Imaging properties of gradient-index lenses

Florian Bociort
# Contents

1. Introduction 9

2. Ray-tracing in inhomogeneous media 15
   2.1 General inhomogeneous media 15
      2.1.1 Differential equation of light rays in inhomogeneous media 15
      2.1.2 Invariants 18
      2.1.3 Refraction at a surface 20
   2.2 Ray-tracing in axial gradients 23
      2.2.1 Finite rays 23
      2.2.2 Special cases 24
      2.2.3 Paraxial approximation 26
   2.3 Ray-tracing in radial gradients 28
      2.3.1 Differential equations of light rays in radial gradients 28
      2.3.2 Paraxial approximation 31
      2.3.3 Special cases 35
   2.4 Analytic ray-tracing method for finite rays in radial gradients 38
      2.4.1 Decoupling the differential ray equations 39
      2.4.2 Ray path determination by power series expansion 41
      2.4.3 Curved end surfaces 47
      2.4.4 Test of accuracy 49
   2.5 Paraxial approximation for rotationally-symmetric optical systems 53
      2.5.1 Positions of object and image 53
      2.5.2 The nature of the paraxial approximation 55

3. Aberrations of gradient-index lenses 59
   3.1 Primary aberrations of rotationally-symmetric optical systems 61
      3.1.1 Types of primary aberrations 61
      3.1.2 Stop-shift formulae 66
   3.2 Decomposition of the transverse aberration in surface and transfer contributions
      3.2.1 The quasi-invariant 70
      3.2.2 Curved end surfaces 72
   3.3 Surface contributions of gradient-index lenses to the Seidel aberration coefficients
      3.3.1 Derivation of the Seidel coefficients 75
      3.3.2 Third-order terms at refraction 76
      3.3.3 Ordinary and inhomogeneous surface contributions 80
3.4 Transfer contributions of axial gradient-index lenses to the Seidel coefficients
   3.4.1 Third-order terms at transfer 84
   3.4.2 Axial transfer contributions 86
   3.4.3 Special cases 88

3.5 Transfer contributions of radial gradient-index lenses to the Seidel coefficients
   3.5.1 Third-order terms at transfer through radial gradients 92
   3.5.2 Transfer contributions of positive and negative gradients 94
   3.5.3 Transfer contributions of shallow gradients 97
   3.5.4 Test of the results 97
   3.5.5 Special case 99

3.6 Primary chromatic aberrations of gradient-index lenses 101
   3.6.1 The chromatic quasi-invariant 101
   3.6.2 Surface contributions 103
   3.6.3 Transfer contributions of axial gradients 104
   3.6.4 Transfer contributions of radial gradients 106

4. Aberration correction in optical systems with GRIN lenses 109
   4.1 Extended thin-lens approximation for radial gradient-index lenses 110
      4.1.1 Paraxial approximation 111
      4.1.2 Primary aberration coefficients 114
      4.1.3 Effects of the change of the lens parameters on the Seidel aberrations 124
   4.2 Thin-lens approximation for axial gradient-index lenses 139
   4.3 Simple optical systems with gradient-index lenses 143
      4.3.1 Aplanatic correction of a single radial gradient-index lens 143
      4.3.2 Correction of all Seidel aberrations with two gradient-index lenses 151

5. Fabrication of gradient-index media 155

6. Summary 157

Appendix A: Eighth-order analytic ray-tracing formulae for radial gradient-index lenses
   A.1 Formulae for ray position, direction and optical path length 159
      A.1.1 Summary of previously introduced formulae 159
      A.1.2 Formulae derived automatically 160
   A.2 Formulae for the z-coordinate of the ray-surface intersection point 166
   A.3 Comparison of the analytic ray-tracing results with CODE V 168
Appendix B: Summary of the formulae for Seidel and chromatic aberration coefficients

B.1 Ordinary surface contributions 171
B.2 Axial gradients 172
B.3 Positive and negative radial gradients 173
B.4 Shallow radial gradients 174

Appendix C: Tables of symbols and conventions of notation 175

C.1 Symbols 175
C.2 Conventions of notation 178

References 179
Zusammenfassung 187
1. Introduction

Gradient-index (GRIN) lenses are lenses fabricated from optically inhomogeneous materials, i.e. from materials where the refractive index varies from point to point. Studies in the past two decades have shown that the use of inhomogeneous materials in optical design may lead to a substantial reduction of the number of components as compared to purely homogeneous optical systems having the same specifications. Therefore, lens design with GRIN lenses alone or together with homogeneous lenses receives presently considerable attention.

If the refractive index distribution inside the GRIN lens is rotationally symmetric, many of the familiar imaging properties of homogeneous optical systems are retained. Thus, in the paraxial approximation, a conjugate image plane exists for any object plane, the transverse magnification being the same for all pairs of object and image points. Beyond paraxial approximation, the aberration types are also the same as for homogeneous lenses.

For the design of optical imaging systems, up to now two types of GRIN media having both a rotationally symmetric refractive index distribution have been found to be particularly useful: the axial gradients, where the refractive index is a function of the coordinate along the symmetry axis, and the radial gradients, where the index is changing with the distance to the axis.

In both cases, the presence of the inhomogeneous medium has two kinds of effects upon the imaging properties of the lens:

i) Surface effects. At the refraction of a ray at a spherical end surface, the local refractive index is a function of the height of the ray-surface intersection point. The local refractive index change produces an effect on the ray aberrations which is similar to that produced by the local change of surface curvature in the case of an homogeneous aspherical surface.

ii) Effects due to the ray transfer through the GRIN medium. Because of the inhomogeneity of the medium, the ray path inside the lens is curved. Thus, unlike homogeneous media, transfer through GRIN media influences the monochromatic and chromatic aberrations of the rays. In radial gradients, the medium contributes also to the total power of the lens. Therefore, with radial GRIN media focusing can be obtained also if the lens has plane end faces.
As compared to homogeneous media, varying the refractive index distribution inside the lens introduces additional degrees of freedom in the lens design process. These additional possibilities for controlling the aberrations can be used to improve the system performance over the equivalent homogeneous system. Alternatively, the number of elements required to meet fixed performance criteria can be reduced (for radial gradients usually by about a factor of three / 19/).

The fact that inhomogeneous media can have remarkable imaging properties became apparent as early as 1854 when James Clerk Maxwell demonstrated that a medium having a suitable spherical refractive index distribution can provide a nontrivial example of an absolute optical instrument in which every point of a region of space is sharply imaged in a corresponding image point. This example, presently only of academic interest, is known in the literature as Maxwell's "fish-eye" / 8/.

Inhomogeneous media occur frequently in nature. Examples are the earth atmosphere and the eyes of animals and insects. At the end of the previous century, studies of insect eyes have lead the entomologist Karl Exner and his brother Sigmund Exner to the investigation of imaging through inhomogeneous media. In 1886 they have found that cylindrical glass plates having a radial refractive index distribution are capable of focusing light and calculated the focal length of these plates / 20/. They have requested Otto Schott to make several inhomogeneous cylinders of this type.

Schott produced the GRIN material, the first one that has ever been fabricated, by rapidly cooling molten glass. The resulting glass was inhomogeneous but also anisotropic and the glass plates obtained therefrom acted like diverging lenses / 20/. Around 1900, R.W. Wood made from gelatine simple radial GRIN lenses with plane end faces and obtained with them sharp images of bright objects. However, because of the technological difficulties of producing GRIN materials and the absence in the first half of this century of the required computing power for tracing rays through GRIN media, the further development of GRIN optics was slow.

GRIN optics was reborn around 1969 with the successful formation of gradients in glass by diffusive ion exchange. Since that time a large amount of research has been devoted to this field. Many of the important historical and recent papers on GRIN optics have been collected in Ref. / 37/. At present, a considerable effort is spent on improving the fabrication capabilities of GRIN lenses. New processes are investigated in order to increase the overall index change and the depth of the gradients and to improve the control of the index profile shape, e.g. the sol-gel process and the fusion of thin layers of glass.
Nowadays, GRIN lenses are used in fiber couplers, compact disk objectives and endoscopes. Arrays of radial GRIN rods allow the design of very compact copiers and fax machines /35/. However, lens design studies and an appreciable number of patents (mainly from Japan and U.S.) show the potential usefulness of GRIN lenses also for traditional types of optical systems such as cameras, microscopes and binoculars. It is reasonable to expect that GRIN lenses will be employed in an increasing number in such systems if suitable gradient materials become available.

While the recent activity in GRIN optics has lead to considerable progress in the areas of design of various applications and of material fabrication, also a further development of the general theoretical foundations underlying lens design with gradients became necessary. This will be attempted in the present study, which aims to yield contributions in the following two directions:

i) the development of mathematical tools analogue for GRIN lenses to those currently used in homogeneous lens design,

ii) the acquirement of insight into the capabilities of GRIN lenses for correcting aberrations as well as into the inherent limitations of these capabilities.

During the lens design process, after some rough estimates have been made on the basis of the paraxial approximation, the parameters specifying the construction of the system are varied until the aberrations lie within certain boundaries. As in the homogeneous case, the formalism for analysing and correcting the aberrations of optical systems with GRIN lenses can be developed at three levels of approximation, which rely on finite ray tracing, primary aberrations and thin lens theory (Fig. 1.1).

