Design and Correction of Optical Systems

Lecture 9: Optimization and correction
2018-06-11
Herbert Gross
<table>
<thead>
<tr>
<th>Week</th>
<th>Date</th>
<th>Topic</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>09.04</td>
<td>Basics</td>
<td>Law of refraction, Fresnel formulas, optical system model, raytrace, calculation approaches</td>
</tr>
<tr>
<td>2</td>
<td>16.04</td>
<td>Materials and Components</td>
<td>Dispersion, anormal dispersion, glass map, liquids and plastics, lenses, mirrors, aspheres, diffractive elements</td>
</tr>
<tr>
<td>3</td>
<td>23.04</td>
<td>Paraxial Optics</td>
<td>Paraxial approximation, basic notations, imaging equation, multi-component systems, matrix calculation, Lagrange invariant, phase space visualization</td>
</tr>
<tr>
<td>4</td>
<td>30.04</td>
<td>Optical Systems (Yi Zhong)</td>
<td>Pupil, ray sets and sampling, aperture and vignetting, telecentricity, symmetry, photometry</td>
</tr>
<tr>
<td>5</td>
<td>07.05</td>
<td>Geometrical Aberrations</td>
<td>Longitudinal and transverse aberrations, spot diagram, polynomial expansion, primary aberrations, chromatical aberrations, Seidels surface contributions</td>
</tr>
<tr>
<td>6</td>
<td>14.05</td>
<td>Wave Aberrations (Johannes Stock)</td>
<td>Fermat principle and Eikonal, wave aberrations, expansion and higher orders, Zernike polynomials, measurement of system quality</td>
</tr>
<tr>
<td>7</td>
<td>28.05</td>
<td>PSF and Transfer function</td>
<td>Diffraction, point spread function, PSF with aberrations, optical transfer function, Fourier imaging model</td>
</tr>
<tr>
<td>8</td>
<td>04.06</td>
<td>Further Performance Criteria</td>
<td>Rayleigh and Marechal criteria, Strehl definition, 2-point resolution, MTF-based criteria, further options</td>
</tr>
<tr>
<td>9</td>
<td>11.06</td>
<td>Optimization and Correction</td>
<td>Principles of optimization, initial setups, constraints, sensitivity, optimization of optical systems, global approaches</td>
</tr>
<tr>
<td>10</td>
<td>18.06</td>
<td>Correction Principles I</td>
<td>Symmetry, lens bending, lens splitting, special options for spherical aberration, astigmatism, coma and distortion, aspheres</td>
</tr>
<tr>
<td>11</td>
<td>25.06</td>
<td>Correction Principles II (subs/shift)</td>
<td>Field flattening and Petzval theorem, chromatical correction, achromate, apochromate, sensitivity analysis, diffractive elements</td>
</tr>
<tr>
<td>12</td>
<td>02.07</td>
<td>Optical System Classification</td>
<td>Overview, photographic lenses, microscopic objectives, lithographic systems, eyepieces, scan systems, telescopes, endoscopes</td>
</tr>
<tr>
<td>13</td>
<td>09.07</td>
<td>Special System Examples</td>
<td>Zoom systems, confocal systems</td>
</tr>
</tbody>
</table>
Contents

1. Principles of nonlinear optimization
2. Basic mathematics in local nonlinear optimization
3. Optimization in optical design
4. Global optimization methods
5. Miscellaneous aspects
Basic Idea of Optimization

- Topology of the merit function in 2 dimensions
- Iterative down climbing in the topology
Optimization

Complex topology of the merit function
Mathematical description of the problem:

- n variable parameters
- m target values
- Jacobi system matrix of derivatives, Influence of a parameter change on the various target values, sensitivity function
- Scalar merit function
- Gradient vector of topology
- Hesse matrix of 2nd derivatives

\[
\begin{align*}
\vec{x} \\
\vec{f}(\vec{x}) \\
J_{i,j} = \frac{\partial f_i}{\partial x_j} \\
F(\vec{x}) &= \sum_{i=1}^{m} w_i \cdot [y_i - f(\vec{x})]^2 \\
g_j &= \frac{\partial F}{\partial x_j} \\
H_{j,k} &= \frac{\partial^2 F}{\partial x_j \partial x_k}
\end{align*}
\]
Nonlinear Optimization

