kept confidential by the developers. For instance, the damping is implemented in all cases in a much more sophisticated way, that includes empirical information based on lens design experience. Depending on the specific design problem, because of the different choices they make, some programs perform more satisfactorily than others.

This “black-box” character of software has also been a significant practical drawback for the present research which is dependent on these commercial ray tracing software. For instance, CODE V and Zemax do not allow the user to control the damping. The damping is chosen automatically as low as possible in order to achieve a fast convergence. However, earlier research at TU Delft has shown that low damping leads in many lens design programs to fractal basins of attraction, which make the optimization process unstable [15]. It is suspected that also for some of the challenges encountered in the present research (e.g. in Section 5.2.2) a much higher damping value would be desirable. Therefore it is necessary to continue the present investigations with software where the algorithms can be adapted as needed.

2.2.4. GLOBAL OPTIMIZATION

When multiple types of parameters, such as curvature, thickness, glass types, etc., are considered as variables in an optical system, the complexity of the multi-dimensional merit function increases. Especially when the number of lens elements increases, multiple local minima are expected to exist and the results of the optimization strongly depends on the choice of the starting point. For simple systems, first-order calculation and aberration analysis can be done to determine a starting point. The result is usually effective. For systems with more complexities (e.g. more than 5 elements with all curvatures and thickness used as variables), conventional practice tends to start from an existing system which is relatively good and has similar specifications to the design requirement. The database which is used for searching for these starting point varies from designer to designer. The definition of the “similarity” between the chosen starting point and the design requirement is normally determined by the designer’s experience. The consequence of such complexity and uncertainty is that the design process can often be trapped in a sub-optimal minimum.

Growing attention has been given in the field of global optimization method following the increase of the computational power of computers. Instead of following the topography to converge to a minimum, global optimization methods apply strategies to cover multiple basins of attraction and attempt to reach to the global minimum. Some of the global optimization methods have been applied to optical design problems and show promising results. In this thesis, we have been using Global Synthesis from CODE V [16][17], which is proven very successful among the commercial lens design programs. Unfortunately, it uses proprietary algorithms which are not disclosed in publications. In the present work, the results obtained with our methods have been compared only with those of Global Synthesis. To make the difference between other methods and ours clear, a few other global optimization methods with known principles will be briefly described below. On high level, they can be divided into two categories: 1) methods that apply strategies to move to a different basin of attraction when trapped. 2) methods that start from multiple basins of attraction, and then use mutual information to converge to the optimal location.

SIMULATED ANNEALING

Simulated annealing is a stochastic approach to the solution of complex problems and it is essentially a search method driven by biased random walk [18]. It is inspired by the thermodynamics and the configuration of an alloy during cooling. When applied to optical design, the method is based on the idea that a given optical system can be thought of as being in some energy state. This energy state is lower when the system is "better (i.e. the merit function has a low value)", and is higher when the lens is "worse (i.e. the merit function value is high)". Different from the iterative method, such as the method of damp least-square presented in the previous section, the step from the starting point is generated randomly. With the merit function value at the starting point \mathbf{x}_0 is MF_0 and at point $\mathbf{x}_0 + \Delta\mathbf{x}_r$ is $MF_0 + \Delta MF$, the new variable set is accepted based on a probability function [19] given by

$$P(MF_0 + \Delta MF) = \begin{cases} 1, & \Delta MF < 0, \\ e^{-\frac{\Delta MF}{T}}, & \Delta MF > 0, \end{cases} \quad (2.23)$$

where T refers to the *temperature* for the annealing and can be adjusted to tune the probability for accepting the proposed step.

The method accepts location with a probability where the merit function value increases. In adaptive simulated annealing methods, the acceptance probability is changed during the optimization process. When maximizing searching space is prioritised, a large increase in the merit function value is accepted. When further reduction of the merit function value is prioritised, the probability is changed such that large increase in the merit function value is unlikely to happen [19]. In practice, trial and error is required to set the optimal parameters for the algorithm to function as expected.

ESCAPE FUNCTION

The method of escape function, as the name suggests, is to escape from the stagnation of a local minimum. It is implemented in the optical design software OSLO [20] as a

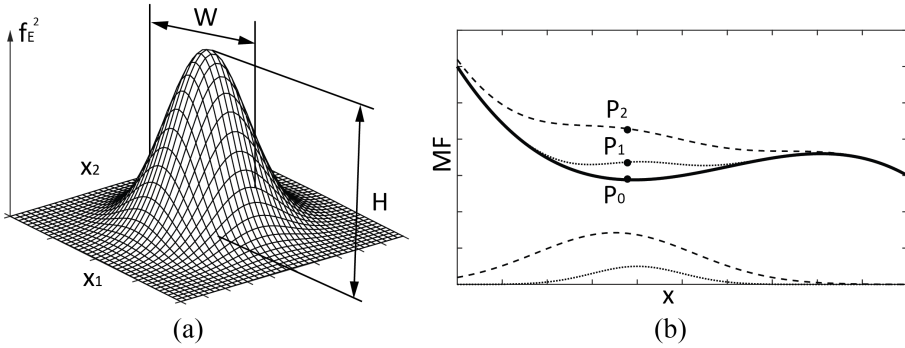


Figure 2.4: Explanation of the method of escape function. (a) a 2D escape function illustration, where H modifies the height and W modifies the width of the function. (b) P_0 is a minimum for the original 1-D merit function (solid line). By applying an escape function (dotted line), P_0 is lifted to P_1 . However, it cannot help the optimizer to escape to a new basin of attraction. When applying a different escape function (dashed line), P_0 is lifted to P_2 . The optimizer escapes to the next basin of attraction.

global optimizer. In addition to the merit function, the method uses an escape function to modify the landscape. The escape is given by [21],

$$f_E = \sqrt{H} \cdot \exp\left\{-\frac{1}{2W^2} \sum_j [\mu_j (x_j - x_{jL})]^2\right\}, \quad (2.24)$$

where x_j is the j -th variable in the merit function, x_{jL} is its value at the obtained local minimum, and μ_j is the scale factor for the j -th variable. The escape function has a form of a multi-dimensional Gaussian function, where H modifies the height and W modifies the width of the function (Figure 2.4). Figure 2.4(b) uses a one-dimensional case to illustrate the escape function with two different settings. The solid line represents the original one-dimensional merit function. After local optimization routine, the minimum P_0 is obtained. In order to escape from the current basin of attraction, the first escape function setting lifts P_0 to P_1 , as indicated by the dotted line. However, the optimizer remains in the same basin of attraction due to an insufficient modification of the local landscape. The second escape function setting lifts P_0 to P_2 . When local optimization routine is applied for P_2 , the optimizer escapes from the original basin of attraction. In practice, an automatic feedback loop is designed to adjust the escape function settings to increase the efficiency for minimum searching [21].

GENETIC ALGORITHMS

Genetic algorithms are inspired by the process of natural selection which belongs to the larger class of an evolutionary algorithm. Similar to biological reproduction, genetic algorithms define population in the optimization landscape. Crossover and mutation occur among the population through generations. The fitness of survival of a given individual is determined within the population. Individuals that have a high fitness will be selected by the algorithm.

In the context of optical design and optimization, the merit function value of an optical system configuration is equivalent to its fitness of survival. Solutions to optical sys-

tems are thought of as particular individuals of a hypothetical form of life [22]. The continuous variables of the optical system are encoded as bit strings (e.g. of 0 and 1). An entire optical system is encoded with these binary numbers, seen as a “chromosome” in which each variable is considered a “gene”.

A genetic algorithm typically starts with a population of a certain number (e.g. 50) of “chromosomes”, generated randomly either from entire search space, or from regions where good solutions are expected to be found based on previously existing knowledge.

In crossover, from two “parents” A and B, a “child” C is generated e.g. by splitting the two “chromosomes” at the same position and by taking in C e.g. the first part of A and the second part of B. The “child” then shares many of the characteristics of its “parents”. A mutation would happen e.g. when one or more single “genes” of a “chromosome” X are multiplied by a random factor.

Most such changes lead to a performance that is worse than that of A, B and X. But when an improvement is detected, then the resulting system replaces a system with poor performance in the population.

Evolution happens in generations. The population of candidate solutions evolves toward better solutions, because in each generation good solutions are more likely to be selected for the next generation. However, it is necessary to occasionally select poorer solutions as well, but with a lower probability than the good ones.

When the different parameters for mutation, crossover, and other operations are chosen adequately, the average merit function for the population decreases from generation to generation. Unfortunately, in some applications in lens design the quality of the results is critically dependent on a good choice of these parameters.

Different from the method of simulated annealing or escape function, genetic algorithms begin with sampling a larger range of the optimization landscape. Information in different parts of the optimization landscape are exchanged to search for the global optimum. However, because of its combinatorial nature, the convergence to a minimum in the continuum variable space is not efficient. Hybrid algorithms where genetic algorithms are combined with a local optimizer (e.g. damped-least square), are thus recommended to improve the efficiency [22]. Global optimization in the commercial lens design software Zemax uses a combination of genetic algorithm, damped least-square, and some other expert system heuristics to search for new design solutions [23].

PARTICLE SWARM OPTIMIZATION

Particle swarm optimization was introduced as an analogue for the interaction of individuals in a swarm or a flock of birds. A collection of candidate solutions, called particles, move around in the optimization space according to simple mathematical rules. The goodness of a particle's location is given by the value of the merit function. For each particle, three vectors are used to represent it in the optimization space: the current location of the particle, the velocity of the particle, and its previous best position. During each iteration, every particle moves from its original location to a new location considering the velocity vector and the best location vectors. After sufficient iterations, the swarm as a whole is likely to move close to an optimum in the optimization landscape [24].

Similar to the genetic algorithm, particle swarm optimization starts with initializing a population randomly located in the optimization space. Each particle starts with an

initial velocity. The position vector and velocity vector of the particles are updated after each iteration as the particles move. Parameters associated with the particle movement such as initial velocity, movement inertia, and acceleration with respect to the global best positions, need to be determined based on trial and error.

As for optical design, each system configuration represents a particle in the optimization space. Particle swarm optimization has been demonstrated in several design cases and it is especially helpful in handling a large number of variables such as freeform systems [24]. On the other hand, a hybrid algorithm of combining particle swarm optimization with local optimizer has been proposed [25]: Particle swarm optimization is used for the discrete optimization of the optical material choices, while the local optimizer is applied to obtain the minimum given the selected material choice.

2.2.5. SADDLE POINT METHOD

In addition to minimum and maximum, a saddle point is also a type of stationary (or critical) point existing in the optimization landscape. Gradient vanishes at all stationary points. When computing the eigenvalues of the Hessian matrix of these stationary points, a minimum has its signs of the eigenvalues all positive, a maximum has its signs of the eigenvalues all negative, and the saddle point has both positive and negative signs of its eigenvalues. Morse Index is defined as the number of the negative signs of the eigenvalues, and the Morse Index of a saddle point is always larger than zero. For example, if there is one negative eigenvalue of the Hessian matrix for a stationary point, it is then a saddle point with a Morse Index value of 1. For an N -dimensional optimization landscape, the Morse Index value of its saddle points can vary from 1 to $N-1$. The value of the Morse Index indicates the number of directions, following the eigenvectors, along which the merit function value decreases. An example of a saddle point in the 2-dimensional landscape is given in Figure 2.5(a).

Figure 2.5(b) shows an example of a 2-dimensional landscape. If we use an analogue of a mountain view scenario, minima are the *bottom* of the valleys and maxima are the *top* of the hills. Saddle points are located on the passes between the valleys.

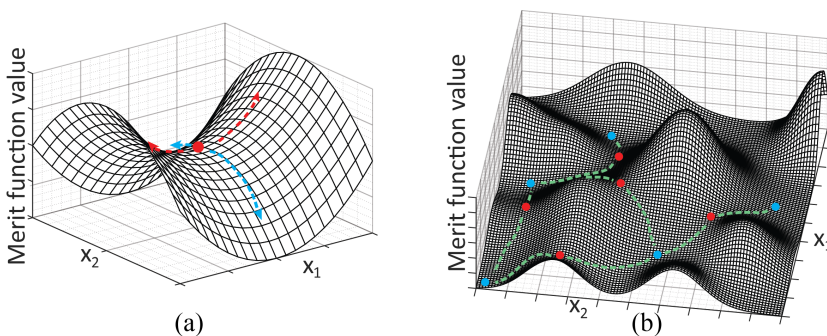


Figure 2.5: Illustration of the 2D saddle points. (a) A 2D saddle point (horse saddle, with a Morse Index value of 1). It is a maximum in one direction and a minimum along the other direction. (b) Example of a 2D landscape where minima are connected via saddle points.

The aforementioned global optimization methods treat the optimization landscape

as different basins of attraction. When a larger number of basins are covered, the chance of obtaining a global solution increases. Methods such as simulated annealing and escape function apply strategies to switch basins of attraction when a local minimum is encountered. Methods such as genetic algorithm and particle swarm start with a population distributed in the optimization space. Strategies were applied to search for the basin with the best fit minimum. As a result, the task of searching for local minima can be then replaced by the task of searching for basins of attraction. The saddle point is of special interest when the task is to search for different basins of attraction. As indicated in Figure 2.5(b), saddle points with a Morse Index of 1 are located at the boundaries between different basins. Once such a saddle point is found, two basins of attraction can be reached afterwards.

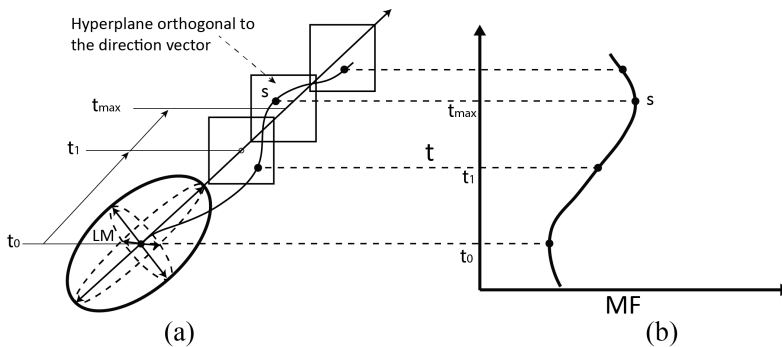


Figure 2.6: Illustration of constrained optimization to detect saddle points in the optimization landscape. (a) Saddle point detection in high-dimensional space. Minimization is performed on hyperplanes which are orthogonal to the chosen direction. (b) Example of MF value along a search direction. The saddle point detection starts from an existing local minimum. The directions around the local minimum are chosen towards the directions of the eigen vectors of the Hessian matrix. The value of the MF will first increase. After finding the saddle point, the MF value will decrease.

SADDLE POINT DETECTION

NETwork of Local MINima (NETMIN), a tool developed in Delft University of Technology, detects saddle points with Morse Index 1 in the optimization landscape. NETMIN uses constrained local minimization to search saddle points starting from a local minimum. As illustrated in Figure 2.6, a set of directions is defined around a local minimum. These directions can be determined by the eigenvectors of the Hessian matrix, which are computed at the local minimum. Around the local minimum, the surfaces along which the merit function is constant have the shape of ellipsoids [26]. The eigenvectors of the Hessian matrix correspond to the directions of the half-axes of the ellipsoids. Along each direction, a set of hyperplanes orthogonal to the direction vector, can be defined. For an N -dimensional space, these hyperplanes have a dimension of $N - 1$. The distance between the local minimum and the hyperplane is indicated by t . On each hyperplane, a constrained minimization can be performed. The positions of the constrained minima as a function of t is symbolically shown in Figure 2.6(a). The merit function value

(MF) as a function of t is shown in Figure 2.6(b). When t is sufficiently small, the merit function value of the constrained minimum on the hyperplane increases. At a certain $t = t_{max}$, the merit function value reaches its maximum value (point s in Figure 2.6[a]). This point is a maximum along the vector direction and a minimum on the $(N - 1)$ -dimensional hyperplane, hence a saddle point with Morse Index 1.

When a saddle point with Morse Index 1 is detected, a point "on the other side" of the saddle point can be chosen to start a local optimization routine. The optimizer searches in a different basin of attraction to find a new local minimum. NETMIN combines local optimizer and saddle point detection algorithms to reveal the network of local minima and saddle points in the optimization landscape. NETMIN has been used to study triplet network ([27]) and six-mirror systems ([26]). In Chapter 4 of this thesis, we also use NETMIN obtained results as a reference.

The computation of NETMIN is expensive. For each step along the chosen direction, a constrained optimization has to be performed. In practice, multiple directions have to be searched to find all saddle points around a local minimum. Parallel processing can be enabled to facilitate the search.

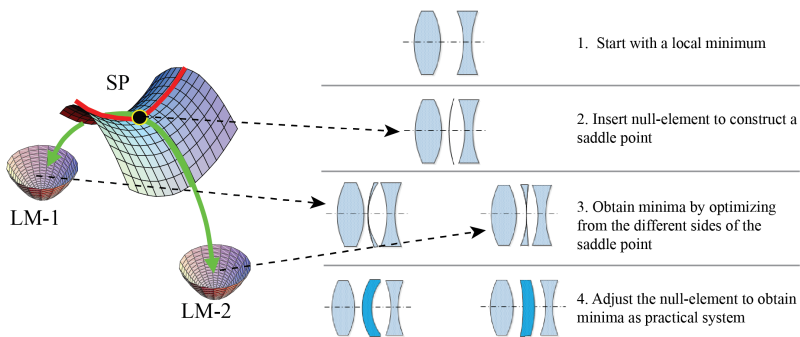


Figure 2.7: Illustration of the saddle point construction process. The right side shows an example of constructing a saddle point on a doublet minimum. Two triplet minima result from this process. The left is an example of how a saddle point leads to two minima in a 2D optimization space. In steps 2 and 3, the optimization space remains the same. In steps 1 and 2, the optimization space has different dimensions (step 1 does not include the null-element, and step 4 adds extra variables).

SADDLE POINT CONSTRUCTION

In the context of lens design, research has shown that in most cases saddle points can be constructed directly instead of performing a saddle point detection algorithm [8] [28]. It was firstly found that the saddle point systems obtained via the saddle point detection algorithm always contain low power meniscus lenses or air spaces. Once the thickness of these meniscus are reduced to zero, it is possible to show theoretically that these systems become a saddle point in the design landscape (detailed explanation is given in Chapter 3).

An illustration of the saddle point construction is given in Figure 2.7. The construction of a saddle point is indicated by steps 1 and 2: It starts from a local minimum and

adds an extra lens element to the existing system. When a saddle point system is obtained in step 2, step 3 and step 4 show how to obtain two new local minima from the saddle point system. The detailed explanation of saddle point construction is given in Chapter 3.

2.3. CONCLUSION

In this chapter, we use the example of a magnifying glass to demonstrate how a lens can be designed on paper. However, nowadays optical systems are becoming too complicated to perform such analysis and be designed from scratch.

The major concepts such as basins of attraction, starting point, local and global optimization are explained for the optimization part of lens design. We show that the optimization algorithms rely almost exclusively on generally applicable mathematical models and use little or no specific knowledge about the optical system (and its design landscape). We briefly mentioned SPC at the end and the special structure of the lens design landscape it reveals. In the following three chapters, we will focus on the SPC method and its effectiveness in different design scenarios.

REFERENCES

- [1] R. R. Shannon, *The Art and Science of Optical Design* (Cambridge University, 1997) p. 354, Shannon's comment on global search method.
- [2] E. Hecht, *Optics* (Pearson, 2012).
- [3] H. Gross, H. Zügge, M. Peschka, and F. Blechinger, in *Handbook of Optical Systems*, Vol. 1 (Wiley-VCH, Weinheim, 2007).
- [4] Wikipedia, *Optical lens design*, https://en.wikipedia.org/wiki/Optical_lens_design?oldformat=true, Accessed: 2018-12-11.
- [5] K. P. Thompson, *The earliest history of computer-aided optical design on large computers: the previously classified work by James G. Baker 1945-1954*, Proc. SPIE **6668**.
- [6] D. P. Feder, *Automatic optical design*, *Appl. Opt.* **2**, 1209 (1963).
- [7] H. Gross, H. Zügge, M. Peschka, and F. Blechinger, *Principles of optimization*, in *Handbook of Optical Systems*, Vol. 3 (Wiley-VCH, Weinheim, 2007).
- [8] M. van Turnhout, *A systematic analysis of the optical merit function landscape: towards improved optimization methods in optical design*, Ph.D. dissertation (Delft University of Technology), 63 (2009).
- [9] I. Livshits and V. Vasilyev, *Q and A tutorial on optical design*, *Advanced Optical Technologies* **2**, 31 (2013).
- [10] D. Shafer, *Doing more with less*, Proc. SPIE **2537**, 11 (1995).
- [11] W. Smith and I. Genesee Optics Software, *Modern Lens Design: A Resource Manual*, Optical and electro-optical engineering series (McGraw-Hill, 1992).

- [12] W. J. Smith, *Modern Optical Engineering* (McGraw-Hill, 2000).
- [13] J. C. Miñano, P. Benítez, W. Lin, J. Infante, F. Muñoz, and A. Santamaría, *An application of the SMS method for imaging designs*, *Opt. Express* **17**, 24036 (2009).
- [14] J. Meiron, *Damped least-squares method for automatic lens design*, *J. Opt. Soc. Am.* **55**, 1105 (1965).
- [15] M. van Turnhout and F. Bociort, *Instabilities and fractal basins of attraction in optical system optimization*, *Opt. Express* **17**, 314 (2009).
- [16] T. G. Kuper and T. I. Harris, *Global optimization for lens design: an emerging technology*, *Proc. SPIE* **1781**, 15 (1993).
- [17] J. R. Rogers, *Using global synthesis to find tolerance-insensitive design forms*, *International Optical Design*, TuA4 (2006).
- [18] S. W. WELLER, *Simulated annealing and lens design*, *Optics News* **13**, 20 (1987).
- [19] G. W. Forbes and A. E. W. Jones, *Towards global optimization with adaptive simulated annealing*, *Proc. SPIE* **1354** (1991), 10.1117/12.47874.
- [20] Lambda Research Corporation, OSLO, <https://www.lambdaresearch.com/oslo/> (Retrieved from 2019).
- [21] M. Isshiki, *Global optimization with escape function*, *International Optical Design Conference 1998*, **3482**, 104 (1998).
- [22] K. E. Moore, *Algorithm for global optimization of optical systems based on genetic competition*, *Proc. SPIE* **3780** (1999), 10.1117/12.363781.
- [23] Ansys Zemax OpticStudio, *Ansys Zemax OpticStudio User Manual (version 2022 R2.02)*, www.zemax.com.
- [24] C. Menke, *Application of particle swarm optimization to the automatic design of optical systems*, *Proc. SPIE* **10690** (2018), 10.1117/12.2311610.
- [25] D. Guo, L. Yin, and G. Yuan, *New automatic optical design method based on combination of particle swarm optimization and least squares*, *Opt. Express* **27**, 17027 (2019).
- [26] O. Marinescu and F. Bociort, *Network search method in the design of extreme ultraviolet lithographic objectives*, *Appl. Opt.* **46**, 8385 (2007).
- [27] P. van Grol, F. Bociort, and M. van Turnhout, *Finding order in the design landscape of simple optical systems*, *Proc. SPIE* **7428**, 11 (2009).
- [28] M. van Turnhout, P. van Grol, F. Bociort, and H. P. Urbach, *Obtaining new local minima in lens design by constructing saddle points*, *Opt. Express* **23**, 6679 (2015).

3

THE SADDLE POINT CONSTRUCTION (SPC) METHOD

The Saddle Point Construction (SPC) method is one of the major methods investigated in this research. Several variations of the methods for constructing the saddle point system have been introduced through the years [1][2][3]. Instead of searching for alternative local minima, the SPC method constructs saddle points as intermediate steps to obtain new local minima. This is achieved by adding extra variables to the existing system. With this approach, it is shown in this chapter that it is possible to obtain new solutions in a systematic way.

3.1. SADDLE POINT AND DESIGN LANDSCAPE

For a given optical design problem, once the merit function and the number of variables N have been determined, an N -dimensional design (optimization) space is defined. The optimization of an optical system then is equivalent to finding the minimum of the merit function in the N -dimensional design landscape. As mentioned in the introduction, the optimization problem of optical design is a nonlinear problem, where multiple local minima are present in the design landscape. As a result, searching for one or several good local minima among many becomes a challenge for optical design.

For the simple case of only two variables, the minima can be viewed as the bottoms of the valleys in the design landscape. On the other hand, maxima are the mountain tops, which are usually not observed in the lens design landscape. Minima and maxima are both critical (or stationary) points. As explained in Section 2.2.5, there are also saddle points, which can be used as proxies to find new local minima.

In a 2D space, a 2D saddle point is a "saddle" where along one direction in variable space, the value of the function decreases with respect to the saddle point value, while along the perpendicular direction, it increases. Mathematically, the so-called Morse Index is used to distinguish minima, maxima and saddle points. The Morse Index is defined as the number of negative eigenvalues of the Hessian matrix H at the critical

point, when the Hessian matrix H is non-degenerate. A negative eigenvalue indicates that along the direction defined by the corresponding eigenvector of the Hessian matrix, the critical point is a maximum. Therefore, in an N -dimensional space, the Morse Index of a minimum and a maximum is 0 and N respectively. A saddle point can have any Morse Index between 1 and $N-1$.

In this research, the saddle point with a Morse Index of 1 is of particular interest. This kind of saddle point has one direction, along which the value of the merit function decreases, while in $N-1$ other directions the merit function increases. This means that a saddle point of Morse Index 1 is a local minimum in all but one direction. When the optimization follows the two opposite directions in which the merit function descends, two local minima can in principle be found (Figure 3.1(b)). As a result, if there is a way to rapidly obtain these saddle points in the design landscape, it provides an approach to find more local minima. In fact one finds two local minima from a saddle point of Morse Index 1.

3.2. THE SPC METHOD

In this section, the SPC method is introduced and explained. As the name suggests, the method constructs saddle points in the design landscape. These are all saddle points with Morse Indices of 1¹. Instead of optimizing from an arbitrary point in the landscape, optimizing from a saddle point can systematically lead to two local minima.

The saddle points are constructed by adding extra variables to the merit function that has attained the existing local minimum using the original number of variables. In principle, the variables can be any parameter in the lens design model. It should be mentioned that the most effective variables that have been used so far are the lens curvatures.

There are a general version and a special version of SPC. The general version is flexible with the insertion position of the null element and a scan of the computed derivatives of the merit function is required to identify the saddle points. The special version is restricted with the positions for inserting a null element. However, a saddle point can be directly obtained without additional steps. In the following texts and rest of the thesis, we mainly focus on the general version of SPC. The special version will be briefly explained and more details about it can be found in Reference [1].

3.2.1. THE GENERAL VERSION OF SPC

The SPC method always starts with an optimized lens system that is a minimum of a merit function defined by N arbitrary variables (x_1, x_2, \dots, x_N) . A saddle point can then be constructed in a design space with $N + 2$ variables by adding a lens with zero thickness and with surfaces having the same curvatures (see Figure 3.1(a)). Such a lens has no optical effect and hence the merit function is the same after insertion of this element as before. The two curvatures are the two new variables, so that the total number of variables has been increased to $N+2$. For certain values of the curvatures of the inserted

¹In the special version of the SPC, it can be proven that the Morse Index of the constructed saddle point is 1. In the general version of the SPC, it is shown in the numerical experiments that the Morse Index of the constructed saddle point is 1.

surfaces, the resulting system is a saddle point, which is a minimum along $N + 1$ directions in the variable space and a maximum along one direction (i.e., mathematically it will have a Morse index of 1 [2]). The details of how such values are obtained will be explained further in the following paragraphs. Along the direction for which the SP is a maximum, two minima can be obtained systematically, as shown for simplicity in a 2D case in Figure 3.1(b).

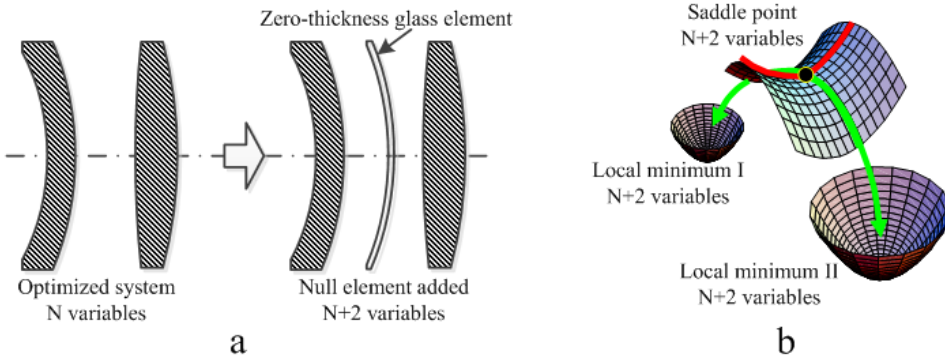


Figure 3.1: Illustration of the SPC. (a) An SP in $(N + 2)$ -dimensional space obtained by insertion of a zero-thickness glass meniscus. (b) A Morse index 1 saddle point in the 2D case.

The pair of surfaces with equal curvatures can be either a zero-thickness lens added to the original minimum or a zero-thickness air space inside a lens. Both the added lens and the added air space with zero thickness, do not affect the ray paths, and hence the merit function value of the original system is not changed. Such an element of zero thickness is therefore called a null element. We can write the merit function MF in a N -dimensional space as

$$MF^N(x_1, x_2, \dots, x_N).$$

At a local minimum $M0(x_1^0, x_2^0, \dots, x_N^0)$, we have

$$\frac{\partial}{\partial x_i} MF^N(x_1^0, x_2^0, \dots, x_N^0) = 0, \text{ for } i = 1, 2, \dots, N. \quad (3.1)$$

$\frac{\partial}{\partial x_i} MF^N(x_1^0, x_2^0, \dots, x_N^0)$ denotes the partial derivatives of MF^N with respect to variable x_i , evaluated at the point $(x_1^0, x_2^0, \dots, x_N^0)$. Furthermore, the Morse Index of the Hessian at $(x_1^0, x_2^0, \dots, x_N^0)$ is zero, i.e. the eigenvalues of the matrix

$$H_{ij}^N = \frac{\partial^2}{\partial x_i \partial x_j} MF^N(x_1^0, x_2^0, \dots, x_N^0), \quad (3.2)$$

are all non-negative.

Next we introduce a null element with common curvature $c_1 = c_2$ of its two surfaces. Additional variables are then introduced as follows:

$$\begin{aligned} x_{N+1} &= c_1 + c_2, \\ x_{N+2} &= c_1 - c_2. \end{aligned} \quad (3.3)$$

Because the introduction of the null-element does not alter the physical property of the optical system, therefore does not have any effect on the merit function, we have the following relation between the merit functions with N and $N + 2$ dimensions:

$$MF^{N+2}(x_1, x_2, \dots, x_N, x_{N+1}, x_{N+2} = 0) = MF^N(x_1, x_2, \dots, x_N). \quad (3.4)$$

Equation 3.1 and Equation 3.4 implies

$$\begin{aligned} \frac{\partial}{\partial x_j} MF^{N+2}(x_1^0, x_2^0, \dots, x_N^0, x_{N+1}, x_{N+2} = 0) \\ = \frac{\partial}{\partial x_j} MF^N(x_1^0, x_2^0, \dots, x_N^0) \\ = 0, \end{aligned} \quad (3.5)$$

for all $j = 1, 2, \dots, N$, and all x_{N+1} .

Furthermore Equation 3.4 implies

$$\frac{\partial}{\partial x_{N+1}} MF^{N+2}(x_1, x_2, \dots, x_N, x_{N+1}, x_{N+2} = 0) = 0, \quad (3.6)$$

for all x_{N+1} . Equation 3.6 applies to $(x_1^0, x_2^0, \dots, x_N^0)$, we have

$$\frac{\partial}{\partial x_{N+1}} MF^{N+2}(x_1^0, x_2^0, \dots, x_N^0, x_{N+1}, x_{N+2} = 0) = 0, \quad (3.7)$$

for all x_{N+1} .

Next we vary x_{N+1} to find a point x_{N+1}^0 for which

$$\frac{\partial}{\partial x_{N+2}} MF^{N+2}(x_1^0, x_2^0, \dots, x_N^0, x_{N+1}^0, x_{N+2} = 0) = 0. \quad (3.8)$$

When such a point, x_{N+1}^0 , can be found (and this happen often in practice, but not always, as will be seen in later examples), we conclude that

$$(x_1^0, x_2^0, \dots, x_N^0, x_{N+1}^0, 0)$$

is a critical point of MF^{N+2} .

Now Equation 3.4 implies

$$\begin{aligned} \frac{\partial}{\partial x_i \partial x_j} MF^{N+2}(x_1^0, x_2^0, \dots, x_N^0, x_{N+1}^0, x_{N+2} = 0) \\ = \frac{\partial}{\partial x_i \partial x_j} MF^N(x_1^0, x_2^0, \dots, x_N^0), \end{aligned} \quad (3.9)$$

for $1 \leq i, j \leq N$.

Furthermore, from Equation 3.6

$$\frac{\partial}{\partial x_i \partial x_{N+1}} MF^{N+2}(x_1^0, x_2^0, \dots, x_N^0, x_{N+1}^0, x_{N+2} = 0) = 0, \quad (3.10)$$

for all $1 \leq i \leq N + 1$.

So we have for the Hessian in $N+2$ dimensional space

$$\left(H_{ij}^{N+2} \right) = \begin{bmatrix} \left(H_{ij}^N \right) & & 0 & \frac{\partial^2 M_{F^{N+2}}}{\partial x_1 \partial x_{N+2}} \\ & & \vdots & \vdots \\ 0 & \dots & 0 & \frac{\partial^2 M_{F^{N+2}}}{\partial x_{N+1} \partial x_{N+2}} \\ \frac{\partial^2 M_{F^{N+2}}}{\partial x_{N+2} \partial x_1} & \dots & \frac{\partial^2 M_{F^{N+2}}}{\partial x_{N+2} \partial x_{N+1}} & \frac{\partial^2 M_{F^{N+2}}}{\partial x_{N+2}^2} \end{bmatrix}. \quad (3.11)$$

The last column and last row are, in general, nonzero, hence there are in general nonzero eigenvalues. Numerical experiments show that typically the eigenvalues of H_{ij}^N are close to a subset of N eigenvalues of H_{ij}^{N+2} , but they are not the same. In many cases in which the eigenvalues of the Hessian matrix have been computed numerically the Morse Index turned out to be 1. A mathematical proof of the fact that the Morse Index is always one has not yet been found, but no exception has been found numerically.

Obtaining saddle point systems with the general version of the SPC includes the following steps:

1. Insert a null element into the existing minimum.
2. Compute the derivatives (numerically) of the merit function with respect to the curvature of the inserted null element. The null element has two identical curvatures. The computed derivative of the two curvatures have opposite signs.
3. Select the curvature values where the derivative of the merit function equals zero. Systems with these curvature values are the saddle point systems.

These steps are referred to as "performing an SPC scan" in this thesis. In our research, the SPC scans are performed in CODE V using the MACRO function. Figure 3.2 shows an example of the results of such an SPC scan. An SPC scan in a chosen position can lead to multiple saddle point systems. There are four zero crossings in Figure 3.2(b) indicating that four saddle points are found. When the saddle points are found, initial systems for subsequent local optimization can be obtained by choosing for each zero crossing in Figure 3.2(b) two systems, one to the left, one to the right of the saddle point. Local optimization, using, e.g., a Damped Least-Square (DLS) algorithm, will then lead to two minima, one on each side of the "saddle," as shown in Figure 3.1(b). To obtain the broadest variety of new minima with SPC, in general both zero-thickness lenses and zero-thickness air spaces are necessary. For simple systems, there are examples of minima that can be obtained with one of these two types of null elements but not with the other one. Finding the saddle points is in principle much less time consuming than DLS optimization, because it only involves the evaluation of the derivative of the MF for the one-dimensional sequence of scan points according to Equation 3.8, whereas local optimization involves many iterations where a Jacobian matrix is evaluated. The zero-thickness condition for the null element is not a severe limitation, as it may seem, because in the resulting minima the distances between the surfaces (and the glass of the new lens) can be easily changed as desired. Once the two local minima are obtained with

the SPC, other parameters of the new lens (thickness, aspheric coefficients, etc.) can be made available as variables to further minimize the merit function.