As a rough description of the structure of this study, each main part (Chapters 2, 3 and 4) will be devoted to one of these approximation levels. Starting with finite ray-tracing, at each further level of approximation additional insight in the lens properties will be gained at the expense of quantitative accuracy.

In Chapter 2, after a review of light propagation in inhomogeneous media, paraxial and finite ray-tracing in axial and radial gradients will be described. While at present in lens
design with gradients finite ray-tracing is performed mostly by means of numerical methods /50/, in this study very accurate analytical methods will be developed for determining the ray path within the medium by power series expansions. It is known that analytical ray-tracing can have a substantial speed advantage over numerical methods /28/. Here, however, analytical ray tracing will be used mainly as starting point for determining expressions for aberration coefficients.

For axial gradients the derivation of ray-tracing formulae will be done in a rather straightforward manner starting from the first principles. For radial gradients, however, a method for decoupling the nonlinear differential ray equations will be first developed. Then, a perturbation method will be used for computing successive approximations for the ray path and optical path length. Moreover, for the first time, for radial gradients an analytical method will be used also for determining the intersection point of the curved rays within the medium with the spherical end surfaces. Thus, the entire ray-tracing process will be performed analytically.

Unlike other similar methods known from the literature /28/, the present derivation method can be easily translated into a computer algebra code for symbolic mathematics software, enabling thus in principle the calculation of contributions of any order of approximation. Here, contributions up to the eighth order will be calculated for all quantities of interest (Appendix A), which seems sufficient for tracing rays accurately enough in all practical applications at the presently attainable level of control of the refractive index distribution.

Chapter 3 will be devoted mainly to the derivation of expressions for the Seidel and chromatic paraxial aberration coefficients (i.e. the primary aberration coefficients) for axial and radial GRIN lenses. After more than two decades of work for developing adequate computation methods, at present the computation of these coefficients for practical applications is done only by using numerical methods or by some analytical approximations for special cases /58/. Despite of the considerable practical importance of these coefficients for lens design with gradients, because of the complexity of these methods, no commercially available optical design software provides values for the primary aberration coefficients of GRIN lenses. (Code V computes values for the surface contributions to the Seidel coefficients, but not for the transfer contributions /15/.)

In this study, a new method for deriving aberration coefficients for rotationally symmetric gradients on the basis of analytical ray tracing will be developed. Unlike previous at-
tempts, with the new method simple and accurate analytical expressions will be obtained for all primary aberrations of axial and radial GRIN lenses.

In fact, for axial gradients, expressions for the primary aberration coefficients have already been published by Sands /47/, /49/. However, because the derivation itself is complex and the final form of the results as well, it seems that these expressions are nowhere in use. In this study, expressions for the primary aberration coefficients will be obtained, which are in principle equivalent with the results of Sands, but have a simpler form.

For the first time, short analytic formulae for the primary aberrations will be obtained also for radial GRIN lenses. In this case, the derivation will consist of two stages. In the first stage, large algebraic expressions will be obtained as a first raw form of the aberration coefficients. For shortening these expressions, a heuristic method based on symmetry requirements will be developed. This method will turn out to be very powerful and will lead to remarkably simple final expressions for all primary aberrations. All formulae which are necessary for the numerical computation of the primary aberrations of axial and radial GRIN lenses will be summarized in Appendix B.

In order to obtain qualitative insight into aberration correction with GRIN lenses, in Chapter 4 generalizations of the well-known thin lens approximation will be developed both for radial and axial GRIN lenses.

At present, the understanding of the effect of transfer through the medium on the primary aberrations of axial and radial gradients as can be found in the literature is not very satisfactory (excepting, for radial gradients, the cases of Petzval curvature and axial color /2/). In this study, for both types of GRIN lenses, methods will be developed for investigating this effect by evaluating the lowest order contributions of the medium to primary aberrations. It will be then found that considering together the lowest order contributions from the surfaces and medium yields aberration expressions for the entire lens which have the nearly the same structure as in the homogeneous case.

For radial gradients, comparison with exact results will show that, for lenses which are not too thick and gradients which are not too strong, these approximate formulae give a good qualitative description of the variation of the Seidel aberrations when lens parameters are changed. On the other hand, the approximate formulae give the explicit dependence of the primary aberrations on the lens parameters and thus offer much more insight into the lens properties. Thus, it will be shown that one of the inherent limitations of the
correction capabilities occurring in the case of homogeneous lenses can be found also for radial gradients. For instance, if a lens is corrected both for spherical aberration and coma, then, within the domain of validity of the so-called extended thin lens approximation, it cannot be corrected for astigmatism, too.

For axial gradients it will be found that, within the domain of validity of the thin lens approximation, the primary aberrations of the GRIN lens are equivalent to those of a pair of homogeneous aspherical lenses in contact, having a common plane surface and having refractive indices and Abbe numbers equal to the corresponding axial GRIN values at the two end surfaces.

Simple examples of aberration correction will then be given where thin lens theory, primary aberrations and finite ray tracing are successively applied. In an example of apllanatic correction with a single radial GRIN lens, it will be shown how the thin lens approximation can be used to explain features of optimized solutions obtained with CODE V.

Finally, in Chapter 5 a brief description of the most important types of fabrication methods of GRIN lenses will show the present limits of the producible values of the lens parameters.

In the above description of the structure of this study, emphasis has been given to the new results. However, since important results known from the literature have been derived there in a variety of styles and notations, they will be newly derived here in order to unify the treatment and to ensure a reasonable degree of mathematical self-consistency of this study. It will be attempted in various cases to give a more concise and/or more general derivation than in the original source. It is therefore hoped that some of the new derivations of the older results will also shed some new light on the subject. For the convenience of reading, a list of the main symbols and their meaning is given in Appendix C. For other symbols, which are defined by formulae, a reference to the definition formula will be made whenever these symbols appear.

It should also be noted that, while the main formalism will be developed both for axial and for radial gradients, radial gradients will be treated in more detail. This is because the new results of this study will be mostly for this type of gradients. Thus, numerical examples will be given only for radial gradients.
2. Ray-tracing in inhomogeneous media

Tracing of paraxial and finite rays in optical systems are essential tools in optical design. Paraxial ray-tracing is used in aberration correction as a basis of comparison to indicate how large the aberrations of finite rays i.e. their departures from their ideal location are. The paraxial approximation is strictly valid only for rays whose path is situated in the immediate neighbourhood of the optical axis. However, for well-corrected optical systems the aberrations are small and paraxial ray-tracing gives a good general description for the imagery of the system. The relationships between system data and the positions and sizes of objects and images are determined therefore from paraxial calculations. When the general form of the optical system is known, but the aberrations have to be reduced to acceptable tolerance levels, finite ray-tracing is used for analysing the system's performances and for refining them during the optimisation process.

In this chapter, ray-tracing methods for gradient-index (GRIN) lenses are presented. For general refractive index distributions, the differential ray equations must be solved numerically. For axial and radial refractive index distributions paraxial ray-tracing formulae are derived and for finite ray-tracing, as an alternative to numerical methods, approximate analytic formulae are found.

2.1 General inhomogeneous media

2.1.1 Differential equation of light rays in inhomogeneous media

In this section several forms of the differential equation describing the ray path in general inhomogeneous isotropic media are derived.

Since the refractive index of inhomogeneous media is varying from point to point, the ray path in these media has a more or less complicated geometry. As an example, the curved path of a skew ray inside a thick gradient medium with radial refractive index distribution is symbolically shown in Fig. 2.1.1. Consequently, as compared with the case of homogeneous media, in GRIN media an additional computational problem must be solved: the accurate determination of the ray path inside the lens.

A simple derivation of the light ray equation relies on the fact that the direction of light rays is everywhere normal to the corresponding wave front element /18/.
In an isotropic medium with a variable refractive index \( n(R) \), the phase difference between two neighbouring points along the ray path, having the position vectors \( R \) and \( R + dR \), is

\[
d\Phi = \frac{2\pi}{\lambda} n(R) ds ,
\]

(2.1.1)

where \( \lambda \) is the wavelength in vacuum, and \( ds \) denotes the arc length between the two points. If the ray direction at the point \( R \) is given by the direction vector \( a \), then we have

\[
dR = ads
\]

(2.1.2)

with \( R = (x, y, z) \) and \( a^2 = 1 \), such that

\[
ds^2 = dx^2 + dy^2 + dz^2
\]

(2.1.3)

On the other hand

\[
d\Phi = \frac{\partial \Phi}{\partial x} dx + \frac{\partial \Phi}{\partial y} dy + \frac{\partial \Phi}{\partial z} dz = \nabla \Phi \cdot dR
\]

and it follows from (2.1.1) and (2.1.2) that

\[
\nabla \Phi a = \frac{2\pi}{\lambda} n(R).
\]

(2.1.4)

Since \( \nabla \Phi \) and \( a \) have the same direction (that of the normal to the wave front element) we have \( \nabla \Phi a = |\nabla \Phi| a \) and \( |\nabla \Phi| a = \nabla \Phi \). By multiplying both sides of Eq. (2.1.4) with \( a \) we obtain

\[
\nabla \Phi = \frac{2\pi}{\lambda} n(R) a.
\]

(2.1.5)
and after derivation with respect to $ds$
\[
\frac{2\pi}{\lambda} \frac{d}{ds} (na) = \frac{d}{ds} \nabla \Phi = \nabla \Phi \frac{d\Phi}{ds}.
\] (2.1.6)

According to Eq.(2.1.1) the right hand side of Eq. (2.1.6) can be rewritten as
\[
\frac{2\pi}{\lambda} \nabla n.
\]

Consequently we obtain
\[
\frac{d}{ds} (na) = \nabla n
\] (2.1.7)
or by using (2.1.2):
\[
\frac{d}{ds} \left( n \frac{dR}{ds} \right) = \nabla n
\] (2.1.8)

Eq. (2.1.8) is a commonly used vector form of the differential equations of the light rays /8/.

