A SYSTEMATIC ANALYSIS OF THE OPTICAL MERIT FUNCTION LANDSCAPE

TOWARDS IMPROVED OPTIMIZATION METHODS IN OPTICAL DESIGN

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. J. T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op

dinsdag 14 april 2009 om 15.00 uur

door

Maarten VAN TURNHOUT

natuurkundig ingenieur
geboren te Rotterdam.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. H. P. Urbach.

Copromotor: Dr. F. Bociort.

Samenstelling promotiecommissie:

Rector Magnificus,	voorzitter
Prof. dr. H. P. Urbach,	Technische Universiteit Delft, promotor
Dr. F. Bociort,	Technische Universiteit Delft, copromotor
Prof. dr. I. L. Livshits,	State University of Information Technologies, Mechanics and Optics, St. Petersburg, Rusland
Prof. dr. T. D. Visser,	Technische Universiteit Delft
Prof. dr. ir. A. Gisolf,	Technische Universiteit Delft
Prof. dr. ir. C. Roos,	Technische Universiteit Delft
Dr. C. Menke,	Carl Zeiss AG, Duitsland

This research was financially supported by the Dutch Technology Foundation STW (project 06817).

ISBN 978-90-78314-11-0

Copyright © 2009 by M. van Turnhout

All Rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, mechanical, photocopying, recording or otherwise, without prior written permission of the author.

A free electronic version of this thesis can be downloaded from:
http://www.library.tudelft.nl/dissertations

Summary

A Systematic Analysis of the Optical Merit Function Landscape: Towards Improved Optimization Methods

A major problem in optical system design is that the optical merit function landscape is usually very complicated, especially for complex design problems where many minima are present. Finding good new local minima is then a difficult task. Over the past decades, significant progress in the field of global optimization has led to development of powerful software packages, which facilitates the optimization task. However, avoiding poor-quality optical systems remains very challenging, particularly if the number of optimization variables is large.

In this thesis, we try to make the optimization task easier by analyzing characteristics of the optical design space. However, without focusing on particular features, it becomes practically unmanageable to examine properties of the optimization solution space. We show that by considering saddle points with a Morse index of 1, a certain degree of order is present in the optical design space. This order is also observed when we change the dimensionality of the optimization problem.

In order to determine the initial configurations that lead to a given local minimum, the so-called basins of attraction for that minimum should be considered. The set of all starting configurations that are attracted to a local minimum is the basin of attraction for that minimum. Computing the basins of attracting is very time-consuming, and it is difficult to visualize them in a multidimensional merit function space. Moreover, we show that the basin shapes highly depend on the local optimization method.

Saddle points are critical points in the merit function landscape which always remain on the boundaries of basins of attraction, independent of the used optimization method. They can be used to systematically travel from one basin of attraction to another. Because detecting saddle points is more difficult than detecting local minima, we present a method, which we call Saddle-Point Construction (SPC), that can be used to construct saddle points in a simple way. We prove that, if the dimensionality of the optimization problem is increased in a way that satisfies certain mathematical conditions (the existence of two independent transformations that leave the merit function unchanged), then a local minimum is transformed into a saddle point.

By using SPC, lenses are inserted in an existing design in such a way that subsequent optimizations on both sides of the saddle point result in two different system shapes, giving the designer two choices for further design. In spite of theoretical novelty, the practical implementation of the method is very simple. We discuss three simple exam-

ples that illustrate the essence of the method. In these examples, we use a simple and efficient version of the SPC method. The method can be used in essentially the same way for arbitrary systems.

We also present a generalized version of the SPC method, so that the restrictions, which are needed for the special version, are removed. We show how, by performing a succession of one-dimensional calculations, many local minima of a given global search can be systematically obtained from the set of local minima corresponding to systems with fewer lenses. As a simple example, the results of the Cooke triplet global search are analyzed. In this case, the vast majority of the saddle points found by saddle point detection can in fact be obtained in a much simpler way by saddle point construction, starting from doublet local minima.

Many optical design programs use some form of the damped least-squares method for local optimization. We show that damped least-squares algorithms, with maximized computational speed, can create sensitivity with respect to changes in initial conditions. In such cases, starting points, which are very close to each other, lead to different local minima after optimization. Computations of the fractal capacity dimension show that the basins of attraction have a fractal structure.

Fractals increase the complexity of the optimization problem, and they are a source of unpredictable behavior during the design process. One of the goals of this thesis is to make optical system designers aware of the possibility of the presence of instabilities in the optimization process. The inherent complexity of the design landscape limits the sizes of the basins of desirable solutions. In addition, instabilities in optimization, if present, further increase complexity and decrease predictability by mixing basins of desirable and undesirable solutions together. We show that a better understanding of the unexpected instabilities can be obtained by using low damping values in a damped least-squares method.

Finally, we explore the use of a damped least-squares method for a purpose that goes beyond local optimization. In lens design, damped least-squares methods are typically used to find the nearest local minimum to a starting configuration in the merit function landscape. We show that a low merit function barrier, which separates an unsatisfactory solution from a neighboring one that is better, can be overcome by using low damping and by allowing the merit function to temporarily increase. However, such an algorithm displays chaos, chaotic transients and other types of complex behavior. A successful escape of the iteration trajectory from a poor local minimum to a better one is associated with a crisis phenomenon that transforms a chaotic attractor into a chaotic saddle. The results also enable a better understanding of peculiarities encountered with damped least-squares algorithms in conventional local optimization tasks.

Delft, March 2009 Maarten van Turnhout

Samenvatting

Systematische analyse van het optische foutfunctielandschap: Streven naar verbeterde optimalisatiemethoden in optisch ontwerp

Een groot probleem in optisch systeemontwerp is dat het optische foutfunctielandschap gewoonlijk erg gecompliceerd is, vooral in het geval van ontwerpproblemen met vele lokale minima. Het is dan moeilijk om goede nieuwe lokale minima te vinden. In de afgelopen decennia heeft een aanzienlijke vooruitgang op het gebied van globale optimalisatie er toe geleid dat er nu krachtige software pakketten zijn die het optimaliseren makkelijker maken. Het blijft echter een grote uitdaging om optische systemen met een slechte afbeeldingskwaliteit te ontwijken, zeker wanneer het aantal optimalisatievariabelen groot is.

In dit proefschrift proberen we het optimalisatieprobleem eenvoudiger te maken door de optische ontwerpruimte te analyseren. Het is echter praktisch onmogelijk om eigenschappen in de oplossingsruimte van het optimalisatieprobleem te onderzoeken, zonder ons te concentreren op specifieke kenmerken. We tonen aan dat er een zekere orde in de optische ontwerpruimte aanwezig is wanneer we naast lokale minima ook zadelpunten met een Morse index waarde van 1 beschouwen. Deze orde zien we ook wanneer we de dimensionaliteit van het optimalisatieprobleem veranderen.

Om de systeemconfiguraties te bepalen die bij de optimalisatie naar een bepaald lokaal minimum gaan, beschouwen we het zogenoemde aantrekkingsgebied van dat minimum. De verzameling van alle beginconfiguraties die worden aangetrokken tot een lokaal minimum is het aantrekkingsgebied van dat minimum. Het berekenen van aantrekkingsgebieden kost echter veel rekentijd en het is erg moeilijk om ze in een multidimensionale ruimte voor te stellen. Bovendien hangen de vormen van de aantrekkingsgebieden in hoge mate af van de lokale optimalisatiemethode.

Zadelpunten zijn kritische punten in het foutfunctielandschap die altijd op de grenzen van de aantrekkingsgebieden liggen, onafhankelijk van de gebruikte optimalisatiemethode. Ze kunnen worden gebruikt om op een systematische manier van het ene aantrekkingsgebied naar het andere te gaan. Zadelpunten zijn echter moeilijker te detecteren dan lokale minima. We tonen aan dat met een speciale methode, die we zadelpuntconstructie (Saddle-Point Construction, SPC) noemen, zadelpunten eenvoudig geconstrueerd kunnen worden. Een lokaal minimum wordt in een zadelpunt getransformeerd wanneer de dimensionaliteit van het optimalisatieprobleem wordt verhoogd en er aan bepaalde wiskundige condities wordt voldaan (namelijk het bestaan van twee onafhankelijke transformaties die de foutfunctie onveranderd laten).

Met SPC worden lenzen op zo'n manier in een bestaand lensontwerp gevoegd dat achtereenvolgende optimalisaties aan beide zijden van het zadelpunt resulteren in twee verschillende systeemvormen, die de ontwerper twee keuzes geven voor toekomstig ontwerp. We bespreken drie simpele voorbeelden die de essentie van de methode illustreren. In deze voorbeelden maken we gebruik van een simpele en efficiënte versie van de SPC methode. Op dezelfde manier kan de methode ook gebruikt worden voor algemene systemen.

Naast de speciale versie van de SPC methode, beschrijven we ook een algemene versie, waarvoor de beperkingen van de speciale versie niet gelden. We laten zien hoe door het uitvoeren van een reeks van ééndimensionale berekeningen vele lokale minima van een bepaalde globale zoektocht systematisch kunnen worden gevonden uit lokale minima die één lens minder hebben. Als eenvoudig voorbeeld analyseren we de resultaten van de globale zoektocht van het Cooke Triplet. In dit voorbeeld kan de overgrote meerderheid van de zadelpunten, die gevonden worden met een detectiemethode, ook eenvoudiger worden verkregen door het toepassen van zadelpuntconstructie op de lokale minima van het doublet.

Veel optische ontwerpprogramma's gebruiken een bepaalde vorm van de gedempte kleinste-kwadratenmethode voor lokale optimalisatie. We laten zien dat gedempte kleinste-kwadratenalgoritmes een grote gevoeligheid voor de begincondities kunnen hebben wanneer ze voor rekensnelheid geoptimaliseerd zijn. In zulke gevallen gaan startpunten die zeer dicht bij elkaar liggen bij de optimalisatie naar verschillende lokale minima. Berekeningen van de fractale dimensie laten zien dat de aantrekkingsgebieden een fractale structuur hebben.

Fractalen vergroten de complexiteit van het optimalisatieprobleem en zijn een bron van onvoorspelbaar gedrag tijdens het ontwerpproces. Een van de doelen van dit proefschrift is om optische systeemontwerpers bewust te maken van de mogelijkheid dat er instabiliteiten in het optimalisatieproces aanwezig kunnen zijn. De inherente complexiteit van het ontwerplandschap beperkt de grote van aantrekkingsgebieden van gewenste oplossingen. Vervolgens vergroten instabiliteiten de complexiteit van het optimalisatieprobleem en verkleinen ze de voorspelbaarheid door het met elkaar vermengen van aantrekkingsgebieden van gewenste en ongewenste oplossingen. We tonen aan dat een beter begrip van instabiliteiten kan worden verkregen door gebruik te maken van lage dempingswaardes in een gedempte kleinste-kwadratenmethode.

Tot slot hebben we het gebruik van een gedempte kleinste-kwadratenmethode onderzocht voor een reden die verder gaat dan lokale optimalisatie. In lensontwerp worden gedempte kleinste-kwadratenmethoden vaak gebruikt om het lokale minimum te vinden dat het dichtst bij een gekozen beginconfiguratie ligt. We laten zien dat een lage foutfunctiebarrière, die een onbevredigend minimum scheidt van een naburig beter lokaal minimum, overwonnen kan worden door één of meerdere iteraties met verminderde demping uit te voeren en door toe te staan dat de foutfunctie tijdelijk stijgt. In sommige gevallen vertoont zo'n algoritme echter chaotisch gedrag, chaotische transients en ander soort van complex gedrag. Een succesvolle ontsnapping van het iteratietraject uit een slecht lokaal minimum naar een beter minimum gaat samen met een bijzonder verschijnsel waarbij een chaotische attractor in een chaotische zadel transformeert. De onderzoeksresultaten geven een beter begrip van de eigenaardighe-

den die kunnen optreden bij gedempte kleinste-kwadratenalgoritmes in traditionele lokale optimalisaties.

Delft, maart 2009 Maarten van Turnhout

Contents

Summary **iii**

Samenvatting **v**

1 Introduction **1**
1.1 Aberrations in optical system design 2
1.2 Optical system design 4
1.3 Goal and outline of this thesis 8

2 Optical system optimization **11**
2.1 Introduction 11
2.2 Local optimization 13
2.2.1 Steepest descent 14
2.2.2 Conjugate gradient 14
2.2.3 Newtonian methods 16
2.2.4 Damped least-squares 16
2.3 Global optimization 18
2.3.1 Simulated annealing 19
2.3.2 Genetic algorithms 19
2.3.3 Global Explorer 20
2.3.4 Detecting saddle points 21
2.4 Conclusions 23

3 Finding new local minima by constructing saddle points **25**
3.1 Introduction 25
3.2 Systems of thin lenses in contact 27
3.3 SPC in a nutshell 31
3.4 SPC: theory 33
3.5 SPC: examples 39
3.5.1 Example 1: Generating doublets from a singlet 40
3.5.2 Example 2: Generating quintets from a quartet 41
3.5.3 Example 3: Obtaining a Double Gauss design 42
3.6 Non-degenerate and degenerate merit function behavior during SPC 43
3.7 Conclusion 46

4 Looking for order in the optical design landscape **47**
4.1 Introduction 47
4.2 Generalizing the SPC method 49
4.3 Generalized SPC: examples 51
4.4 Why are there so many minima? 53
4.5 Why do non-NESP's exist? 60
4.6 Conservation of topological degree 63
4.7 Conclusion 65

5 Instabilities and fractal basins of attraction **67**
5.1 Introduction 67
5.2 Method for stable behavior 69
5.3 Fractal basins of attraction 72
5.3.1 Two-dimensional doublet optimization problem 72
5.3.2 Seven-dimensional optimization problem 77
5.3.3 Saddle-Point Construction example 79
5.4 Instabilities in optimization 82
5.5 Conclusion 85

6 Chaotic behavior **87**
6.1 Introduction 87
6.2 Algorithm with adaptive damping for searching beyond local minima . . . 88
6.3 Period doubling route to chaos 90
6.4 Escaping from a poor local minimum 95
6.5 Instabilities in conventional damped least-squares optimization 100
6.6 Conclusion 101

7 Conclusions **103**

A Examples of Sadde-Point Construction in the special case **107**
A.1 Example 1: Generating doublets from a singlet 107
A.2 Example 2: Generating quintets from a quartet 110
A.3 Example 3: Obtaining a Double Gauss design 113

Bibliography **117**

Acknowledgments **123**

About the author **127**

List of publications **129**

Chapter 1

Introduction

Maxwell's work shows that the behavior of light is a manifestation of the laws of electromagnetism. Almost all optical phenomena can be understood in terms of electromagnetic wave fields as described by Maxwell's equations. However, in most cases, rigorous solutions of Maxwell's equations are impractical in optical design [1]. Optical configurations [1] are usually too complex to describe their behavior without simplifications in the description of the propagation of light. Fortunately, in many cases, the full electromagnetic or even wave description of light is not needed to calculate the performance of optical systems.

The field of geometrical optics gives the necessary approximations. In geometrical optics, single rays are used for the description of light, and diffraction effects are neglected. Rays are considered as paths along which radiation energy travels, and which are normal to the wavefronts [2]. The ray approximation is valid if the objects with which light interacts are much larger than the wavelength of light. In many cases, disregarding the finiteness of the wavelength (which is usually in the order of 10^2 nm) is a good approximation [3]. Even though diffraction and interference are not longer observable, the simplifications made in geometrical optics more than compensate the inaccuracies.

The direction of ray propagation is altered by reflection and refraction, and optical systems use those effects to form images. An optical system is defined by a number of parameters, such as the curvatures (or radii) of the surfaces, the distances between them, the materials that are used in the construction of the system, etc. The challenge for an optical designer is to find the best suitable values for all parameters in order to obtain a configuration with the desired image quality.

The design of optical systems has been practiced for many years. Until the late 1940's, all optical designs had to be made by hand calculation, which took many hours of work and patience. After the introduction of the computer, a lot of computation time was saved in the design process. Nowadays, the computer plays an important role in optical design. Finding good quality optical systems is mainly done by using powerful optimization algorithms from various optical design software packages.

Modern optical design is mostly based on the laws of geometrical optics. Although this

[1] Here, the term 'configuration' means the basic form of the system, which includes the number of elements, and their distribution within the lens system.

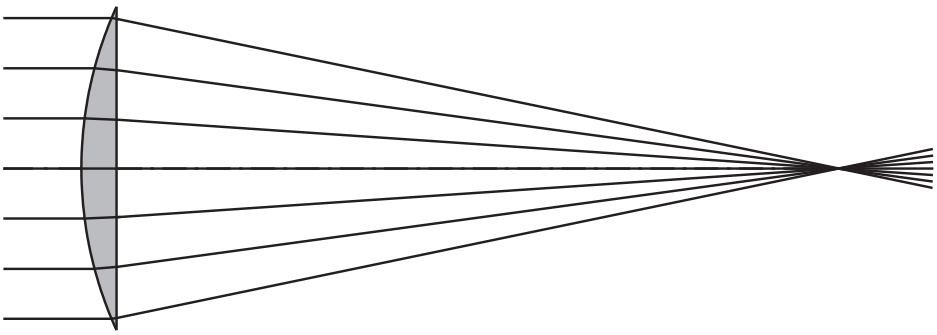


Figure 1.1: *A perfect point image: all incoming rays are focused to a single point on the optical axis.*

simplifies the design process a lot, it does not make the optimization issue an easy problem to solve. Richard P. Feynman said about geometrical optics [4]:

> *"Geometrical optics is either very simple, or else it is very complicated."*

Our research is an attempt to understand and to explain the 'very complicated' part. In this thesis, we try to find properties that introduce a degree of order in the design space, making optical system optimization less complicated.

In the next section, we give a very brief introduction on aberrations in optical system design. Aberrations are defects that reduce the imaging quality of optical systems. In Section 1.2, we describe the design of optical systems, where the goal is to minimize the aberrations to a satisfactory level. Section 1.3 describes the research topics and the outline of this thesis.

1.1 Aberrations in optical system design

To evaluate the performance of an optical system, rays should be traced through the system. Ray tracing describes the transport of rays by applying the law of refraction at the first optical surface, locating where the transmitted rays strike the next surface, then applying the refraction law again, and so on until the rays hit the image surface. The law of refraction (also known as Snell's law) is given by:

$$n \sin i = n' \sin i', \tag{1.1}$$

where n and n' are the refractive indices before and after refraction of a ray; i and i' are the angles of incidence and refraction at the interface between the media with refractive indices n and n', respectively.

In a perfect point image, all the image-forming rays intersect each other exactly in a single point (see Figure 1.1), which is equivalent to the definition that, in case of perfect imaging, all wavefronts are spherical. In the domain of paraxial (also called Gaussian) optics, rays are considered close enough to the optical axis to ensure that all terms higher than quadratic can be neglected. This makes wavefronts indistinguishable from

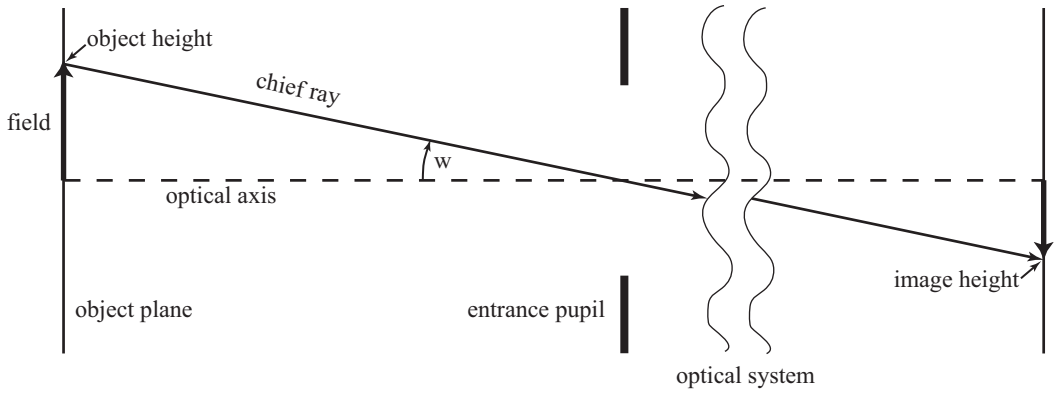


Figure 1.2: *Relationships between different forms of field specification.*

a sphere, implying aberration-free imaging. By considering the angles of the rays to be infinitely small, the sine functions in Equation (1.1) can be approximated by Taylor series, truncated after the first term:

$$n\, i = n'\, i'. \tag{1.2}$$

In geometrical optics, the quality of an optical system is characterized by ray aberrations, which depend on aperture and field. A field can be defined by an object angle (angle w in Figure 1.2), by the object height, or by the paraxial (or real) image height.

For optical systems with sufficiently small field and numerical aperture, the transmitted rays have very small angles and heights relative to the optical axis. When the light is monochromatic, this paraxial case results in an aberration-free image. For larger field and aperture sizes, we obtain larger ray angles and heights. Those deviations from the paraxial domain give rise to monochromatic aberrations. They are typically described as a Taylor expansion in terms of aperture and field. The coefficients in the expansion describe the type and magnitude of the existing aberrations.

We can express the deviation from ideal image formation as wavefront aberration or as transverse ray aberrations. The wavefront aberration W is a function in terms of optical-path difference, which measures the departure of the imaging wavefronts from their ideal spherical shape. The transverse ray aberration Ξ expresses the distance between the true intersection point of the ray with the image plane and the Gaussian image point. By developing a series expansion of the wavefront or transverse aberration, we are able to study aberrations analytically.

Best known are the expressions for the surface contributions to the third-order aberration coefficients of symmetric systems [5]. The third-order aberrations (also called primary or Seidel aberrations) are the five lowest order monochromatic aberrations. They are often used as a quick estimation of the aberrations in an optical system, and they can help in understanding the basic dependencies of aberrations on system parameters.

Let a ray through an optical system be defined by the two vectors $\boldsymbol{\sigma} = (\sigma_x, \sigma_y)$ and $\boldsymbol{\tau} = (\tau_x, \tau_y)$, where the components of $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ are the normalized aperture and field coordinates, respectively. In the case of rotationally symmetric optical systems, we can

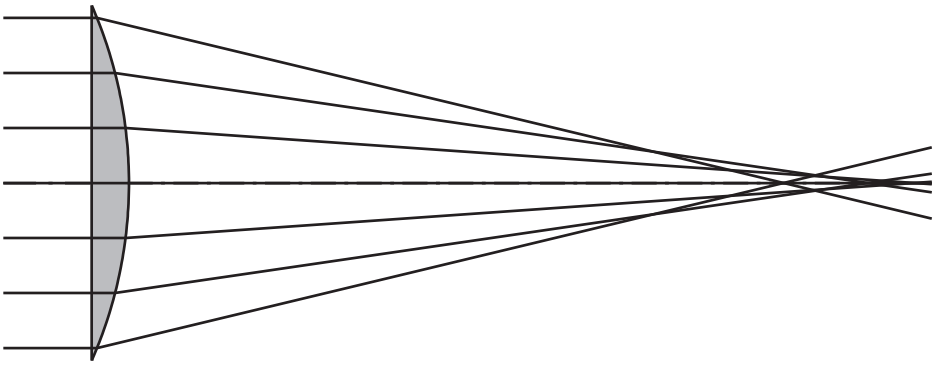


Figure 1.3: *A lens that suffers from spherical aberration ($\boldsymbol{\tau} = (0,0)$). The rays that strike the lens further from the axis focus closer to the lens than rays that enter closer to the axis. The resulting image is not a point, but a spot.*

write the third-order transverse aberration $\boldsymbol{\Xi}_{(3)}$ as [5]:

$$\boldsymbol{\Xi}_{(3)} = \frac{1}{nu}\left\{\left[S_I(\boldsymbol{\sigma}^2) + 2S_{II}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau}) + (S_{III}+S_{IV})\boldsymbol{\tau}^2\right]\boldsymbol{\sigma} + \left[S_{II}\boldsymbol{\sigma}^2 + 2S_{III}(\boldsymbol{\sigma}\cdot\boldsymbol{\tau}) + S_V\boldsymbol{\tau}^2\right]\boldsymbol{\tau}\right\}, \tag{1.3}$$

where $\boldsymbol{\sigma}^2 = \sigma_x^2 + \sigma_y^2$, $\boldsymbol{\sigma}\cdot\boldsymbol{\tau} = \sigma_x\tau_x + \sigma_y\tau_y$, $\boldsymbol{\tau}^2 = \tau_x^2 + \tau_y^2$, and u is the marginal ray angle in the image space. The five coefficients S_I, S_{II}, S_{III}, S_{IV}, and S_V are called the five Seidel aberration sums: spherical aberration, coma, astigmatism, field curvature (also called the Petzval sum), and distortion. Figure 1.3 shows an example of a lens that suffers from spherical aberration.

The third-order aberrations of an optical system are represented as a sum of contributions of all surfaces. These contributions are independent and they can be computed from paraxial ray data. Only higher order aberrations are affected by aberrations at the preceding surfaces. If it is impossible to eliminate all Seidel aberrations in a certain optical system, they can only be controlled and balanced to provide the best possible result for that system and application.

Because the refractive index of any medium other than vacuum varies with wavelength (also known as dispersion), aberrational properties of any refracting optical system are functions of wavelength. Therefore, chromatic aberrations exist when different wavelengths are used. Their magnitude is typically comparable to that of the third-order aberrations [6].

1.2 Optical system design

In practice, the formation of perfect images is almost impossible. In general, there will always be some aberrations that reduce the imaging quality. In fact, aberrations are inherent shortcomings of a lens, even when it is made of the best glass without manufacturing defects. Balancing of aberrations is then the underlying principle for the correction of optical systems [6]. The subsystems in an optical system may exhibit aberrations, but they can partially compensate each other so that the overall aberration of the entire system is minimal.

The goal of optical system design is to find the list of system parameters that satisfies given design targets within tolerances determined by the application. This is not an easy task, because optical systems can have many parameters and the relationship between the system parameters and the image defects is highly nonlinear. Modern optical system software packages have powerful optimization algorithms, which alleviates the optimization problem. They take a starting design and automatically produce a new optimized design.

The optical designer has to specify the following requirements before optimizing a starting configuration [7]:

1. A function of the system parameters that expresses the quality of the system at every design stage. This function is often called the merit function.

2. A subset of system parameters (the optimization variables) that are automatically changed during optimization.

3. A set of constraints to limit the variation domain of the optimization variables.

Starting from an initial configuration having N optimization variables, local optimization iteratively reduces the value of the merit function by changing the variables until it arrives at (or comes close enough to) a (constrained) minimum in the N-dimensional merit function landscape.

We describe a point in the merit function landscape by the vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$, where the components are the N optimization variables. At a local minimum, the gradient of the merit function (MF) with respect to the optimization variables vanishes:

$$\nabla MF(\mathbf{x}) = \left(\frac{\partial MF}{\partial x_1}, \frac{\partial MF}{\partial x_2}, \ldots, \frac{\partial MF}{\partial x_N} \right) = 0. \tag{1.4}$$

A small change in the optimization variables will then only lead to an increase of MF. In general, there are many local minima in the MF landscape, and the attempt to find the best one among those minima is called global optimization.

The merit function expresses the behavior of an optical system into a single number, such that a smaller value yields a better system performance. It considers both image defects and constraints for the system. In the next chapter, we mathematically define the merit function. The merit function can be visualized as a mountain scenery above the variable space, where the local minima are located in the valleys, separated by hills. A starting system for optimization can be thought of as a ball placed at a specific location in this landscape. By using a local optimization algorithm, the ball rolls down until it reaches a point where the slope is zero in all directions.

Saddle points are the passes in the mountain scenery. The merit function around these points behaves like the two-dimensional surface of a horse saddle. Figure 1.4 shows a typical saddle point with a Morse index value of 1 in a merit function landscape with only two variables. As in the case of local minima, the gradient of the merit function vanishes at saddle points. It can be seen that such a two-dimensional saddle point is a maximum in one direction (the downward direction, shown by the thick white curve

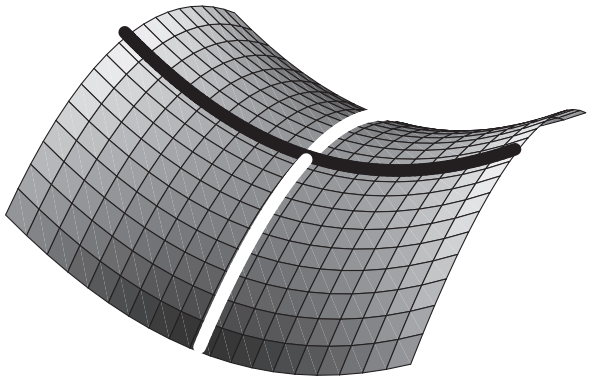


Figure 1.4: *Two-dimensional saddle surface in the merit function landscape. The saddle point is situated at the crossing of the thick black and white curves.*

in Figure 1.4), and a minimum along the other direction (the upward direction, shown by the thick black curve), which is perpendicular to the first one. If we have more than two variables, some directions at the saddle point are upward, and some are downward directions. The value of the Morse index is given by the number of downward directions. In the introduction of Chapter 3, we will come back to this discussion.

In Figure 1.5(a), we show a projection of the merit function on a two-dimensional variable space of a monochromatic doublet [2]. The curvature of the second surface is plotted along the vertical axis, and the curvature of the third surface is plotted along the horizontal axis. The curvature of the first surface is kept constant, and the last surface is used to keep the effective focal length constant. A plot very similar to Figure 1.5(a) has already been published in Reference [8, Figure 7].

In the example we present here, there are four local minima (A, B, C, and D), surrounded by equimagnitude contours along which the merit function remains constant. Minimum B has a high merit function value. The system shapes of the four doublet local minima are shown in Chapter 5. As expected, the best three local minima are located almost on the contours for third-order spherical aberration (SA) equal to zero (thick gray curves) and close to the zero coma curves (dashed gray), see Figure 1.5(b). The three small black points in Figure 1.5 are saddle points in the merit function landscape, which are exactly situated where two merit function equimagnitude lines cross.

By moving from a local minimum over a saddle point, a neighboring local minimum can be found. The problem is that the detection of saddle points (finding the mountain passes) is a very difficult mathematical problem. In this thesis we show that many saddle points can be obtained in a very simple way.

In order to understand the essence of optimization, we make very small steps in the direction of steepest descent (i.e. in the direction of $-\nabla MF$). The steps are always perpendicular to the equimagnitude contours. Figure 1.5(b) shows the resulting six paths in the steepest descent direction. All paths end up in one of the four local minima. In Chapter 5, we compute the regions of starting doublet configurations which lead to the same local minimum after local optimization.

[2] Axial color correction will be included for systems discussed in Chapter 4.

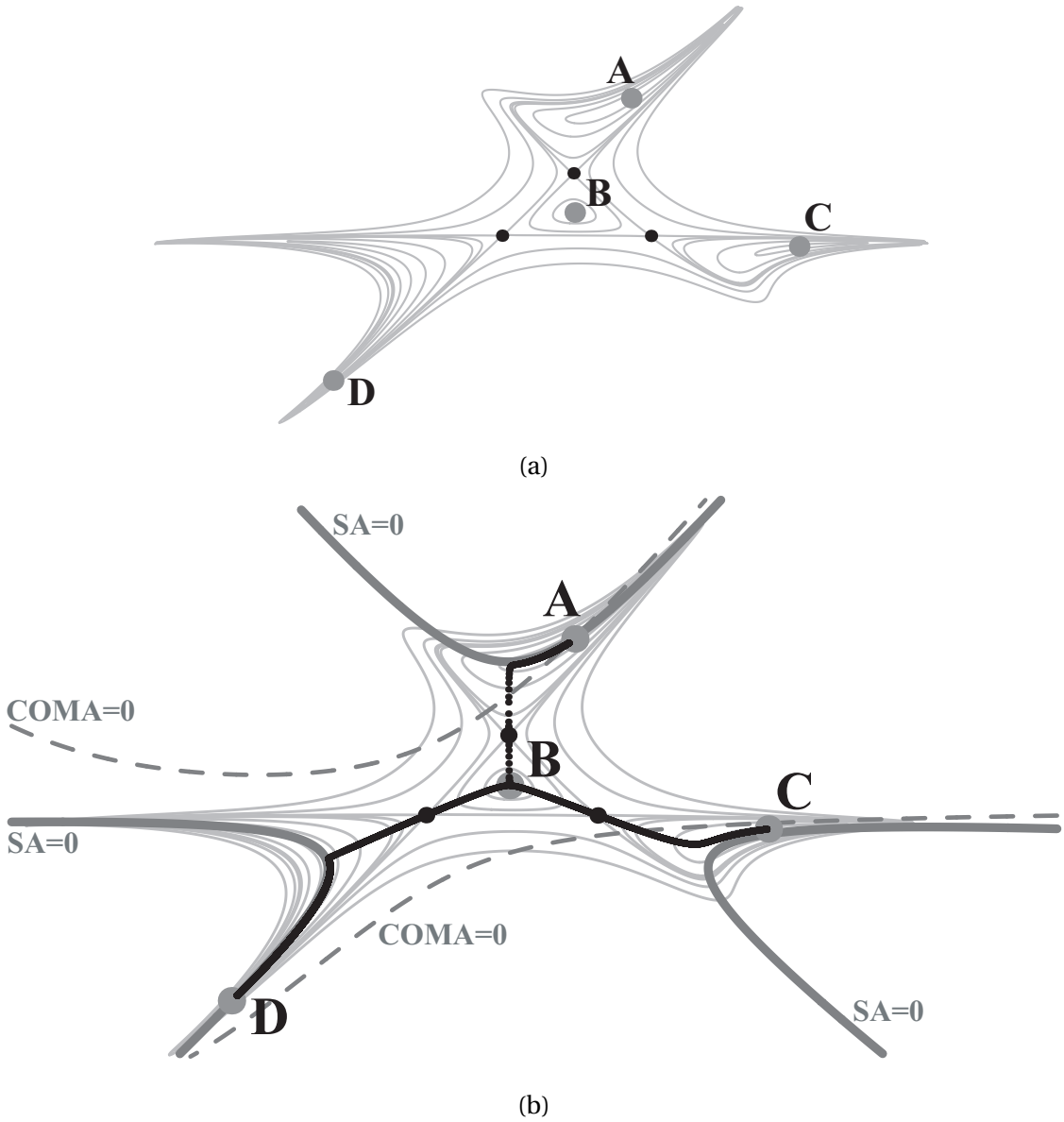


Figure 1.5: *Topography for a typical value of the first curvature of a two-dimensional monochromatic doublet landscape (f number* 3*, field of* 3 *degrees, and the lens thicknesses are small). (a) MF equimagnitude contours (light gray), four local minima (large gray points), and three saddle points (small black points). (b) Superposition of Figure 1.5(a) with the curves for zero SA (thick gray) and coma (dashed gray). The paths from the saddle points to the local minima (thick black dotted curves) are in the direction of the steepest descent.*

Choosing a good starting configuration is perhaps one of the most difficult problems in optical system design. The choice has traditionally been done on the basis of previous experience, patents or lens databases (adapted to the new specifications), intuition, and often a considerable amount of trial and error [7]. Optical system designers often find it useful to start with known configurations and increase the complexity only when necessary to solve a problem [9].

Due to the enormous increase in computer power, which can now be applied to optimization, it might appear that computers can now automatically solve all optimization problems in a reasonable amount of time. Unfortunately, this is not true. Although over the past decades, significant progress in the field of global optimization has led to development of powerful software packages, avoiding poor-quality optical systems remains very challenging, particularly if the number of optimization variables is large. Besides, inputs from human designers remain essential to the design process. For example, the designer has to supply the correct requirements and has to monitor the optimization process to ensure that the design evolves towards the envisaged goal.

1.3 Goal and outline of this thesis

It is often stated that optical system design is both a science and an art. The art of designing optics relies heavily on the experience of the designer, and contains a significant component of trial and error. This thesis is an attempt to reduce the trial and error part, making optical system design more systematic. By analyzing optical merit landscapes, we try to find characteristics that can improve optimization methods.

In optical system optimization, the correction of aberrations is usually done by using numerical optimization software. One of the major difficulties is the presence of many local minima in the merit function landscape. A good understanding of which initial configurations lead to a given local minimum, can help to solve the multiple minima problem. In order to determine such initial configurations, the so-called basin of attraction for a given minimum should be considered. The set of all starting configurations that are attracted to a local minimum is the basin of attraction for that minimum.

Let us assume for the moment that local optimization methods work without irregularities, so that the basins are compact regions in the merit function landscape. When we know all the basins of attraction, the multiple local minima problem would be solved. However, examining basins of attractions introduces other difficulties.

First, computing the basins of attracting is very time-consuming, and it becomes nearly impracticable for a large number of variables. Second, even when we have enough computational power, the visualization of high-dimensional basins is very difficult. Third, as will be shown in Chapter 5, the shape of the basins highly depend on the local optimization method. For example, the basins obtained with a steepest descent method are significantly different from the basins obtained with damped least-squares methods. Saddle points with a Morse index value of 1 are points in the merit function landscape that always remain on the basin boundary, independent of the used optimization method.

Because the high-dimensional merit function space is usually very complicated, it is unmanageable to examine properties of the optimization solution space without focusing on particular features. Past research has shown that the local minima in the optical merit function landscape form a network [10]. All local minima are linked through optimization paths generated from saddle points having a Morse index of 1. Hence, for a successful detection of all local minima, optimizing each Morse index 1 saddle point on both sides is sufficient.

The authors of Reference [10] have used Morse index 1 saddle points to develop a new type of global optimization that could be applied to systematically approach the problem of detecting multiple minima. The method detects the saddle points by a constrained local optimization. A drawback of this method is that the detection of saddle points is more difficult than detecting local minima. Research on specific characteristics of saddle points have shown that most of them can be constructed in a simple way. In stead of detecting saddle points, it is computationally much more effective to construct them. We present here a method, which we call Saddle-Point Construction (SPC), that facilitates the construction of saddle points.

This thesis shows that, when the dimensionality of the optimization problem is increased in a way that satisfies certain mathematical conditions (the existence of two independent transformations that leave the merit function unchanged), then a local minimum is transformed into a saddle point. Subsequent optimizations on both sides of the saddle point result in two different local minima.

Until now, we assumed that local optimization methods are stable methods without any irregular behavior. Unfortunately, this assumption is not always correct. We will demonstrate that, in certain situations, the frequently used damped least-squares algorithms can be a source of instabilities. In such cases, the basins of attraction have very complicated structures, which will be proved to be fractal. In a fractal region, starting points, which are very close to each other, lead to different local minima after locally optimizing them.

Fractals increase the complexity of the optimization problem, and they are a source of unpredictable behavior during the design process. When designers obtain a system shape after local optimization, which is not the expected one, it is usually believed that the cause lies in the inherent complexity of the design landscape. This is indeed a major source of unpredictability, but not the only one.

One of the goals of this thesis is to make optical system designers aware of the possibility of the presence of instabilities in the optimization process. Sensitivity to initial conditions can influence the result that will be obtained after optimization. The inherent complexity of the design landscape limits the sizes of the basins of desirable solutions. In addition, instabilities in optimization, if present, further increase complexity and decrease predictability by mixing basins of desirable and undesirable solutions together.

This thesis is organized as follows. In Chapter 2, the basic mathematical aspects of nonlinear numerical optimization are presented. We discuss several locally working algorithms, as well as some global optimization algorithms. Chapter 3 discusses a simple and efficient version of the SPC method, which is developed during this Ph.D. research. Detailed examples can be found in Appendix A. In Chapter 4, the SPC is generalized so that the restrictions, which are needed for the special version of the SPC method, are removed. We also show that a certain degree of order is present in the optical design landscape, and that this order manifest itself at different levels. The emphasis will be on understanding relationships between minima and saddle points when the dimensionality of the optimization problem is changed. In Chapter 5, we study the behavior of the damped least-squares method. We show that choosing low damping factors for the sake of increasing computational speed can create sensitivity with respect to initial conditions. In Chapter 6, we explore the use of a damped least-squares method

for a purpose that goes beyond local optimization. The results obtained with such a method enable a better understanding of peculiarities encountered with damped least-squares algorithms in conventional local optimization tasks. Finally, the conclusions on the main chapters are summarized in Chapter 7.

Chapter 2

Optical system optimization

2.1 Introduction

In the field of optical system design, the aim of optimization is to find a system configuration that fulfills certain design targets within tolerances determined by the application. The system to be designed is modeled as a point in a multi-dimensional space in which the variables are constructional parameters of the system. Typically, there are more targets than optimization variables, and only a least-squares solution can be found, where all targets are met collectively as close as possible. The merit function of the problem, which expresses the quality of an optical system into a single number, must therefore be minimized [7, 11].

We define the merit function (*MF*) of an optimization problem as the weighted sum of squares of operands that describe the design targets [12]:

$$MF(\mathbf{x}) = \sum_{j=1}^{m} w_j \left[\tilde{f}_j(\mathbf{x}) - \tilde{f}_{tar,j} \right]^2, \tag{2.1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ describes a point in the N-dimensional variable space, $\tilde{f}_j(\mathbf{x})$ are the operands with target values $\tilde{f}_{tar,j}$ and positive weighting factors w_j. *MF* is then a single number giving the difference between the various operands and their target values. Examples of operands are ray displacements (e.g. transverse ray aberrations of individual rays), aberration coefficients, lens parameters to be held within certain limits, etc. The weights and targets can be absorbed in the definitions of the operands, so that we have:

$$MF(\mathbf{x}) = \mathbf{f}^T(\mathbf{x}) \cdot \mathbf{f}(\mathbf{x}), \tag{2.2}$$

where $\mathbf{f}$ is a vector having the operands as components. The main goal of the optimization problem is to find the vector $\mathbf{x}$ that minimizes *MF* to a satisfactory level. For simplicity, in the rest of this chapter, we only consider cases without any constraints. Constraints can be added by using Lagrange multipliers or by handling them as components of the merit function.

In this research, we have mainly used the default merit function of the optical design program CODE V [13], which is based on the weighted transverse ray aberrations for a

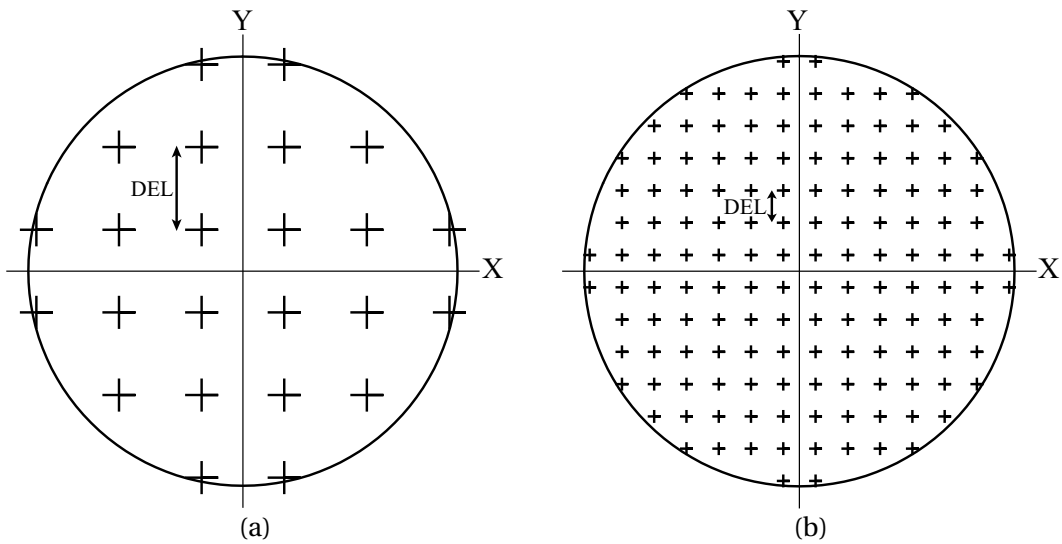


Figure 2.1: *The number of rays in the grid and the location of the grid relative to the pupil boundary is determined by the particular value of the parameter 'DEL'. (a) The default value (DEL* = 0.385*) for spherical surfaces provides a 24 ray pattern in the entrance pupil with the outermost rays in the grid lying nearly at the edge of the pupil. (b) DEL* = 0.15*, which corresponds to 140 rays.*

rectangularly spaced grid (in the entrance pupil) of rays traced at each wavelength and field. Optimization is carried out simultaneously over all fields. See Figure 1.2 for an illustration of different field specifications.

The ray grid is determined by the value of a parameter (called 'DEL' in CODE V), which defines the ray spacing in a pupil of normalized radius 1. In Figure 2.1, we show two examples of ray grids with different values for 'DEL'. For optical systems with spherical surfaces, CODE V uses 24 rays, see Figure 2.1(a). Figure 2.1(b) shows a grid of 140 rays. Note that rays near the boundary of the entrance pupil typically have large displacements and therefore have a considerable negative influence on the merit function. When a small number of rays is used over the entire entrance pupil, the optimization can become unstable, because then the rays near the boundary of the pupil have relatively more negative impact on the merit function, than in the case when more rays are used. When the operands of the merit function are transverse ray aberrations, the total number r_{tot} of rays traced through the system is given by:

$$r_{\text{tot}} = r_1 r_2 r_3, \tag{2.3}$$

where r_1 is the number of rays through the entrance pupil (the crosses in Figure 2.1), r_2 the number of fields, and r_3 the number of wavelengths.