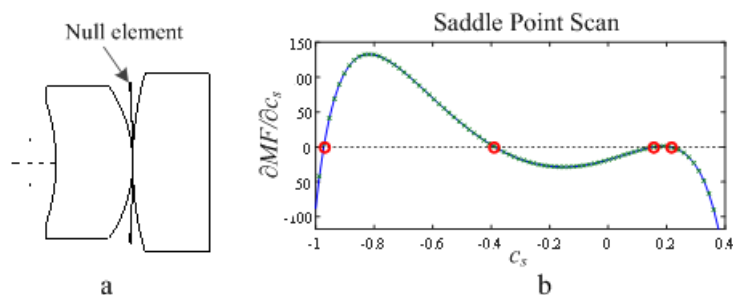


Figure 3.2: Example of an SPC scan. (a) The insertion of a glass null element between two lenses exhibiting a known minimum value of the merit function. (b) The SPC scan finds four saddle points in this search. The red circles indicate the corresponding values of the null element curvatures.

3.2.2. THE SPECIAL VERSION OF SPC

Different from performing a scan using the general version of SPC, the special version of SPC directly constructs the saddle points by adding null elements with specific configuration to the existing system. The way in which the special version of SPC works is described in Reference [1]. The null element is inserted in contact with (i.e. at zero axial distance from) an existing surface (reference surface) in the original minimum. The two curvatures and the optical material of the null element have to be the same as the ones at the reference surface. In this case, three surfaces have the same curvatures and two zero-thickness spaces between them, as illustrated in Figure 3.3. It has been shown in Reference [1] that the new system created in this way is a critical point² as well as a saddle point with a Morse Index of 1³.

Compared to the general version of SPC, the special version is a rapid way to construct saddle point systems. However, the constructing positions are limited to the locations of the existing surfaces. As a result, some of the saddle points found by the general version can not be obtained by the special version of SPC. In this research, the general version of SPC is mainly used for acquiring new minima. When an SPC scan is performed on an existing surface, the saddle point constructed by the special version of SPC can also be obtained.

²As a critical point, all components of the gradient must be zero. Hence it satisfies Equations 3.8.

³Saddle points with Morse Index 2 can also be easily constructed: As shown in Reference [4], if the special version of SPC is applied by inserting not only one, but two null elements, both of them in contact with an existing surface in an existing minimum, then the resulting critical point will have a Morse index of 2. Examples have shown how Morse Index 2 saddle points can be used in lens design in the same reference. However, the way to use them is very different from the Morse 1 saddle points used in the present work. The question whether constructing Morse Index 2 saddle points is useful enough for practical usage in lens design or not requires further study.

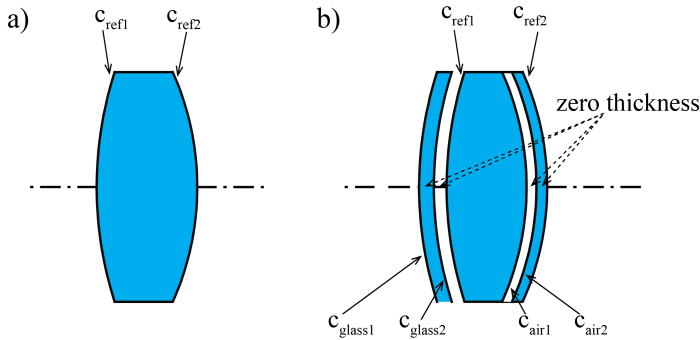


Figure 3.3: Illustration of the special version of SPC. a) A lens before adding any null element. b) A glass null element has been attached to the surface on the left. An air null-element has been attached to the surface on the right. $c_{glass1} = c_{glass2} = c_{ref1}$, $c_{air1} = c_{air2} = c_{ref2}$. The special version of the SPC usually adds one glass null element or one air null element. For demonstration purpose, different types of null elements are added to the lens, and the thickness of the null elements are also nonzero in the drawings.

3.3. RECOMMENDATIONS FOR APPLYING SPC

As mentioned in the previous section, the SPC adds two extra curvatures to an existing optical system. It is straightforward, if the design goal is to add new elements to the system. Nevertheless, when applying the SPC in the practical design scenario, the technique can be used in different ways to achieve different design goals. Some scenarios are listed below as recommendations.

3.3.1. ADDING LENS ELEMENTS

The direct way of applying SPC is to add a lens element to the existing system. This can be achieved by either inserting a glass null element or an air null element. Essentially, both approaches introduce two curvature variables to the system. Adding a glass null element can be treated as adding a lens element, while adding an air null element can be regarded as splitting an existing lens. The two approaches are described in the following two paragraphs.

GLASS NULL ELEMENT - ADDING A LENS

The Cooke triplet is chosen here as an example to demonstrate how SPC is applied. The position of the insertion of the glass null element is indicated by the dashed line in Figure 3.4. In principle, any position in the spaces between the lenses can be selected for the insertion. The choice of the position can be guided by the conventional lens design method. Examples of combining a traditional design method and the SPC method are given in Chapter 4.

In Figure 3.4, the insertion position is chosen between the middle lens and the right lens. As shown in the SPC scan curve, three zero-crossing points of the derivative of the merit function are found, hence three saddle points are found. Three new solutions with the added lens are listed in the same figure. Numerical experiments show that in simple cases neighbouring saddle points typically have a common minimum linked to

them, i.e. M saddle points lead to $M + 1$ minima. Based on the scan curve in Figure 3.4, we would expect four solutions instead of three to be produced after optimizing from the three saddle points. It is the case when the minima are in the form of having an element with zero-thickness. However, when the thickness of the inserted element is increased, a solution will disappear or merge with other solutions (an example is given in the next chapter in Figure 4.7). This is the situation in this example, and only three minima remained once the thickness is increased to a practical value. Among the three solutions, the added lenses present different optical powers. The original lens elements of the system remain almost the same powers compared to the status before insertion. In this example, two systems have better merit function values (left system $43.27 \mu\text{m}^2$, right system $47.80 \mu\text{m}^2$) than the original one ($49.95 \mu\text{m}^2$). The merit function used here is the default transverse ray function in CODE V. The default transverse ray function is a widely-used merit function in CODE V because it can be computed easily. For each ray it uses the length of the transverse aberration vector, and its value is related to the image spot size. The choice of the rays and their weights is done automatically by CODE V such that maximum accuracy can be obtained for a given size of the ray grid in the pupil. For instance, for a lens system with 5 variable curvatures, typically about 60 rays are chosen automatically by the lens design software to construct the transverse ray merit function. Alternative merit functions that use the optical paths of different rays are also widely used, but their computation is slightly more complicated. Merit functions using the modulation transfer function are also used, but because they are computationally very intensive they are used typically for the final polishing of local minima.

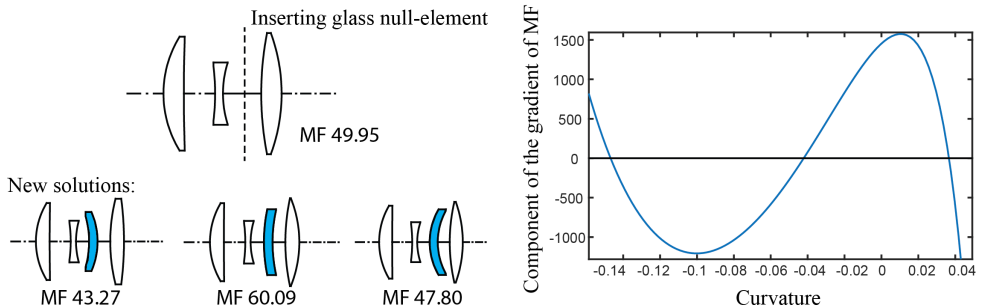


Figure 3.4: Adding lens element by adding glass null element using the SPC. The dashed line indicates the position where the null element is inserted. Three saddle points lead to four solutions with zero-thickness element (not shown). When the thickness of the inserted element is increased, three solutions with one lens more are obtained. The added lens elements are highlighted. Merit Function (MF) values are listed for each system, and the unit is μm^2 .

AIR NULL ELEMENT - SPLITTING A LENS

Alternatively, adding lenses can also be realized by splitting lenses. With the SPC, inserting an air null element into a lens is equivalent to splitting this lens. In Figure 3.5, the most right lens element was chosen to be split. The dashed line indicates where the lens was split. The derivative as function of the curvature of the inserted air space is plotted in the right part of the figure. It is seen that there are four saddle points. As with the

situation in glass element insertion case, they lead to three solutions instead of five. The solutions and their merit function values are listed in Figure 3.5. When evaluating the merit function values, it is seen that the three new systems all have better merit function values than the original one. The inserted air spaces are highlighted.

Compared with the solutions produced by inserting glass null element, the solution on the right in Figure 3.5 for air null-element insertion presents a similar system shape as the solution on the left in Figure 3.4 for glass null-element insertion. The similarity is shown by the same power distribution of each element in the systems⁴. The power distribution (i.e. the succession of surface powers) can be estimated visually by looking at the succession of the signs of the surface curvatures (i.e. whether a surface is curved to the right or to the left). The power distribution gives a rough indication of how similar two systems are. When adjusting the thickness and airspace, the two systems converge to the same solution.

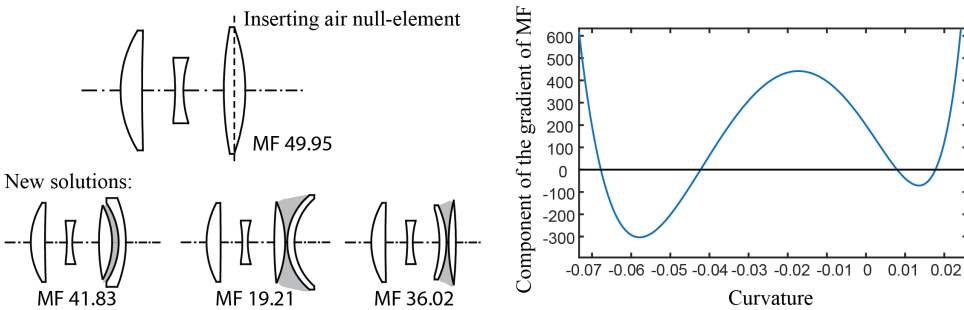


Figure 3.5: Adding a lens element by adding an air null element using the SPC (splitting lenses). The dashed line indicates the SPC scan position. Four saddle points lead to three solutions with one lens more. The added air spaces are shaded. MF value of each system is listed, and the unit is μm^2 .

3.3.2. SWITCHING TO ALTERNATIVE MINIMA

THE EXTRACT AND ADD OPERATION

In practical lens design, it is usually important to obtain a better solution without changing the number of variables (lens elements). Adding extra elements in the system often means that the costs of the system increase and that the system becomes less compact.

When using the SPC, with some subsequent steps described below, it is possible to switch between local minima with the same number of variables. Illustrated in Figure 3.6, the first step is to choose which lens to modify. The middle one of the Cooke triplet was chosen in this example. The next step is to extract this element from the original system as follows. The thickness of the chosen element is gradually reduced to zero while in each step the system is optimized with respect to the remaining variables. Small steps should be applied for the change of the thickness to prevent the optimization moving to a different basin of attraction. When the basin of attraction of the optimization changes,

⁴For the solution on the right in Figure 3.5 and the solution on the left in Figure 3.4, the succession of the signs of the surface curvatures (from left to right) is: positive, no sign (zero), negative, positive, negative, negative, positive, and negative.

a sudden change in the merit function value can be observed. After the thickness is reduced to zero, the curvatures of the two surfaces are usually different. Next, the curvatures can be changed gradually until they become equal and the system is optimized with respect to the remaining variables after each step.

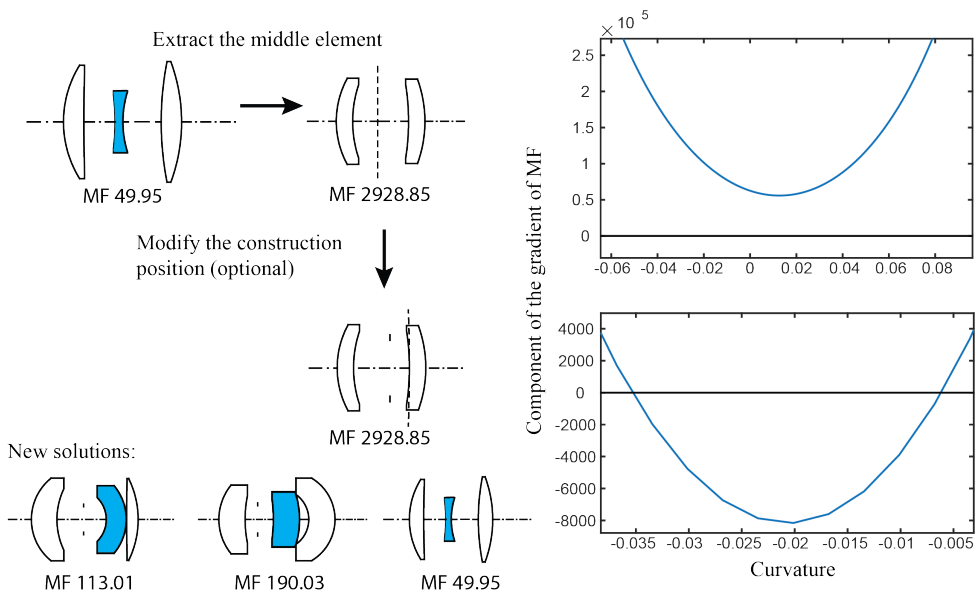


Figure 3.6: Switching to different solutions with the SPC. The middle lens element was chosen to be first extracted, and the SPC was performed at the same location. However, no saddle point was found in that location. The scanning location was then shifted right. Two saddle points and three solutions were obtained. MF value of each system is listed, and the unit is μm^2 .

Once the thickness of the chosen lens has been made zero and the two curvatures have been made equal by this procedure, the null element can be extracted from the system without changing the value of the merit function. In fact, we have now a system with one lens less which is a local minimum for the reduced number of variables. The system is ready to be inserted with a null element in order to add back the extracted lens and obtain new local minima. The null element is usually inserted at the same location where the lens is extracted and the derivative of the merit function with respect to the curvature is scanned to search for saddle points. In Figure 3.6, the curve plot on the top shows the SPC scan when it is performed at the position of the extraction. It is seen that the scan shows no saddle point. The subsequent step is to change the position where the null elements is inserted in the space between the two lens elements. Where the arrow points down in Figure 3.6, the position of the insertion of a null element is chosen to coincide with the left surface of the right lens. The curve of the derivative has two zero-crossings, which indicates two saddle point systems. One of the saddle points can be explained by the special version of the SPC [1], where the curves of the inserted element are identical to the surface which the null element is attached to.

Three solutions were obtained from the two saddle points. After relaxing all variables

and optimizing, the solution on the right in Figure 3.6 becomes identical to that of the original Cooke triplet. The other two solutions are new solutions. None of the two shows better merit function values compared to the Cooke triplet design.

THE ADD AND EXTRACT OPERATION

It was shown that with an extract-and-add strategy, one can switch systematically from one local minimum to other minima, with the same number of variables. An intermediate step is necessary to first reduce the number of the variables in the original problem (by extracting a lens).

It is also possible to switch to different solutions using the SPC by first increasing the number of variables: an add-and-extract strategy. However, it is not really recommended since it requires more decisions on where to add or extract in the intermediate steps. For an extract-and-add strategy, it is demonstrated in the previous section that the designer only needs to decide where to extract an element. In contrast, an add-and-extract strategy will work as follows: The designer first has to decide where to add an element to the original system. Then, after the SPC operation, usually more than one solution is produced. In the simplest case, the designer can decide to extract the same element for all of them. However, it is also possible to decide to extract different elements from the solutions obtained after SPC. In either way, the add-and-extract strategy will require more decision steps than the extract-and-add strategy.

3.4. PRACTICAL FLOW OF IMPLEMENTING SPC IN LENS DESIGN

In Section 3.3 above, we have explained several scenarios to add lens or air null elements in existing designs in order to perform SPC scans and obtain new solutions. At a basic level, the SPC procedure is simple and remains the same in all cases. It can be described as the following steps:

1. Start with the optimized local minimum for which we want to further lower the merit function value. If the merit function necessary to satisfy all practical requirements is too complex, simplify it and reoptimize the system with the simplified merit function. Make sure that the optimization quality (e.g. number of cycles used) is higher than what one would usually use for practical optimization in lens design. For a complex system, if obtaining a starting minimum with a residual gradient that is sufficiently low takes too long, reduce the number of variables.
2. Insert a glass or air null element at a certain position in the system.
3. Compute the scan curve using Equation 3.8 and find all its zero crossings.
4. For each zero crossing, perturb the resulting saddle point on both sides of the “saddle”. For N zero crossings we then obtain $2N$ starting points for subsequent local optimization. Locally optimize the $2N$ starting points.
5. In the resulting local minima adapt the axial thickness that results from the null element such that all edge thickness violations are corrected.
6. Select the best solution, or several of good solutions, by examining the merit function value as well as the system drawing (to avoid solutions with too much stress). If for SPC the merit function has been simplified, reoptimize the selected good solution(s) with the original merit function that contains all necessary operands.

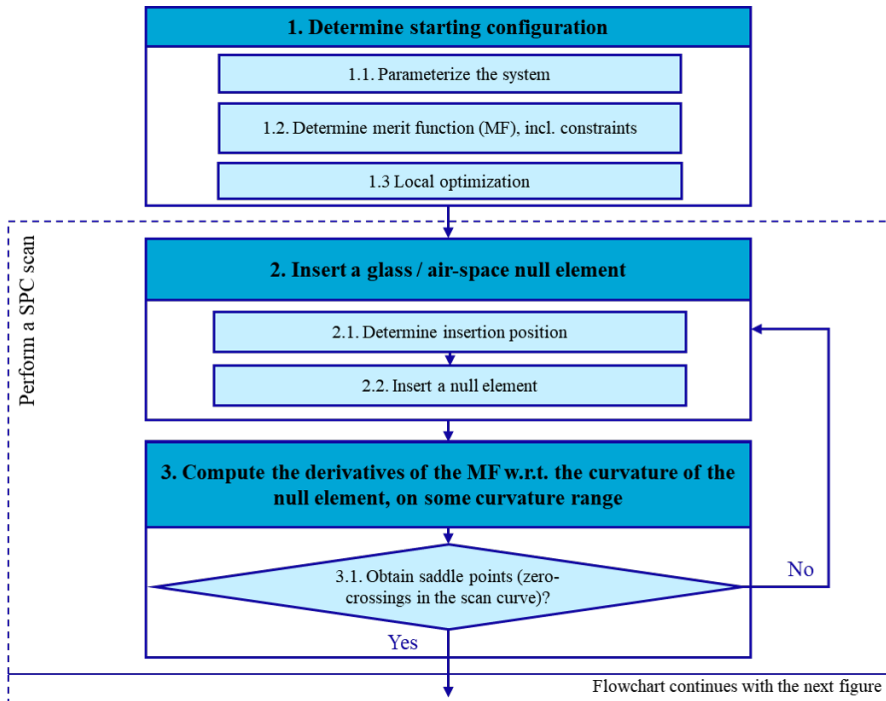


Figure 3.7: Flowchart of performing a SPC scan. Once saddle point systems are obtained, Figure 3.8 shows the next stage of how to obtain minima with SPC.

Experience at TU Delft and elsewhere has shown that in many cases even following these basic steps of SPC can produce valuable new designs. However, the basic approach may miss some solutions. As with all other local and global optimization algorithms in lens design, optimal performance is obtained by adapting some details and parameters by using empirical information depending on the specific application. The steps mentioned above are mapped into the flowchart in Figures 3.7 and 3.8, together with some more sub-steps. We briefly describe below how these refined steps can help finding more solutions. Some of the descriptions require detailed explanation, which will be elaborated in the following chapters. With this flowchart, we hope to make it easier for readers to implement the method themselves.

The flowchart in Figure 3.7 consists of two major blocks: determine the starting configuration and perform a SPC scan. As stated previously, it is required to start with a local minimum to perform a SPC scan. The configuration of merit function and constraints can impact the effectiveness of SPC. This is elaborated in Section 5.3.2 and Section 5.3.4 of Chapter 5. One of the practical recommendations is to use minimal and essential constraints in the merit function.

Step 2.1 in Figure 3.7 indicates that decisions need to be made on where the insertion position should be. This can be a conscious choice made by the designer or some automated process sampling the positions in the lens system. In Section 5.3.3 of Chap-

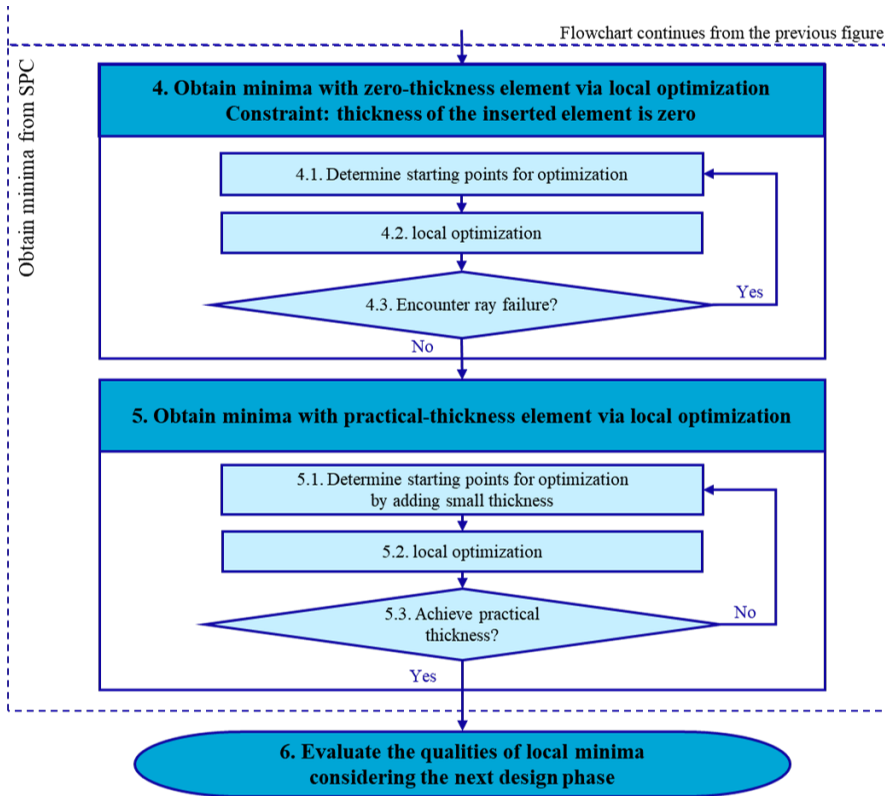


Figure 3.8: Flowchart of obtaining minima from SPC. After performing a SPC scan (Figure 3.7), two additional steps (5, and 6 in the flowchart) are required to obtain practical solutions.

ter 5, we demonstrate two examples of combining designers' experience with SPC to obtain additional good results. The algorithm on automating the element insertion is not studied in detail in this thesis. For inserting null element, our experience shows that sampling at positions in contact with existing lens elements and at positions in the middle between two lens elements is sufficient to achieve good results. Section 3.2.1 of this chapter explains the theoretical background of step 3 in Figure 3.7. The results achieved within this thesis use programmed scripts in CODE V to numerically compute the derivatives for SPC scans.

The flowchart in Figure 3.8 explains the steps of obtaining minima once the saddle point systems are constructed. The sub-steps under step 4 explain how to obtain the minima with zero-thickness element. This step is vital and may be overlooked. The purpose of this step is to ensure the local optimizer follows two descending directions predicted by the constructed saddle points. An unintended practice would be optimizing while making the thickness of the inserted element as a variable. It would alter the optimization landscape (compared to the one where the saddle points are constructed) in an unpredictable way. As a result, the assumption of the two descending directions

can be invalid. Chapter 4 shows a network of saddle points and minima in a specific scenario: All consecutive lens elements are in contact with each other (zero axial distance) and the inserted element is an air null element. In this way, we create a practical use case to study the network staying at step 4.

When optimizing from the saddle points, choice needs to be made on the starting points (step 4.1) to ensure that the two descending directions are captured. An detailed example on the choice of starting points near the saddle points has been given in Section 5.2.2 of Chapter 5. In the same section, we also show that, when increasing the thickness of the inserted element (step 5 in Figure 3.8), it is recommended to use small increments if one wants to maintain the system shape of the solution⁵.

The final results from SPC should be several design candidates which will be evaluated considering the next phase of the design. For instance, this can either be reoptimizing the system with additional constraints or another cycle of SPC to further improve system performance (several examples are available in Chapters 4 and 5).

3.5. CONCLUSION

In this chapter, the SPC method is explained in detail. The method constructs a saddle point system with Morse Index 1 in a higher dimensional space (from dimension N to dimension $N + 2$). Optimizing from a saddle point of Morse Index 1 can systematically lead to two local minima. The general version of the SPC involves a scan of the inserted curvatures to search for the saddle point. It is not guaranteed that the scan will generate saddle points. In the special version of the SPC, a null element attached to an existing surface is added where a saddle point system can be immediately obtained [1].

By performing a scan to find systems where the first order derivative of the merit function vanishes, the general version of SPC usually finds more than one saddle point, therefore, often several solutions are obtained. When such a scan of the derivative happens at the same position as an existing surface, the saddle point predicted by the special version of the SPC will also be included.

A flowchart is provided (Figures 3.7 and 3.8) to illustrate the steps required for obtaining minima with SPC. Simple examples are given to explain how SPC can be performed in practice: An SPC scan can be done either by scanning a glass null element (adding a lens element) or scanning an air null element (splitting a lens element). Depending on the scan positions, the two approaches can produce similar solutions (e.g. Figures 3.4 and 3.5).

Switching to different local minima using the SPC without increasing the number of elements (i.e. variables), can be very useful. This is realized by first extracting a lens element, and then performing an SPC scan at the extracted position. Figure 3.6 demonstrates the steps where extra solutions are obtained.

In the following two chapters, the SPC is used in the way described in this chapter. A detailed study of the saddle points - minima networks using SPC is presented in Chapter 4. Several practical examples given in Chapter 5, demonstrate how these approaches can be used in lens design.

⁵If there is no preference to keep the system shape, one can also make the thickness as a variable with or without a constraint.

REFERENCES

- [1] F. Bociort and M. van Turnhout, *Finding new local minima in lens design landscapes by constructing saddle points*, *Opt. Eng.* **48(6)**, 063001 (2009).
- [2] M. van Turnhout, P. van Grol, F. Bociort, and H. P. Urbach, *Obtaining new local minima in lens design by constructing saddle points*, *Opt. Express* **23**, 6679 (2015).
- [3] Z. Hou and F. Bociort, *Reducible complexity in lens design*, *Proc. SPIE* **9626**, 8 (2015).
- [4] F. Bociort and P. van Grol, *Systematics of the design shapes in the optical merit function landscape*, *Proc. SPIE* **7717**, 14 (2010).

4

LENS DESIGN LANDSCAPE EXPLORATION WITH SIMPLE SYSTEMS

In a lens design landscape, from a saddle point system with Morse index 1, optimization will lead to two solutions of local minima. When performing a SPC¹, it can find several saddle point configurations, which subsequently lead to multiple local minima.

Therefore, it is an intriguing question whether or not the local minima in the design landscape are all connected by the saddle points of Morse index 1. If this is the case, then by constructing sufficiently many saddle points, in principle all local minima of the design landscape can be obtained.

This chapter is mainly based on a research paper[1] published in 2016. It is shown that in the design landscape of a wide-angle pinhole lens and in closely related optimization landscapes, all good local minima found by other methods can be obtained easily with a succession of one-dimensional searches (SPC scans) starting from simpler systems. By replacing high-dimensional searches with a succession of one-dimensional searches, the design efficiency can be increased significantly. By combining this method with conventional design methods, the wide-angle pinhole lens can be designed starting from a single lens.

In a 2004 research paper [2], Bociort et al. show that when some general mathematical requirements are satisfied, the minima and saddle points in the lens design landscape form a network. The saddle points in this network can be theoretically predicted, hence from the saddle points, minima are obtained. Research conducted by Van Grol et al.[3] claims that there is a fundamental network for a lens design problem with a certain number of lenses. In their paper, examples of the fundamental network of doublet and triplet

Parts of this chapter have been published in Applied Optics 55, 10449 (2016) [1].

¹The definition of "performing an SPC scan" is given in Chapter 3.2.1.

systems are given. The authors predict that in practical design cases, the types of system shapes² of the solutions will not exceed the types of system shapes from the fundamental network. In a later paper from 2010[4][5], Bociort shows that a toy model based on third-order spherical aberration can be used to mathematically explain the fundamental network of the triplet.

For the assumption that the dominant aberration is spherical aberration, and that the lenses are not too thick, it is possible to predict the existing solutions within the network. However, in practical cases, aberrations other than spherical aberration will also determine the topology of the merit function landscape. In this case, it is mathematically too complex to build an analytical model to predict the behaviour of the merit function landscape.

Despite the complexity to use an analytic model to predict the merit function landscape, earlier research shows [6][7] that in the lens design space, a special structure is present that makes the lens design problem different from a general global optimization problem: certain saddle points existing in the landscape are reducible to minima of simpler systems with one lens element less. Using this structure, the SPC [7] method is developed to obtain new design solutions. In the conventional approach, trial and error is needed to look for different solutions (if there is more than one) by adding one extra lens element to the existing system. With the SPC, there is a systematic procedure to search for available solutions by adding new elements to the system. When starting from a minimum with N variables, with SPC it is possible to systematically find minima in an $(N + 2)$ -dimensional variable space by using one-dimensional searches, rather than $(N + 2)$ -dimensional searches. Important questions are, however, how many minima existing in the landscape can be obtained in $(N + 2)$ -dimensional search spaces, by replacing high-dimensional searches with one-dimensional searches? What percentage of them can be found? Is it possible to at least obtain the good solutions? In this chapter, the performance of SPC is evaluated for more general problems, especially those of practical interests.

4.1. WIDE-ANGLE PINHOLE LENS

A wide angle pinhole lens, designed by Irina Livshits from ITMO, has been chosen to study the design landscape using SPC. For surveillance applications, this kind of lens, has in recent years, become very popular. Instead of a lensless optical element, the pinhole lens here refers to a lens having an aperture stop with a small diameter placed in front of the system. All lenses are in contact, one lens is cemented, and three different glass materials are used [see Figure 4.1(a)]. The design specifications are given in Table 4.1 and the surface specifications are provided in Appendix A.1. In Figure 4.1(b) a 2D image simulation made with CODE V shows that, despite its simplicity, the system has an imaging quality that is adequate for its intended purpose. The system can also be adapted for spectral imaging applications [8]. For optimization we use the default CODE

²In a lens system, each lens element has its lens form: biconcave, biconvex, meniscus, etc. The type of the system shape of a lens system is determined by the combination of the individual lens forms. For example, a biconcave-biconvex doublet is one type of system shape, and a biconcave-meniscus is a different type of system shape.

V merit function that is based on transverse ray aberrations. The value of the merit function (called in CODE V error function) for this system is $5.68 \mu\text{m}^2$. It is a composite value, scaled so that it is the mean square of the weighted image radius. The value has considered a setting with three wavelengths as shown in Table 4.1. In the following text in this chapter, values of the merit function are presented without adding the unit.

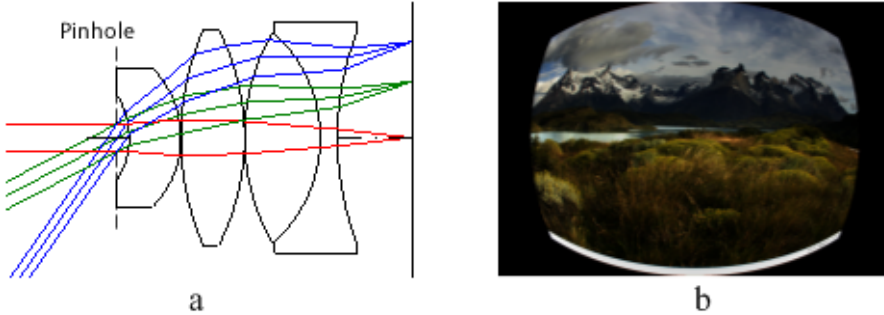


Figure 4.1: Wide-angle pinhole lens. (a) Lens drawing. (b) 2D image simulation of its imaging quality.

Table 4.1: System specification

Entrance Pupil Diameter (EPD, mm)	0.8
(Full) Field Of View (FOV, °)	110
Effective Focal Length (EFL, mm)	3.5
Wavelength (nm)	644, 546, 480

4.2. SWITCHING BETWEEN LOCAL MINIMA

To find other possible solutions for the same specifications, two different global optimization methods are used, Global Synthesis from CODE V [9][10] and a saddle point detection method (the program NETMIN³). In this example, for simplicity only the surface curvatures are used as variables. Minor edge thickness violations, which can be easily corrected in a later stage, are acceptable in this analysis.

It is of interest to find the so-called stable solutions. Stable solutions means that the corresponding local minima do not easily appear and disappear when specifications (such as size of the aperture, field of view [FOV], or effective focal length [EFL]) are changed. Three stable solutions are found, denoted in Figure 4.2 by M1, M2, and M3. They exist for a wider range of specifications, e.g., not only for the FOV of 110° as shown in Table 4.1, but also for a FOV of 90° , and not only for an EFL of 3.5 mm (for which the corresponding merit function values are 5.68, 11.51, and 11.09, respectively) but also for an EFL of 3.0 mm (with merit function values of 6.80, 23.20, and 16.70). These solutions have no large edge thickness violation and relatively low merit function values.

³NETMIN is a tool developed in TU Delft to search for new local minima based on saddle point detection method. Detailed description can be found in Chapter 2.2.5.

From the lens drawings of these solutions, we observe that the second element (from the left) differs significantly between them. A few solutions with low stability and large merit function values can also be found. They exist for instance either at EFL 3.5 mm or at EFL 3.0 mm, but not at both. Since these solutions with low stability appear or disappear easily when the EFL is changed, experience shows that their basins of attraction are small⁴ and therefore the risk that the optimization gets trapped in these basins is rather low. The solutions with low stability will therefore be ignored in what follows. No global optimization method can guarantee with absolute certainty that all minima are found. However, the landscapes examined in this chapter are simple (up to six curvature variables) and extensive numerical investigations with different methods have been performed. This gives us confidence that all minima having a basin of attraction large enough to be relevant for practical purposes are found with the methods being used.