- Linearized environment around working point
  Taylor expansion of the target function
  \[ \bar{f} = f_0 + J \cdot \bar{x} \]

- Quadratical approximation of the merit function
  \[ F(\bar{x}) = F(\bar{x}_0) + J \cdot \Delta \bar{x} + \frac{1}{2} \Delta \bar{x} \cdot H \cdot \Delta \bar{x} \]

- Solution by linear Algebra
  system matrix \( A \)
  cases depending on the numbers of \( n / m \)
  \[ A^+ = \begin{cases} A^{-1} & \text{if } m = n \\ \left( A^T A \right)^{-1} \cdot A^T & \text{if } m > n \text{ (under determined)} \\ A^T \cdot \left( AA^T \right)^{-1} & \text{if } m < n \text{ (over determined)} \end{cases} \]

- Iterative numerical solution:
  Strategy: optimization of
  - direction of improvement step
  - size of improvement step
Linear algebra of optimization:

1. Exact solvable
2. Over determined: more parameter than targets
   solution not unique
3. Under determined: more targets than parameter
   best approximation solution

\[
\begin{align*}
A^+ &= \begin{cases} 
\frac{A^{-1}}{m=n \text{ (exact)}} \\
(A^T A)^{-1} \cdot A^T & \text{if } m > n \text{ (over determined)} \\
A^T \cdot (AA^T)^{-1} & \text{if } m < n \text{ (under determined)}
\end{cases}
\end{align*}
\]
Derivative vector in merit function topology:
Necessary for gradient-based methods

Numerical calculation by finite differences

Possibilities and accuracy

\[ g_{jk} = \frac{\partial f_j(\bar{x})}{\partial x_k} = \nabla_{x_k} f_j(\bar{x}) \]

\[ g_{jk} = \frac{f_{j}^{\text{right}} - f_j}{\Delta x_k} \]

- forward
- central exact
- backward

\[ f_j(x_k), f_{j-1}, f_j, f_{j+1}, f_{j+1}(x_k) \]
Optimization Principle for 2 Degrees of Freedom

- Aberration depends on two parameters
- Linearization of sensitivity, Jacobian matrix
  Independent variation of parameters
- Vectorial nature of changes:
  Size and direction of change
- Vectorial decomposition of an ideal step of improvement,
  linear interpolation
- Due to non-linearity:
  change of Jacobian matrix, next iteration gives better result
Optimization Damping

- Damping with factor $l$

$$\Delta x_j = -\left(J_{ji} \cdot J_{ij} + \lambda \cdot I_{kk}\right)^{-1} \cdot J_{jk} \cdot f_k$$

- Damping defines the orientation and the size of the improvement step

Ref: C. Menke
Steepest Descent Method

- Control function

\[ \Phi = \bar{f}_0^T \cdot \bar{f}_0 + 2 \bar{g}^T \cdot \Delta \bar{x} + \lambda \cdot \Delta \bar{x}^T \cdot \Delta \bar{x} \]

- Gradient method with steepest descent

\[ \bar{g} + \lambda \cdot \Delta \bar{x} = 0 \]

\[ \Delta \bar{x} = -\frac{\bar{g}}{\lambda} = -\nabla_{\bar{x}} F(\bar{x}) \]

- Changing directions: zig-zac-path with poor convergence

- Optimal damping of step size

\[ \lambda = \frac{\bar{g}^T \cdot J^T \cdot J \cdot \bar{g}}{\bar{g}^T \cdot \bar{g}} \]
Effect of constraints