In general, the merit function is highly nonlinear. This nonlinearity often gives rise to many local minima in the merit function landscape. Finding deep local minima, if possible the global minimum, is an essential part of the design process. In this chapter, we discuss the main aspects of optimization algorithms. In Section 2.2, we discuss local optimization methods, and in Section 2.3, we give a short overview of popular global optimization methods in optical system design.

2.2 Local optimization

Since the operands are typically nonlinear functions of $\mathbf{x}$, it cannot be assumed that only one step will solve the optimization problem. Therefore, we have to use iterative numerical optimization methods to solve the problem in several steps. When we represent a lens system having N optimization variables as a single point in a N-dimensional merit function space, then the optimization process consists of moving from one point to another, reducing the merit function after each step, until a minimum is reached. Close enough to a minimum, further iterations will not produce any significant changes in the system parameters and the process is called convergent [14].

At every iteration step, local optimization methods approximate the nonlinear optimization problem by a local quadratic problem. The solution of the previous step is used as a starting point for the next step, which updates the local approximation. In a local quadratic problem, the operands of the merit function are approximated around an initial point $\mathbf{x_0}$ according to a Taylor expansion of the second order:

$$f_j(\mathbf{x}) = f_j(\mathbf{x_0}) + \Delta\mathbf{x}^T \cdot \nabla f_j(\mathbf{x_0}) + \frac{1}{2}\Delta\mathbf{x}^T \cdot H_j \cdot \Delta\mathbf{x}, \tag{2.4}$$

where f_j is the j-th component of $\mathbf{f}$, and $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x_0}$, which describes the size and direction of the step of improvement at a certain iteration, and $H_j = H_j(\mathbf{x_0})$ is the Hessian matrix of f_j in $\mathbf{x} = \mathbf{x_0}$; its elements are:

$$\left(H_j\right)_{i,k} = \left.\frac{\partial^2 f_j}{\partial x_i \partial x_k}\right|_{x_i = x_{0,i}, x_k = x_{0,k}}.$$

The computational effort can be simplified by ignoring or approximating the Hessian matrix H_j.

A control function $\Phi(\mathbf{x})$ is used to minimize the merit function. By using Equation (2.4), and by omitting the constraints, we can write the control function for $\mathbf{x}$ close to $\mathbf{x_0}$ as [12]:

$$\Phi(\mathbf{x}) = \mathbf{f_0}^T \cdot \mathbf{f_0} + 2\Delta\mathbf{x}^T \cdot \mathbf{J}^T \cdot \mathbf{f_0} + \frac{1}{2}\Delta\mathbf{x}^T \cdot \mathbf{H} \cdot \Delta\mathbf{x} + \lambda\Delta\mathbf{x}^T \cdot \Delta\mathbf{x}, \tag{2.5}$$

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 $\mathbf{f}$ in $\mathbf{x} = \mathbf{x_0}$ (with elements: $J_{ij} = \left(\partial f_i / \partial x_j\right)\big|_{x_j = x_{j,0}}$), $\mathbf{H} = 2\left(\mathbf{J}^T \cdot \mathbf{J} + \sum_{j=1}^{m} f_j(\mathbf{x_0}) \cdot H_j\right)$, which is the Hessian of MF in $\mathbf{x} = \mathbf{x_0}$, and λ is the damping parameter. For λ sufficiently large, the quadratic term $\lambda\Delta\mathbf{x}^T \cdot \Delta\mathbf{x}$ ensures that the second order approximation in Equation (2.4) remains valid.

To find a minimum of Φ at a certain iteration, it is necessary that the gradient of Φ with respect to the optimization variables vanishes:

$$\nabla\Phi(\mathbf{x}) = 2\mathbf{J}^T \cdot \mathbf{f_0} + \mathbf{H} \cdot \Delta\mathbf{x} + 2\lambda\Delta\mathbf{x} = 0. \tag{2.6}$$

There are different methods to solve this equation. Since we have a succession of iterations, it is common practice to make some 'risky' approximations at each step, hoping that the 'error' introduced in this way is corrected by future iterations.

2.2.1 Steepest descent

The first local optimization methods were based on the steepest descent method, which only uses first-order derivatives of *MF*. When we ignore the Hessian **H**, then solving Equation (2.6) for $\Delta\mathbf{x}$ gives us:

$$\Delta\mathbf{x} = -\frac{\mathbf{J}^T \cdot \mathbf{f_0}}{\lambda} = -\frac{\nabla MF}{2\lambda}, \tag{2.7}$$

where $\nabla MF = 2\mathbf{J}^T \cdot \mathbf{f_0}$. Equation (2.7) shows that the step taken by the steepest descent method is along the negative gradient of the merit function, and can be optimized to obtain a maximum improvement in the merit function. However, this simple method has drawbacks, which causes the algorithm to oscillate, especially near a local minimum with elongated equimagnitude contours. This is illustrated by the zig-zag form of the optimization path in Figure 2.2. The zig-zag form appears because each gradient is orthogonal to the previous gradient, which makes the steepest descent method to be very slow in such cases.

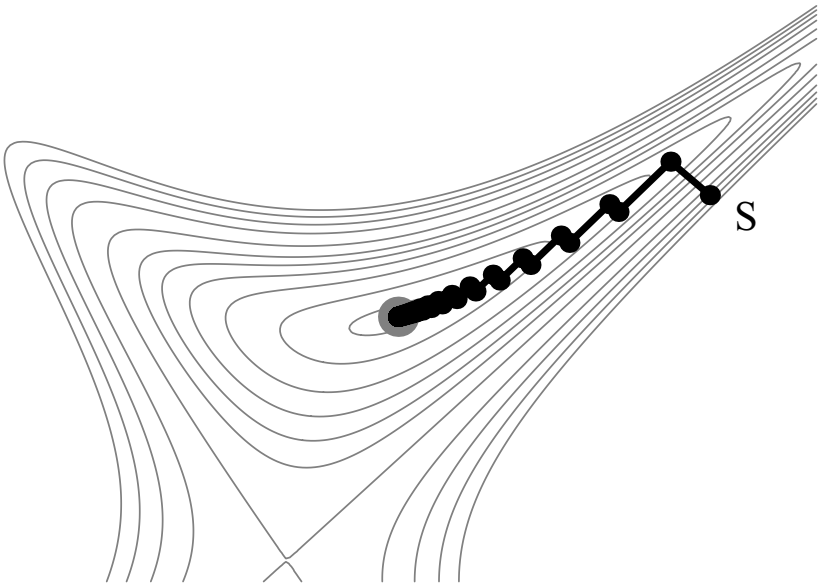


Figure 2.2: *Zig-zag path (black) of the steepest descent method, started in point S, in the neighborhood of a doublet local minimum (big gray point). The merit function (with gray equimagnitude contours) is based on a simplified thin-lens model, in which only third-order spherical aberration and third-order coma are considered.*

2.2.2 Conjugate gradient

The conjugate gradient method is much more effective than steepest descent, because it avoids the zig-zag steps by using information of the previous steps. The first iteration of this algorithm is in the direction of steepest descent, and the following iterations are along directions that are conjugate to the previous direction. At each iteration, the step size is controlled by λ, which is chosen such that a maximum improvement in *MF*

is obtained. Conjugate directions are independent of each other, which means that a minimization along one direction does not change the other directions.

In the Fletcher-Reeves method, the conjugate search directions $\mathbf{d}_i$ are based on first order derivatives of MF. The step at iteration i is given by:

$$\Delta \mathbf{x}_i = \lambda_i \mathbf{d}_i, \tag{2.8}$$

where the search directions $\mathbf{d}_i$ are given by:

$$\mathbf{d}_0 = -\nabla MF(\mathbf{x}_0), \tag{2.9}$$

$$\mathbf{d}_{i+1} = -\nabla MF(\mathbf{x}_{i+1}) + \frac{\nabla MF(\mathbf{x}_{i+1})^T \cdot \nabla MF(\mathbf{x}_{i+1})}{\nabla MF(\mathbf{x}_i)^T \cdot \nabla MF(\mathbf{x}_i)} \mathbf{d}_i, \quad \text{for } i = 0, 1, 2, 3, \ldots. \tag{2.10}$$

When we have a quadratic function of N variables, convergence to the minimum is theoretically guaranteed at the end of the minimization along the $(N-1)$-th conjugate direction (the first step is in the direction of steepest descent). However, for non-quadratic functions, it is suggested to restart the algorithm with a steepest descent after every N iterations [15].

Figure 2.3 shows the path for the same optimization problem as Figure 2.2 but in the case of the conjugate gradient algorithm restarted after every second iteration. Although a zig-zag pattern can still be observed (because we restart the algorithm after each second iteration with a steepest descent), the conjugate gradient method converges much faster (in 10 iterations) to the local minimum than the steepest descent method used in Figure 2.2 (which needed 64 iterations) [1].

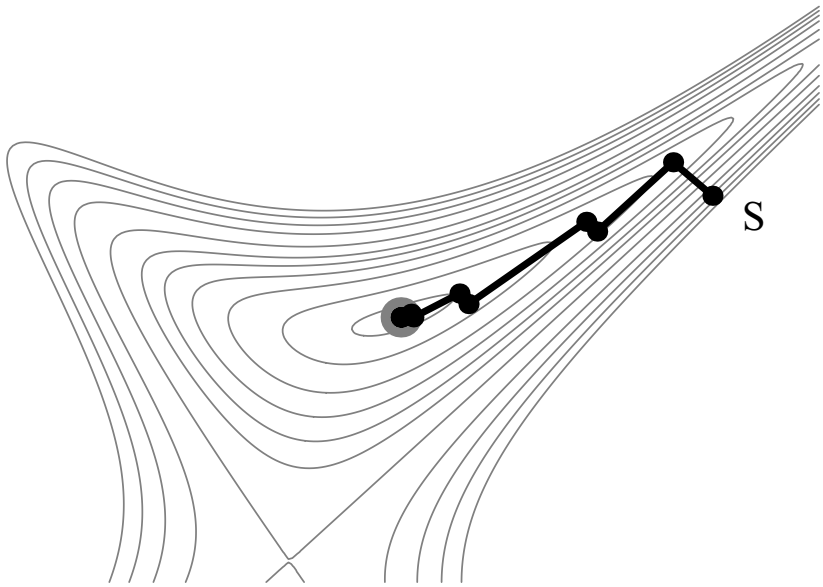


Figure 2.3: *Optimization path of the conjugated gradient method, restarted after every second iteration.*

[1] Both figures were made by using Gurdal's Mathematica tutorial *Fletcher-Reeves conjugate gradient algorithm*, which can be downloaded from Reference [16].

2.2.3 Newtonian methods

Newton's method (also known as the method of Newton-Raphson) solves Equation (2.6) (with $\lambda = 0$) for $\Delta\mathbf{x}$ without neglecting the Hessian of MF. Therefore, it describes the local merit function landscape better than the methods of steepest descent and conjugate gradient, and it converges rapidly in the neighborhood of a local minimum. Since computing the Hessian takes more computation time, it is desirable to use Newton's method with an approximation for the Hessian.

The Gauss-Newton method is such a modification of Newton's method. In the Gauss-Newton method, second-order derivatives of the operand components are excluded, and $\mathbf{H}$ is approximated by $2\mathbf{J}^T \cdot \mathbf{J}$. Without damping term, the step $\Delta\mathbf{x}$ is given by:

$$\Delta\mathbf{x} = -\left(\mathbf{J}^T \cdot \mathbf{J}\right)^{-1} \cdot \mathbf{J}^T \cdot \mathbf{f_0}. \tag{2.11}$$

An example of another Newton-like method is the Quasi-Newton method. This method does not exclude the Hessian H_j, but it uses an approximation for it. After each iteration step, the changes in the gradient of MF are used to update the approximation along the search direction.

When the operands are linear, the optimization problem is solved in only one step, which is given by Equation (2.11). However, because we typically have nonlinear problems, we approximate an optimization problem by a local quadratic problem with the hope that future iterations will correct the previous approximation.

2.2.4 Damped least-squares

The damped least-squares method is particularly suitable to solve nonlinear least-squares problems, in which the merit function is given by Equation (2.2). There are various versions of it, and for many computer programs the details are not even disclosed. In the damped least-squares method, the steps taken by using the Gauss-Newton method with $\lambda = 0$ are reduced by adding a diagonal matrix with λ as diagonal elements to the Hessian approximation (λ works as an additive damping). The step $\Delta\mathbf{x}$ is then given by:

$$\Delta\mathbf{x} = -\left(\mathbf{J}^T \cdot \mathbf{J} + \lambda\mathbf{I}\right)^{-1} \cdot \mathbf{J}^T \cdot \mathbf{f_0}. \tag{2.12}$$

Multiplicative damping can also be used, which uses, instead of $\lambda\mathbf{I}$, a diagonal matrix of individual damping constants. The damping does not only influence the optimization step, but forces the numerical algorithm to be stable as well [12].

If λ is regarded as independent variable, the angle between the changes in the variables, $\Delta\mathbf{x}$, and $-\nabla MF$ is a monotonically decreasing function of λ, with the angle going to zero when λ goes to infinity. By changing λ, the algorithm has a behavior between the Gauss-Newton method ($\lambda \to 0$) and the method of steepest descent ($\lambda \to \infty$) [17].

In damped least-squares methods, the operands are considered to be linear in a first approximation, and the damping factor λ has been originally introduced with the purpose of limiting the change of the variables to ensure that this linear approximation remains valid. Although in some situations damped least-squares methods may show

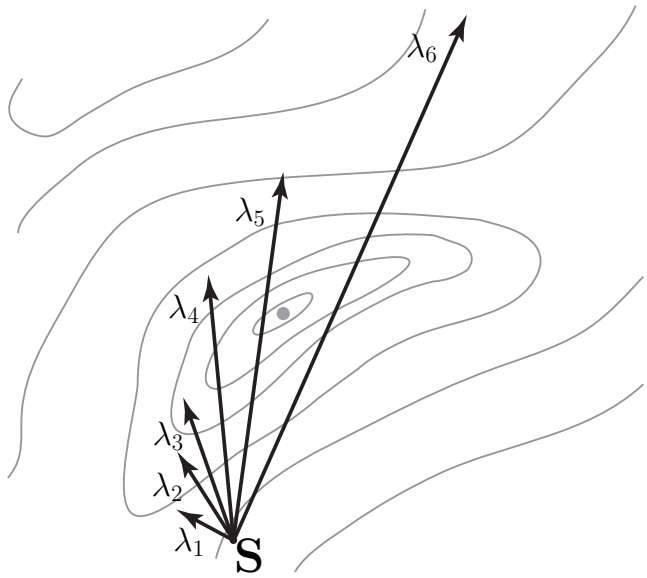


Figure 2.4: *Searches started in point 'S' for six different damping factors. The figure is based on the first figure in Reference [18].*

slow convergence, many designers know how to manipulate the damping factor to accelerate the convergence [11]. By changing the damping factor, not only the step size is changed, but also its direction. This behavior is schematically drawn in Figure 2.4 for a simple case with two variables.

Figure 2.4 shows the merit function equimagnitude contours (gray curves), a local minimum (gray point), and the starting configuration at 'S', from which six steps are taken with different damping factors. A very large damping factor in Equation (2.12) results in a small step in the direction of steepest descent (the black arrow indicated with λ_1), while a small damping factor results in a large step and generally not in the steepest descent direction. When the step is too large (as in the case of using λ_6), the linear approximation is not valid anymore and the result has a negative effect on the merit function. For $\lambda = 0$, we obtain the Gauss-Newton step given by Equation (2.11).

In general, there is no damping factor that works best in all cases. A damping factor that it is optimal in one problem, may work not so well in a different problem. Because every optical design program has his unique way of choosing the optimal damping, there is a large diversity in their local optimization algorithms [11].

The Levenberg-Marquardt damped least-squares algorithm dynamically controls the damping factor λ during optimization, and it has been used successfully in the search for local minima in the merit function landscape of optical systems [11,12]. The method uses two loops to step through the merit function landscape. At every optimization step, an outer iteration loop calculates the Jacobian to find the search direction in which the value of merit function decreases. An inner loop uses the known matrices to optimize λ so that the maximum decrease of *MF* is achieved. As will be shown in Chapter 5, such strategies for choosing λ that envisage only the largest decrease in *MF* can create sensitivity with respect to initial conditions and fractal basin boundaries.

2.3 Global optimization

The progress of optimization algorithms in optical system design is closely connected to the power of available computational tools. First, in the second half of the 1950's, designers started with using computers for local optimization methods based on the steepest descent method. Those methods were then abandoned by the 1960's after the great success of damped least-squares and conjugate gradient methods [19]. As shown in the previous section, local optimization methods use local gradients to determine the direction in which the merit function decreases. Therefore, they usually find a local minimum nearest to the starting configuration, and this optimum may not be the envisaged solution. Once the design is trapped in a local minimum, it cannot escape by further continuing the local optimization process.

Over the past decades, the available computer power increased impressively. This raised the interest in developing global optimization methods, which have the capability to overcome the merit function barriers around local minima [19–25]. For optical designs for which the complexity is not too high, present-day global optimization algorithms are valuable tools for finding a good (perhaps even the best) solution among the many local minima that are found in the merit function landscape. For moving an optical configuration from one local minimum to another, these methods rely almost exclusively on generally applicable mathematical algorithms, rather than on specific optical properties of the design landscape. However, when the number of components is growing, even local optimization becomes time consuming, and it becomes increasingly difficult to apply such tools straightforwardly.

The simplest global optimization method is a systematic evaluation of the merit function on a multidimensional grid of starting configurations [8]. The main problem of this method is that the computation time increases exponentially with the number of variables. By reducing the number of evaluation points, and optimizing these points locally to find the nearest minimum, the computation effort decreases. The challenge is to find a set of starting configurations without sacrificing the possibility of success.

Global Synthesis, which is the 'black-box' global optimization algorithm in the optical design program CODE V, has shown to be a powerful tool for finding good solutions in the merit function landscape. In some cases, radical changes in local optimization algorithms can move the solution into a region with smaller merit function values [11, 26]. Besides, there are various 'tricks' that can be used to find solutions beyond the nearest local minimum, such as making small changes to lens parameters, changing weights in the merit function, switching merit functions during local optimization, or reducing the damping factor in damped least-squares methods [11, 26, 27]. In the latter case, the algorithm overcomes a merit function barrier by running one or more iterations with reduced damping in which the merit function increases. In Chapter 6, we use low damping as an empirical strategy to escape from a poor local minimum.

The most popular methods in global optimization of optical systems are based on simulated annealing and genetic principles, usually combined with damped least-squares for further locally optimizing the found designs [12].

2.3.1 Simulated annealing

Simulated annealing is an iterative method based on the idea that a lens configuration can be regarded as being in some energy state, and the behavior of the method is controlled principally by the 'temperature' T [28]. The energy state corresponds to the merit function value of an configuration, and the goal of optimization is to minimize the energy (the optimization procedure corresponds to the thermodynamic cooling process). With simulated annealing, a random step is generated at each cycle, and depending on the merit function at the associated trial point, the step will be accepted or not [20, 28]. The step is always accepted if the merit function at the trial point is reduced. When the merit function increases by an amount of ΔMF, there is a probability equal to $e^{-\Delta MF/T}$ that the step will be accepted.

In adaptive simulated annealing methods, the acceptance probability of a step is changed during the optimization process. In the first stage, large jumps are possible to maximize the search space, but in the final stage, the probability should be changed to ensure that improved solution areas are not missed [20]. The correct choice of parameters in the algorithm is one of the major problems in simulated annealing, and this usually requires a process of trial and error.

2.3.2 Genetic algorithms

Genetic algorithms regard an optical configuration as an imaginary individual with a genetic code [22, 29]. Each design parameter is then modeled as part of this genetic code, and the merit function corresponds to the fitness for survival of the individual. Genetic algorithms are used in optimization problems with pre-specified number and arrangement of lenses [23].

A genetic algorithm typically starts with generating a random group of strings of numbers (the population), representing the values of the parameters of the optical system. Natural selection and genetic recombination processes are used to generate the next generation. Individuals with low merit function values have a better fitness and their genetic code has therefore higher influence on successive generations.

To generate the next generation, some individuals are selected depending on their fitness function value, some pairs of individuals exchange randomly chosen parts of their genetic code, and there is a low probability that some numbers in the genetic code of a few individuals are changed by mutation. Finally, the least fit individuals are removed or replaced with new random solutions (which introduce 'new blood' in the population). The resulting population is used to produce the following generation, hoping that fit 'parents' (i.e. the configurations with low merit function values) will produce 'offspring' with similar or better characteristics. During the evolution process, the population gradually fulfills the target criteria to a greater degree. After a finite number of generations, the best individual is regarded as the solution of the optimization problem.

Genetic programming, which is a branch of genetic algorithms, creates a population of computer programs as solutions of the problem, and it can start from 'scratch' (i.e. without starting from a good design, and without pre-specified number and arrange-

ment of lenses). The computer programs that represent optical lens systems are created by functions (e.g. functions to insert lens surfaces, to increase thicknesses, to set glass types, etc.) and terminals (e.g. the variables and constants), which are constrained by a structure specifying the allowable combinations of the functions and terminals. Genetic programming was used successfully as an automated invention machine that recreated (or slightly improved upon) patented lens systems [23, 30–33].

Figure 2.5 shows an example with two different functions (func1 and func2) and nine terminals (the lens parameters), which create two segmented lenses. The function func1 is a connective function, and func2 inserts a surface, distance, and fills the space to the right of the added surface with a specified glass type.

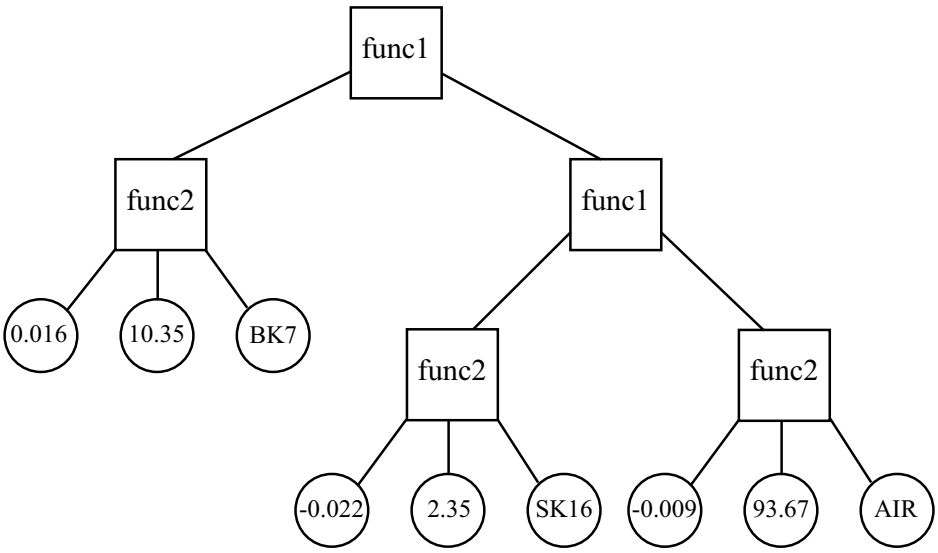


Figure 2.5: *Genetic programming. An optical system is created by functions (func1 and func2) and terminals.*

2.3.3 Global Explorer

The Global Explorer method of Isshiki [21, 34], which is implemented in the optical design program OSLO [35], uses the damped least-squares method in combination with an escape function, which is treated as an additional aberration. The escape function has a form of a Gaussian function, which raises the value of the merit function in the vicinity of a local minimum so that the design can get out of that local minimum to find another solution, see the dashed curve in Figure 2.6. The problem is to choose the settings of the algorithm such that the escape overcomes the merit function barrier.

Figure 2.6 shows a simple one-dimensional illustration of the escape function for two different settings, and their result after adding them to the merit function. The settings used for the dashed curve were sufficient to overcome the merit function barrier on the right-hand side of P1. However, when the height of the escape function is too large (the gray curve), we are trapped in P2 if we only use the merit function with escape function. We can escape from P2 by continuing the optimization with the 'old' merit

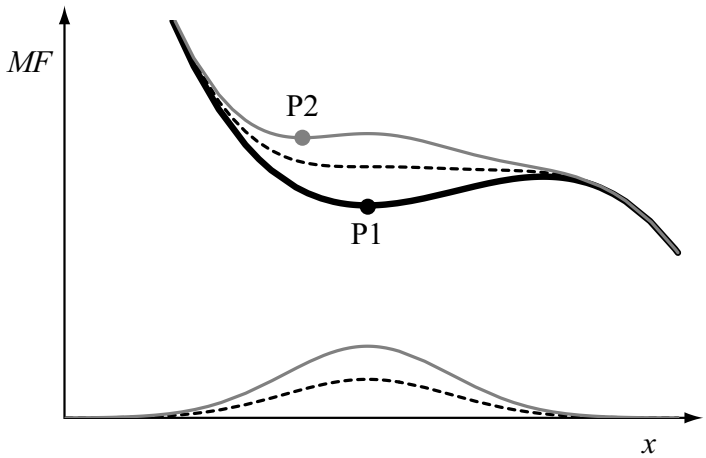


Figure 2.6: *Two different settings for the escape function (dashed black curve and continuous gray curve). The original MF without escape function is shown by the thick black curve. The two escape functions show that the correct choice of settings is essential for escaping from local minimum P1.*

function, where the escape function is not included. Either the optimization process results in a new minimum, or it returns to P1, in which case the escape function was not successful and should be modified for another attempt.

2.3.4 Detecting saddle points

The saddle-point detection method described in Reference [10] uses saddle points to travel from one basin of attraction to another. The method is based on the idea that the local minima form a network in the optical merit function space, in which they are all linked through optimization paths generated from saddle points having a Morse index of 1 (as illustrated in Figure 1.5).

In Reference [10], this network property has been used to develop a new type of global optimization that could be applied to systematically approach the problem of detecting multiple minima. The method uses constrained local optimization to detect Morse index 1 saddle points. For a relatively simple global optimization problems, the entire network of minima has been detected and visualized by two-dimensional representations, which allow a comprehensible view of the relationship between the various minima without having to deal with aspects such as the dimensionality of the optimization problem. For determining the links between two minima, two points close to the saddle point, but on opposite sides, are used to optimize the saddle point on both sides along the downward direction. Figure 2.7 shows an example of such a network. Other examples will be given in Chapter 4.

In Figure 2.7, we show the detected network of *MF* local minima and Morse index 1 saddle points for a monochromatic split doublet search. The systems have small equal distances between surfaces and all three independent curvatures are used as variables (the last surface is used to keep the effective focal length constant). By locally reoptimizing the monochromatic minima with appropriate glasses for color correction, we

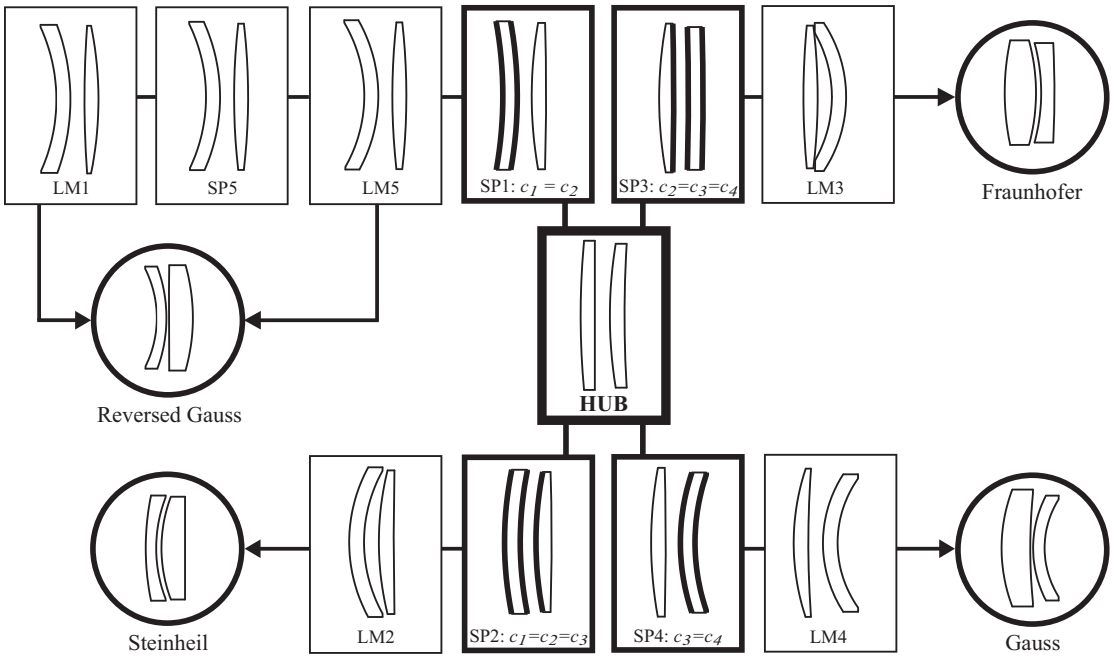


Figure 2.7: *Network of local minima (LM) and saddle points (SP) for a monochromatic split doublet search (f number* 5*, field of* 3 *degrees, and* $n = 1.5$*). The lines between rectangles show how these systems are linked in a network. For the saddle points, the surfaces with nearly equal curvatures have been drawn with thick lines. The basic achromatic doublet shapes are shown within circles. The arrows indicate which monochromatic local minima lead to them, after local reoptimization for color correction (the glasses and distances between the surfaces were adjusted).*

obtain four local minima that strongly resemble the well-known shapes of four possible solutions for split achromatic doublets [36, 37]: Fraunhofer, Steinheil, Gauss, and Reversed Gauss solution.

Note that saddle point SP5 and local minimum LM5 have almost the same shape, and the *MF* differences between them turn out to be low. In the network, the pair of systems formed by saddle point SP5 and local minimum LM5 is therefore less robust than the rest of the network. The other systems in Figure 2.7 are robust. For instance, their shape and the links between them are not affected by minor changes in the merit function definition. We will come back to this issue in Chapter 4.

The network detection method can not only reproduce the results of known global optimization algorithms, but it can also provide additional insight into the topography of the merit function landscape. A drawback of this method is that the detection of saddle points is computationally very expensive (in the order of hours for simple systems). In Chapters 3 and 4, we present a method that can be used to construct saddle points within a few seconds instead of detecting them.

2.4 Conclusions

In this chapter, we briefly discussed the main aspects of the optimization process in optical system design. Since the operands are typically nonlinear functions, we have to use iterative numerical optimization methods to locally solve the optimization problem in several steps. At each step, it is common practice to make some 'risky' approximations hoping that the 'error' introduced in this way is corrected by future iterations. However, once the design is trapped in a local minimum, it cannot escape by further continuing the local optimization process.

Global optimization methods have the capability to overcome the merit function barriers around local minima. Although global optimization methods are powerful tools to explore the design space, they tend to be very time-consuming if the dimensionality of the optimization problem is large. Besides, it is not guaranteed that they will find the global minimum.

Chapter 3

Finding new local minima in lens design landscapes by constructing saddle points

3.1 Introduction

Recent research and our studies on the network structure of the set of local minima have shown that not only local minima, but also saddle points are useful for understanding the merit function landscape of optical systems [10, 38–42]. Minima, saddle points and maxima are all critical points; i.e. the gradient of the merit function vanishes at these points. An important property of (non-degenerate) critical points is the so-called Morse index.

Intuitively, one can think about a two-dimensional saddle point (the surrounding surface has the shape of a horse saddle), which is a minimum along a certain direction and a maximum along the perpendicular direction (see Figure 1.4). Similarly, critical points in an N-dimensional optimization problem have a set of mutually orthogonal directions; along some of these directions the critical points are minima, along the other ones (called downward directions) they are maxima. The Morse index is the number of downward directions. Thus, for minima and maxima the Morse index is 0 and N, respectively, and saddle points have a Morse index between 1 and $N-1$.

When critical points merge, they are called degenerate. The determinant of the Hessian matrix $\mathbf{H}$ of the merit function is then zero. Assuming that the critical point is non-degenerate, the value of the Morse index (*MI*) is given by the number of negative eigenvalues of $\mathbf{H}$. A negative eigenvalue means that along the direction defined by the corresponding eigenvector of $\mathbf{H}$ the critical point is a maximum. As $\mathbf{H}$ is a square symmetric matrix, all eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ are real and the eigenvectors corresponding to different eigenvalues are mutually orthogonal. The determinant $\det\mathbf{H}$ is then given by:

$$\det\mathbf{H} = \prod_{j=1}^{N} \lambda_j. \tag{3.1}$$

For non-degenerated points, all eigenvalues of **H** are nonzero.

For optimization problems, saddle points with Morse index 1 are of special interest [10]. They are maxima in one direction, which one can visualize as the downward direction of a two-dimensional saddle point, and they are minima in the remaining $N-1$ directions, which are all very similar to the upward direction in the two-dimensional case. As in a two-dimensional situation, choosing two points close to each other, but on opposite sides of the saddle and starting local optimizations at those points, will lead to two distinct minima. An illustration will be given in Figure 3.3.

As has been shown in References [38, 42], if a local minimum is known, new local minima can be found by detecting Morse index 1 saddle points in the vicinity of the known minimum, and then by optimizing the configurations on the other side of these saddle points. A drawback of this method is that detecting Morse index 1 saddle points without a-priori information about them is computationally more expensive than finding local minima.

As a computationally effective alternative to saddle point detection, we present a new method, which we call Saddle-Point Construction (SPC). In this chapter, we study a simple and efficient version of the SPC method. This method can be used with success even in the case of very complex systems with many variables and constraints, because it can lead to new system shapes with only a small number of local optimizations [43, 44]. In Chapter 4, we discuss the general case where the restrictions discussed in Section 3.4 are removed.

With the (special or general) SPC method, saddle points are created by inserting a lens into an existing optical configuration, which is already a minimum in its variable space. Lens designers frequently insert lenses into their designs and, in the traditional way, one new system shape results after optimization. However, when a lens is inserted with SPC, two distinct system shapes result and for further design one can choose the better one. By inserting lenses according to the SPC method, and then, if necessary, by extracting lenses, new local minima can be obtained effectively for optical design tasks of arbitrary complexity.

Saddle points already exist in simple systems consisting of thin spherical lenses in contact, where the only variables are curvatures. In Section 3.2 we study these systems with emphasis on a remarkable property that survives generalization. As will be shown in Section 3.4 and Chapter 4, this property can be used to generate saddle points for arbitrary optical systems. Since we use in this chapter mathematical ideas which are rather new in optical system design (such as that of a saddle point in a high-dimensional variable space), we will first give an intuitive description of the SPC method in Section 3.3. In Section 3.4, we will then discuss a simple and efficient version of the SPC method rigorously, and we will prove that a system constructed according to our receipt is a saddle point with Morse index 1. Three simple examples that illustrate the essence of the SPC method in Section 3.5 show that the practical implementation of the method is actually very easy. In Section 3.6, additional mathematical properties of SPC will be investigated. In Chapter 4, we present the generalized version of the SPC method.

3.2 Systems of thin lenses in contact

Systems of thin spherical lenses, in which all lens thicknesses and air spaces between lenses are set equal to zero, and for which the merit function (*MF*) includes only third-order spherical aberration (SA), are simple enough to be studied analytically in detail. This model, which relies on third-order aberration theory, may seem oversimplified. However, as will be seen below, it enables us to understand a mechanism that is also present in systems having a much higher complexity.

If in a power series expansion for the transverse aberration of a given ray we keep only the SA contribution and neglect the rest, then for all rays the transverse aberration will be proportional with SA. When the *MF* is a weighted sum of squares for the transverse aberrations of a given set of rays, it will be proportional with SA^2. In our simplified model, Morse index 1 saddle points are solutions of the system of equations $\nabla\mathrm{SA} = 0$ (∇ has as components the partial derivatives with respect to the variables).

When the imaging is monochromatic, all glasses have refractive index n, and the object is at infinity, we can write the SA of a doublet as [39]:

$$\mathrm{SA} = h^4\left(f_0 + \frac{n-1}{n} f_1 f_2 f_3\right), \tag{3.2}$$

where the functions f_0, f_1, f_2 and f_3 are given by:

$$f_0 = \left(\frac{n}{n-1}\right)^2 - \frac{2n+1}{n-1}c_1 + \frac{n+2}{n}c_1^2, \tag{3.3}$$

$$f_1 = c_1 - c_2 - \frac{1}{n-1}, \tag{3.4}$$

$$f_2 = c_2 - c_3, \tag{3.5}$$

$$f_3 = -\frac{n(2n+1)}{n-1} + (-2n^2+n+4)c_1 + 2(n^2-1)c_2 + (n+2)c_3. \tag{3.6}$$

In these formulas, c_1, c_2, and c_3 are the curvatures of the first, second, and third surface, respectively. The fourth curvature, which does not appear explicitly, is used to keep the total power equal to 1; h is the marginal ray height at the lens. Since the aperture stop is placed at the lens, h is equal to half the stop diameter.

When we use the first curvature c_1 as a control parameter, the condition $\nabla\mathrm{SA} = 0$ [$\nabla = (\partial/\partial c_2, \partial/\partial c_3)$] has four solutions:

$$\left(f_1 = 0,\ f_2 = 0\right), \tag{3.7}$$

$$\left(f_1 = 0,\ f_3 = 0\right), \tag{3.8}$$

$$\left(f_2 = 0,\ f_3 = 0\right), \tag{3.9}$$

$$f_3 = (n+2)f_2 = n(2n+1)f_1, \tag{3.10}$$

of which Equations (3.7)–(3.9) are saddle points [39]. These three saddle points are located at the three intersections of the straight lines $f_1 = 0$, $f_2 = 0$ and $f_3 = 0$, taken in pairs of two. Along the lines, SA remains unchanged [see Equation (3.2)]. One of the two-dimensional saddle points has thus three equal curvatures, $c_2 = c_3 = c_4$ ($f_1 = f_2 = 0$),

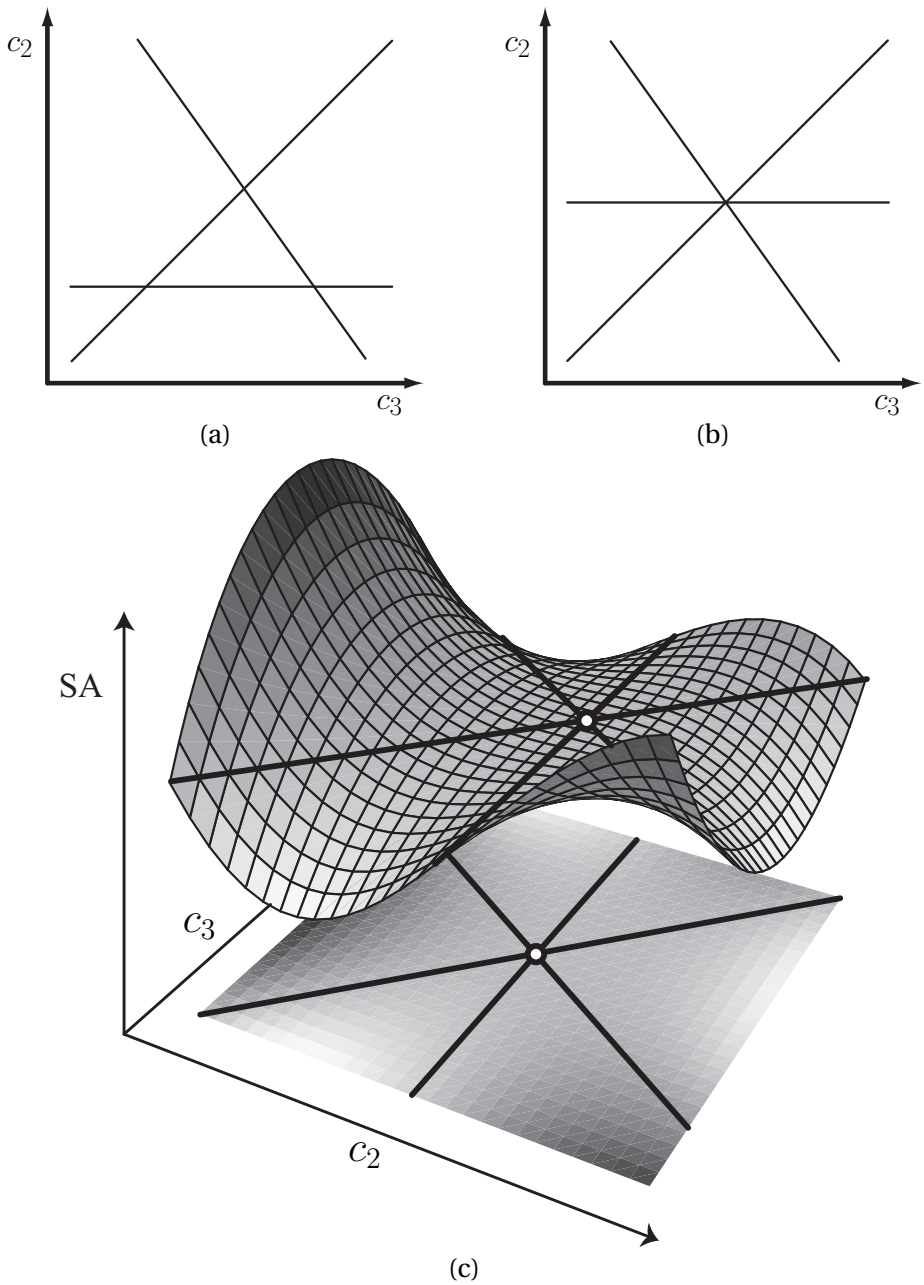


Figure 3.1: *SA equimagnitude lines that pass through the saddle points for a thin-lens doublet with* $n = 3/2$, $h = 1/20$, *and total power equal to* 1. *(a)* $c_1 = 0$, *(b)* $c_1 = c_{1,crit} = 24/7$, *(c) shape of the monkey saddle.*

and the other two have two equal curvatures, $c_2 = c_3$ ($f_2 = 0$) and $c_3 = c_4$ ($f_1 = 0$), respectively.

Using the first curvature in the system as a control parameter (and not as a variable) is useful for studying the relationship between the critical points. Figure 3.1(a) shows the SA equimagnitude lines $f_1 = 0$ (horizontal line, corresponding to $c_3 = c_4$), $f_2 = 0$ (left oblique line, corresponding to $c_2 = c_3$), and $f_3 = 0$ (right oblique line) when we use $c_1 = 0$. In the middle of the triangle formed by the three saddle points, we find a fourth solution with $\nabla \mathrm{SA} = 0$. For values of the control parameter c_1 that are not too large, this

point (which we call the 'hub' of the network) is a minimum. Despite the fact that merit function of the hub is poor, this minimum plays an essential role for the connectivity of the network. Our analytic results show that as long as the first curvature is lower than a certain critical value, all saddle points are linked on one side with the same hub. In the rest of this thesis, we refer to local minima that have more than two links as 'hubs'.

When c_1 increases, the distances between the saddle points decrease, until c_1 reaches a critical value, where the hub, the surrounding Morse index 1 saddle points and, when they exist, also saddle points with a higher Morse indices, merge into a single degenerate critical point. Within the frame of the simplified model, the critical value is given by:

$$c_{1,\text{crit}} = \frac{n(2n+1)}{(n-1)(n+2)}. \tag{3.11}$$

The SA landscape around such a point takes then the peculiar shape sometimes called a 'monkey saddle' [45]. For doublets with zero thickness and a slightly shifted critical value of c_1, the monkey saddle can also be observed using a merit function where aberrations other than SA play a role [39].

For a thin-lens doublet with $n = 1.5$, the critical point is for instance given by $c_{1,\text{crit}} = 24/7 \approx 3.43$. When $c_1 = c_{1,\text{crit}}$ the three equimagnitude lines pass through the same point [Figure 3.1(b)], and the SA landscape around this point has the shape of a monkey saddle [Figure 3.1(c)]. Remarkably, close enough to the crossing points of the lines $f_1 = 0$, $f_2 = 0$, and $f_3 = 0$, the SA behavior in Figure 3.1(a) is non-degenerate. However, the behavior of SA tends to become degenerate far away from the crossing points. The surface has then the shape of a monkey saddle. We will give another example of degeneracy in Section 3.6.

It is important to note that $f_1 = 0$ and $f_2 = 0$ remain equimagnitude lines even if other aberrations, including higher order ones, are included. This property is an essential ingredient of the SPC method, which will be explained in detail in Section 3.4.

If all three independent curvatures of a thin-lens doublet are used as variables, $\nabla\text{SA} = 0$ has five solutions: four SA Morse index 1 saddle points and a minimum (the hub). With the first curvature now variable, the curvatures of the four saddle points follow the same pattern. For two of them three successive curvatures are equal and for the other two saddle points two curvatures are equal (the first two curvatures and the last two curvatures, respectively), see Table 3.1.

Table 3.1: *The curvatures of the four SA Morse index 1 saddle points of a thin-lens doublet with $n = 3/2$, and object at infinity.*

	c_1	c_2	c_3	c_4
SP1	$-12/7$	$-12/7$	$12/7$	$-2/7$
SP2	$12/7$	$12/7$	$12/7$	$-2/7$
SP3	$12/7$	$-2/7$	$-2/7$	$-2/7$
SP4	$12/7$	$-2/7$	$22/7$	$22/7$

The network linking of the three-dimensional thin-lens doublet is similar to the linking shown in Figure 2.7. In fact, the existence of the robust systems in Figure 2.7 can be ex-

plained with our simplified merit function model. This simplified model predicts that, on one side, all the robust saddle points are linked with the hub. On the other side, the robust saddle points are linked with local minima that begin to resemble the familiar achromatic doublet shapes.