It will be seen in the next sections that in GRIN media it is useful to describe the ray direction by the vector
\[
a = na
\] (2.1.9)

which can be called "optical ray vector" because its components are the optical direction cosines with respect to the three coordinate axes, $a = (\xi, \eta, \zeta)$, It follows from (2.1.9) that $|a| = n$ i.e.
\[
\xi^2 + \eta^2 + \zeta^2 = n^2
\] (2.1.10)

From (2.1.7) we obtain
\[
da = \nabla nds
\] (2.1.11a)

The change of the optical ray vector between two arbitrary points along the ray path is therefore given by
\[
\Delta a = \int_{s_1}^{s_2} \nabla nds
\] (2.1.11b)

Eq. (2.1.8) can be put into a more compact form which is useful both for analytical and for numerical computations. Multiplication of both sides of Eq. (2.1.8) by $n$ yields
\[
n \frac{d}{ds} \left( n \frac{dR}{ds} \right) = n \nabla n = \frac{1}{2} \nabla (n^2)
\]

By introducing as independent variable
\[
d\tau = \frac{ds}{n}
\] (2.1.12)

we obtain
For tracing rays in optical systems with GRIN lenses, most computer programs known at present use numerical iterative solutions of the light ray equation in the form of Eq.(2.1.13). For that purpose the lens thickness is divided into a large number of small intervals. The ray path is then determined step by step using standard numerical methods e.g. Runge-Kutta. (see Sharma et al. /50/) For axial and radial GRIN media analytical methods for finding the ray path will be given in Sections 2.2.1 and 2.4.1.

2.1.2 Invariants

If the refractive index distribution has special symmetries then certain invariants exist which are conserved during ray propagation in the medium /25/. It will be seen in the following sections that these invariants are very useful for the analytical determination of the ray paths.

For rotationally symmetric refractive index distributions, let us choose the z-axis along the symmetry axis of the gradient. By writing the components of Eq. (2.1.11b) explicitly we obtain for the change of the optical direction cosines

\[ \Delta \xi = \int_{1}^{2} \frac{\partial n}{\partial x} \, ds \quad , \]  
\[ \Delta \eta = \int_{1}^{2} \frac{\partial n}{\partial y} \, ds \quad , \]  
\[ \Delta \zeta = \int_{1}^{2} \frac{\partial n}{\partial z} \, ds \quad . \]  

Consider now two special cases, both of considerable practical interest:

i) In the case of radial refractive index distributions, the refractive index varies only perpendicularly to the symmetry axis, \( n(R) = n(r^2) = r^2 + y^2 \). Consequently \( \partial n / \partial z = 0 \) and it follows from (2.1.14c) that the optical direction cosine with respect to the symmetry axis is invariant along any ray

\[ \zeta = \zeta_0 = \text{const} \quad . \]  

ii) In the case of axial refractive index distributions, the refractive index varies along the symmetry axis, \( n(R) = n(z) \). Therefore \( \partial n / \partial x = \partial n / \partial y = 0 \) and it follows from (2.1.14a,b) that

\[ \xi = \xi_0 = \text{const} \quad , \]  
\[ \eta = \eta_0 = \text{const} \quad . \]
In Eqs. (2.1.15-16) \( \xi_0, \eta_0 \) and \( \zeta_0 \) are the initial values of the optical direction cosines inside the gradient medium and are determined by the refraction at the first surface of the lens as shown in the next section (Eq.(2.1.25)).

An additional invariant occurs as a consequence of rotational symmetry in all media with a refractive index distribution of the form \( n(R) = n(r, z) \). It follows from the vector multiplication of Eq. (2.1.11a) by \( R \) that

\[
da \times R = \nabla n \times R ds . \tag{2.1.17}
\]

The left hand side of Eq. (2.1.17) is \( da \times R = d(a \times R) - a \times dR = d(a \times R) \), because \( a = n dR/ds \) is always parallel to \( dR \). The integration of Eq. (2.1.17) yields

\[
\Delta(a \times R) = \int_1^2 \nabla n \times R ds . \tag{2.1.18}
\]

Let us now examine the z- component of the vectorial equation (2.1.18)

\[
\Delta(\xi_0 - \eta_0) = \int_1^2 \left( \frac{\partial n}{\partial x} y - \frac{\partial n}{\partial y} x \right) ds . \tag{2.1.19}
\]

Since we have

\[
\frac{\partial n}{\partial x} = \frac{\partial n}{\partial r} \frac{\partial r}{\partial x} = \frac{\partial n}{\partial r} \frac{x}{\sqrt{x^2 + y^2}}
\]

\[
\frac{\partial n}{\partial y} = \frac{\partial n}{\partial r} \frac{\partial r}{\partial y} = \frac{\partial n}{\partial r} \frac{y}{\sqrt{x^2 + y^2}}
\]

the right hand side of Eq. (2.1.19) vanishes and we obtain the invariant

\[
\xi_0 - \eta_0 = \eta_0 x_0 - \eta_0 x_0 = \text{const} . \tag{2.1.20}
\]

where the initial values of the ray coordinates have been also denoted by the index 0.
2.1.3 Refraction at a surface

Snell's refraction law can be derived directly from Eq.(2.1.11b). As known from vector analysis, the gradient of a function is a vector which is perpendicular to the surface of constant values of the function, i.e. the gradient has the direction of the "most rapid variation" of the function. At refraction at a surface separating two different media, the refractive index is discontinuous and \( \nabla n \) becomes a delta function. The direction of the gradient, however, is that of the unit vector \( \mathbf{N} \) along the normal to the surface at the point of incidence.

Let us now examine Eq.(2.1.11b) in the case where the two points giving the integration limits are situated in the immediate neighbourhood of the surface, before and after refraction. After a vector multiplication of Eq.(2.1.11b) by \( \mathbf{N} \) the right hand vanishes and we obtain Snell's law

\[
\Delta a \times \mathbf{N} = 0 \quad (2.1.21a)
\]
i.e. / 59/

\[
n' a' \times \mathbf{N} = na \times \mathbf{N} \quad (2.1.21b)
\]

In Eq. (2.1.21b) the quantities after refraction are denoted by a prime and those before refraction are left unprimed.

By denoting the angles between \( \mathbf{N} \) and the direction vectors before and after refraction (i.e. the angles of incidence and refraction) by \( \theta \) and \( \theta' \), it follows for the absolute values of both sides of (2.1.21b) that

\[
n' \sin \theta' = n \sin \theta . \quad (2.1.22)
\]

For ray-tracing it is convenient to rewrite the refraction law in a form expressing the optical ray vector after refraction explicitly. A vector multiplication at left of Eq. (2.1.21a) by \( \mathbf{N} \) yields

\[
\Delta a \times \mathbf{N} = 0 . \quad (2.1.23)
\]

It is known from vector algebra that for three arbitrary vectors we have

\[
\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) .
\]

The left hand side of Eq.(2.1.23) can be rewritten as

\[
\mathbf{N} \times (\Delta a \times \mathbf{N}) = \Delta a (\mathbf{N}^2) - \mathbf{N} (\mathbf{N} \Delta a) = \Delta a - \mathbf{N} (\mathbf{N} \Delta a) ,
\]

and therefore

\[
a' = a + \mathbf{N} (\mathbf{N} a' - \mathbf{N} a) . \quad (2.1.24)
\]

Because
\[
\mathbf{Na'} = \mathbf{N}n'\mathbf{a'} = n' \cos i' = \sqrt{n'^2 - n'^2 \sin^2 i'} = \sqrt{n'^2 - n^2 \sin^2 i} = \sqrt{n'^2 - n^2 + n^2(1 - \sin^2 i)} = \sqrt{n'^2 - n^2 + (Na)^2}.
\]

Eq. (2.1.24) finally reads
\[
a' = a + N\left(\sqrt{n'(R)^2 - n(R)^2} + (Na)^2 - Na\right), \quad (2.1.25)
\]

For GRIN lenses, note that in Eq. (2.1.25) the refractive index at the incidence point having the position vector \( \mathbf{R} \) is changing across the surface.

![Fig. 2.1.2 The normal vector to a spherical surface](image)

For determining the form of \( \mathbf{N} \) in the case of refraction at a spherical surface, note that in this case \( \mathbf{N} \) is oriented towards the centre of the surface. It can be easily seen from Fig. (2.1.2) that if the system of coordinates has its origin at the vertex of the surface, then the components of \( \mathbf{N} \) are
\[
(N_x, N_y, N_z) = (-\rho x, -\rho y, 1 - \rho z) \quad (2.1.26)
\]
where the surface curvature \( \rho \) ensures that \( \mathbf{N}^2 = 1 \).