4

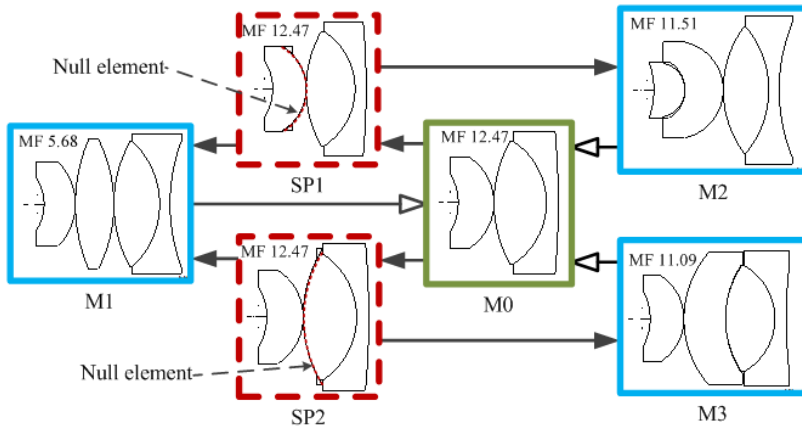


Figure 4.2: Switching between minima (blue boxes) using SPC. Extraction of the second element from M1, M2, or M3 results in the system M0 (hollow arrows). Starting from M0, two routes, each having an SP lead to three different minima (solid arrows). The two saddle points shown in dashed red boxes can be found both by the general and by the special version of SPC. In SP1, the null-element has the same curvature as the second surface of M0; in SP2, it has the same curvature as the third surface of M0 (because of two zero thicknesses, we see there three overlapping surfaces, marked by red dashed line).

Optimization can be trapped in the sub-optimal minima M2 and M3, which are significantly worse than M1. However, the global minimum M1 can be rapidly obtained with SPC from both of them. In M2 or M3, the second element is extracted first and then the SPC method is applied (suitable candidate lenses for extraction are in general weaker-power lenses that seem to have little function). After extraction and optimization, both solutions become the M0 system shown in the green box in Figure 4.2. Then two saddle point systems, SP1 and SP2 in Figure 4.2 (in the red dashed boxes), were constructed from M0 by adding a null element in between the first and the second lens (i.e., at the position of extraction). From each saddle point system, two local minima can be

⁴The size of a basin of attraction can be expressed as the size of the set of starting points in the variable space that after optimization, converge to a given solution.

found after optimization. In this case, both saddle point systems lead on one side to the same minimum, the global minimum M1. On the other sides, the two saddle points lead to M2 and M3. In fact, from any of these three minima, the other two can be found by a switching operation. In this example, different ways of using the SPC method described in Chapter 2 are demonstrated: firstly, the SPC method can systematically find new solutions with more lenses (M1, M2, and M3) from a simpler system (M0); secondly, extracting and then adding lenses with the SPC can lead to different minima in a systematic way. This possibility of switching between different local minima is a typical property of the SPC.

4.3. DESIGNING A PINHOLE SYSTEM STARTING FROM A SINGLE LENS

In the previous section, by using SPC, we show how different designs can be found systematically by increasing the number of lenses. By using this approach in combination with traditional methods, it is possible to design an optical system starting with just one lens. This part shows how a pinhole lens similar to the one in Figure 4.1(a) can be obtained from a single lens. The specifications are listed in Table 4.1. The same merit function as before, is used.

A global optimization is performed starting from a plane parallel plate. Only one singlet solution is obtained, with a merit function value of 576.30, as shown on the left in Figure 4.3(a). This singlet serves as a basic element, providing optical power to the system [11]. On this singlet, SPC is performed by adding a zero-thickness glass element at the back surface. Two doublet minima are found, with merit function values 164.81 and 1783.94 (the latter one disappears if we replace EFL 3.5 mm by, e.g., EFL 3.0 mm). The better one (the first one) is chosen for the design step.

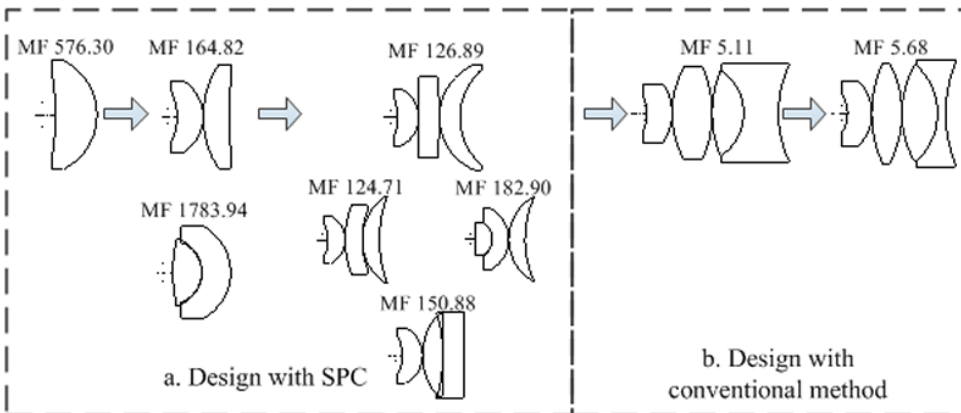


Figure 4.3: Designing a wide-angle pinhole lens from one lens by combining SPC with conventional methods.

With the same strategy, SPC is performed on four different surfaces of the doublet with a merit function value of 164.82 and four triplet minima are found (with merit func-

tion values 124.71, 126.89, 150.88, and 182.90, respectively). Only one type of glass is used for the steps mentioned above. In the next step, we continue with the best two minima, i.e., the local minima found with the lowest MF values, using conventional design techniques: glasses are changed and the third lens is replaced with a cemented doublet to correct chromatic aberrations. The glasses are selected to be the same as the original system designed by Irina Livshits. When the glass are modified, the best minimum (a merit function value of 124.71) of the triplets encounters ray failure. However, the system with a merit function value of 126.89 still remains stable after changing the glass. Optimization with all curvatures and thicknesses as variables leads to a system with a merit function value of 5.11 [Figure 4.3(b)]. By readjusting the thickness of the lenses for a more compact size, a system with the merit function value of 5.68 was obtained [see Figure 4.3(b)], which is identical to the original design in Figure 4.1(a).

4

SPC is useful for dealing with minima created by monochromatic aberrations. It cannot deal directly with minima created by chromatic aberrations. However, since typically chromatic aberrations are less non-linear than monochromatic ones (e.g., unlike the Seidel aberrations, for axial chromatic aberration the thin-lens expressions are linear), the chromatic aberrations tend to create fewer new minima than the monochromatic ones. As shown in this example, in the design process SPC can be easily combined with traditional approaches to handle chromatic correction.

In addition, four triplet minima are found with the SPC method in the search using only one glass, as shown in Figure 4.3(a). Both Global Synthesis⁵ and saddle point detection (NETMIN) finds only three minima (missing the one with the merit function value of 182.90). Therefore in this example, the SPC method has the advantage of being able to find more minima than the two methods used for comparison.

4.4. DECOMPOSING A HIGH-DIMENSIONAL SEARCH FOR NEW MINIMA IN A SUCCESSION OF ONE-DIMENSIONAL SEARCHES

It is mentioned in the previous section that the SPC found more system candidates compared to the Global Synthesis of CODE V and the saddle point detection algorithm. However in the examples, the number of minima found in the landscape is small. In this section it is investigated whether the performance of SPC is still good when more minima are present in the landscape.

In order to generate more minima, the FOV of the triplet system is reduced to 60° while the other specifications remained as that given in Table 4.1. The reason why a reduced field leads to more minima will be discussed in the next section. As in Figure 4.3(a) only one type of glass is used, the variables are only the curvatures. The same default CODE V merit function is used, and, for the purpose of this study, the control of edge thickness is disabled. All lenses are in contact (i.e., all air spaces between lenses have zero thickness). The thicknesses of all lenses in a minimum system are set equal in order to avoid the multiple appearance of essentially the same minima (with similar curvatures, but different lens thicknesses) that would unnecessarily complicate the

⁵It is a proprietary global optimization algorithm of CODE V. The algorithm starts from the given system and explore solutions space for new configurations in a systematic manner.

study [12].

For a better understanding of the results, in addition to the comparison of minima, it is also useful to compare the saddle points obtained using SPC with those obtained using the saddle point detection algorithm (realized by the program NETMIN). For SPC there are the options of inserting in the existing minimum either a glass null-element (i.e., inserting a lens) or an air null-element (i.e., splitting a lens). If the zero-thickness glass element was used, the saddle points resulting from the SPC scan would still have a zero-thickness lens, whereas the saddle points detected with NETMIN would have finite thicknesses for all lenses. This is because that the NETMIN starts from existing minima with non-zero thickness element and searches for the saddle points. The saddle points obtained from the two methods can not be directly compared due to the thickness differences of the lens element. In our study, all lens elements are set to be in contact. In this case, the saddle points detected with NETMIN will have zero air spaces. It is therefore better for comparison to study the performance of the SPC when air null elements are used. Splitting lenses with SPC leads to saddle points with a zero distance between lenses, which are directly comparable with the saddle points from NETMIN. For performing SPC with a zero-thickness air element, the thickness of the lens, which will be split, is first doubled. Then the system is re-optimized to a minimum, and a zero-thickness air element is inserted in the middle of the lens. As mentioned in section 3.3, the example of Cooke triplet shows that the minima are different between applying zero-thickness glass and zero-thickness air space constructions. Therefore, both approaches are recommended to be used. However, in this example, it turns out that SPC with zero-thickness glass does not give more local minima than using zero-thickness airspace. For simplicity and better comparison, the SPC result using zero-thickness air space is shown.

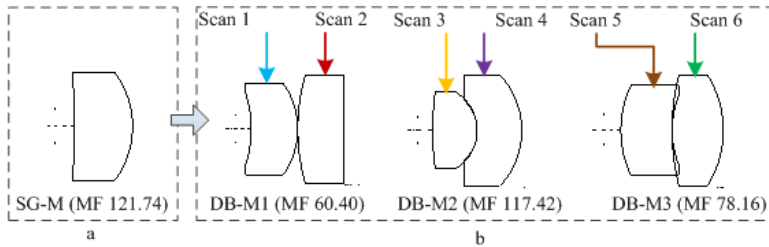


Figure 4.4: SPC by inserting a zero-thickness air element. (a) Singlet solution (SG-M). (b) The three doublet minima (DB-M) that result from the singlet are the starting minima for the scans shown in Figure 4.5.

As previously, we start with a single optimized lens [SG-M shown in Figure 4.4(a)] and three doublets are obtained with the SPC [shown in Figure 4.4(b)]. The three doublets DB-M1, DB-M2, and DB-M3 have a total of six lenses. Each of the six lenses is used to generate an SPC scan which leads to triplet minima. The colored arrows in Figure 4.4(b) indicate the lenses that are split (after the thickness doubling that is not shown here) in the corresponding scan.

Figure 4.5 shows the minima and saddle point systems found in the triplet design

space with NETMIN and Global Synthesis (Global Synthesis has found nine out of 10 minima, missing the minimum M6). Five curvatures are used as variables, and the curvature of the last surface is controlled by a *Solve* in CODE V to keep the effective focal length constant.

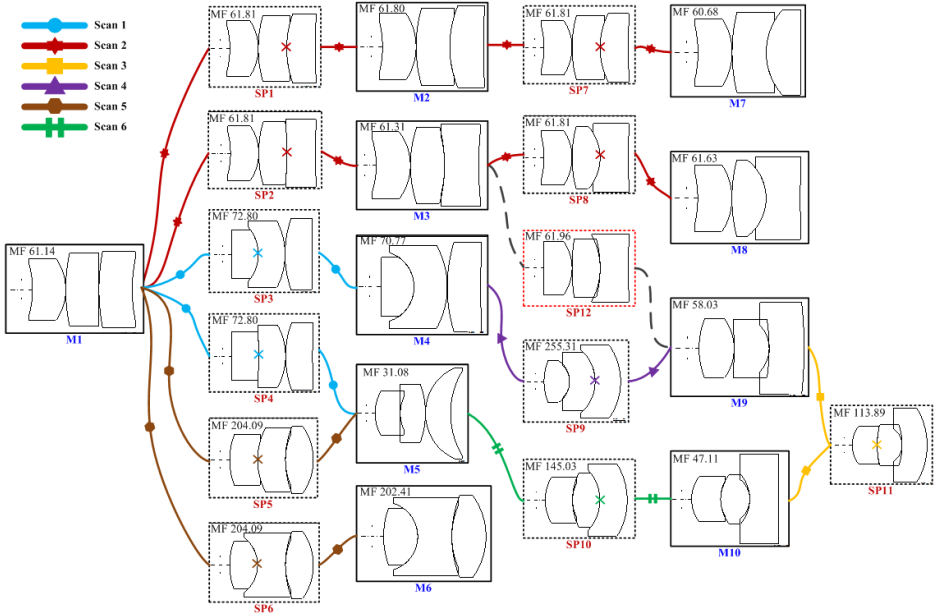


Figure 4.5: Network of minima (solid boxes, named as M#) and saddle point systems (dashed boxes, named as SP#) in the 60° landscape, where the other specifications are the same as in Table 4.1. Eleven saddle point systems result from six SPC scans (each color represents one scan). Optimization starting at these saddle point systems leads to all 10 minima that are obtained with NETMIN and Global Synthesis. If in any of these 11 saddle points we remove the null-element of the corresponding scan (this does not affect the ray paths or the merit function), we obtain the starting doublet of the scan. The null elements are marked by crosses with the corresponding color. This figure shows why the lens design landscape is different from general global optimization landscapes: because of the close relationship that exists between local minima of a design problem (here the 10 triplet minima) and local minima with one lens less (here the starting doublets).

A significant result of this search is that all 10 triplet minima shown in Figure 4.5 are also obtained with SPC starting from the three doublets in Figure 4.4. Eleven saddle point systems are obtained from the six one-dimensional SPC scans. Different scans find different numbers of saddle points. For instance, in Scan 2 there are four saddle points (SP1, SP2, SP7, and SP8 in Figure 4.5), whereas in other scans, only one or two saddle points are found. The 10 minima are resulted from the 11 saddle point systems by optimising on both sides of the saddle. Saddle point systems and minima obtained from the same scan are connected in Figure 4.5 with the same colored lines. In the figure, when we gradually increase the FOV, M1 corresponds to the system leading to the final design in the 110° case [merit function value 126.89 in Figure 4.3(a)]. However, in this case, its merit function value of 61.14 is not the smallest one in Figure 4.5.

NETMIN finds the extra saddle point system SP12, which cannot be found by SPC.

This saddle point system connects with a dashed link the minima M3 and M9 in Figure 4.5. However, there is sufficient redundancy in the SPC approach, and both M3 and M9 also result from other saddle point systems which are found by SPC. Therefore the lack of ability of SPC to find SP12 is not critical.

Since there are five variables, with general global optimization tools this search has to be performed in a five-dimensional space. In this example, however, all minima found with other methods are also found with SPC, where the five-dimensional search is replaced by six one-dimensional searches. Replacing a high-dimensional search for new minima by a succession of one-dimensional searches reduces the complexity of the search significantly.

The example discussed above shows the utility of the feature of the optical merit function landscape that enables the decomposition of the search for many of the minima in simpler steps. The feature means that minima and saddle points are all connected, and the saddle points can be constructed by SPC. This example has the advantage that in this case, the feature mentioned above can be observed in a pure form, without interference from other features (e.g. the minima are not connected by saddle points that can be constructed) of the landscape. However, in general, other features which deserve further study to gain a better understanding, may also play a role. Therefore, it cannot be expected that we can always find all minima using SPC. Nevertheless, even when other features are present, examples studied so far show that SPC can lead to satisfactory results.

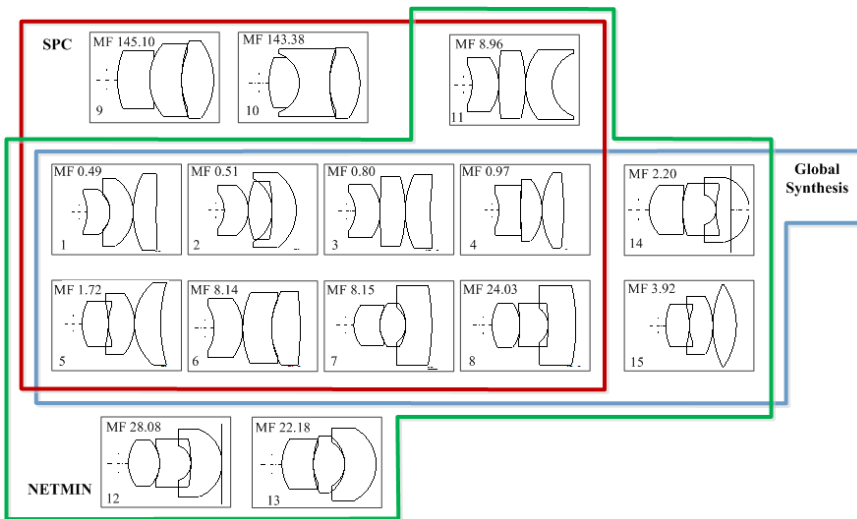


Figure 4.6: Minima found with SPC (red box), Global Synthesis (blue box), and NETMIN (green box) in a monochromatic run. Minima 12 with Merit Function (MF) value 28.08 and Minima 14 with MF value 2.20, whose image planes are inside or coincide with the last lens, are unphysical. The five minima with the lowest MF value (numbered from 1 to 5) are found by all three methods. Edge thickness violations is disabled in this search.

For instance, when the search described in Figure 4.5 is repeated with modified wave-length specifications, SPC finds 11 minima and Global Synthesis finds 10 (see Figure 4.6).

Among these minima, eight, including the best five, are found by both methods. For the four minima that are missed by SPC, two of them are unphysical (minima 12 and 14, they have either negative or zero back focal length that cannot be corrected, but should be prevented with an additional constraint). These unphysical solutions also have a merit function more than 4 times and more than 50 times that of the best solution. The other two minima missed by SPC (minima 13 and 15) have merit function values 40 times and 8 times that of the best solution. NETMIN finds all four solutions missed by SPC, and Global Synthesis finds two of them.

Despite the fact that SPC cannot find four (poor-quality) minima directly, it turns out that by using the switching strategy described earlier, it is easy to escape from any of them if optimization is trapped. Eliminating a lens from any of these poor-quality minima leads to one of the three doublets. From these doublets, better minima can be obtained with SPC.

However, in order to understanding the potential and the limitations of SPC, it is useful to examine why SPC is not able to reach these four minima. SPC is based on the assumption that small changes in lens thicknesses do not affect good designs significantly, i.e., that good minima continue to exist as minima when a reasonably small non-zero thickness is replaced by zero. This assumption is similar to the one at the basis of the well-known thin-lens aberration theory. In earlier days of lens design, primary aberration formulas using zero thickness were used to qualitatively predict the existence of good designs.

It turns out that the minima that are not found by SPC only exist for certain non-zero thickness values and do not have a zero-thickness equivalent. However, the existence of these minima does not contradict our assumption that using SPC can obtain satisfactory solutions, because these are not among the best solutions. In the numerical experiment, final solutions have the same thickness for each lens element (1.5 mm). Since in the SPC procedure the practical design is always obtained after increasing the zero-thickness of a null element, it is not possible to reach minima that disappear when the lens corresponding to the null element has zero thickness. In NETMIN and Global Synthesis searches, the lens thicknesses are kept constant. This typical behavior is illustrated in Figure 4.7(a) and (b). In the system in Figures 4.7(a) found by SPC, the thickness of the first lens is 1.5 mm. By starting with a zero-thickness null-element as the first lens as in Figure 4.7(b), increasing the thickness leads to the system in Figure 4.7(a). By decreasing the thickness of the first lens of the system in Figure 4.7(a), the system in Figure 4.7(b) is obtained. In contrast, in the system in Figure 4.7(c) found by NETMIN but not by SPC, when the thickness of the first lens is decreased ray failure occurs and a system with zero-thickness element cannot be obtained. However, systems like the one in Figure 4.7(c) have low stability. A small perturbation on the system in Figure 4.7(c) will lead to the system in Figure 4.7(a).

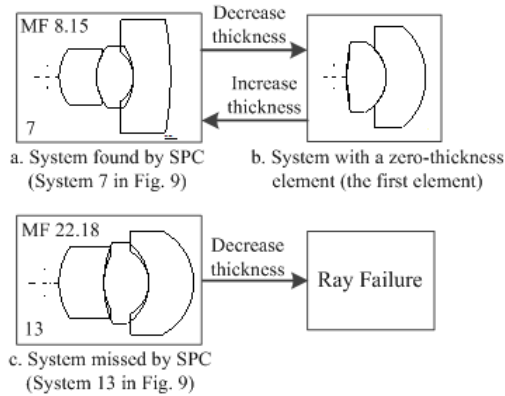


Figure 4.7: Increasing the thickness of the first element of system (b) results in system (a), and vice versa. System (b), which is obtained by the SPC, is a minimum with a zero-thickness element. System (c), which is missed by the SPC, does not have a corresponding zero-thickness pair. A small perturbation on system (c) will lead to system (a).

4.5. EFFECT OF CHANGING THE FIELD OF VIEW ON THE LANDSCAPE

In Section 4.3 and 4.4 it is found that when the field of view is reduced, more minima can be found in the design space. This happens for two reasons. First, when design specifications (e.g., FOV or EPD) are modified, the distances in the design space between minima and saddle points change and therefore the merit function landscape also changes. As can be observed in numerical experiments, when a minimum and a saddle point with Morse Index 1 collide, they both disappear. Alternatively, such a pair can appear when specifications are changed. Similarly, two minima with a saddle point between them can be replaced by one minimum, and two saddle points with a minimum between them can be replaced by an SP. This can be explained mathematically by the conservation of the topological degree of the merit function landscape [13] [14]. The topological degree of a critical point is the sign of the determinant of the Hessian matrix in that point. The degree of a critical point is thus $(-1)^{MorseIndex}$. A minimum has a topological degree of +1 and a saddle point with Morse Index 1 has a topological degree of -1. The sum of the topological degrees remain constant under perturbation such as changing the FOV. Second, when the FOV is increased ray failure appears more easily, and the region of ray failure expands in the design space. Because certain minima and saddle points which exist for lower FOV cross the border of the ray failure region and disappear, fewer of them are found at large FOV.

When varying the FOV of the systems shown in Figure 4.5, the number of solutions changes with the FOV as shown in Figure 4.8. Three different FOVs are chosen: 110°, 90° and 60°. Six solutions were obtained for 110°, five for 90° and ten for 60°. From Figure 4.8, it is seen that some system shapes exist in all the three FOV configurations. However, other system shapes only exist in the specific FOV range. For FOV 60°, five new systems, which did not exist for a larger field (second row for the 60° systems in Figure 4.8). The

dashed boxes in Figure 4.8 mark the missing systems which exist in other FOV specifications. As mentioned in the previous paragraph, the appearance and disappearance of the systems can be explained by the change of the merit function landscape.

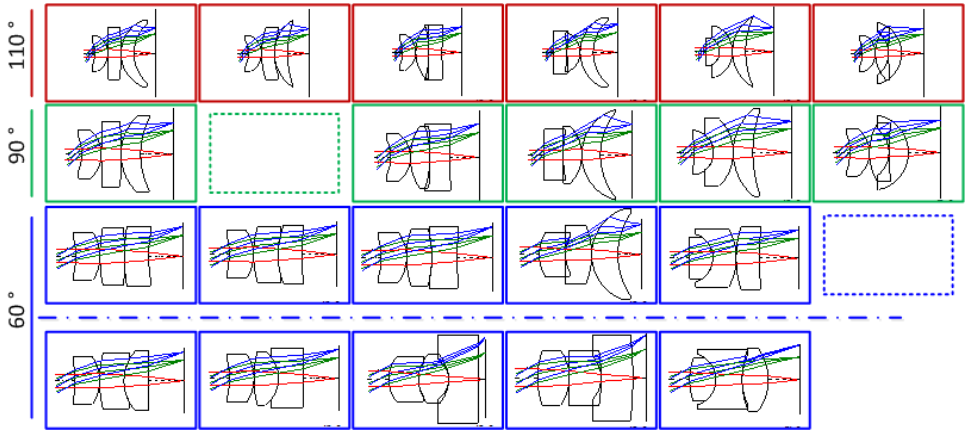


Figure 4.8: For different field of view, the number of solutions is different. Systems with 110° FOV have six solutions marked by the red colour; Systems with 90° have five solutions marked by the green colour; Systems with 60° marked by the blue colour. The dashed boxes mark the missing systems.

These phenomena can also be observed by examining the SPC scan curves. Figure 4.9 refers to Scan 2 of Figure 4.5 where five minima are linked by four saddle points, and the FOV is 60° . In Figure 4.9, the scans are performed in different FOV settings: the FOV of the system increases from 50° [Figure 4.9(a)] to 90° [Figure 4.9(d)]. With the increase of the FOV, the ray failure region (shown as shaded area) is seen to expand. Two saddle points are found at 70° [Figure 4.9(c)], however, at 90° [Figure 4.9(d)] the most left saddle point disappears into the ray failure region.

A different kind of change is also observed. At 50° [Figure 4.9(a)] four saddle points can be found. However, at larger fields the third and fourth saddle point first move closer [Figure 4.9(b)], merge and then disappear [Figure 4.9(c) and (d)].

By optimizing from the saddle points, at 50° it is found that the last two saddle points lead on one side of the saddle to the same minimum. While changing the FOV in small steps, it is possible to keep track of the two saddle points and one minimum by keeping the norm of the gradient to zero via minimization. The process is elaborated in Figure 4.10. When the FOV increases from 50° to 60° , it is seen that the Euclidean distances between the saddle points and minimum reduce. Hence the critical points are getting closer in the high-dimensional design space. At 70° , it is no longer possible to find saddle points via an SPC scan. However, it is still possible to obtain the merged point by keeping the norm of the gradient to zero at the vicinity of the previous critical points. From Figure 4.10, it is seen that the original two saddle points and one minimum merge into one saddle point with a merit function value of 74.92. It is a saddle point without any null elements, hence it cannot be constructed using SPC. This shows how two saddle points and one minimum merge into one saddle point while the FOV increases.

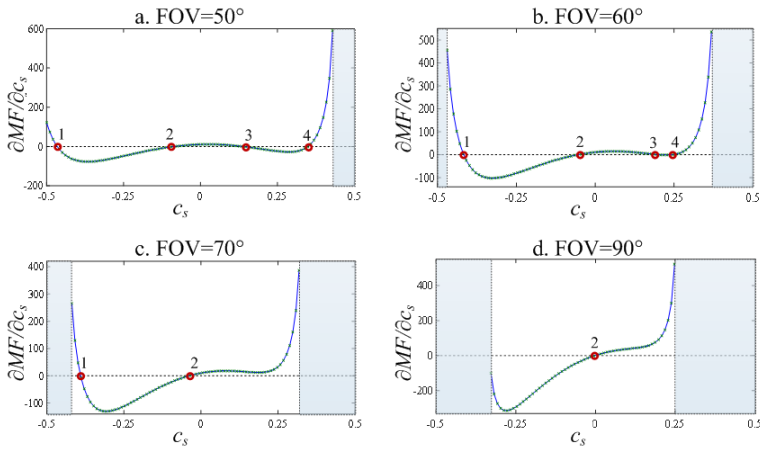


Figure 4.9: SPC scan curves for different values of the FOV. Shaded areas are ray failure regions. The scan curves are obtained in the same way as the one shown in Figure 3.2. Saddle point curvatures are marked by red circles. With the increase of the FOV, saddle point 1 disappears in the ray failure region. SP3 and SP4 merge with the minimum between them, and are replaced by a saddle point (not shown) that is not constructible with SPC.

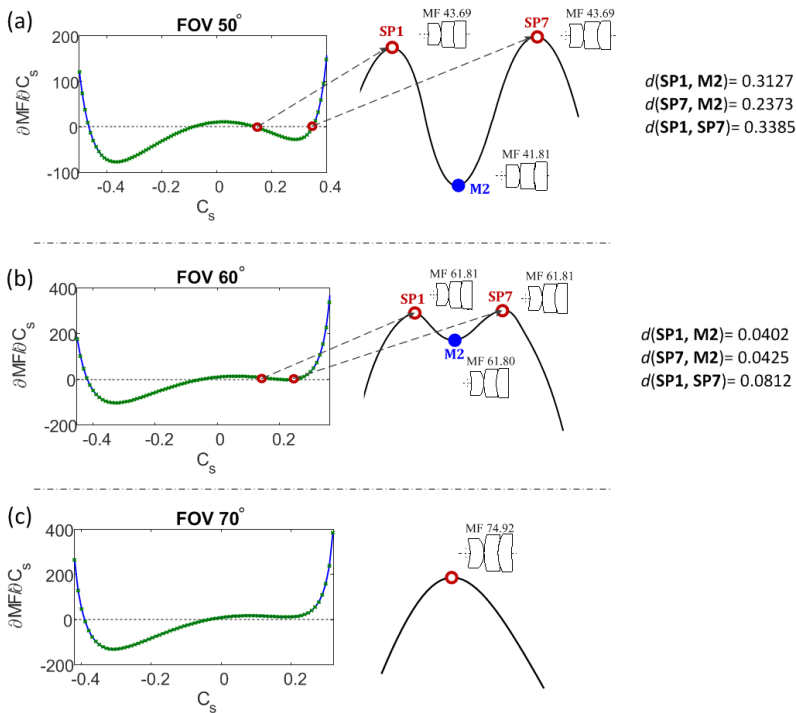


Figure 4.10: Merging of two saddle points and one minimum when the FOV increases, from Figure (a) to Figure (c). In the SPC scan curves, the two saddle points get closer and then disappear. The two saddle points and one minimum in between merge into one SP. The decreasing Euclidean distances also indicate the merging.

A different phenomenon is observed when FOV increases. In Figure 4.11, the two saddle points on the left side of the scan curve are investigated: At 54° , SP8 and SP2 both lead to M9 when optimizing on one side. However, when the FOV increases to 56° , the two saddle points are connected to M3 instead of M9. The minimum M9 is connected to M3 via SP12. M3 and SP12 do not exist when the FOV is 54° . We see that when the FOV increases, a pair of local minimum (M3) and saddle point (SP12) start to emerge. Their merit function values are very close to each other (54.06 vs. 54.07) at 56° FOV. When the FOV increases to 60° , the difference between merit function values increases (61.31 vs. 61.96) as well as the Euclidean distance between M3 and SP12 (increases from 0.1148 to 0.2611).

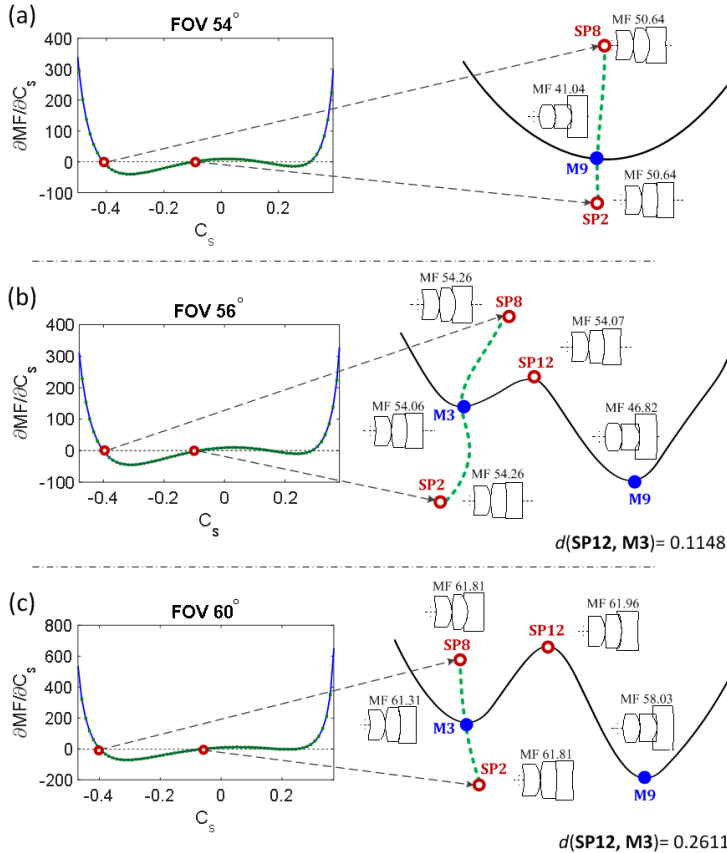


Figure 4.11: A pair of minimum - saddle point emerges when FOV increases. From (a) to (c), the FOV increases. M3 and SP12 start to appear and form the link M3-SP12-M9. The connection between saddle points and minima are altered: at 54° , SP8 and SP12 connect to M9; at 56° , SP8 and SP12 connect to M3.

Besides merging or splitting, it is also found that in some special situations, when the FOV increases, both approaching and distancing⁶ happens for the same pairs. Figure

⁶When mentioned merging, we observed that a pair of points approach each other and reach a situation where they cannot be distinguished anymore. When mentioned splitting, we observe that one point splits into two

4.12 illustrates a scenario when a glass null element was used as an SPC scan element. The FOV is increased from 60° to 110° , where three steps (60° , 90° and 110°) are plotted in the figure. At 60° , two saddle points were found by the scan. Both of them lead to a common minimum at one side (illustrated on the right of the scan curve). With the increasing of the FOV, the two saddle points become closer in the design landscape: The Euclidean distance reduces from 0.0785 to 0.0023 when the FOV increases from 60° to 90° . At this phase, a much smaller step size of the scan is necessary to be able to find the saddle point. After the FOV increased to 110° , the Euclidean distance between the two increases from 0.0023 to 0.0217. As a result, the tendency of the merging of the two saddle points changes to splitting.

It is seen in the aforementioned three examples that the change of the design landscape can be dynamic and unpredictable. When a certain design specification changes, it essentially alters the design landscape. This change may affect the number of saddle points and minima which can be obtained with the SPC approach. Nevertheless, for the given examples, redundancy of the saddle point - minimum links also reduces the risk of missing solutions with the SPC approach. All the good solutions were obtained with the SPC method.

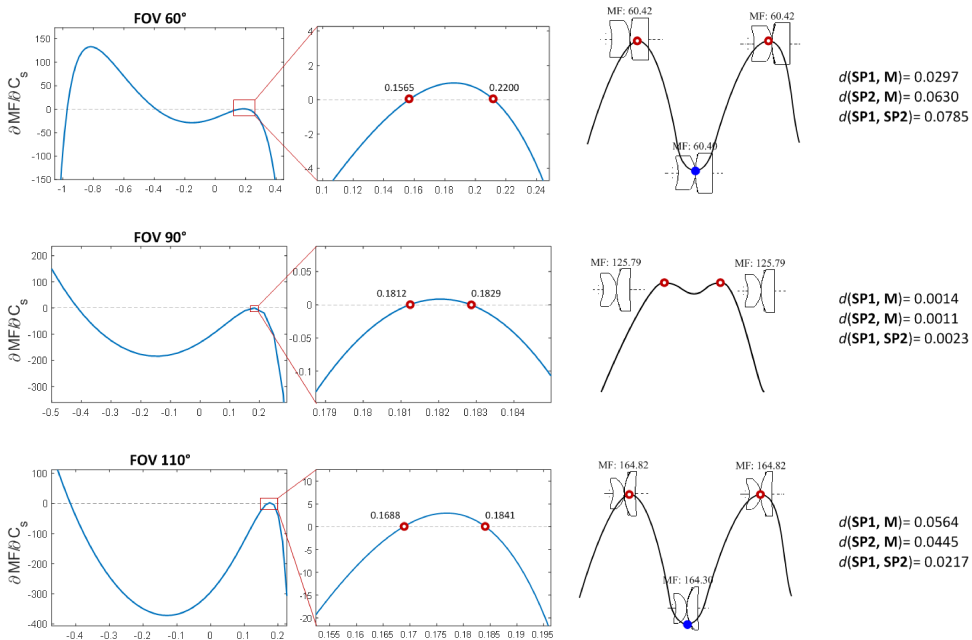


Figure 4.12: A glass null-element SPC scan is performed on DB-M1 from Figure 4.4. The scan position is between the two lens elements. When the FOV increases from 60° to 110° , the two saddle points on the right first get closer, and then further apart. This phenomenon can be observed from the SPC scan curve as well as from the Euclidean distances between the two.

and they move away from each other. When this happens in a continuous way with a change of the system setting (e.g. FOV), approaching and distancing is used to describe the behaviour of the two points.