Table 3.2 reproduces the exact values of the curvatures of four saddle points of a monochromatic triplet with refractive index $n = 1.5$, and object at infinity. The first five curvatures were used as independent variables, the sixth one was used to keep the focal length unity.

Table 3.2: *The curvatures of four Morse index 1 saddle points of a thin-lens triplet with $n = 3/2$, object at infinity.*

c_1	c_2	c_3	c_4	c_5	c_6
6/7	6/7	6/7	−1/7	11/7	4/7
6/7	−1/7	−1/7	−1/7	11/7	4/7
6/7	−1/7	11/7	11/7	11/7	4/7
6/7	−1/7	11/7	4/7	4/7	4/7

With the same simplified model, the values of the curvatures of the hubs can also be computed analytically. The monochromatic thin-lens doublet hub has for instance:

$$c_1 = 6/7, \quad c_2 = -1/7, \quad c_3 = 11/7, \quad c_4 = 4/7. \tag{3.12}$$

Note that the curvatures of the triplet saddle points in Table 3.2 can be very simply derived from the parameters of the doublet local minimum given by Equation (3.12). The triplet curvature values are those given by Equation (3.12), but for each saddle point, one of these curvatures appears three times successively. Physically, this numerical property can be interpreted as follows: the four triplet saddle points are obtained by inserting successively one meniscus lens at the four surfaces of the doublet hub. In all cases, the meniscus lens has equal curvatures, and their common value is equal to the curvature of the original doublet surface where the meniscus is introduced. It therefore results from general formulas for systems of thin lenses in contact that certain saddle points with $N+2$ surfaces can be obtained in the same way from hubs with N surfaces. This happens when all curvatures are used as variables, and also when some of them are used as control parameters. Later on, we refer to those saddle points as 'null-element saddle points' (NESP's).

Degenerate critical points such as the monkey saddle shown in Figures 3.1(b) and (c) exist in higher dimensions as well. In the case of a triplet with four variable curvatures and the first curvature used as control parameter, the higher-order equivalent of the monkey saddle is a critical point that is 16-fold degenerate. When the first curvature is below the critical value, this degenerate critical point splits into a local minimum (a hub), five NESP's and ten saddle points with Morse index 2.

For the monochromatic doublet, triplet and quartet, Fulcher has derived the curvatures of the minimum, which we call hub, several decades ago [46]. These systems turn out to be relaxed designs [47, 48]. While the hub doublet has a poor imaging quality, the Fulcher monochromatic quartet with $n = 1.5$ is a remarkable system. By optimizing the

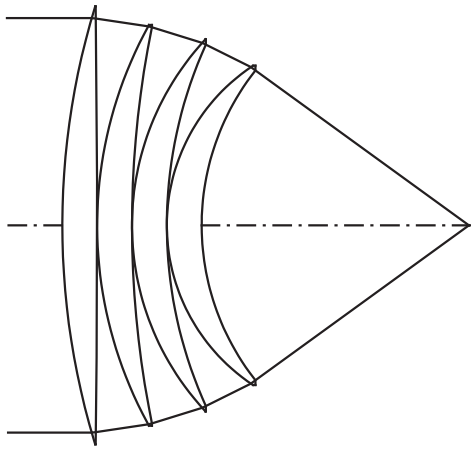


Figure 3.2: *A Fulcher-type design.*

thin-lens Fulcher specifications and adding adequate lens thicknesses, we obtain for axial imaging at a numerical aperture of 0.6 a Strehl Ratio of 0.999. Note in Figure 3.2 the smooth bending of the ray paths and the similarity with elements in designs of lithographic objectives. In fact, hubs also exist for high-quality lithographic designs [43]. It is worth investigating in detail whether, as it seems, there is a more general correlation between the hub structure and relaxation.

3.3 SPC in a nutshell

In this section, we give an intuitive description of the SPC method. This method can be used with a broad class of optical merit functions. For example, with a merit function based on transverse aberrations (root-mean-square spot size), wavefront aberration, aberration coefficients, etc. For optimizations that must include properties that are not frequently included as optimization operands, the applicability of SPC must be examined separately by using the same reasoning as shown in Section 3.4. If the mathematical conditions for SPC are not satisfied, such unusual operands can be omitted in the first stage. After the two local minima on both sides of the saddle point are obtained, those properties can be included in the merit function again.

Figure 3.3 gives an intuitive illustration of the SPC method. We start with an optimized system with N surfaces. For clarity, only one variable is shown in dark gray in the upper left part of the figure. The original local minimum can have any number and type of variables (e.g. thicknesses, curvatures, etc.). In this system, we insert a meniscus lens with zero axial thickness and equal variable curvatures. Such a meniscus disappears physically and does not affect the path of any ray or the merit function of the system. Therefore, we call this meniscus a 'null-element'. For simplicity, we discuss here the case when the element to be inserted is a lens with spherical surfaces, but the method also works with mirrors and aspherical surfaces [43, 44]. Although, when inserted, the null-element does physically nothing, it comes with two new variable curvatures, which, when changed during optimization, allow the merit function to decrease.

For some specific values for these two curvatures (as will be explained in Section 3.4),

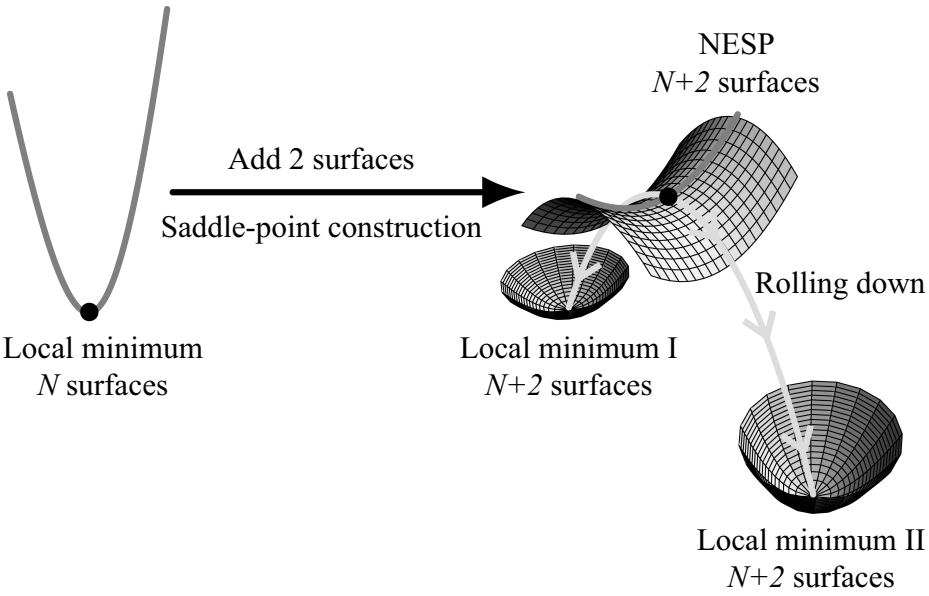


Figure 3.3: *Obtaining two local minima with $N+2$ surfaces from a local minimum with N surfaces via a NESP.*

the null-element transforms the original local minimum into a saddle point with Morse index 1 in the variable space with increased dimensionality. Because such a saddle point contains a null-element lens, we call this saddle point a 'null-element saddle point' (NESP). The new 'downward' direction, shown in Figure 3.3 by the lighter gray curve passing through the saddle point, and a new 'upward' direction (not shown, but similar to the dark gray curve through the saddle point) appear in the new variable space (with a dimension increased by 2). Along the downward direction, the new system is a maximum. Since, as will be shown in the next section, the Morse index is 1, despite of the fact that typically we have much more than two variables, the surrounding surface of the saddle point resembles very much a two-dimensional horse saddle. If we choose two points close to the saddle, but situated on opposite sides, and then optimize them, the optimizations 'roll down' from the saddle and arrive at two distinct local minima. The optimization variables are those of the starting local minimum plus the two curvatures of the null-element.

In the following two sections, we discuss SPC in the special case when the inserted null-element is in direct contact with one of the surfaces of the original local minimum (the reference surface) and when the glass of the new meniscus is the same as that at the reference surface. If in the final design the glass of the new lens must be different and/or the lens must be placed at a certain distance from a given surface of an existing minimum, then those adjustments can be done once the two minima on both sides of the NESP have been obtained. In Chapter 4, we present the generalized version of the SPC method, where the glass and distance restrictions mentioned above are removed, and the curvatures of the null-element meniscus to be inserted can be computed numerically.

3.4 SPC: theory

In the special case, a certain restriction on glass and axial thicknesses is needed. The null-element is inserted in contact with (i.e. at zero axial distance from) an existing surface (called reference surface) in the original local minimum. This reference surface should have a variable curvature, and the glass of the null-element should be the same as the glass of the reference surface. The conditions for insertion position and glass type are not as restrictive as they might seem, because once the two minima on both sides of the NESP have been obtained, the distances between surfaces and the glass of the lens resulting from the null-element can be changed as desired. In this section, we show that when the curvatures of the null-element are equal to the curvature of the reference surface, we obtain a NESP.

Assume that the reference surface is the k-th surface of an optimized N-dimensional system [Figure 3.4(a)]. The value of the merit function is MF_{ref} and the curvature of the surface in the starting minimum is c_{ref}. After the surface, we introduce a null-element in contact with the k-th surface (i.e. we have two consecutive zero axial thicknesses). When the curvatures c_{k+1} and c_{k+2} of the null-element lens are varied, but are kept equal, then the inserted meniscus remains a null-element, and the merit function of the new system with $N+2$ surfaces remains equal to MF_{ref} [Figure 3.4(b)].

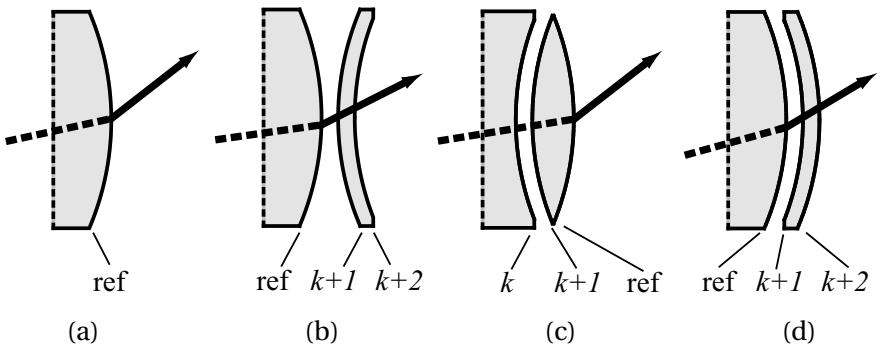


Figure 3.4: *(a) The k-th surface with curvature c_{ref} of a system that is a local minimum. The path of any ray is left unchanged by (b) a null-element meniscus with equal curvatures $c_{k+1} = c_{k+2}$ added after surface k, or (c) a thin air meniscus with equal curvatures $c_k = c_{k+1}$ before the surface with curvature $c_{k+2} = c_{ref}$. (d) In the special case that all three successive curvatures are equal to c_{ref}, and the null-element lens has the same glass as the k-th surface, we have a saddle point. For clarity, the thin lens and the air space before it are drawn with nonzero thicknesses in all figures.*

Similarly, if the curvatures c_k and c_{k+1} are varied, but are kept equal, the air space before the lens becomes a null-element. When, in addition, we have $c_{k+2} = c_{\text{ref}}$, the merit function of the new system remains again unchanged by the insertion and equal to MF_{ref} [Figure 3.4(c)]. Note that, although the curvature c_k of the original minimum with N surfaces is varied, the second curvature of the new lens takes its role and ensures that MF_{ref} remains unchanged.

When $c_k = c_{k+1} = c_{k+2} = c_{\text{ref}}$, we are in the special situation shown in Figure 3.4(d). To

examine this case, consider the following transformations:

$$c_k = c_{\text{ref}}, \quad c_{k+1} = c_{k+2} = u, \tag{3.13}$$

$$c_k = c_{k+1} = v, \quad c_{k+2} = c_{\text{ref}}, \tag{3.14}$$

where in both transformations all other variables of the original minimum are kept unchanged. The transformations given by Equations (3.13) and (3.14) describe two lines in the variable space of the new system with $N+2$ surfaces. The position of individual points along these lines is given by the parameters u and v. The two lines intersect for $u = v = c_{\text{ref}}$, which gives us:

$$c_k = c_{k+1} = c_{k+2} = c_{\text{ref}}, \tag{3.15}$$

where all other variables have the values of the original minimum. As shown above, along both lines the merit function is invariant and equal to MF_{ref}. Therefore, the system with the property described by Equation (3.15) has also a merit function value equal to MF_{ref}.

In the above analysis, we assumed that a thin lens with surfaces $k+1$ and $k+2$ is inserted after the k-th surface in an existing design that is a local minimum. However, Equations (3.13)–(3.15) are also valid if a thin lens with surfaces k and $k+1$ is placed before the $k+2$-th surface in an existing minimum. In this case, the invariant lines given by Equations (3.13) and (3.14) are related to the 'null-airspace' lens and the 'null-glass' lens that are formed, respectively.

In optical optimization problems, paraxial equality constraints are frequently used. For instance, the effective focal length is kept constant. Below, we analyze the properties of the intersection point given by Equation (3.15) in the variable space of the new system with $N+2$ surfaces. We show that this crossing point is a saddle point with Morse index 1 or 2 if no constraints are used, and with Morse index 1 when the same paraxial constraint is used both in the existing minimum, and during SPC.

For examining the crossing point given by Equation (3.15) in the $N+2$-dimensional merit function landscape, it is sufficient to investigate a three-dimensional subspace of this landscape, defined by the variables c_k, c_{k+1} and c_{k+2} (see Figure 3.5). The coordinate system in Figure 3.5(a) is given by:

$$x = c_k - c_{\text{ref}},\ y = c_{k+1} - c_{\text{ref}},\ \text{and}\ z = c_{k+2} - c_{\text{ref}}. \tag{3.16}$$

In the unit cube shown in Figure 3.5(a), the invariant lines given by Equations (3.13) and (3.14) are oriented along the vectors:

$$\mathbf{OA} = (1,1,0), \quad \text{and} \quad \mathbf{OB} = (0,1,1), \tag{3.17}$$

respectively. It is convenient to rotate the coordinate system so that the points in the plane OAB can be parameterized by only two numbers (instead of three). Since $OA = OB = AB$, the angle between **OA** and **OB** is 60 degrees, and the two lines cannot be used as axes in a rectangular coordinate system. However, an orthogonal axis system in the plane OAB can be easily constructed. The new axes x' and y' are then oriented along

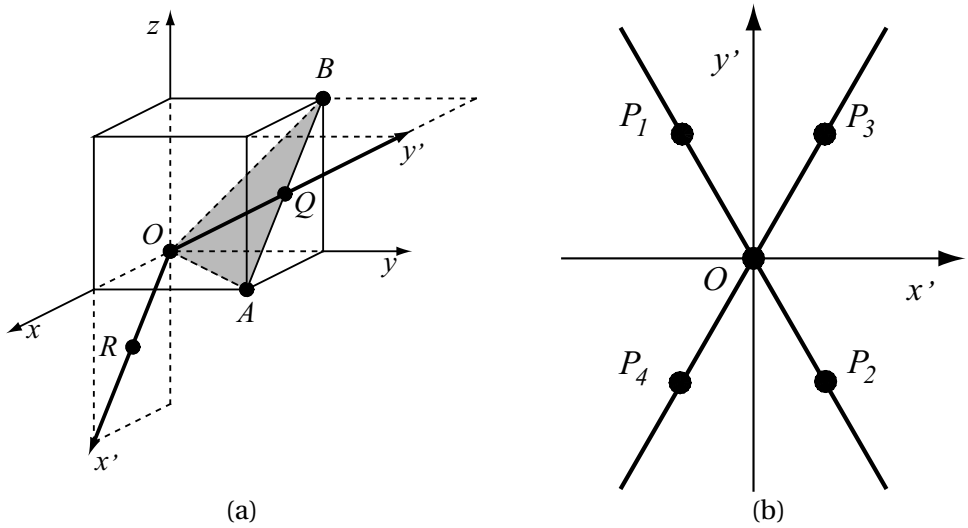


Figure 3.5: *(a) Plane OAB (gray) defined by the invariant lines given by Equations (3.13) and (3.14) in the subspace defined by the variables c_k, c_{k+1}, and c_{k+2}. Note that point R also belongs to this plane. (b) The two invariant lines (continuous lines) and the orthogonal axes x' and y' in the same plane. The two lines cross at the saddle point O, the origin of the coordinate system (x', y').*

the vectors:

$$\mathbf{OR} = \frac{1}{2}(\mathbf{OA} - \mathbf{OB}) = \frac{1}{2}(1, 0, -1), \tag{3.18}$$

and

$$\mathbf{OQ} = \frac{1}{2}(\mathbf{OA} + \mathbf{OB}) = \frac{1}{2}(1, 2, 1), \tag{3.19}$$

see Figure 3.5(a). By computing the vector product $\mathbf{OR} \times \mathbf{OQ}$, it can be seen that the axis z' orthogonal to the plane OAB (not shown) is then oriented along the vector $\mathbf{v} = (1, -1, 1)$. By taking unit-length vectors along OQ and OR, the position of an arbitrary point in the plane OAB is given by:

$$x' \frac{2}{\sqrt{2}}\mathbf{OR} + y' \frac{2}{\sqrt{6}}\mathbf{OQ} = \left(\frac{x'}{\sqrt{2}} + \frac{y'}{\sqrt{6}}, \sqrt{\frac{2}{3}}y', -\frac{x'}{\sqrt{2}} + \frac{y'}{\sqrt{6}} \right). \tag{3.20}$$

The three curvatures for the points in the plane OAB are then given by:

$$c_k = c_{\text{ref}} + \frac{x'}{\sqrt{2}} + \frac{y'}{\sqrt{6}}, \tag{3.21}$$

$$c_{k+1} = c_{\text{ref}} + \sqrt{\frac{2}{3}}y', \tag{3.22}$$

$$c_{k+2} = c_{\text{ref}} - \frac{x'}{\sqrt{2}} + \frac{y'}{\sqrt{6}}. \tag{3.23}$$

If in a plane two lines, along which a function is constant, cross, then the crossing point is a two-dimensional saddle point of that function. In the plane OAB, the two invariant

lines given by Equations (3.13) and (3.14) intersect in O [Figure 3.5(b)]. We consider the case when the null-element lens with curvatures c_{k+1} and c_{k+2} is placed in contact with the reference surface k of the N-dimensional original local minimum, in which $c_k = c_{\text{ref}}$.

We start by showing that O is critical point in the plane OAB. It is well known that the direction of the most rapid variation of a function is orthogonal to the direction along which the function is constant. Therefore, at each point along the invariant lines, the projection on OAB of the gradient of the merit function is orthogonal to the invariant lines. At the intersection point, the projection of the gradient must be zero, because it cannot point in two different directions.

Assuming that O is non-degenerate, the point O is then a two-dimensional saddle point in the plane OAB. Along one of the orthogonal axes in the plane OAB (x' or y'), the point O is a minimum, along the other one it is a maximum. For a maximum or a minimum, the equimagnitude contours for merit function (MF) values close to MF_{ref} are ellipses, which reduce to a point for $MF = MF_{\text{ref}}$. The point O cannot be a two-dimensional maximum or minimum, because here the equimagnitude contours are crossing straight lines. We will come back to the assumption of non-degeneracy in Section 3.6.

Because for the original local minimum the merit function derivative with respect to c_k was already zero, the point O is also a minimum along the direction of Ox, outside OAB [1]. In the new coordinate system with axes x', y' and z', the direction of Ox is orientated along the vector $\mathbf{r}' = \left(r'_x, r'_y, r'_z\right) = (3,1,2)$. Since O is a minimum along the direction of Ox, the derivative of MF with respect to x', y' and z' is zero along the direction of $\mathbf{r}'$:

$$\nabla_{\mathbf{r}'} MF = \frac{\partial MF}{\partial x'} r'_x + \frac{\partial MF}{\partial y'} r'_y + \frac{\partial MF}{\partial z'} r'_z = 0. \tag{3.24}$$

In the plane OAB, the partial derivatives $\frac{\partial MF}{\partial x'}$ and $\frac{\partial MF}{\partial y'}$ at O vanish. Because $r'_z \neq 0$, the partial derivative $\frac{\partial MF}{\partial z'}$ should be equal to zero as well to satisfy Equation (3.24).

Since the partial derivatives with respect to x', y', and z' are zero, the merit function derivatives with respect to y and z must both be zero at O. In addition, because the variables of the original local minimum other than c_k are kept unchanged, the merit function derivatives with respect to them remain zero as well. Therefore, all components of the gradient of the merit function vanish at O, making O a critical point.

We have shown that O is a maximum in one direction in the plane OAB, that it is a minimum in the orthogonal direction in that plane and that it is also a minimum with respect to the variables of the original local minimum other than c_k. The only direction that remains to be studied is that of Oz'. If there are no equality constraints, along this direction O can be a minimum or a maximum, and the Morse index is then 1 or 2, respectively [2].

[1] The above analysis is also valid if a thin lens with surfaces k and $k+1$ is placed before the $k+2$-th surface in an existing minimum. Then, the point O is a minimum along the direction of Oz.

[2] In case there are no constraints, it can be investigated numerically whether O is a minimum along the direction Oz'. Otherwise, the Morse index will be 2 and the procedure for generating local minima from the saddle point must be adapted, but we have not encountered such situations yet.

In the following analysis, we demonstrate that when we use paraxial constraints, Oz' cannot be an additional upward direction and the Morse index remains 1, because these constraints are violated along the direction of Oz'. We show that if the original minimum satisfies a paraxial constraint, then all points in the plane OAB satisfy this constraint, because for these points paraxial ray paths at surfaces other than k, $k+1$ and $k+2$ are left unchanged.

By using the relations given by Equation (3.16), we can write the total power K_{tot} of surfaces k, $k+1$ and $k+2$ that are changed in Figure 3.4 as:

$$\begin{aligned} K_{\text{tot}} &= K_k + K_{k+1} + K_{k+2} \\ &= (n-1)\left(-c_k + c_{k+1} - c_{k+2}\right) \\ &= (n-1)\left(-x + y - z - c_{\text{ref}}\right) \\ &= (1-n)\left(c_{\text{ref}} + x v_x + y v_y + z v_z\right), \end{aligned} \tag{3.25}$$

where n is the refractive index of the null-element meniscus and of the lens with which it is in contact, and $\mathbf{v} = (1,-1,1)$ is the normal vector to the plane OAB. Note however that

$$x v_x + y v_y + z v_z = 0 \tag{3.26}$$

is the equation for the plane passing through the origin, perpendicular to Oz' (which is oriented along the vector $\mathbf{v}$), i.e. for the plane OAB. The total power K_{tot} of surfaces k, $k+1$ and $k+2$ remains constant in the plane OAB, and is equal to the power $(1-n)c_{\text{ref}}$ of surface k of the original local minimum (in which $c_k = c_{\text{ref}}$).

Because of zero axial thicknesses between surfaces k, $k+1$ and $k+2$, the ray heights at surfaces $k+1$ and $k+2$ of any paraxial ray are all equal to the ray height h_k at surface k, and the angle u_{k+2} after the null-element lens is given by:

$$u_{k+2} = n u_{k-1} - h_k K_{\text{tot}}, \tag{3.27}$$

where u_{k-1} is the ray angle before refraction at surface k. Note that if K_{tot} is kept constant, u_{k+2} remains constant as well, and the entire paraxial ray path remains unaffected.

The above analysis shows that the paraxial properties of the entire system remain invariant when c_k, c_{k+1}, and c_{k+2} are changed in the plane OAB, and that the paraxial constraint is violated in the direction of Oz', normal to the plane OAB. Therefore, O is a NESP with Morse index 1 *in the paraxially constrained variable space* [3].

Before inserting the meniscus lens, the original local minimum should be properly optimized so that the residual gradient of the merit function is sufficiently close to zero. By inserting the null-element in such a way that a saddle point is created, we again obtain a system which has zero gradient. For optimizing this system, we first construct two starting points on opposite sides of the saddle. This can be done for instance by

[3] If constraints on real rays are used instead of paraxial ones, it is not expected that the Morse index will change, because then one of the eigenvalues of the Hessian matrix should change from positive to negative, see Section 3.1. The real ray properties are not so different from the paraxial ray properties to expect a change in sign of an eigenvalue.

slightly perturbing two consecutive surfaces, which are in contact in the NESP, with a small change $\pm\epsilon$ in the surface curvatures:

$$c_{k+1} = c_{k+2} = c_{\text{ref}} \pm \epsilon, \quad c_k = c_{\text{ref}}, \tag{3.28}$$

[the points P_1 and P_2 in Figure 3.5(b)], or alternatively:

$$c_k = c_{k+1} = c_{\text{ref}} \pm \epsilon, \quad c_{k+2} = c_{\text{ref}}, \tag{3.29}$$

[the points P_3 and P_4 in Figure 3.5(b)], where c_{ref} is the curvature of the reference surface in the starting local minimum. Perturbing the saddle point in this way has the advantage that in both starting systems we still have a (glass or air) null-element which does not affect the ray paths. Therefore, if the original local minimum already satisfies the required optimization constraints, the starting points will satisfy those constraints automatically. Note that the procedure remains valid if radii are used instead of curvatures.

Since a saddle point has zero gradient, the two starting points have a small value of the gradient for small values of ϵ. In order to obtain the desired outcome of the subsequent local optimization, the residual gradient of the original local minimum must be significantly smaller than the ϵ-dependent gradient at the two starting points.

By optimizing the two points obtained either with Equation (3.28) or (3.29) with all variables, we obtain two different local minima. If ϵ is chosen in the correct range, and the basins of attraction are well behaved [4], the same pair of local minima is obtained from Equation (3.28) or (3.29). For this reason, using only one pair of points, either the ones given by Equation (3.28) or (3.29), is sufficient.

We frequently observe that optimization algorithms have on one side of the saddle some difficulty to go down. After locally optimizing such points for the first time, we have to run the optimization routine for a second time before the algorithm converges to a local minimum. More research is necessary to fully understand this behavior.

Once the two minima on both sides of the NESP have been obtained, the distances between surfaces and the glass of the lens resulting from the null-element can be changed as desired. Technically, it is easier to increase thicknesses in the two resulting minima than in the saddle point itself. However, in Chapter 4, we show that many NESP's continue to exist as saddle points in the merit function landscape when in them the thin-lens thickness is increased.

[4] It should be noted that this way of choosing starting points on both sides of the saddle can conflict, in certain situations, with choices that have been made in the software implementation of local optimization algorithms. The saddle points are special points on the boundary between the basins of attraction that correspond to the two adjacent local minima. The basin of attraction is the set of starting points that lead to the same minimum after local optimization. Ideally, the two starting points should be points situated deeper within the two basins. However, the basin shapes depend on implementation details of the local optimization algorithm (e.g. damping method), as will be shown in Chapter 5. We have found examples where the two starting points are also close to the boundary between the two basins. In such cases, the outcome of local optimization started at those points can become less predictable. We have found examples where only one pair of points seems to be affected, either the one obtained with Equation (3.28) or the one with Equation (3.29). The other pair is still well-behaved and is adequate for SPC. Other choices of the pair of points on opposite sides of the saddle are also possible.

3.5 SPC: examples

In this section, we illustrate the special case of the SPC method with three examples. A detailed description of all necessary steps is given in Appendix A. For CODE V, lens files for all examples and a macro that creates the NESP are available via our website [49]. All three examples can be executed in a very short time. By optimizing two points at opposite sides of the saddle [the points given by Equation (3.28) or (3.29)] with all variables, we obtain two different local minima with zero distances between the surfaces inherited from the null-element. In these local minima, we gradually increase the zero distances to the desired values. The steps should be small enough to avoid jumps to other local minima. After each thickness increment, we reoptimize the system. During the optimizations, we keep the thicknesses constant.

The examples are independent of the used optical design software. We have tested all examples in CODE V and ZEMAX [50], obtaining the same results. In these examples, we use a merit function, which is based on transverse aberrations (root-mean-square spot size) with respect to the chief ray. In the merit function, all wavelengths and fields have a weight factor of unity. The object is placed at infinity and the image is kept at its paraxial position. The control of edge thickness violation was disabled when optimizing the two starting points situated on both sides of the saddle. Figure 3.6(a) shows the glass and air edge thicknesses of an optical system. We have an edge thickness violation when surfaces cross each other, see for example Figure 3.6(b). With two zero thicknesses, an edge thickness violation can easily appear, but may disappear when the zero thicknesses are increased in the resulting local minima.

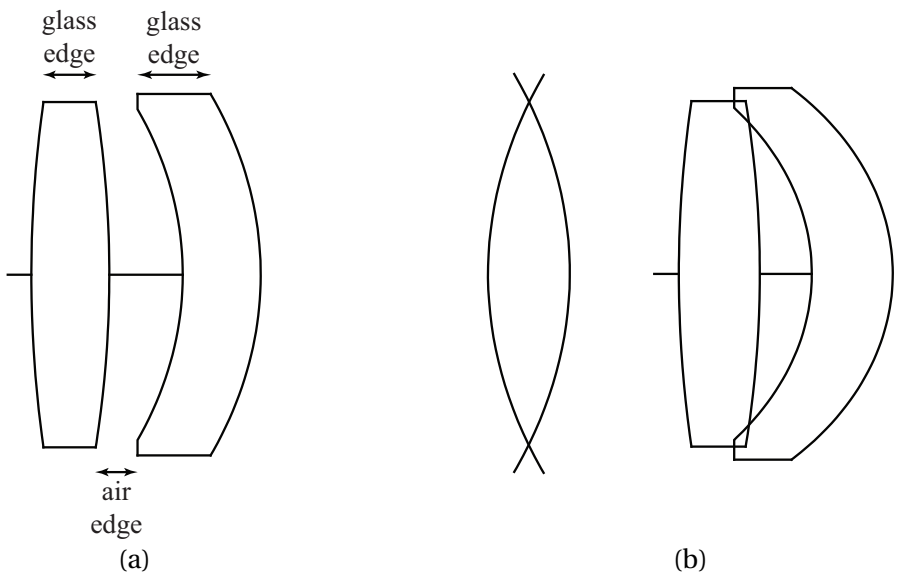


Figure 3.6: *(a) Glass and air edge thicknesses. (b) Two examples of edge thickness violations.*

3.5.1 Example 1: Generating doublets from a singlet

This very simple example serves two purposes: it illustrates the special case of the SPC method, which can be used in essentially the same way in all cases (including very complex systems), and also shows an advantage of SPC. When using SPC to insert a lens so that a NESP is created, two systems result after optimization. Inserting or splitting a lens in the traditional way results in a single system, which is not necessarily the better one. In this example, the better solution obtained from a singlet with SPC is missed when splitting the same singlet in a traditional way.

We start with an optimized singlet (f number 5, field of 5 degrees) with both curvatures as variables [Figure 3.7(a)]. A constraint is used to keep the effective focal length constant. Next, we insert a null-element meniscus, which has two variable curvatures and the same glass as the singlet, in contact with the second surface. When the two curvatures c_3 and c_4 of the null-element are equal to the second curvature (c_2) of the singlet, the obtained system is a doublet NESP. Note that, because of the two zero distances and the property $c_2 = c_3 = c_4$, the last three surfaces are overlapping and the lens drawing for the NESP is indistinguishable from the one for the original local minimum shown in Figure 3.7(a).

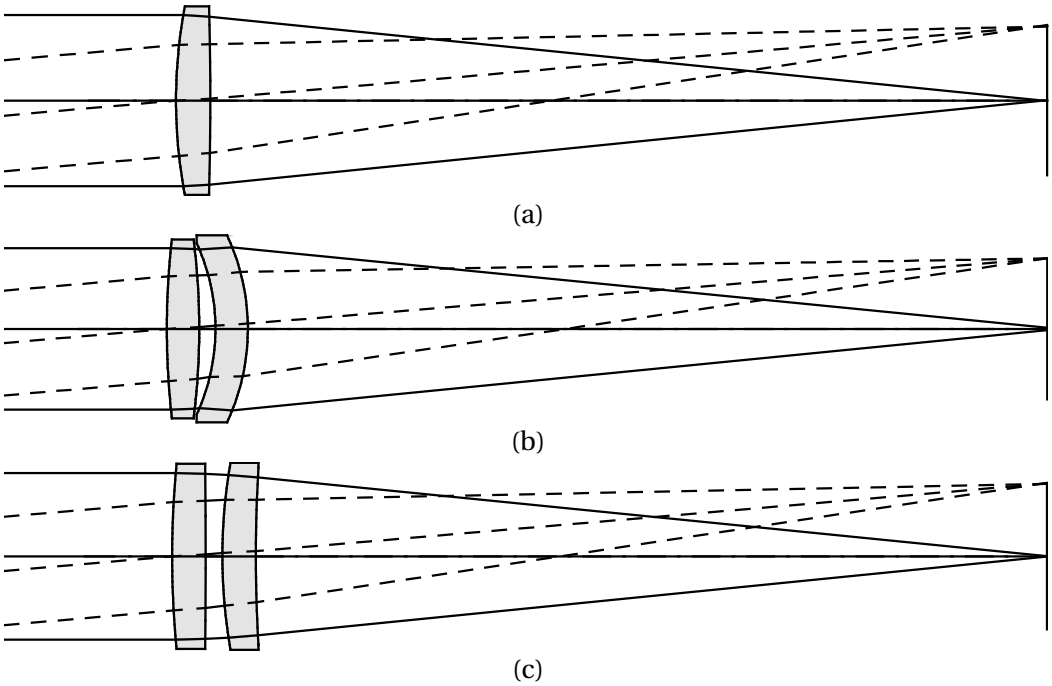


Figure 3.7: *SPC at the second surface of a singlet. (a) The starting singlet system and (b–c) the two resulting doublet local minima after increasing the distances between the last three surfaces (see text).*

By perturbing the null-element in the NESP with a small curvature change ϵ (and keeping the other variables unchanged), either with Equation (3.28) or (3.29), we obtain two systems situated on both sides of the doublet saddle. Optimizing these two systems results in two local minima with zero distances between the last three surfaces. We increase the zero thicknesses in small steps, and we optimize the obtained system after

each step. During the optimizations, we keep the thicknesses constant. Figures 3.7(b) and (c) show the two resulting doublet local minima after we increased the thicknesses in the way as described above.

After a glass change that makes axial color correction possible, reoptimization of the system shown in Figure 3.7(b) has lead to the air-spaced Fraunhofer-type configuration (Figure 3.8). When the starting singlet was split in a traditional way and then reoptimized, the resulting system was the system shown in Figure 3.7(c). However, this system has a poorer imaging performance than the system shown in Figure 3.7(b).

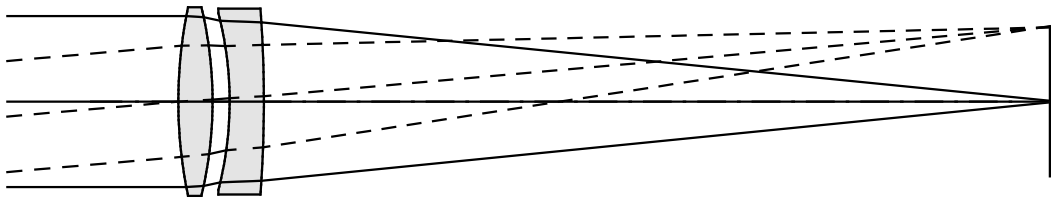


Figure 3.8: *Air-spaced Fraunhofer-type design.*

3.5.2 Example 2: Generating quintets from a quartet

In this example, we construct quintet NESP's in three different ways with the procedure described in Section 3.4. From the six resulting quintet minima, three of them (one resulting from each NESP) are identical, which illustrates the property we call 'convergence' that we have frequently observed with SPC: the same final design can be obtained in several different ways.

As a starting system, we use an optimized monochromatic quartet (f number 2, field of 14 degrees), with the first seven curvatures used as variables [Figure 3.9(a)]. The last surface is used to keep the effective focal length constant. All lenses have the same glass.

We first construct a quintet NESP by inserting a null-element (with variable curvatures and the same glass as the existing lenses) in contact with the second surface of the quartet. We perturb the saddle according to Equation (3.28) or (3.29), and after optimization and increasing of thickness of the second lens of the quintet, we obtain the local minima shown in Figures 3.9(b) and (c). In Figure 3.9(b), the axial distance between the first two lenses has also been increased in order to remove edge thickness violation. Note that as in the case of Example 1, we increase the thicknesses in small steps, and we optimize the obtained system after each step. During the optimizations, we keep the thicknesses constant.

Similarly, we construct two other quintet NESP's by inserting a thin-lens meniscus at the first and third surface of the quartet, respectively. From the resulting four quintet local minima, two of these turn out to be identical to the one shown in Figure 3.9(c). This is a very simple example of convergence to the same system via three different design routes. For configurations that have this property, if for any reason a design route, that should be successful, accidentally misses the goal (e.g. sometimes instabilities in local optimization influence the outcome, see Chapter 5), the same goal can be achieved via

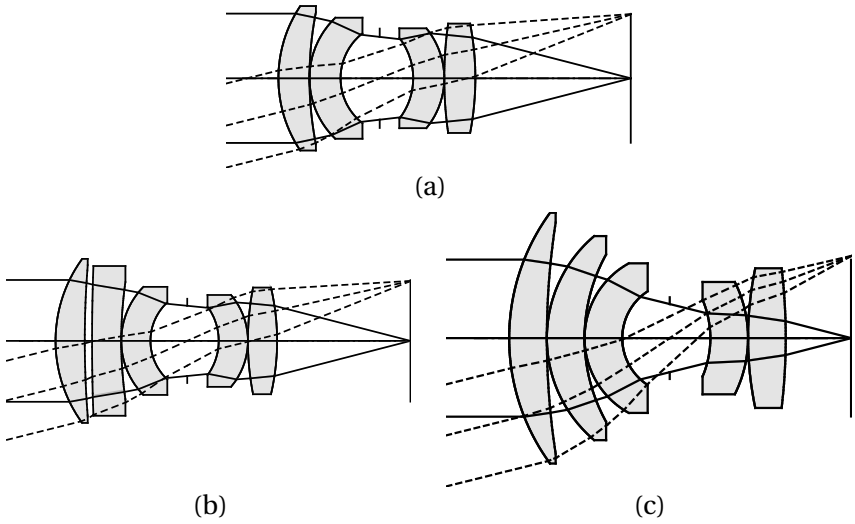


Figure 3.9: *SPC at the second surface of a quartet. (a) The starting quartet system and (b–c) the two resulting quintet local minima.*

another design route of the same kind. For more complex examples of convergence, see for instance Reference [43].

3.5.3 Example 3: Obtaining a Double Gauss design

Frequently, experienced designers observe that the shape of a design is probably not the best possible one. From the optimized system shown in Figure 3.10 (f number 3.33, field of 14 degrees), we want to obtain a Double Gauss shape. Here we show how we can move from one local minimum to a better one by first inserting, and then by extracting a lens. The optimization variables are all lens curvatures, except the curvatures of the two plane cemented surfaces which are kept unchanged, and the last curvature, which is used to keep the effective focal length constant. First, we use the SPC method at the second surface of the starting system to construct a NESP.

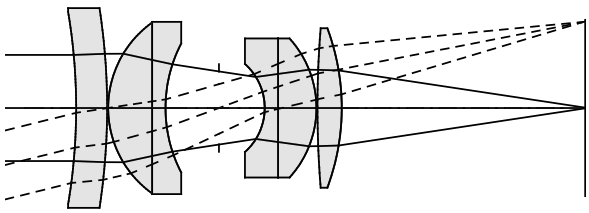


Figure 3.10: *Starting system in Example 3.*

On one side of the NESP, we optimize the point corresponding to point P_1 in Figure 3.5(b) [i.e. we take the + sign in Equation (3.28)], and we obtain the local minimum shown in Figure 3.11(a). As mentioned earlier, edge thickness control is temporarily disabled. Therefore, the second lens, which in Figure 3.11(a) seems to be a negative lens, has actually positive power.

Then, we gradually increase the thickness of the second lens to the same value as that

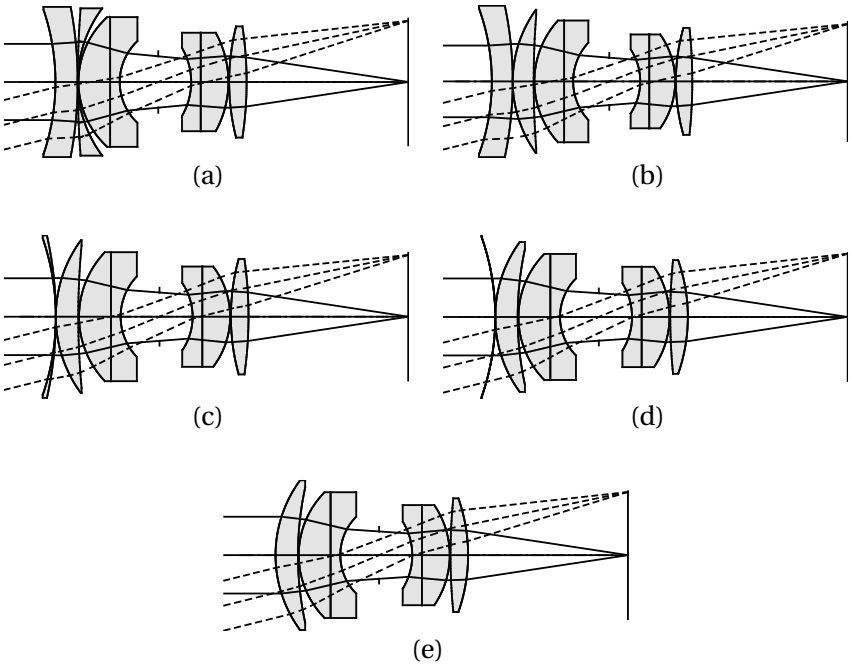


Figure 3.11: *SPC at the second surface in the lens system shown Figure 3.10. The resulting system with* 12 *surfaces after (a) optimizing the point* $c_{k+1} = c_{k+2} = c_{ref} + \epsilon$*, (b) increasing the thickness of the second lens, (c) decreasing the thickness of the first lens to zero, and (d) setting both curvatures of this lens equal. (e) After removing the first lens, the resulting system resembles the well-known Double Gauss design.*

of the first lens (in the same way as described in the previous examples). The resulting lens system with 12 surfaces is shown in Figure 3.11(b). Next, we remove the first lens in two steps: after gradually decreasing its thickness to zero [Figure 3.11(c)] and reoptimizing the system after each decrease in thickness (the thickness is not used as a variable during optimization), we make both curvatures of the first lens equal. Finally, we reoptimize the system (with the first two curvatures fixed), see Figure 3.11(d). In this way way, the first lens becomes a null-element, which can be removed without changing any ray path. The resulting system is a local minimum with the same number of lenses as the starting one, but with a much lower merit function value [Figure 3.11(e)]. The final shape resembles the well-known Double Gauss design. Similar techniques have been applied successfully in the design of state-of-the-art lithographic objectives [44]. Sometimes, after SPC, some lens in the resulting optimized system seems to have no role any more. Such a lens is then a good candidate for removal, even when it is situated farther away than in the above case from the position of SPC insertion.

3.6 Non-degenerate and degenerate merit function behavior during SPC

For a better understanding of the properties of SPC, Figure 3.12(a) shows the equimagnitude contours (i.e. the contours along which the merit function is constant) close to

the NESP constructed at the second surface in the previous example, computed numerically for points in the plane OAB given by Equations (3.21)–(3.23). The NESP with the property $c_2 = c_3 = c_4$ is located exactly in the middle of the figure, there where two equimagnitude lines cross. In the near vicinity of the crossing point $c_2 = c_3 = c_4$, the MF landscape in the plane OAB has the shape of a two-dimensional saddle. The dark regions correspond to regions with low MF values. Figure 3.12(b) shows the MF values along the two axes in Figure 3.12(a). The thin curve in Figure 3.12(b) corresponds to the horizontal axis, and the thick curve corresponds to the vertical axis. Note that the thin curve is minimal at the position of the NESP, while the thick curve has two minima, and a local maximum at the position of the NESP. The thick curve indicates the two sides of the saddle where the MF initially decreases. Further away from the crossing point, the MF increases in all directions of the plane OAB.