The equations (2.1.25) and (2.1.13) enable the tracing of finite rays in optical systems containing homogeneous and GRIN lenses. However, the determination of the ray-surface intersection point for GRIN lenses must still be discussed. This computation is more complex than for homogeneous lenses, because the ray path inside the GRIN lens is curved. A convenient numerical method for this purpose has been developed by Sharma and Ghatak /52/. Their method is based on the approximation of the ray path near the surface by a set of nonlinear parametric equations. By inserting these equations into the equation of the surface, the problem of finding the intersection point is reduced to solving numerically an algebraic (cubic) equation. For radial GRIN lenses, an analytical method to determine the ray-surface intersection point will be given in Section 2.4.3.
Consider now the paraxial approximation of Eq. (2.1.25). In this approximation, $x, y$ and the corresponding optical direction cosines can be regarded as small quantities. Therefore, paraxial ray-tracing formulae can be obtained by considering in the corresponding finite ray-tracing formulae only the lowest order terms in these variables. (For the derivation of aberration coefficients a more detailed analysis of the paraxial approximation is necessary. This will be done in Sec. 2.5.) In what follows we denote the paraxial approximation for $x, y, \xi, \eta$ (and also for any functions of these variables) by a tilde $\sim$ over the corresponding quantity.

The angle of the paraxial ray with the optical axis is denoted by $u$ and its height by $h$. We follow the notation and sign conventions currently used at the Technical University of Berlin /22/, /23/. The sign convention adopted for paraxial ray angles with the optical axis is that they are of opposite sign to the corresponding optical direction cosines. If the ray path is situated in the $xz$ plane, then

$$\tilde{x} = h = \tilde{x}' = h' \tilde{\xi} = -nu \tilde{\xi}' = -n' u'$$

Considering only lowest order contributions, $n - n'$ can be replaced by $n' - n$. In this case $n$ and $n'$ are the values of the refractive index on the optical axis, prior and after refraction. Inserting Eqs. (2.1.26) and (2.1.27) into the $x$- component of Eq. (2.1.15) we obtain

$$n' u' = nu + h \rho \Delta n \quad .$$

In Gauss-matrix notation, the refraction of the paraxial ray is described by

$$\begin{pmatrix} n' u' \\ h' \end{pmatrix} = B \begin{pmatrix} nu \\ h \end{pmatrix} ;$$

$$B = \begin{pmatrix} 1 & \rho \Delta n \\ 0 & 1 \end{pmatrix}$$

Thus, the paraxial refraction formulae are precisely the same as in the case of homogeneous media /22/. In the paraxial approximation, the refraction invariant (2.1.22) reads

$$n' i' = n' h \rho - n' u' = ni = nh \rho - nu$$
2.2 Ray-tracing in axial gradients

2.2.1 Finite rays

In axial gradients, i.e. when \( n(R) = n(z) \), the invariance of the optical direction cosines \( \xi \) and \( \eta \) according to Eq.(2.1.16) enables the determination of the ray path directly, without solving the equation of light rays.

The components of \( a = n \frac{dR}{ds} \) read

\[
\begin{align*}
\xi_0 &= n \frac{dx}{ds} \\
\eta_0 &= n \frac{dy}{ds} \\
\zeta &= n \frac{dz}{ds}
\end{align*}
\]  

(2.2.1)

Dividing the first two equations (2.2.1) by the third one yields

\[
\frac{dx}{dz} = \frac{\xi_0}{\zeta}, \quad \frac{dy}{dz} = \frac{\eta_0}{\zeta}
\]  

(2.2.2)

According to Eq. (2.1.10) we have

\[
\frac{1}{\zeta} = \pm \frac{1}{\sqrt{n^2(z) - \xi_0^2 - \eta_0^2}}
\]  

(2.2.3)

Usually we assume that the ray propagates from left to right and therefore keep the plus sign in Eq. (2.2.3) (For large values of \( \xi_0 \) and \( \eta_0 \) an exception will be shown in the next section.) Since Eq. (2.2.3) is a function of \( z \) only, Eqs. (2.2.2) can be integrated directly. The solutions read

\[
\begin{align*}
x(z) &= x_0 + \xi_0 A(\xi_0, \eta_0, z) \\
y(z) &= y_0 + \eta_0 A(\xi_0, \eta_0, z)
\end{align*}
\]  

(2.2.4)

where A depends on the form of the refractive index distribution and on the initial direction of the ray:

\[
A(\xi_0, \eta_0, z) = \int_{z_0}^{z} \frac{1}{\sqrt{n^2(z) - \xi_0^2 - \eta_0^2}} dz
\]  

(2.2.5)

For general axial GRIN media, the quadrature (2.2.5) must be carried out numerically. An example where an analytical solution can be obtained will be discussed in the next section.

In order to separate the dependence of Eq. (2.2.5) on the initial ray direction from that on the refractive index distribution we expand Eq. (2.2.3) into a power series
\[ \frac{1}{\sqrt{n^2(z) - \xi_0^2 - \eta_0^2}} = \frac{\omega}{\sqrt{1 - C_{\psi\theta} \omega^2}} = \omega + \frac{1}{2} C_{\psi\theta} \omega^3 + \frac{3}{8} C_{\psi\theta}^2 \omega^5 + \frac{5}{16} C_{\psi\theta}^3 \omega^7 + \ldots, \]  

(2.2.6)

where the following abbreviations have been introduced:

\[ \omega = \frac{1}{n(z)}, C_{\psi\theta} = \xi_0^2 + \eta_0^2. \]  

(2.2.7)

Integrating Eq. (2.2.6) yields

\[ A(\xi_0, \eta_0, z) = Z_1(z) + \frac{1}{2} (\xi_0^2 + \eta_0^2) Z_3(z) + \frac{3}{8} (\xi_0^2 + \eta_0^2)^2 Z_3(z) + \frac{5}{16} (\xi_0^2 + \eta_0^2)^3 Z_3(z) + \ldots \]  

(2.2.8)

where for odd integers p the parameters \( Z_p \)

\[ Z_p(z) = \int_0^z n(z)^{p-1} dz, p = 1,3,5,\ldots,2k+1,\ldots. \]  

(2.2.9)

can be determined by quadrature from the axial index distribution. Thus, for not too large values of \( \xi_0 \) and \( \eta_0 \) the ray path of arbitrary finite rays inside the axial gradient is determined by Eqs. (2.2.4) and (2.2.8)-(2.2.9).

### 2.2.2 Special cases

For certain axial refractive index distributions the integrals appearing in the ray tracing formulae of the previous section can be evaluated in closed form. For example, if

\[ n(z) = n_0 \sqrt{1 + \alpha z} \]  

(2.2.10)

then exact formulae for the ray path can be obtained /25/. Special cases often provide additional insight into more general behaviour of solutions. We use this example to illustrate the possibility of total reflection inside an axial gradient.

From Eq. (2.1.10) we have

\[ \zeta^2(z) = n_0^2 (1 + \alpha z) - \xi_0^2 - \eta_0^2 \]  

(2.2.11)

and, by differentiating

\[ 2\zeta d\zeta = n_0^2 \alpha dz. \]

Consequently, introducing \( \zeta \) as a new variable of integration into Eq. (2.2.5) yields

\[ A(\xi_0, \eta_0, z) = \frac{1}{\zeta} d\zeta = \int_0^{\zeta(z)} \frac{1}{\zeta} \frac{2\zeta d\zeta}{n_0^2 \alpha} \]

which finally gives

\[ A(\xi_0, \eta_0, z) = \frac{2}{n_0^2 \alpha} (\zeta(z) - \xi_0) \]  

(2.2.12)

Let us now consider the case when \( \alpha < 0 \) (i.e. refractive index decreasing with depth) and \( \xi_0 \) and \( \eta_0 \) are large. Assume \( \zeta_0 > 0 \). However, if the medium is thick enough, because of the decreasing refractive index the ray reaches a certain value \( z = z_r \), where

\[ \zeta^2(z_r) = n_0^2 (1 + \alpha z_r) - \xi_0^2 - \eta_0^2 = 0 \]  

(2.2.13)
i.e. the optical ray vector is perpendicular to the z-axis. In order to analyse what happens with the ray after reaching this point, we use the light ray equation Eq.(2.1.13). Inserting Eq. (2.2.10), the z-component of Eq. (2.1.13) reads

$$\frac{d^2 z}{d\tau^2} = -\frac{1}{2} n_0^2 |\alpha| = \text{const}$$  \hspace{1cm} (2.2.14)

Because \(\frac{d^2 z}{d\tau^2} < 0, \zeta = n \frac{dz}{ds} = \frac{dz}{d\tau}\) after decreasing from positive values to zero, becomes negative along the ray path. As shown previously, for given values of \(\xi_0\) and \(\eta_0\) Eq. (2.2.3) is satisfied by two values of \(\zeta\) having opposite sign. Therefore, inserting Eqs. (2.2.11) and (2.2.12) into Eqs. (2.2.4), the solution has two branches which together form a parabola

$$x(z) = x_0 + \frac{2\xi_0}{n_0\alpha} \left( \pm \sqrt{n_0^2 (1 + \alpha \zeta) - \xi_0^2 - \eta_0^2} \right)$$

$$y(z) = y_0 + \frac{2\eta_0}{n_0\alpha} \left( \pm \sqrt{n_0^2 (1 + \alpha \zeta) - \xi_0^2 - \eta_0^2} \right).$$  \hspace{1cm} (2.2.15)

The ray first propagates from left to right along the positive branch up to \(z = z_r\) and then changes to the negative branch, propagating from right to left. The change of sign of the \(z\)-component of the optical direction cosine while the other two components remain unchanged means that a total reflection of the ray occurs inside the gradient.