4.6. DISCUSSION AND CONCLUSION

The design landscape of a wide-angle pinhole lens and its closely related optimization landscape are investigated in this chapter. Several observations that result from this study have a more general validity. Switching between minima with SPC introduced in Subsection 3.3.1, is applied to the pinhole lens as shown in Figure 4.2. It can be used in other design tasks as well to avoid the trapping of the optimization in a sub-optimal solution. Combining SPC with conventional design methods, as shown in Figure 4.3, can lead to good complex designs starting from simpler ones [15].

The lens design landscape is different from general global optimization problems because of the close relationship that often exists between the local minima of a design problem and local minima with one lens less. Figure 4.5 is particularly useful to illustrate this structure, due to the special choices that are made in the corresponding search (lenses in contact, use of air null elements). The systems SP1–SP11 are triplet saddle points, but if the air null-element of the corresponding scan (a pair of surfaces that has no influence at all on the rays or on the merit function) is removed, they become doublet minima. However, when these doublet minima with an extra null-element are optimized on both sides of the resulting saddle, they lead to the triplet minima M1–M10. In the search shown in Figure 4.5, these 10 triplet minima are all minima found in the landscape with other methods such as NETMIN and CODE V that are used for comparison. A scenario where this structure holds for all minima has been previously encountered [3]. The corresponding landscape is dominated by spherical aberration. The wide angle examples discussed here show that the scenario where this structure is present for most of the minima, can be much more general.

The example shown in Figure 4.5 has the advantage that the special structure can be observed in a pure form, where all minima and saddle points are connected and the saddle points can be constructed using SPC. Interference of other features can also be present as seen in the examples: Figure 4.7 indicates that some minima do not have a corresponding system with a zero-thickness element, therefore cannot be obtained following the SPC procedure; Figure 4.9 shows that saddle points which cannot be constructed with SPC are present. In these examples, using SPC cannot always find all minima. Nevertheless, Figure 4.6 shows an example where even in the presence of the aforementioned features of the landscape, the five solutions with the lowest merit function value found with other methods can still be found with SPC.

The example shown in Figure 4.9 is useful for understanding the transition in the local landscape where saddle points that can be constructed using SPC become ones that cannot be constructed using SPC. In an extreme case where none of the saddle points can be constructed using SPC, the landscape becomes a standard global optimization landscape which does not have the special structure we mentioned. The replacement of the SP3 and SP4 in Figure 4.9 by a saddle point, which cannot be constructed with SPC, shows that the landscape is locally transformed into such a standard global optimization landscape. The process is elaborated in Figure 4.10. However, the transition can also happen in the opposite direction if the process (in this case, decreasing the FOV) is reversed: A usual saddle point splits into two saddle points and a minimum in between, and in this case, the two saddle points can be constructed using SPC. This complexity of

the design landscape dynamic is further explored in Figure 4.11 and Figure 4.12: When the FOV is increasing, a pair of minimum - saddle point appears; Two saddle points and a minimum first tend to merge and then tend to separate again.

Based on the examples shown in this chapter, the complex dynamic of the design landscape can already be observed. However, the structure that enables SPC to obtain good solutions is stable. For systems that are more complex than those presented in this chapter, the utility of SPC will be determined by its practical success, rather than by a detailed study of the entire landscape, which becomes difficult. Systems with bigger complexities are investigated in the next chapter.

The replacement of high-dimensional searches by several one-dimensional searches plus local optimization can lead to a reduction of the complexity of the search. It can be essential for practical purposes. Compared to global optimization methods where multiple starting points are firstly chosen heuristically and then local optimization is applied, SPC operates in a more systematic way that starting points are provided based on constructed saddle points. In computer programs where the search for new solutions is automated, the general version of SPC could be combined with other methods. In the case of the special version of SPC this has already been achieved in the commercial software SYNOPSIS [16]. Excepting simple problems, no known global optimization method can guarantee finding all (good) solutions in a reasonable time, and SPC is no exception. However, high-quality designs have already been obtained with the special [17][18] and with the general versions [15] of SPC. SPC will finally be validated if many high-quality designs are obtained with this method.

REFERENCES

- [1] Z. Hou, I. Livshits, and F. Bociort, *One-dimensional searches for finding new lens design solutions efficiently*, *Appl. Opt.* **55**, 10449 (2016).
- [2] F. Bociort, A. Serebriakov, and M. van Turnhout, *Saddle points in the merit function landscape of systems of thin lenses in contact*, *Proc. SPIE* **5523**, 11 (2004).
- [3] P. van Grol, F. Bociort, and M. van Turnhout, *Finding order in the design landscape of simple optical systems*, *Proc. SPIE* **7428**, 11 (2009).
- [4] F. Bociort and P. van Grol, *Systematics of the design shapes in the optical merit function landscape*, *Proc. SPIE* **7717**, 14 (2010).
- [5] F. Bociort, *Why are there so many system shapes in lens design?* *Proc. SPIE* **7849**, 13 (2010), this paper introduces the simplified toy model to explain the fundamental triplet network, and shows the increasing number of local minima in lens design.
- [6] F. Bociort and M. van Turnhout, *Finding new local minima in lens design landscapes by constructing saddle points*, *Opt. Eng.* **48(6)**, 063001 (2009).
- [7] M. van Turnhout, P. van Grol, F. Bociort, and H. P. Urbach, *Obtaining new local minima in lens design by constructing saddle points*, *Opt. Express* **23**, 6679 (2015).

- [8] M. Strauch, I. Livshits, F. Bociort, and H. Urbach, *Wide-angle spectral imaging using a fabry-pérot interferometer*, *Journal of the European Optical Society - Rapid publications* **10** (2015).
- [9] T. G. Kuper and T. I. Harris, *Global optimization for lens design: an emerging technology*, *Proc. SPIE* **1781**, 15 (1993).
- [10] J. R. Rogers, *Using global synthesis to find tolerance-insensitive design forms*, *International Optical Design*, TuA4 (2006).
- [11] I. Livshits and V. Vasilyev, *Q and A tutorial on optical design*, *Advanced Optical Technologies* **2**, 31 (2013).
- [12] Z. Hou and F. Bociort, *Reducible complexity in lens design*, *Proc. SPIE* **9626**, 8 (2015).
- [13] M. van Turnhout, *A systematic analysis of the optical merit function landscape: towards improved optimization methods in optical design*, *Ph.D. dissertation (Delft University of Technology)*, 63 (2009).
- [14] T. H. Koornwinder, *The topological degree of a mapping*, *Nonlinear Analysis* **Vol. 1**, 7 (1976).
- [15] I. Livshits, Z. Hou, P. van Groel, Y. Shao, M. van Turnhout, P. Urbach, and F. Bociort, *Using saddle points for challenging optical design tasks*, *Proc. SPIE* **9192**, 8 (2014).
- [16] D. C. Dilworth, *Novel global optimization algorithms: binary construction and the saddle-point method*, *Proc. SPIE* **8486**, 8486 (2012).
- [17] O. Marinescu and F. Bociort, *Saddle-point construction in the design of lithographic objectives, part I: method*, *Opt. Eng.* **47(9)**, 093002 (2008).
- [18] F. Bociort, M. Van Turnhout, and O. E. Marinescu, *Method of designing a projection system, lithographic apparatus and device manufacturing method*, (US 7714307 B2, May. 2010).

5

APPLYING SPC IN COMPLEX SYSTEMS

If you cannot measure it, you cannot improve it¹.

Lord Kelvin

In Chapters 3 and 4, suggestions are given on how to apply SPC to simple design problems. In this chapter, SPC is applied to some more complicated practical design problems.

5.1. INTRODUCTION

For a simple optical design system, such as a doublet or a triplet, the number of variables is rather small. It is possible to use different design (optimization) strategies (including SPC) to search for all or most of the local minima. So for these relatively simple systems, the results produced by SPC can be compared with those produced by other methods. This was done in Chapter 4, where the design network of a simple wide-angle pin-hole lens [2] shows that SPC is effective in escaping from poor local minima and reaching the best minimum. In a specific case, SPC is able to find all the solutions found in the design landscape that other methods also find. The number of variables used in this study was five (only curvatures). However, in a practical problem where a larger number

Parts of this chapter were presented in the SPIE conference Optical Systems Design (2018, Frankfurt) and published in Proc. SPIE **10690**, Optical Design and Engineering VII, 1069007 (2018) [1].

¹The quote is popularly rephrased from the original quote "When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be."

of variables is used, it is difficult to successfully repeat a similar study. One of the reasons is that the number of local minima increases dramatically with the number of variables. It is not possible to find every minimum in a complex design landscape, therefore when we compare the SPC results with the results obtained with different methods, one should focus on the ability of SPC to find good design solutions.

Despite the preceding observations, the SPC method can be applied in a complex lens design problem. Attempts have been made to apply the special version (Chapter 3.2.2) of SPC to some complex objective designs[3][4], and the results of general version (Chapter 3.2.1) has also been shown to produce good systems[5]. However, in all cases, the details of how SPC should be applied effectively is not explained, therefore it is difficult to quantify the performance of the method, and reproducing the same results from the mentioned studies is not possible.

In this chapter, we derive recommendations for how to apply SPC to several practical design problems, namely: 1) A wide-angle system consisting of two groups of lenses where SPC can only be used for a limited number of positions in the system; 2) Two microscope objectives, where designer experience and SPC produce additional solutions; 3) A complex lithographic objective where applying SPC shows improvement in the intermediate design stage.

5.2. INVESTIGATION OF A WIDE-ANGLE LENS WITH SIX ELEMENTS

Wide-angle lenses are broadly used in various applications, including photographic cameras, surveillance cameras and projectors. To increase the FOV of an optical system and while maintaining an image of good quality, a certain amount of complexity of the system is necessary. We consider here a wide-angle lens with a moderate complexity to study the performance of SPC as an approach to switch between solutions with the same number of variables. The system consists of six lenses including one cemented surface. There are eleven surfaces, hence eleven variables in terms of curvatures. The design specifications of the lens are shown in Table 5.1 and the surface specification is available in Appendix A.2. The maximum FOV is 120°. The overall length of the system is 38.75mm, which is small enough for integration into surveillance systems. The Modulated Transfer Function (MTF) curve shows that the modulation stays above 0.4 for 100 cycles/mm at full FOV (5.1). The system is not optimized for distortion which can be corrected with image processing afterwards (e.g. digital dewarping [6]).

Table 5.1: System specification of the wide-angle lens

(Full) Field Of View (FOV, °)	120
Effective Focal Length (EFL, mm)	3.0
F Number	2.0
Overall Length (mm)	38.75
Wavelength (nm)	644, 546, 480

5.2.1. DESIGN A WIDE-ANGLE LENS WITH SPC

The wide-angle lens is first designed with a conventional method by using surfaces with special properties (concentric and aplanatic) [7]. Surfaces, at which the chief ray has everywhere a normal incident angle, are free from coma and astigmatism. We call such an optical surface a concentric surface. The aplanatic surfaces in the system do not introduce spherical aberration and coma. In contrast, as will be seen in what follows, SPC can be used successfully in lens design without the need of considering the detailed aberrational properties of individual surfaces. In the system plot of Figure 5.1, it is shown that the six lenses are divided into two groups. A front group with three lenses is responsible for expanding the FOV. A rear group containing a cemented element is the focusing group. The two parts of the system were first designed separately and then combined to form the complete lens system. To combine the two parts, the front group has the stop at the end of the lens group, and the rear group has the stop in front of it. The front group is designed as an afocal system with an angle reduction ratio of 0.33. The FOV on the object side is designed to be 120° and the FOV on the image side is 40° . The rear group is designed as a focusing system with a FOV of the same 40° . The diameter of the common stop of the two parts, which is the aperture stop of the whole system, is 4.5 mm.

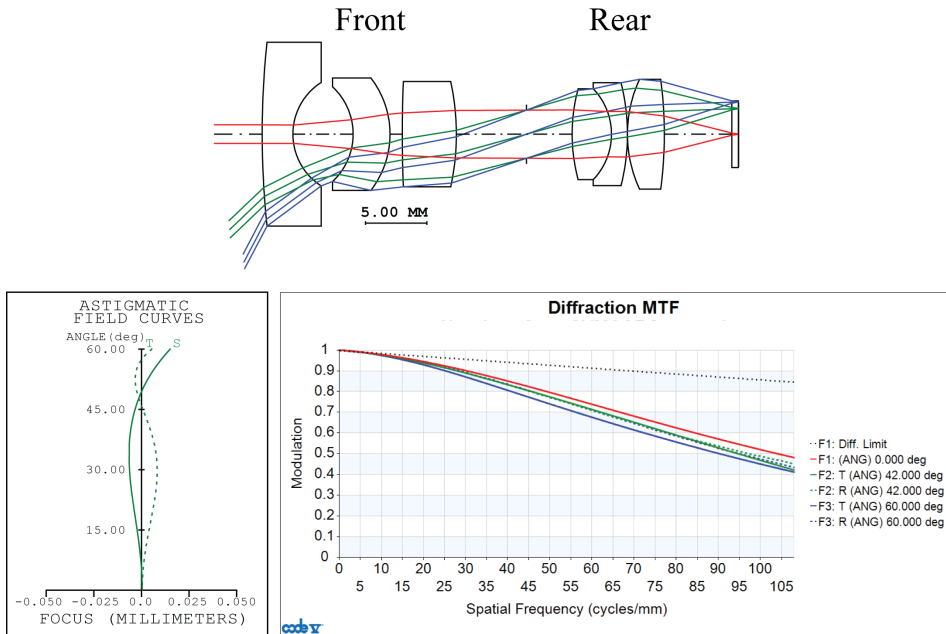


Figure 5.1: System plot of a wide-angle lens and its performance.

With the conventional method, the design of the front group started with a single lens using a flat and a concentric surface, given that the stop is at the right side of the lens. A moderate FOV is first used. Then lenses are added with either concentric or aplanatic surfaces. We gradually increase the FOV in small steps and optimize the system after each step to keep the aberrations small. Experimenting with different values has led to

the conclusion that a good choice for its angle reduction ratio is 0.33. Values that differ significantly from 0.33 lead to ray failure, or cause that the specifications cannot be met. An angle reduction ratio of 0.33 is used as a constraint in the merit function in the later steps. The optimization variables are the surface curvatures, spacing and thicknesses of the lenses. The rear group is designed separately by starting with a cemented flat surface doublet and two concentric surfaces. Another lens is added and a different glass is used to further reduce the monochromatic and chromatic aberration. The front group and rear group are then combined to optimize for the final solution.

When a solution is obtained by the conventional approach, SPC can be used to investigate alternative solutions. Since the system is combined from two separate parts, one way is to apply SPC to both the front and rear parts alone, without the other part being present. After that, different combinations of the front and rear parts can be investigated. Another way of applying SPC is to search for new solutions for the combined system.

After applying SPC to the front group and rear group separately, we obtain one solution for the front group and three solutions for the rear group. The next step is to combine the front group option with the rear group options. Note in Figure 5.1 that for the front group the aperture stop of the system is the exit pupil of the group, and that for the rear group it is the entrance pupil. Therefore, recombining the two parts as shown in Figure 5.2 has been done by ensuring that the exit pupil of the first group, and the entrance pupil of the second group are both located at the position of the aperture stop of the entire system. After that, the whole system has been reoptimized. The merit function used in the three cases (front group, rear group, entire system) is of the same type (i.e. it is based on transverse ray aberrations), however, the choice of the rays used in the construction of the merit function is different. The ray choice has been made automatically by the software depending on the different field specifications (the aperture stop size is the same in all three cases). All three cases provide the same final solution as illustrated in Figure 5.2.

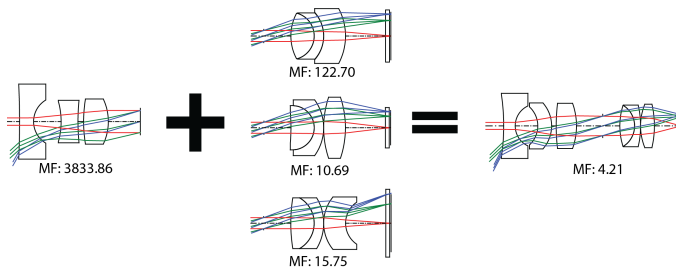


Figure 5.2: Combinations of the front part (conventional approach) and the three rear parts (conventional with SPC). All converged to one solution. Merit Function (MF) values of the systems are provided.

Once the combined system is obtained, SPC is performed to search for alternative solutions with the same amount of variables. As explained in Chapter 3, the first step is to extract lenses from the original system. Because the first lens is responsible for collecting the large angle rays, extracting it from the system while keeping the large FOV

is not possible. The only possible lenses to extract and perform SPC scans are the second, third and last elements. The current SPC cannot be performed on cemented elements for switching to a different solution. It is because a cemented lens requires the two sub-elements to be in contact (having a shared surface). A null element cannot be created in such a situation while varying the curvatures for the SPC scan. The merit function based on transverse aberration in CODE V is used in this example. It is a composite value, scaled so that it is the mean square of the weighted image radius (unit: μm^2). We keep the paraxial marginal ray exit angle as -0.25 rad on the last surface to ensure that the EFL is equivalent to 3.0 mm (the entrance pupil diameter is 1.5 mm). Ten curvatures are used as variables. Six different minima are found in the design network illustrated in Figure 5.3. A global optimization run (Global Synthesis of CODE V) is also performed. No additional local minima are found compared to the result of SPC. From Figure 5.3, it is noticed that the major difference in the systems are in the front group. The system shape stays the same for the rear group in all six solutions.

The relationship between different systems is indicated by arrows in Figure 5.3. Black arrows show that via SPC, two systems are connected in both ways. If M1 and M2 are connected in both ways, it means that by using SPC, M2 can be found starting from M1, and vice versa. Blue arrows indicate a one-way connection. For example, starting from M2, it is possible to find M3 via SPC. However, M3 does not lead to M2 by using SPC. Normally, the expected situation is a two-way connection. However, a one-way connection can appear in two scenarios: In the first scenario, a system is unstable because it has high stress and higher value of merit function (e.g., for M6, the high stress is indicated by consecutive large surface curvatures in the front group). The stability relates to the size of the basins of attraction. An unstable local minimum has a relative smaller basin of attraction compared to others. Therefore, it is easy to get out of this kind of poor local minimum with SPC. However, it is not easy (and not necessary for practical purposes) to reach them. The second scenario happens when two minima are very close to each other in the design space (e.g. M1 and M2). The landscape has relative smaller variation in this region around the two close minima. In Figure 5.3, systems M1 and M2 (best two systems) have similar system shapes and close values of the merit function. Another system, such as M3 or M4, obtained by switching from M2, can easily, when switched back, go to M1.

As shown in Figure 5.3, all minima are connected within one network. It means that starting from any of the minima in the network, there is always a route to reach any other system. It can also be observed that the less good solutions (M3, M4, M5, and M6) are all directly connected with best two systems (M1 and M2). This implies that if the design process was trapped in one of the poor solutions, applying SPC once would be sufficient to switch to one of best two systems.

5.2.2. STEPWISE ANALYSIS OF SPC FOR A WIDE-ANGLE SYSTEM

VISUALIZATION OF THE SADDLE POINT

In Chapter 3, recommendations have been given for applying SPC in steps. Examples are presented in this section to demonstrate how SPC is applied to the wide-angle system. The system M3 in Figure 5.3 is selected to demonstrate how SPC is practically performed.

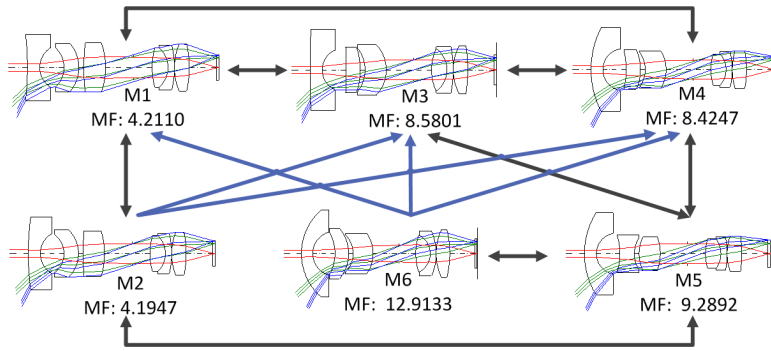


Figure 5.3: A network of minima for the wide-angle lens. All systems are connected with SPC. Global optimization using CODE V did not find additional systems. The transverse ray aberration in CODE V is used as the merit function.

5

First, by extracting (details are described in Chapter 3.3.1) the second element from M3, a minimum with one lens less is obtained. Then, a null element is inserted at the same position where the lens is extracted and an SPC scan 3.2.1 is performed on the null element. This null element is not in contact with the existing lens elements in front of or behind it. With the SPC scan, curvatures of the null-element are chosen such that saddle point systems are obtained. Figure 5.4 shows the SPC scan curve (Figure 5.4(b)) and the corresponding saddle point system (Figure 5.4(a)) which is found by the scan. In an SPC scan curve, the crossing of the curve and the horizontal axis indicates the curvature of the saddle point system (the black dot in Figure 5.4(b)). In this case, only one saddle point system is found via the scan and the curvature of the saddle point system is -0.170 mm^{-1} .

The saddle point system obtained using SPC scan is usually a saddle point with a Morse Index of 1. This means that there exists a 2D plane in the high-dimensional design space where from the obtained saddle point, along one direction the value of merit function decreases, and along another (orthogonal to the previous one) the value increases. Given this 2D plane, the saddle point can be visualized (Figure 5.4(c)).

The value of the merit function for the saddle point equals the value of the merit function at the minimum before adding the extra two variables. If the direction where the merit function decreases can be determined, then the 2D plane to visualize the saddle point can be easily chosen. Numerically, this direction can be determined by minimizing on a hyperplane, which is orthogonal to the SPC scan line along which the merit function value is constant. In Figure 5.5(a), we use a 3D case to demonstrate how such a hyperplane is found. The 3D space has the coordinates (x_1, η, ξ) . The variables are explained in Figure 5.5(b): $x_1 = c_1$, where c_1 is the curvature variable which does not belong to the null element; $\eta = c_2 + c_3$, $\xi = c_2 - c_3$, where c_2 and c_3 are the curvatures of the null element. Point $S(x_{1,m}, \eta_s, 0)$ is the saddle point system containing the null element. Point A is a point on the line close to Point S where the merit function value is constant. And the value of the merit function is the same as Point S. An example of the merit function landscape in the x_1, η plane is given in Figure 5.5(c). Along the line SA, the value of

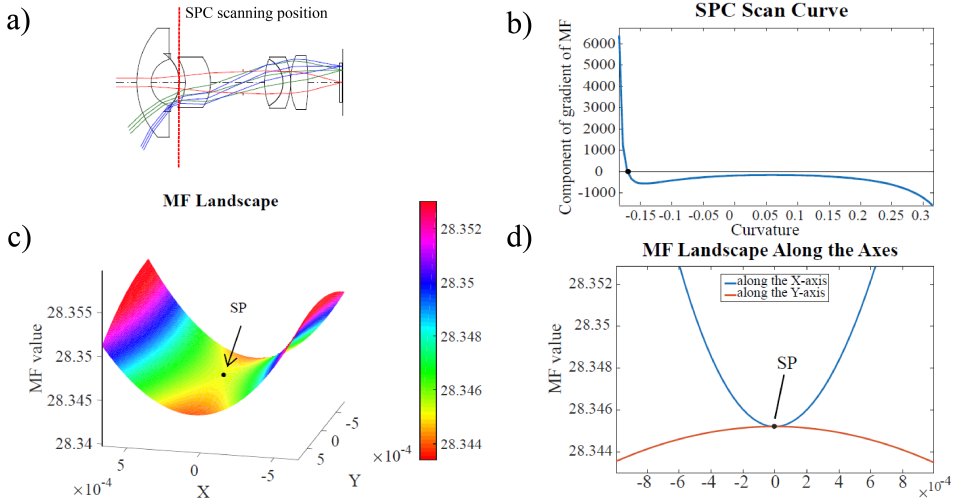


Figure 5.4: a) the system we choose for the SPC scan. b) the SPC scan curve. c) the saddle point visualized in the chosen 2D plane (Figure 5.5). d) the descending and ascending direction around the saddle point in the same 2D plane. The saddle point is marked with a black dot.

the merit function is constant, hence, to find the direction of descent in a N -dimensional space, we have to move in the $(N - 1)$ -dimensional hyperplane which is perpendicular to the line SA . It is the 2D plane $P1$ as shown in Figure 5.5(a). On this plane, B is a point where the merit function value is minimized. We will have $MF_B < MF_S$. With points A , B and S , a 2D plane $P2$ can be determined in which we can plot both the directions of ascent and descent. Given the two vectors on the 2D plane, \vec{SA} and \vec{SB} , an orthogonal base can be calculated with the Gram-Schmidt process:

$$\beta_1 = \left\| \vec{SA} - \frac{\langle \vec{SA}, \vec{SB} \rangle}{\langle \vec{SB}, \vec{SB} \rangle} \vec{SB} \right\|, \quad \hat{u}_1 = \frac{\beta_1}{\|\beta_1\|} \quad (5.1)$$

$$\beta_2 = \vec{SB}, \quad \hat{u}_2 = \frac{\beta_2}{\|\beta_2\|}, \quad (5.2)$$

\hat{u}_1 and \hat{u}_2 are the two orthogonal unit vectors of the 2D plane. The 2D plane with the visualized saddle point can be plotted as shown in Figure 5.4(c).

In Figure 5.4(d), the plots along the X and the Y direction show that the value of the merit function is increasing in one direction and decreasing in the other direction. Optimization routes following the descending directions on both sides of the saddle (on the Y-axis shown in Figure 5.4(d)) lead to two distinct local minima.

CHOICES OF STARTING POINT OF LOCAL OPTIMIZATION NEAR THE SADDLE POINT

Visualizing a saddle point in the 2D plane $P2$ is good for demonstration and helps to choose the optimization direction, however, in practice, it is computationally expensive

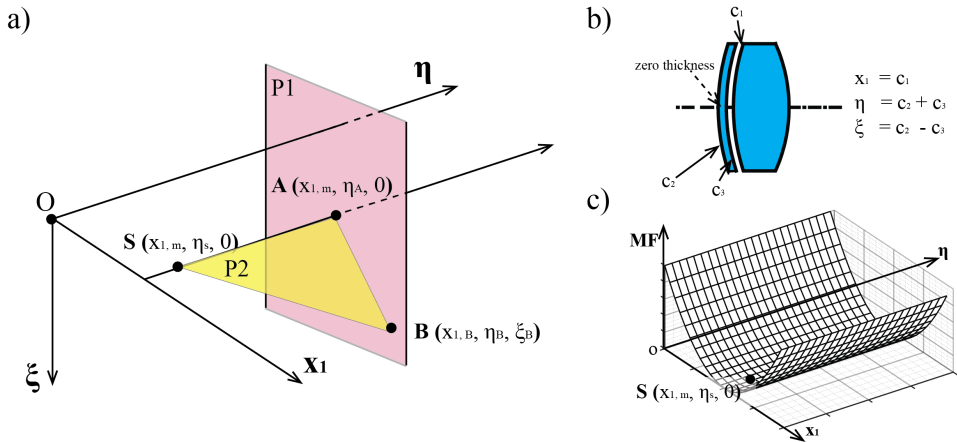


Figure 5.5: a) The construction of a 2D plane P2 to visualize the saddle point S in a 3D dimensional example. Point A is a point on the SPC scan line. P1 is a hyperplane perpendicular to line SA. Point B is where the merit function value is minimized on P1. b) Explanation about variables and coordinates. When $\xi = 0$, the x_1, η plane includes all the null element systems. c) Example of a merit function landscape in the x_1, η plane. $x_{1,m}$ indicates the original minimum values before insertion of the null element.

5

to look for the 2D plane for each saddle point. As mentioned previously, a common practice is to change the system along the SPC scan line, which means changing the curvatures of the inserted null-element with a small positive and negative value (this value is chosen empirically). The modified systems are the starting points for optimization. The optimization includes two curvature variables in addition to the variables where the minimum (without the null element) is obtained. In the experiment, choosing different values for the curvature change may lead to different minima. This is demonstrated in Figure 5.6, where different values were chosen for the null-element curvature as a starting point to find the corresponding local minimum. In order to better illustrate the phenomenon, the scan range chosen here (from -0.2 to 0.2) was larger than the curvature changes normally used in practice. From Figure 5.6, it is observed that four different minima were obtained with different values chosen as the starting curvature. The corresponding system shapes and merit function values are shown in Table 5.2.

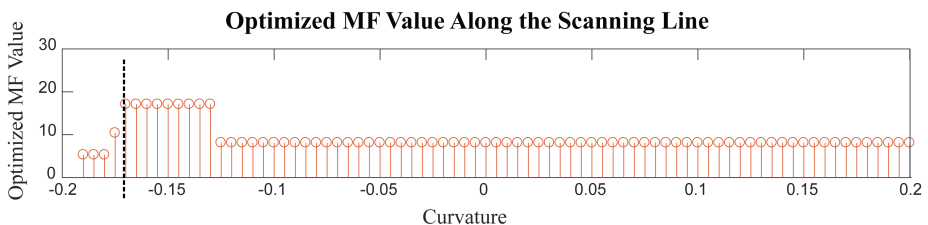
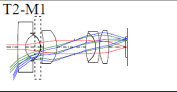
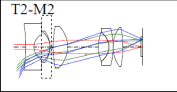
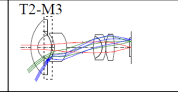
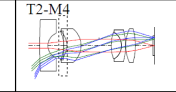
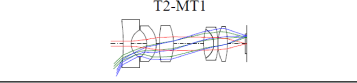



Figure 5.6: Optimized MF value along the scan line. The vertical dashed line indicates the saddle point curvature. The stepsize is 0.05.

Table 5.2: Result of the optimization along the scan line (dashed boxes mark the zero-thickness elements).

Curvature	[-0.190, -0.175]	[-0.175, -0.170]	[-0.170, -0.120]	[-0.120, 0.2]
MF of systems with zero thickness element	5.30	10.40	17.08	8.12
Plot of systems with zero thickness element				
Plot of systems after increasing thickness				

In Figure 5.6, with a curvature step size of 0.05, we see that different local minima are obtained around the curvature value of the saddle point. It indicates that near the saddle point, the local optimization is still sensitive to the choice of the starting point. To better understand the basin of attractions near the saddle point, we have also investigated the cases where the starting point is selected on a 2D plane instead of the one-dimensional scanning line. The constructed saddle point is located in this 2D plane. We have used two different approaches to determine such 2D planes. The first method is to find the plane defined by Equation 5.1 and 5.2 (Plane P_2 in Figure 5.5). On this plane, the saddle point can be visualized and in principle selecting starting points on the opposite side of the saddle point leads to two different local minima. The second method is to construct a plane directly using the two newly inserted variables (two curvatures of the null element). The first method serves as the reference since it clearly captures the ascending and descending directions around the saddle point. However, computing such a plane needs additional effort. The second method is a computation-wise simple way. By comparing the two, we would like to study if a simple approach is sufficient to obtain reliable starting points around the saddle point.

The results of the optimization on the aforementioned 2D planes are shown in Figure 5.7. These plots can be seen as the plot of the basins of attraction. The same colour indicates that all starting points from that region (basin) end in the same local minimum after local optimization. The left side of Figure 5.7 shows the basins of attraction in a relatively large region, from which the complexity of the optimization landscape can be observed. On the right side of the plot, the region around the saddle point is shown enlarged. In plots generated by both approaches, the scan line sits on the borders between two basins: the green part converging to the minimum with a merit function value of 8.12 (T2-M4 in Table 5.2) and the red part converging to the minimum with a merit function value of 5.30 (T2-M1 in Table 5.2). In both cases, we also observe that if the starting points are chosen on the scan line, the optimization may converge to the minimum with a merit function value of 17.08 (yellow color) or the minimum with a merit function value of 10.40 (navy blue color). On the two sides of the saddle point, the red and green regions represent the two basins where the two descending directions from the saddle point are intended to lead to. However, in this case, choosing starting points on the scan line may lead to minima different from the intended ones (even with higher merit function values), which are unwanted.

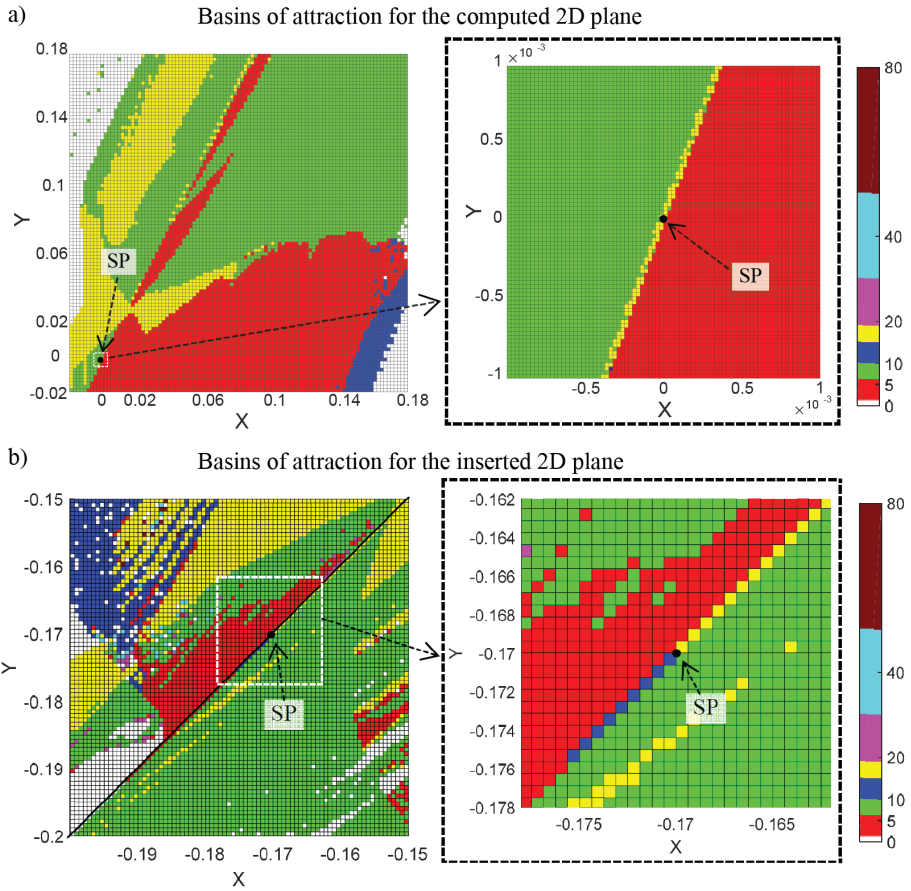


Figure 5.7: Basins of attraction around the saddle point in 2D planes. a) Computed 2D plane where x and y are the variables defined by equation 5.1 and 5.2; b) 2D plane given by the two newly inserted variables. The saddle points (SP) are marked by black dots.

Practically, it is still recommended to choose the starting points on the 2D plane defined by the two inserted variables instead of constructing the 2D plane as in Figure 5.5. As explained in the texts above Figure 5.5, the latter 2D plane requires minimization on a $N - 1$ dimensional hyperplane, which is perpendicular to the SPC scan line. This results as an additional step of constrained optimization of which the computation time increases with the number of dimensions. From the plots in Figure 5.7, it is evident that a better way to choose the starting points is to avoid the scan line, for example, by choosing a point on a line that is perpendicular to the scan line. It is also seen that the starting points should not be too far away from the saddle point. Otherwise, the two basins around the saddle point will be missed. In Figure 5.7(b), the saddle point has coordinates $(-0.1702, -0.1702)$. The closest starting point where the optimization converges to an unwanted minimum is $(-0.1713, -0.1675)$. The distance between this starting point

and the SP is 0.0029, which is 1.7% of the absolute value of the curvature at the saddle point. It indicates in this example that even with a small step, optimization can go to a local minimum which is not expected. Practically, it suggests that more than two starting points around the saddle point are needed such that the minima connected to this saddle point is obtained.