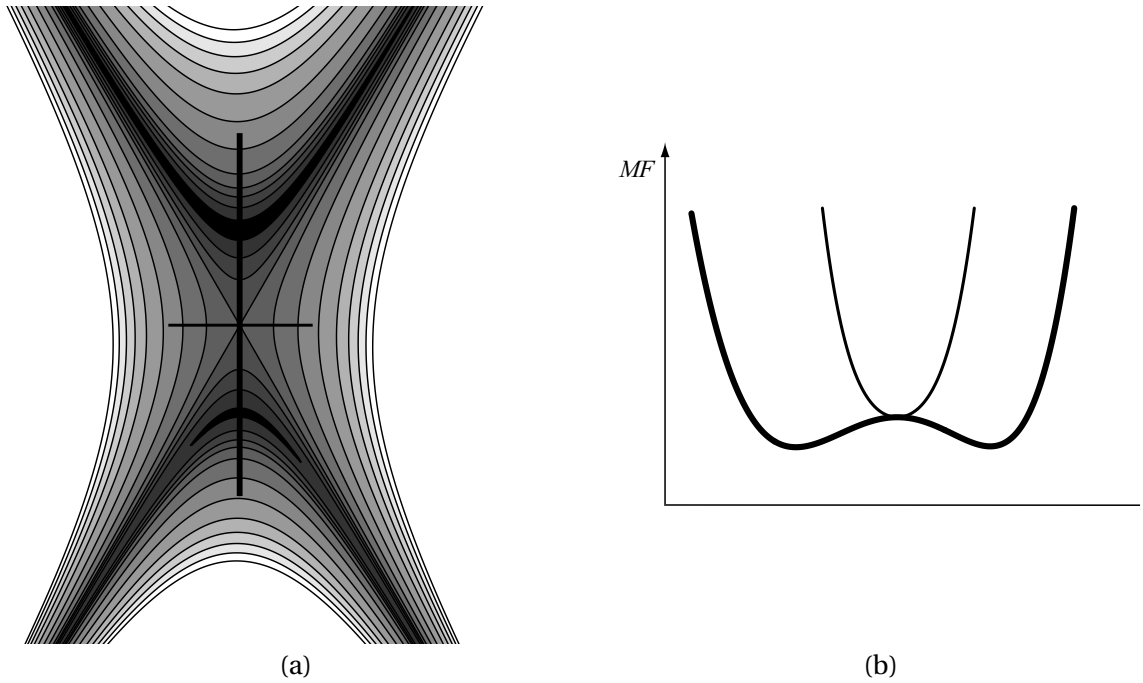


Figure 3.12: *(a) Equimagnitude contours in the plane OAB. A null-element meniscus is inserted at zero distance after the second surface of a lens system with* 10 *surfaces [see Figure 3.10]. The glass of the null-element is the same as the glass of the first lens. The dark regions correspond to low MF values. (b) MF plots along the horizontal axis (thin line) and vertical axis (thick line) in Figure 3.12(a).*

This example of merit function behavior in the plane OAB illustrates important properties of the SPC method. In Section 3.4, it was assumed that the saddle point O in the plane OAB is non-degenerate. In the absence of degeneracy, when two equimagnitude lines cross in a plane, the crossing point is a saddle point. For instance, for $f(x, y) = xy$ (the plot of this function is very similar to the saddle surface in Figure 3.3) the crossing equimagnitude lines for $f(x, y) = 0$ are the two lines $x = 0$ and $y = 0$ which correspond to the two axes of the coordinate system. The origin [where $f(0,0) = 0$] is then a saddle point: it is a minimum along the line $x = y$ where $f(x, y) > 0$ for all nonzero x and y values and a maximum along the line $x = -y$ where $f(x, y) < 0$.

However, from a mathematical point of view, two crossing equimagnitude lines do not necessarily indicate the presence of a saddle point. For instance, for $g(x,y) = x^2y^2$ (see Figure 3.13) the same two axes of the coordinate system are also equimagnitude lines for $g(x,y) = 0$, but the origin with $g(0,0) = 0$ cannot be a saddle point, because we have $g(x,y) > 0$ for all points that are not on the axes and there is no direction along which the origin is a maximum. In fact, all points on the two axes are critical points: they are minima in one direction, and are flat (i.e. g is constant) in the orthogonal one. Such critical points are called non-isolated because in arbitrarily small vicinities of any critical point we can find other critical points. Because of the flatness in one direction, non-isolated critical points are degenerate, but mathematically, as Poston and Stewart write, they are "in a strong sense extremely uncommon, so for many purposes may be ignored" [51].

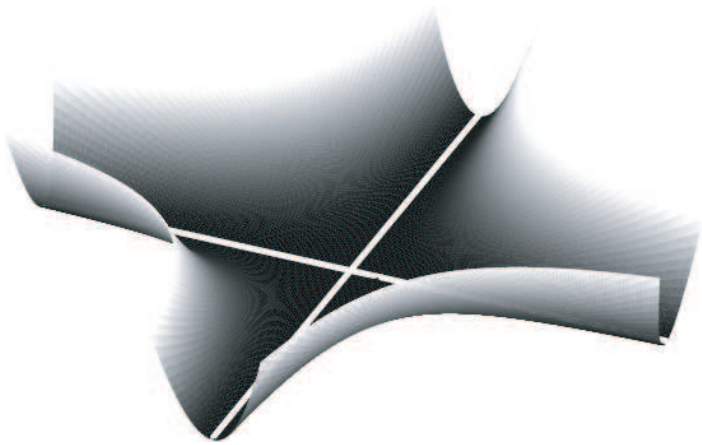


Figure 3.13: *The function $g(x,y) = x^2y^2$. The two white lines correspond to $x = 0$ and $y = 0$. All points on these two lines have zero gradient.*

The remarkable property revealed by Figure 3.12 is that, while close enough to the crossing point of the invariant lines given by Equations (3.13) and (3.14) the *MF* behavior is non-degenerate, far away from the crossing point the behavior of *MF* tends to become degenerate. In the latter case, the points on the invariant lines tend to become non-isolated critical points (i.e. they are minima perpendicular to the invariant lines, and, obviously, these points are flat along the invariant lines) [5].

While from a mathematical point of view one may always expect a non-degenerate region around the point constructed with SPC, in practical situations one can encounter situations where the numerically observed behavior is practically degenerate. It is well known that, if a lens is inserted at an inadequate position in an optimized system, subsequent optimization cannot decrease the merit function significantly. The thick curve with two minima and a maximum in Figure 3.12(b) gives an intuitive idea about what can be expected in the entire multi-dimensional variable space around the saddle point.

Further research is needed in order to understand SPC behavior at a position where the system does not 'need' extra degrees of freedom for improvement. If the non-degenerate region close to the saddle point (not only in the *OAB* plane, but in the

[5] In the thin-lens doublet example in Section 3.2, we also encountered a tendency to degenerate behavior far away from the crossing points.

entire variable space) is too small and the degenerate behavior is dominant, the positions and heights of the two minima on both sides of the saddle point can be so close to those of the saddle point itself that SPC becomes useless. Identifying the best insertion positions for SPC in an existing design is an important issue that requires further research. In the Double Gauss example, the insertion position for which SPC was successful was determined on the basis of previous knowledge about the desirable outcome.

3.7 Conclusion

One of the major difficulties in present-day global optimization is that the computing time increases significantly when the dimensionality of the optimization problem is increased. The Saddle-Point Construction (SPC) method for finding new local minima suffers much less from this drawback.

In this chapter, a simple and efficient version of the SPC method is presented. We prove that, if the dimensionality of the optimization problem is increased in a way that satisfies certain mathematical conditions (the existence of two independent transformations [Equations (3.13) and (3.14)] that leave the merit function unchanged) then a local minimum is transformed into a saddle point.

In lens design, we transform a local minimum into a saddle point by adding a 'null-element' meniscus (which does not affect the path of any ray or the merit function of the system) in contact with an existing lens. The null-element comes with two new variables (the two surface curvatures), and when the values for these curvatures are equal to the curvature of the contact surface (the reference surface), and it is made from the same glass as the lens at the reference surface, it transforms the local minimum into a 'null-element saddle point' (NESP) in the variable space with increased dimensionality. After optimization, two new local minima with both two zero distances result from the NESP. When both local minima are obtained, the zero distances can be increased to the desired values.

Lens designers frequently insert lenses into their designs and, in the traditional way, one new system shape results after optimization. However, when a lens is inserted with SPC, two distinct system shapes result and for further design one can choose the better one. By inserting lenses according to the SPC method, and then, if necessary, by extracting lenses, new local minima for optical systems of arbitrary complexity can be obtained very rapidly.

In principle, SPC should also be applicable in other optimization problems, where it is possible to define a 'null-element' and to find two independent transformations, similar to Equations (3.13) and (3.14), that leave the merit function unchanged (e.g. in thin-film optimization). However, in applications other than lens design more research is needed to investigate the practical utility of SPC.

Chapter 4

Looking for order in the optical design landscape

4.1 Introduction

In this chapter, we show that a certain degree of order is present in the optical design landscape, and that this order manifest itself at different levels. The order is best observed when we consider not only local minima of the merit function, but saddle points with Morse index 1 as well. As described in Chapters 1 and 3, the merit function around these saddle points behaves like the two-dimensional surface of a horse saddle. Irrespective of the number of variables, such a point, when perturbed on both sides of the 'saddle', leads after local optimization to two different local minima (see Figure 3.3).

For a design landscape with a fixed number of variables (i.e. the dimensionality of the problem remains constant), all minima and saddle points with Morse index 1 are linked in a network when certain quite general conditions are met. An example of such a network is presented in Figure 4.1 for the global search corresponding to the well-known Cooke triplet [10, 52]. The shown results are collected from a set of runs with different specifications (runs with object at infinity, runs for symmetric problems with transverse magnification -1, the field varies between 20 and 33 degrees). Individual runs may show minor departures from this figure, e.g. a pair consisting of a minimum and a neighboring saddle point may disappear or such an extra pair may appear. Since the network is symmetric, pairs of systems, which look like mirror images of each other, are grouped together for clarity in the same box. The minima are drawn within thick-line boxes and the saddle points within thin-line boxes. The two downward paths of local optimization started on both sides of the saddle at a saddle point $s_{i\text{-}j}$ lead to minima m_i and m_j, as indicated by the continuous lines between systems. The dashed lines indicate a possible instability of the link when specifications are changed. Then, on one side of the saddle point $s_{i\text{-}(j \text{ v } k)}$ the downward path may lead either to minimum m_j or to m_k. ('v' denotes the Logical OR.) Extra information in this figure will be explained in Sections 4.3, 4.4, and 4.5.

For systems that are sufficiently simple, the entire network can be detected link by link, where each link between two local minima contains a Morse index 1 saddle point. How-

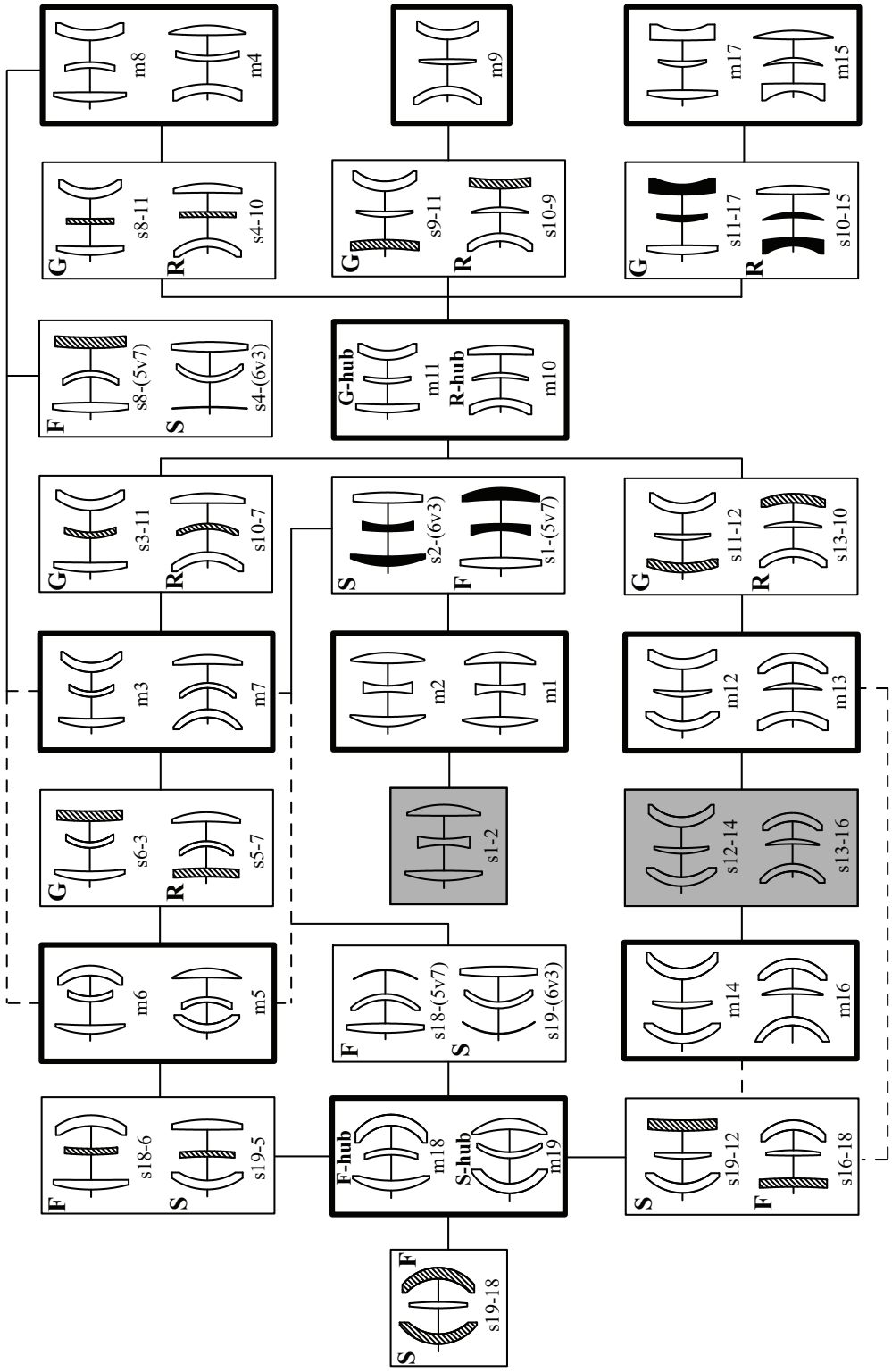


Figure 4.1: *Network of the global search corresponding to the Cooke triplet (m1 and m2) for which the variables are the surface curvatures. The results are collected from a set of runs with different specifications.*

ever, when the number of components increases, this detection approach becomes difficult and very time-consuming.

A different level of order can be found when we compare global searches in design landscapes of different dimensions. For instance, starting from a system which is a minimum with N optimization variables, a saddle point with $N+2$ optimization variables can be constructed by inserting in the system a 'null-element' lens in the special way as shown in Chapter 3. Here, we discuss the general case where the restrictions mentioned in Chapter 3 are removed. The generalized version of the Saddle-Point Construction (SPC) method typically finds more than one saddle point configuration for a certain position of the inserted null-element. Some of these saddle points lead to local minima that cannot be found via saddle points constructed with the special version. Furthermore, the generalized SPC method can also be used to construct saddle points with a 'null-air' lens that cannot be made with the special version.

The SPC method will be generalized in Section 4.2, and examples will be given in Section 4.3. In Sections 4.3, 4.4 and 4.5, we will discuss the network shown in Figure 4.1 in more detail. In Section 4.4, the emphasis will be on understanding relationships between minima and saddle points when the dimensionality of the optimization problem is changed. Although many saddle points can be obtained by using the SPC method, for some saddle points the method does not work. Their existence will be explained in Section 4.5. In Section 4.6, we introduce the concept of topological degree, with which the appearance and disappearance of local minima and saddle points can be better understood.

4.2 Generalizing the SPC method

In Chapter 3, we have shown how saddle points can be constructed in the special case when the inserted lens is in direct contact with one of the surfaces of the original local minimum (the reference surface) and when the glass of the new lens is the same as that at the reference surface. If in the final design the glass of the new lens must be different and/or the lens must be placed at a certain distance from a given surface of an existing minimum, then those adjustments can be done once the two minima on both sides of the so-called 'null-element saddle point' (NESP) have been obtained.

With the generalized version of the SPC method, the glass and distance restrictions mentioned above are removed, and the curvatures of the lens to be inserted can be computed numerically. The difference in glass properties and the distance to the existing surface can be interpreted as a 'perturbation' of the previous 'ideal' situation described in Chapter 3. In this section, we show mathematically that as long as this 'perturbation' is not too large, the saddle point continues to exist and that it will keep the same Morse index.

We consider an optical system that is a local minimum with N optimization variables (called in what follows the 'old' variables). We want to insert a lens in the system in such a way that we create a saddle point. The two new surfaces will have the indices $k+1$ and $k+2$ in the system. For simplicity, we consider the case when the lens surfaces

are spherical, with the curvatures c_{k+1} and c_{k+2} as two 'new' optimization variables, although the method can be easily generalized to aspherical surfaces, as has been shown for extreme ultraviolet systems in Reference [43]. In Chapter 3, we have shown that in the special case when the new lens is introduced in contact with a surface (with index k) of an already existing lens with the same glass, a saddle point is created if, for instance, the new lens has zero thickness and equal curvatures having the values:

$$c_{k+1} = c_{k+2} = c_k, \tag{4.1}$$

and all N old variables are left unchanged. Since a zero-thickness meniscus with equal spherical surface curvatures $c_{k+1} = c_{k+2}$ disappears physically, we called it a 'null-element', see also Chapter 3.

We consider now the case when a null-element is inserted at an arbitrary position in the system and has an arbitrary glass. The merit function is an arbitrary optical merit function (MF). Since a null-element does not affect the path of the light rays, MF remains unchanged and equal to the merit function of the original local minimum with N variables for any value of the null-element curvature, i.e. along an entire straight line in the variable space:

$$c_{k+1} = c_{k+2} = u, \tag{4.2}$$

where u is variable and all old variables are kept constant. With a null-element, the partial derivatives of the merit function with respect to the old optimization variables remain unchanged, i.e. equal to zero. Saddle points are critical points, which means that for obtaining a saddle point the partial derivatives with respect to the new variables $\partial MF/\partial c_{k+1}$ and $\partial MF/\partial c_{k+1}$ must also vanish. Since along the line given by Equation (4.2) we have:

$$dMF = \frac{\partial MF}{\partial c_{k+1}} dc_{k+1} + \frac{\partial MF}{\partial c_{k+2}} dc_{k+2} = 0, \tag{4.3}$$

and $dc_{k+1} = dc_{k+2}$, it follows that:

$$\frac{\partial MF}{\partial c_{k+1}} = -\frac{\partial MF}{\partial c_{k+2}}, \tag{4.4}$$

which means that both new components of the gradient vanish simultaneously. Thus, for finding the saddle points, we find numerically the values of the null-element curvature for which:

$$\frac{\partial MF}{\partial c_{k+1}} = 0. \tag{4.5}$$

The critical points that are solutions of Equation (4.5) are saddle points (and have therefore a direction along which MF can decrease) because they cannot be local minima (assuming that the solutions are non-degenerate [1]). For local minima in the variable space, the equimagnitude hypersurfaces of MF around them are ellipsoids that reduce to a single point when MF has the value that corresponds to the minimum. However,

[1] A discussion about non-degeneracy is given in Section 3.6.

solutions of Equation (4.5) are situated on equimagnitude hypersurfaces that contain an entire straight line, given by Equation (4.2), along which MF remains constant.

Note that, because of Equation (4.4), for finding a NESP only one equation with one unknown c_{k+1} must numerically be solved. This technique is considerably more efficient than detecting a saddle point in the optical design space without a-priori knowledge of optical system properties, as in the case of the saddle points shown in Figure 4.1. When a saddle point is found, we take on the line given by Equation (4.2) two points on opposite sides of the saddle, having:

$$c_{k+1} = c_{k+2} = c_s \pm \epsilon, \tag{4.6}$$

where ϵ indicates a small curvature change and c_s is a solution of Equation (4.5). Then, these two points are optimized and two different local minima will result. Finally, in these solutions the zero distance between surfaces $k+1$ and $k+2$ can be increased to the desired values. This can be done in the same way as described in the examples in Section 3.5: We increase the zero thicknesses in small steps, and we optimize the obtained system after each step. During the optimizations, we keep the thicknesses constant.

4.3 Generalized SPC: examples

We have implemented the generalized SPC method in the macro language of the optical design program CODE V. In all examples shown in this chapter, the merit function is this program's default error function based on the transverse ray aberration computed with respect to the chief ray. A simple example is shown in Figure 4.2. For a doublet local minimum, in which a null-element meniscus has been inserted after the first lens, $\partial MF/\partial c_3$ is plotted as a function of $c_3 (= c_4)$. The curve has a parabolic shape and cuts the axis ($\partial MF/\partial c_3 = 0$) in two points.

In this example, the reason why Equation (4.5) has two solutions can be easily understood by examining the drawings of the systems corresponding to the two saddle points. For the saddle points s10-7 and s4-10, the null-element curvature is almost the same as c_2 and c_5, respectively, despite the fact that the null-element is placed at some distance from the neighboring surfaces and that its glass differs from the glasses of the first and last lens. In fact, if for s10-7 the distance between the second and third surface is gradually decreased to zero, and the glass of the null-element is gradually changed into the glass of the first lens, then the saddle point has the property $c_3 = c_4 = c_2$. Similarly, s4-10 can be continuously transformed into a saddle point with $c_3 = c_4 = c_5$.

From the two NESP's, three local minima m4, m7 and m10 are obtained by using Equation (4.6) and optimizing the perturbed points (m10 results from both saddle points). When the zero lens thickness is increased, the three minima become the systems indicated by black arrows after reoptimization. The same three nonzero thickness minima (m4, m7 and m10) and two nonzero thickness saddle points (s10-7 and s4-10) linked to them can also be obtained with the network-detection software. By comparing the NESP's with the corresponding nonzero thickness saddle points, we observe that with the exception of the thickness of the middle lens the systems are essentially identical.

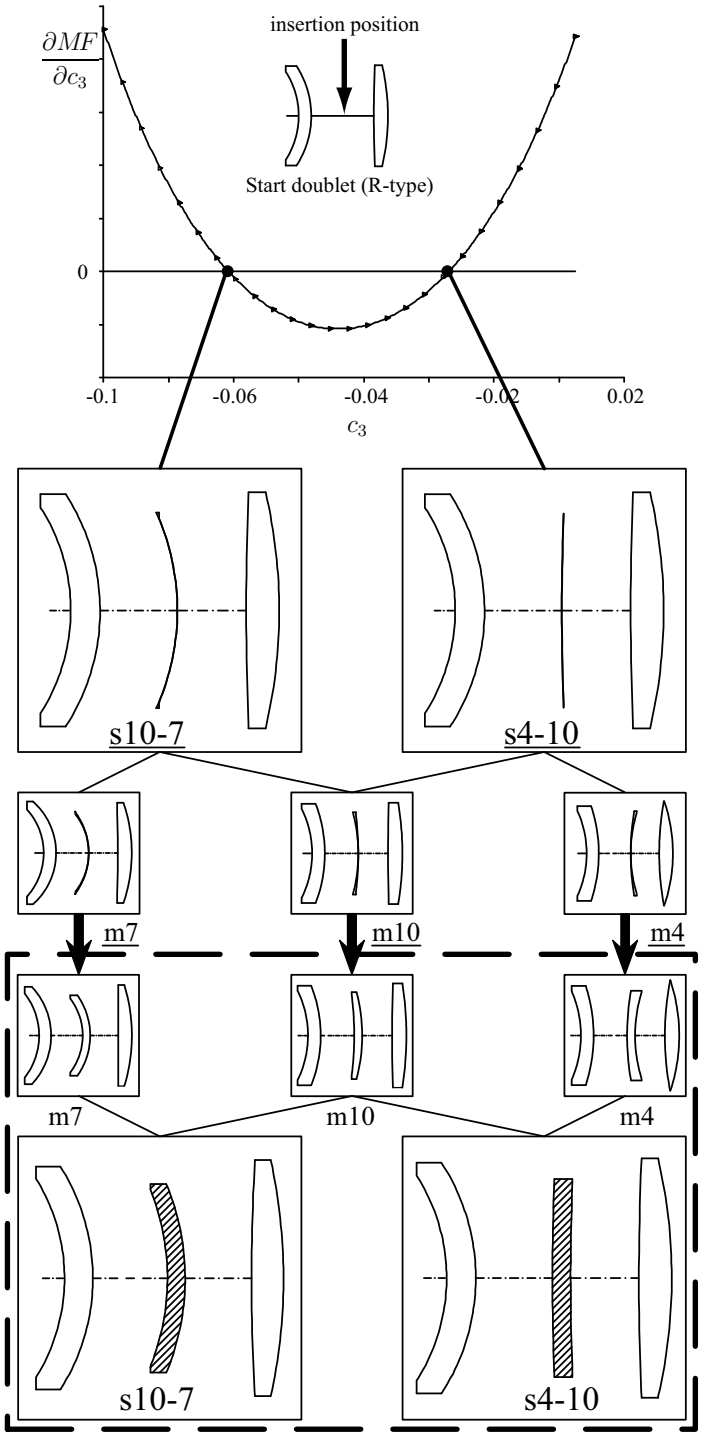


Figure 4.2: *Construction of a triplet saddle point by inserting a null-element meniscus into a doublet local minimum with variable curvatures. The aperture, field, and wavelength specifications, the glass types and constant distances between surfaces are those for a Cooke triplet global search with object at infinity. The systems in the dashed box can be found in Figure 4.1, their counterparts with a zero-thickness lens are underlined.*

Thus, the fact that one lens in the saddle points constructed with this method has zero thickness is not an obstacle for theoretical analysis or practical applications, because from NESP's the same minima with nonzero thickness can always be obtained as from the corresponding nonzero thickness saddle points. It is just technically easier to increase the thickness in a local minimum than in a saddle point.

In Section 3.2, we have studied simple systems consisting of thin spherical lenses in contact, and we have shown that the curvatures of certain thin-lens saddle points have the property of three consecutive equal curvatures. This property is used in the special version of the SPC method. By using the generalized SPC method, we can also construct NESP's which cannot be continuously transformed into saddle points with three consecutive equal curvatures. Those NESP's can lead to new local minima, which cannot be found with saddle points constructed with the special version of the SPC method. For instance, by using the SPC method in the special case, we can construct the thin-lens doublet NESP's SP2 and SP3, which are given in Table 3.1, but not SP1 and SP4. These last two NESP's can be constructed with the generalized SPC method. As shown in Figure 2.7 for a doublet with small distances between the surfaces, SP1 and SP4 lead to new local minima.

A similar characteristic can be observed for the thin-lens triplet NESP's. From the doublet hub [see Equation (3.12)], we can construct six triplet NESP's. Four of them (see Table 3.2) can be constructed with the special version of the SPC method, and the remaining two (given in Table 4.1 [2]) can be obtained by using the generalized SPC method.

Table 4.1: *The curvatures of two of the six Morse index 1 saddle points of a thin-lens triplet with $n = 3/2$, object at infinity. The remaining four saddle points are given in Table 3.2.*

c_1	c_2	c_3	c_4	c_5	c_6
$-6/7$	$-6/7$	$6/7$	$-1/7$	$11/7$	$4/7$
$6/7$	$-1/7$	$11/7$	$4/7$	$16/7$	$16/7$

Another example can be found in Figure 4.1. Saddle points s9-11 and s10-9 can be continuously transformed into NESP's which cannot be constructed with the special version of the SPC method. These two NESP are essential because local minimum m9 can only be found via those two saddle points.

4.4 Why are there so many minima in the Cooke triplet global search?

In global optimization, the focus is usually on the quality, given by the value of their merit function, of the systems that correspond to different local minima. We believe, however, that the existence of local minima and their quality are distinct issues to be investigated. For instance, the network structure (i.e. the local minima, the saddle points,

[2] The given curvatures are analytical solutions of $\nabla SA = 0$.

and the links between them) are often remarkably stable when specifications change, but the best local minima for some specifications do not necessarily remain the best ones for other specifications.

As a simple example, we compare the networks corresponding to an achromatic air-spaced doublet global search in two different situations: the situation when the first lens has a flint glass and the second lens a crown glass, and the situation when the glass order is reversed. For thin achromatic doublets, for which the optimization variables are the surface curvatures, it is well known that there are four doublet shapes (G, S, F and R, see Figure 4.3) for which the important aberrations (spherical aberration, coma and axial color) can be simultaneously corrected [37]. Another doublet minimum, H, also exists, but has poor imaging qualities, because not even spherical aberration is corrected [47]. For axial color correction, the crown glass element must be the one with positive power, so with crown glass first the best solutions are F and G and with flint first R and S are best, as can be also observed in Figure 4.4. In Figure 4.4, a system shown with its corresponding name between brackets performs worse than the system in the other network with the same name, but without brackets. The given values below each doublet are CODE V merit function values. Remarkably however, the two networks are identical, i.e. all five minima exist in both cases and are linked in the same way via the corresponding saddle points, despite the fact that the solutions which have good quality are not the same.

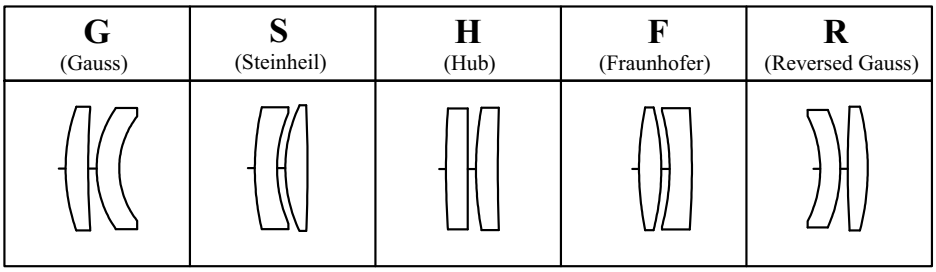


Figure 4.3: *Minima in the achromatic doublet global search: G (Gauss), S (Steinheil), F (Fraunhofer), R (Reversed Gauss) and H (Hub).*

The networks in Figure 4.4 were obtained with our network detection software. Note that all saddle points have the shape of a positive lens plus a meniscus. We also obtained all of them from singlets in a way similar to that shown in Figure 4.2. For example, Figure 4.5 shows how we can construct the network on the right in Figure 4.4. Note that SP3 and SP4 can be continuously transformed to NESP's that can also be obtained by using the SPC method in the special case. However, we have to use the generalized SPC method to construct SP1 and SP4. These two NESP's can be continuously transformed to the saddle points SP1 and SP4 given in Table 3.1.

In general, when we change control parameters (i.e. system specifications and parameters that remain constant during optimization), the network structure remains unchanged until a control parameter reaches a certain threshold value. At threshold, local minima and saddle points appear or disappear in the networks. For many network links, the parameter ranges between threshold values are large, and therefore, despite significant changes of the control parameters, the corresponding local minima and saddle points continue to exist, with usually only minor modifications in the shape of the

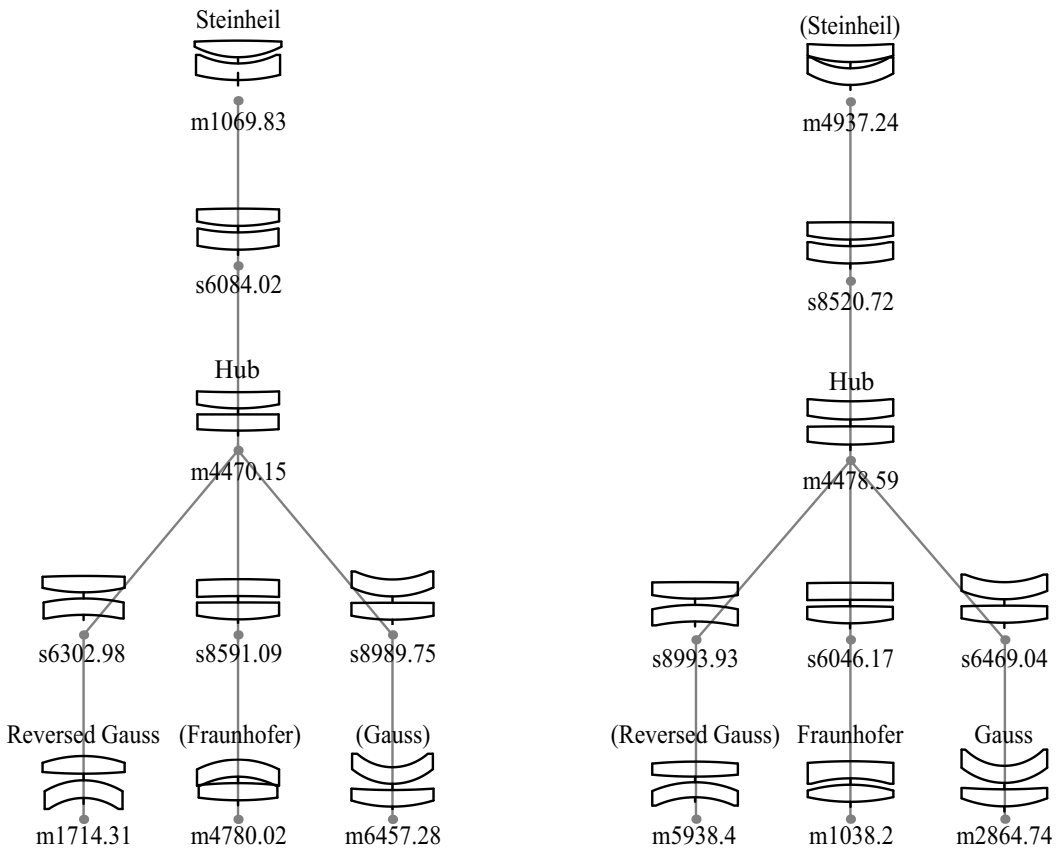


Figure 4.4: *The networks of local minima and saddle points for an air-spaced doublet. On the left: flint-crown glass order; on the right: crown-flint glass order. The 'm' systems are local minima, and the 's' systems are saddle points with Morse index* 1. *The given values are CODE V merit function values.*

systems. For example, the doublet minima (especially G, S, F and R) are remarkably robust. In most cases, they continue to exist as local minima when the aperture and field specifications, as well as the distances between surfaces and the glass types, are gradually changed to Cooke triplet values, although the merit function deteriorates severely (see Figure 4.6). The NESP's constructed as shown in Chapter 3 and in the previous sections, together with the local minima that are linked to them and still have zero distances between certain surfaces, continue to exist in most cases when the zero distances between surfaces are increased to values that are not too large.

Since we believe that understanding stable network structures is important, we ignore for the moment the imaging quality of the various minima and focus in what follows on why multiple local minima exist at all and on how they can be obtained in the given problem. If in a one-dimensional optimization problem a single minimum was expected, but two local minima were observed, the reason for this would become clear if one could understand why a maximum exists between the minima. In a multi-dimensional optimization problem, a saddle point with Morse index 1 is a maximum in one direction in the variable space, and a minimum in all other orthogonal directions. Since minima can be viewed as resulting from saddle points by letting the optimization roll down on both sides of the 'saddle', in order to understand why multiple local

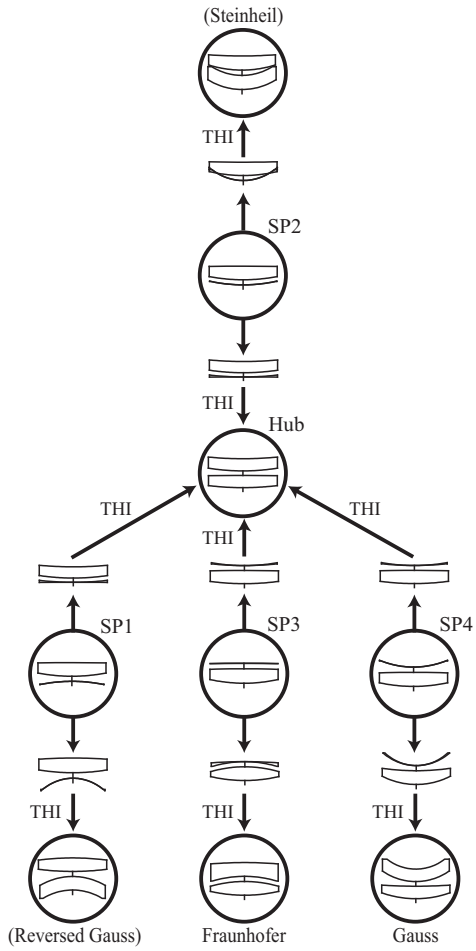


Figure 4.5: *Constructing a doublet network from a singlet by using the generalized SPC method. The encircled configurations correspond to the doublets shown in the network on the right in Figure 4.4. The arrows show which local minimum we obtain after optimization. Eight arrows (indicated with 'THI') result from optimizing doublet local minima in which the thickness of the null-element that results from a NESP is increased.*

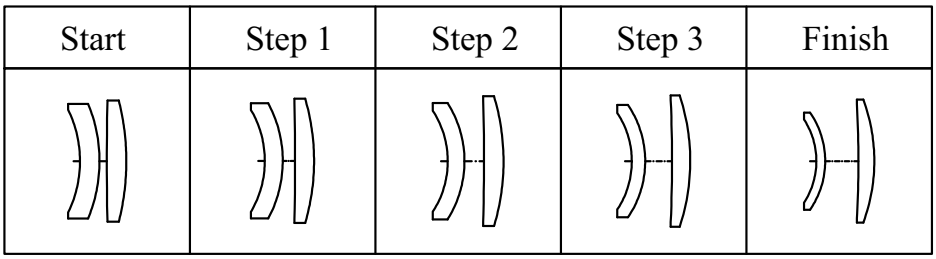


Figure 4.6: *A minimum from Figure 4.3 is changing shape when specifications are gradually changed in several steps.*

minima exist, it is sufficient to understand why saddle points exist.

According to this line of reasoning, the presence of minima such as m4, m7 and m10 in Figure 4.2 (which can be detected not only with the network search, but with other methods, such as Global Synthesis of CODE V, as well) can be viewed as a consequence of the presence of saddle points such as s10-7 and s4-10 which are linked to them. These saddle points can be viewed as resulting from NESP's (s10-7 and s4-10) which in turn are the two solutions of Equation (4.5), applied to a doublet local minimum plus a null-element that has been inserted between the doublet lenses. In this way, the presence of a minimum in a doublet merit function landscape leads to the presence of saddle points (and of minima resulting from them via optimization) in a triplet landscape.

For studying the relationship between minima with N variables and saddle points with $N+2$ variables, it is useful to examine simple situations first, in which the entire set of saddle points can be found, and see how many of them result from local minima with fewer lenses. The corresponding minimum with N variables will be called a precursor of the saddle point with $N+2$ variables and of the two minima linked to it. For this reason, we return now to the Cooke triplet global search shown in Figure 4.1, which was obtained with our network detection software. It turns out that out of the 26 saddle points, 19 can be obtained from doublet local minima plus a null-glass meniscus , and 4 can only be constructed with a null-airspace meniscus. The exceptions are the three saddle points shown in gray boxes, which we could not obtain from doublets with a null-element. In some sense they seem less important because they do not lead to new interesting local minima. In Section 4.5, we discuss the 'gray boxed' saddle points. We refer to them as non-NESP's.

Two sets of doublet local minima have been obtained by removing one of the Cooke triplet lenses (the first, or the second, or the third) and globally optimizing the remaining doublet with variable curvatures for Cooke triplet specifications. With our generalized saddle point construction method, 19 triplet saddle points (those that have a hatched lens in Figure 4.1) result from the doublets of types G, S, F, and R (see Figure 4.3). By ignoring the hatched lens in the saddle point drawings, we can easily recognize the doublet type indicated in the figures. Keep in mind that because of different specifications, the systems in Figure 4.3 change shape as shown in Figure 4.6. In all cases, the hatched lens is a meniscus with almost equal curvatures. For zero thickness this lens becomes the null-element used in the SPC method, as shown in Figure 4.2. In a few drawings, a thin lens is drawn instead of a hatched one, indicating that the corresponding saddle point exists only for small thickness values, and that the saddle point disappears in some of the network runs where the thickness of that lens happens to be larger than the threshold value. One saddle point, s19-18, has two hatched lenses, which shows that it can be obtained in two different ways from doublets.

Four saddle points, which are indicated with two black lenses, result by adding a null-airspace lens into one of the doublet lenses. The initial doublet types can be recognized when we glue the black lenses together. For example, the lenses in the two saddle points s2-(6v3) and s1-(5v7) that lead to the two Cooke triplets (m1 and m2) have some power, and none of them can be continuously transformed into a null-element meniscus. However, both saddle points can be constructed by inserting a 'null-airspace'

lens in a particular doublet configuration. Saddle points s2-(6v3) and s1-(5v7) result by adding a null-air lens in the first lens of a type S doublet and in the second lens of a type F doublet, respectively. In order to calculate the curvatures of the null-airspace lens, we have to use the SPC method in the general case.

When we try to find saddle points s2-(6v3) and s1-(5v7) for instance by adding a null-element lens between the two lenses of a H-type doublet, Equation (4.5) does not have real solutions, i.e. the parabola-like curve in Figure 4.2 does not cut the axis. However, by starting from the H-type doublet plus null-element lens and performing a one-dimensional succession of local optimizations along the line given by Equation (4.2) with values of u close to the minimum of the parabola, we were able to reach m1 and m2 and also the local minima on the other side of s2-(6v3) and s1-(5v7). Remarkably, since these two saddle points have Morse index 1, they repel the solution in one direction (the downward direction of the 'saddle'), but attract the solution in all other directions. Therefore, the optimization first converges towards them and then diverges, reaching the local minima. If the optimization is stopped in early stages, zero-thickness systems close to saddle points are obtained that resemble s2-(6v3) and s1-(5v7), as shown in Figure 4.7.

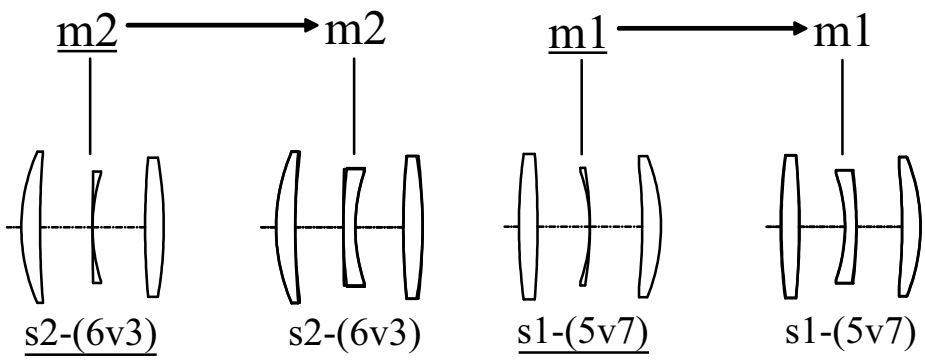
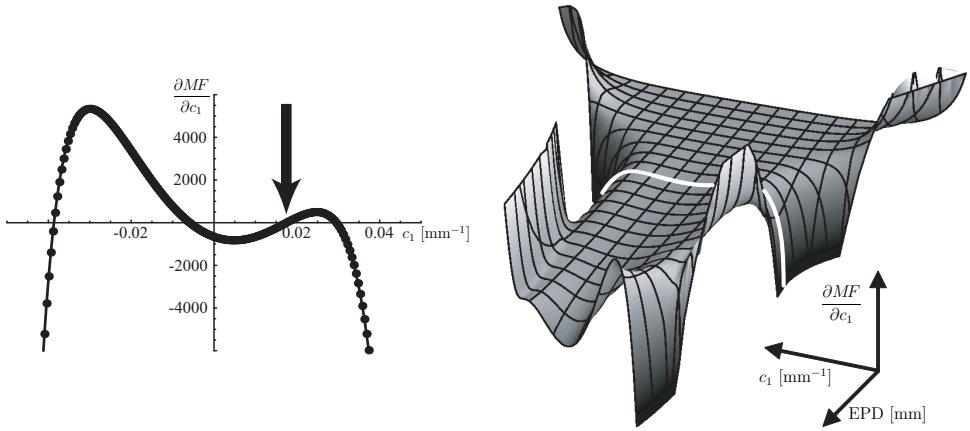


Figure 4.7: *For certain starting values of $c_3 = c_4 = u$, the optimization algorithm leads after a certain number of cycles to systems that strongly resemble the saddle points s2-(6v3) and s1-(5v7). For both saddle points, the left drawing shows the thin-lens intermediate optimization result (underlined). The right drawing actually shows the superposition of two drawings: the exact saddle point found with our network search, and the drawing on the left in which the thickness of the middle lens has been increased to facilitate comparison. Note that for both saddle points the two drawings on the right side are almost indistinguishable, indicating that the optimization has arrived very close to the real saddle points.*

The ∇MF-curve in Figure 4.2 has a parabolic shape. However, the shape is not always like that. The shape of the ∇MF-curve depends on the starting local minimum, the position of the inserted null-element, and the system parameters such as aperture and field. In some cases, the curve does not cut the axis. When we have more than two lenses, the ∇MF-curve has a complicated shape, and more than two intersection points are possible. Figure 4.8(a) shows an example of such a curve. In a local minimum with seven variables [the configuration shown in Figure 3.11(e)], a null-element meniscus has been inserted in contact with the first lens surface, and $\partial MF/\partial c_1$ is plotted as a function of $c_1 (= c_2)$. The curve cuts the axis ($\partial MF/\partial c_1 = 0$) in four points [which are the



(a) (b)

Figure 4.8: *$\partial MF/\partial c_1$ for a configuration with nine variables, made by inserting a null-element in contact with the first lens surface of the system shown in Figure 3.11(e)]. (a) EPD = 35 mm, field = 2 degrees. (b) While the field was kept constant at 2 degrees, the EPD was changed from 15 mm to 61.3 mm. The white curve corresponds to the curve shown in Figure 4.8(a).*

four solutions of Equation (4.5)]. The glass of the null-element is the same as the glass of the first lens in the old configuration. Note that the NESP, which corresponds with the third intersection point (indicated with the arrow), can also be obtained by using the SPC method in the special case. In Figure 4.8(b), we show the $\partial MF/\partial c_1$-surface when we vary the entrance pupil diameter (EPD). The field is 2 degrees and is kept constant. For small EPD values (EPD $\leq$ 24 mm), the curve has a parabolic shape. Further research is necessary do fully understand the shape of the surface.