Effect of Constraints on Optimization

initial point

global minimum

local minimum

$x_1$

constraint

$x_1 < 0$

path without constraint

path with constraint

0

initial point

local minimum

global minimum

constraint $x_1 < 0$
Boundary Conditions and Constraints

- Types of constraints
  1. Equation, rigid coupling, pick up
  2. One-sided limitation, inequality
  3. Double-sided limitation, interval

- Numerical realizations:
  1. Lagrange multiplier
  2. Penalty function
  3. Barrier function
  4. Regular variable, soft-constraint
Local working optimization algorithms

- **Nonlinear optimization methods**
  - Methods without derivatives
    - Simplex
    - Conjugate directions
  - Derivative based methods
    - Single merit function
    - Descent methods
      - Steepest descents
      - Variable metric
    - No single merit function
      - Adaptive optimization
      - Nonlinear inequalities
  - Least squares
    - Undamped
      - Line search
    - Damped
      - Additive damping
      - Multiplicative damping
      - Orthonormalization
      - Second derivative
    - Conjugate gradient
    - Davidon Fletcher
Local Optimization Algorithms

- Gauss-Newton method
  Normal equations
  System matrix

- Damped least squares method (DLS)
  Damping reduces step size, better convergence without oscillations

- ACM method according to E. Glatzel
  Special algorithm with reduced error vector

- Conjugate gradient method
  Reduction of oscillations

\[ \Delta \bar{x} = -\left( J^T \cdot J \right)^{-1} \cdot J^T \cdot \Delta \bar{f} \]

\[ A = \left( J^T \cdot J \right)^{-1} \cdot J^T \]

\[ \Delta x_j = \left( J_{ij}^T \cdot J_{ij} + \lambda^2 \cdot I_{ij} \right)^{-1} \cdot J_{ij}^T \cdot \Delta f_i \]

\[ \Delta x_j = J_{ij}^T \cdot \left( J_{ij} \cdot J_{ij}^T \right)^{-1} \cdot \Delta f_i \]
Adaptation of direction and length of steps:
rate of convergence

Gradient method:
slow due to zig-zag
Speed of Optimization...

- Very slow
- Slow
- Quasi fast
- Fast
- Very fast
- Extreme fast
Optimization and Starting Point

- The initial starting point determines the final result
- Only the next located solution without hill-climbing is found
- Existing solution modified
- Literature and patent collections
- Principal layout with ideal lenses
  successive insertion of thin lenses and equivalent thick lenses with correction control

- Approach of Shafer
  AC-surfaces, monochromatic, buried surfaces, aspherics
- Expert system
- Experience and genius
First Paraxial Layout

- Fixation of distances and focal power distribution
- What is different in both initial setups?
  - final FoV identical
  - aperture identical
  - length nearly the same

Ref: B. Böhme
Pre-Calculations

- Zero-order properties of the system:
  - focal length
  - magnification
  - pupil size and location
  - size/length of the system, image location

- Pre-Calculation of the system structure, which is independent of lens bendings,
  Analytical conditions for these 3rd order corrections,
  Determination of focal power distribution
  - symmetry
  - field flattening correction
  - achromatism, apochromatism
  - distortion-correction
  - anastigmatism

- Lens bending with 3rd order lens contributions
  - spherical aberration
  - coma
  - astigmatism
Initial Conditions

Valid for object in infinity:

1. Total refractive power

\[ s_1 = -\infty \]

\[ F' = \sum_{m=1}^{M} \omega_m \sum_{n=1}^{N} F'_{nm} \]

2. Correction of Seidel aberrations

2.1 Dichromatic correction of marginal ray
axial achromatical

\[ \frac{F'}{\nu} = \sum_{m=1}^{M} \omega_m^2 \sum_{n=1}^{N} \frac{F'_{nm}}{\nu_{nm}} \]

2.2 Dichromatic correction of chief ray
achromatical lateral magnification

\[ \frac{F'}{\bar{\nu}} = \sum_{m=1}^{M} \omega_m \omega_{pm} \sum_{n=1}^{N} \frac{F'_{nm}}{\nu_{nm}} \]