It is not clear in the presented example why on the scan line the optimization might lead to other solutions than the expected ones. It has been demonstrated in previous research done by van Turnhout [8] (Page 81)[9] that when the damping factor of the damped least square algorithm is sufficiently large, there will be less artefacts shown in the plot of the basins of attraction. A hypothesis is that on the scan line there will also be no artefact locations that cause convergence to a different local minimum. In the example presented here, there is no control over the damping factor because the local optimizer of CODE V is used, hence, it is not possible to identify the cause of the behaviour along the scan line.

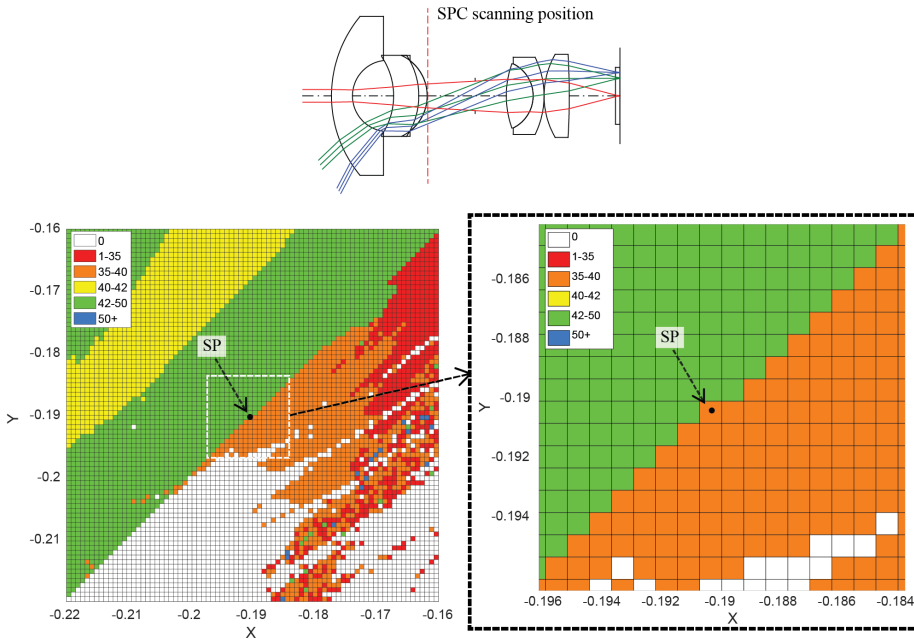


Figure 5.8: Basins of attraction on the 2D plane defined by the inserted variables at a different scanning position compared to Figure 5.7. Along the scan line (diagonal line in the plot) around the saddle point, there are no starting points for which local optimization leads to minima other than the two next to the saddle point.

It follows that by using a different position for the extraction of a lens and subsequently applying SPC, it is found in Figure 5.8 that a different behaviour occurs along the scan line. In this case, the third lens element was extracted. As shown in Figure 5.8, along the scan line there are no starting points that lead to different local minima other than the expected two. Local optimization will either converge into a local minimum in

the green region with merit function value $MF = 44.53$, or to a local minimum in the orange region with $MF = 37.82$. The white region ($MF = 0$) indicates ray failure. The closest starting point with respect to the saddle point, at which the system encounters a ray failure, is at $(-0.1893, -0.1960)$. The coordinates of the saddle point are $(-0.1903, -0.1903)$. The distance between the saddle point and this starting point is 0.0058, which is 3.0% of the absolute value of the curvature of the saddle point. This is a small distance. It again indicates in this example that optimization can become unstable if the starting points around the saddle points are not chosen properly. Both examples show that the starting point for optimization should be within a small distance (less than 1.7% of the saddle point curvature) from the saddle point. In practice, the optimization landscape will differ per use case. Our practical recommendation would be starting with 1% distance from the saddle point and choosing points perpendicular to the scan line. If ray failure is encountered, a smaller distance can be tested (a simple automated feedback loop can be made for such a task).

DISCUSSION ON LOCAL MINIMA AND THE THICKNESS OF THE INSERTED ELEMENT

Until now, the mentioned minima are the minima with zero thickness for the inserted element. To obtain practical solutions, it is necessary to increase the thickness of the zero-thickness element to some nonzero values. In the example of this wide-angle lens, the thickness of the zero-thickness element in each local minimum is restored to the value of the original lens element before extraction. The number of the zero-thickness local minima and the number of local minima after increasing the thickness, are not the same. In Table 5.2, both the zero-thickness minima and the minima after increasing the thickness are listed. When the thickness is increased, T2-M1 merges with T2-M2, and so do T2-M3 and T2-M4. The final results become T2-MT1 and T2-MT2. T2-MT1 corresponds to M1 in Figure 5.3 with a merit function value of 4.20. T2-MT2 corresponds to M3 in Figure 5.3 with a merit function value of 8.58.

Figure 5.9 shows the change of merit function value when the thickness of the inserted element is increased in small steps (0.05 mm). In the left plot, the merit function value of T2-M3 changes drastically (shown as a jump) around thickness 2 mm. The system plots just before and after the peculiar value of the thickness point are also illustrated in the same figure. Looking at the first lens element, it is seen that the system became very stressed (large curvature values) just before the thickness increased to 2 mm. Ray failure occurs when the thickness is further increased. In case of ray failure, optimization in CODE V adapts certain proprietary algorithms to modify the system in order to be able to trace the rays. Optimization is then continued to obtain a new solution. In this example, the system T2-M3 merges with T2-M4 when the thickness is increased to more than 2 mm. On the right side of Figure 5.9, the merit function value of T2-M4 changes continuously with increasing thickness. The same phenomenon occurs with the other two systems: system T2-M2 merges with T2-M1 right after increasing the thickness. As previously illustrated in Figure 5.7, T2-M1 and T2-M4 correspond to the two expected local minima. It is worth mentioning that the two expected local minima are the two stable ones as shown in this example.

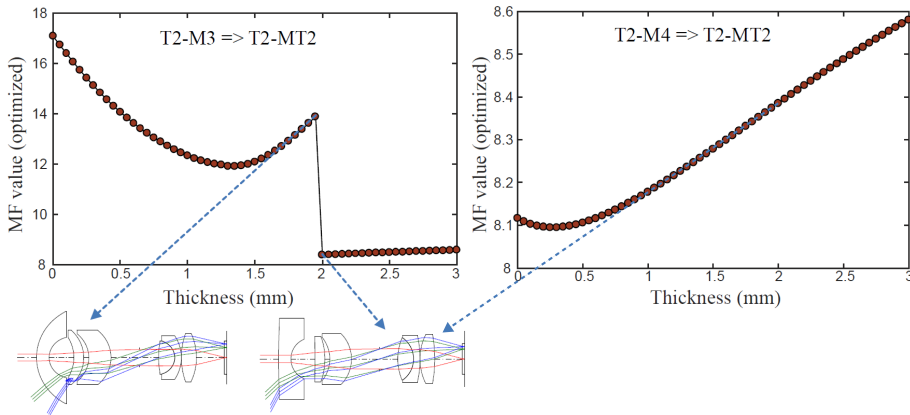


Figure 5.9: Change of the optimized value of the merit function while increasing the thickness of the zero-thickness element. T2-M3 and T2-M4 correspond to systems when the thickness is zero. T2-MT2 corresponds to the system when the thickness equals 3 mm. The discontinuity of the left plot indicates the instability due to ray failure around a minimum in the design space: The jump appears because when the system on the left (T2-M3) ceases to exist because of ray failure, optimization in CODE V leads automatically to the system on the right. Systems like the left one are usually poor local minima.

SECTION SUMMARY

With a step by step analysis of applying SPC in this example, it is seen that, for some cases, choosing starting points on the SPC scan line may lead to unexpected local minima (zero-thickness). This may be caused by the specific landscape around the saddle point or numerical artefacts from the optimization algorithm. Even though these problems do not always occur, examples suggest that a starting point should be chosen outside the scan line to prevent trapping in unwanted local minima. In the plotted basin of attraction, there are features that make the optimization around the saddle point unstable: optimization can reach to an unexpected local minimum or ray failure region. Therefore, choosing more than two starting points with proper distances to the saddle point is needed. In our experiment, we use four starting points: two are on the scan line and two are perpendicular to the scan line. The used distances to the saddle point curvatures are 2.5% of the absolute value of the saddle point curvature (having a zero curvature is rarely the case in practice). In addition to the local optimization, experimental results also reveal that, in some cases, zero-thickness minima obtained in the intermediate steps of SPC are not stable and disappear after increasing the thickness of the inserted element. For the example given here, the final result (two finite-thickness minima) is less dependent on the precise choice of the starting points chosen after obtaining the saddle point.

5.3. COMBINING DESIGNERS' KNOWLEDGE AND SPC TO IMPROVE DESIGN EFFICIENCY

In lens design practice, it is common that multiple design techniques are applied to effectively obtain an optimal solution. The technique can be using a cemented doublet to reduce chromatic aberration, controlling the exit marginal ray angle to keep EFL constant, using a certain optimization method to find better local minima, etc. The SPC method provides a systematic approach to rapidly obtain local minima. However, it does not tell you where the SPC should be applied in an optical system. A straightforward way is to apply many SPCs in a system to generate all the possible local minima. However, this is not always preferred because at a certain phase of the design process, the designer does not really expect a drastic change of the system. In this section, we explore the option of combining designer knowledge and SPC. A UV dry microscope objective is used to show what can be achieved with this approach.

5.3.1. ULTRAVIOLET (UV) MICROSCOPE DRY OBJECTIVE

As a method to search for alternative solutions in lens design, SPC can be automated. In such a scenario, SPC could be used as a global search tool, where, after setting up all the specifications, the program would generate potential solutions for further selection. However, one of the challenges in using SPC, is to choose the proper position for inserting the null element. Conventional design approaches encounter a similar dilemma, namely: when a design process gets stagnated, how to modify the system (adding, extracting or changing) to improve the system? Optical designers usually use their unique insight and experience based on monitoring the aberration, ray properties etc. to make decisions on how to change the system. These approaches and methods used can be very different from person to person, therefore, setting up an algorithm to try out all promising looking options is difficult. If the optical system is relatively simple, it is possible to combine the SPC with approaches based on the designer's knowledge. To use SPC, the modification of the system should be adding or replacing lens elements in the existing system. In the following sections, we will provide two examples of how to combine SPC with conventional approaches.

An ultra-violet (UV) microscope objective is chosen as a starting configuration. The microscope objective is taken from an existing patent[10] where its specification is reported. The variation designed for wavelength 325 nm is selected from the patent and reproduced in CODE V. The system consists of seven lens elements. The material used in CODE V is fused silica (LithosilQ from SCHOTT) with a refractive index of 1.4816 at 325 nm. The specification and performance of the system are listed in Table 5.3. With a barrel lens having a focal length of 250 mm, the magnification of the composite microscope becomes 100. The seven lens elements of the system can be divided into three function groups according to a categorization method proposed by Zhang and Gross [11]: the first lens on the left forms the rear group; the second and third lenses form the middle group, and the rest of the lens elements form the front group. The objective is shown in Figure 5.10. The front group is defined as the one facing to the object in the application, however, during the design it is convenient to work reversely, hence the front group is on the right side in Figure 5.10. According to the same author, existing microscopes can be

categorized into three types depending on the functions performed by different groups of lenses. The distribution of the spherical aberration in different groups shown in Figure 5.10 indicates that the middle and rear group compensate each other to correct the spherical aberration. This correction behaviour belongs to one of the types described in [11].

Table 5.3: System specification and performance of the microscope objective

Effective Focal Length (mm)	2.5
Field (real image height, mm)	0.1
Numerical Aperture (NA)	0.90
Operating Wavelength (nm)	325
Working Distance (WD, mm)	0.6
Overall Length (mm)	25
RMS Wave Front Error ($m\lambda$)	49
Strehl Ratio at 0 mm	0.976
Strehl Ratio at 0.1mm	0.849

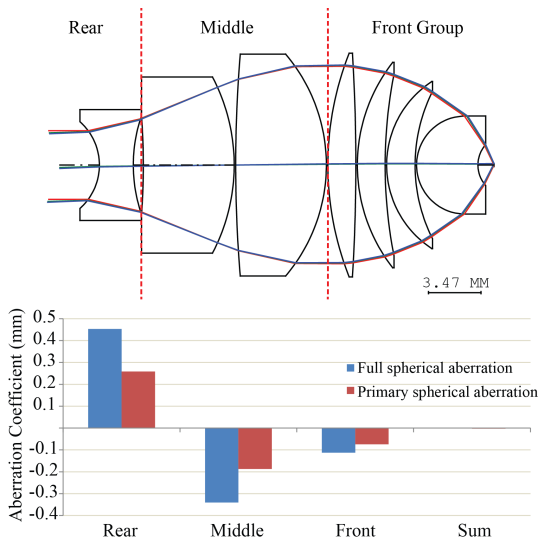


Figure 5.10: System plot of the UV microscope objective [10]. The objective can be divided into three groups of lenses. The rear group and middle group compensate each other to correct the spherical aberration [11].

The current system already has a good performance and a compact size, its specifications are listed in Table 5.3. However, if some of the specifications are made more demanding, a certain structural modification of the original system is necessary to keep a satisfactory performance. It is then interesting to investigate how SPC could help to

find these modifications. The most simple case to investigate is to add an additional lens element.

In contrast to a simple design problem, a complex system often needs several constraints to fulfill the design requirements. Before showing the design examples, the handling of constraints in SPC is first discussed in the following section.

5.3.2. CONSTRAINTS IN OPTICAL DESIGN SOFTWARE

The wide-angle lens shown in the previous section is a moderately complex system where SPC can be used in a limited number of positions. The system only has one requirement on the specifications that needs to be controlled, namely the EFL. It was controlled by setting the value of the exit angle α of the marginal ray on the last optical surface. Given the EPD and the EFL, the exit marginal ray angle α on the last optical surface can be set in CODE V as:

$$\tan \alpha = \frac{EPD}{2 \times EFL}, \quad (5.3)$$

Certain design requirements, such as EFL, can be used to force selected parameters to be dependent variables which are solved explicitly to keep the chosen requirements satisfied. In automatic design (optimization), this also makes it possible to reduce the number of independent variables, and thereby save computation effort.

However, in lens design practise, most of the design requirements cannot be explicitly controlled in the aforementioned way. Constraint function such as equality constraint, inequality constraint, penalty function, etc. are then necessary to be implemented in order to fulfill the various requirements. In lens design software, there are usually different constraint functions for the user to choose so that it is possible to control the system design in different ways. In CODE V for example, there are four different types of constraint functions can be used as shown in Table 5.4.

As seen in Table 5.4, both equality and inequality constraints are handled with the Lagrangian multiplier method. We will show in the following part that these two options treat the constraint by adding new variables to the optimization problem. We will also explain that, instead of the original merit function, the Lagrange function should be used for the SPC scan. On the other side, the penalty and weighted function methods treat the constraint by directly modifying the merit function. One can keep using the modified merit function for SPC scans.

One of the advantages of using a Lagrange multiplier rule is that it treats the constraint separately from the merit function, allowing the merit function to be based solely upon image quality-related criteria. For the case of an equality constraint, the optimization of the merit function can be written as follows:

$$\begin{aligned} &\text{minimize } MF(\mathbf{x}), \\ &\text{subject to } g(\mathbf{x}) = 0, \end{aligned} \quad (5.4)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_N)$ is a vector describing a point in the N -dimensional variable space. Equation $g(\mathbf{x}) = 0$ is an equality constraint that the variables have to satisfy. Given that both MF and g have continuous first partial derivatives, a new variable λ called a

Table 5.4: Different types of constraints in CODE V[12]

Constraint Mode	Explanation
Equality constraint (=)	Always active. Handled with Lagrangian Multiplier Method. Constrains exactly to target value of the constraint.
Inequality constraint (>, <)	Only active when the optimization violates the targeted boundary. Handled with Lagrangian Multiplier Method.
Weighted function (WTC)	Only used for targeted constraint. Includes a weighted function in the merit function.
Penalty function (PTC)	Only used for targeted constraint. Includes a penalty function in the merit function. The penalty function is defined in CODE V with more tuning options than the weight function. It is essentially the same way of constraint implementation as the weight function.

Lagrange multiplier is introduced, and the Lagrange function is defined by

$$\mathcal{L}(\mathbf{x}, \lambda) = MF(\mathbf{x}) - \lambda \cdot g(\mathbf{x}). \quad (5.5)$$

A critical point of Equation 5.5 has zero gradient and satisfies the following equations

$$\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \lambda) = \nabla_{\mathbf{x}} MF(\mathbf{x}) - \lambda \nabla_{\mathbf{x}} g(\mathbf{x}) = 0, \quad (5.6a)$$

$$\nabla_{\lambda} \mathcal{L}(\mathbf{x}, \lambda) = g(\mathbf{x}) = 0. \quad (5.6b)$$

From Equation 5.6b, it is seen that once a critical point of $\mathcal{L}(\mathbf{x}, \lambda)$ is found, the equality constraint is automatically satisfied. Equation 5.6a indicates that the gradient of the merit function can be expressed as a linear combination of the gradient of the constraints at the critical point. The Lagrange multiplier λ acts as the coefficients. The Lagrange multiplier rule states that at the critical point, the value of the constrained merit function is either a maximum or minimum. In this way, the constrained optimization problem in Equation 5.4 is converted to an unconstrained optimization problem. The Lagrange Multiplier method converts the original N -dimensional problem into a $N + 1$ -dimensional problem by adding one more variable λ . At the critical points of $\mathcal{L}(\mathbf{x}, \lambda)$, $\mathcal{L}(\mathbf{x}, \lambda)$ equals $MF(\mathbf{x})$. Therefore, the smallest value of $\mathcal{L}(\mathbf{x}, \lambda)$ at the critical point gives the minimum of the constrained merit function.

Before performing the SPC scan and adding the null element, the system needs to be firstly optimized to a local minimum. In CODE V, when equality or inequality constraints are applied, the Lagrange Multiplier method is used. The optimizer converges to a critical point satisfying Equations 5.6a and 5.6b.

We would like now to investigate the possibility of $\nabla_{\mathbf{x}}MF(\mathbf{x}) = 0$ at a critical point of $\mathcal{L}(\mathbf{x}, \lambda)$. According to Equation 5.6a, satisfying $\nabla_{\mathbf{x}}MF(\mathbf{x}) = 0$ requires either of the two following conditions: $\lambda = 0$ or $\nabla_{\mathbf{x}}g(\mathbf{x}) = 0$. Given Equation 5.6b is valid, either condition $\lambda = 0$ or $\nabla_{\mathbf{x}}g(\mathbf{x}) = 0$ indicates that the local minimum of the original unconstrained merit function (MF) coincides with the local minimum of the Lagrange function. Therefore, the constraint is already fulfilled at the local minimum of the original merit function. It indicates Lagrange function is in practice not needed for this situation and this situation is not in general valid when Lagrange function is applied. As a results, $\nabla_{\mathbf{x}}MF(\mathbf{x}) \neq 0$ is the typical situation for a critical point of $\mathcal{L}(\mathbf{x}, \lambda)$. It indicates that the unconstrained merit function is not at its local minimum.

When performing the SPC scan, two additional variables (c_1 and c_2 , curvatures of the null element) are added to the existing system. In ray tracing software like CODE V, the user usually only sees the value of the merit function. In a case that Lagrange Multiplier method is used, if the SPC scan of the saddle point uses the unconstrained merit function, the construction of the saddle point will not start from a local minimum. The precondition for performing the SPC is thus not satisfied and saddle points will not be obtained. As a recommendation, when Lagrange Multiplier method is used, the SPC scan should be performed using the Lagrange function.

The inequality constraint mentioned in Table 5.4 should take the same situation into consideration: The constraint becomes active when the optimization violates the boundary, and the computation for SPC scan should use the Lagrange function; The constraint remains inactive if the optimization is operated within the defined boundary, and the original merit function is used for SPC scan.

When constraints are implemented using a penalty function or a weighted function method, the preconditions for performing SPC are satisfied. Both approaches add new terms to the merit function with explicit expressions. The altered merit function is used to optimize the system to a local minimum and the same function is used to perform SPC.

The modified merit function with a weight function term can be written as the following

$$MF_W(\mathbf{x}) = MF(\mathbf{x}) + \sum_{i=1}^m W_i(\mathbf{x}), \quad (5.7a)$$

$$W(\mathbf{x}) = (wt \cdot \Delta C)^2. \quad (5.7b)$$

where the summation indicates that multiple weight functions can be added, wt is the weighting factor set in CODE V, and ΔC is the departure of the constraints from its target. Using the penalty function replaces the weight function term in Equation 5.7a with a penalty function described as the following in CODE V

$$P = wt * \left(\frac{\Delta C}{half\ width} \right)^{\frac{exponent}{2}}, \quad (5.8)$$

where *half width* and *exponent* are two additional variables for tuning the penalty function. If the *half width* is set to 1 and the *exponent* is set to 4, the penalty function in Equation 5.8 is the same as the weight function in Equation 5.7b. The two approaches for implementing the constraints are essentially identical. They both provide

a less rigid control compared to the Lagrange Multiplier method. The penalty function mode is ideal for the "soft constraints", which are constraints that have an acceptable range of values [12].

As described in this section, applying constraints using Lagrange Multiplier method needs additional engineering treatment to be compatible with SPC scans. We prefer to use weighted or penalty function for its simplicity of keeping track of the same merit function. In the rest of this chapter, where different constrained systems are investigated, the constraints are implemented with the weighted function.

5.3.3. APPLYING SPC TO INDUCE STRUCTURAL CHANGES IN LENS

When trapped in a local minimum, it is useful to introduce structural change in order to obtain different local minima with significant changes in the values of the merit function. The structural change refers to changes such as using a lens with a positive power instead of a negative one, adding a new element into the system, etc. In this section, we focus on two cases where structural changes are needed for the systems to meet a higher requirement. In the first case, the NA of the system is further increased from 0.90 to 0.95. In the second case, the working distance of the system is increased from 0.6 mm to 2.5 mm. We show how SPC can be effectively applied to obtain new local minima.

INCREASING NA OF THE OBJECTIVE

For a high-resolution microscope objective, high NA is always required. The first example is to increase the NA of the microscope objective from 0.90 to 0.95. The rays can still be traced after the modification, i.e. in the case considered no ray failure occurs after the increase of the NA. The system is then optimized with a wavefront error function. An EFL constraint and a working distance constraint are applied with the equality constraint mode. Figure 5.11 shows that the change of the system shape² is mainly in the thickness of the elements. The bending³ of each element remains almost unchanged. From Table 5.5, it is seen that the performance of the system deteriorates after increasing the NA. A major cause is that the spherical aberration (especially the high-order spherical aberration) increases with increasing NA. In traditional approaches, structural changes to the system would be made by which the number of degrees of freedom increases to be able to reduce the aberration. These kinds of structural change could, for example, be replacing a spherical surface by an asphere, adding extra lens elements, or splitting existing lens elements. In the following examples, the introduced changes are adding or splitting lens elements.

With the separation of the system into different function groups, it is possible to understand the function of different parts of the system which provides insight when structural change is needed. It follows from Figure 5.10 that the rear and middle group compensate each other to correct the spherical aberration. A reasonable strategy is then to introduce structural changes in these two groups, namely by adding or splitting elements in the rear and the middle group of the system. Based on this insight, four dif-

²A system shape is determined by the bending of each lens element, the thicknesses of the lens elements and the distance between them.

³The definition of the bending factor is given in Equation 2.10.

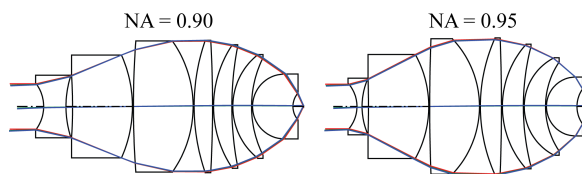


Figure 5.11: Change of the system shape after increasing the NA from 0.90 to 0.95.

Table 5.5: Change of the system performance with increasing NA

Numerical Aperture	0.90	0.95
RMS Wavefront Error (m λ)	49	111
Strehl Ratio at 0 mm	0.976	0.833
Strehl Ratio at 0.1 mm	0.849	0.386

5

ferent positions are chosen in the rear and the middle group to add or split elements in the usual way. The resulted system shapes are shown in Figure 5.12. It is worth pointing out that with the traditional approach of adding or splitting lens, each operation only yields one solution. The performance of the resulting systems are listed in Table 5.6. The data from the table reveal that out of four, one (the number 4 system) has satisfactory performance. For a satisfactory performance, the criteria used here are that the Strehl ratio is above 0.8 for all fields and the wavefront RMS is below $\lambda/14$ [10], for which value the system is considered to be diffraction-limited. In Figure 5.12, the inserted or split element is highlighted by shading. For operation ①, the insertion of the lens results in an additional shell lens in the front group. It gives four new degrees of freedoms (two curvatures, one lens thickness, and an air space thickness) for optimization. However, the performance of the optimized system is not satisfactory. Operation ② splits one of the lens elements in the middle group and thereby distributing its original power to two lens elements. The same happens in the case of operation ③. Both operations ② and ③ produce only slightly better results than operation ①, however, neither of them is good enough. Splitting the lens element in the rear group in operation ④ produces a different system shape compared to the other three and the original system. A reverse-bended element is created (second element from the left), which forms a narrow air gap with the element to the right of it. This kind of air gap is effective for introducing high-order spherical aberrations, which balances the high-order spherical aberrations introduced by other parts of the system [13].

Table 5.6: System performance for the systems shown in Figure 5.12

No.	1	2	3	4
RMS Wavefront Error (m λ)	119	117	99	41
Strehl Ratio at 0 mm	0.736	0.769	0.832	0.985
Strehl Ratio at 0.1 mm	0.428	0.453	0.548	0.882

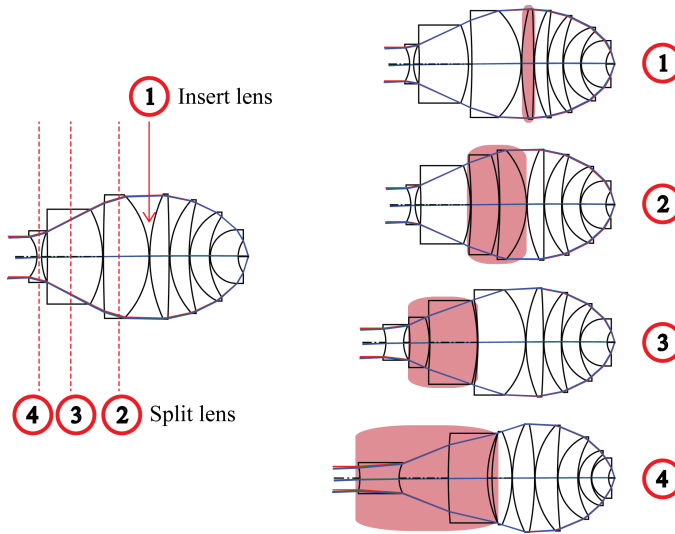


Figure 5.12: Introducing structural change in the traditional way for system performance improvement. Four operations have been executed on the NA 0.95 system resulting in four different results. The inserted or split element is highlighted by shading.

Based on the insight on where the change should be introduced in the system, SPC scans were performed at four positions in the system. Figure 5.13 presents the results from these SPC scans. From four scans, a total number of eighteen solutions were obtained. Besides four solutions which are identical to the ones obtained with the traditional method, fourteen extra solutions were obtained. Their performance is listed in Table 5.7. Out of the eighteen solutions, nine reach satisfactory results. Solution ④A and ④D even show better performance than the best one obtained by the traditional method (No.4 in Table 5.6). Among all the solutions, ①A and ②B have very close system shapes. One is the result from the left bending of the glass lens, and the other is the result of the right bending of the air lens. As explained in Chapter 3, there are two different ways to apply SPC, which lead to the same solution in this case.

Compared to the traditional method, more than 4 times as many systems are obtained for a comparable amount of computational load with SPC than with the traditional method. Besides the good solutions which are similar to the ones obtained with the traditional method, SPC also found eight additional solution with satisfactory performance.

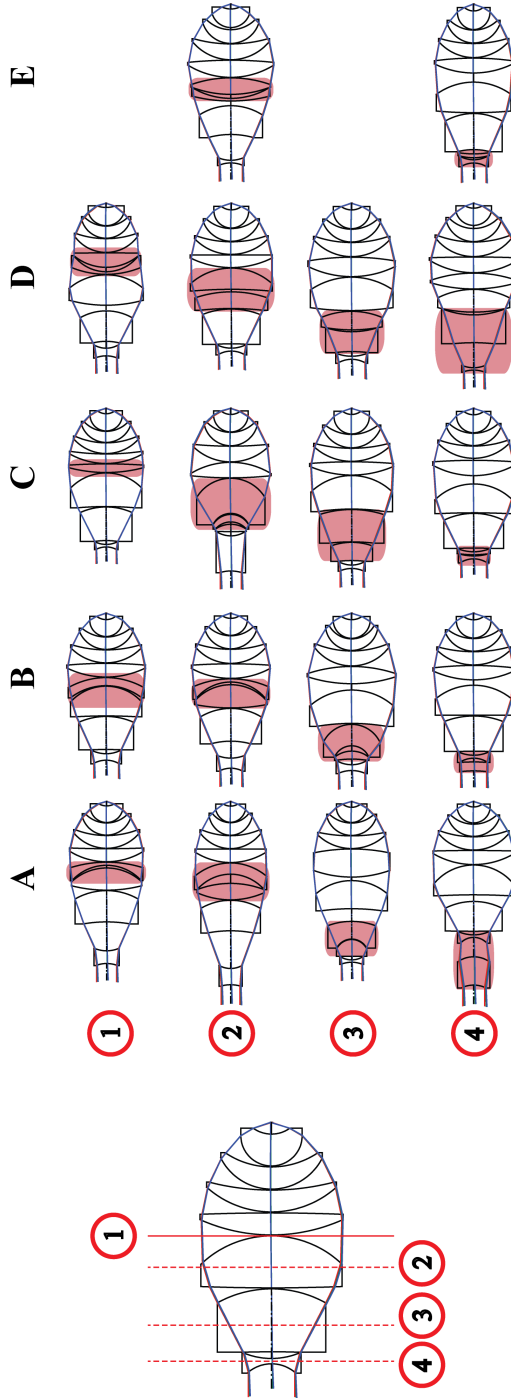


Figure 5.13: Eighteen solutions from four SPC scans. The dashed line indicates constructing an air lens inside a lens element (equivalent to splitting lens). The solid line indicates constructing a glass element in the air space (equivalent to add lens). In comparison with Figure 5.12, solution ①C, ②D, ③C, and ④D are the ones obtained with conventional approach.

Table 5.7: System performance for the systems shown in Figure 5.13

	1	2	3	4	
RMS WFR	65	59	60	26	
SR(0)	0.862	0.942	0.958	0.998	A
SR(0.1)	0.829	0.807	0.808	0.945	
RMS WFR	109	67	52	135	
SR(0)	0.876	0.941	0.960	0.604	B
SR(0.1)	0.378	0.756	0.847	0.319	
RMS WFR	98	45	104	124	
SR(0)	0.826	0.989	0.779	0.661	C
SR(0.1)	0.505	0.874	0.540	0.371	
RMS WFR	48	102	89	35	
SR(0)	0.950	0.828	0.820	0.993	D
SR(0.1)	0.859	0.444	0.625	0.894	
RMS WFR		43		135	
SR(0)		0.982		0.789	E
SR(0.1)		0.853		0.251	

RMS WFR = RMS Wavefront Error ($m\lambda$); SR(0) = Strehl Ratio at 0 mm; SR(0.1) = Strehl Ratio at 0.1 mm.

INCREASING THE WORKING DISTANCE OF THE OBJECTIVE

A second structural change is introduced in the case of an increase of the working distance of the original system (the system on the left in Figure 5.11), from 0.6 mm to 2.5 mm. This requires that the first lens element on the right is aplanatic-concentric instead of plano-concentric, causing a necessary decrease of power of the first element. In order to maintain the performance, this decrease of optical power need to be compensated. This is done by increasing the optical power of the lens element after the first one and those in the middle group. With the limited number of degrees of freedom, the performance of the system strongly deteriorates. Figure 5.14 shows how the system changed after increasing the working distance from 0.6 mm to 2.5 mm. The OverAll Length (OAL) of the system increases from 25 mm to 76 mm. The diameter of the maximum element increases from 6.6 mm to 17 mm. A longer air gap emerges after the first element on the left caused by the increase of the size (diameter) of the lens elements. The third lens element from the left is compressed into a thinner element. The bending of each element still remains close to that of the original system. Table 5.8 gives the change of the performance with the increase of the working distance.

5

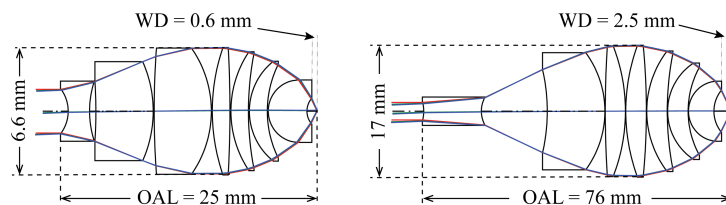


Figure 5.14: Change of the system shape after the increasing of the working distance from 0.6 mm to 2.5 mm.

Table 5.8: Change of the system performance with increasing working distance

Working distance (WD, mm)	0.6	2.5
RMS wavefront error (mλ)	49	135
Strehl ratio at 0 mm	0.976	0.254
Strehl ratio at 0.1 mm	0.849	0.019

To compensate for the decreased optical power of the first element, a common practice to introduce changes to the system for system improvement is to introduce a structural change in the middle or front group. By splitting or adding lens elements, newly added degrees of freedom will help to correct the aberrations. Two insertion positions and one splitting position were chosen for adding new degrees of freedom to the system. Figure 5.15 presents the results of the three operations: two insertions led to the same solution; the splitting shows a different system with good performance. The data of the system are listed in Table 5.9.

Table 5.9: System performance for the systems shown in Figure 5.15

No.	1 & 2	3
RMS Wavefront Error ($m\lambda$)	100	41
Strehl Ratio at 0 mm	0.833	0.975
Strehl Ratio at 0.1 mm	0.507	0.902

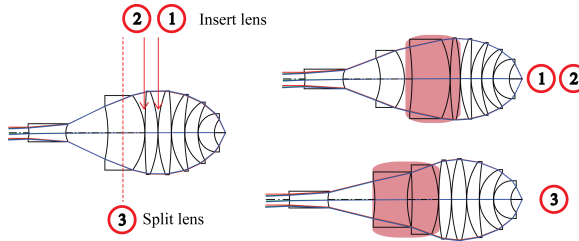


Figure 5.15: Introducing structural change in the traditional way to improve the performance of the system. Insertions at position ① and ② result in the same solution.