Finally, by examining how systems that have the same precursor are arranged in the Cooke triplet network, we observe an interesting structure, shown in Figure 4.9. The systems (minima and saddle points) that have the same precursor turn out to be all connected to each other by network links. With the exception of the three saddle points in gray, the entire network is partitioned in regions in which all systems result from the same doublet type F, S, G, or R. Because of symmetry, Figure 4.9 shows two regions: F–S, G–R. In these regions, many saddle points are linked to the same minimum, called the 'hub' of that region. In fact, in the doublet networks in Figure 4.4, we encounter a similar feature: all doublet saddle points (they can be obtained from the same singlet) are linked to the same minimum, the doublet hub (H). For certain specifications, hubs that have hubs as precursors can be found, e.g. the Fulcher system [46]. Such 'super' hubs have, we believe, interesting properties.

The doublet regions in the Cooke triplet network overlap at certain local minima (which we call linking minima) thus ensuring the connectivity of the entire network. Because, as shown in Figure 4.1, a few network links (drawn dashed) are unstable, for some specifications instead of the main linking minimum (shown with overlapping hatching) another minimum (in brackets) takes over the role of linking minimum. The systems m5 and m6 in Figure 4.1, which in Figure 4.9 are shown as linking minima for three different doublet regions, are particularly interesting. By adding lenses to them via SPC, we could obtain the Double Gauss design.

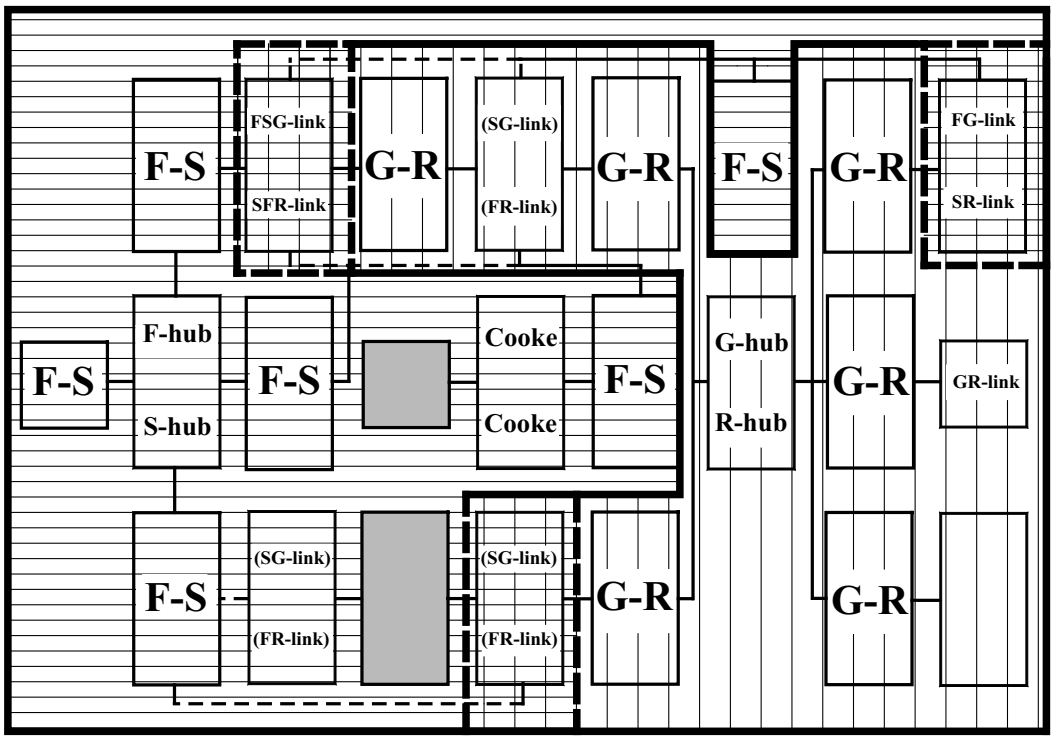


Figure 4.9: *Partition of the Cooke triplet network in distinct connected regions, such that the systems in each region have the same doublet precursor. The systems in the boxes (drawings not shown) are the same as in Figure 4.1. Different regions are marked with different hatching and are separated by thick continuous lines. Several local minima that have more than one precursor form links between distinct regions. Linking minima are shown in thick dashed boxes in which the hatching is an overlap of the hatching of the regions they link.*

4.5 Why do non-NESP's exist?

In order to understand the existence of saddle points which cannot be constructed with the SPC method (e.g. the saddle points shown in the gray boxes in Figure 4.1), we study the merit function landscape of a very simple monochromatic optical system (f number 3, field of 3 degrees), consisting of two thin lenses with equal glasses in contact. The curvatures of the second and third surfaces (c_2 and c_3, respectively) are used as variables. The first curvature of this thin-lens doublet is kept constant, and the last curvature is used to keep the effective focal length constant at 100 mm. The merit function landscape of this thin-lens doublet contains five local minima. By gradually increasing the distances between the lens surfaces, the thin-lens local minima can be continuously transformed into the G, S, H, and F type doublets in Figure 4.3 [3]. Besides the five minima, we find four saddle points. Three of them are NESP's. Note that, because the first curvature is fixed, the type R doublet cannot be obtained.

Since the effective focal length remains constant, astigmatism and Petzval blur remain constant. The stop is placed at the lens system, so distortion vanishes as well. Because the doublet is monochromatic, chromatic aberrations are also not present. Due to the

[3] Two thin-lens doublet minima transform into the F type doublet.

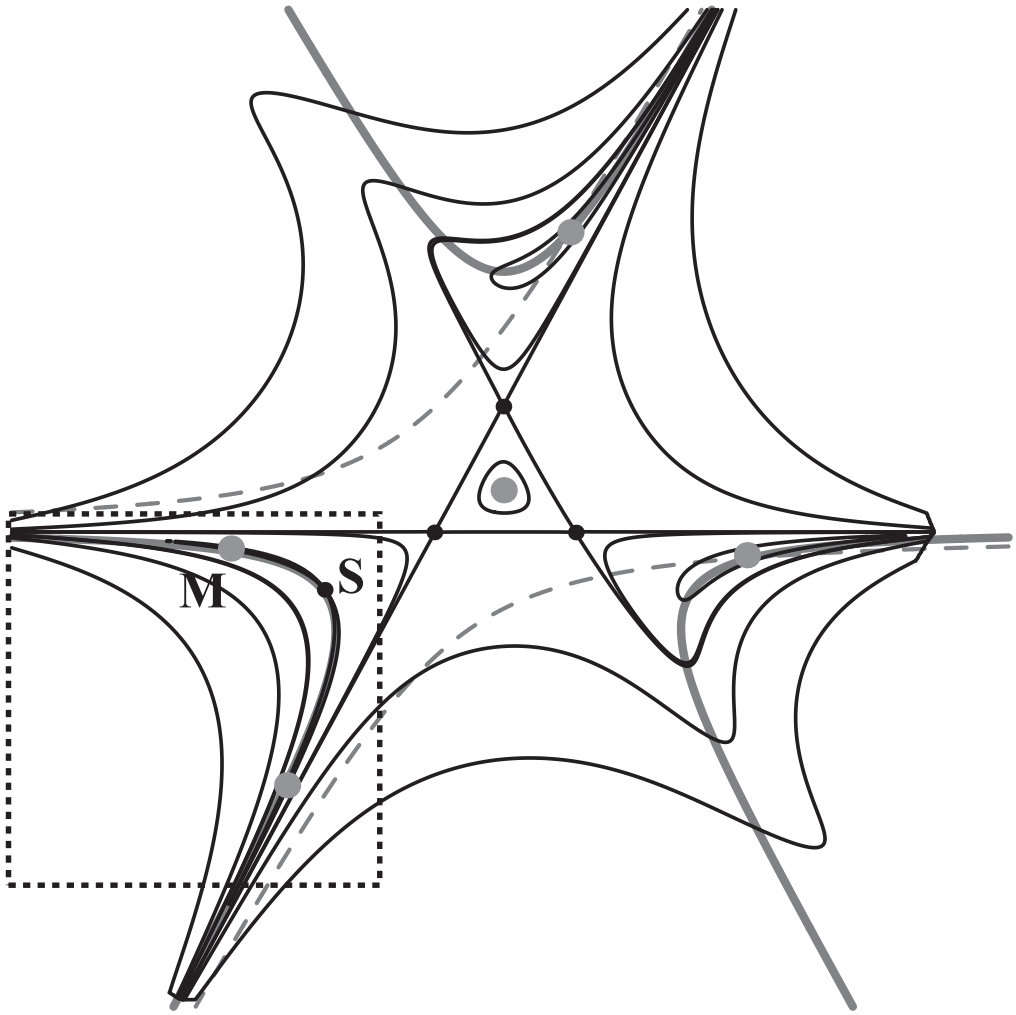


Figure 4.10: *Topography of a two-dimensional monochromatic thin-lens doublet landscape (f number* 3*, and field of* 3 *degrees) for a typical value of the first curvature. Shown are MF equimagnitude contours (black), five local minima (large gray points), four saddle points (small black points), three curves for zero SA (thick gray), and two curves for zero coma (dashed gray). One of the SA* = 0 *curves is in the dashed box, and is almost totally hidden by overlapping merit function equimagnitude contours. This SA* = 0 *curve is close to the two local minima (minimum M is one of them) and saddle point S.*

large aperture, spherical aberration is the predominant aberration in this doublet example. The field is small, but not negligible, so third-order coma influences the system performance as well.

Figure 4.10 shows the topography of the *MF* landscape, with five local minima (large gray points), four saddle points (small black points), the *MF* equimagnitude contours (black), and the curves of zero third-order spherical aberration SA (thick gray curves) and zero third-order coma (dashed gray curves). The best four local minima are located almost on one of the curves of SA = 0, and close to a zero third-order coma curve. Saddle point 'S' is the only saddle in this merit function landscape that cannot be constructed with our SPC method. We call this a non-NESP.

For the following discussion, we focus on the lower left part of Figure 4.10, indicated by the dashed box. The existence of saddle point S can be understood by considering the three most dominant aberrations: SA, third-order coma, and fifth-order spherical aberration. When we use a simplified merit function model, in which only SA is considered, then the curve of SA = 0 shown in the dashed box contains a curve of solutions of the optimization problem (the gray SA=0 curve in the dashed box, which is close to the two local minima and saddle point S). By also taking coma into account, this degeneracy can be removed. This leads to two local minima close to both sides of the SA = 0 curve. Depending on the aperture, fifth-order spherical aberration influences that result. The interplay between third-order coma and fifth-order spherical aberration results in one or two local minima, which are close to the curve of SA = 0.

We remark that the saddle point S and minimum M are less robust, and can easily disappear when we change some parameters that have not been used for optimization. For example, when we slightly increase the distances between the surfaces, both critical points merge into a non-critical point. The resulting topography is then similar to the one shown in Figure 1.5 [4]. In this example, we have the so-called 'fold' catastrophe [53]. When a saddle point disappears, one of the two local minima linked to it always disappears as well. In Section 4.6, we show that then the topological degree remains constant.

Another example of the fold catastrophe can be found when we, for instance, decrease the distances in the doublet configurations shown in Figure 2.7. Saddle point SP5 and local minimum LM5 have almost the same shape, and the *MF* differences between them turn out to be low. In the network, the pair of systems formed by saddle point SP5 and local minimum LM5 is therefore less robust than the rest of the network, and disappears when we decrease the distances between the surfaces. The local minimum and the saddle point, which were linked with the disappeared pair of systems (LM1 and SP1, respectively), will then be linked together. In this way, the connectivity of the network will not be affected.

The best two local minima in Figure 4.1 are m1 and m2, which have the well-known shape of the Cooke triplet. Both minima are slightly asymmetric and form a mirror pair, whereas the saddle point s1-2 (a non-NESP) between them is symmetric. However, for large numerical apertures, these two minima will merge into a single symmetric one, which looks similar to s1-2 [7]. In Catastrophe Theory, this behavior is known as the 'cusp' catastrophe [53]. This phenomenon of appearing and disappearing is illustrated in Figure 4.11.

When we change control parameters such as aperture, field, or system parameters that have not been used for optimization, the non-NESP's such as saddle point S in Figure 4.10, saddle point SP5 in Figure 2.7, or saddle point s1-2 in Figure 4.1, tend to appear and disappear in the landscape more easily than saddle points that can be constructed with the (generalized) SPC method. In this research, we noticed that NESP's are typically more robust and therefore more fundamental than non-NESP's. In the merit function landscape, non-NESP typically form low barriers separating two local minima in the merit function landscape. These bariers can easily be overcome by using for instance a method based on simulated annealing. A slight change in the local optimization al-

[4] In Figure 1.5, we have used a different aspect ratio.

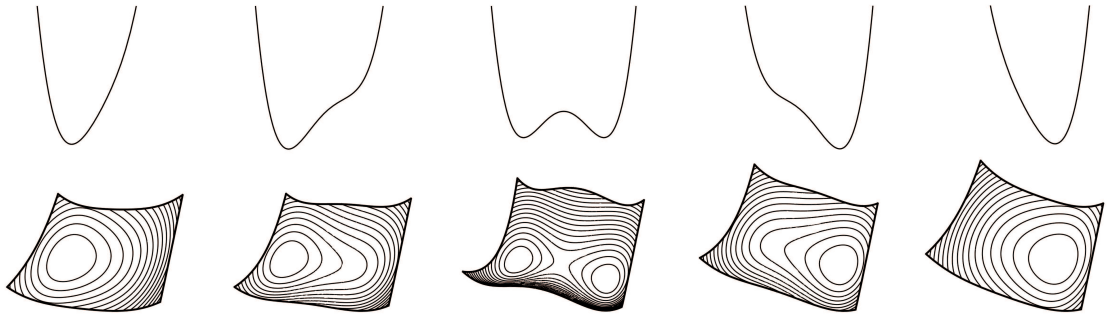


Figure 4.11: *One- and two-dimensional illustrations of the 'cusp' catastrophe. By changing a control parameter, we can change the topography of the MF from having one local minimum (left), via two local minima separated by a saddle point (middle), to one local minimum (right) which is different from the starting one.*

gorithm (e.g. use a lower damping factor, see Chapter 6 [5]) can also be sufficient to overcome such barriers.

In Chapter 3, we have already shown a third type of catastrophe: the so-called 'elliptic umbilic' catastrophe. When a control parameter reaches a critical value, the saddle points and the hub merge into a single degenerate critical point. The merit function landscape around such a point takes then the peculiar shape of a 'monkey saddle'. In the spirit of Catastrophe Theory, it can be said that the non-degenerate saddle points are 'shaken loose' from the monkey saddle when the control parameter deviates from its critical value.

4.6 Conservation of topological degree

Mathematically, there is a conservation property for the set of all critical points (minima, saddle points, and maxima) that are present in the merit function landscape. Using the concept of topological degree, the appearance and disappearance of local minima and saddle points can be better understood. In Reference [54], a well written introduction is given about the concept of topological degree.

When there are no degenerate points, the Hessian matrix $\mathbf{H}\left(\mathbf{x}^j\right)$ of the merit function MF in critical point $\mathbf{x}^j$ is an invertible matrix for all m critical points $\mathbf{x}^j$ $(j = 1, 2, \ldots, m)$ in the N-dimensional variable space Ω. Hence, the determinant $\det \mathbf{H}\left(\mathbf{x}^j\right)$ is nonzero. In that case the total topological degree of the derivative ∇MF of MF in Ω is defined by [54]:

$$\deg(\nabla MF, \Omega, \mathbf{0}) = \sum_{j=1}^{m} \operatorname{sign} \det \mathbf{H}(\mathbf{x}^j), \tag{4.7}$$

[5] In Section 6.4, we show that low damping in a damped least-squares algorithm can be used to overcome a low merit function barrier that separates a poor local minimum from a better one.

where

$$\text{sign}(x) = +1, \quad \text{if} \quad x > 0,$$
$$\text{sign}(x) = -1, \quad \text{if} \quad x < 0,$$

and $\det \mathbf{H}\left(\mathbf{x}^j\right)$ is given by the product of all eigenvectors of $\mathbf{H}\left(\mathbf{x}^j\right)$, see also Equation (3.1). Note that the value of $\deg(\nabla MF, \Omega, \mathbf{0})$ is always an integer.

A critical point with Morse index 1 has one negative eigenvalue (see also Section 3.1), and the sign of $\det \mathbf{H}$ is therefore negative. For critical points with Morse index 2 (having two negative eigenvalues) the sign of $\det \mathbf{H}$ is positive, and so on. Thus, we can rewrite Equation (4.7) as:

$$\deg(\nabla MF, \Omega, \mathbf{0}) = \sum_{j=1}^{m} (-1)^{MI^j}, \tag{4.8}$$

where MI^j is the Morse index of critical point $\mathbf{x}^j$.

When there is a degenerate point $\mathbf{x}^j$, i.e. $\det \mathbf{H}\left(\mathbf{x}^j\right) = 0$ for some of the critical points in Ω, we can not use Equations (4.7) and (4.8). However, it can be shown that in this case the degree of ∇MF in Ω can still be uniquely defined. Here, we omit the details and refer to Reference [54].

The conservation of topological degree applied to the derivative of the merit function implies the following. Assume that the optical system depends on a parameter t that is fixed. For example, t could be the focal distance, numerical aperture, some distance between two surfaces, or some curvature that is kept fixed; t may also be a set of parameters that is kept fixed. Suppose that the merit function (or more appropriately ∇MF) continuously depends on t, and that for all t of interest the derivative ∇MF does not have critical points on the boundary $\partial\Omega$ of the variable space. Then the topological degree is independent of t [54]. This implies that if for certain t critical points are born, these points must have Morse indices such that the sum of their topological degrees vanishes. Otherwise, a critical point has appeared (or disappeared) through the boundary of Ω. However, the individual critical points are typically unstable: they can change position inside Ω with changing t. They can be born or disappear for certain t. As shown in Figure 4.11, critical points can also merge or appear inside Ω for certain values of t. But in both cases, the topological degree remains constant. This holds for all critical points (degenerate and non-degenerate).

When a local minimum disappears, also a saddle point that connects this minimum to another minimum disappears. According to Equation (4.8), the topological degree is then conserved. More complex scenarios are also possible, as long as they do not violate the conservation of topological degree. For example, for a certain value of t, two saddle points with a Morse index value of 1 and one local minimum can merge into a single Morse index 1 saddle point in the new merit function landscape.

We remark that the practical importance of critical points decreases with the Morse index. For manufacturing lens systems, local minima (having a Morse index value of zero) are the most important critical points in the merit function landscape. Saddle points with a Morse index value of 1 are the most important saddle points, because they determine the relationships between local minima.

Degree theory is part of topology and as such it can be applied to many parts of physics, including optics. For example, in the analysis of phase singularities in the electromagnetic field, degree theory can be used to explain the birth and collapse of phase singularities [55].

4.7 Conclusion

In the complex topography of the optical merit function landscape a surprising 'hidden' structure has been found. This structure is best observed when we compare the results of global searches with different dimensionality and when we consider not only local minima, but saddle points as well. We have shown that for simple systems, many saddle points result from adding null-elements to minima with a lower dimensionality.

In this chapter, we have shown the network of the global search corresponding to the Cooke triplet, which was obtained with our network search algorithm. This algorithm is generally applicable to any optimization problem with continuous variables and does not use information about specific optical system properties. However, we have shown that the vast majority of the triplet saddle points and local minima obtained with our network search can be obtained in a much simpler and faster way by using the SPC method, which does use such extra information.

Although the theory behind the SPC method contains concepts which are new in lens design, the technique is very straightforward, and both versions (for the special and general case) can be easily integrated with traditional design techniques. With the generalized SPC method, we can find all robust saddle points in the merit function landscape of doublet and triplet lens designs. A small number of saddle points (the non-NESP's) cannot be constructed with our SPC method. They typically have low merit function barriers, which can easily be overcome by using for instance a method based on simulated annealing.

Thinking in terms of saddle points is still unfamiliar to most optical system designers, but we believe that the potential for discovering new solutions and for improving design productivity justifies the effort for understanding such new methods and for combining them with traditional design methodology.

Chapter 5

Instabilities and fractal basins of attraction in optical system optimization

5.1 Introduction

There are typically many local minima in the optical merit function landscape, and the solution that will be obtained after local optimization is critically dependent on the choice of the initial configuration. In order to determine the initial configurations that lead to a given local minimum, the so-called basin of attraction for that minimum should be considered. The set of all starting configurations that are attracted to a local minimum is the basin of attraction for that minimum [56] (see Figure 5.1). A plot of the basins for all starting points after a finite number of iterations can be interesting for optical system design, because the basins show the sensitivity of the optimization result with respect to the starting values.

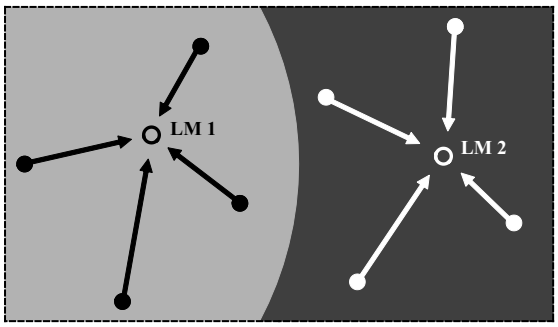


Figure 5.1: *Two neighboring basins of attraction. After local optimization, all initial points in one basin lead to the same local minimum (LM1 or LM2).*

The damped least-squares method is particularly suitable to solve nonlinear least-squares problems, in which the merit function is the weighted sum of squares of operands that describe the design targets (see Chapter 2). Many optical design programs use damped least-squares methods for local optimization in various forms with

great success [11, 12]. In damped least-squares methods, the operands are considered to be linear in a first approximation; the damping factor was originally introduced with the purpose of limiting the change of the variables to ensure that this linear approximation remains valid. However, in optimization programs, the damping factor is frequently chosen to increase computational speed. For example, it could be chosen such that the merit function decrease per optimization step is maximized.

In this chapter, we show that choosing low damping factors for the sake of increasing computational speed can create sensitivity with respect to initial conditions. In these cases, the basins of attraction have very complicated structures with non-integer fractal dimensions, similar to those encountered in research on dynamical systems [56–61]. One of the goals of this chapter is to make optical system designers aware of the possibility of presence of such instabilities in the optimization process, because they can influence the result that will be obtained after optimization.

It is well-known that a change in the optimization strategy or in the settings of the local optimization algorithm can lead to a different result, even when starting from the same configuration [12]. When the optimization results are distinct for different optimization methods, then the basins of attraction obtained with these methods are also different. We observe this property when we use different damped least-squares software implementations to obtain the basins of attraction for the same optimization problem.

Since computing the basins of attraction is very time-consuming, we use a simple lens configuration to illustrate the appearance of fractal behavior. A monochromatic doublet with only two variables will be the main example of this chapter. However, we also show that fractal behavior can be found in practical situations with more than two variables as well. In addition, by using an algorithm that suffers from fractal behavior, fractal basins can appear after applying our saddle-point construction method.

To have maximum stability, we first use a differential equation to go down the gradient of the merit function (Section 5.2). The basins of attraction obtained with this method are then used as a reference for the basins obtained with different programs using some form of a damped least-squares method (Section 5.3).

In order to explain the peculiar behavior we observed in commercially available codes, where we have no control over the damping, we study this behavior in more detail with the program OPTSYS, written by Joseph Braat. This program gives us more control over the damping. We will demonstrate that the choice of the damping factor is essential in obtaining stable results.

To further understand the fractal basin shapes, we examine the steps taken by the optimization algorithms of two commercially available programs in the variable space of the doublet example (Section 5.4). In addition, we demonstrate a procedure to change the step sizes in such a way that the basin boundaries become regular. With regular basin boundaries, the instabilities in the optimization process disappear.

5.2 Method for stable behavior

One technique used to find minima in the merit function (*MF*) landscape is to continuously follow the gradient of the merit function downwards, by solving the differential equation:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} = -\frac{\nabla MF}{\|\nabla MF\|}, \tag{5.1}$$

where the vector $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ describes a point in the N-dimensional solution space, and $\mathrm{d}s = \sqrt{\|\mathrm{d}\mathbf{x}\|^2}$ is the arc length along the downward path. Although this method is rather slow, it gives maximum stability to the minimization process and does not display fractal behavior in our examples. The basins of attraction obtained with this method are used as a reference for the basins generated by three different optimization algorithms, all of which are based on some version of the damped least-squares method: OPTSYS, and two commercial codes CODE V and ZEMAX. All these programs adapt the damping factor in the damped least-squares method automatically, but OPTSYS allows the user to specify the maximum value of the damping factor to be used.

The main example in this chapter is a monochromatic doublet (f number 3), with curvature c_2 of the second surface and curvature c_3 of the third surface used as variables. All specifications are given in Tables 5.1 and 5.2. The first curvature of the doublet is kept constant, and the last curvature is used to keep the effective focal length constant at 100 mm. The stop is placed at the first surface. The image plane is kept at the paraxial image position.

Table 5.1: *Specifications for the monochromatic doublet example.*

Entrance pupil diameter (mm)	33.333
Effective focal length (mm)	100.000
Paraxial image height (mm)	5.241
Field points (deg)	0.000, 2.000, and 3.000

Table 5.2: *Curvatures, thicknesses, and refractive indices of the starting points for the monochromatic doublet example. For curvatures c_2 and c_3 the variation domain is given in brackets.*

Surface #	Curvature (mm^{-1})	Thickness (mm)	Refractive index
object	0	infinity	1.
1	0.020	10.346	1.618
2	$(-0.025, 0.040)$	1.000	1.
3	$(-0.045, 0.075)$	2.351	1.717
4	solve	solve	1.
image	0	0.000	1.

Depending on the software used, optimizing the two variable curvatures results in four or five different local minima, which are shown in Figure 5.2. Minima A–D are stable local minima, and minimum B has a high merit function value. For the settings given

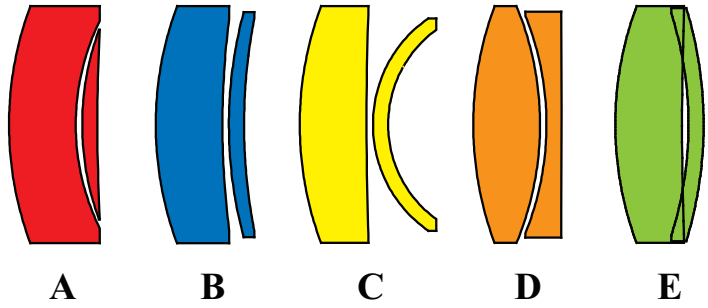


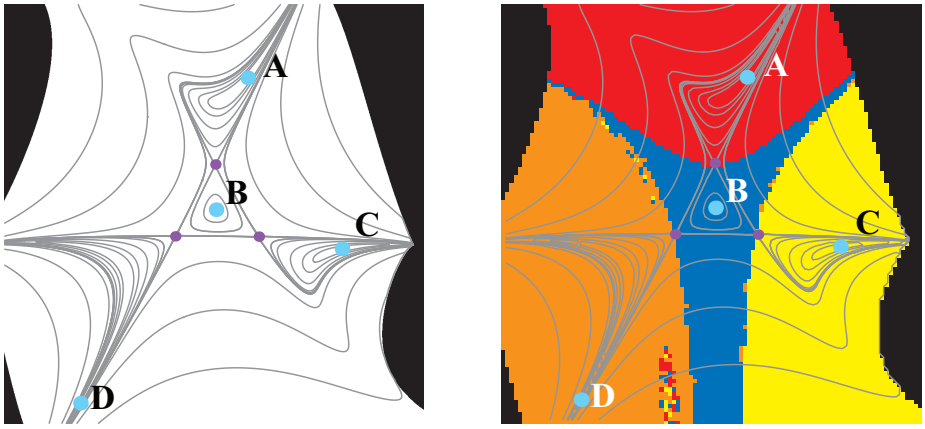
Figure 5.2: *Five local minima for a two-dimensional monochromatic doublet optimization problem. The colors correspond to the colors used for the basins of attraction in subsequent figures.*

in Tables 5.1 and 5.2, minimum E (which also suffers from some edge thickness violation) is less stable, and appears or disappears easily when we change some parameter or the software. The control of edge thickness violation was disabled in this example, in order to eliminate the constraints from the list of possible causes of the peculiar behavior described in what follows. Detailed analysis shows no qualitative distinction in this respect between regions of the variable space where edge thickness constraints are satisfied or violated.

In Figure 5.3, we show the merit function equimagnitude contours (i.e. the contours along which the merit function remains constant) of the default merit function of CODE V, which is based on transverse ray aberrations (root-mean-square spot size) with respect to the chief ray. In the merit function, all fields have a weight factor of unity. A similar plot has already been given in Figure 1.5 (the aspect ratio is different in both figures). In Figure 5.3, we also plotted the positions of the minima A–D (large blue dots) and of the saddle points (small purple dots), which are situated at the crossing of two segments of the equimagnitude merit function contours [10] (see also Chapter 3). By taking the merit function of minimum B as reference, the relative merit function values for A, C and D are 0.01, 0.08 and 0.02, respectively. The three saddle points have relative merit function values close to 1.2. In this and following figures, curvature c_2 is plotted along the vertical axis, and curvature c_3 is plotted along the horizontal axis.

For the settings used in Figure 5.3(a), minimum E did not exist. However, with slightly modified settings, minimum E can appear in CODE V as well. For example, when we decrease the glass thickness of the second lens to 0.3 mm, minimum E also appears in the merit function landscape made by CODE V. Note that minimum E also exist when all distances between the surfaces are equal to zero. See for instance minimum M in Figure 4.10.

To compute the basins of attraction for the doublet, we first associate colors to the various minima that result from local optimization. We then specify a grid of equally spaced starting points in a two-dimensional plane (the variation domain is given in Table 5.2). At every grid point, we start to optimize the corresponding configuration until it arrives in one of the local minima. Depending on which local minimum we obtain, we color the starting point with the corresponding color for that minimum (see Figure 5.2). Thus, the location of the point shows the initial value of the variables, and its color shows to



(a) (b)

Figure 5.3: *(a) Merit function equimagnitude contours of the two-dimensional monochromatic doublet example (specifications are given in Tables 5.1 and 5.2). The larger blue points A, B, C, and D correspond to the local minima shown in Figure 5.2. The smaller purple points are saddle points. In all figures of the same kind, the compact black regions contain configurations that suffer from ray failure. (b) Basins of attraction for the same doublet example (on a grid of* 101×101 *points), obtained by using differential equation (5.1) in a first stage of the optimization process in CODE V. The colors correspond to the local minima as shown in Figure 5.2.*

which local minimum it converges to. Starting points for which the initial configuration suffers from ray failure (a ray misses a surface or total internal reflection) are shown in black.

To achieve maximum predictability of the downward path from the starting point to the local minimum, we first use differential equation (5.1). Using CODE V, we implement a fourth-order Runge-Kutta method with an adaptive step size in the macro language to solve Equation (5.1). We determine the basins of attraction by choosing each point of a grid of 101×101 points as starting points. Since solving Equation (5.1) is very time-consuming, we use a finite number of steps. In a second stage, we continue with the default local optimization routine of CODE V.

For the vast majority of grid points, after using Equation (5.1), the starting point for the default optimization routine in the second stage is close enough to one of the local minima, and the damped least-squares method becomes stable. Figure 5.3(b) shows the resulting basins of attraction for the doublet example. These will be our references for the following runs.

The basins shown in Figure 5.3(b) are compact regions in the merit function space, and the boundaries between them are smooth curves. However, there are some minor artifacts, mainly along the basin boundaries, particularly prominent close to the upper and lower part of the boundary between the basins of minimum B and D. The artifacts are caused by the fact that we only use a finite number of steps with differential equation (5.1). When we increase the number of steps, and switch to default optimization closer to the minimum, the artifacts disappear. However, the cost for doing this is a considerable increase in computation time. Note that the saddle points in the merit function landscape are situated exactly on the basin boundaries.

5.3 Fractal basins of attraction

In this section, we demonstrate the fractal behavior of various forms of the damped least-squares method with plots of two-dimensional cuts of the basins of attraction. First, we discuss the basins of attraction of the two-dimensional monochromatic doublet optimization problem, and then we show that we obtain similar behavior for a typical design problem with seven variables.

5.3.1 Two-dimensional doublet optimization problem

When we use the program OPTSYS with the settings given in Tables 5.1 and 5.2, we obtain all five doublet local minima which are shown in Figure 5.2. Figure 5.4 shows the basins that are obtained with OPTSYS for two different values of the maximum damping factor. By changing the maximum damping factor, the shapes of the basins change as well. When the maximum damping is sufficiently large, we obtain the basins shown in Figure 5.4(a), which look quite similar to the ones shown in Figure 5.3(b).

(a) (b)

Figure 5.4: *Basins of attraction for the two-dimensional monochromatic doublet optimization problem, obtained with OPTSYS, for different values of the maximum damping factor on a grid of* 401×401 *points. The specifications of the doublet are given in Tables 5.1 and 5.2. The colors correspond to the local minima as shown in Figure 5.2. (a) Large damping, (b)* 10 *times smaller damping.*

It is known that the behavior of the damped least-squares optimization with high damping begins to resemble the behavior of the steepest descent method [12]. Therefore, it is not surprising that the basin boundaries in Figure 5.4(a) are qualitatively of the same kind as the ones shown in Figure 5.3(b). Note that the differential equation method used in Figure 5.3(b) is essentially a steepest descent with infinitesimal step size.

By decreasing the damping, the regular basin shapes change into more complicated ones with increasingly diffuse boundaries. Since less damping allows the algorithm to take larger steps in the merit function landscape, the number of ray failures increases. When ray failure occurs during optimization, OPTSYS stops the optimization process.

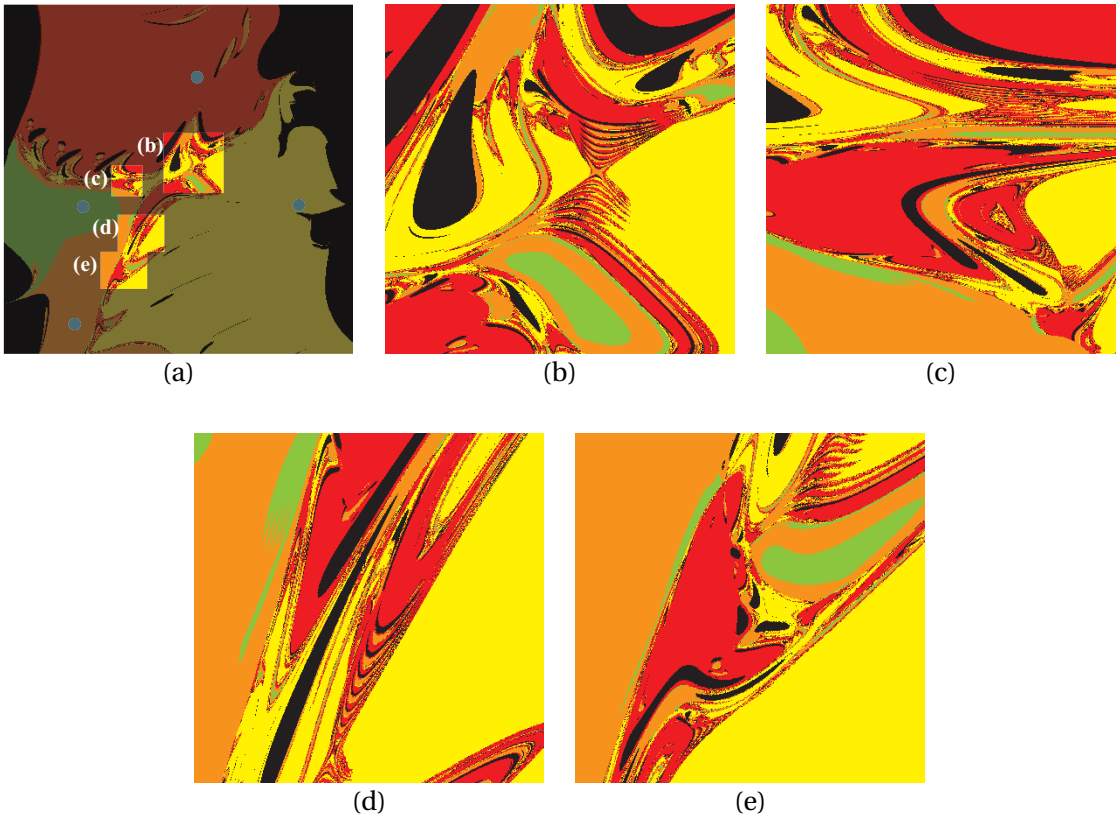


Figure 5.5: *Magnified regions of basin boundaries shown in Figure 5.4(b). Figure 5.5(a) shows the positions of these regions within the parameter domain shown in Figure 5.4(b).*

(CODE V and ZEMAX can seemingly escape from these situations.) The corresponding starting points are then drawn in black. As will be discussed in more detail in the next section, there are more black points in Figure 5.4(b) (low damping) than in Figure 5.4(a). When the maximum damping is too small, the algorithm does not converge to minimum B anymore, and the corresponding blue basin completely disappears [see Figure 5.4(b)]. This peculiar behavior will be studied in detail in Section 6.4.

Several regions of basin boundaries from Figure 5.4(b) are shown magnified in Figure 5.5. When we examine these magnified figures, we notice the presence of very fine-scale interwoven structures. By using the concept of non-integer or fractal dimension [56, 60, 62], we can quantitatively investigate an important property of all basins, and show that the basin structure is fractal. There are several ways to compute fractal dimensions. In this chapter, we use a simple method known as the box-counting method to calculate the so-called capacity dimension. A simple illustration of this method can also be found in Reference [63].

In order to compute the fractal capacity dimension D for one basin, we cover the figure with a set of grids having mesh sizes ϵ, and count for each ϵ the number of grid boxes $N(\epsilon)$, which contain at least one point of the considered basin. For small values of ϵ, the total number of grid boxes N needed to cover the basin is inversely proportional to ϵ^D:

$$N(\epsilon) = k\epsilon^{-D}, \tag{5.2}$$

where k is a constant. As ϵ is made smaller, the area of the considered basin covered

by the grid boxes approaches the true area of that basin. The capacity dimension is defined as the value of D in Equation (5.2), in the limit $\epsilon \to 0$ [56, 60, 62].

For points, lines, and surfaces, the capacity dimension is identical to the topological dimension, i.e. for a single point, $N(\epsilon) = 1$ regardless of ϵ, giving $D = 0$ [62]. For a line with length L, we have $N(\epsilon) = L/\epsilon$, giving $D = 1$, and for a surface of area S, we obtain $N(\epsilon) = S/\epsilon^2$, which yields $D = 2$.

As an example of computing fractal dimensions for basins of attraction, Figure 5.6 shows the orange basin shown in Figure 5.5(b), covered with five different mesh sizes. It is convenient to use a sequence of grids where the initial mesh size ϵ_0 is the smallest, and the remaining mesh sizes increase by a factor of 2 from one grid to the next: $\epsilon_0 = \epsilon_1/2 = \epsilon_2/4 = \ldots = \epsilon_m/2^m$, with $m = 0, 1, 2, 3, 4$, corresponding to grids with 400×400, 200×200, 100×100, 50×50, and 25×25 points, respectively. When a grid box contains at least one point of the orange basin, we color the complete box gray. The other basins are not considered and are colored white.

Let a be the length of the two sides of Figure 5.5(b), then the mesh sizes are given by $a/400$, $a/200$, $a/100$, $a/50$ and $a/25$, respectively. We can rewrite Equation (5.2) as:

$$\frac{N(\epsilon_m)}{N(\epsilon_0)} = \left(\frac{\epsilon_m}{\epsilon_0}\right)^{-D} = \left(\frac{1}{2}\right)^{mD}, \tag{5.3}$$

where the constant k in Equation (5.2) is now given by $k = N(\epsilon_0)\epsilon_0^D$. Solving Equation (5.3) for mD gives us:

$$mD = \log_2 \frac{N(\epsilon_0)}{N(\epsilon_m)}. \tag{5.4}$$

Note that when we replace ϵ by the dimensionless quantity $\hat{\epsilon}_m$:

$$\hat{\epsilon}_m = \frac{\epsilon_m}{\epsilon_0}, \tag{5.5}$$

then Equation (5.2) can be rewritten such that the constant k is absorbed in the definitions of $\hat{\epsilon}_m$ and the quantity $\hat{N}(\hat{\epsilon}_m)$:

$$\hat{N}(\hat{\epsilon}_m) = \frac{N(\epsilon_m)}{N(\epsilon_0)} = \frac{N(\hat{\epsilon}_m \epsilon_0)}{N(\epsilon_0)}. \tag{5.6}$$

With these new dimensionless quantities, we can rewrite Equation (5.2) in a form that can often be found in the literature [56, 60, 62]:

$$D = \lim_{\hat{\epsilon} \to 0} \frac{\ln \hat{N}(\hat{\epsilon})}{\ln 1/\hat{\epsilon}}. \tag{5.7}$$

Equation (5.4) shows that when we plot the base 2 logarithms of $N(\epsilon_0)/N(\epsilon_m)$ versus m, then the slope of the line that fits the data gives us an estimation for D. For all fractals studied in this thesis, the data points $\left(m, \log_2 N(\epsilon_0)/N(\epsilon_m)\right)$ turn out to be very close to a straight line. After fitting the data for the orange basin, we obtain an almost straight line with slope 1.48. The data with linear fit are shown in Figure 5.7.

By using Equation (5.4), we calculate the capacity dimensions D of all remaining basin structures shown in Figure 5.5 (listed in Table 5.3). The dimensions vary between $D =$

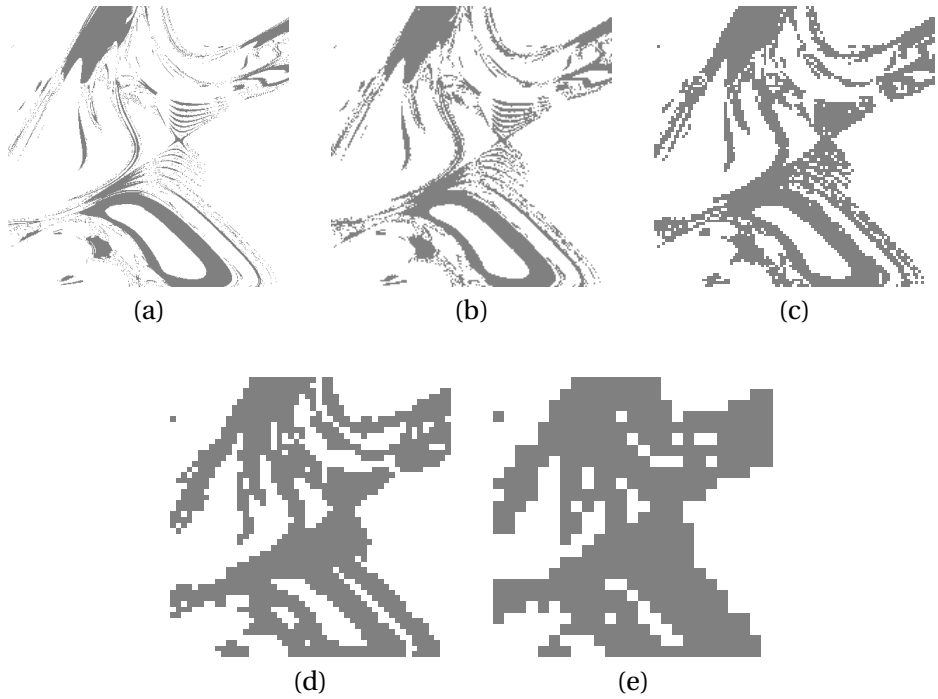


Figure 5.6: *Grid boxes that contain at least one point of the orange basin shown in Figure 5.5(b), for five different grids with (a)* 400×400, *(b)* 200×200, *(c)* 100×100, *(d)* 50×50, *and (e)* 25×25 *points. The values of* $N(\epsilon)$ *are* 23981, 8734, 3328, 1149, *and* 375, *respectively.*

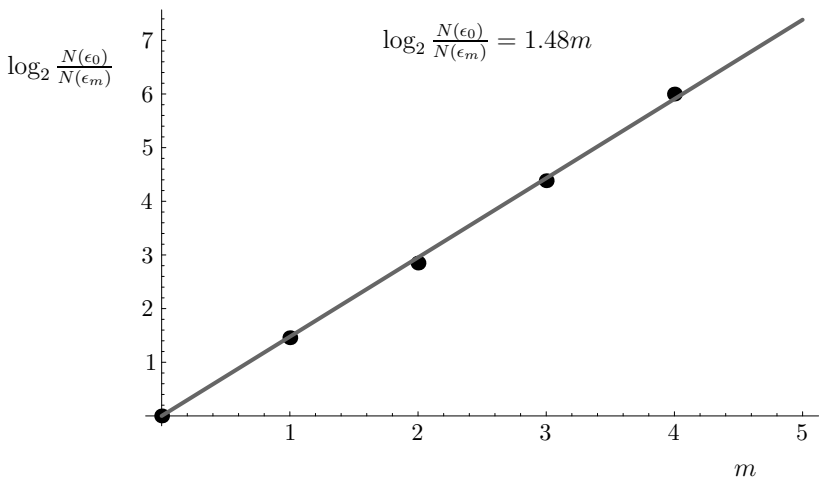


Figure 5.7: *Calculation of the capacity dimension D for the orange basin shown in Figure 5.5(b). The slope of the straight line, which fits the data, gives* $D = 1.48$ *[see Equation (5.4)].*

1.16 and $D = 1.87$. The fact that D is not an integer (or very close to an integer) demonstrates that the basin boundaries are fractal [56, 60].