Numerical integration at ray-tracing can be avoided also for other axial distributions, e.g. for

$$n(z) = n_0(1 + \alpha z)^q$$  \hspace{1cm} (2.2.16)

where \(q\) is an arbitrary nonzero number. In this case \(t = 1 + \alpha z\) can be used as a variable of integration in Eqs. (2.2.9)

$$Z_p(z) = \frac{1}{n_0^p \alpha} \int^{\alpha z} dt \frac{dt}{(1 + \alpha z)^{pq}}.$$  \hspace{1cm} (2.2.17)

We obtain generally (excepting the case \(pq = 1\))

$$Z_p(z) = -\frac{1}{n_0^p \alpha (pq - 1)} \left( (1 + \alpha z)^{-pq} - 1 \right).$$  \hspace{1cm} (2.2.18a)

and for \(pq = 1\) we have

$$Z_p(z) = \frac{1}{n_0 \alpha} \ln(1 + \alpha z).$$  \hspace{1cm} (2.2.18b)

Using Eqs. (2.2.18a, b), (2.2.8) and (2.2.4) the ray path inside these axial gradients can be thus calculated by means of power series expansions in the variable \(\xi_0^2 + \eta_0^2\).

**2.2.3 Paraxial approximation**
In the paraxial approximation, by keeping only the first term in the power series expansion (2.2.8), Eqs. (2.2.4) become linear and, as for homogeneous media, the ray path projections onto the planes xz and yz can be described independently from each other

\[ \bar{x}(z) = \bar{x}_0 + \xi_0 Z_1(z) \]
\[ \bar{y}(z) = \bar{y}_0 + \eta_0 Z_1(z) \]  

(2.2.19)

Since the two expressions in (2.2.19) are similar, only the ray projection onto one of these planes (e.g. xz) will be considered in detail in what follows.

Consider now the transfer of a paraxial ray through an axial gradient medium of thickness d. The angle of the ray with the optical axis is denoted by u and its height by h. Quantities before transfer are denoted by unprimed symbols and after transfer by primed ones. Consequently, we have

\[ \bar{x}_0 = \bar{x}(0) = h, \bar{x}(z) = \bar{x}(d) = h', \xi_0 = -nu, \xi(z) = -n'u' \]  

(2.2.20)

The function \( Z_1(d) \) can be written using the average value of the reciprocal of the refractive index as:

\[ Z_1(d) = d\left\langle n^{-1}\right\rangle \]
\[ \int \frac{dz}{n(z)} \]  

(2.2.21)

From Eqs. (2.1.16) and (2.2.19) we have

\[ h' = h - d\left\langle nu \right\rangle \]
\[ n'u' = nu \]  

(2.2.22)

or, in Gauss-matrix notation

\[ \begin{pmatrix} n'u' \\ h' \end{pmatrix} = G_z \begin{pmatrix} nu \\ h \end{pmatrix} \]
\[ G_z = \begin{pmatrix} 1 & 0 \\ -\left\langle n^{-1}\right\rangle d & 1 \end{pmatrix} \]  

(2.2.23)

It is remarkable that the Gauss-matrix of the transfer through an axial GRIN medium has the same form as that for a homogeneous medium, provided that the reciprocal of the refractive index is replaced by the average value (2.2.21).

The second of the Eqs. (2.2.22) indicates that at transfer through an axial GRIN medium one of the paraxial transfer invariants is the product nu. Consider now a second paraxial ray, having an angle \( w \) and height \( m \) prior to transfer and \( w' \) and \( m' \) after transfer. In homogeneous media, the quantity...
\[ m'n'u' - h'n'w' = mnu - hnw = H \]  

(2.2.24)

is the well-known paraxial invariant of the system. It can be easily shown that Eq. (2.2.24) holds also at transfer through axial gradients. Indeed, Eq (2.2.23) holds also for the second paraxial ray. By grouping the column vectors describing the two rays prior and after transfer into 2x2 matrices, we have

\[
\begin{pmatrix}
  n'u' & n'w' \\
  h' & m'
\end{pmatrix} = G \begin{pmatrix}
  nu & nw \\
  h & m
\end{pmatrix}.
\]  

(2.2.25)

The left hand side of Eq. (2.2.24) is the determinant of the left hand side of Eq. (2.2.25). It is known from linear algebra that the determinant of the product of two quadratic matrices is equal to the product of the determinants. According to Eq. (2.2.23) \( \det G_z = 1 \) and thus Eq. (2.2.24) is proven.
2.3 Ray-tracing in radial gradients

2.3.1 Differential equations of light rays in radial gradients

In order to find the differential equations of light rays in the case of radial GRIN media, i.e. when \( n(R) = n(r^2) = x^2 + y^2 \) we start by writing the components of the general vectorial ray equation (2.1.13) as

\[
\frac{d^2x}{d\tau^2} = \frac{1}{2} \frac{\partial}{\partial x} n^2(r^2) \\
\frac{d^2y}{d\tau^2} = \frac{1}{2} \frac{\partial}{\partial y} n^2(r^2) \\
\frac{d^2z}{d\tau^2} = \frac{1}{2} \frac{\partial}{\partial z} n^2(r^2)
\]  

(2.3.1)

Since \( \frac{\partial n^2}{\partial z} = 0 \), it follows from the third of the Eqs.(2.3.1), that \( \frac{dz}{d\tau} = n \frac{dz}{ds} = \zeta_0 = \text{const} \), as already shown with Eq. (2.1.15). Hence

\[
d\tau = \frac{dz}{\zeta_0} \tag{2.3.2}
\]

It is therefore possible to use \( z \) as the independent variable in the first two of Eqs.(2.3.1):

\[
\frac{d^2x}{dz^2} = \frac{1}{2\zeta_0^2} \frac{\partial}{\partial x} n^2(r^2) \\
\frac{d^2y}{dz^2} = \frac{1}{2\zeta_0^2} \frac{\partial}{\partial y} n^2(r^2)
\]  

(2.3.3)

For the optical direction cosines \( \xi = n \frac{dx}{ds} = \frac{dx}{d\tau} \) and \( \eta = \frac{dy}{d\tau} \) it follows from Eq. (2.3.2) that

\[
\xi = \zeta_0 \frac{dx}{dz} \\
\eta = \zeta_0 \frac{dy}{dz} \tag{2.3.4}
\]

Because the refractive index distribution is rotationally symmetric, it is possible to rewrite Eqs.(2.3.3) in cylindrical coordinates /25/. However, previous studies indicated that both for finding approximate analytic solutions for the ray path /28/ and for deriving aberration coefficients /10/, Cartesian coordinates seem to be more suitable. In both cases, the author of this study shares this opinion.

Since
\[
\frac{\partial n^2}{\partial x} = \frac{\partial n^2}{\partial r} \frac{\partial r}{\partial x} = \frac{\partial n^2}{\partial r} \frac{x}{r} \\
\frac{\partial n^2}{\partial y} = \frac{\partial n^2}{\partial r} \frac{\partial r}{\partial y} = \frac{\partial n^2}{\partial r} \frac{y}{r}
\]

is follows from Eq. (2.3.3) that

\[
\frac{d^2 x}{dz^2} = \frac{x}{2r}\frac{\partial}{\partial r} n^2(r^2) \\
\frac{d^2 y}{dz^2} = \frac{y}{2r}\frac{\partial}{\partial r} n^2(r^2)
\]  

Eq. (2.3.5) enables the determination of the ray path of skew rays in GRIN media with arbitrary radial refractive index distribution.

Assume that the refractive index distribution is given by a power series expansion with respect to the distance to the optical axis \( r \). Each radial refractive index distribution with a nonzero quadratic term can be written as

\[
n^2(r^2) = n_0^2 \left(1 - kr^2 + N_2 k^2 r^4 + N_4 k^4 r^6 + N_6 k^6 r^8 + \ldots\right)
\]  

(2.3.6)

It will be seen later that this particular form for the second and the higher order terms is convenient for writing the formulae for both ray-tracing and the Seidel coefficients as simple as possible. The special case \( k=0 \) (the so-called shallow gradients) will be discussed separately.

As will be seen in the next section, the refractive index on the optical axis \( n_0 \) and the coefficient of the quadratic term \( k \) having the dimensionality of length\(^{-2} \) determine the paraxial properties of the medium. The dimensionless coefficients \( N_2, N_4, N_6, \ldots \) occurring in the higher order terms are additional parameters that influence the aberrations of the system. For example \( N_4 \) appears in the expressions of the Seidel aberration coefficients (Section 3.5). In order to achieve a prescribed accuracy at ray tracing, the number of terms which must be considered in Eq. (2.3.6) increases with the distance of the ray path to the optical axis i.e. with the size of aperture and field.