SPC scans are applied at the same positions where a lens is added or split. The result is shown in Figure 5.16. From three operations of SPC, ten systems were obtained. The performance of these 10 systems is listed in Table 5.10. Out of the ten systems, six have low wavefront RMS value and Strehl ratio above 0.8 for both fields. By carefully examining the system plots in Figure 5.16, it is seen that some systems have similar system shapes. For example, systems ①B, ②B, and ③B have very close bending for each lens element. The differences among the three are in the lens thickness and air space variations. System ②A and system ③A also are very similar. When compared with the results from the conventional approach, system ③D in Figure 5.16 and system ③ in Figure 5.15 have similar system shape (especially with respect to the front group on the right) and essentially the same RMS wavefront error (39 and 41 $m\lambda$) for practical purpose.

In this example where the working distance is increased, with the conventional approach, two systems were obtained with three operations. One out of three (33%) has good performance. Using SPC, ten systems were obtained with the same amount of operations as needed in the traditional approach. Six out of ten (60%) give satisfactory performance.

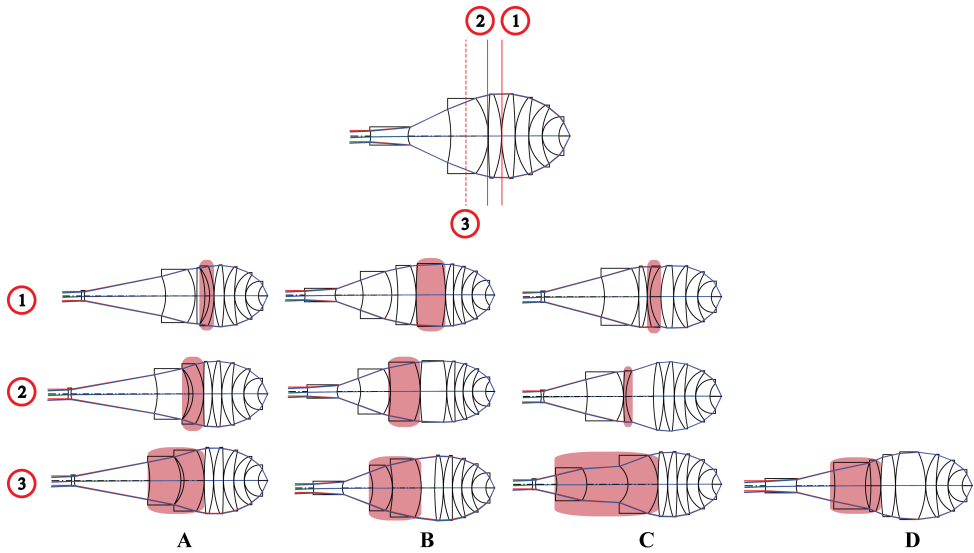


Figure 5.16: Systems obtained using three SPC scans performed at the marked locations. Solid lines (① and ②) indicate insertion of glass null-element. Dashed line (③) indicates insertion of air null-element. In total, ten systems were obtained. The modified parts are highlighted.

Table 5.10: System performance for the systems shown in Figure 5.16

	1	2	3	
RMS WFR	55	49	60	
SR(0)	0.952	0.976	0.946	A
SR(0.1)	0.835	0.863	0.807	
RMS WFR	133	101	179	
SR(0)	0.645	0.779	0.410	B
SR(0.1)	0.309	0.511	0.131	
RMS WFR	95	43	41	
SR(0)	0.732	0.971	0.983	C
SR(0.1)	0.642	0.901	0.896	
RMS WFR			39	
SR(0)			0.978	D
SR(0.1)			0.910	

RMS WFR = RMS Wavefront Error (mλ); SR(0) = Strehl Ratio at 0 mm; SR(0.1) = Strehl Ratio at 0.1 mm.

5.3.4. FURTHER CONSIDERATION WHEN APPLYING SPC TO COMPLEX SYSTEM

When working with complex systems, there are some additional factors which have to be considered in order to properly use SPC. These additional factors usually appear because of the increasing complexity of multi-variable optimization and the design approaches in complex lens design. Increased complexity here refers to the increased number of variables used.

QUALITY OF THE LOCAL MINIMUM AND THE SPC SCAN CURVE

One of the known issues occurring with complex systems is that often when the number of variables increases, the optimization basin around the local minimum becomes a vast shallow region where absolute convergence can not be achieved and no proper local minimum can be found, even when very large number of optimization cycles are performed [14]. It is observed that, when the local minima in N dimensional space were not computed accurately, there are problems finding the saddle points in the $N + 2$ dimensional space with the SPC method. An example is given for the case where the working distance of the microscope objective is increased. Before proceeding with the SPC scan, the system has to be well-optimized to a local minimum. The merit function is defined by the default transverse ray aberration in CODE V. Two weight constraints were used: an EFL constraint of 2.5 mm and an edge thickness constraint of 2.5 mm on the first surface from the right, both with a weight of 0.01. Fourteen curvatures of the surfaces were used as variables. A threshold for the improvement of the merit function was defined as a criteria to stop the optimization. The threshold was set as 10^{-15} , which means that when the decrease of the merit function value is less than 10^{-15} , the optimization will stop.

Without human interference, the optimization for this system ran with 5500 cycles (or optimization iterations), and it took 20 minutes (CPU: Dual-core 3.20GHz) to finish. This implies that there must be a very shallow region around the local minimum where the value of the gradient is close to zero. In the practice of lens design, the designer probably will manually stop the optimization before it runs for 5500 cycles. However, Figure 5.17 demonstrates that the decision of when to stop the optimization affects the number of saddle points that are found. The SPC scan was performed on the fourth surface from the left, as shown in Figure 5.17. In the same figure, different scan curves are plotted with different numbers of optimization cycles applied. Up to 2500 cycles, there is no crossing between the SPC scan curve and the "0" horizontal axis. From 5500 cycles up to 30,000 cycles, the scan curves are closely positioned giving very similar values of the saddle points. From the zoomed plot in Figure 5.17, the five curves intersecting the horizontal axis are from 0.00466 to 0.00472. The example shows that, to properly perform SPC scan, it is necessary to compute the local minimum very accurately, even if the local optimization may take many more steps than usual. In this case, with insufficient optimization, two saddle points may be missed. Consequently, the three systems in row ② (the second row) of Figure 5.16 would not be obtained.

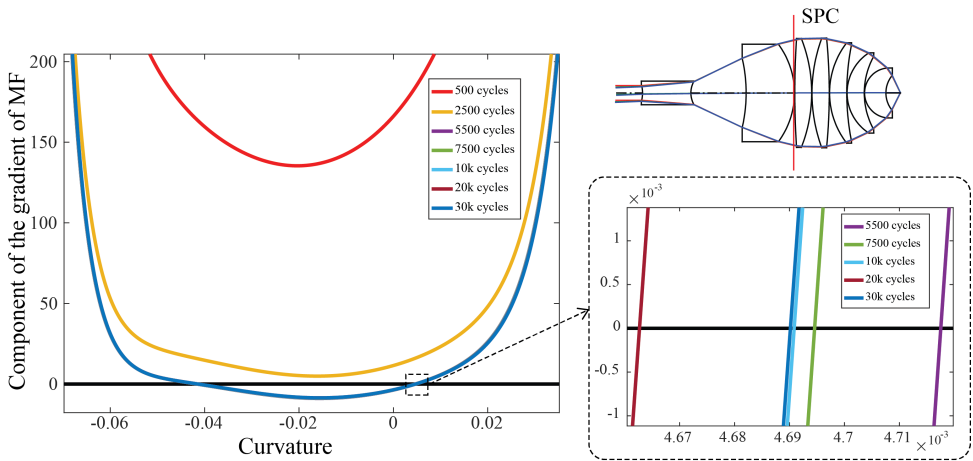


Figure 5.17: SPC scan curves at the same position with different numbers of optimization cycles

5

CHOICE OF WEIGHTS AND CONSTRAINTS

Other aspects which also require attention concerning SPC scans are the choices of constraints and their weights used in the merit function. As explained in Section 5.3.2, using weighted function (or penalty function, essentially identical to weighted function) for constraints in the merit function is compatible with SPC scan. Adding weighted constraints will alter the topography of the optimization landscape and this change is affected by the number of constraints and their weights. In this section, we explain that the choice of constraints and their weights, change the shape of the SPC scan curve significantly. In some cases, the number of obtained saddle points varies for different combinations of constraints and weights.

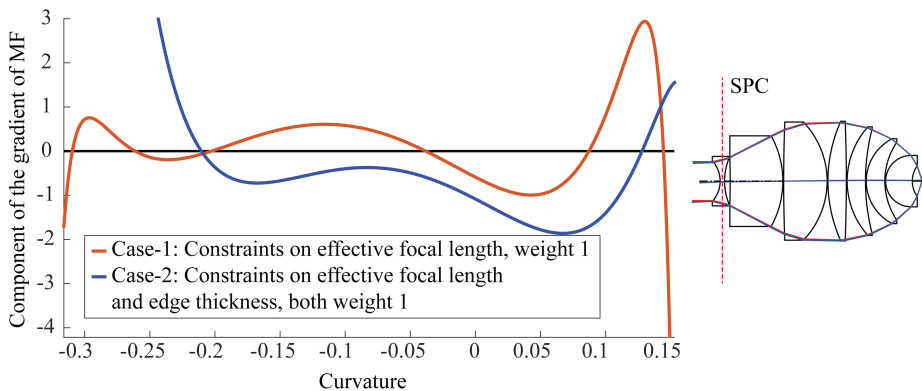


Figure 5.18: Change in the saddle point scan curve when different constraints are applied.

In Figure 5.18, the scan curves of two choices of constraints are shown. Case - 1 (the red-coloured line in Figure 5.18) was only with an EFL constraint, and Case-2 was using both an EFL constraint and an edge thickness constraint. As seen in the figure, the scan curve of Case-1 has six zero-crossings with the horizontal axis, whereas that of

Case-2 has only two zero-crossings. From the comparison in Figure 5.18, we also see that adding simple constraints (Case - 1) during SPC generates more local minima than applying several constraints simultaneously (Case - 2).

In practice, obtaining more local minima via SPC is preferred since it generates more solution candidates for evaluation, hence increasing the chance of finding a satisfactory solution. For the given example, Case-1 did not include the edge thickness requirement in the merit function. If this requirement is not met, it is necessary to add the constraint on the edge thickness and optimize from the minima acquired in the next step. In Case-2 the two necessary constraints on the EFL and the edge thickness are taken account of from the start, hence, the acquired minima are the final solutions. Design procedure and the obtained local minima are shown for both cases in Figure 5.19. For Case-1, seven local minima are acquired from six saddle points. For Case-2, three minima are obtained from two saddle points. Judging by the system shapes, the three minima from Case-2 correspond to the local minima in the set obtained in Case-1. The pairs of corresponding local minima are highlighted with different colours in Figure 5.19. It is seen that MA-4 and MB-2, which belong to the same system shape, have very similar merit function values. However, the two pairs MA-1, MB-1 and MA-6, MB-3 have bigger differences in their merit function values, with Case-1 showing smaller merit function values. Although the system shapes are similar, due to the complex topography of the design landscape in the case of many variables, the merit function values can be quite different. The same observation was described in Reference [15].

The choice of the weights associated with the constraints can also alter the SPC scan curve. We observe that as far as the weights are not extreme (neither too loose nor unnecessarily tight), the SPC scan curve remains relatively stable. An example is given in Figure 5.20, where different scan positions are chosen using the same system from the previous example. Only the EFL constraint is applied here with different weight number from 0.001 to 10. In Figure 5.20, it is shown that when the weight is 0.001 or 10, the SPC scans produce curves that differ from those in the other three cases. Weight 0.001 gives four saddle points, weight 10 gives two saddle points, and the other three cases give three saddle points with relatively similar SPC scan curves. The target value of the EFL is 2.50 mm. When a weight of 0.001 is used, the optimized system has an EFL of 2.65 mm (6% off the target). For a weight of 0.01, the EFL is stabilized around 2.53 mm. When the weight is larger than 0.1, the optimized system has an intended EFL of 2.50 mm.

According to the derivation of the special case of SPC [16], when the space between the constructed surface and another optical surface is zero (or close to zero), a saddle point can be found with the curvatures of the constructed surface identical (or similar) to that optical surface. In this example, when the weight is 0.01, 0.1 and 1, the most left and most right saddle points represent the special case saddle points. However, when the weight is too big or too small, the design landscape is altered in a way that the special case of saddle point cannot be obtained via the scan (weight 10 case missed the left one). One of the hypotheses is that there are some numerical inaccuracy when computing the local minima and saddle points with these extreme weights. More investigations are needed to understand these situations. Based on the observations in this example, for practicing the SPC, we would suggest to avoid using extreme weight values.

From the two examples, it is seen that the way to implement the constraints affects

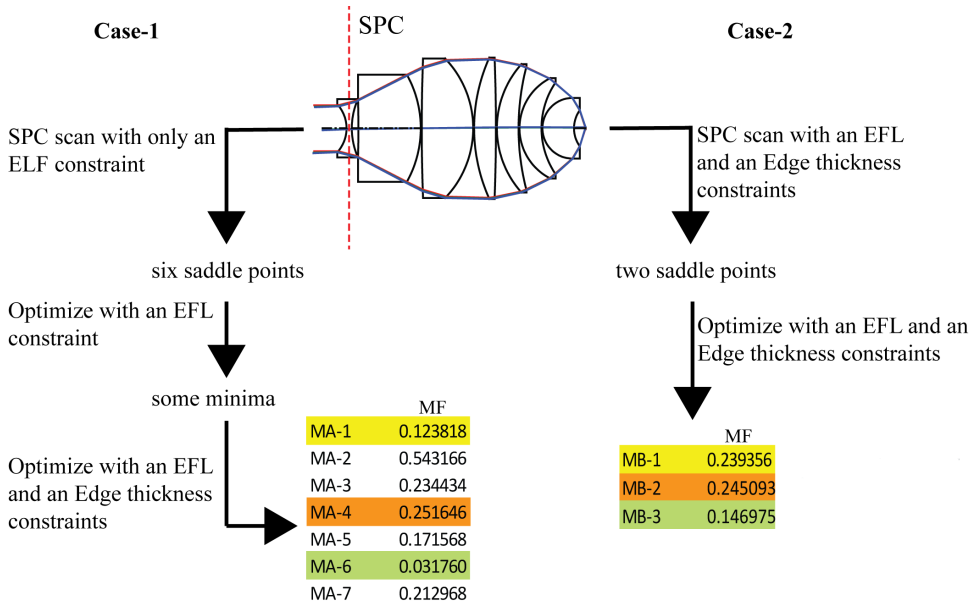


Figure 5.19: Two different design processes leading to different sets of local minima (minima with zero-thickness element). On the left: saddle point scan with only the EFL constraint, with the edge thickness constraint applied only in the last stage. In total seven minima are obtained. On the right: SPC scan with both EFL and edge thickness constraints incorporated followed by local optimization using both constraints. This led to three local minima. Systems with similar shapes are given the same colour.

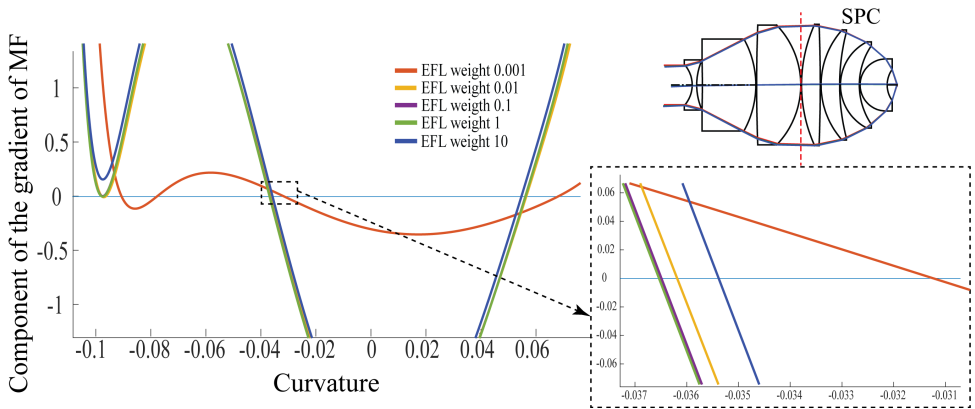


Figure 5.20: Change in the saddle point scan curve when different constraint weights are applied.

the number of saddle points acquired from SPC. This will lead to differences in the obtained minima. Regarding the constraints applied, the example shows that starting with a simple constraint generates more saddle points than applying several constraints at the same time. It is consistent with the design rule that constraints are usually added

progressively. If all constraints are added at the same time, using the Lagrange Multiplier method adds multiple new variables to the optimization problem, therefore can generate more local minima and make the global optimization more difficult. When using the weight function, there is lack of systematic knowledge on how to adjust multiple weights in one merit function. Constraints that are not massively violated should not be included in the SPC scan, but should be applied only on the resulting minima. We recommend avoiding extreme values of weight when using SPC with a merit function containing weighted constraints. In general, the weights should be adapted during the design process, depending on the problem. When applying SPC, it is necessary to keep the weights the same during the application of SPC and the following local optimizations. After obtaining the local minima, the weights can be adapted when the design process is continued.

5.4. APPLYING SPC TO A PHOTO-LITHOGRAPHIC OBJECTIVE

Lithographic systems represent one of the most sophisticated types of optical systems nowadays. With an imaging objective consisting of more than twenty lens elements, nanometer resolution has to be realized. These systems are used in the semiconductor industry to produce integrated chips [17], which boost all kinds of technology advances in the 21st century.

For these kinds of complicated optical systems with a large number of variables, the number of possible solutions increases drastically. These systems suffer more easily from ray failures. Varying variables in the system can easily cause failures in ray tracing. Normally, a new design starts from an existing one, where the system specification is close to the design requirement. The system is then modified by the designer based on certain design guidelines [7][18][19] or his/her own experience. Trial and error always evolves before obtaining a good solution.

A Deep Ultra Violet (DUV) lens from an existing patent [20] is selected. We choose the design from the embodiment described in Figure 5 of this patent. The system parameters are given in Table 5 in the patent. We use this DUV objective as an example to demonstrate how SPC can be applied in a complex case. The DUV objective is designed to work at 193 nm wavelength. The numerical aperture is 0.85, the image height is 14.02 mm and the magnification is -0.25. The system is optimized for distortion and telecentricity. The wavefront RMS error is $3.26 m\lambda$. The objective consists of twenty-three lens elements, where eight surfaces are aspheric. The system plot is presented in Figure 5.21.

This objective has been previously studied with the special version of SPC. In the special version, no curvature scan is used. The designer decides where to insert the null-element. This is always done on an existing surface. A saddle point system is directly obtained by the way the null element is inserted [16]. Two new solutions can be obtained by optimizing from the saddle point. In the study of Marinescu[21][3], examples are given to use the special version of SPC to add lenses to the original system to obtain different local minima. Afterwards, other lenses are chosen to be deleted. In this way, the system maintains similar performance with a smaller number of optical elements. However, the strategy for where to add or extract the lenses is not explained in detail.

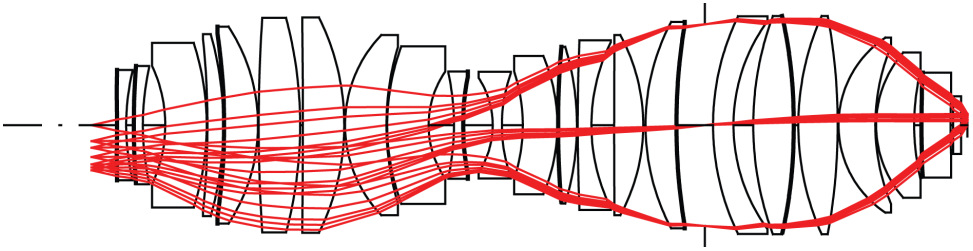


Figure 5.21: DUV objective lens from [20]. It consists of, in total, twenty-three lens elements. Eight surfaces are aspheric surfaces marked by thickened line.

This question remains unanswered by our research, and it makes the design of large systems more complicated. Marinescu's method has two steps: 1) choose the place to add a lens with SPC; 2) choose lenses to be extracted. The design process is a combination of SPC and conventional methods. The designer needs to make the decision twice on where to add and extract the lens. In the following examples, we will show that with SPC, the designer only needs to decide once where to perform SPC to achieve performance improvement.

The sixth lens (from the left) is chosen to be replaced using SPC following the example in Marinescu's thesis (where she chose to add a null-element next to it). The first step was to remove the chosen lens by gradually reducing the thickness of the lens while optimizing the system, at the same time to keep the system staying in the same basin of attraction. Once the thickness is reduced to zero, constraints are applied to make the element a meniscus null-element (equal curvature with zero-thickness). In this way, the element could be removed without affecting the system. In practise, the element does not need to be physically removed, since the next step is to perform an SPC scan at the same location.

As demonstrated in the previous section, the choice of constraints in the merit function may affect the obtained result. The recommendation is to start with a relative simple merit function. The complexity of the merit function can be increased after the SPC process. A default unconstrained merit function based on wavefront aberration⁴ is used for the phase of the SPC scan: the variables used are only the curvatures of the spherical surfaces (thirty six variables, the last element is a plane-parallel plate). Since the air spaces and lens thicknesses are not varied and the system is already very "tight", an unconstrained optimization does not drastically change the system. With the default wavefront aberration merit function, the original system has a merit function value of 0.002915, and a wavefront aberration of 3.26 $m\lambda$. After unconstrained optimization, the merit function value becomes 0.002391, and the wavefront aberration increases to 3.53 $m\lambda$. The merit function is calculated assuming that the best focus is determined by the paraxial calculation (ideal focus plane). The wavefront aberration is given by the wavefront analysis in CODE V, which traces the real rays and determines best focus plane using the real rays. Therefore, due to the different choices of the focal planes, a de-

⁴In CODE V, the merit function is based on wavefront variance. It comes from the weighted combination of the optical path difference (OPD) and the X and Y wavefront slope values.

creased merit function can bring an increased wavefront aberration. The computation technique used in the merit function is aimed for facilitating the optimization. When the sixth lens element was extracted, the merit function deteriorated to 0.013822, and the wavefront aberration increased to 12.39 m λ .

Performing an SPC scan at the position where the sixth lens is extracted, yielded six saddle points. These six saddle points lead to twelve minima with a zero-thickness element. After increasing the thickness, the twelve systems converge into four solutions. The reason for the convergence of the solutions is that when the thicknesses increase, local minima can encounter ray failure or merge with other local minima. An example of ray failure caused by thickness increase is explained in Figure 5.9. In Figure 5.22, the first two steps of performing an SPC scan on the lithographic objective are illustrated.

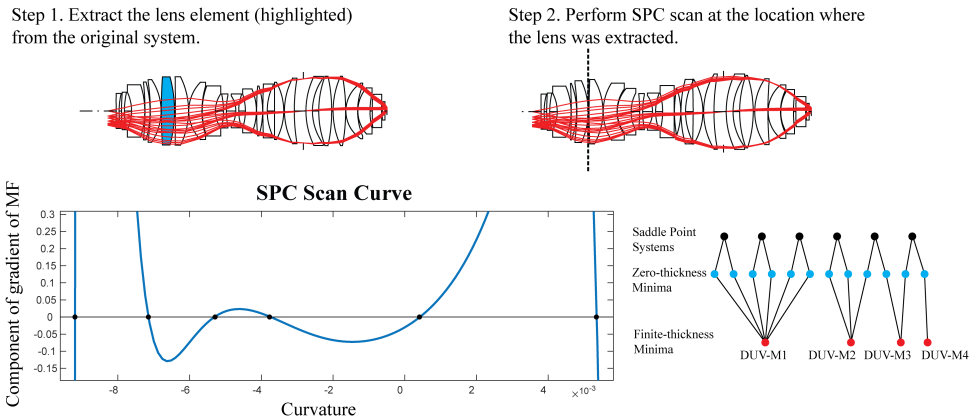


Figure 5.22: Step 1, Extract the selected element (highlighted); Step 2, Perform SPC at the location of the extraction (dashed line location). With the SPC scan, six saddle points were obtained which led to four local minima.

The system plots of the four new systems are shown on the left of Figure 5.23. The extracted and then constructed lens elements are highlighted. It is observed that the new solutions mostly differ from each other in the lenses close to the position of extraction and where the SPC scan is performed. By comparing the systems, we see that the highlighted lens and its neighbouring lens differ in the lens bending between different systems. For example, DUV-M1 has meniscus - meniscus - biconvex around the highlighted lens and DUV-M4 has meniscus - biconvex - meniscus around the highlighted lens. It is also worth noticing that there is no system identical to the original system at the top row of Figure 5.23. The constructed lens of DUV-M4 shows similar bending shape as the original system, however, the neighbouring right element has a meniscus shape which is different from the plano-convex shape in the original system. One can not expect to find a system that is identical to the original system because the solutions are obtained by optimization without constraints. The next step (Step 3) as shown in Figure 5.23 is to add constraints to the merit function and minutely optimize every solution obtained from the SPC scan. Constraints like magnification, distortion control and edge thickness etc. were added to the original merit function. Weights and number of constraints

were tailored for each system in order to reach a solution which has good performance as while as fulfill the constraints. Extra variables such as aspheric coefficients, air spaces and lens thicknesses were also used in the last stage. The corresponding solutions are listed on the right side of Figure 5.23 as DUV M1' - M4'. Judging from the system plots, it is seen that most of the differences between the systems are still localized around the constructed element. Compared to the previous stage, the major change of the further optimized solutions are the thicknesses of the lenses and the air spaces between them. This implies that in the fine tuning optimization stage, the system stays in the same basin of attraction when the curvature variables are varied.

Step 3. Apply various constraints (magnification, distortion and edge thickness etc.) to further optimize the system including more variables (aspheric surfaces, air spaces and thicknesses etc.).

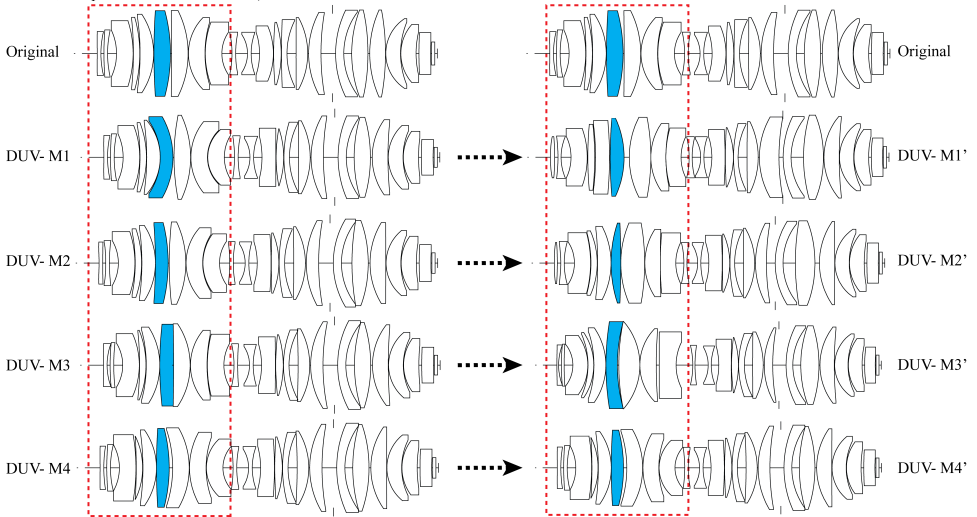


Figure 5.23: Step 3 of performing SPC on a lithographic objective: applying constraints to the systems obtained from the SPC and further optimize the systems using more variables (thickness, aspheric terms, etc.). The optimized systems are evaluated on their performances.

In Table 5.11, the performances of the four optimized solutions are listed. While minimizing the wavefront aberration (in RMS), the distortion and the telecentricity are controlled at the same time. Among the four new solutions, DUV-M3' shows better wavefront RMS (29% better than the original), distortion (16%), and telecentricity at the image side (94%). The telecentricity at the object side is worse than the original one (15%). The other three new solutions show some mixed performances, with some properties better than the original system, but others worse. For a practical case, more properties have to be considered in designing a lithographic system, e.g. thermal sensitivity, manufacturability and cost. In our example, the assessing criteria have been limited to those mentioned above. The results show that with SPC and the design flow used, new solutions in a complex lithographic system can be systematically found. In the example discussed, systems with better and worse performance as compared to the original, are obtained.

Table 5.11: Performance of the optimized solutions shown in Figure 5.23

	Wavefront RMS (mλ)	Distortion (nm)	Object Side Telecentricity* (mrad)	Image Side Telecentricity* (mrad)
Original	3.26	<1.00	<14.36	<19.18
DUV-M1'	2.96	<1.62	<17.60	<17.31
DUV-M2'	3.36	<1.43	<17.98	<5.84
DUV-M3'	2.32	<0.84	<16.56	<1.15
DUV-M4'	9.91	<2.39	<57.01	<10.05

* Telecentricity here means the deviation from telecentricity.

5.5. CONCLUSION

In this chapter, multiple examples are given to demonstrate how SPC can be applied in practical lens designs. It is recommended to integrate the SPC method with conventional strategies. In this way, more solutions can be effectively produced and from them the designer should then choose the best. In the example of a wide-angle lens, where the system is divided into two groups, we show that SPC can be applied to the individual group of lenses or to the wide-angle lens combined from the two groups of lenses. Applying SPC in this wide-angle lens design helps to produce alternative solutions in the landscape. For complex system like the UV microscope objective, guided by the designer's knowledge, SPC shows better efficiency than the conventional approaches in terms of obtaining good solutions. In order to design a lithographic objective with multiple constraints, it is discussed how SPC can be integrated in a two-step design process. New solutions with local changes are obtained with mixed performances in the sense that the obtained systems, compared to the original system, are better in some aspects and worse in others.

With the wide-angle lens, detailed steps for applying the SPC to switch to new solutions are given. The basin plots in Figure 5.7 reveal that in some cases the obtained minima depend on the choice of the optimization starting point around the saddle point. This dependency can be a property of the design landscape, or it can be caused by the numerical configuration (e.g. choices of the damping factor) of the local optimizer. Further research is necessary to understand this phenomenon. For the given example, this dependency between the starting point and the local minima with zero-thickness element did not influence the final results after increasing the thickness for the zero-thickness minima (Table 5.2).

Factors that will alter the SPC scan curve are investigated. These factors are: 1) When the number of variables increases (fourteen for the UV microscope objective), a shallow local minimum should be very well optimized to ensure that the correct SPC scan curve is obtained; 2) When constraint type and weights are varied, the corresponding SPC scan curves are different. We observe that more constraints and more rigid weights (leading to a stressed system) result in fewer saddle points than fewer constraints and more flexible weights (which leads to a relaxed system). Therefore, it is recommended to implement only the necessary constraints at the SPC scan stage. Extreme values of these constraints should also be avoided. It is better to add the other constraint later in the design process.

In conclusion, to design a complex system, applying SPC following the guidelines presented in this chapter will help to generate (intermediate) solutions effectively. When implementing the SPC method in a software package to serve as a lens optimization

tool, the cases discussed in this chapter should be taken into account (e.g. automatically verify the SPC scan curve stability) to assist the design process.

REFERENCES

- [1] Z. Hou, I. Livshits, and F. Bociort, *Practical use of saddle-point construction in lens design*, *Optical Design and Engineering VII*, **10690**, 8 (2018).
- [2] Z. Hou, I. Livshits, and F. Bociort, *One-dimensional searches for finding new lens design solutions efficiently*, *Appl. Opt.* **55**, 10449 (2016).
- [3] O. Marinescu and F. Bociort, *Saddle-point construction in the design of lithographic objectives, part 2: application*, *Opt. Eng.* **47(9)**, 093003 (2008).
- [4] S. M. Zhen Cao, Yanqiu Li, *Grouping design method of catadioptric projection objective for deep ultraviolet lithography*, *Opt. Eng.* **56(2)**, 025102 (2017).
- [5] I. Livshits, Z. Hou, P. van Grol, Y. Shao, M. van Turnhout, P. Urbach, and F. Bociort, *Using saddle points for challenging optical design tasks*, *Proc. SPIE* **9192**, 8 (2014).
- [6] F. E. Sahin and A. R. Tanguay, *Distortion optimization for wide-angle computational cameras*, *Opt. Express* **26**, 5478 (2018).
- [7] I. Livshits and V. Vasilyev, *Q and A tutorial on optical design*, *Advanced Optical Technologies* **2**, 31 (2013).
- [8] M. van Turnhout, *A systematic analysis of the optical merit function landscape: towards improved optimization methods in optical design*, *Ph.D. dissertation (Delft University of Technology)*, 63 (2009).
- [9] M. van Turnhout and F. Bociort, *Instabilities and fractal basins of attraction in optical system optimization*, *Opt. Express* **17**, 314 (2009).
- [10] W. Vollrath, *Method for manufacturing ultraviolet microscope dry objectives and microscope objectives manufactured in accordance with this method*, (1995), US Patent 5,440,422.
- [11] Y. Zhang and H. Gross, *Systematic design of microscopic lenses*, *Proc. SPIE* **10590**, 12 (2017).
- [12] Synopsis, Inc., *CODE V® (version 11.0) Optimization Reference Manual*, www.synopsis.com.
- [13] Y. Zhang and H. Gross, *Systematic design of microscope objectives. Part II: Lens modules and design principles*, *Advanced Optical Technologies* **8**, 349 (2019).
- [14] B. Brixner, *Lens design and local minima*, *Appl. Opt.* **20**, 384 (1981).
- [15] P. van Grol, F. Bociort, and M. van Turnhout, *Finding order in the design landscape of simple optical systems*, *Proc. SPIE* **7428**, 11 (2009).

- [16] F. Bociort and M. van Turnhout, *Finding new local minima in lens design landscapes by constructing saddle points*, *Opt. Eng.* **48(6)**, 063001 (2009).
- [17] T. Matsuyama, Y. Ohmura, and D. M. Williamson, *The lithographic lens: its history and evolution*, *Proc. SPIE* **6154**, 14 (2006).
- [18] D. Shafer, *Doing more with less*, *Proc. SPIE* **2537**, 11 (1995).
- [19] Z. Cao, Y. Li, and S. Mao, *Grouping design method of catadioptric projection objective for deep ultraviolet lithography*, *Opt. Eng.* **56(2)**, 025102 (2017).
- [20] R. H. W. Ulrich and H. Rostalski, *Fig. 5: projection objective for the wavelength 193 nm, refractive projection lens*, (2003), German patent WO2003075096.
- [21] O. E. Marinescu, *Novel design methods for high-quality lithographic objectives*, Ph.D. dissertation (Delft University of Technology) , 77 (2006).

6

SIMULTANEOUS MULTIPLE SURFACE (SMS) 2D DESIGNS AS STARTING POINTS FOR LENS OPTIMIZATION

The first blow is half the battle.

An 18th century English proverb.

6.1. INTRODUCTION

Choosing a sufficiently good starting configuration is crucial to every kind of optical design. For simple systems which require a limited amount of optical elements, analytical method such as paraxial calculation or aberration analysis can be used to determine a starting point. For systems with demanding requirements (e.g. larger FOV, high image quality) it is usually more convenient to select a starting configuration based on existing designs documented either in lens design books [2][3][4] or in various patents. Once the starting point is determined, the system is parameterized in design software and further improved in order to fulfill various constraints and performance requirements. In most cases, additional degrees of freedom are needed to enhance the system performance.

The most typical approach in conventional lens design to add degrees of freedoms is to introduce new lens elements. We have shown in the previous chapters that one aspect of the SPC provides a systematic and rapid approach to improve system performance by adding new lens elements. However, adding new elements in the system usually means

Parts of this chapter have been published in Optics Express **25**, 32474 (2018) [1].

bigger volume and higher cost, which are disadvantages for practical applications. Besides using additional lens elements, another way is to introduce new degrees of freedom on existing lens surfaces. For rotational symmetric systems, aspheric surfaces can be used instead of spherical surfaces. For non-rotational symmetric systems, freeform surfaces (surfaces with arbitrary shape) can be used. They permit new dimensions in optical performance by eliminating aberrations without added weight and size, thus fulfilling the need for an ever-increasing performance balanced with an equal desirable goal of ever smaller and lighter optics. Experience shows that aspheric and freeform surfaces are often more challenging to manufacture and incorporate into design than the conventional spherical surfaces because of their complex geometric shape. Nonetheless, provided the added benefit, the development of precise manufacturing techniques makes aspheric and freeform surfaces no longer a luxury and therefore can be encountered quite often in modern designs. In this chapter, we would like to focus on the aspheric surface since it shares more similarity with spherical surfaces. Freeform will not be discussed.