Note that for the basins of attraction of a given local minimum, the value of D for the entire figure is different from the value of D for the individual parts of the figure. This shows that the exact value of the fractal dimension also depends on the specific choice

Table 5.3: *Capacity dimensions of the basins of attraction shown in Figures 5.5(a)–(e). The letters A, C, D, and E correspond to the basins of the corresponding local minima, and 'RF' refers to starting configurations for which ray failure occurs during optimization. Recall the basin of minimum B has disappeared in Figure 5.5.*

	5.5(a)	5.5(b)	5.5(c)	5.5(d)	5.5(e)
A	1.86	1.65	1.74	1.65	1.69
C	1.87	1.75	1.66	1.83	1.81
D	1.60	1.48	1.60	1.66	1.77
E	1.56	1.25	1.16	1.16	1.32
RF	1.71	1.51	1.35	1.34	1.21

of the size of the variation domain of the starting point parameters (c_2 and c_3 in this case). Therefore, the emphasis in this chapter is on the fact that the fractal dimension of the basins we examine is not an integer, rather than on the specific value of the fractal dimension in these cases. The fractal nature of the basins implies that the same complex mixture of basins shown in Figure 5.5 is also present on arbitrarily small scales. The line in Figure 5.7 can be extrapolated to the left and right as much as desired.

For the same doublet problem, we also examined the basin boundaries obtained with commercial optical design software. We used CODE V and ZEMAX with their default damped least-squares optimization algorithms and merit functions. In CODE V, for each initial set of curvatures, the system was iterated with a maximum of 100 optimization cycles (in ZEMAX with 30 cycles). For the settings we use here, minimum E (see Figure 5.2) does not exist in the merit function landscape of CODE V or ZEMAX.

The resulting basins of attraction for the two-dimensional doublet optimization problem are shown in Figure 5.8. In Figures 5.8(a)–(e), we used CODE V on a grid of 1001×1001 points, and we used ZEMAX on a grid of 401×401 points in Figure 5.8(f). The lens data of the doublet in Figures 5.8(a)–(c) and (f) are given in Tables 5.1 and 5.2. In Figure 5.8(d), the thickness of the second lens is changed to 0.3 mm. In the doublet used for Figure 5.8(e), all distances are changed to zero, and both glasses are set equal ($n = 1.717$) [1]. Figures 5.8(b)–(c) show magnified regions of Figure 5.8(a), indicated with white rectangles, where we zoom in on some parts of the basin boundary between minima A and D. Note the very complex shapes and the fine-scale structure in the basins of attraction, which is present on all scales. The basin boundaries show very finely interwoven basins, indicating the fractal nature of the boundaries. When we compare Figures 5.8(a), (d)–(f) (obtained with two commercial programs), with Figure 5.4(b) (obtained with OPTSYS), we observe certain similarities. However, since the two commercial programs can handle ray failure, if it occurs during optimization, better than OPTSYS, Figures 5.8(a), (d)–(f) have less black points than Figure 5.4.

The capacity dimensions D of the basins of attraction shown in Figure 5.8 are listed in Table 5.4. Except the blue basin in Figure 5.8(e), all basins have a capacity dimension between 1 and 2, indicating the fractal nature of the basin boundaries. In the regions where the compactness of a basin increases, its capacity dimension also increases. The blue basin of the thin-lens doublet hub in Figure 5.8(e) has a capacity dimension close

[1] This thin-lens doublet was also used as an example in Section 4.5.

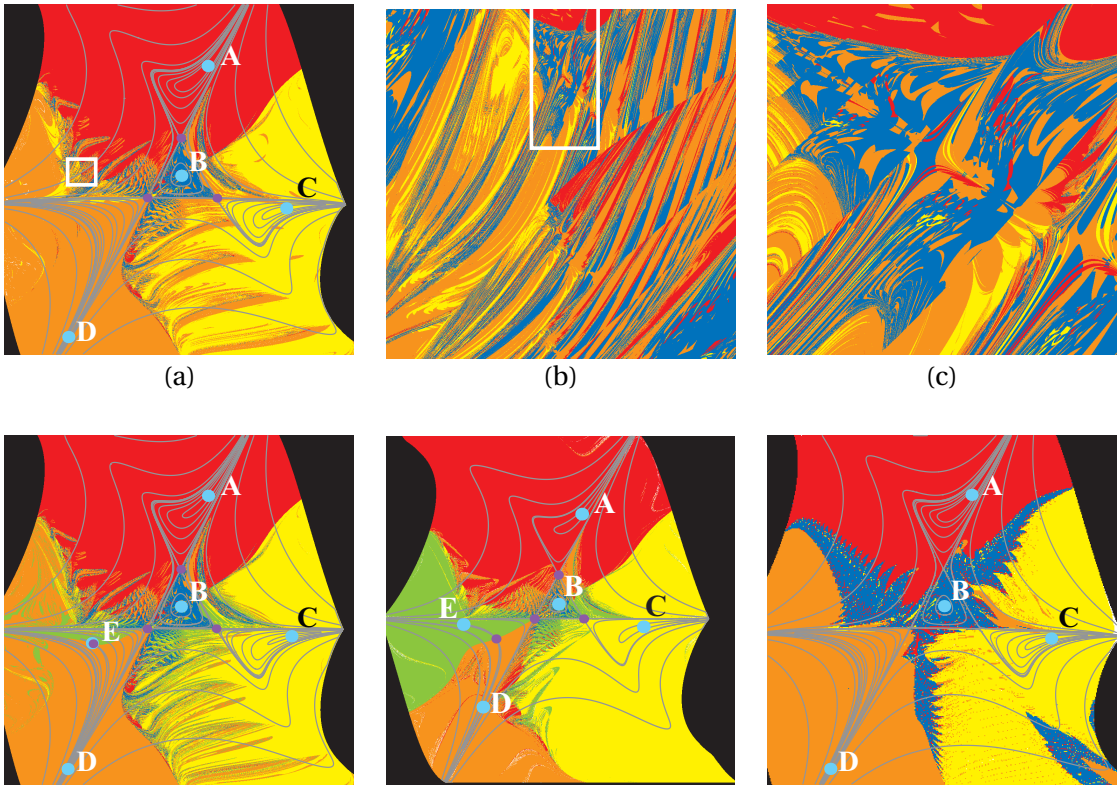


Figure 5.8: *Basins of attraction for the two-dimensional monochromatic doublet optimization problem obtained with CODE V (a–e) and Zemax (f). (The gray contours are the equimagnitude contours of the merit function.) (a) Doublet with settings as given in Tables 5.1 and 5.2 and default optimization of CODE V. (b) Magnification of the small white rectangle in a). (c) Magnification of the small white rectangle in b). (d) Same as a), but with the thickness of the second lens equal to* 0.3 *mm. (e) Same as a), but all distances are equal to zero, and both glasses are equal (n =* 1.717*). (f) Same as a), but with the damped least-squares optimization of ZEMAX.*

to 1. Note the remarkable symmetry that is present in Figure 5.8(e).

Figure 5.8 shows a substantial difference between the basin shapes obtained with the algorithms implemented in the two commercial programs and our much slower algorithm based on Equation (5.1). In this example, the finely interwoven fractal basins reveal the existence of extreme sensitivity to initial configurations in the default optimization routines of CODE V and ZEMAX, which causes unpredictable results in the optimization process. A small change in the starting configuration of a system, which is initially in a fractal region, can lead to a different local minimum after optimization.

5.3.2 Seven-dimensional optimization problem

The fractal behavior shown in the doublet example was also frequently found in basin plots for systems with more than two variables. As an example, we show the basins of attraction, obtained with CODE V, for a typical optimization problem with seven vari-

Table 5.4: *Capacity dimensions of the basins of attraction shown in Figures 5.8(a)–(f). The letters A–E correspond to the basins of the local minima.*

	5.8(a)	5.8(b)	5.8(c)	5.8(d)	5.8(e)	5.8(f)
A	1.88	1.32	1.69	1.89	1.87	1.77
B	1.32	1.60	1.76	1.30	0.94	1.69
C	1.80	1.26	1.45	1.77	1.89	1.81
D	1.81	1.67	1.69	1.71	1.74	1.74
E	-	-	-	1.24	1.73	-

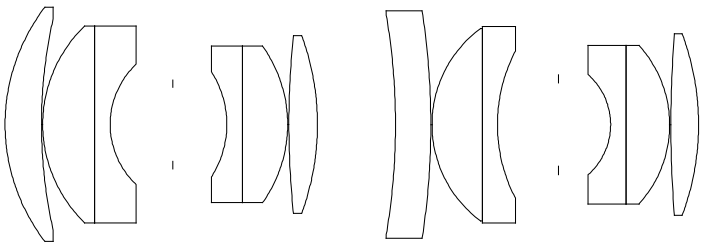
ables, where the goal is to design a Double Gauss system (without vignetting), see Figure 5.9(a).

We use the default merit function of CODE V with all lens curvatures as optimization variables, except for the curvatures of the two plane cemented surfaces, which are kept unchanged, and the last curvature, which is used to keep the effective focal length constant.

In Figure 5.10, we show a two-dimensional cut of the seven-dimensional basins of attraction for this optimization problem. We choose a two-dimensional set of starting points, for which the curvatures of the first and second surfaces are varied within a certain range. The starting values for the other variables are identical for all these points.

We observe the presence of three local minima. In addition to the expected Double Gauss solution (the red basin in Figure 5.10), we also obtain the minimum shown in Figure 5.9(b) (the green basin) and a third minimum (the blue basin). The third minimum resembles the system shown in Figure 5.9(b), except that the first lens has stronger curvatures. The Double Gauss design has a merit function value approximately 30 times smaller than the values for the other two systems.

Note that the basins of attraction for the minima again have irregular shapes. The fractal dimensions computed for the red, green, and blue basins are: 1.87, 1.85, and 1.88, respectively. Although not shown in this work, we have computed fractal dimensions for basin boundaries of optimization problems having an even larger number of variables than in the examples shown here.



(a) (b)

Figure 5.9: *(a) Double Gauss system (without vignetting). (b) Another local minimum which is found after optimizing a two-dimensional set of starting points (see text).*

Figure 5.10: *Basins of attraction for a seven-dimensional Double Gauss optimization problem. The initial curvature of the first surface is plotted vertically, and the initial curvature of the second surface is plotted horizontally.*

5.3.3 Saddle-Point Construction example

With the present implementation of local optimization in existing software packages, the fractal behavior of the basin boundaries can sometimes also be observed in the neighborhood of null-element saddle points (NESP's), which are made with the Saddle-Point Construction (SPC) method discussed earlier in Chapter 3. Here, we show part of the basins for two monochromatic quartet NESP's (f number 2), which are constructed by adding a null-element meniscus in contact with the first and second lens of the triplet shown in Figure 5.11(a), where the second lens of the quartet NESP's is the inserted null-element. The two NESP's have the properties $c_3 = c_4 = c_2$, and $c_3 = c_4 = c_5$, respectively. All specifications are given in Tables 5.5 and 5.6. The last curvature of the quartet is used to keep the effective focal length constant at 100 mm, and the remaining seven curvatures are used as variables. The stop is placed at the first surface. The distance between the last surface and the image has a paraxial image distance solve.

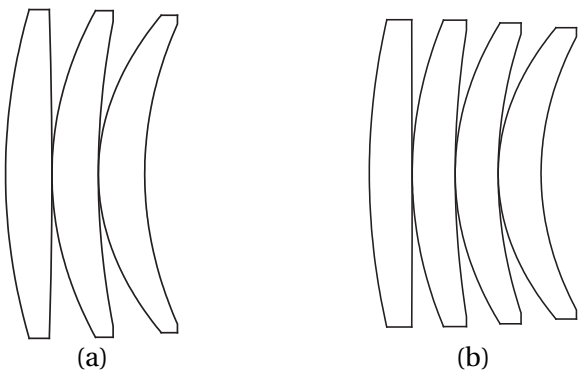


Figure 5.11: *(a) Starting triplet minimum for SPC. (b) Quartet minimum (Fulcher-type design), which corresponds to the red basins shown in Figure 5.12 (after increasing the thickness of the second lens).*

In Figure 5.12(a), we show a two-dimensional cut of the merit function landscape near the saddle point with the property $c_3 = c_4 = c_2$. The starting points are all in the plane

Table 5.5: *Specifications for the monochromatic quartet example.*

Entrance pupil diameter (mm)	50.00000
Effective focal length (mm)	100.00000
Paraxial image height (mm)	3.49210
Fields points (deg)	0.0000, 1.41452, and 2.00001

Table 5.6: *Curvatures, thicknesses, and refractive indices of the starting points for the monochromatic quartet example. The variable curvatures are indicated with 'V', and solves are indicated with 'S'. When $c_3 = c_4 = c_2$, or $c_3 = c_4 = c_5$, the obtained configuration is a NESP.*

Surface #	Curvature (mm^{-1})		Thickness (mm)		Refractive index
object	0.00000		infinity		1.
1	0.00576	V	4.00000		1.51680
2	-0.00065	V	0.00000		1.
3	c_3	V	0.00000		1.51680
4	c_4	V	0.00000		1.
5	0.01061	V	4.00000		1.51680
6	0.00421	V	0.00000		1.
7	0.01584	V	4.00000		1.51680
8	0.00951	S	93.06504	S	1.
image	infinity		0.0000		1.

defined by the invariant lines $c_2 = c_3$ and $c_3 = c_4$ (the plane OAB in Figure 3.5). The variation domain for the two coordinates x' and y' is $(-0.005, 0.005)$ [see Equation (3.20)]. Since the distances between the surfaces 2, 3 and 4 are equal to zero, the value of the merit function does not change along the lines $c_2 = c_3$ and $c_3 = c_4$. As we have shown in Chapter 3, the saddle point is exactly at the crossing of the two lines. Along both invariant lines, the variation domain of the corresponding two equal curvatures is $(-0.0047, 0.0034)$.

The basins in the figure on the left of Figure 5.12(a) are obtained by using OPTSYS with sufficient damping. The boundary between the two basins is a smooth curve through the saddle point. Near the basin boundary, there are points that lead after optimization to a third local minimum, indicated with purple. By using CODE V, the resulting shapes of the basin boundaries drastically change into much more complicated ones, as shown in the figure on the right. Besides the change in shape, also the number of basins increase.

A similar behavior can be seen in Figure 5.12(b), where we examine the basins near the other saddle point (when $c_3 = c_4 = c_5$). In this case, we consider the plane defined by the two invariant lines $c_3 = c_4$ and $c_4 = c_5$. Along both invariant lines, the variation domain of the corresponding two equal curvatures is $(0.0065, 0.015)$.

Figure 5.11(b) shows the minimum corresponding to the red basin, after increasing the thickness of the second lens. The shape of this minimum is very similar to the shape of the monochromatic quartet, for which Fulcher has derived the curvatures several

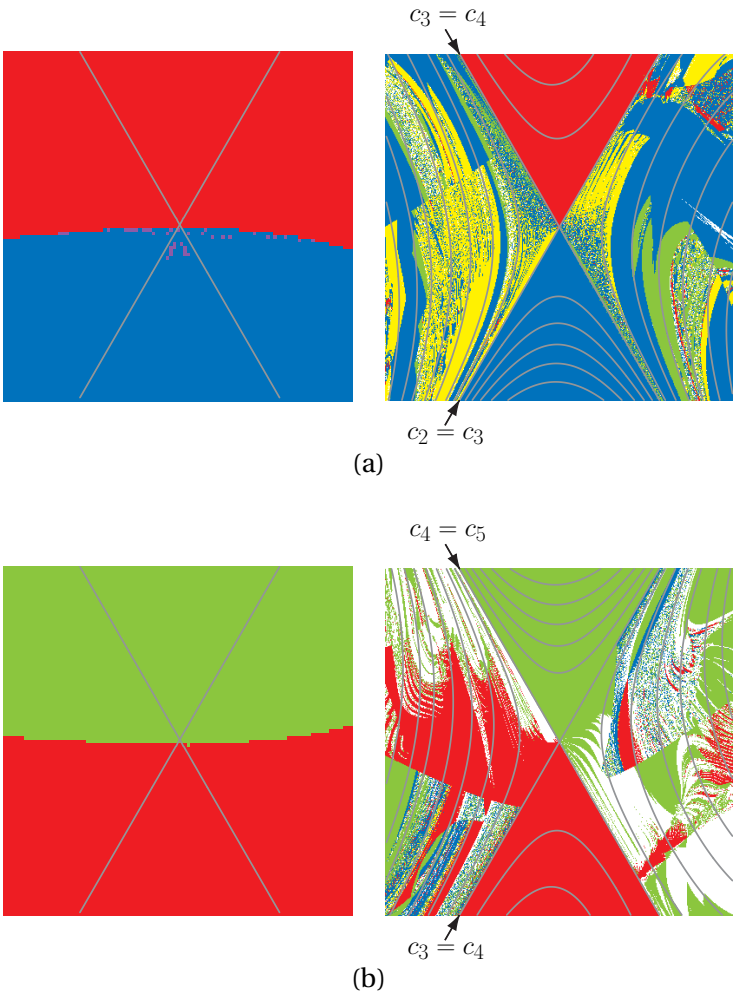


Figure 5.12: *Basins of attraction for the 7-dimensional monochromatic quartet optimization problem (the settings are given in Tables 5.5 and 5.6), shown in a two-dimensional cut of the merit function landscape obtained with OPTSYS (left), using sufficient damping, and CODE V (right). White points indicate configurations that have poor convergence. (a) Near the saddle point with the property $c_3 = c_4 = c_2$. (b) Near the saddle point with the property $c_3 = c_4 = c_5$.*

Table 5.7: *Capacity dimensions of the basins of attraction shown in Figure 5.12.*

	5.12(a)	5.12(b)
red	1.69	1.73
green	1.36	1.63
blue	1.74	1.23

decades ago [46] (see also Figure 3.2). Although the merit function value of the quartet in Figure 5.11(b) is approximately twice as large as the values for the other quartet local minima (indicated with different basin colors), this Fulcher-type design has almost vanishing third-order spherical aberration. The capacity dimensions of the red, green, and blue basins are listed in Table 5.7.

When we perturb the null-element in the NESP as described in Section 3.4 [either with

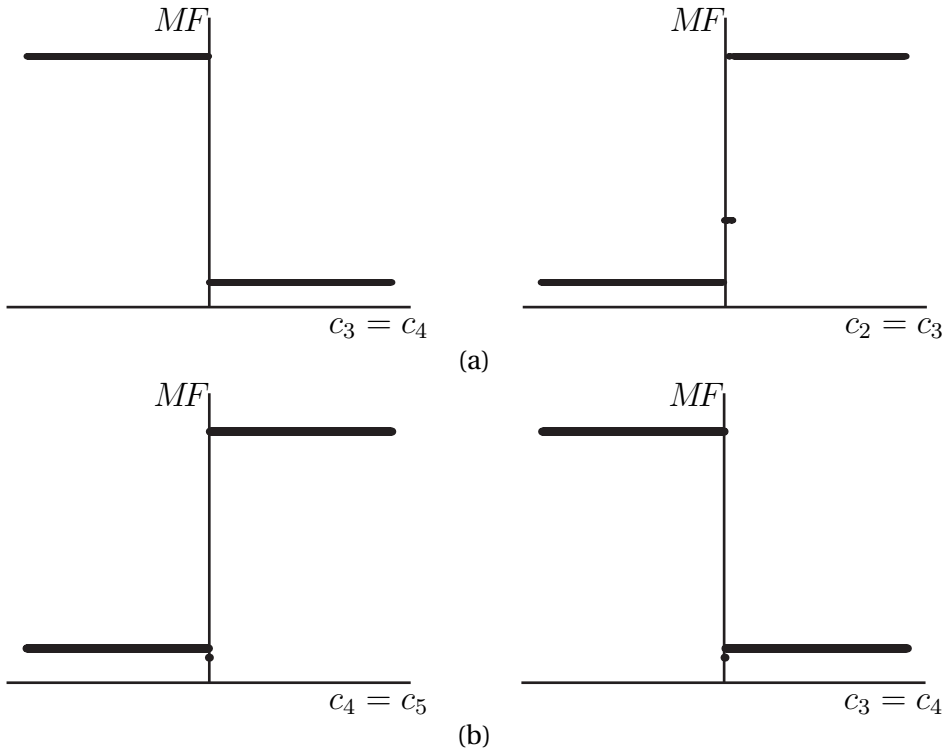


Figure 5.13: *Values of MF after locally optimizing 400 points along both pairs of the invariant lines (the two lines that cross each other at the saddle point) shown in (a) Figure 5.12(a), (b) Figure 5.12(b).*

Equation (3.28) or (3.29)], it is not clear from Figure 5.12 which local minima we obtain after optimizing the perturbed points. The basin boundaries of configurations after SPC typically coincide with the two invariant lines, as shown in Figure 5.12. Fortunately, in many cases, optimizing the perturbed points results in a predictable optimization behavior. By examining the basins of attraction along both pairs of invariant lines in Figure 5.12, we only notice some irregularities in the near vicinity of the saddle point (see Figure 5.13), which corresponds to a small curvature change (ϵ) of $\epsilon \leq 0.0002\text{mm}^{-1}$.

For avoiding the undesirable behavior shown in the basin plots on the right in Figure 5.12, increasing the damping during the first few cycles in local optimization started on both sides of the NESP would solve the problem without significantly increasing the computational effort. The starting points for the optimization after SPC are then far away from the basin boundaries. As shown in the basin plots on the left in Figure 5.12, higher damping changes the orientation of the basin boundary significantly.

5.4 Instabilities in optimization

To understand why basin boundaries can be so complicated, we examine the sequences of steps in the two-dimensional variable space of the monochromatic doublet that are taken by the optimization routines of CODE V, OPTSYS, and ZEMAX. Examples of optimization paths are shown in Figure 5.14. For observing details better, these figures can

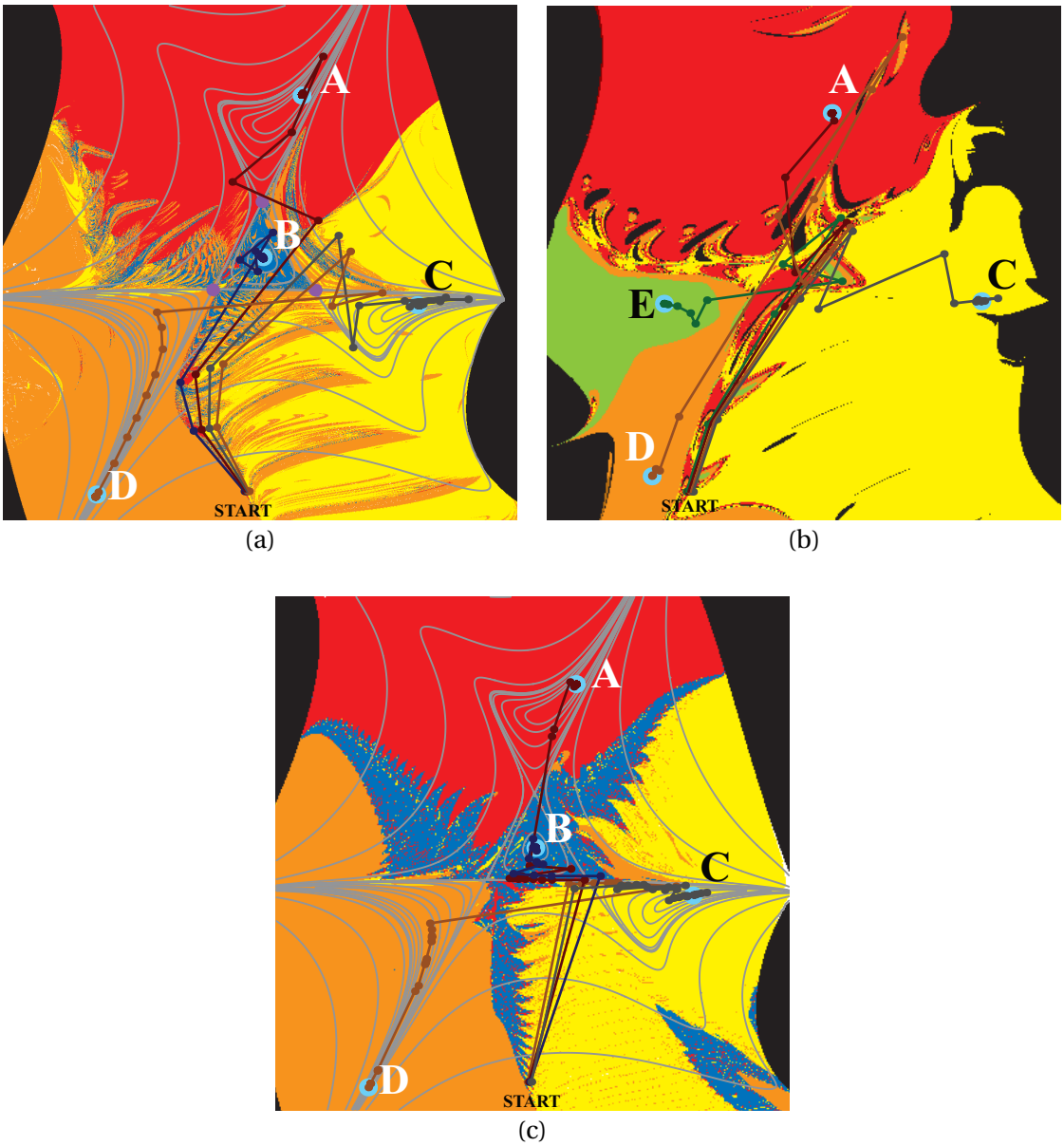


Figure 5.14: *Optimization paths (dark colored lines) in the variable space of the doublet, obtained with (a) default optimization of CODE V, (b) OPTSYS, (c) damped least-squares optimization of ZEMAX. Four starting points for optimization are chosen very close to each other at the position 'START', and the iterations are shown as dots.*

be enlarged on-screen in the electronic version of this thesis. The starting points for optimization are indicated with the label 'START', and the result after each optimization cycle is shown in the figure by a dot. The dots are connected with lines, showing the direction of the optimization path in the merit function landscape.

In Figure 5.14, we see a remarkable behavior of optimizations started in a fractal region. In all three examples, the four starting points are very close to each other, but they converge to four different local minima after optimization. The iterations are first attracted, and then repelled by equimagnitude lines that pass through two of the three saddle points that are present in the merit function landscape. The behavior in the

vicinity of those lines is similar to the behavior in the vicinity of the chaotic saddle (see Section 6.4). Before reaching the attracting-repelling lines, the distance between the trajectories increases rapidly with the number of iterations. The extreme sensitivity to initial conditions occurs because, when they have reached the attracting-repelling lines, the trajectories are already far apart from each other and therefore, after the trajectories are repelled, they converge towards different local minima, despite of the fact that they have their origin in almost the same point. This type of behavior is known as a chaotic transient [56].

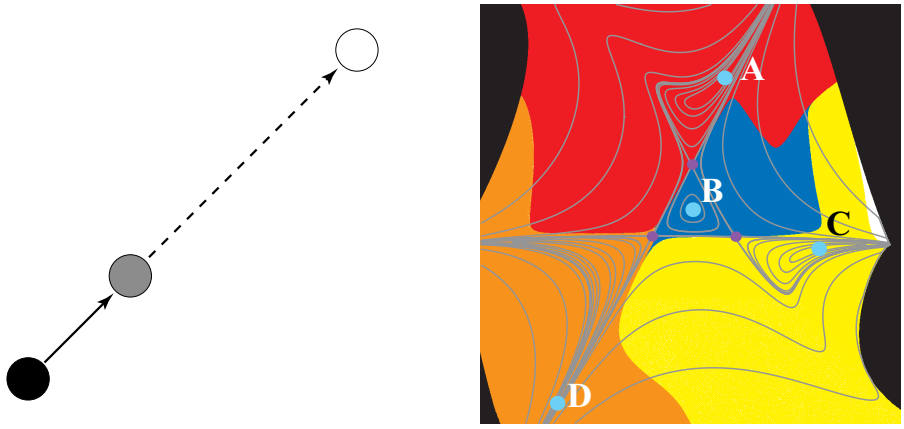
Note that the sequence of iterations always remains inside the same basin as for the starting point, independent of the basin shape. For example, when the starting point is in the orange basin, all following iterations are also in the orange basin, and the iterations will finally lead to the orange minimum (minimum D). If the basins have a fractal structure, it can sometimes be difficult to notice this property.

The attracting-repelling line that plays the key role in Figure 5.14(c) is the almost horizontal *MF* equimagnitude line. The large number of iteration points in the middle of the figure indicates the presence of a horizontal attracting-repelling line there. For each point, the optimization iterates a number of small steps close to this line, until it is suddenly repelled by the line, and converges to one of the four local minima. In Chapter 6, we will study this phenomenon in more detail.

In the middle of Figure 5.14(b), where the curvatures have moderate values, we encounter unexpected ray failure (black points). This is because the jumps are so large, that for some iterations the solution reaches the compact black regions (close to the borders of the figure) where ray failure occurs (due to large curvatures).

The fact that in large domains of the parameter space, arbitrarily close starting points can converge to different solutions, can be undesirable in practical design work. It is therefore worth investigating how the behavior of optimization can be stabilized. The irregular character of the paths shown in Figure 5.14 and the similarity between the basins obtained with OPTSYS [Figure 5.4(b)], and the basins obtained with CODE V and ZEMAX [Figures 5.8(a), (d)–(f)], indicate that the damping, which is used in these programs, is too low for stability in this example. The results shown in Figure 5.4 suggest that in order to make the boundaries more regular, a higher damping factor in the optimization algorithm is needed. However, in the versions of CODE V and ZEMAX available at the time of this writing, the user does not have the possibility to influence the damping factor, which is computed automatically.

To prevent the optimization algorithm from jumping unpredictably, we have implemented an external damping factor in the macro language of CODE V. Our external damping factor does not change the direction of the optimization step, but it decreases the step size. The direction of the variable change at each iteration is still determined by the optimization algorithm. Our procedure is illustrated in Figure 5.15(a). Note that this technique is not equivalent to a change of the automatic internal damping of the optimization algorithm, but it removes the instabilities in the optimization process. For a discussion of a typical way to implement internal damping, which is to be preferred to external damping, see for example Reference [12]. Our macro language implementation of external damping is about 6 times faster than our differential equation macro, but about 20 times slower than the standard optimization that is internally implemented in



(a) (b)

Figure 5.15: *(a) Illustration of the external damping procedure. The starting point is shown in black, and the result obtained after one optimization step with CODE V is shown in white. The result after external damping is shown in gray. The external damping factor damps the optimization in such a way that it does not change the direction, but only shortens the step size. (b) Basins of attraction for the doublet after applying our external damping to the default optimization method of CODE V.*

CODE V.

Figure 5.15(b) shows the basins of attraction in CODE V for the doublet with high external damping. In comparison with Figure 5.8(a), the basin boundaries in Figure 5.15(b) are regular and the basins are compact, which confirms our hypothesis that in Figure 5.14(a) the chaotic transients are caused by a damping that is too low to produce a stable result. However, unlike internal damping, adjusting the external damping does not lead to basin shapes that are similar to the reference basins shown in Figure 5.3(b).

When we compare Figures 5.3(b), 5.8(a), and 5.15(b), we observe that the saddle points in the merit function landscape are unaffected by the choice of algorithm or damping. Although the boundary shapes vary considerably when the optimization algorithm or the applied damping are altered, the saddle points always remain points on the basin boundaries.

5.5 Conclusion

A fundamental way to understand when optimization leads to a certain local minimum rather than to a neighboring one, is to look at the basin of attraction for that minimum. The examples discussed in this chapter show that the basin structure and boundaries depend on the implementation details of the algorithm that is used and on numerical parameters, such as the damping factor. However, the saddle points in the merit function landscape are always on a basin boundary.

We show that, in certain situations, damped least-squares optimization methods can have an unpredictable behavior. In our examples, there are regions in the merit function landscape where starting points, which are very close to each other, lead to differ-

ent local minima after optimization. This instability illustrates an extreme sensitivity to change of initial configurations. Computations of the capacity dimension then show that the basin boundaries have a fractal shape. When optimization is started in a fractal region, it first displays a chaotic transient path in the merit function landscape, before it finally converges to a local minimum.

By using damped least-squares methods with sufficient damping, we can obtain smooth basin boundaries even in optimization problems that have a natural propensity for fractal basins. Depending on the specific optimization problem and optimization method that is used, instabilities may be absent or present in different degrees. If these instabilities are present, then the main cause for them is to be found in the optimization algorithm.

The concept of damping factor was originally introduced with the purpose of limiting the step size during optimization, so that certain mathematical approximations in the least-square algorithm remain valid. Many optimization algorithms presently use much lower values of the damping factor. Although the exact details may vary from one program to another, the automatic choice of the damping factor is typically made such that the merit function decrease per optimization step is maximized. However, as we have shown, the gain in computation speed is at the cost of stability. Strategies for choosing the damping factor, which turn out to be optimal close to local minima, are not necessary optimal far away from local minima or near the basin boundaries.

When designers adapt an existing design to new specifications, they would usually prefer to obtain a system shape and correction properties that are similar to ones that are already known. A higher degree of predictability is also desirable when one attempts to reach the design goal in several stages, and at each stage, one wants to preserve what has been achieved in previous stages. When the outcome is not the expected one, it is usually believed that the cause lies in the inherent complexity of the design landscape. This is indeed a major source of unpredictability. As we have shown in this chapter, even in a very simple optimization problem with only two variables, we find no less than 4-5 solutions. In addition, the number of local minima increases rapidly with the number of optimization variables. However, here we show another possible cause. One of the goals of this thesis is to make optical system designers aware of the possibility of the presence of instabilities in the optimization process. Sensitivity to initial conditions can influence the result that will be obtained after optimization. While the inherent complexity of the design landscape limits the sizes of the basins of desirable solutions, instabilities in optimization, if present, further increase complexity and decrease predictability by mixing basins of desirable and undesirable solutions together.

Chapter 6

Chaotic behavior in an algorithm to escape from poor local minima in lens design

6.1 Introduction

Optical system designers are usually confronted with many local minima in the merit function landscape of their design problem. A major challenge is to find good solutions among these minima. Increasingly elaborate and powerful global optimization methods [19–25, 42] can provide a remedy when local optimization produces an unsatisfactory solution. However, simpler empirical techniques can also be effective for escaping from poor local minima, especially when the merit function barrier that must be overcome is low. Such techniques include reoptimization after making small changes to lens parameters, changing weights in the merit function, switching merit functions during local optimization, and modifying the conditions under which local optimization algorithms operate, for example by modifying the damping factor or even changing the method of local optimization [11, 26, 27].

In the past decades, various versions of the damped least-squares algorithm have been successfully used in optical system design [11, 12]. If used in the conventional way, damped least-squares algorithms search only locally for a minimum and are therefore computationally much less expensive than global optimization methods.

In this chapter, we use low damping in a damped least-squares algorithm as an empirical technique to escape from poor local minima, and we temporarily allow configurations with higher merit function than for the initial configuration. For such empirical techniques, simplicity is their main asset and raison d'être, given the fact that they are not as powerful as global optimization methods. We have chosen this particular strategy because it was very simple to implement it in an existing code.

Since we allow the merit function to increase during the search process, the algorithm does not even always converge and more complicated scenarios including chaos are also observed in the asymptotic regime. A simple example (the monochromatic doublet with two variables, which has also been used in Chapter 5) will show that when this

strategy does lead to a better configuration on the other side of a merit function barrier, this does not happen in a straightforward way. The new configuration will be obtained after a complex process including a chaotic transient. Because of the novelty within the framework of optical design methodology of the mechanism we have observed, the emphasis in this chapter will be on the anatomy of this complex mechanism, rather than on the practical aspects of such a method.

The change of the damped least-squares algorithm discussed here is also useful for a different, but not less important purpose. In Chapter 5, we have shown unexpected instabilities in the behavior of damped least-squares local optimization, even in computer programs that are widely used at present. A better understanding of these instabilities is necessary in order to assess whether, or in what degree, these instabilities affect design productivity. We will show that the present algorithm change facilitates the study of certain features of these instabilities.

Our modified implementation of the Levenberg-Marquardt damped least-squares algorithm is shortly described in Section 6.2. In Sections 6.3 and 6.4, we describe the behavior of this algorithm when different values of the damping factor are used. Section 6.3 shows how a chaotic attractor is formed, and Section 6.4 shows that a successful escape from a poor local minimum to a better one is associated with destruction of a chaotic attractor via what is called a 'crisis' in chaos theory. In Section 6.5, we use the obtained results for a better understanding of the instabilities observed in the behavior of conventional damped least-squares local optimization when used with low damping.

6.2 Algorithm with adaptive damping for searching beyond local minima

In this work, we use the program OPTSYS written by Joseph Braat, in which the Levenberg-Marquardt damped least-squares method is implemented as described in detail in Reference [64]. We first describe briefly the conventional use of this code for local optimization. During optimization, the damped least-squares algorithm dynamically controls the damping factor. The algorithm uses two loops to step through the merit function landscape. At every optimization step, an outer iteration loop calculates the Jacobian $\mathbf{J}$ of the optimization problem to find the search direction in which the value of merit function decreases. This matrix $\mathbf{J}$ is subjected to a singular value decomposition (SVD) with complete orthogonalization of both the operand space (dimension m) and the variable space (dimension n).

An inner loop uses the known matrices to optimize the damping factor λ. Successive linear damped solutions are then obtained by solving the system

$$(\mathbf{J} - \lambda_k \mathbf{I})\ \Delta\mathbf{x}_k = -\mathbf{f}, \tag{6.1}$$

for different values of λ_k, where the index k denotes the cycle number of the inner loop; $\Delta\mathbf{x}_k$ is the excursion from the current point in the variable space, $\mathbf{I}$ is the identity matrix, zero padded where needed given the dimensions (m, n) of the system and $\mathbf{f}$ is the vector of operand values to be reduced to zero or certain target values. The vector of operand

values is extended when necessary with the weighted values of constraint violations. For more details, see Reference [64]. Note that Equation (6.1) can be transformed into Equation (2.12). First, we have to ignore the damping term $\lambda_k \mathbf{I}$, and multiply both sides of Equation (6.1) with $\mathbf{J}^T$. Then, in the resulting formula, we include the damping term $\lambda_k \mathbf{I}$ again, but now $\mathbf{I}$ is an identity matrix with dimensions (n, n).

Because every optical design program has his unique way of choosing the damping method, there is a large diversity in the implementation details of local optimization [11]. In OPTSYS, the choice of the damping factor λ_k is done as follows. For the k-th iteration in the inner loop, λ_k is given by:

$$\lambda_k = p\, S_1\, 10^{-(k-1)/a}, \tag{6.2}$$

where S_1 is the largest singular value delivered by the SVD and p and a are values to be provided by the user. In this work, we have chosen a fixed value $a = 10$, but the value of p has been varied. In the following two sections, p will be the control parameter that determines whether the algorithm converges or whether its behavior is more complex. During optimization, the algorithm monitors the decrease in the squared value $|\mathbf{f}_k|^2$ of the merit function as a function of the counter k and stops the iteration of the inner loop when the merit function starts increasing again. For a one-dimensional problem, the whole procedure of choosing the most optimal damping factor is shown in Figure 6.1.

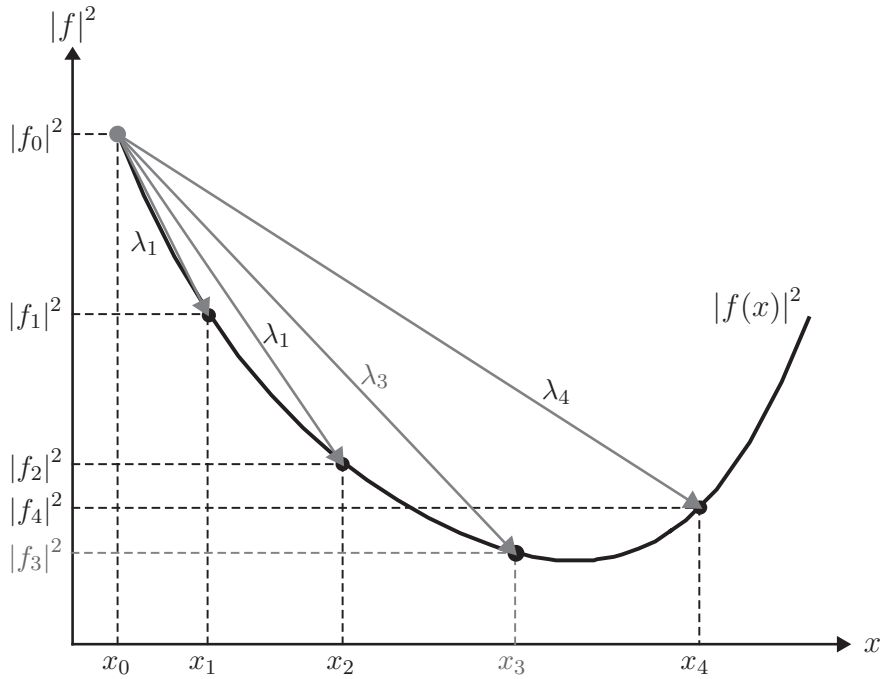


Figure 6.1: *Optimization of the damping value λ in OPTSYS for a one-dimensional problem. The quasi-minimal linear solution x_3 is used as a starting point for the next (outer) optimization cycle.*

To keep the calculation time short, the algorithm does not carry out an interpolation to find the solution $\mathbf{x}$ corresponding to the exact local minimum of the merit function. The quasi-minimal linear solution $\mathbf{x}_k$ is used to update the system variables. In the

one-dimensional example shown in Figure 6.1, this solution is obtained for $k = 3$. The new working point is then used to obtain the Jacobian that serves for the next (outer) optimization cycle.

A slight change in the conventional form of this algorithm is sufficient to overcome, for low damping, the merit function barrier that separates a poor local minimum from a better one, as will be seen in Section 6.4. We iterate the inner loop at least two times and we allow the merit to increase, but only in the first step of the inner loop. After the first two iterations, the process is stopped when $|\mathbf{f}_{k+1}|^2 > |\mathbf{f}_k|^2$ [1].

6.3 Period doubling route to chaos

When we want to escape from a poor local minimum and achieve convergence towards a different, hopefully better one, we first have to destabilize the convergence towards the original local minimum. An intermediate stage, which is discussed in this section, is the formation of a chaotic attractor. Depending on the starting configuration, the stages between convergence and chaos can be different. For simplicity, in this section we illustrate the well-known period doubling route to chaos.

We examine the behavior of the algorithm described in the previous section for a simple two-dimensional monochromatic doublet optimization problem (f number 3, field of 3 degrees), with the curvature of the second and third surfaces (c_2 and c_3, respectively) as variables. More details about this doublet optimization problem are given in Chapter 5. For sufficiently large damping, optimizing the two variable curvatures results in five different local minima (A, B, C, D, and E), which are shown in Figure 5.2. In Section 6.4, we will show an example in which we escape from a poor local minimum of this doublet. The observed behavior depends on three factors: the choice of the initial configuration, the value of the damping parameter p, and the number of iterations that have been performed. Therefore, in this chapter, we present three sorts of numerical results, in which one of these factors is varied.

In this section, we are especially interested in the result of the iteration process in the asymptotic regime (i.e. after a large number of outer iterations). When p is sufficiently large, the algorithm is convergent, as expected. For our example, in order to show which configuration leads to a certain local minimum, we compute the basin of attraction for that local minimum. (The set of all starting configurations that are attracted to a local minimum is the basin of attraction for that minimum [56, 61], see also Chapter 5.)

Figure 6.2 shows the basins of attraction for $p = 0.002$ [2]. In this and in the following figures of the same kind, curvatures c_2 and c_3 are plotted along the vertical and horizontal axis, respectively. We find five different local minima (see Figure 5.2), indicated by blue

[1] In the inner loop, the modified algorithm does not monitor the starting value of the merit function. After the first (inner) iteration, the merit function is calculated, and this will be used as the 'starting' value. Because the algorithm does not compare the merit function value after the first (inner) iteration with that of the initial system, the inner loop always performs at least two iterations. When the merit function value after the second (inner) iteration is higher than the value obtained after the first (inner) iteration (the 'starting' value), the algorithm stops the iteration of the inner loop, and uses pS_1 as damping factor.

[2] We have already shown the basins of attraction for $p = 0.002$ on a grid of 401×401 points in Figure 5.4.

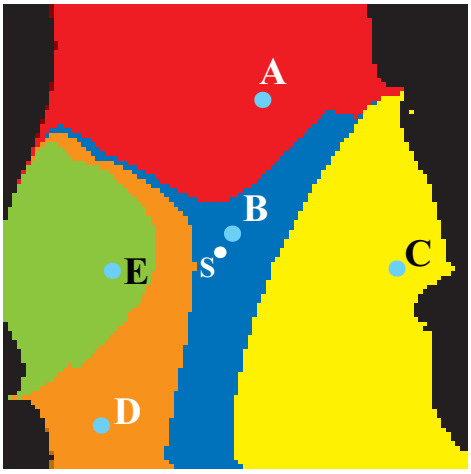


Figure 6.2: *Basins of attraction for the five local minima of the two-dimensional monochromatic doublet optimization problem, obtained with damping parameter* $p = 0.002$ *on a grid of* 101×101 *points. Each grid point is iterated* 999 *times. The starting point 'S' will be used in Section 6.4.*

dots and the letters A, B, C, D, and E. Compared to the other minima, minimum B has a high merit function value.