2.3 Field flattening
Petzval

\[ \frac{F'}{n} = \sum_{m=1}^{M} \sum_{n=1}^{N} \frac{F'_{nm}}{n_{nm}} \]

2.4 Distortion correction according
to Berek

\[ 0 = \sum_{m=1}^{M} \omega_{pm} \sum_{n=1}^{N} F'_{nm} \]

3. Tri-chromatical correction
Secondary spectrum

\[ \frac{F' P}{\nu} = \sum_{m=1}^{M} \omega_m^2 \sum_{n=1}^{N} \frac{P_{nm} F'_{nm}}{\nu_{nm}} \]
Development in Optics

- System development flow chart
Goal of optimization:
Find the system layout which meets the required performance targets according to the specification.

Formulation of performance criteria must be done for:
- Apertur rays
- Field points
- Wavelengths
- Optional several zoom or scan positions

Selection of performance criteria depends on the application:
- Ray aberrations
- Spot diameter
- Waveform description by Zernike coefficients, rms value
- Strehl ratio, Point spread function
- Contrast values for selected spatial frequencies
- Uniformity of illumination

Usual scenario:
Number of requirements and targets quite larger than degrees of freedom,
Therefore only solution with compromise possible
Optimization in Optical Design

- **Merit function:**
  Weighted sum of deviations from target values

- **Formulation of target values:**
  1. fixed numbers
  2. one-sided interval (e.g. maximum value)
  3. interval

- **Problems:**
  1. linear dependence of variables
  2. internal contradiction of requirements
  3. initial value far off from final solution

- **Types of constraints:**
  1. exact condition (hard requirements)
  2. soft constraints: weighted target

- **Finding initial system setup:**
  1. modification of similar known solution
  2. Literature and patents
  3. Intuition, genius and experience

\[ \Phi = \sum_{j=1,m} g_j \cdot (f_j^{ist} - f_j^{soll})^2 \]
Characterization and description of the system delivers free variable parameters of the system:

- Radii
- Thickness of lenses, air distances
- Tilt and decenter
- Free diameter of components
- Material parameter, refractive indices and dispersion
- Aspherical coefficients
- Parameter of diffractive components
- Coefficients of gradient media

General experience:
- Radii as parameter very effective
- Benefit of thickness and distances only weak
- Material parameter can only be changes discrete
- Special problem in glass optimization: finite area of definition with discrete parameters $n$, $\nu$
- Restricted permitted area as one possible constraint
- Model glass with continuous values of $n$, $\nu$ in a pre-phase of glass selection, freezing to the next adjacent glass
Constraints in the optimization of optical systems:

1. Discrete standardized radii (tools, metrology)
2. Total track
3. Discrete choice of glasses
4. Edge thickness of lenses (handling)
5. Center thickness of lenses (stability)
6. Coupling of distances (zoom systems, forced symmetry,...)
7. Focal length, magnification, working distance
8. Image location, pupil location
9. Avoiding ghost images (no concentric surfaces)
10. Use of given components (vendor catalog, availability, costs)
Constraints on thickness values and distances

- Maximum center thickness of air: $MXCA$
- Minimum center thickness of air: $MNCA$
- Minimum edge thickness of air: $MNEA$
- Maximum center thickness of glass: $MXCG$
- Minimum center thickness of glass: $MNCG$
- Minimum edge thickness of glass: $MNEG$
Illustration of not useful results due to non-sufficient constraints
- Typical in optics:
  Twisted valleys in the topology
- Selection of local minima
- Typical merit function of an achromate
- Three solutions, only two are useful
Merit function of an achromate

- **Weight** $w$ of chromatic aberration
  - $w = 10$
  - $w = 10000$
  - $w = 100000$