In common practices, the aspheric surfaces are created by utilizing the existing surfaces of an already designed spherical systems. The designer chooses which surface to be aspherized based on his/her observation. The introduced aspheric surface is then often a "mild" surface, which is described by a small polynomial order. However, if the performance improvement of system is not as expected, the designer may consider using more polynomial terms or choose more surfaces to be aspherized. The increasing choices of the degrees of freedom often trap the system in a poor local minimum (Figure 6.1[a]). In such cases, the designer's experience is often a crucial factor when trying to get out of a poor minimum in search for a good solution. In the case of utilizing aspheric surfaces, the aim is to add the correct design freedoms such that the new starting point with more degrees of freedom can lead to better system performance (Figure 6.1[b]).

Several global optimization methods such as genetic algorithms [5], evolution strategies [6], particle swarm [7] and simulated annealing [8] have been adapted to different optical design problems to avoid getting trapped in a local minimum. One can add the aspheric variables to the existing ones and run a global optimization to find out what the could be the best performance. However, these methods are based on generally applicable mathematical models, and designers understand little about the optical systems behind them.

Alternatively, for non-spherical systems, direct design methods such as the Simultaneous Multiple Surfaces (SMS) method allow to find a starting point with which a subsequent local optimization can lead to a good minimum [9][10]. Even though SMS is capable of designing good aspheric imaging systems, there is no study comparing it directly with other design strategies and describing the characteristics of different methods.

In this chapter, we consider rotationally symmetric designs, for which SMS 2D is used [11]. SMS 2D may utilize rays in the meridian plane or skew rays [12], but, in any case, obtaining a 2D profile of the surfaces. The method involves simultaneous calculation of N optical surfaces using N one-parameter bundles of rays for which specific conditions are imposed. The number of ray bundles can be greater than the number of surfaces to design when the footprints of the design bundles do not occupy the full SMS surfaces. This can be controlled by the size of the aperture stop or the relative distance between

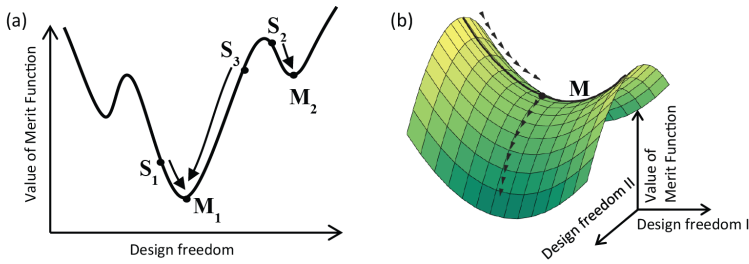


Figure 6.1: An illustration of an design landscape in 2D (a) and 3D (b). The plot (a) on the left shows that the obtained local minimum is sensitive to the starting point: Choosing S_2 as the starting point leads to a sub-optimal local minimum M_2 , while choosing S_1 and S_3 leads to a better local minimum. The plot (b) on the right shows that by adding the right design freedom, a local minimum M in lower dimension can progress further to improve system performance.

the aperture stop and the optical surfaces as demonstrated in [13][14][15]. SMS also has a freeform version called SMS 3D. However, these cases are not discussed in this thesis. We start by briefly describing the SMS method applied in this thesis. Then the design obtained with SMS is used as a starting point, followed by a single-step optimization of all parameters. The final design is compared with other designs obtained using spherical starting point and single-step, two-step, stepwise and global optimizations. Two designs are considered, a simple one and a slightly more difficult one with a larger aperture and wider field.

6.2. STARTING POINT WITH THE SMS METHOD

The standard SMS procedure for designing aspheric surfaces involves only meridian rays. It consists of two steps: selection of central segments of the surfaces and recursive generalized Cartesian oval calculation [12][16]. We will focus on SMS designs using only meridian rays. Skew rays are then controlled in a subsequent optimization.

The illustration in Figure 6.2 gives an example of how SMS works. The figure shows one of the approaches on how to design a two-surface aspheric imaging system. The following aspects need to be defined as the starting conditions: 1) Coordinates of the object points (O_1 and O_2) and image points (I_2 and I_1). The points in each pair are symmetric to each other with respect to the optical axis. 2) Vertex position (P_0) of one of the surfaces and its normal (along the optical axis); thickness of the lens (it is approximated by the projection of P_0-P_1 on the optical axis in Figure 6.2). 3) The refractive indices of the lens (n' and the medium outside the lens n_0).

With all the above parameters defined, the SMS starts with tracing the ray from O_1 to P_0 . Then the ray is refracted to P_1 , then to I_1 (the normal at P_1 can now be calculated). Once the first complete ray is traced from O_1 to I_1 , the optical path distance (OPD) from O_1 to I_1 is determined. Because of the symmetry, a perfect imaging system will produce the OPD from O_2 to I_2 identical to that from O_1 to I_1 . When tracing the ray from I_2 to P_1 , the refracted angle of the ray is known. Given the OPD and the end point as O_2 , the position of P_2 and the local normal can be calculated.

Now, it is clear that the next step would be tracing the ray from **O1** to **P2**. The aforementioned steps can be iterated such that a chain of points is constructed. These points are positioned sparsely on the chain. To get dense sampled points, a curve can be interpolated with two neighboring points (e.g. **P0** and **P2**) and their normals. A new point can be selected on this curve with its normal. Starting with the new point, we can perform the ray-tracing step again to create denser sampled points filling the gap between the original two points. The same procedure can be repeated with all the sparse points firstly created. After the dense sample points are available on both sides, surfaces can be created by fitting these points.

This is only a brief explanation to show the concept of SMS. More details on different techniques for starting conditions, surface fitting and constructions are explained in more depth in [11].

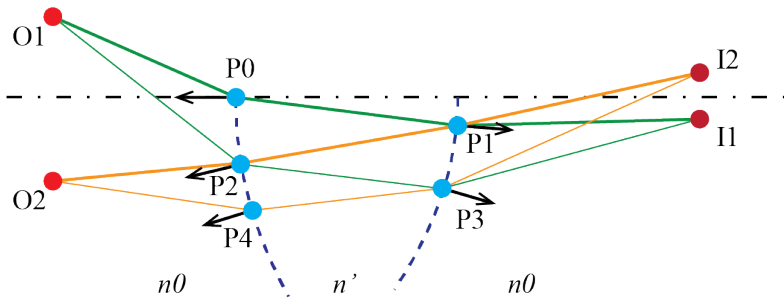


Figure 6.2: Illustration of an SMS process to simultaneously construct two surfaces. **O1** and **O2** are pre-selected points on the object side. **I1** and **I2** are the ones on the image side. **P0** - **P4** are the examples of the constructed points which are used to build the optical surfaces. Arrows indicate the direction of the surface normal at each point.

In this chapter, two designs are studied. Both designs consist of four aspheric surfaces. The two designs are distinguished by the F/Number (aperture size) and the FOV where a smaller F/Number and larger FOV is considered as a more demanding design. With the SMS 2D approach, four meridian bundles were used. It is different from the two surfaces case in the way that the starting conditions include the parts of the surfaces near the axis, which are calculated with paraxial optics [16]. In the case at hand, the four meridian ray bundles selected for the SMS design correspond to the rays emitted from four object points placed symmetrically with respect to the optical axis at infinity. The image points are located in the same way at a finite distance to match the specified focal length. The field angles associated with the object points are selected so that their Root Mean Square 2D (RMS2D) distribution curves (defined as the RMS spot diameter calculated using only meridian rays in the field) present a ripple shape over the field of view as shown in the right-hand plot in Figure 6.3 and explained in detail in [10].

Applying a standard SMS 2D procedure, for the design directions we simultaneously calculate the set of points and normals. SMS profiles are then fitted into a Forbes Qcon polynomial [17], and parameterized in the CODE V software. This system is now ready

to be optimized for the whole field of view, as in the example shown in Figure 6.3.

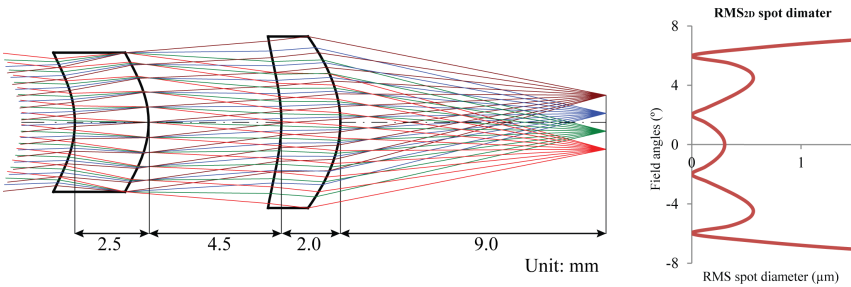


Figure 6.3: SMS 2D design introduced to CODE V after fitting the surfaces to Forbes Qcon polynomials. The change of the spot diameter through the fields is shown on the right side. The four imaging points which are used during the SMS are minimized for their spot diameter. An aperture stop is positioned at the right surface of the first lens element (when counting from left to right).

6.3. COMPARING RESULTS OF DIFFERENT DESIGN APPROACHES

Two lens systems were designed and optimized for monochromatic light ($\lambda = 587.56 \text{ nm}$) using the SMS method and other approaches mentioned in the introduction. Both lenses are made of PMMA ($n = 1.4918$ at the used wavelength) and have four aspheric surfaces in total. The two systems are referred as system 1 and system 2 where their specifications are given in Table 6.1. The object is located at infinity. System 1 has both a smaller field of view (FOV) and aperture, and system 2 has both a larger FOV and aperture.

Table 6.1: System Specification

	Image F/Number	Effective Focal Length (EFL)	Maximum Half Field of View (MHFOV)	Position of the Aperture Stop
System 1	2.24	8.60 mm	7.50°	At surface 2
System 2	1.77	10.60 mm	11.50°	Air-space between the two lenses

Three design approaches are used for the comparison: 1) SMS design with a subsequent one-step optimization; 2) optimization from a spherical system and adding aspheric coefficients, either with multiple coefficients in one-step, two-step (each step with multiple coefficients) or one-by-one (stepwise, only one coefficient per step); 3) global optimization with Global Synthesis (GS) in CODE V. Since we do not intend to compare different global optimization algorithms in this paper, we choose for comparison with SMS only GS that is known to perform very well in comparison with other global optimization and search methods in lens design [18][19]. For comparison, the systems designed all use Qcon polynomials (Appendix A.4) to represent the aspheric surfaces. The variables are the curvatures, conic constants and the coefficients of the Qcon polynomials. To better represent the surface constructed by the SMS method, we use up to the 12th order for system 1, and 16th order for system 2. The vertices are fixed in all cases. For non-GS optimization, we use the default CODE V transverse ray aberration merit

function. Two types of constraints are used. One is to constrain the EFL, and the other is to keep the normalization radii¹ the surfaces to a value exceeding the size of the surface aperture by 5%. The local optimization is done within CODE V based on a damped least square method. Vignetting was readjusted every 20 cycles, since the surface shape may vary significantly during the optimization process. Iterations are stopped when the merit function value drops to less than $1^{-5} \mu\text{m}^2$ (convergence is then assumed). For GS optimization, the same merit function and constraints are applied. However, vignetting can only be adjusted after the GS is finished.

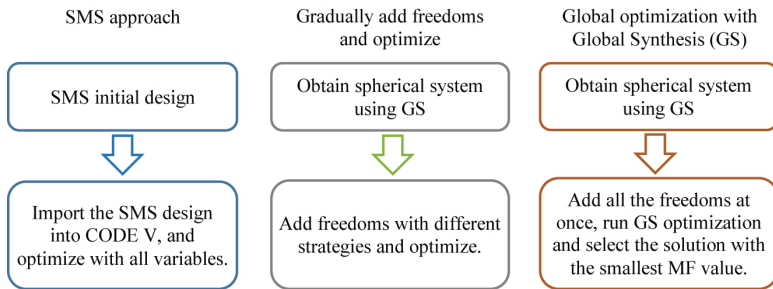


Figure 6.4: Flowchart for our different design approaches.

6

The SMS 2D procedure enforces stigmatic imaging for meridian rays at a few discrete design fields. After the system is designed with SMS, it is imported into CODE V to be optimized for better performance throughout the whole field of view and the full pupil. The surface parameters of the two SMS constructed systems are given in Table A.3 and Table A.4 in the appendix. The left flowchart in Figure 6.4 shows the design flow with SMS.

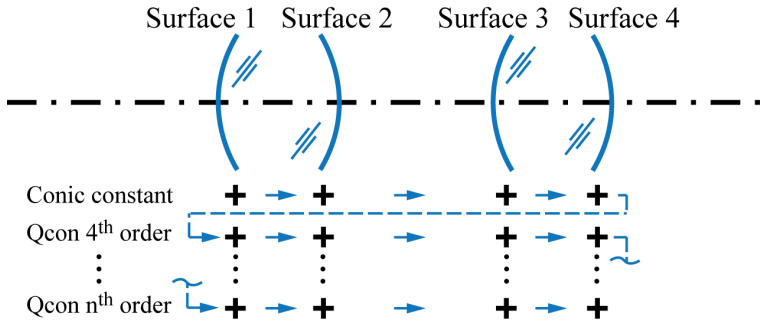


Figure 6.5: The strategy for extensive stepwise optimization. The blue shapes represent the optical surfaces. Each cross represents a freedom on the surface described by the Qcon polynomial. The blue arrows indicate how the freedoms are added step by step up to the highest order coefficients used for Qcon polynomials of all surfaces.

The second approach is conventional and intuitive. We start with a spherical system.

¹The normalization radius r_n is a parameter defined in Qcon polynomial which is used to normalize the radial distance r . The details are given in Appendix A.4.

The vertices of each surface, the lens thicknesses and air-spaces are fixed. The system specifications shown in Table 6.1 are kept the same for the spherical system. We used the global synthesis (GS) of CODE V to search through the solutions space. It took less than 10 minutes to finish because of the very small number (four) of variables used. Only one solution is obtained. Next, the aspheric coefficients are added as new variables. There are different strategies for adding variables to the system. The direct way is to free all the variables at once and optimize. For an optimization landscape with multiple minima, the optimized solutions will depend on the chosen starting point. Designers will not have control of the optimization route, and it is easy to get trapped in a poor local minimum. A common practice of experienced designers is therefore to add variables step by step. In this chapter, we apply three different strategies for adding new variables: a single-step optimization, a two-step optimization, and an extensive stepwise approach which will be explained below. A two-step approach begins by adding all the conic coefficients at once and then optimize to a local minimum with conic surfaces. Subsequently, the higher-order coefficients of the Qcon polynomials are added all together and optimized. The extensive stepwise approach requires more steps by adding one freedom in each step and then optimize again: We start by freeing the conic constant on surface 1. After optimization, new freedom is added on the next surface. The steps are illustrated in Figure 6.5. A higher-order coefficient is added as a variable to surface 1 only after surfaces 2, 3 and 4 have been given step by step the polynomial coefficient with the same order as variables. The system has been re-optimized in each step. In practice, one can add the new degrees of freedom to the system in different sequences, keeping in mind that on each surface, the coefficients should be added in increasing order. Since it is not known which stepwise strategy will lead to a better solution, in this paper, we choose an easy stepwise strategy for demonstration.

In addition to the two approaches mentioned above, a global optimization method based on CODE V's GS is used for comparison as well. The proprietary algorithm of GS can help the designer to automatically explore the solution space and synthesize new configurations from an (almost) arbitrary starting point [20]. To apply GS, we use the spherical system mentioned above as a starting point. All the degrees of freedom are added at once to the system. GS is executed with constraints on EFL and normalization radii. A time limit of two hours is set to the GS (i. e. GS stops after two hours if it is not finished). The resulting system with a reasonable system shape and the smallest merit function value, is chosen for the comparison.

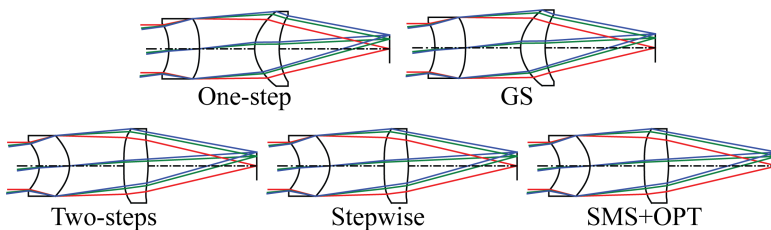


Figure 6.6: System 1: the system shapes obtained using different design approaches.

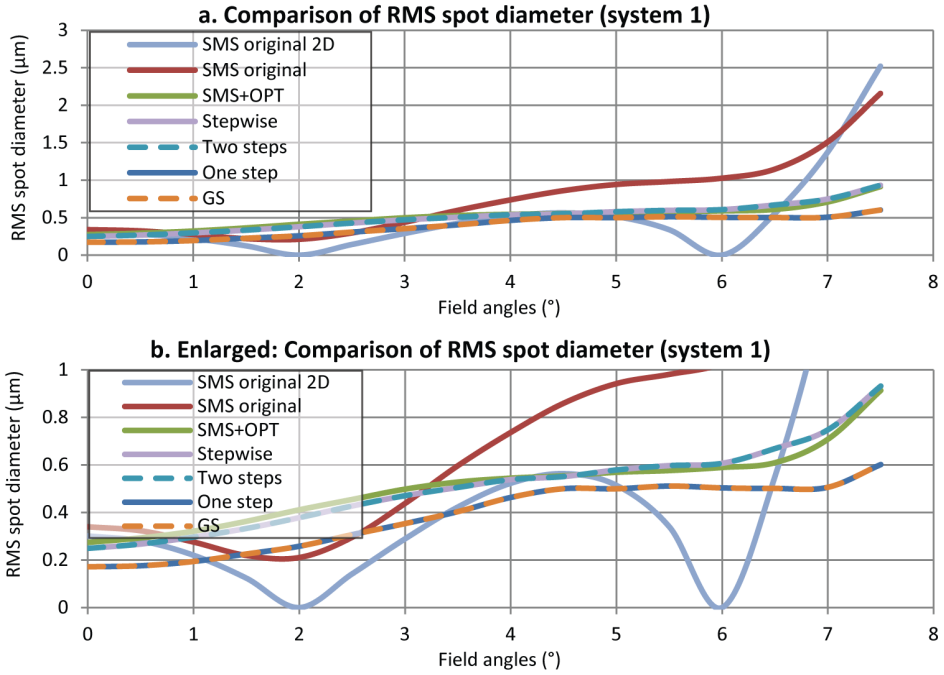


Figure 6.7: RMS spot diameter distribution curves for different design approaches: complete curves (a); enlarged vertical axis (b). The RMS spot diameter values of the starting spherical system vary from 60 to 96 μm and are not shown in the graph.

6.3.1. SYSTEM 1

The designed results for system 1 are presented in Figure 6.6. By examining the shapes of the surfaces, we observe two groups of systems: the first group consisting of the solutions obtained by GS and one-step optimization and the second group consisting of the solutions obtained by the two-step, stepwise and the SMS+OPT approaches. This is consistent with the two groups of RMS spot curves in Figure 6.7. It indicates that the solutions from five different strategies end in two different minima in the design space.

In Figure 6.7, we see that all systems generated by the different approaches show good performance with the maximum RMS spot diameter smaller than $1 \mu\text{m}$ at the field of 7.5° . The starting SMS system has a relatively large RMS spot diameter when the field is increased. Since SMS 2D construction involves only meridional rays, the RMS spot diameter is zero for the design fields (2° and 6°) only for those meridional rays (SMS original 2D curve in Figure 6.7). In Figure 6.7, a minimum is observed at 2° for the SMS original curve, however, at 6° there is no obvious minimum. Since the RMS spot diameter of SMS original in Figure 6.7 results both from meridional and skew rays, it is reasonable to expect that when the whole pupil is considered, skew rays start to have a larger influence on the spot diameter. The optimization of the system constructed by the SMS balances the spot diameter over the field by considering the rays from the whole pupil.

Except the SMS original and SMS original 2D curves, in Figure 6.7, GS and one-step

form one group and SMS+OPT, two-step and stepwise form the other group. This is consistent with the system drawings shown in Figure 6.6. The results from GS and one-step group are slightly better than the results of the other group. It can be observed that the curves of GS and one-step almost overlap. The same happens for two-step and stepwise curves in the second group. The overlap of the curves indicates that the optimization strategies lead to the same local minimum in the optimization landscape. The curve of SMS+OPT only has a very small difference in field dependence of the RMS compared to the stepwise and two steps curves, and therefore can be considered to correspond to the same local minimum. Because of the smaller values of aperture and field considered here, we have many degrees of freedom, and it is easy to achieve diffraction limited performance. In this situation, several good minima exist in the solution space as illustrated in Figure 6.8. Even though the designer still has no control during a one-step optimization, the chance for the optimization reaching a good minimum is higher than in a more difficult design problem.

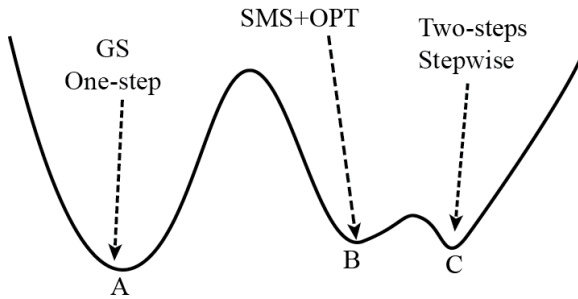


Figure 6.8: An illustration of the design landscape for system 1. We have found three equally good local minima which form two groups. Solution A (MF 0.0301) forms the first group. Solution B (MF 0.0461) and Solution C (MF 0.0450) form the second group.

Figure 6.8 shows a comparison between the efficiency of the strategies considered in this chapter. The GS result is not included because of its very different approach compared to the others: instead of optimizing from one starting point, it searches for different starting points, applies local optimizations to them, and produces multiple solutions [20]. We then usually choose the system with the smallest merit function value and perform an extra local optimization to make sure that it converges. In this case, we cannot easily use the number of optimization cycles to represent the efficiency of GS. Three solutions are generated with GS after two hours (Intel i5-3470 dual-core @3.20 GHz system). In comparison, it takes around 30 seconds for a one-step optimization to run 1000 cycles on the same computer. In Figure 6.9, the starting points are shown at cycle 0. The two groups have their average merit function values as 0.0454 and 0.0301 (the merit function value of the GS result is 0.0301). An MTF analysis shows that all five systems are diffraction limited at the wavelength used here. This is consistent with our statement that the design problem is easy and multiple equally good minima are present. Therefore, this example shows that in simple design problems like system 1, simple methods can easily lead to a good solution.

6.3.2. SYSTEM 2

System 1 has a relatively small pupil and field, which reduces the difficulty of the design. With the same approach, we used SMS to construct system 2, which has both a larger pupil and field of view. Figure 6.10 shows the five plots of the systems produced by the five mentioned optimization approaches. In contrast with the solutions of system 1 in Figure 6.6, four different system shapes are now obtained, where stepwise and SMS+OPT resulted in the same system shape. The operation with GS for system 2 is not as straightforward as that for system 1 which will be explained below. It is also worth noting that the aperture stop of system 2 is placed in between the two lens element as shown in Figure 6.10. This is to impose the requirement that, when using SMS, only four ray bundles are used for a four-surface design. It turned out that other positions of the aperture stop will lead to the use of more than four ray bundles which will not be consistent with the SMS design approach in system 1. In system 1, four ray bundles (on the left-hand of Figure 6.3) are used to construct four surfaces. The result is that four points have zero RMS spot diameter at the image plane (on the right-hand of Figure 6.3). It is demonstrated in the previous research [13][14][15] that given certain conditions, it is possible to use more than N ray bundles to construct N optical surfaces using the SMS method. This is realized by adjusting the size of the aperture stop or the distance between the aperture stop and optical surfaces. As a result, the footprints of the original design ray bundles can leave unoccupied patches on optical surfaces. These unoccupied patches can be used to couple additional ray bundles.

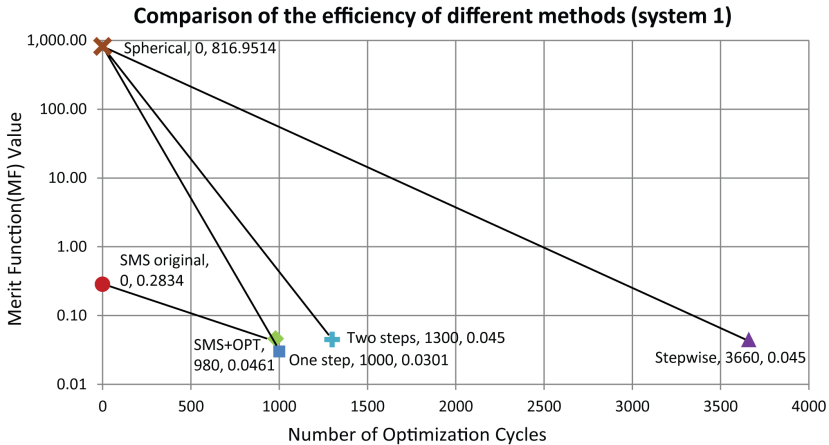


Figure 6.9: Comparison of the efficiency of different design methods analyzed, evaluated as the merit function value versus the number of cycles performed. The merit function value of the system obtained with GS is 0.0301.

Optimization with conic constants as variables ended with systems having an extremely large conic constant. Therefore, we first fixed the conic constant on all the surfaces to -1 , and ran the GS which resulted in twelve solutions. After that, the conic constants on all the surfaces of the twelve systems were freed at once as variables. Subsequent local optimization, which is the same as the one used in the other four strategies,

is applied to each of the twelve systems. The system with the smallest merit function value is chosen as the final result. In Figure 6.11, the curves of the RMS spot diameter as a function of the field angle can be again divided into two groups. One-step and two-step optimization generate poorer solutions than the other three approaches. SMS with optimization, stepwise optimization and GS obtain good solutions with close RMS dependence over the field (Figure 6.11(b)). Consistent with the same system shape in Figure 6.10, the RMS curves of SMS+OPT and stepwise in Figure 6.11 almost overlap with each other. On the other hand, the result of GS with a different system shape produces a different RMS curve, where the RMS spot diameter is bigger under 7.8° and smaller above 7.8° .

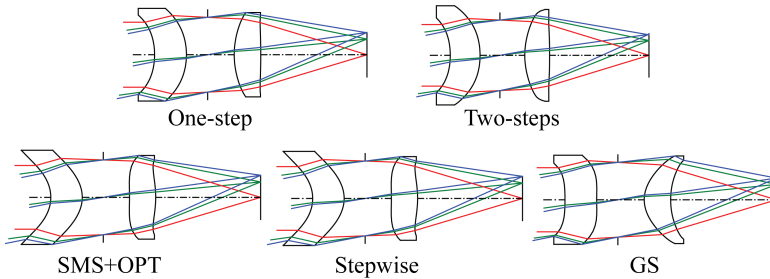


Figure 6.10: System 2: the system shapes obtained using different design approaches.

Figure 6.12 shows the comparison of the efficiency of different methods. Stepwise needs a large number of optimization cycles (8100 cycles) to converge. With a similar merit function value, fewer optimization cycles are needed for SMS optimization. This shows the SMS constructed starting point is closer to the minimum in the optimization landscape. Both one-step and two-step optimization are trapped in poor solutions with merit function values at least nine times of the best solution. From the two examples we studied, to design a simple system with small field and aperture, all approaches, including a one-step optimization, are able to obtain good solutions. However, when the field and aperture increase, simple strategies like one-step and two-step start to fail by getting trapped in poor local minima as shown in the example of system 2. A stepwise approach leads to a good local minimum with a large number of optimization cycles. An SMS constructed starting point located closer to a good solution shows its advantage in both cases where fewer optimization cycles are needed to reach a good local minimum.

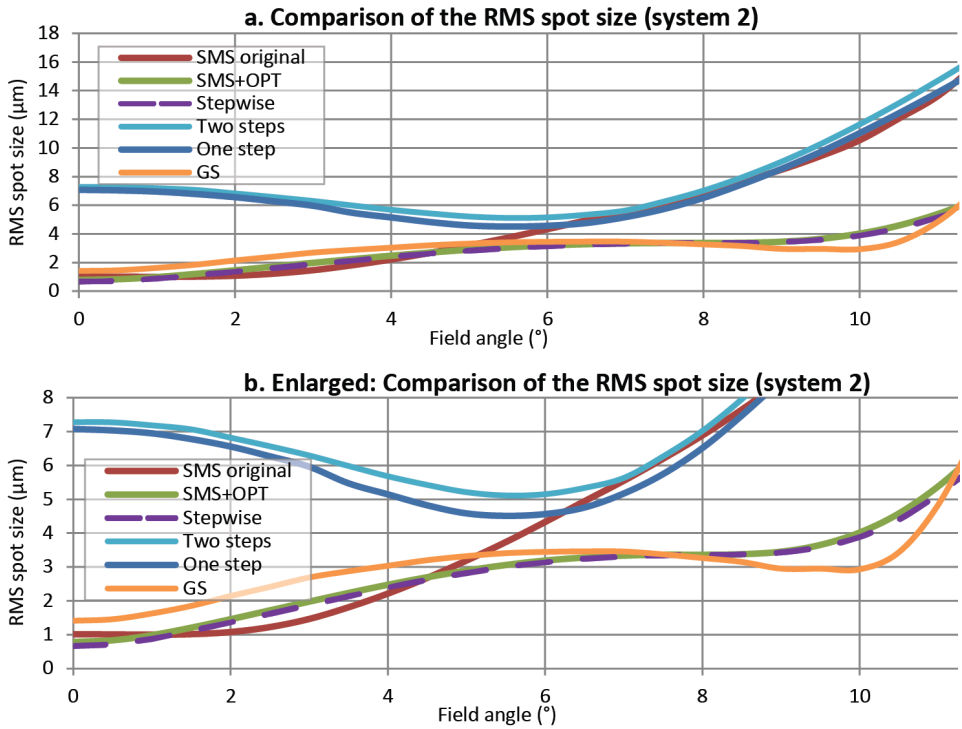


Figure 6.11: RMS spot diameter curves for system 2 using different design approaches: complete curves (a); enlarged vertical axis (b). The RMS spot diameter values of the starting spherical system vary from 124 to 142 μm from the center to the full field and are not shown in the graph.

6.4. DESIGN LANDSCAPE FOR ASPHERIC SYSTEMS

In our examples, we notice that when aspheric variables are not introduced, there is only one solution in the optimization landscape given the boundary condition for rays that are traced. However, when aspheric coefficients are added, new minima start to appear.

After adding the conic constant as a variable, multiple minima appear in the design landscape. We use the GS of CODE V in this case to search for different solutions. For system 1, we obtain four minima (Figure 6.13), and for system 2 three (Figure 6.14). The following step consisted of adding all other higher-order aspheric coefficients at once and locally optimizing the systems from the conic minima. For system 1, four different minima in conic space converge to three minima. In Figure 6.13, two of the three minima (M1-P, M3-P) are identical to the ones shown in Figure 6.6. M2-P is one of the results obtained with the GS that is not shown in Figure 6.6. GS generates multiple solutions at once and the one with the smallest merit function is chosen. For system 2, three solutions result from the three conic minima. In Figure 6.14, M1-P and M2-P are same as the results of the two-step approach and the SMS with optimization, respectively. M3-P is a local minimum which is not found by the previously discussed five approaches we used. In terms of merit function value, M3-P is almost seven times of M2-P (the smallest merit

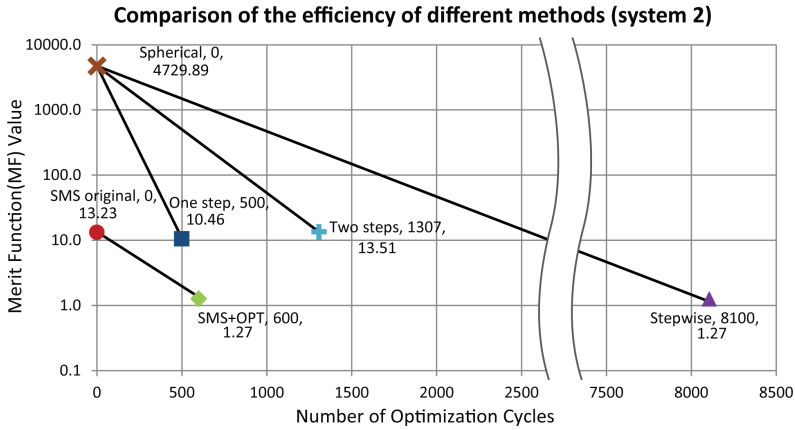


Figure 6.12: Comparison of the efficiency of different design methods. The starting points are shown at zero cycles and are connected with the straight lines with the results obtained using different design methods. The merit function value of the GS result is 1.03.

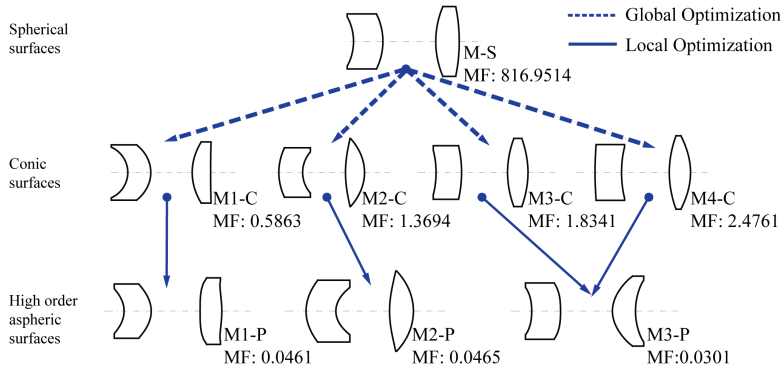


Figure 6.13: Evolution of the minima with the change of aspheric coefficients for system 1. Four minima are found with conic surfaces. They become three minima when optimized with aspheric coefficients up to 12th order. All three solutions have been obtained with the approaches we mentioned previously. M1-P and M3-P are the two solutions identical to the ones shown in Figure 6.6. M2-P is one the results obtained by GS (not the best in terms of merit function value).

function found), and approximately two-thirds of M1-P.

The results show the dynamic behavior of the optimization landscape, namely, the number of minima depends on the number and type of variables used. In the two systems we study in this chapter, the introduction of the conic surface creates new minima in the design landscape, and adding higher-order aspheric coefficients deepens the original minima. For a simple system, adding higher-order aspheric coefficients does not complicate the landscape. With local optimization, the conic solutions lead to the good solutions with higher-order aspheric coefficients. In the case of system 2 where aperture and field are large, however, adding higher-order aspheric coefficients complicates the landscape by creating new local minima. That is the reason why simple strategies are

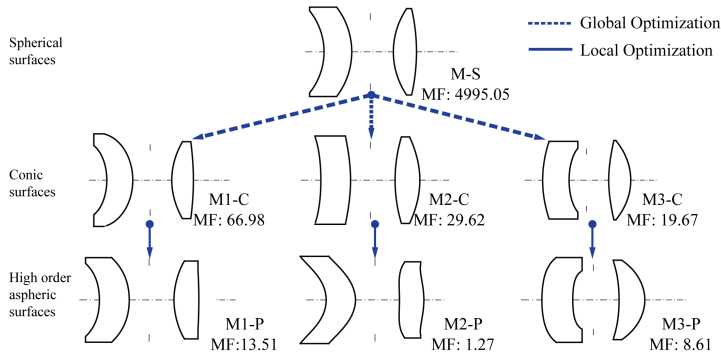


Figure 6.14: Evolution of the minima with the change of aspheric coefficients for system 2. Three minima are found with conic surfaces. Adding higher-order aspheric coefficients up to 16th order to these minima results in three different solutions. M1-P is the solutions that is also found by the two-step approach, and M2-P is the solution that is also found by the SMS with optimization in Figure 6.10.

trapped in poor local minima. In both cases, the SMS constructed starting points are already close to one of the good local minima in the landscape. Therefore, local optimizations with few optimization cycles lead to good solutions if, as starting value, the system constructed with SMS is used.