The basins of attraction are computed by using a grid of 101×101 equally spaced starting points in the two-dimensional variable space, and at each grid point, we start to optimize the corresponding configuration, until it arrives in one of the local minima. To each of the five local minima, we associate a different color: red, blue, yellow, orange, and green to local minima A, B, C, D, and E, respectively. Depending on which local minimum we obtain, we color the starting point with the corresponding color for that minimum. The location of the point shows the initial value of the variables, and its color shows to which local minimum it converges to. Starting points for which the initial configuration suffers from ray failure (a ray misses a surface or total internal reflection) are shown in black. If ray failure occurs during optimization, we stop the optimization process and color the corresponding starting point black as well. Note that in Figure 6.2, the basins for all five minima have smooth boundaries.

When we decrease p, we observe types of behavior that are much more complex than the convergent behavior towards the minima shown in Figure 6.2. In the asymptotic regime, the limit of the sequence of iterations is then not necessarily a point, but can be a set of points. For example, for a certain starting configuration in the former basin of local minimum C, below a critical value of $p \approx 0.000138$, the algorithm does not converge to local minimum C anymore. Figure 6.3 shows how the algorithm oscillates between different points in the variable space when we decrease the value for p. For each value of p, we computed 999 outer iterations. The values of curvature c_3 of the last 899 iterations are plotted superimposed as function of p (p decreases along the horizontal axis). Since we are interested in the asymptotic behavior, the first 100 outer iterations are discarded in the figure in order to avoid transitory features.

At $p_1 \approx 0.000138$, we observe a so-called pitchfork bifurcation [62]. Minimum C, which was stable for higher values of p, becomes unstable, but a stable pair of points (both

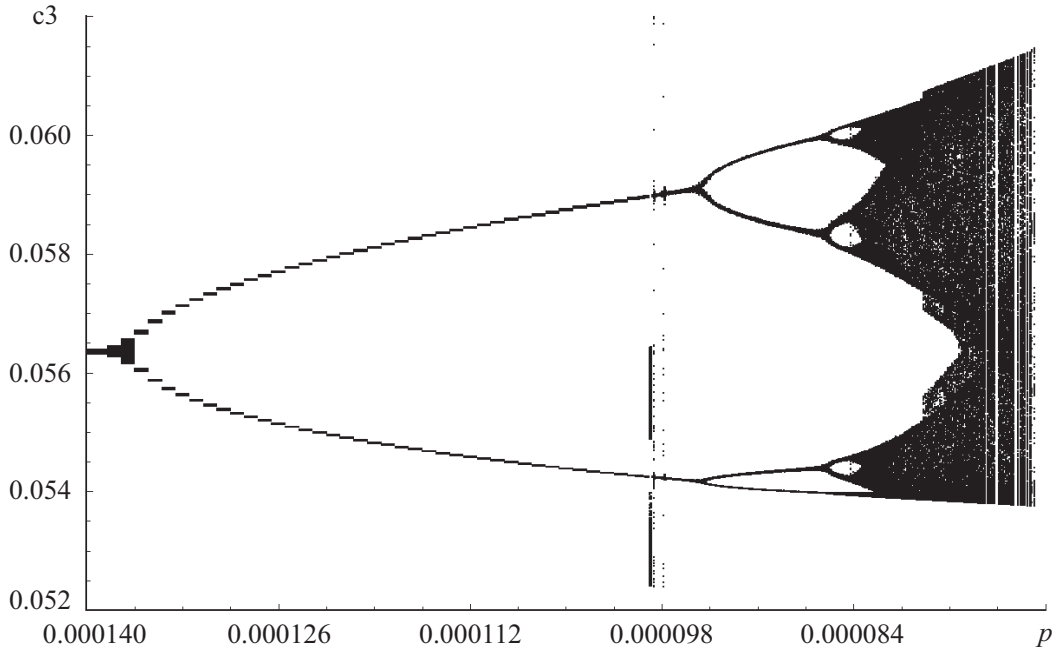


Figure 6.3: *Period doubling route to chaos. Curvature c_3 of the points in the asymptotic regime for a starting configuration in the basin of doublet local minimum C (see Figure 6.2) is shown as function of the damping parameter p.*

of them distinct from C) appears. For p between 0.000138 and 0.000095, the algorithm alternates in the asymptotic regime between the upper and lower branch resulting from the bifurcation, which correspond to two points with different values of the merit function, without ever converging to any of them. That means we have a periodic attractor consisting of two points. An attractor is the set of points in the two-dimensional variable space towards which starting configurations in the corresponding basin of attraction approach asymptotically in the course of the outer iteration process.

In the asymptotic regime, the iteration process has now a period of two. By starting the algorithm at one of the attracting points, each second iteration produces the starting point as result: after one (outer) iteration, we obtain the other point of the attractor; after the next iteration, the system is moved back to the original point, and so on. (In Figure 6.3, the points on vertical lines close to $p = 0.000100$ can be discarded.)

When p is decreased further, the two points bifurcate again into four points at $p_2 \approx 0.000095$ (period of four), then at $p_3 \approx 0.000086$ into eight points (period of eight). In the latter case, only seven out of eight distinct points can be easily distinguished in Figure 6.3, because two points in the attracting set have approximately the same value for c_3. Examining the values of curvature c_2 shows that these points are distinct. Note that the values of p where the bifurcations take place, become closer and closer to each other, until p reaches a critical value at which a chaotic set of attracting points appears. This behavior is known as the period doubling route to chaos, the most common of several routes to chaos for a nonlinear dynamical system [62].

The pitchfork bifurcations in Figure 6.3 are similar to the ones found in the logistic map and in experiments in many areas, including hydrodynamics, electronics and laser

physics. In Figure 6.3, the ratio of the first two bifurcation intervals

$$\delta = \frac{p_2 - p_1}{p_3 - p_2} \approx 4.78 \tag{6.3}$$

is close to the value of the universal Feigenbaum constant ($\delta = 4.6692\ldots$) encountered for functions that approach chaos via period doubling [65]. After a range of chaotic behavior, for values of p smaller than approximately 0.000071, the algorithm encounters ray failures and the iterative process is stopped.

The period doubling route to chaos is also illustrated in Figure 6.4. Figures 6.4(a)–(c) show periodic attractors consisting of two, four and eight points, respectively. The figures on the left hand side show the iteration trajectories in the two-dimensional variable space. Curvatures c_2 and c_3 are plotted along the vertical and horizontal axis, respectively. The iteration points shown in gray are transients, and local minimum C is shown as a large gray dot.

The figures on the right hand side show the values of c_2 as function of the number of iterations, including the transients (gray dots). After the transients, Figures 6.4(a)-(c) show horizontal lines, and each line corresponds to the c_2 value for a point in the periodic attractor. As mentioned above, attracting points that have equal values of c_2 cannot be distinguished. For example, two horizontal lines overlap each other in Figure 6.4(c). Figure 6.4(d) shows a remarkably complex pattern, corresponding to the transition region between order (period doubling) and chaos. In Figure 6.4(e), the set of attracting points forms a chaotic attractor, which is similar to those studied in nonlinear dynamics [60].

It is important to note that in the examples shown in Figures 6.3 and 6.4 the points of a certain periodic or chaotic attractor do not have the same value of the merit function. Figure 6.5 shows that the points of periodic or chaotic attractors obtained with low damping have a merit function value that is higher than or equal to the merit function value of the original minimum that was destabilized (dashed gray line). An algorithm change that allows the merit function to increase is therefore essential for nonconvergent behavior.

By not allowing the merit function to increase during the process, in the conventional local optimization algorithms the set of attracting points must have the same merit function value in the asymptotic regime. Therefore, an attractor consisting of several points with different values of the merit function cannot exist. In the asymptotic regime, we have then a point attractor, which means that the final result is convergence. However, as shown in Figure 5.14, the iterations of conventional local optimization algorithms based on damped least-squares methods sometimes first display an unpredictable chaotic path in the merit function landscape, before they finally converge to a local minimum. This type of behavior is known as a chaotic transient [56]. In the next section, we give another example of a chaotic transient. In Section 6.5, we use the obtained results for a better understanding of the instabilities observed in conventional damped least-squares local optimization algorithms.

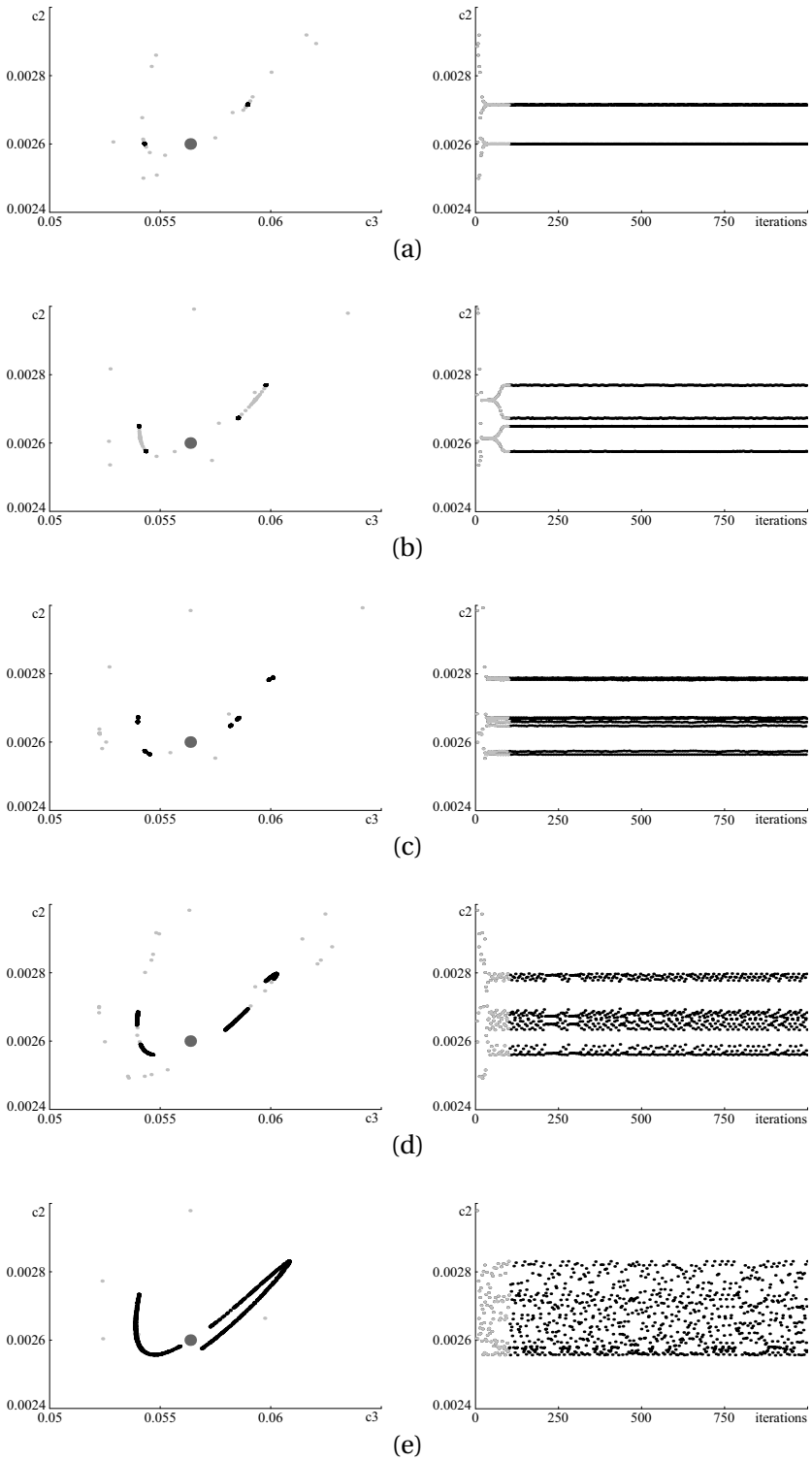


Figure 6.4: *Figures on the left: Iteration trajectories in the two-dimensional variable space* (c_3, c_2) *of a monochromatic doublet for five different values of damping parameter* p*. The starting configuration is in the former basin of local minimum C (see Figure 6.2). The large gray point corresponds to local minimum C (obtained with sufficient damping), and the iterations shown in gray are considered as initial transients. Figures on the right: The evolution of curvature* c_2 *as function of the number of iterations. (a)* $p = 0.0001$*, (b)* $p = 0.00009$*, (c)* $p = 0.0000854$*, (d)* $p = 0.0000833$*, (e)* $p = 0.0000777$*.*

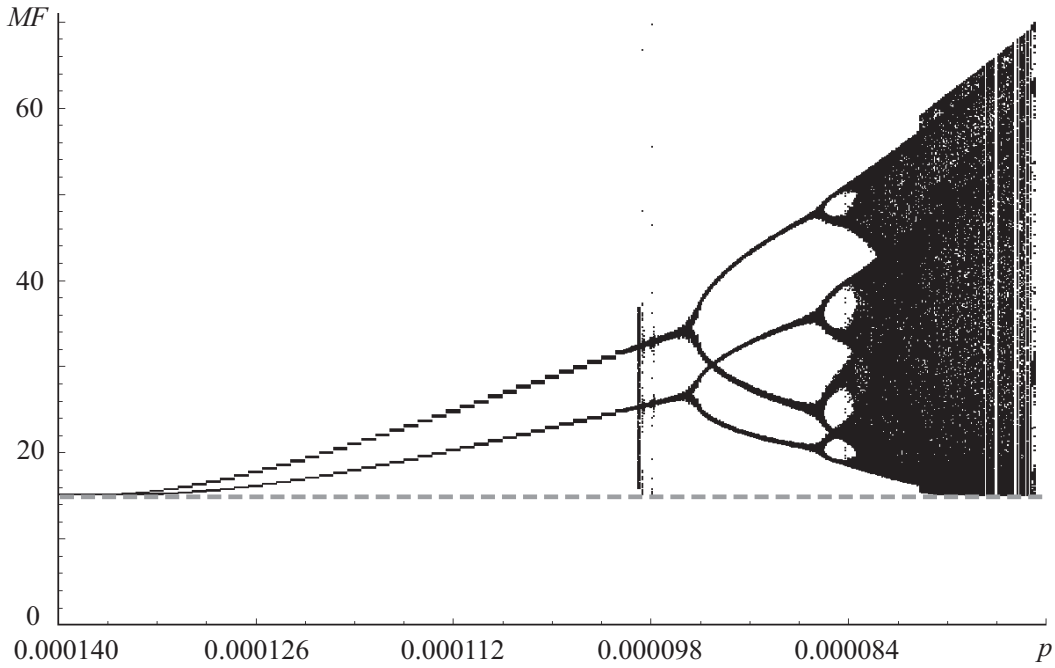


Figure 6.5: *Merit function (MF) values (in relative units) for the points shown in Figure 6.3.*

6.4 Escaping from a poor local minimum

In this section, we show that the algorithm described in Section 6.2 may improve the result of local optimization that would otherwise converge to a poorer solution (local minimum B). When we use sufficiently low damping (i.e. p small enough), the algorithm overcomes the merit function barrier between local minima B and A, and finally converges to the better minimum A. After a chaotic attractor is formed, further decrease of damping destroys the chaotic attractor and the iteration trajectory moves off to a different region of the variable space, where it converges to minimum A. However, before the iterations converge to minimum A, first a chaotic transient is observed.

Figure 6.6 shows the basins of attraction for the monochromatic doublet for four low values of p. We used a grid of 101×101 points, where each grid point is iterated 999 times. Note that the basins of minima A, C, D, and E remain compact. However, the basin of minimum B, which has the largest value of the merit function, completely disappears for $p = 0.0006$. First, for a variation domain of p including the value $p = 0.0009$, the algorithm approaches a periodic attractor consisting of two points [the white points P_1 and P_2 in Figure 6.6(a)], which are close to the former point B. After a sufficiently large number of iterations n, the succession of each second iteration converges to P_1 (the regions that lead to this point after n large and odd are shown in dark blue) or to P_2 (light blue).

For even lower values of p, the number of attracting points increases until a chaotic attractor is formed [white points in Figure 6.6(b)]. The purple region shows the set of starting configurations that are attracted to the chaotic attractor. When the damping is lowered further, the process iterates over the merit function barrier and converges to

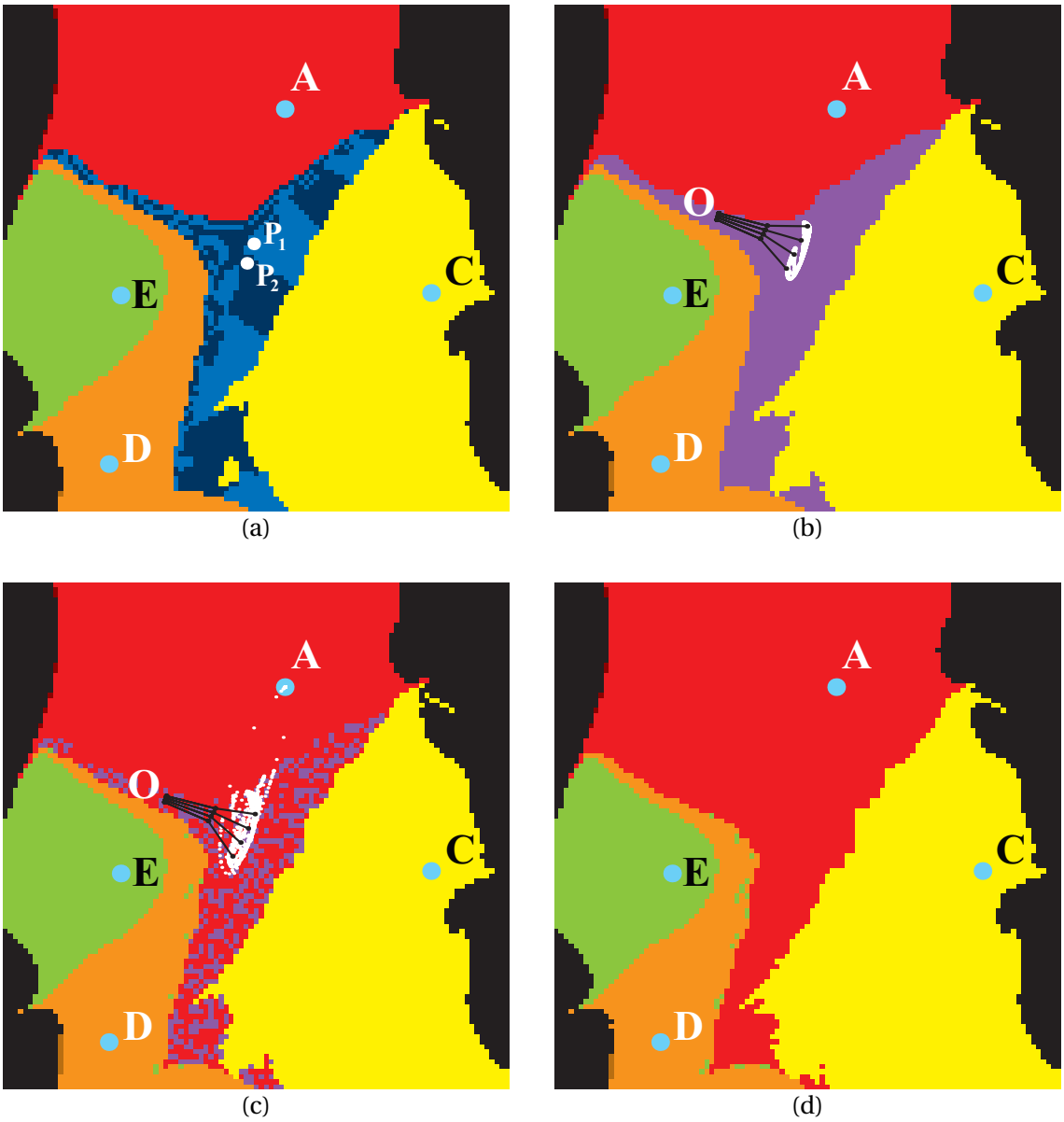


Figure 6.6: *Basins of attraction obtained for different values of the damping parameter p on a grid of* 101×101 *points. Each grid point is iterated* 999 *times. (a)* $p = 0.0009$, *(b)* $p = 0.00075$, *(c)* $p = 0.000679$, *and (d)* $p = 0.0006$. *The black lines starting close to point O in Figures 6.6(b) and (c) will be explained in Section 6.5.*

local minimum A [Figures 6.6(c) and (d)]. In chaos research, one commonly encounters situations called 'crises' where, after the control parameter passes through a critical value, the chaotic attractor becomes unstable [60]. The same phenomenon is observed here.

When we choose as starting point for instance the white point indicated with 'S' in Figure 6.2, or the point 'O' in Figure 6.6(c) (in the set of four neighboring black points, O is the second one from below), the algorithm first displays a chaotic transient. Then, it oscillates for many iterations in a domain of the variable space that looks very similar to a set of attracting points, but which finally turns out to be unstable. In chaos research, such a set of points resulting from a chaotic attractor that became unstable after a 'crisis' is called a chaotic saddle [66]. Starting configurations for which the algorithm converges to minimum A are shown in red. The white points in Figure 6.6(c) show the set of iteration points of the chaotic transient and of the subsequent escape to minimum A, when we choose point O as starting point (the first two iterations started from O are shown in black and connected with black lines). For the starting configurations colored in purple, the total number of 999 iterations was not sufficient for $p = 0.000679$ to escape from the chaotic behavior. When we increase the number of iterations, the configurations corresponding the the purple points also reach minimum A. When p becomes sufficiently low ($p = 0.0006$), all starting configurations in the former basin of minimum B converge within 999 iterations to minimum A [Figure 6.6(d)].

The transition to chaos, the chaotic-like behavior, and the chaotic transients are also illustrated in Figures 6.7 and 6.8. For the starting point S (see Figure 6.2), Figure 6.7 shows the transition from convergence (to minimum B) to chaos and then to convergence to minimum A after the 'crisis'. The first 300 iterations are discarded and curvature c_2 of the last 699 iteration points are plotted as a function of p. The horizontal lines on the left- and right-hand side of Figure 6.7 correspond to the c_2 values of minima B and A, respectively.

Figure 6.8 shows the iteration trajectories for a set of values of p in the (c_3, c_2)-space (left), and the values of c_2 for the iteration points as function of the number of iterations (right). The iterations shown in gray are considered to be initial transients, and local minimum B is shown as a large gray point.

In comparison to Figure 6.3, the route to chaos in Figure 6.7 is a more unusual one. First, we observe a bifurcation, which is then followed by trifurcation. Note that the figure shown on the right in Figure 6.8(a), where two stable attracting points can be observed, also seems to suggest the presence of six unstable fixed points. When p decreases, these six points become stable and the algorithm continues with oscillations between these six points [Figure 6.8(b)].

In Figure 6.8(c), the complex pattern of iterations shows that we are near the transition to chaos. For slightly smaller values of p, the iteration points form a chaotic-like set in the variable space [Figure 6.8(d)]. After the 'crisis', this set of points becomes an unstable chaotic saddle. The algorithm escapes from it and moves to a different region of the variable space [Figure 6.8(e)]. The convergence to local minimum A is illustrated by the horizontal line at the top right of Figure 6.8(e). Certain details are different when we choose another point in the former basin of minimum B as starting point.

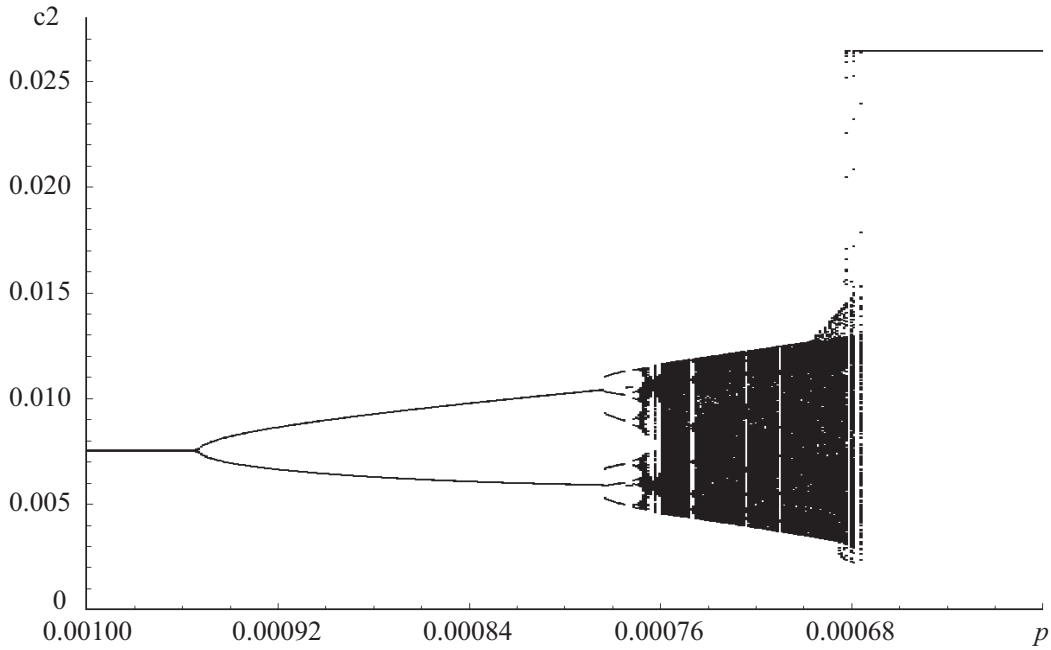


Figure 6.7: *Attracting points for the starting configuration S in the basin of local minimum B (see Figure 6.2). The first* 300 *iterations have been discarded. Curvature* c_2 *of the iterations between* 301 *and* 999 *has been plotted as function of* p.

Note in Figures 6.7 and 6.8(a)–(d) how the size of the stable periodic or chaotic attractors increases when the damping is reduced. For the chaotic attractor, the basin of attraction is the purple region shown in Fig. 6.6(b). As can be seen in Figure 6.6(c), the 'crisis' occurs when the growing chaotic attractor reaches the chaotic basin boundaries. After this so-called 'boundary crisis' [56, 66], the white iteration points leave the mixed red-purple region [essentially the basin of the chaotic attractor shown in Fig. 6.6(b)] and enter the compact red region (the basin of minimum A). The final result is then convergence to minimum A.

As in the case of Figure 6.5, Figure 6.9 shows that before the 'crisis' all points of periodic or chaotic attractors have a merit function which is higher than or equal to that of minimum B (dashed gray line). However, after the 'crisis' and after escaping from the chaotic saddle, the merit function decreases significantly and becomes that of minimum A (horizontal black line in the lower right part of the figure). A behavior similar to the one described above can also be observed when the algorithm escapes from minimum E to the somewhat lower minimum D. However, we then encounter ray failure for many values of p after the 'crisis'.

In the chaos regions of Figure 6.7, bands are formed where only a small number of points appear to form the attractor. Similar bands are formed in the chaos region of the logistic map, where trifurcations can also be observed. However, in the diagram shown in Figure 6.7, the trifurcation appears outside the chaos region.

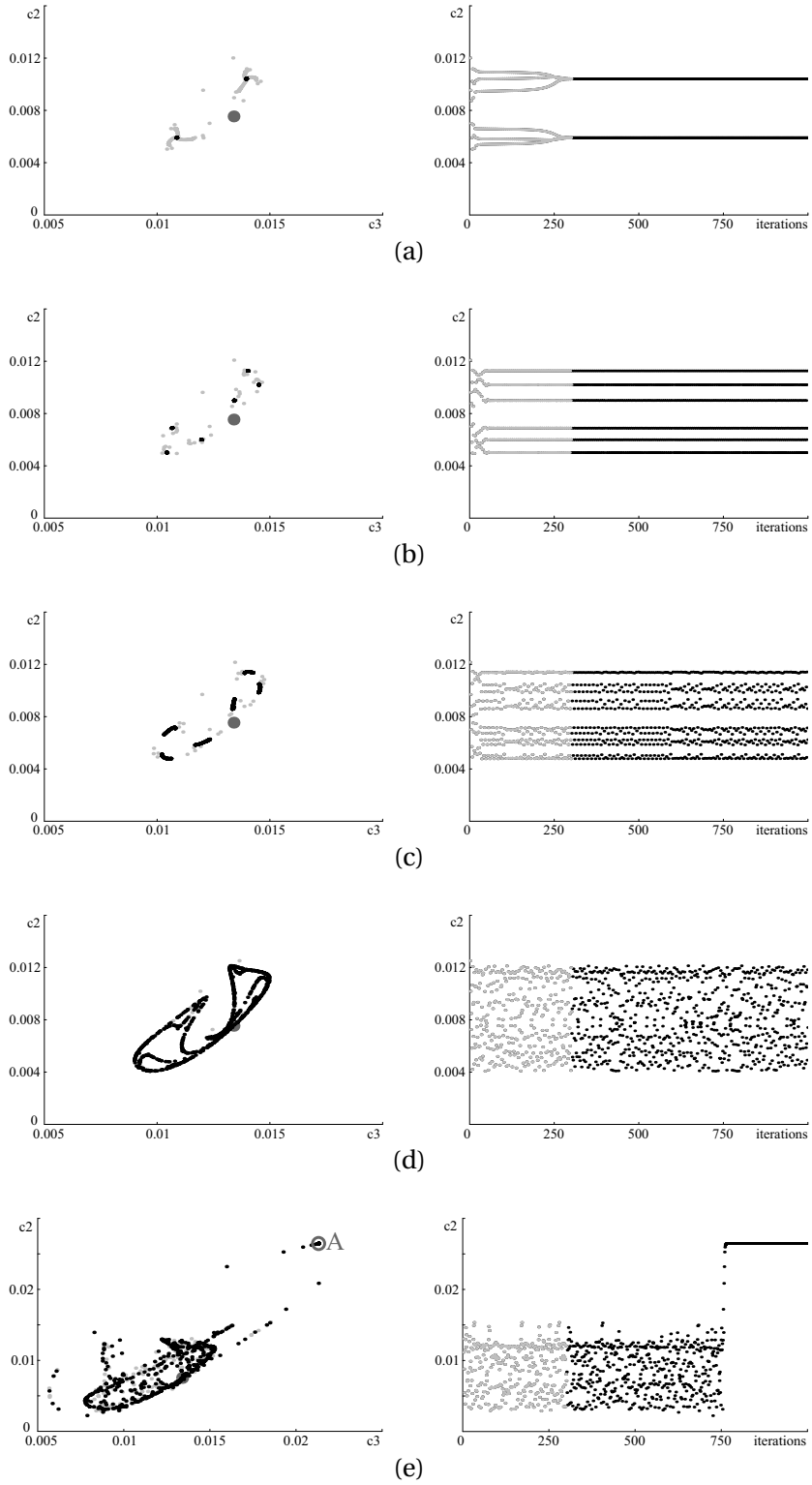


Figure 6.8: *Figures on the left: Iteration trajectories in the two-dimensional variable space* (c_3, c_2) *of a monochromatic doublet for five different values of damping parameter* p*. The starting configuration is point S in the former basin of local minimum B (see Figure 6.2). The large gray point corresponds to local minimum B (obtained with sufficient damping), and the iterations shown in gray are considered as initial transients. Figures on the right: The evolution of curvature* c_2 *as function of the number of iterations. (a)* $p = 0.000784$*, (b)* $p = 0.000776$*, (c)* $p = 0.000768$*, (d)* $p = 0.000730$*, (e)* $p = 0.000679$*.*

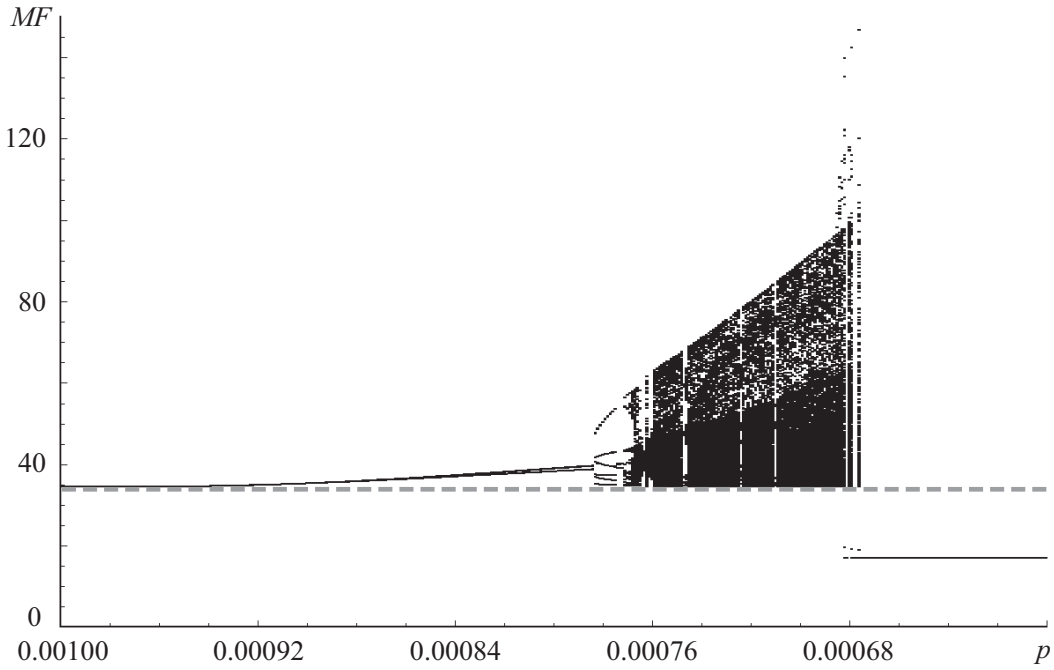


Figure 6.9: *Merit function (MF) values (in relative units) for the points shown in Figure 6.7.*

6.5 Instabilities in conventional damped least-squares optimization

In Chapter 5, we have shown that optical design problems exist where conventional local optimization, as it is implemented in commercial optical design programs, is unstable if low damping is used. Then, starting points close to each other converge to different local minima. As we have shown there, if present, instabilities decrease the degree of predictability of the design process, and therefore a better understanding of the mechanisms by which such instabilities can arise is desirable.

While the detailed mechanisms of these instabilities require further study, below we show that certain features mentioned in Chapter 5 can be understood by examining Figures 6.6(b) and (c). (If necessary, for observing details better, these figures can be enlarged on-screen in the electronic version of this thesis.) These figures show the second and third iteration for a set of four starting points close to each other (the black lines). The four trajectories are attracted by the chaotic attractor in Figure 6.6(b) and initially also by the chaotic saddle in Figure 6.6(c). While for the settings of Figure 6.6(b) the attractor is stable and all four trajectories remain there indefinitely, in Figure 6.6(c), the previous attractor that is now a chaotic saddle first attracts, then repels the trajectories, which finally arrive at minimum A. We note that, until the four trajectories reach the chaotic attractor in Figure 6.6(b) and the chaotic saddle in Figure 6.6(c), the distance between them increases significantly after each iteration.

A similar behavior can be observed in Figures 5.14(a) and (c), where essentially the same doublet optimization problem as the one discussed in this chapter is analyzed with two different widespread commercial optical design programs (see Chapter 5 for

full details). The distance between several starting points close to each other increases significantly, until the trajectories reach a domain of the variable space that is close to a certain equimagnitude contour line (i.e. a curve along which the merit function is constant) that first attracts, then repels the trajectories [3]. For instance, a large number of iteration points in the middle of Figure 5.14(c) indicates the presence of a horizontal attracting-repelling line there.

The extreme sensitivity to initial conditions occurs because, when they have reached the attracting-repelling lines, these trajectories are already far apart from each other and therefore, after the trajectories are repelled, they converge to different local minima, despite of the fact that they have their origin in almost the same point.

Since the domains close to the attracting-repelling lines play a key role in the mechanism of the instabilities, it is important to understand their nature. However, studying chaotic transients directly in commercial programs is difficult, because there the chaotic transients are short. Since it can be used to increase the duration of the chaotic transients, the modification of the damped least-squares algorithm described in this chapter is helpful for an analogy. As shown in chaos research, there is a fundamental relationship between the presence of chaotic transients and the existence of a chaotic saddle [66]. This fact, and the fact that domains close to the attracting-repelling lines act in a way that is similar to what is proven to be a chaotic saddle in Figure 6.6(c), suggest that these domains contain a chaotic saddle (or parts of it) as well.

6.6 Conclusion

Our simple example shows that a merit function barrier in the optical design landscape can be overcome by using a damped least-squares method with reduced damping, in which the merit function is allowed to increase. Like most other empirical strategies of this kind, success is not guaranteed (ray failure sometimes occurs). However, when the present method is successful, the complex mechanism we have observed is one that is frequently encountered in chaos research. Because such mechanisms depend not on model details but rather on some general properties of the model, we believe that these results are relevant not only for the present algorithm with its specific implementation choices, but for other possible algorithms of the same kind as well.

To escape towards a better local minimum, it is necessary to destabilize first the convergence of the algorithm towards the original poor minimum, and we have seen that this can be achieved by lowering the damping. However, before convergence to a different minimum occurs, decreasing the damping leads to the successive stabilization and then destabilization of a sequence of periodic and chaotic attractors and to the growth of the size of the attractor. The distance between the attractor and its basin boundary decreases, and when the attractor touches its basin boundary, the attractor is destroyed by a boundary crisis.

[3] In Reference [67], a movie shows the history (in green) obtained with CODE V data of a set of ten starting points that are very close to each other. For each starting point, the iteration trajectories are first attracted, and then repelled by domains close to certain equimagnitude contour lines (the red and purple lines in the movie).

For damping values that are lower than the critical crisis value, after a chaotic transient, the final result is convergence to a minimum which is distinct from the original one. The sequence of iterations is first attracted, then repelled by the chaotic saddle that results from the chaotic attractor after the crisis.

The chaotic transients shown in Section 5.4 suggest that such complex behavior is more general, and also occurs with conventional damped least-squares algorithms, if they are designed to maximize speed. By not allowing the merit function to increase (significantly) during local optimization, the occurrence of periodic or chaotic set of points with different values of the merit function, that might prevent convergence, is not possible, and conventional damped least-squares methods are convergent, as expected. However, in early stages of optimization, conventional low-damping algorithms can create chaotic transients, which could add an extra element of unpredictability in the design process. Because of their temporary character, chaotic transients are more difficult to be studied than chaos itself. The modification of the damped least-squares algorithm used here was also a mean to increase the duration of the chaotic transient, and even to stabilize the chaotic phase. Therefore, this modification is helpful for a better understanding of qualitative features of the unexpected behavior in widespread optical design programs.

The position in the variable space and the geometry of domains that act like chaotic saddles are important factors that determine whether instabilities in optimization are present or not. As we have seen in Figure 6.6(c), the presence of chaotic transients and of a chaotic saddle do not necessarily lead to instabilities in optimization. The growing chaotic attractor in Figure 6.6(b), that becomes a chaotic saddle in Figure 6.6(c), exceeds its own basin boundary in only one region, and in the final phase of the chaotic transient the optimization converges to the minimum that has its basin on the other side of that region. Virtually all starting points in the basin of the former chaotic attractor converge finally to the same minimum and instabilities are not present. However, in the examples shown in Section 5.4, starting points close to each other do converge to different minima because domains acting like chaotic saddles first attract neighboring trajectories (and at the same time enlarge the distances between them) and then repel them so that they leave the domain in different regions that are far away from each other.

Chapter 7

Conclusions

It is often stated that optical system design is both a science and an art. The art of designing optics relies heavily on the experience of the designer, and contains a significant component of trial and error. In this thesis we have tried to reduce the trial and error part, making optical system design more systematic.

The optical merit function is in general highly nonlinear, which typically gives rise to many local minima in the merit function landscape. The number of local minima increases rapidly with the number of optimization variables. The solution that will be obtained after local optimization is then critically dependent on the choice of the initial configuration. We have shown that even a very simple optimization problem with only two variables has in some cases no less than five solutions.

In present-day optical system design, it is tacitly assumed that all local minima are a disordered set of points in the merit function landscape without relationships between them. However, by considering saddle points with a Morse index value of 1, we have shown that there is order in the merit function topography. Saddle points are critical points in the merit function landscape which always remain on the boundaries of basins of attraction, independent of the used optimization method. They can be used to systematically travel from one basin of attraction to another. The basin of attraction of a local minimum is the set of initial conditions which lead to that minimum after optimization.

One of the major difficulties in present-day global optimization is that the computing time increases significantly when the dimensionality of the optimization problem is increased. In this thesis, we have presented a new method for finding local minima that suffers much less form this drawback. This method, which we call Saddle-Point Construction (SPC), changes the existing system shape (to find a new local minimum) by changing the dimensionality of the problem.

We have proven that, if the dimensionality of the optimization problem is increased in a way that satisfies certain mathematical conditions (the existence of two independent transformations that leave the merit function unchanged) then a local minimum is transformed into a saddle point. In lens design, we transform a local minimum into a saddle point by adding a 'null-element' meniscus (which does not affect the path of any ray or the merit function of the system) in the optical system. We have shown that

saddle points can be constructed in the special case when the inserted null-element lens is in direct contact with one of the surfaces of the original local minimum (the reference surface) and when the glass of the new lens is the same as that at the reference surface. With the generalized version of the SPC method, the glass and distance restrictions mentioned above are removed, and the curvatures of the lens to be inserted can be computed numerically. We have shown that this is a simple one-dimensional problem.

The saddle points obtained with the SPC method are called 'null-element saddle points' (NESP's). The null-element comes with two new variables (the two surface curvatures), which enable the merit function to decrease after optimization. Lens designers frequently insert lenses into their designs and, in the traditional way, one new system shape results after optimization. However, when a lens is inserted with SPC, two distinct system shapes result after optimization and for further design one can choose the better one. By inserting lenses according to the SPC method, and then, if necessary, by extracting lenses, new local minima for optical systems of arbitrary complexity can be obtained very rapidly. We have discussed examples that illustrate the essence of the method, which can be used in essentially the same way for arbitrary systems. For simple systems, many saddle points result from adding null-elements to minima with a lower dimensionality.

Although the theory behind the SPC method contains concepts which are new in lens design, the technique is very straightforward, and both versions (for the special and general case) can be easily integrated with traditional design techniques. With the generalized SPC method, we can find all robust saddle points in the merit function landscape of doublet and triplet lens designs. A small number of saddle points (the non-NESP's), which cannot be constructed with our SPC method. They typically have low merit function barriers, which can easily be overcome by using for instance a method based on simulated annealing.

Thinking in terms of saddle points is still unfamiliar to most optical system designers, but we believe that the potential for discovering new solutions and for improving design productivity justifies the effort for understanding such new methods and for combining them with traditional design methodology. In principle, SPC should also be applicable in other optimization problems, where it is possible to define a 'null-element' and to find two independent transformations, similar to the ones we use in optical system optimization, that leave the merit function unchanged. For example, this is the case in thin-film optimization. In applications other than lens design, more research is needed to investigate the practical utility of SPC. However, this was not part of this research project.

Optical system designers usually assume that local optimization methods have a predictable behavior. However, we have shown that, in certain situations, the frequently used damped least-squares optimization methods can have an unpredictable behavior. The examples discussed in this thesis show that the basin structure and boundaries depend on the implementation details of the local optimization algorithm that is used and on numerical parameters, such as the damping factor. In our examples, there are regions in the merit function landscape where starting points, which are very close to each other, lead to different local minima after optimization. This instability illustrates

an extreme sensitivity to change of initial configurations. Computations of the capacity dimension then show that the basin boundaries have a fractal shape. When optimization is started in a fractal region, it first displays a chaotic transient path in the merit function landscape, before it finally converges to a local minimum.

By using damped least-squares methods with sufficient damping, we have obtained smooth basin boundaries even in optimization problems that have a natural propensity for fractal basins. Depending on the specific optimization problem and optimization method that is used, instabilities may be absent or present in different degrees. If these instabilities are present, then the main cause for them is to be found in the optimization algorithm.

The concept of damping factor was originally introduced with the purpose of limiting the step size during optimization, so that certain mathematical approximations in the least-square algorithm remain valid. Many optimization algorithms presently use much lower values of the damping factor. Although the exact details may vary from one program to another, the automatic choice of the damping factor is typically made such that the merit function decrease per optimization step is maximized. However, as we have shown, the gain in computation speed is at the cost of stability. Strategies for choosing the damping factor, which turn out to be optimal close to local minima, are not necessary optimal far away from local minima or near the basin boundaries.

Once a design is trapped in a local minimum, it cannot escape by further continuing the local optimization process. To escape towards a better local minimum, it is necessary to destabilize the convergence of the algorithm towards the original poor minimum. We have shown that a low merit function barrier in a design landscape can be overcome by using a damped least-squares method with reduced damping, in which the merit function is allowed to increase. Like most other empirical strategies of this kind, success is not guaranteed (ray failure sometimes occurs). However, when the present method is successful, the complex mechanism we have observed is one that is frequently encountered in chaos research. Because such mechanisms depend not on model details but rather on some general properties of the model, we believe that these results are relevant not only for the present algorithm with its specific implementation choices, but for other possible algorithms of the same kind as well.