We now insert Eq. (2.3.6) into Eqs. (2.3.5) Because

\[
\frac{1}{2r} \frac{\partial}{\partial r} n^2(r^2) = n_0^2 \left(-k + 2N_2 k^2 r^2 + 3N_4 k^4 r^4 + 4N_6 k^6 r^6 + \ldots\right)
\]

we obtain

\[
\frac{d^2 x}{dz^2} + \frac{n_0^2 k x}{\xi_0^2} = \frac{n_0^2 k x}{\xi_0^2} \left(2N_2 k r^2 + 3N_4 k^2 r^4 + 4N_6 k^4 r^6 + \ldots\right)
\]

\[
\frac{d^2 y}{dz^2} + \frac{n_0^2 k y}{\xi_0^2} = \frac{n_0^2 k y}{\xi_0^2} \left(2N_2 k r^2 + 3N_4 k^2 r^4 + 4N_6 k^4 r^6 + \ldots\right)
\]  

(2.3.7)
Consider first the case $k > 0$. (Positive gradient: In this case, for not too thick lenses the gradient medium has a positive contribution to the total power of the lens, as shown later.) It is convenient to introduce a new independent variable

$$t = \frac{n_0 g}{\zeta_0} z,$$  \hspace{1cm} (2.3.8)

where the abbreviation

$$g = k^{1/2}$$  \hspace{1cm} (2.3.9)

was used. Since we have

$$\frac{d^2 x}{dz^2} = \frac{n_0^2 g^2}{\zeta_0^2} \frac{d^2 x}{dt^2}.$$

Eqs. (2.3.7) can be rewritten as

$$\begin{align*}
\ddot{x} + x &= x \left[ 2N_4 (gr)^2 + 3N_6 (gr)^4 + 4N_8 (gr)^6 + \ldots \right] \\
\ddot{y} + y &= y \left[ 2N_4 (gr)^2 + 3N_6 (gr)^4 + 4N_8 (gr)^6 + \ldots \right].
\end{align*}$$  \hspace{1cm} (2.3.10)

where a dot indicates $d/dt$. It follows from Eqs. (2.3.4) and (2.3.8) that the optical direction cosines can be rewritten as

$$\begin{align*}
\xi &= n_0 g \dot{x} \\
\eta &= n_0 g \dot{y}.
\end{align*}$$  \hspace{1cm} (2.3.11)

In Sec. 2.4 the equations (2.3.10) will be the starting point for the analytical determination of the path of finite skew rays in radial GRIN media. In the next section the paraxial ray-tracing formulae will be derived from them.

According to Eq.(2.1.15) the optical direction cosine with respect to the optical axis $\zeta_0$ is invariant along any ray. This is also true for any function of $\zeta_0$, e.g. for

$$\mu = \frac{1}{2} \left( 1 - \frac{\zeta_0^2}{n_0^2} \right).$$  \hspace{1cm} (2.3.12)

/6/ (We have introduced the coefficient 1/2 for the convenience of subsequent calculations.) Since, according to Eq.(2.1.10), we have

$$\zeta_0^2 = n_0^2 \left[ 1 - (gr)^2 + N_4 (gr)^4 + N_6 (gr)^6 + N_8 (gr)^8 + \ldots \right] - \xi^2 - \eta^2$$  \hspace{1cm} (2.3.13)

Eq. (2.3.12) can be rewritten as

$$\mu = \frac{1}{2} \left[ \frac{1}{n_0^2} \left( \zeta^2 + \eta^2 \right) + g^2 r^2 - \left( N_4 (gr)^4 + N_6 (gr)^6 + N_8 (gr)^8 + \ldots \right) \right].$$  \hspace{1cm} (2.3.14)

For the interpretation of the invariant (2.3.14) we consider its paraxial approximation (i.e. only the lowest order terms).:

$$\bar{\mu} = \frac{1}{2} \left[ \frac{1}{n_0^2} \left( \bar{\zeta}^2 + \bar{\eta}^2 \right) + g^2 \left( \bar{x}^2 + \bar{y}^2 \right) \right].$$  \hspace{1cm} (2.3.15)

(Recall that the tilde~ denotes the paraxial approximation of a given quantity.) Since for positive gradients $\mu$ increases with increasing ray height and inclination, this invariant
indicates how close the ray path to the optical axis is. This feature of \( \mu \) will be used in Sec 2.4 for finding approximate analytic solutions of Eqs. (2.3.10).

For negative gradients, i.e. when \( k < 0 \), we can set instead of Eq. (2.3.9)

\[
\hat{g} = (-k)^{1/2}
\]

(2.3.16)

and an analogue formalism with that for \( k > 0 \) can be developed. However, it will be seen in Sec. 2.4 that the approximate ray tracing formulae contain only analytical functions, and we find it more convenient to use the same ray-tracing formulae both for positive and negative gradients. According to Eq. (2.3.9) \( g \) is imaginary for \( k < 0 \) and therefore some intermediate results have complex values, but the final results are always real. Moreover, it will turn out in Sec. 3.5 that the expressions for the Seidel coefficients resulting from the ray-tracing formulae contain only \( g^2 = k \) and therefore no complex quantities arise for \( k < 0 \).

### 2.3.2 Paraxial approximation

Since in the paraxial approximation \( x, y \) and the corresponding optical direction cosines are small, we keep in Eqs. (2.3.10) and (2.3.13) only the terms of lowest order. In this approximation, Eqs. (2.3.10) become linear

\[
\ddot{x} + \dot{x} = 0,
\]

\[
\ddot{y} + \dot{y} = 0,
\]

(2.3.17)

and, since \( \bar{n}^2(r^2) = n_0^2 \), it follows from Eq. (2.3.13) that

\[
\bar{\xi}_0 = n_0.
\]

(2.3.18)

As in Sec. 2.2.3, only the projection of the ray onto the plane \( xz \) will be considered further. The general solution of the first of the Eqs. (2.3.17) is

\[
\bar{x}(t) = C_1 \cos t + C_2 \sin t,
\]

(2.3.19)

and its first derivative is

\[
\dot{\bar{x}}(t) = C_1 \cos t - C_2 \sin t.
\]

(2.3.20)

It follows from Eqs. (2.3.11) and (2.3.20) that

\[
\bar{\xi}(t) = n_0 g C_2 \cos t - n_0 g C_1 \sin t.
\]

(2.3.21)

The two integration constants \( C_1 \) and \( C_2 \) are determined from the initial conditions:

\[
\bar{x}_0 = \bar{x}(0) = C_1
\]

\[
\bar{\xi}_0 = \bar{\xi}(0) = n_0 g C_2.
\]

Finally we obtain
Consider now the transfer of a paraxial ray through a radial GRIN medium of thickness $d$. As in Sec. 2.2.3, the ray data will be the angles $u$ and $u'$ before and after transfer and the corresponding heights $h$ and $h'$. We have

$$
\tilde{x}_0 = \tilde{x}(0) = h, \tilde{x}(z) = \tilde{x}(d) = h', \tilde{\xi}_0 = -n_0u, \tilde{\xi}(z) = -n_0u' \tag{2.3.23}
$$

Hence Eq.(2.3.22) reads

$$
h' = h \cos gd \frac{n_0u}{n_0g} \sin gd
$$

$$n_0u' = n_0u \cos gd + n_0gh \sin gd
$$

i.e.

$$u' = u \cos gd + h \sin gd
$$

$$h' = \frac{u}{g} \sin gd + h \cos gd \tag{2.3.24}
$$

or, in the Gauss-matrix notation

$$
\begin{pmatrix} n_0u' \\ h' \end{pmatrix} = G_r \begin{pmatrix} n_0u \\ h \end{pmatrix},
$$

$$
G_r = \begin{pmatrix} \cos gd & n_0gh \sin gd \\ \frac{1}{n_0g} & \sin gd & \cos gd \end{pmatrix}. \tag{2.3.25}
$$

For $k<0$ $g$ is imaginary, as noted before. Since Eqs.(2.3.24-25) are often needed, a separate treatment of this case is convenient in order to avoid programming with complex numbers.

Trigonometric functions with imaginary arguments can be replaced by the corresponding hyperbolic functions. From the definitions

$$
\sin x = \frac{1}{2i} (e^{ix} - e^{-ix}) \quad \sinh x = \frac{1}{2} (e^x - e^{-x})
$$

$$
\cos x = \frac{1}{2} (e^{ix} + e^{-ix}) \quad \cosh x = \frac{1}{2} (e^x + e^{-x})
$$

it follows that, for arbitrary $x$

$$
\sin ix = i \sinh x
$$

$$
\cos ix = \cosh x
$$
From (2.3.9) and (2.3.16) we have \( g = -i\hat{g} \). Thus the transfer equations for negative gradients read in real form

\[
\begin{align*}
    u' &= u \cosh \hat{g} d - h \hat{g} \sinh \hat{g} d \\
    h' &= -\frac{u}{\hat{g}} \sinh \hat{g} d + h \cosh \hat{g} d
\end{align*}
\]

(2.3.26)

The difference between the transfer formulae in the two cases \( k > 0 \) and \( k < 0 \) disappears if the trigonometric and hyperbolic functions are expanded into power series. Setting

\[
E_1(kd^2) = \cos gd, \quad E_2(kd^2) = \frac{\sin gd}{gd}
\]

(2.3.27)

the Gauss-matrix (2.3.25) reads

\[
G_r = \begin{pmatrix}
E_1(kd^2) & n_0kdE_2(kd^2) \\
\frac{d}{n_0}E_2(kd^2) & E_1(kd^2)
\end{pmatrix}
\]

(2.3.28)

From the series expansions of Eqs. (2.3.27) it can be observed that \( E_1 \) and \( E_2 \) depend, as expected, only on \( g^2 = k \):

\[
\begin{align*}
E_1(kd^2) &= 1 - \frac{1}{2!}kd^2 + \frac{1}{4!}(kd^2)^2 - ... \\
E_2(kd^2) &= 1 - \frac{1}{3!}kd^2 + \frac{1}{5!}(kd^2)^2 - ...
\end{align*}
\]

(2.3.29)

In Sec. 4.2.1 the equations (2.3.28-29) will be the starting point for the derivation of a consistent thin-lens approximation for radial GRIN lenses.