6

6.5. CONCLUSIONS

Two different optical systems comprising of two lenses are analyzed. In the analysis of the design landscape of the systems, we observe that with aspherising the surfaces, new minima emerge into the landscape. In both examples, introducing lower order aspheric coefficients (conic constant) to all the surfaces brings new local minima into the landscape, and adding higher-order aspheric coefficients deepens the minima. In a simple design with both small aperture and field (system 1), adding higher-order coefficients does not complicate the landscape by creating new local minima. In contrast, for a different design with large aperture and field (system 2), higher-order aspheric coefficients introduce new local minima, where a design process can be trapped.

We compare the SMS design approach with other design techniques for both system 1 and system 2. In the case of system 1 which is relatively simpler, all design techniques including a one-step optimization obtain solutions close to the diffraction limit. Nevertheless, surface shapes of the final designs are different which suggest that their optimization paths ended in different basins of attractions.

For the case of system 2, SMS with posterior optimization, the stepwise approach and GS find optimal solutions, while the one-step and two-step approach only obtain poor local minima. After carefully analyzing the three obtained results, we conclude that they belong to two different minima. Since designing system 2 is more demanding than system 1, less robust methods such as one-step optimization and two-step optimization do not lead to good minima. Despite the fact that three approaches finding good solutions, stepwise optimization usually takes the largest amount of optimization cycles. GS

works well for the simple case. However, for system 2, effective constraints (in this case, fixed conic constant) and further optimizations have to be implemented into GS to get reasonable solutions.

In practice, lens design problems are very different depending on the application and constraints. While we cannot guarantee that SMS will find good solutions in other design problems, the present results are encouraging. In both design problems system 1 and system 2, SMS with posterior optimization gives good solutions and we can see from Figure 6.9 and Figure 6.12 that in the two examples considered, SMS constructs good starting points which are already close to the good solutions. We believe that with this and similar comparative studies, new insight can be obtained about the characteristics of different design methods.

REFERENCES

- [1] Z. Hou, M. Nikolic, P. Benítez, and F. Bociort, *SMS2D designs as starting points for lens optimization*, *Opt. Express* **26**, 32463 (2018).
- [2] R. Kingslake, *Lens Design Fundamentals* (Academic Press, 1978).
- [3] W. J. Smith, *Modern Optical Engineering* (McGraw-Hill, 2000).
- [4] R. E. Fisher and B. Tadic-Galeb, *Optical System Design* (McGraw-Hill, 2000).
- [5] K. E. Moore, *Algorithm for global optimization of optical systems based on genetic competition*, *Proc. SPIE* **3780** (1999), 10.1117/12.363781.
- [6] J. Nagar, S. D. Campbell, and D. H. Werner, *Apochromatic singlets enabled by metasurface-augmented grin lenses*, *Optica* **5**, 99 (2018).
- [7] C. Menke, *Application of particle swarm optimization to the automatic design of optical systems*, *Proc. SPIE* **10690** (2018), 10.1117/12.2311610.
- [8] G. W. Forbes and A. E. W. Jones, *Towards global optimization with adaptive simulated annealing*, *Proc. SPIE* **1354** (1991), 10.1117/12.47874.
- [9] L. Wang, *Advances in the simultaneous multiple surface optical design method for imaging and non-imaging applications*, Ph.D. dissertation (Universidad Politécnica de Madrid) (2012).
- [10] W. Lin, P. Benítez, J. C. Miñano, J. M. Infante, G. Biot, and M. de la Fuente, *SMS-based optimization strategy for ultra-compact swir telephoto lens design*, *Opt. Express* **20**, 9726 (2012).
- [11] J. Chaves, *Introduction to Nonimaging Optics* (CRC, 2015) Chap. 8.
- [12] W. Lin, P. Benítez, J. C. Miñano, J. Infante, and G. Biot, *Progress in the SMS design method for imaging optics*, *Proc. SPIE* **8128**, 8 (2011).

- [13] P. Benítez, J. C. Miñano, M. Nikolic, J. Liu, J. Infante, and F. Duerr, *Conditions for perfect focusing multiple point sources with the SMS design method*, *Proc. SPIE* **9191** (2014), 10.1117/12.2062594.
- [14] F. Duerr, Y. Meuret, and H. Thienpont, *Potential benefits of free-form optics in on-axis imaging applications with high aspect ratio*, *Opt. Express* **21**, 31072 (2013).
- [15] F. Duerr, P. Benítez, J. C. Miñano, Y. Meuret, and H. Thienpont, *Analytic free-form lens design in 3D: coupling three ray sets using two lens surfaces*, *Opt. Express* **20**, 10839 (2012).
- [16] J. C. Miñano, P. Benítez, W. Lin, J. Infante, E. Muñoz, and A. Santamaría, *An application of the SMS method for imaging designs*, *Opt. Express* **17**, 24036 (2009).
- [17] G. W. Forbes, *Shape specification for axially symmetric optical surfaces*, *Opt. Express* **15**, 5218 (2007).
- [18] T. G. Kuper and T. I. Harris, *Global optimization for lens design: an emerging technology*, *Proc. SPIE* **1781**, 15 (1993).
- [19] D. Shafer, *Global optimization in optical design*, (1994).
- [20] Synopsis, Inc., *CODE V® (version 11.0) Optimization Reference Manual*, www.synopsis.com.

7

CONCLUSION AND OUTLOOK

7.1. CONCLUSION

In this thesis, we mainly explore the practicality of applying Saddle Point Construction (SPC) as a lens design strategy and reveal the dynamic nature as well as the complexity of the lens design (optimization) landscape. As explained in the thesis, SPC provides a way to systematically obtain new solutions for lens design. SPC requires searching for saddle points with one-dimensional scans and it does not increase the computation effort significantly. Our understanding of the method and practical experience are summarized into a flowchart (Figures 3.7 and 3.8), which illustrates the steps for performing SPC scans and obtaining new solutions.

SPC can be used as a global search tool when combined with local optimization routines. This is demonstrated in Chapter 4 and together with the first wide-angle lens example in Chapter 5. It can also be used as a tool to effectively provide a small pool of solution candidates without drastic modification of the system configuration (UV microscope objective and photo-lithographic objective in Chapter 5). SPC requires the insertion of a null element, but does not predict itself the optimal position for insertion. Therefore, SPC reduces the search significantly, but not completely.

To use SPC as a global search tool, we show that for systems containing a relatively small number (e.g. less than six) of lens elements, it is possible to start from an existing local minimum to search through the saddle points - minima network for a better solution (Figure 4.5). An original global optimization problem in a five-dimensional space is reduced to six one-dimensional scans and several local optimization routines. We observe that not all the saddle points can be constructed using SPC. Nevertheless, because of the redundancies in the saddle points - minima network, all minima can still be reached with SPC in this example.

When the design requirements (e.g. wavelength requirements, FOV) changes, the design landscape changes accordingly. It results as a changed number of saddle points and minima as well as changed links in the network. At the latter part of Chapter 4, we elaborate how the FOV change would affect the saddle points and minima. One of the

observations is that new minima can appear through this process, and SPC may miss these minima. However, these minima are usually solutions with high merit function values (worse performance). In the design cases we have investigated, using SPC can always obtain the best solution(s). Therefore, we argue that missing solutions with high merit function values is practically acceptable.

What is always valid for SPC is that, given a known local minimum, using SPC is effective to provide several other local minima. These new local minima can be compared with the original local minimum to search for better solutions. This is especially useful when considering a dynamic design landscape.

The dynamic aspect of the lens design landscape comes from the nature of the design activity - it is an interactive and continuous process between the design subject and the evaluator (designer). In Chapter 6, we show that for an ideal case where the design landscape can be constructed statically (using SMS to directly generate a starting point with all aspheric surfaces), the chance of arriving at a global minimum is high. The rest of the optimization strategies which involve gradually following a changing design landscape, are less effective: they either cannot find a satisfactory minimum or require a large amount of optimization cycles compared to SMS. In Chapter 4, we dedicate a section to discuss how a changing FOV can affect the number of solutions (and saddle points). In the example of the UV microscope objective (Chapter 5), we demonstrate how the use of the constraints could affect the design landscape by altering the number of saddle points which can be constructed. The rule of thumb is that a less constrained system has a higher chance of producing more solution candidates.

In the example of the photo-lithographic objective in Chapter 5, we demonstrate a strategy to effectively obtain solution candidates. It can be summarized as the following steps: Step one, apply the least possible number of constraints in the design; Step two, perform SPC scans at desired positions; Step three, restore all necessary constraints to further optimize the candidates for evaluation; Step four, repeat the above steps when the designer intends to escape from a local minimum.

Lens design is a science as well as an art of using and combining different design strategies and tools to obtain a satisfactory design. It is always with the good understanding of the design problem and the characteristics of tools, that a designer will be able to improve the efficiency of the design. Given the described systematic and adaptive aspects of SPC, we believe it can certainly be a useful tool in the lens designer's toolbox.

7.2. OUTLOOK

In the present work SPC has been studied by using state-of-the art commercial lens design software. It was shown that SPC is capable of producing valuable lens design results even for complex systems. In the present stage of SPC development, and with commercial software as available at present, the application of SPC is not always without challenges. Commercial lens design software is fine-tuned for standard design techniques, e.g., its optimization settings are optimal for obtaining local minima efficiently, but not for working with saddle points. For a better understanding of the present SPC challenges and for more specific practical recommendations on how to approach them, it is necessary to continue these investigations with software where the algorithms can be adapted, i.e. with software that allows full access to the source code. The conclusions will then hopefully also make the use of SPC with commercial software easier. Since past research at TU Delft has already led to algorithm improvements in commercial software, it is possible that the developers of commercial software will adapt their code to make it more suitable for design work with the SPC method. Given the present status, we summarize our recommendations for further activities in the following three aspects.

I. MAKE THE SPC METHOD MORE AVAILABLE TO THE DESIGN COMMUNITY

One of the purposes of this thesis is to depict better how SPC can be practically used in lens design. Given the presented insights in this thesis, it is important for SPC to be used more frequently in practical assignments to make it useful for lens design. In the commercial lens design software SYNOPSIS™[1], the special version of SPC is already implemented. It would be valuable if the general version of SPC and the recommendations in this thesis could also be integrated in such a commercially available software. Following are some of the engineering aspects which need to be clarified before making SPC a robust lens design and optimization tool:

1. To use SPC as a global search method, a strategy to determine the number and position of the SPC scans is necessary. In this thesis, we look at simple examples and used trial and error to determine the SPC scan positions. However, for an automated global search SPC algorithm, it should be done automatically to ensure the maximum number of saddle points;
2. When optimizing from a constructed saddle point, the neighbourhood around the saddle point should be carefully evaluated to choose the starting points for consecutive local optimization. This is related to the observation that there can be more than two basins of attraction around a constructed saddle point (Figure 5.7);
3. Before performing a SPC scan, the local minimum should be sufficiently optimized (minimized);
4. When restoring the thickness of a zero-thickness element, the thickness should be changed in practically relevant small steps such that the system is less likely to switch to a different basin of attraction¹;

¹There are cases when the solution in the original basin of attraction is very stressed and further increasing the thickness easily switches the solution to a different basin of attraction. The change becomes visible in the system drawing. If this happens, this is not a problem in most cases, because usually optimization reaches a better minimum.

5. An option should be available to allow the designer to use SPC either for all lenses in the system, or only for a group of lenses.

II. CONTINUE ASSESSING THE EFFECTIVENESS OF SPC IN LENS DESIGN

As aforementioned, the remaining question to answer is: does the saddle points - minima network obtained via SPC always contain the best or the best pool of solutions for lens design? One way to answer it is to see if there is an example to support the negative side of the question (falsification). It requires applying SPC on more design examples and comparing its results with other algorithms to verify if a good solution is missed by SPC.

Another approach would be trying to better understand the relationship between SPC revealed design network and design landscape. From our point of view, a first step would be characterization of the design landscape. For instance, to quantify the sparsity of the saddle points and minima and the merit function altitude of these critical points. Much like how real-world landscapes are characterized (such as plain and mountains), this characterization of the design landscape could reveal its property that makes SPC always (or not always) capture the best local minima. A practical recommendation would be starting with one type of lens systems that corresponds to a certain type of design landscape. The main difficulty of this approach is having an effective metrics to characterize the design landscape. In addition, it also implies a global search tool is required as reference in order to identify all the region of interest in the landscape. Given such a difficult task together with its less foreseeable practical benefit/interest, from our point of view, it is not a very promising direction.

III. EXPLORATION OF VARIOUS TECHNIQUES FOR LENS DESIGN AND OPTIMIZATION

Current research on optical design and optimization tends to treat each method individually and study its effectiveness. As the method does not "see" an optical system but a design landscape, it is a naive intuition of us that a certain method can be more effective than another given a certain type of design landscape (it requires a characterization of the design landscape as we described in II). For instance, in a hilly landscape, SPC is more effective in detecting the "valleys" while in a plain, simulated annealing is more effective in finding the bottom of the "marshes". In SPC, it already combines the saddle point construction with local optimization (damped least-squared) to obtain minima. Other collaborating strategies (e.g. evolutionary algorithm and SPC) can also be explored.

Given the complexity of the design landscape, it is an interesting direction to start combining different design and optimization methods/strategies to see if a collaboration of the methods could further improve the design efficiency.

REFERENCES

- [1] Optical Systems Design (OSD Optics), *SYNOPSIS*TM, <https://osdoptics.com/>, accessed October-2021.

A

APPENDICES

A.1. SURFACE SPECIFICATIONS OF THE WIDE-ANGLE PINHOLE LENSES WITH FOUR ELEMENTS

The original system surface specifications of the wide-angle pinhole lens in Chapter 4 is provided in the table below. The listed surface parameters corresponds to the system in Figure 4.1-a. The design requirements are listed in Table 4.1.

Table A.1: System surface specification of the wide-angle pinhole lens with four lens elements

Surface #	Surface Type	Curvature (1/mm)	Thickness (mm)	Glass refractive index*	Glass Abbe #*
1 (Object)	Sphere	0	Infinity		
2 (Stop)	Sphere	0	0.3500		
3	Sphere	-0.4433	1.5000	1.7477	50.19
4	Sphere	-0.3581	0		
5	Sphere	0.1278	2.0000	1.6152	58.08
6	Sphere	-0.1478	0		
7	Sphere	0.1743	2.0000	1.7477	50.19
8	Sphere	-0.2267	0.5000	1.8138	25.17
9	Sphere	0.1385	2.1892		
10 (Image)	Sphere	0	0		

* The glass material property are from a Russian glass catalog which we cannot retrieve.

A.2. SURFACE SPECIFICATIONS OF THE WIDE-ANGLE PINHOLE LENSES WITH SIX ELEMENTS

The original system surface specifications of the wide-angle pinhole lens in Chapter 5 is provided in the table below. The listed surface parameters corresponds to the system in Figure 5.1. The design requirements are listed in Table 5.1.

Table A.2: System surface specification of the wide-angle pinhole lens with six lens elements

Surface #	Surface Type	Curvature (1/mm)	Thickness (mm)	Glass Type*
1 (Object)	Sphere	0	Infinity	
2	Sphere	0.0146	2.5000	HLAK4L_CDGM
3	Sphere	0.2125	4.9000	
4	Sphere	-0.2091	3.0000	HLAK4L_CDGM
5	Sphere	-0.1491	1.0000	
6	Sphere	0.0109	4.4000	HLAK4L_CDGM
7	Sphere	-0.0645	5.7000	
8 (Stop)	Sphere	0	3.7200	
9	Sphere	0.0745	3.2000	HZPK2_CDGM
10	Sphere	-0.1978	1.3000	HZF52_CDGM
11	Sphere	-0.0584	0	
12	Sphere	0.0844	3.0000	HLAK52_CDGM
13	Sphere	-0.0446	5.5000	
14	Sphere	0	0.5250	HK11_CDGM
15	Sphere	0	0.0200	
16 (Image)	Sphere	0	0	

* The glasses are from the CDGM catalog in CODE V.



A.3. SURFACE SPECIFICATIONS OF THE SMS CONSTRUCTED SYSTEMS

The system parameters of system 1 and system 2 described in Chapter 5 are listed in the following tables. The first two tables provide the parameters of the systems after the SMS construction. The optimized system based on the first two tables are given in the last two tables.

Table A.3: Surface parameters of system 1 constructed with the SMS method

Surface	1 (Object)	2	3 (Stop)	4	5	6 (Image)
Surface type	Sphere	Qcon Asphere	Qcon Asphere	Qcon Asphere	Qcon Asphere	Sphere
Material		PMMA		PMMA		
Curvature (1/mm)	0	-0.3142	-0.3424	-0.1874	-0.2925	0
Thickness (mm)	Infinity	2.50	4.50	2.00	9.00	0
Normalized Radius (mm)		2.20	2.40	2.75	2.90	
K		-1.2937	-1.3707	-7.0733	-0.9741	
QC4		5.9543E-02	2.7731E-02	1.1610E-01	1.2458E-01	
QC6		9.8831E-03	2.8067E-03	-7.3253E-03	-9.2028E-03	
QC8		-4.3303E-04	2.6278E-04	8.3891E-04	9.1018E-04	
QC10		-6.7898E-05	2.8571E-05	-2.4930E-05	-1.9440E-05	
QC12		5.5475E-06	-1.7237E-05	6.7655E-06	9.0658E-06	

QC means the coefficient of Qcon polynomial.

Table A.4: Surface parameters of system 2 constructed with the SMS method

Surface	1 (Object)	2	3	4 (Stop)	5	6	7 (Image)
Surface type	Sphere	Qcon Asphere	Qcon Asphere	Sphere	Qcon Asphere	Qcon Asphere	Sphere
Material		PMMA			PMMA		
Curvature (1/mm)	0	-0.0823	-0.1386	0	-0.0876	-0.2183	0
Thickness (mm)	Infinity	3.00	2.00	2.50	2.50	10.00	0
Normalized Radius (mm)		3.70	3.40		3.50	3.70	
K		-9.2639	-5.7686		6.0225	-0.3232	
QC4		1.4285E-01	1.8381E-01		2.6624E-01	2.1069E-01	
QC6		2.8156E-02	4.0717E-02		1.2446E-02	9.0396E-03	
QC8		-1.6114E-03	3.0800E-03		3.4723E-03	2.4930E-03	
QC10		-1.4590E-04	7.4292E-04		5.4780E-04	4.6335E-04	
QC12		2.0759E-05	1.2620E-04		9.8619E-05	9.9287E-05	
QC14		-1.4991E-06	2.9401E-05		1.7711E-05	2.0624E-05	
QC16		-2.7096E-06	3.8057E-06		3.1363E-06	4.9458E-06	

QC means the coefficient of Qcon polynomial.

Table A.5: Surface parameters of system 1 constructed with the SMS method

Surface	1 (Object)	2	3 (Stop)	4	5	6 (Image)
Surface type	Sphere	Qcon Asphere	Qcon Asphere	Qcon Asphere	Qcon Asphere	Sphere
Material		PMMA		PMMA		
Curvature (1/mm)	0	-0.4307	-0.3952	0.0098	-0.1153	0
Thickness (mm)	Infinity	2.50	4.50	2.00	9.00	0
Normalized Radius (mm)		2.3309	2.5920	3.1730	2.8996	
K		-0.8916	-1.0712	-1.0000	-2.6340	
QC4		1.6274E-02	2.7731E-02	6.8250E-01	3.4386E-01	
QC6		9.7144E-03	2.8067E-03	4.0248E-03	2.6399E-02	
QC8		-4.0643E-04	2.6278E-04	1.0928E-02	8.1496E-03	
QC10		-1.1117E-04	2.8571E-05	4.7472E-04	1.5117E-03	
QC12		-8.8182E-05	-1.7237E-05	4.7380E-05	1.1745E-04	

QC means the coefficient of Qcon polynomial.

Table A.6: Surface parameters of system 2 constructed with the SMS method

Surface	1 (Object)	2	3	4 (Stop)	5	6	7 (Image)
Surface type	Sphere	Qcon Asphere	Qcon Asphere	Sphere	Qcon Asphere	Qcon Asphere	Sphere
Material		PMMA			PMMA		
Curvature (1/mm)	0	-0.2921	-0.3233	0	-0.0407	-0.1148	0
Thickness (mm)	Infinity	3.00	2.00	2.50	2.50	10.00	0
Normalized Radius (mm)		4.1468	4.6408		4.1310	3.7007	
K		-0.8802	-2.0004		-7.7743	-3.1404	
QC4		3.2669E-02	-6.5028E-01		1.3960E+00	6.2198E-01	
QC6		1.0872E-01	1.5641E-01		3.5767E-03	6.6385E-02	
QC8		-5.0938E-03	-5.3001E-03		3.1591E-02	2.6863E-02	
QC10		-1.4142E-03	1.7891E-03		2.5853E-03	8.0621E-03	
QC12		2.8801E-05	-2.5384E-04		1.3342E-03	2.4871E-03	
QC14		-1.4259E-04	2.7349E-04		1.7310E-04	7.4316E-04	
QC16		-1.8754E-04	-7.0801E-05		2.3699E-05	1.1682E-04	

QC means the coefficient of Qcon polynomial.

A.4. DEFINITION OF THE QCON POLYNOMIALS

The Qcon (*con* stands for conic) polynomials in CODE V defines an asphere on a base conic, and are most applicable to "strong" aspheres. The relationship between surface sag and the Qcon polynomials is given by the following equation.

$$z = \frac{cr^2}{1 + \sqrt{1 - (1 + k)c^2r^2}} + u^4 \sum_{m=0}^{13} a_m Q_m^{con}(u^2), \tag{A.1}$$

where z is the sag of the surface parallel to the z -axis (in lens unit). c is the vertex curvature. k is the conic constant. r is the radial distance which equals $\sqrt{x^2 + y^2}$. r_n is the normalization radius which is not explicitly given in the formula. u equals r/r_n . a_m is the m^{th} Qcon coefficient. Q_m^{con} is the m^{th} Qcon polynomial. In CODE V, up to the 13th term of the Qcon polynomial is used. Since the order of the coefficient is expressed in terms of the order of u . The Qcon coefficients in CODE V are expressed as 4th, 6th, ..., 30th orders.

Table A.7: Qcon Polynomials up to the 6th term

Term	Qcon Polynomial
1	1
2	$-(5 - 6x)$
3	$15 - 14x(3 - 2x)$
4	$-\{35 - 12x[14 - x(21 - 10x)]\}$
5	$70 - 3x\{168 - 5x[84 - 11x(8 - 3x)]\}$
6	$-[126 - x(1260 - 11x\{420 - x[720 - 13x(45 - 14x)]\})]$

More details about the Qcon polynomial can be found in [1].

REFERENCES

[1] G. W. Forbes, *Shape specification for axially symmetric optical surfaces*, *Opt. Express* **15**, 5218 (2007).

CURRICULUM VITÆ

Zhe HOU

20-03-1988 Born in Hohhot, Nei Mongol, P. R. China

WORK EXPERIENCE

2019 – Now Metrology Design Engineer, ASML (Eindhoven, the Netherlands)

EDUCATION

2013 – 2023 PhD Candidate
Program: **AD**vanced **OP**tical **SY**stem Design, EU FP7 Marie Curie ITN
Main host:
Delft University of Technology (Delft, the Netherlands, 2013–2018)
Secondment:
Datalogic (Bologna, Italy, Jan–May, 2016)
Polytechnic University of Madrid (Madrid, Oct–Dec, 2015)

2011 – 2013 Master of Science, Applied Physics
Program: **Optical Science and Technology**, EU Erasmus Mundus
Delft University of Technology (Delft, the Netherlands, 2012–2013)
Warsaw University of Technology (Warsaw, Poland, 2011–2012)

2007 – 2011 Bachelor of Optical Engineering
Zhejiang University (Hangzhou, P. R. China)

LIST OF PUBLICATIONS

Refereed publications

1. **Z. Hou**, M. Nikolic, P. Benitez, and F. Bociort, "SMS2D designs as starting points for lens optimization," *Opt. Express* **26**, 32463-32474 (2018).
2. **Z. Hou**, I. Livshits, and F. Bociort, "One-dimensional searches for finding new lens design solutions efficiently," *Appl. Opt.* **55**, 10449-10456 (2016).
3. A. Reyes-Reyes, **Z. Hou**, E. van Mastrigt, R. C. Horsten, J. C. de Jongste, M. W. Pijnenburg, H. P. Urbach, and N. Bhattacharya, "Multicomponent gas analysis using broadband quantum cascade laser spectroscopy," *Opt. Express* **22**, 18299-18309 (2014).

Conference proceedings

1. **Z. Hou**, I. Livshits, and F. Bociort, "Practical use of saddle-point construction in lens design," *Proc. SPIE 10690*, Optical Design and Engineering VII, 1069007 (2018).
2. **Z. Hou** and F. Bociort, "Reducible complexity in lens design," *Proc. SPIE 9626*, Optical Systems Design 2015: Optical Design and Engineering VI, 96260J (2015).
3. I. Livshits, **Z. Hou**, P. van Grol, Y. Shao, M. van Turnhout, P. Urbach, and F. Bociort, "Using saddle points for challenging optical design tasks," *Proc. SPIE 9192*, Current Developments in Lens Design and Optical Engineering XV, 919204 (2014).

Conference contributions

1. **Z. Hou**, Y. Zhang, and F. Bociort, "Design Using Saddle Point Construction in Complex Lens Systems," oral presentation at the European Optical Society Biennial Meeting, Delft, the Netherlands (October 2018).
2. **Z. Hou**, I. Livshits, and F. Bociort, "Practical use of saddle-point construction in lens design," oral presentation at the SPIE Optical Systems Design conference, Frankfurt, Germany (May 2018).
3. **Z. Hou**, "One-dimensional searches for finding new lens design solutions efficiently," Excellent Oral Presentation Award at the International Doctoral Students Conference on "Opportunities and Challenges Arises from Global Technological Revolution", Zhejiang, China (May 2017).
4. **Z. Hou**, F. Bociort, I. Livshits, and H.P. Urbach, "A systematic study on the design landscape of a pin-hole lens," oral presentation at the 10th International Conference on Optics-photonics Design and Fabrication, Weingarten, Germany (March 2016).

5. **Z. Hou**, F. Bociort, and I. Livshits, "Reducible complexity in lens design - replacing high-dimensional search by one-dimensional searches to find new solutions efficiently," poster presentation at the IST Science Day, Delft, the Netherlands (October 2015). **Z. Hou** and F. Bociort, "Reducible complexity in lens design," invited Paper at the SPIE Optical Systems Design conference, Jena, Germany (September 2015).
6. F. Bociort, P. van Grol, I. Livshits, **Z. Hou**, Y. Shao, P. Urbach, "Saddle-point methods for systematic design," the European Optical Society Biennial Meeting, Berlin Adlershof, Germany (September 2014). I. Livshits, **Z. Hou**, P. van Grol, Y. Shao, M. van Turnhout, P. Urbach, and F. Bociort, "Using saddle points for challenging optical design tasks," the SPIE OPTICAL ENGINEERING + APPLICATIONS conference, San Diego, U.S.A. (August 2014).

ACKNOWLEDGEMENT

It has been a long journey for this PhD, and I have to constantly remind myself that it is indeed now the end. Some people say that PhD can be a lonely journey, however, I would not be able to undertake it all by myself. Looking back to the years, I realize that the support I got during this journey was crucial for me to accomplish the PhD. I often think the following analogy: The process is similar to an optimization problem which converges slowly. The support I got through the years helped me to escape from the stagnating "local minima". Step by step, it finally approaches the end of accomplishment. Given the length of this PhD, memories are scattered, and it is difficult to capture every piece of gratitude that I want to express. Nevertheless, I would like to take the opportunity to thank my dear colleagues, mentors, friends and family during this precious journey.

First and foremost, I would like to express my deepest gratitude to my daily supervisor, Dr. Florian Bociort, for his unwavering guidance and support throughout my doctoral studies. I remember those days we had many insightful discussions on papers and my results in Florian's office, where my understanding of the SPC method deepened. As a supervisor, Florian always provides prompt feedback to my questions which I find very helpful during my studies. I extend my appreciation to my promotor, Prof. Paul Urbach, for his constructive and insightful comments, which are instrumental for the quality of the research and this thesis. Paul often asks questions from unique perspectives which I tend to overlook. It helped me tremendously in sharpening my thoughts.

I am also grateful to the members of my thesis committee, especially Prof. Wilbert Ijzerman, for his valuable feedback and constructive criticism on my work.

I am deeply indebted to Prof. Joseph Braat, who helped me a lot at the last phase of my PhD. Joseph helped review most of the chapters and provided valuable comments. In 2019, after I left the group to work in Eindhoven, Joseph invited me to his home to discuss the thesis and motivate me to finalize the work. Besides direct support for the thesis, as an emirates professor, Joseph used to often join the "students' coffee break" in the group and shared anecdotes with us. It added some nice flavours to the monotone research day. Thank you, Joseph!

This PhD project was funded by a Marie-Curie actions initial training networks (ITN) programme, ADOPSYS, where multiple academia and industry partners were involved. Within the programme, I had the chance to learn from experts from different optical design fields. I would like to offer a big thanks to Prof. Irina Livshitz, for her contribution to many initial design examples in this thesis. I am glad that I had the chance to witness how Irina, an experienced optical designer, designed lenses step by step. I would also like to express my sincere thanks to the two secondments host: At CeDInt, Polytechnic University of Madrid, Prof. Juan Miñano, Prof. Pablo Benítez and Dr. Milena Nikolić introduced me to the world of non-imaging optics. They are helpful and always avail-

able for questions. The collaboration resulted in a chapter of this thesis and a journal publication. At Datalogic, Italy, I was supported by Kurt Vonmetz and Federico Canini, where I had the chance to design and test the prototype of a light homogenizer intended for real-life application in bar-code scanners. At this time, I cannot recall many names of the supervisors in ADOPSYS, however, I did enjoy every training and summer school we had, and I am grateful for their roles in the programme. I do remember the names of the 12 ESRs (early-stage researchers). They are Milena Nikolić, Yufeng Nie, Adam Hirst, Simone Sorgato, Olga Baladron-Zorita, Marco Mout, Ton Verhoeven, Rafael Falcon, Lijo Thomas, Nenad Zoric, and Priya Dwivedi. We were like close classmates, joining training and conferences together. I truly appreciate this experience and learned a lot from the team. I do wish we could gather again somewhere in the future.

I also want to thank Dr. Yueqian Zhang, for his contribution to the design examples where I tried to combine SPC with the designer's experience. Yueqian is a great friend and expert in optical design. Besides optics, Yueqian is knowledgeable about culture and history, from which I always learn new things in our discussions.

I am grateful to the staff, faculty and student fellows in the Optics Research Group. Since my master's study, I spent 6 years in the group, which equals 1/5 of my life by the time I left. I would like to first thank Yvonne, who was group assistant when I was in the group. Yvonne was immensely helpful throughout my PhD: she helped with basically all the administrative work which enabled me to have sufficient focus on the research topics. Yvonne is also very kind and cares about the students. When I had low time during my PhD, she was always available to talk and brought positivity to me. I would like to extend my special gratitude to Nandini, who was also pivotal during my PhD journey. Nandini was my master's thesis supervisor and continued to support me during my PhD. She is a reliable mentor. I can always be open in front of her, and get valuable advice. She cares about students, especially the inner side, and her support improves the whole experience of my PhD a lot. Thank you, Nandini! I do want to thank everyone from the group, and I try my best to recall the (first) names: Chris Christie, Lidija Nikolic, Roland Horsten (thanks for handling all the computer stuff!), Thim Zuidwijk, Rob, Peter Somers and Mrs Somers, Jeff Meisner, Aurèle Adam (glad to work with you on the optics courses), Man Xu, Silvania Pereira, Omar El Gawhary, Pascal van Grol (thanks for all the SPC inspiration!) Thomas Liebig, Bob Kruizinga (thanks for the CODEV advice), Nitish Kumar, Goptika Ramanandan, Nishan Kumar, Lei Wei, Luca Cisotto, Alessandro Polo, Katsiaryna Ushakova, Alberto Assafrão, Adonis Reyes Reyes (Thanks for all the good advice and nice vibes), Matthias Strauch (thanks for all the fun non-academic activities), Hamed Ahmadpanahi (good time working in the same office), Yifeng Shao (big thanks for all the help), Ying Tang, Zheng Xi, Peiwen(Doris) Meng, Po-Sheng Chiu (good luck with whatever the next step!), Po-Ju Chen (good luck with finalizing your thesis), Sander Konijnenberg (thanks for the inspiring discussions back in the group and offering help to check the summary translation), Xukang Wei, Zheng Zhu, Yuxin Wang (thanks for the tennis lessons), and those may be I cannot recall by now, however, essential to my wonderful experience in the group. Luckily, I can continue working with some of them in ASML.

I would like to express my special thanks to Dr. Conor Sheil. We met at an SPIE conference, and I was inspired by his presentation on mobile camera design using as-

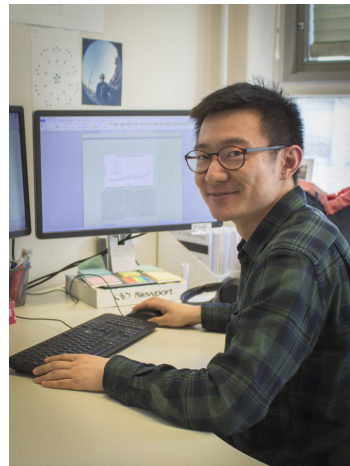
pheric lenses. I sent a follow-up email asking for some information about the design technique, and Conor replied with very detailed procedures containing notes on each step. This kind of positive out-of-expectation brought me warm feelings about the community, where researchers share, collaborate and together make steps.

At the last phase of my PhD, my family and friends played an instrumental part in supporting me mentally and physically in order to reach this very point. Thank you, Zhi Li, my therapist, for helping me regain motivation during the Corona period. Thank you, Hui Lin, for helping me create the beautiful cover art and realize the ideas in my head. Thank you, Jie Li, for making time to prepare the delicious cake I planned for the defence. Jie Li is also active on social media where she being an inspiring example for me as one can make work and life fun at the same time. Thank you, Chenying, Xiaofei, Ye Huang, Yandan, Xuechao, Chen Li, Simon, Weizhe, Wu Sun, and Hanzhen for the good time together skiing, travelling, and cat-socializing. Of course, also a big thanks to my lovely cat friend, Jean-Jean, who literally stayed by my side to accompany me in finalizing the last piece of this work.

Last but not least, I would like to thank, sincerely, my girlfriend Suyuan Zhang, who supported me relentlessly during this period, and my parents, who are always by my side, providing unconditional love. Nothing is taken for granted, and I am truly grateful for everything you did.

Zhe Hou

March 2023, Eindhoven



(photo taken on 2017 at the office of Optics Research Group)



ISBN 978-94-6384-431-4



9 789463 844314

 **TU**Delft