- **Conjugation parameter** $M$
  - $M = 1$
  - $M = 6$
  - $M = 9$

- **Crown glass choice**
  - BK7 / SF5
  - LLF1 / SF5
  - FK54 / SF5

- **Flint glass choice**
  - BK7 / SF5
  - BK7 / SF58
  - BK7 / SK4
Optimization of Achromatic Glasses

- Perfect aplanatic line of corresponding glasses
- For one given flint a line indicates the usefull crown glasses and vice versa
Optimization: Achromate

- Bending of an achromate: spherical aberration
- 3 Parameters: n, X, M

\[ A_s = -\frac{1}{32n(n-1)f^3} \left[ \frac{n^3}{n-1} + \frac{1}{n-1} \left( X + \frac{2(n^2-1)}{n+2} M \right)^2 - \frac{n^2(n-1)}{n+2} M^2 \right] \]

- Merit function

relief
Global Optimization

- No unique solution
- Constraints not sufficient:
  fixed: unwanted lens shapes
- Many local minima with nearly the same performance

reference design : $F = 0.00195$
solution 5 : $F = 0.000266$
solution 11 : $F = 0.000470$
solution 1 : $F = 0.000102$
solution 6 : $F = 0.000273$
solution 12 : $F = 0.000510$
solution 2 : $F = 0.000160$
solution 7 : $F = 0.000296$
solution 13 : $F = 0.000513$
solution 3 : $F = 0.000210$
solution 8 : $F = 0.000312$
solution 14 : $F = 0.000519$
solution 4 : $F = 0.000216$
solution 9 : $F = 0.000362$
solution 15 : $F = 0.000602$
solution 10 : $F = 0.000384$
solution 16 : $F = 0.000737$
Global Optimization: Escape Function

- Method of local minima escape: temporarily added term to overcome local minimum

\[ \Delta F_{esc}(\tilde{x}) = \Delta F_0 \cdot e^{-\beta(F(\tilde{x})-F_0)^2} \]

- Optimization and adaptation of escape parameters

\[ \beta = 1.0 \]
\[ \beta = 0.01 \]
- Saddel points in the merit function topology
- Systematic search of adjacent local minima is possible
- Exploration of the complete network of local minima via saddelpoints
Optimization

- Saddel point method:
  - network of local minima connected by saddel points

- very saddelpoint has two neighboring local minima

Ref: F. Bociort
- Adding a meniscus lens with thickness zero:
  Two additional degrees of freedom
  Generates two adjacent local minima

- Different opportunities to generate a meniscus

Ref: F. Bociort
Optimization

- Network of local minima and saddle points for the triplet system

Ref: F. Bociort
- Network of local minima and saddel points for the triplet system
- Change of setup by thin-meniscus saddelpoint construction and 'walk'

Ref: F. Bociort
Saddel Point Method

- Example Double Gauss lens of system network with saddelpoints
Microscopic Objective Lens Initial System

- a) negative achromate
- b) reversed positive achromate
- c) symmetric to c) positive achromate
- d) non-infinite positive achromate
- e) AC-meniscus lenses
Initial System Influence

- Simple system of two lenses
- Criterion: spot on axis, one wavelength
- Starting with different radii of curvature: completely different solutions

<table>
<thead>
<tr>
<th>No.</th>
<th>Layout</th>
<th>Start radii</th>
<th>Rms spot size</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td></td>
<td>0 - 0 - 0 - 0α</td>
<td>184α</td>
</tr>
<tr>
<td>b</td>
<td></td>
<td>0 - 0 - 0 - 20α</td>
<td>39α</td>
</tr>
<tr>
<td>c</td>
<td></td>
<td>+20 - 0 - 0 - 0α</td>
<td>92α</td>
</tr>
</tbody>
</table>
- Non-uniqueness of design solutions

Ref: J. Bentley, SPIE 5875 (2005)
- Approximate number of spots over the field as a function of the number of lenses. Linear for small number of lenses. Depends on mono-/polychromatic design and aspherics.

- Diffraction limited systems with different field size and aperture.