In a two-dimensional optimization problem for a monochromatic doublet, we have shown a successful escape from a poor local minimum. Before convergence to the better minimum occurs, decreasing the damping leads to the successive stabilization and then destabilization of a sequence of periodic and chaotic attractors and to the growth of the size of the attractor. The distance between the attractor and its basin boundary decreases, and when the attractor touches its basin boundary, the attractor is destroyed by a boundary crisis. After a sequence of iterations, which are first attracted, and then repelled by the chaotic saddle that results from the chaotic attractor, the algorithm converges to a local minimum that is better than the one obtained with normal damping.

By not allowing the merit function to increase (significantly) during local optimization, the occurrence of periodic or chaotic set of points with different values of the merit function, that might prevent convergence, is not possible, and conventional damped least-squares methods are convergent, as expected. However, in early stages of optimization, conventional low-damping algorithms can create chaotic transients, which

could add an extra element of unpredictability in the design process. Because of their temporary character, chaotic transients are more difficult to be studied than chaos itself. The modification of the damped least-squares algorithm used in this thesis was also a mean to increase the duration of the chaotic transient, and even to stabilize the chaotic phase. Therefore, this modification is helpful for a better understanding of qualitative features of the unexpected behavior in widespread optical design programs.

When designers adapt an existing design to new specifications, they would usually prefer to obtain a system shape and correction properties that are similar to ones that are already known. A higher degree of predictability is also desirable when one attempts to reach the design goal in several stages, and at each stage, one wants to preserve what has been achieved in previous stages. When the outcome is not the expected one, it is usually believed that the cause lies in the inherent complexity of the design landscape. This is indeed a major source of unpredictability.

However, in this thesis we have shown another possible cause. One of the goals of this thesis is to make optical system designers aware of the possibility of the presence of instabilities in the optimization process. Sensitivity to initial conditions can influence the result that will be obtained after optimization. While the inherent complexity of the design landscape limits the sizes of the basins of desirable solutions, instabilities in optimization, if present, further increase complexity and decrease predictability by mixing basins of desirable and undesirable solutions together.

Appendix A

Examples of Sadde-Point Construction in the special case

In Chapter 3, we have illustrated the special case of the Saddle-Point Construction (SPC) method with three examples. In this appendix, we give a more elaborate description of those examples. All necessary steps are given in detail, so that they can be reproduced with any optical design software. We have tested all examples in CODE V and ZEMAX, obtaining the same results. For CODE V, lens files for all steps, a macro that creates the null-element saddle points (NESP's), and details specific to CODE V are available via our website [49]. If desired, for obtaining the starting configurations, users of other optical design software can import or adapt the corresponding CODE V files, which are in ASCII format. Automatic conversion to the format of another program is very useful and does most of the work, but sometimes some information (e.g. wavelengths, solves) is lost. Therefore, for the present purpose it is important to compare the result of automatic conversion with the information given in the tables below, and to make changes wherever necessary.

A.1 Example 1: Generating doublets from a singlet

In five steps, we obtain two doublet local minima by starting from a singlet minimum, with specifications given in Tables A.1 and A.2. The curvatures are used as variables, and a constraint is used to keep the effective focal length constant with a value of 100 mm. The distance between the last surface and the image has a paraxial image distance solve (indicated with 'S').

The procedure is as follows (steps 3, 4, and 5 are shown in Figure A.2):

Step 1: Start with the singlet given in Tables A.1 and A.2, and reoptimize (Figure A.1).

Step 2: Insert a 'null-element' (a meniscus with zero thickness) in contact with the second surface (with curvature c_2) of the singlet (the arrow in Figure A.1). The glass of the new lens must be the same as that of the first one. Next, make the two curvatures

Table A.1: *Specifications for the configurations in Example 1. The reference wavelength is indicated with (*).*

Entrance pupil diameter (mm)	20.000
Effective focal length (mm)	100.000
Paraxial image height (mm)	8.749
Wavelenghts (nm)	656.30, 587.60(*), and 486.10
Field points (deg)	0.000, 3.000, and 5.000

Table A.2: *Starting system for Example 1. The variable curvatures are indicated with 'V', and solves are indicated with 'S'.*

Surface #	Curvature (mm^{-1})		Thickness (mm)		Glass type
object	0.00000		infinity		AIR
1	0.01685	V	4.00000		BK7_SCHOTT
2	-0.00255	V	97.70304	S	AIR
image	0.00000		0.00000		

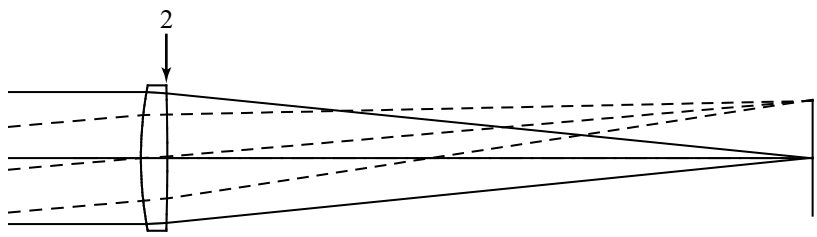


Figure A.1: *Starting system for Example 1.*

(c_3 and c_4) of the null-element equal to c_2. The obtained system is a doublet NESP. Finally, make the new curvatures variable. Note that, because of the zero distances and the property $c_2 = c_3 = c_4$, the lens drawing for the NESP is indistinguishable from the one for the original local minimum shown in Figure A.1.

Step 3: While keeping the first two curvatures of the doublet (i.e. the old variables of the singlet) unchanged, change the two curvatures of the thin lens. Construct one doublet with the property $c_3 = c_4 = c_2 - \epsilon$ (system 1 in Figure A.2), and one doublet with the property $c_3 = c_4 = c2 + \epsilon$ (system 2 in Figure A.2), where ϵ indicates a small curvature change. To avoid that the gradient of the merit function becomes too small for optimization, ϵ should not be chosen too small. In this example, $\epsilon = 0.00003$ mm^{-1} (in the figure, the curvature change is shown exaggerated). In this way, two systems situated on opposite sides of the NESP are obtained which can be further optimized to generate two new solutions.

Step 4: Optimize the two systems that have been obtained after step 3.

Step 5: Gradually increase the thickness of the lens resulting from the null-element and the axial distance between it and the previous lens, until the desired values are ob-

tained. In this example, the thickness of the null-element has been increased in four steps (e.g. thickness = 1, 2, 3, and then 4 mm) and the air distance separating the two lenses of the system in two steps (e.g. distance = 1, and then 2 mm). After each increment, the system should be reoptimized while keeping the thicknesses constant.

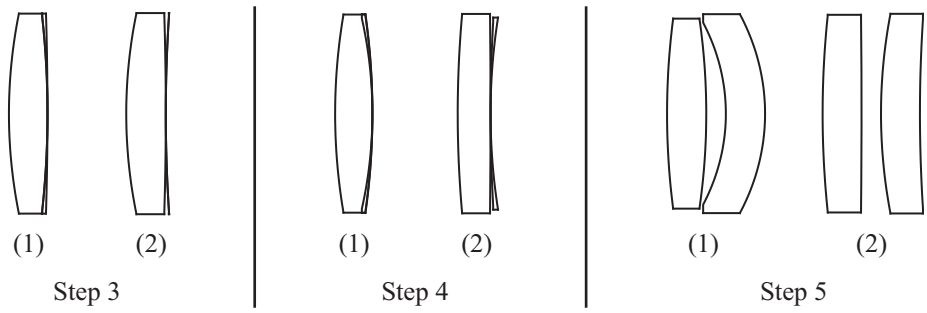


Figure A.2: *Steps 3, 4 , and 5 for SPC in Example 1.*

Since the two glasses of the doublets resulting from step 5 are the same, these systems are not corrected for axial color. However, by glass change we can improve the first of the two systems. For example, if we change the material of the second lens from BK7 to F2, system 1 will become a well-known air-spaced Fraunhofer-type design after reoptimization (Figure A.3).

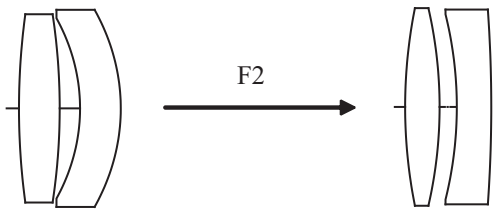


Figure A.3: *Obtaining a Fraunhofer-type design.*

When the starting singlet is split in a traditional way as shown below and then reoptimized, the resulting system will be system 2 in step 5. However, this system has a poorer imaging performance than system 1 in step 5, which can lead to the Fraunhofer-type configuration. This is illustrated in the following four steps (Figure A.4).

Step 1': Start with the optimized singlet used for the SPC.

Step 2': Split the singlet into two lenses with the same thickness (i.e. insert a plate of air with zero thickness and flat surfaces in the middle of the singlet.

Step 3': Optimize the system with the curvatures as variables (the merit function remains the same as above).

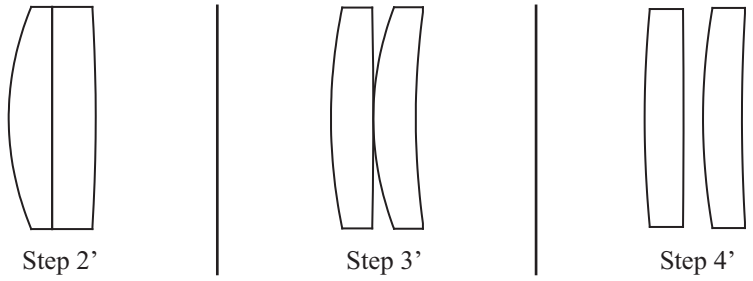


Figure A.4: *Traditional power splitting.*

Step 4': Increase the thickness of the two lenses and the axial distance between them to the same values as that in the two systems obtained with SPC (see step 5 above). In our example, the lens thickness is 4 mm and the air distance is 2 mm.

The final system that results from splitting the singlet is one of the two systems which are obtained with SPC (system 2 in the result of step 5, see Figure A.2). It has a merit function that is worse than that of the other system obtained with SPC (system 1 in the result of step 5, see Figure A.2).

A.2 Example 2: Generating quintets from a quartet

In this example, we construct a quintet NESP from a quartet minimum. The specifications of the quartet which is used as starting system are given in Tables A.3 and A.4.

Table A.3: *Specifications for the configurations in Example 2.*

Entrance pupil diameter (mm)	50.000
Effective focal length (mm)	100.000
Paraxial image height (mm)	24.933
Wavelenght (nm)	656.30
Field points (deg)	0.000, 10.000, and 14.000

The merit function is the same as in Example 1, except that, instead of a constraint within the merit function, the last surface has a solve for the marginal ray exit angle (equal to -0.25) to keep the effective focal length constant. The distance between the last surface and the image has a paraxial image distance solve.

First, we construct a quintet NESP by inserting a null-element at the second surface in this quartet (Figures A.5 and A.6).

Step 1: Start with the quartet given in Tables A.3 and A.4, and reoptimize.

Step 2: Insert a null-element (with variable curvatures) in contact with the second surface of the quartet (arrow 2 in the Figure A.5). The surface curvatures of the null-element are c_3 and c_4, and that of the surface it is in contact with is c_2. Make c_3 and

Table A.4: *Starting system for Example 2. The variable curvatures are indicated with 'V', and solves are indicated with 'S'.*

Surface #	Curvature (mm^{-1})		Thickness (mm)		Glass type
object	0.00000		infinity		AIR
1	0.01981	V	12.00000		SK1_SCHOTT
2	0.00754	V	0.00000		AIR
3	0.03086	V	12.00000		SK1_SCHOTT
4	0.04530	V	15.10786		AIR
5 (STOP)	0.00000		12.92120		AIR
6	-0.03920	V	12.00000		SK1_SCHOTT
7	-0.03291	V	0.00000		AIR
8	0.00671	V	12.00000		SK1_SCHOTT
9	-0.00893	S	60.00533	S	AIR
image	0.00000		0.00000		

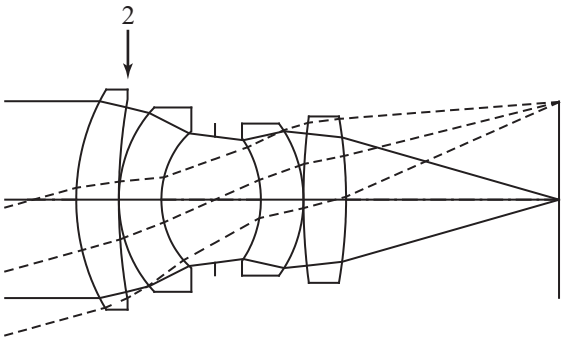


Figure A.5: *Starting system for Example 2.*

c_4 equal to c_2 and the glass of the null-element the same as that of the first lens. The obtained system with $c_2 = c_3 = c_4$ is a (quintet) NESP.

Step 3: While keeping the old variables of the quartet unchanged, change the two curvatures of the null-element as follows: $c_3 = c_4 = c_2 - \epsilon$ and $c_3 = c_4 = c_2 + \epsilon$, where ϵ indicates a small curvature change. In this example $\epsilon = 0.00003$ mm^{-1} (in Figure A.6, the curvature change is shown exaggerated). In this way, two systems on opposite sides of the NESP are obtained.

Step 4: Optimize the two systems that have been obtained after step 3.

Step 5: In the resulting systems, gradually increase the thickness of the null-element (i.e. increase the thickness in small steps and optimize the system after each step) until the desired values are obtained. During the optimizations, the thicknesses should be kept constant. For this example, the thickness of the null-element has been increased in eight steps (e.g. thickness = 1, 2, 3, 4, 6, 8, 10, and then 12 mm). After each increment,

the system should be reoptimized (a few cycles are usually sufficient in the intermediate stages).

Step 6: To remove the edge separation violation in system 1 (see Figure A.6), we slightly increase the axial distance between the first two lenses. Here, this distance was set to 3 mm.

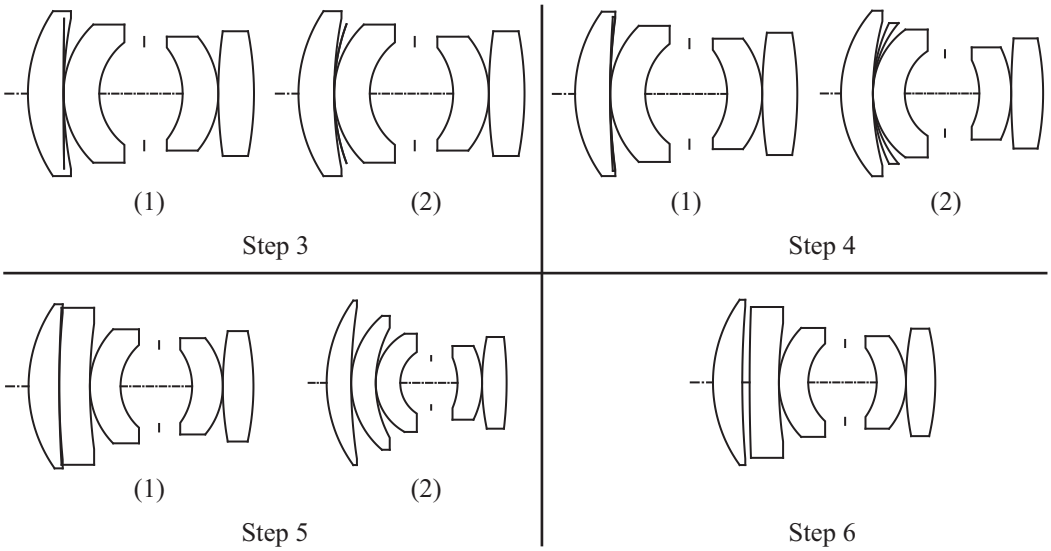


Figure A.6: *Steps 3, 4, 5, and 6 for SPC in Example 2.*

Steps 1–5 described above can also be repeated at the third surface of the quartet (see our website [49] for the resulting system files). When comparing the four systems obtained from these two NESP's, we observe that two of the systems (system 2 after step 5 in Figure A.6 and system 1 after step 5 in Appendix B of Reference [49]) are identical.

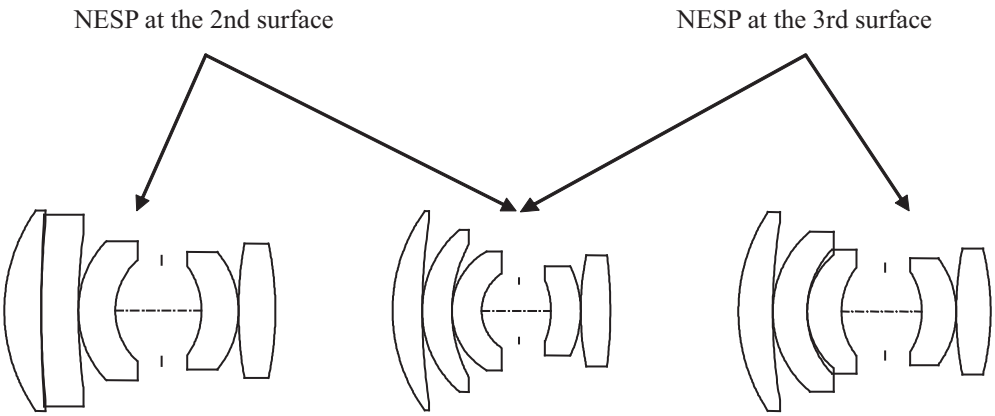


Figure A.7: *The middle system in this drawing can be obtained from both NESP's.*

The middle system in Figure A.7 (we call it a 'hub') can be obtained from other NESP's as well. For instance, by inserting a null-element lens at the first surface of the quartet, concentric with that surface and with the same glass, we obtain a third NESP. For obtaining the 'hub' we choose the side of the NESP where the meniscus lens has a slightly

weaker curvature than that of surface 1, see the configuration shown on the left in Figure A.8. After optimization we obtain the quintet shown on the right in Figure A.8.

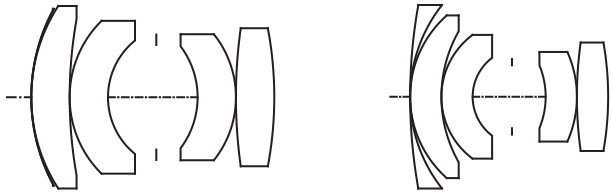


Figure A.8: *Perturbed NESP at first surface (left) and the configuration after optimization (right).*

In the original quartet all lenses have the same thickness. After increasing the thickness of the inserted lens in the system shown on the right in Figure A.8 to the same value as for the other lenses and reoptimizing, we obtain the 'hub' (middle system in Figure A.7) again.

A.3 Example 3: Obtaining a Double Gauss design

In this example, we obtain the Double Gauss design by starting from a configuration with 10 surfaces (Figure A.9). The system parameters of the starting system are given in Tables A.5 and A.6.

Table A.5: *Specifications for the configurations in Example 3. The reference wavelength is indicated with (*).*

Entrance pupil diameter (mm)	30.000
Effective focal length (mm)	100.000
Paraxial image height (mm)	24.933
Wavelenght (nm)	656.30, 587.60(*), and 486.10
Field points (deg)	0.000, 10.000, and 14.000

The object is placed at infinity, and all thicknesses and glass types are constant. The variable radii of curvature are indicated with 'V'. The merit function is the same as in Example 2, and the last surface has a solve for the marginal ray exit angle (equal to −0.15, indicated with 'S'). The distance between the last surface and the image has a paraxial image distance solve. The steps for obtaining the Double Gauss design are shown in Figures A.9 and A.10.

Step 1: Start with the configuration given in Tables A.5 and A.6, and reoptimize (Figure A.9).

Step 2: Insert a null-element (with variable curvatures) in contact with the second surface, indicated with arrow 2 in Figure A.9. Use for the null-element the same glass as that of the first lens.

Table A.6: *Starting system for Example 3. The variable curvatures are indicated with 'V', and solves are indicated with 'S'.*

Surface #	Curvature (mm^{-1})		Thickness (mm)		Glass type
object	0.00000		infinity		AIR
1	-0.00613	V	8.74666		BSM24_OHARA
2	-0.00537	V	0.29818		AIR
3	0.03366	V	12.42423		SK1_SCHOTT
4	0.00000		3.77697		F15_SCHOTT
5	0.02515	V	15.10786		AIR
6 (STOP)	0.00000		12.92120		AIR
7	-0.06005	V	3.77697		F15_SCHOTT
8	0.00000		10.83393		SK1_SCHOTT
9	-0.03422	V	0.29818		AIR
10	0.00353	V	6.85817		SK1_SCHOTT
11	-0.01553	S	68.79122	S	AIR
image	0.00000		0.00000		

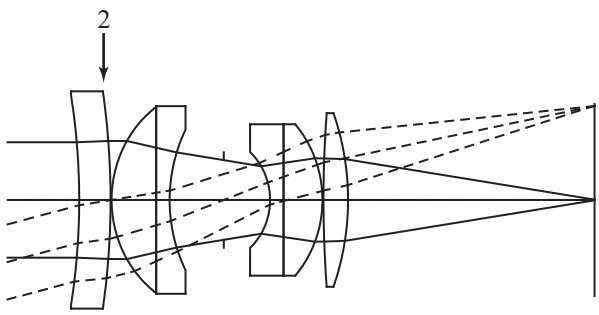


Figure A.9: *Starting system for Example 3.*

Step 3: Change the two curvatures of the null-element as follows: $c_3 = c_4 = c_2 + \epsilon$, where $\epsilon = 0.00003$ mm^{-1} (i.e. reduce the negative curvature; when working with radii, make the two radii a negative number with larger absolute value). For clarity, the curvature change is shown exaggerated in Figure A.10.

Step 4: Optimize the system that has been obtained after step 3.

Step 5: Increase the thickness of lens resulting from the null-element to the same value as that of the first lens (here, this can be done in one step), and reoptimize.

Step 6: Decrease the thickness of the first lens to zero, and reoptimize.

Step 7: Make the curvature of the second surface equal to the one of the first surface, and reoptimize the system with the first two curvatures fixed.

Step 8: Delete the first lens, and reoptimize the system. The resulting system (the last configuration shown in Figure A.10) has a merit function that is much lower than that of the starting one and the shape resembles the well-known Double Gauss design.

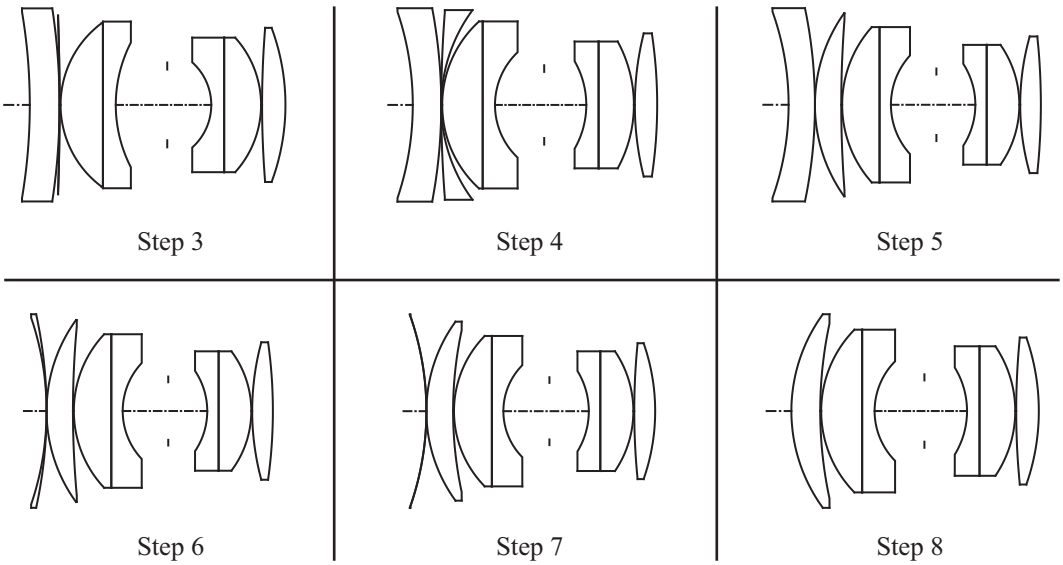


Figure A.10: *Steps 3, 4, 5, 6, 7, and 8 for SPC in Example 3.*

Bibliography

[1] G. W. Forbes and M. A. Alonso, "What on earth is a ray and how can we use them best?" in *International Optical Design Conference 1998*, L. R. Gardner and K. P. Thompson, eds., vol. 3482 of *Proc. SPIE*, pp. 22–31 (1998).

[2] W. T. Welford, *Aberrations of optical systems* (Adam Hilger, Bristol, England, 1986).

[3] H. A. Buchdahl, *An introduction to Hamiltonian optics* (Cambridge University Press, New York, 1970).

[4] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman lectures on physics*, vol. I (Addison-Wesley, Reading, MA, 1963).

[5] P. J. Sands, "Third-order aberrations of inhomogeneous lenses," J. Opt. Soc. Am. **60**, 1436–1443 (1970).

[6] H. Gross, *Aberrations*, vol. 1 of *Handbook of Optical Systems* (Wiley-VCH, Weinheim, 2005).

[7] F. Bociort, *Optical system optimization*, Encyclopedia of optical engineering (Marcel Dekker, New York, 2003).

[8] D. Sturlesi and D. C. O'Shea, "Global view of optical design space," Opt. Eng. **30**(2), 207–218 (1991).

[9] J. Tesar, R. Liang, and M. Mansuripur, "Optical modeling combining geometrical ray tracing and physical-optics software," Opt. Eng. **39**(7), 1845–1849 (2008).

[10] F. Bociort, E. van Driel, and A. Serebriakov, "Networks of local minima in optical system optimization," Opt. Lett. **29**, 189–191 (2004).

[11] D. C. Sinclair, *Optical design software*, vol. 1 of *Handbook of Optics, Fundamentals, Techniques, and Design*, 2nd ed. (McGraw-Hill, New York, 1995).

[12] H. Gross, H. Zügge, M. Peschka, and F. Blechinger, *Principles of optimization*, vol. 3 of *Handbook of Optical Systems* (Wiley-VCH, Weinheim, 2007).

[13] Optical Research Associates, CODE V.

[14] D. P. Feder, "Automatic optical design," Appl. Opt. **2**, 1209–1226 (1963).

[15] J. R. Shewchuk, "An introduction to the conjugate gradient method without the agonizing pain," Tech. Rep. CMU-CS-94-125, School of Computer Science, Carnegie Mellon University

[16] Z. Gurdal, "Fletcher-Reeves conjugate gradient algorithm," URL http://www.esm.vt.edu/~zgurdal/COURSES/4084/4084-Docs/Tutorial.html, accessed 20 March 2009.

[17] J. Pujol, "The solution of nonlinear inverse problems and the Levenberg-Marquardt method," Geophysics **72**, W1–W16 (2007).

[18] X. Cheng, Y. Wang, Q. Hao, and M. Isshiki, "Global and local optimization for optical systems," Optic **117**, 111–117 (2006).

[19] T. G. Kuper and T. I. Harris, "Global optimization for lens design - an emerging technology," in *Lens and optical system design*, H. Zuegge, ed., vol. 1780 of *Proc. SPIE*, pp. 14–28 (1992).

[20] G. W. Forbes and A. E. Jones, "Towards global optimization with adaptive simulated annealing," in *1990 Intl Lens Design Conf*, G. N. Lawrence, ed., vol. 1354 of *Proc. SPIE*, pp. 144–153 (1991).

[21] M. Isshiki, H. Ono, K. Hiraga, J. Ishikawa, and S. Nakadate, "Lens design: Global optimization with Escape Function," Optical Review (Japan) **6**, 463–470 (1995).

[22] K. E. Moore, "Algorithm for global optimization of optical systems based on genetic competition," in *Optical Design and Analysis Software*, R. C. Juergens, ed., vol. 3780 of *Proc. SPIE*, pp. 40–47 (1999).

[23] L. W. Jones, S. H. Al-Sakran, and J. R. Koza, "Automated synthesis of both the topology and numerical parameters for seven patented optical lens systems using genetic programming," in *Current Developments in Lens Design and Optical Engineering VI*, P. Z. Mouroulis, W. J. Smith, and R. B. Johnson, eds., vol. 5874 of *Proc. SPIE*, p. 587403 (2005).

[24] J. P. McGuire, Jr., "Designing easily manufactured lenses using a global method," in *International Optical Design Conference 2006*, G. G. Gregory, J. M. Howard, and R. J. Koshel, eds., vol. 6342 of *Proc. SPIE*, p. 63420O (2006).

[25] J. R. Rogers, "Using global synthesis to find tolerance-insensitive design forms," in *International Optical Design Conference 2006*, G. G. Gregory, J. M. Howard, and R. J. Koshel, eds., vol. 6342 of *Proc. SPIE*, p. 63420M (2006).

[26] M. Laikin, *The method of lens design*, Lens design, 4th ed. (CRC Press, Boca Raton, FL, 1996).

[27] A. Serebriakov, F. Bociort, and J. Braat, "Finding new local minima by switching merit functions in optical system optimization," Opt. Eng. **44**, 100501 (2005).

[28] S. W. Weller, "Simulated Annealing: What good is it?" in *Current developments in optical engineering II*, R. E. Fisher and W. J. Smith, eds., vol. 0818 of *Proc. SPIE*, pp. 265–274 (1987).

[29] F. Cuevas, O. Gonzalez, Y. Susuki, D. Hernandez, M. Rocha, and N. Alcala, "Genetic algorithms applied to optics and engineering," in *Fifth Symposium Optics in Industry*, E. Rosas, R. Cardoso, J. C. Bermudez, and O. Barbosa-García, eds., vol. 6046 of *Proc. SPIE*, p. 60461K (2006).

[30] J. Beaulieu, C. Gagne, and M. Parizeau, "Lens system design and re-engineering with evolutionary algorithms," in *GECCO 2002: Proceedings of the Genetic and Evolutionary Computation Conference*, pp. 155–162 (2002).

[31] S. H. Al-Sakran, J. R. Koza, and L. W. Jones, "Automated re-invention of a previously patented optical lens system using genetic programming," in *Proceedings of the 8th European Conference on Genetic Programming*, M. Keijzer, A. Tettamanzi, P. Collet, J. I. van Hemert, and M. Tomassini, eds., vol. 3447 of *Lecture Notes in Computer Science*, pp. 25–37 (2005).

[32] J. R. Koza, S. H. Al-Sakran, and L. W. Jones, "Automated re-invention of six patented optical lens systems using genetic programming," in *GECCO 2005: Proceedings of the 2005 conference on Genetic and evolutionary computation*, vol. 2, pp. 1953–1960 (2005).

[33] L. W. Jones, S. H. Al-Sakran, and J. R. Koza, "Automated synthesis of a human-competitive solution to the challenge problem of the 2002 international optical design conference by means of genetic programming and a multi-dimensional mutation operation," in *GECCO '06: Proceedings of the 8th annual conference on Genetic and evolutionary computation*, pp. 823–830 (2006).

[34] M. Isshiki, "Global optimization with escape function," in *International Optical Design Conference 1998*, L. R. Gardner and K. P. Thompson, eds., vol. 3482 of *Proc. SPIE*, pp. 104–109 (1998).

[35] Lambda Research Corporation, OSLO.

[36] M. J. Kidger, "Techniques for handling difficult lens design problems," in *Recent Trends In Optical Systems Design; Computer Lens Design Workshop*, C. Londono and R. E. Fisher, eds., vol. 0766 of *Proc. SPIE*, pp. 27–31 (1987).

[37] W. J. Smith, *Modern optical engineering: the design of optical systems* (McGraw-Hill, New York, 2000).

[38] E. van Driel, F. Bociort, and A. Serebriakov, "Topography of the merit function landscape in optical system design," in *Optical Design and Engineering*, L. Mazuray, P. J. Rogers, and R. Wartmann, eds., vol. 5249 of *Proc. SPIE*, pp. 353–363 (2004).

[39] F. Bociort, A. Serebriakov, and M. van Turnhout, "Saddle points in the merit function landscape of systems of thin lenses in contact," in *Current Developments in Lens Design and Optical Engineering V*, P. Z. Mouroulis, W. J. Smith, and R. B. Johnson, eds., vol. 5523 of *Proc. SPIE*, pp. 174–184 (2004).

[40] F. Bociort and M. van Turnhout, "Generating saddle points in the merit function landscape of optical systems," in *Optical Design and Engineering II*, L. Mazuray and R. Wartmann, eds., vol. 5962 of *Proc. SPIE*, p. 59620S (2005).

[41] F. Bociort and M. van Turnhout, "Looking for order in the optical design landscape," in *Current Developments in Lens Design and Optical Engineering VII*, P. Z. Mouroulis, W. J. Smith, and R. B. Johnson, eds., vol. 6288 of *Proc. SPIE*, p. 628806 (2006).

[42] O. Marinescu and F. Bociort, "Network search method in the design of EUV lithographic objectives," Appl. Opt. **46**, 8385–8393 (2007).

[43] O. Marinescu and F. Bociort, "Saddle-point construction in the design of lithographic objectives, part 1: method," Opt. Eng. **47**, 093002 (2008).

[44] O. Marinescu and F. Bociort, "Saddle-point construction in the design of lithographic objectives, part 2: application," Opt. Eng. **47**, 093003 (2008).

[45] E. W. Weissten, *Monkey saddle*, CRC concise encyclopedia of mathematics (CRC Press, Boca Raton, FL, 1999). URL http://mathworld.wolfram.com/MonkeySaddle.html.

[46] G. S. Fulcher, "Telescope objective without spherical aberration for large apertures, consisting of four crown glass lenses," J. Opt. Soc. Am. **36**, 47–53 (1947).

[47] D. Shafer, "Optical design and the relaxation response," in *Recent Trends in Optical Systems Design; Computer Lens Design Workshop*, C. Londono and R. E. Fisher, eds., vol. 0766 of *Proc. SPIE*, pp. 2–9 (1987).

[48] R. E. Fischer and K. L. Mason, "Spherical aberration: some fascinating observations," in *Recent Trends in Optical Systems Design; Computer Lens Design Workshop*, C. Londono and R. E. Fisher, eds., vol. 0766 of *Proc. SPIE*, pp. 53–60 (1987).

[49] URL http://www.optica.tn.tudelft.nl/users/bociort/SPC_guide.zip, accessed 20 March 2009.

[50] ZEMAX Development Corporation, ZEMAX.

[51] T. Poston and I. Steward, *Catastrophe Theory and its Applications* (Dover, Mineola, NY, 1996).

[52] URL http://www.optica.tn.tudelft.nl/users/bociort/networks.html, accessed 20 March 2009.

[53] D. P. L. Castrigiano and S. A. Hayes, *Catastrophe Theory*, 2nd ed. (Westview Press, Boulder, CO, 2004).

[54] T. H. Koornwinder, *The topological degree of a mapping*, vol. 1 of *Nonlinear Analysis* (Mathematisch Centrum, Amsterdam, 1980).

[55] J. Brok, *An analytic approach to electromagnetic scattering problems* (Ph.D. Dissertation, Delft University of Technology, 2007). URL http://www.library.tudelft.nl/dissertations.

[56] E. Ott, *Chaos in dynamical systems*, 2nd ed. (Cambridge University Press, Cambridge, 2002).

[57] C. Grebogi, S. W. McDonald, E. Ott, and J. A. Yorke, "Final state sensitivity: an obstruction to predictability," Phys. Lett. A **99**, 415–418 (1983).

[58] S. W. McDonald, C. Grebogi, E. Ott, and J. A. Yorke, "Fractal basin boundaries," Physica D **17**, 125–153 (1985).

[59] C. Grebogi, E. Kostelich, E. Ott, and J. A. Yorke, "Multi-dimensioned intertwined basin boundaries: basin structure of the kicked double rotor," Physica D **25**, 347–360 (1987).

[60] C. Grebogi, E. Ott, and J. A. Yorke, "Chaos, strange attractors, and fractal basin boundaries in non-linear dynamics," Science **238**, 632–638 (1987).

[61] H. E. Nusse and J. A. Yorke, "Basins of attraction," Science **271**, 1376–1380 (1996).

[62] S. N. Rasband, *Chaotic dynamics of nonlinear systems* (Wiley, New York, 1990).

[63] H.-O. Peitgen, H. Jürgens, and D. Saupe, *Chaos and fractals: New frontiers of science*, 2nd ed. (Springer-Verlag, New York, 2004).

[64] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems* (Prentice-Hall, Englewood Cliffs, NJ, 1974).

[65] B. Davies, *Period Doubling*, Encyclopedia of Nonlinear Science (Routledge, New York, 2004).

[66] Y.-C. Lai and C. Grebogi, "Converting transient chaos into sustained chaos by feedback control," Phys. Rev. E **49**(2), 1094–1098 (1994).

[67] URL http://www.optica.tn.tudelft.nl/users/bociort/paths.avi, accessed 20 March 2009.

Acknowledgments

Although there is only one author mentioned on the cover of this thesis, many people were involved with the completion of this work. Here, I would like to express my gratefulness to all of them.

First of all, I would like to thank Joseph Braat for his confidence in me, and for the encouragement he has given me in the first period of my Ph.D. research. I am very grateful for his help with the program OPTSYS, which enabled us to improve our understanding of the behavior of commercial software tools. Paul Urbach, who took over the role of promotor from Joseph, was of great help in the finishing part of my project. His pragmatic way of solving problems was highly appreciated. I give special thanks to Florian Bociort, my daily supervisor, for guiding me through my master and Ph.D. research. I have learned a lot from him through the many discussions we have had. Although he sometimes drove me nuts in delaying deadlines, the resulting extra scientific input was always improving our work. Without his help, this thesis would not have appeared in its present form.

Here, I would also like to express my sincere thanks to Roland Horsten, whose support was more than just being a 'computer helpdesk'. No matter how busy he was, he always made time to help in all sorts of problems. Thank you Roland! I thank Rob Pols, who has been supplying jokes on daily basis, and introduced me to the world of powerboat racing. I hope that one day I will see his own remote controlled catamaran flying over the water! Yvonne van Aalst and Lucia Heijenga deserve a special word of gratitude. Their help in all sorts of issues, paper work, and missing soccer cards was priceless.

I am very grateful to the Dutch Technology Foundation STW for their financial support of this Ph.D. project. I acknowledge the participants in the biannual STW user committee meetings for many fruitful discussions: Bob Kruizinga from TNO Science and

Industry, Arie den Boef from ASML Netherlands B.V., and Marcel Krijn from Philips Research.

I would also like to thank Herbert Gross and Christoph Menke from Carl Zeiss AG for valuable input and discussions.

My research would have been impossible without the pleasant atmosphere given by all members of the Optics Research Group. It has been a great pleasure to be part of this cheerful group! I first want to thank my (ex-)roommates Oana Marinescu, Alexander (Sasha) Serebriakov, Pascal van Grol, Chiungtze (Joanne) Wang, and Eco van Driel. Without their contributions, this research would not be in its current state. I am greatly indebted to Oana, because she demonstrated that the saddle-point construction method can be applied to extreme ultraviolet mirror systems, and deep ultraviolet lithographic objectives as well. I also want to thank Oana for all the encouragement and help she has given me, and for all the nice discussions we have had. The help Sasha gave me with the CODE V macro language was very much appreciated. I want to thank Pascal for all his support and input in our research. We have all shown that optical system design is not scientifically dead – far from it!

From my colleagues in the Optics Group, I especially want to thank all (ex-)players of the Optics foosball team. Our daily battle for supremacy for the Orb of Greatness was an enjoyable time! Julien, who was never afraid of doing 'risky' things (but please warn us before your next jump into the IJsselmeer!), and always willing to party, was by no question the fun-maker in the group. The conferences I attended with him were the best! Aurèle, who likes a journey through the optical merit function landscape better than a systematic analysis, was always very helpful with air pressure checks. I will not be surprised to see him as the new group leader some day in the future. Flown from America, cowboy Joe ("Hold on to your saddles!"), brought some special humor in the group. Besides a grammar checker, he was also a gifted artist and his 'peanut-butter chaos' will make him famous one day. His hilarious orders to other people always made me laugh. For Pascal, each foosball match was a matter of life and death. His rhetorical question "Should I remind you the score?" was a desperate attempt to wake up his team-mate. Here, I also want to thank Gopakumar for his enthusiastic support of our foosball games.

All the other colleagues in the Optics Group, thank you as well! I want to thank Man and Aura for their great help with LaTeX problems and doctorate regulations, and Sven for all his efforts in showing me how simple things can be done in Adobe Illustrator. I want to give special thanks to Moumoune (Mounir), always ready to help other people (in a peaceful way). I thank Raymond for all his Friday afternoon experiments and his brave attempt to introduce 'spruitjes' during Christmas dinner! I will miss the numerous discussions about photography with Sjoerd (and Roland)! Reshmi, Olaf, Maarten S., Morris, Omar, Alberto, Milija, Nandini, Paul P., Sylvania, Peter, Thim, Marnix, and all ex-members of the Optics group, you all contributed to the nice time I had in the group. Thank you all!

Finally, I want to thank my family and friends. Without their help and moral support, it would never been possible to accomplish this work.

Delft, March 2009 Maarten van Turnhout

Win or lose, sink or swim.
One thing is certain, we'll never give in.
Side by side, hand in hand.
We all stand together.

Paul McCartney & The Frog Chorus – We all stand together (1985).

About the author

Maarten van Turnhout was born in 1979 in Rotterdam, the Netherlands. In 1999, he started to study Applied Physics at Delft University of Technology. His master project was carried out at the Optics Research Group under the guidance of Dr. Bociort, his present copromotor. The subject was saddle points in the merit function landscape of doublet and triplet lens designs. In 2004, he graduated *cum laude*, and continued his research as an associate researcher in the same group. In April 2005, he started his Ph.D. project in the Optics Research Group at Delft University of Technology under supervision of Prof. Dr. H.P. Urbach and Dr. F. Bociort. In this project, the optical merit function landscape was analyzed in order to find characteristics that could improve optimization methods in optical system design. The main results of this research are presented in this thesis, entitled: A Systematic Analysis of the Optical Merit Function Landscape: Towards Improved Optimization Methods in Optical Design.

During his thesis, Maarten was involved in the supervision of two students. He also performed academic tasks in the form of practica and correcting exams for the courses of 'Waves', and as supervisor in the group project-oriented course 'Introduction to Applied Physics'. The results of his Ph.D. studies have been published in several papers and presented at international conferences in Europe and the United States. During the IST Science Day on October 6, 2008 in Delft, Maarten received the Best Poster Prize of the Department of Imaging Science and Technology for the poster "Looking for order in the optical design landscape," because of the best design, content and presentation.

List of publications

Refereed publications

1. M. van Turnhout and F. Bociort, "Instabilities and fractal basins of attraction in optical system optimization," *Opt. Express* **17**, 314–328 (2009).

2. F. Bociort and M. van Turnhout, "Finding new local minima in lens design landscapes by constructing saddle points," submitted to *Opt. Engineering*.

3. M. van Turnhout and F. Bociort, "Chaotic behavior in an algorithm to escape from poor local minima in lens design," submitted to *Opt. Express*.

4. M. van Turnhout and F. Bociort, "The saddle-point construction method in the general case," in preparation.

Patents

1. F. Bociort, M. van Turnhout, and O. Marinescu, "Method of designing a projection system, lithographic apparatus and device manufacturing method" US Patent 2007/0165296A1 (2007).

Conference proceedings

1. 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**, 174–184 (2004).

2. F. Bociort and M. van Turnhout, "Generating saddle points in the merit function landscape of optical systems," *Proc. SPIE* **5962**, 0S1–8 (2005).

3. F. Bociort and M. van Turnhout, "Looking for order in the optical design landscape," *Proc. SPIE* **6288**, 628806 (2006).

4. F. Bociort, M. van Turnhout and O. Marinescu, "Practical guide to saddle-point construction in lens design," *Proc. SPIE* **6667**, 666708 (2007).

5. M. van Turnhout and F. Bociort, "Predictability and unpredictability in optical system optimization," *Proc. SPIE* **6667**, 666709 (2007).

Conference contributions

1. M. van Turnhout and F. Bociort, "Predictability and unpredictability in optical system optimization," oral presentation at the 3rd European Optical Society (EOS) Topical Meeting on Advanced Imaging Techniques, Lille, France (September 2007).

2. M. van Turnhout, F. Bociort, and P. van Grol, "Looking for order in the optical design landscape," poster presentation at IOP Fotonica Evenement, 2008, Nieuwegein, The Netherlands (April 2008).

3. M. van Turnhout, F. Bociort, and P. van Grol, "Looking for order in the optical design landscape," poster presentation at the IST Science Day, 2008, Delft, The Netherlands (October 2008).

4. M. van Turnhout, P. van Grol, and F. Bociort, "Systematically obtaining new local minima in lens design landscapes," oral presentation at the 4rd European Optical Society (EOS) Topical Meeting on Advanced Imaging Techniques, Jena, Germany (June 2009).

5. P. van Grol, M. van Turnhout, and F. Bociort, "Finding order in an optical design landscape," poster presentation at the 4rd European Optical Society (EOS) Topical Meeting on Advanced Imaging Techniques, Jena, Germany (June 2009).

Other

1. F. Bociort and M. van Turnhout, "Saddle points reveal essential properties of the merit-function landscape," SPIE Newsroom (2008).

2. M. van Turnhout, F. Bociort, and P. Urbach, "Nieuwe oplossingen in lensontwerp door het construeren van zadelpunten," Fotonica Magazine (March 2009).

3. M. van Turnhout and F. Bociort, "Are well-known optical design tools sometimes unpredictable?" SPIE Newsroom (to be published, 2009).