For shallow gradients \( k \) tends to zero. In this case we have \( E_1 = E_2 = 1 \) and the Gauss-matrix (2.3.28) has the same form as for a transfer through a homogeneous medium /22/:

\[
U = \begin{pmatrix}
1 & 0 \\
\frac{d}{n_0} & 1
\end{pmatrix}
\]

(2.3.30)

We shall now consider the paraxial invariants at transfer through radial GRIN media. \( H \) is an invariant also for radial gradients. This follows, as shown in Sec. 2.2.3, from the fact that \( \det G_r = 1 \) according to Eq.(2.3.25). Considering a second paraxial ray, we have

\[
m' n_0 u' - h' n_0 w' = m n_0 u - h n_0 w = H
\]

(2.3.31)

Due to the invariant \( \mu \) given by Eq.(2.3.12) which holds for arbitrary finite rays we have additional paraxial transfer invariants. Inserting Eqs.(2.3.23) and \( y = \eta = 0 \) into Eq.(2.3.15) yields

\[
e_1 = kh' - u'^2 = kh^2 + u^2 = \text{const}
\]

(2.3.32)

Similarly, for the second paraxial ray we have
Another useful invariant is
\[ e_3 = kh'm' + u'w' = khm + uw = \text{const} \]  (2.3.34)
(This way of numbering the invariants is more convenient for writing the expressions of aberration coefficients in a systematic form.) However, these four invariants are not independent. It can be easily proven that
\[ e_4 e_3 = e_2^2 + \frac{k}{n_0^2} H^2 \]  (2.3.35)

Direct substitution of the definitions yields:
\[ e_2^2 + \frac{k}{n_0^2} H^2 = (khm + uw)^2 + \frac{k}{n_0^2}(mn_0u - hn_0w)^2 = k^2 h^2 m^2 + u^2 w^2 + 2khuw + km^2 u^2 + kh^2 w^2 - 2khmuw = \]
\[ = k^2 h^2 m^2 + u^2 w^2 + km^2 u^2 + kh^2 w^2 = (kh^2 + u^2)(km^2 + w^2) = e_3 e_4 \]

The invariance of \( e_2 \) follows obviously from the invariance of \( e_4, e_3 \) and \( H \). The invariants (2.3.32-34) have been first obtained by H.H. Hopkins /18/.

### 2.3.3 Special cases

For a few special radial refractive index distributions, the path of certain finite rays can be determined analytically. For parabolic index distributions, exact solutions can be found for all rays. This will be used in Sec.2.4 where a perturbation method is developed which gives approximate analytic solutions for arbitrary near-parabolic index distributions. Another distribution that has received considerable interest in the literature is
\[ n(r) = n_0^2 \sec h^2 (gr) = n_0^2 \left[ 1 - (gr)^2 + \frac{2}{3} (gr)^4 - \frac{17}{45} (gr)^6 + \frac{62}{315} (gr)^8 + \ldots \right] \]  (2.3.36)
because for axial object points it enables aberration-free imaging /14/,/25/, In a thick positive radial GRIN medium having this refractive index distribution, all rays starting from an axial object point are imaged sharply into a succession of image points situated at equal distance between each other. In order to prove this, we consider only meridional rays, i.e. we put \( y = \eta = 0 \) and \( x = r \). For determining the path of these rays, we start from Eqs.(2.1.10) and (2.3.4). In this case we have
\[ \xi_0^2 \left( \frac{dr}{dz} \right)^2 + \xi_0^2 = n_0^2 \sec h^2 (gr). \]  (2.3.37)

Using the following general relations between hyperbolic functions:
\[ \text{sech} x = \frac{1}{\cosh x}, \quad \cosh^2 x - \sinh^2 x = 1, \quad \frac{d}{dx} \sinh x = \cosh x \]
we obtain from Eq. (2.3.37)
\[ dz = \pm \frac{dr}{\sqrt{\frac{n_0^2}{\xi_0^2} \sinh^2(gr) - 1}} = \pm \frac{\cosh gr dr}{\sqrt{\frac{n_0^2}{\xi_0^2} - \cosh^2 gr} - 1} = \pm \frac{d(sinh gr)}{\sqrt{\frac{n_0^2}{\xi_0^2} - 1 - \sinh^2 gr}} \]  

Introducing \( t = \sinh gr \) as a variable of integration in the right hand side of Eq.(2.3.38) yields

\[ \pm gz = \int_{\sinh gr_0}^{\sinh gr} \frac{dt}{A} = \int_{\sinh gr_0}^{\sinh gr} \frac{dt}{\sqrt{A^2 - t^2}} \]  

Noting that

\[ \int \frac{dt}{\sqrt{A^2 - t^2}} = \arcsin \frac{t}{A} \]  

we obtain from Eq. (2.3.39)

\[ E(z) = \sinh gr = A \sin(B \pm gz), B = \arcsin \frac{\sinh gr_0}{A} \]  

Finally, the ray path of meridional rays is given by

\[ r(z) = \frac{1}{g} \arcsinh E(z) = \frac{1}{g} \ln \left( E(z) + \sqrt{1 + E(z)^2} \right) \]  

Since \( E(z) \) is periodic, \( r(z) \) is also periodic. The spatial period

\[ z_p = \frac{2\pi}{\xi} \]  

does not depend on the initial ray direction and is therefore the same for all meridional rays. If at \( z = 0 \) we have an axial object point i.e. \( r_0 = 0 \), then all rays starting from this point with \( \zeta_0 > 0 \) pass through the axial points having the coordinates \( z = z_p, z = 2z_p, \ldots \)

Consider now the optical path length of meridional rays inside this medium. Because \( ndz = \zeta_0 ds \) we have in radial gradients

\[ L = \int n ds = \int \frac{1}{\xi_0} n^2 dz \]  

From Eqs. (2.3.36) and (2.3.42) we obtain after some algebra

\[ n^2 \left( r(z)^2 \right) = \frac{n_0^2}{1 + E(z)^2} \]  

We insert now Eqs. (2.3.45) and (2.3.41) into Eq.(2.3.44). From the integration tables we find that this type of integrals can be evaluated in closed form as

\[ \int \frac{dx}{1 + c^2 \sin^2 ax} = \frac{1}{a \sqrt{1 + c^2}} \arctan \left( \sqrt{1 + c^2 \tan ax} \right). \]

Hence, we obtain

\[ L = \frac{n_0}{g} \left\{ \arctan \left[ \frac{n_0}{\xi_0} \tan(\pm gz + B) \right] - \arctan \left[ \frac{n_0}{\xi_0} \tan B \right] \right\} \]  

As an example of aberration-free imaging, consider a so-called quarter-pitch rod lens, i.e. a lens with flat end faces and a thickness \( d \) chosen such that the ray travels one quarter of
the period inside the lens. It can be seen from the second of the paraxial transfer formulae (2.3.24) that, for any positive index distribution, if

$$gd = \frac{\pi}{2},$$ (2.3.47)

then paraxial rays which are parallel to optical axis in the object space are focused on the axis at the second surface of the quarter pitch rod lens, i.e. an axial object point at infinity has its image at the second surface (See Fig. 2.3.1.a). If the object is situated on the first surface, then its image is at infinity (See Fig. 2.3.1.b).

If the refractive index distribution is given by Eq. (2.3.36) and the thickness by Eq. (2.3.47) consider a point source situated on the axis at the first surface, as shown in Fig. 2.3.1.b. Because of $r_0 = 0$ we have in Eqs. (2.3.41) and (2.3.46) $B = 0$. Because of Eq. (2.3.47) the first square bracket in Eq. (2.3.46) tends to infinity and we have

$$L = \frac{n_0 \pi}{g} = \frac{n_0 d}{2}$$ (2.3.48)

i.e. the optical path length of all rays inside the lens is equal to its value along the optical axis and the wave aberration vanishes. For this kind of applications, the index distribution (2.3.36) producing aberration-free imaging is sometimes called "ideal distribution".

![Fig. 2.3.1 Axial imaging by a quarter-pitch rod lens: a) focusing, b) collimating](image)

In practice, the output of a diode laser or an optical fibre can be collimated by a quarter-pitch lens and then be refocused into another fibre or onto a detector. For this purpose, this type of lenses have been produced by Nippon Sheet Glass. However, in practical applications the source has finite dimensions and therefore the optimal refractive index distribution is not exactly that given by Eq. (2.3.36). Moreover, it has been found that sometimes it is convenient to make the lens slightly shorter than one quarter pitch and to allow for a short space between the fibre and the lens as an additional parameter to improve image quality (e.g. $d=0.23$ pitch to compensate dispersion in a collimating system designed for more than one wavelength) /55/.
2.4 Analytic ray-tracing method for finite rays in radial gradients

We now describe a new method to derive algebraic formulae for tracing skew rays in radial gradient-index media. In the next three sections, we obtain power series for the ray position and direction, optical path length and coordinates of the intersection point of the curved ray with the spherical end surface of the GRIN lens. In Chapter 3., these formulae are the starting point for the derivation of Seidel aberration coefficients.

As noted in Sec. 2.1.1, at the present time the optical design with GRIN lenses is based mainly on the iterative solution of the differential ray equations (2.1.13) inside the lens /50/, /51/, /52/. As in the case of axial media (Sec. 2.2.1), for radial GRIN lenses, an alternative ray-tracing method is to find approximate analytic solutions of these equations. In this case the ray position and direction and the optical path length are given as functions of the initial ray data and gradient profile parameters. Such formulae have been first obtained by Streifer and Paxton /40/, /56/ and Marchand /26/, /28/. For ray-tracing, the analytic formulae can be faster than iterative methods /28/ and have also been used for the derivation of aberration formulae for radial GRIN lenses with plane faces (Wood lenses) /27/ and for the determination of the refractive index profile from measurements /42/, /43/, /61/.

If accurate results are required, analytic formulae become lengthy and the computation of higher order contributions by hand in these formulae is cumbersome. However, recent progress of computer algebra software makes this type of calculations quite manageable. Therefore, in order to obtain high-order ray-tracing formulae what was needed was a systematic derivation method that makes it possible to distinguish clearly between the main ideas of the derivation and calculations that should be automated on the computer. Such a method is described in what follows.

In Ref. /6/ we have developed a new method for the derivation of analytic skew ray tracing formulae in radial GRIN media. Instead of attempting to solve the coupled ray equations (2.3.10) directly, first a differential equation for an intermediate variable is solved (Sec. 2.4.1). By substituting this solution in the ray equations the complexity of the skew-ray calculations is considerably reduced (Sec 2.4.2).

The resulting formulae for the ray position and direction and optical path length are power series where each coefficient can be calculated from the coefficients of preceding orders. The concise statement of the derivation steps makes it possible to perform the last
stage of the lengthy computations of the formulae automatically on a computer using widespread symbolic mathematics software.

Prior to our study / 6/, analytic formulae have been known only for Wood lenses. For the first time, the new method also enables the derivation of analytic expressions for the coordinates of the intersection point of the curved ray with the spherical end surface of the radial GRIN lens. Thus, the range of applicability of the class of analytic methods is extended, the method being now suitable for arbitrary radial GRIN lenses.

We have calculated in all these formulae the expressions of terms up to the eighth order (i.e. terms due to the eighth power of the radius in the refractive index distribution (2.3.6)). These formulae are valid both for positive and for negative radial GRIN media. In what follows, only the derivation method will be described in detail. Because of their length, the resulting ray-tracing formulae are listed separately in Appendix A. Formulae of the same order of approximation but only for ray position and direction in positive radial GRIN media have been previously described in Refs. / 40/,/ 56/. Simultaneously with us / 6/, Sakamoto has published an article describing eighth-order analytic formulae for the optical path length / 4/. At the present moment, however, his published results are only for meridional rays, while ours are for arbitrary skew rays.

2.4.1 Decoupling the differential ray equations

The equations (2.3.10) are coupled, second-order nonlinear ordinary differential equations. Streifer and Paxton / 40/ and Marchand / 28/ developed approximate analytical methods to solve these equations directly. The main idea of our method is as follows: We have found that it is possible to derive a differential equation for an intermediate variable \( \chi \)

\[
\chi = (gr)^2
\]  

(2.4.1)
as shown in the remainder of this section, and divide the computation of the solution of Eqs. (2.3.10) into two steps. First, the equation for \( \chi \) is solved obtaining \( \chi \) as a function of the independent variable \( t \) given by Eq.(2.3.8). By introducing this solution in Eqs(2.3.10) it can be seen that these equations are reduced to two independent linear differential equations with variable coefficients

\[
\begin{align*}
\ddot{x} + f(t)x & = 0 \\
\ddot{y} + f(t)y & = 0 \\
f(t) & = 2N_s\chi(t) + 3N_s\chi^2(t) + 4N_s\chi^3(t) + ...
\end{align*}
\]  

(2.4.2)
and can then be solved much more comfortably. Moreover, for arbitrary skew rays, the optical path length can be computed directly from the solution for $\chi$. Using Eqs. (2.3.44), (2.3.6) and (2.3.8) we obtain

$$L = \int n ds = \frac{1}{\xi_0} \int n^2 dz = \frac{n_0}{g} \Phi(t)$$

(2.4.3)

$$\Phi(t) = \int (1 - \chi(t') + N_4 X^2(t') + N_6 X^4(t') + \ldots) dt'$$

Thus, unlike Sakamoto’s method /44/, our method enables the computation of the optical path length in the general case of skew rays.

By denoting $\hat{\chi} = \tau^2$, i.e. $\chi = g^2 \hat{\chi}$, let us now show that for any radial refractive index distribution, $\hat{\chi}$ satisfies the equation

$$\frac{d^2}{dz^2} \hat{\chi} = 2 \frac{d}{d \hat{\chi}} \left[ \hat{\chi} \left( \frac{n^2(\hat{\chi})}{\xi_0^2} - 1 \right) \right].$$

(2.4.4)

By noting that

$$\frac{d^2}{dz^2} \chi \bigg|_{\chi = g^2 \hat{\chi}} = 2 \chi \frac{d^2 x}{dz^2} + 2x \frac{d^2 x}{dz^2}$$

and that a similar relation holds for $y$, we have

$$\frac{d^2}{dz^2} (x^2 + y^2) = 2 \left( \frac{dx}{dz} \right)^2 + 2x \frac{d^2 x}{dz^2} + 2 \left( \frac{dy}{dz} \right)^2 + 2y \frac{d^2 y}{dz^2}$$

(2.4.5)

Multiplying the first of the Eqs. (2.3.5) by $x$ and the second by $y$ and summing up the results yields

$$x \frac{d^2 x}{dz^2} + y \frac{d^2 y}{dz^2} = x^2 + y^2 \frac{dn^2(r^2)}{dr^2} = \hat{\chi} \frac{d}{d \hat{\chi}} \left( \frac{n^2(\hat{\chi})}{\xi_0^2} \right)$$

(2.4.6)

where $d \hat{\chi} = 2 r dr$ was used. From Eqs. (2.3.4) and (2.1.10) we obtain

$$\frac{dx}{dz} + \frac{dy}{dz} \left( \frac{\xi}{\xi_0} \right)^2 + \left( \frac{\eta}{\xi_0} \right)^2 = \frac{n^2(\hat{\chi})}{\xi_0^2} - 1$$

(2.4.7)

Substituting Eqs. (2.4.6) and (2.4.7) into Eq. (2.4.5) yields

$$\frac{d^2 \hat{\chi}}{dz^2} = 2 \left( \frac{n^2(\hat{\chi})}{\xi_0^2} - 1 + \hat{\chi} \frac{d}{d \hat{\chi}} \frac{n^2(\hat{\chi})}{\xi_0^2} \right)$$

(2.4.8)

Finally, Eq. (2.4.4) is obtained by rearranging the right hand side of Eq. (2.4.8).

For the refractive index distribution (2.3.6) and the variable $\chi$ given by Eq. (2.4.1) we replace in Eq. (2.4.4) the independent variable $z$ by $t$ given by Eq. (2.3.8). After some simple calculations, the differential equation for $\chi$ can be written as

$$\hat{\chi} + 4 \chi = 4 \mu + 6N_4 \chi^2 + 8N_6 \chi^3 + 10N_8 \chi^4 + \ldots$$

(2.4.9)

where $\mu$ is the transfer invariant given by Eq. (2.3.12).
2.4.2 Ray path determination by power series expansion

As a first step, we find an approximate solution of Eq. (2.4.9). The solution will be sought as the series
\[ \chi(t) = \mu \chi_0(t) + \mu^2 \chi_1(t) + \mu^3 \chi_2(t) + \mu^4 \chi_3(t) + \ldots \] (2.4.10)
with respect to the invariant \( \mu \) which indicates, as can be seen from Eq. (2.3.15), how close the ray path to the optical axis is. For simplicity, we require that
\[ \chi(0) = \mu \chi_0(0), \dot{\chi}(0) = \mu \ddot{\chi}_0(0) \] (2.4.11)
i.e. the higher order contributions and their first derivatives vanish for \( t=0 \). Substituting Eq. (2.4.10) into Eq. (2.4.9) and equating the terms with equal powers of \( \mu \) we arrive at a set of equations for the coefficients \( \chi_i, i=0,1,2,3 \ldots \)

The first equation in this set, obtained from the terms containing \( \mu \) at the first power,
\[ \ddot{\chi}_0 + 4 \dot{\chi}_0 - 4 = 0 \] (2.4.12)
has the solution
\[ \chi_0(t) = a \cos 2t + b \sin 2t + 1 \] (2.4.13)
where \( a \) and \( b \) are found from the initial conditions. It follows from Eqs. (2.4.11), (2.3.11) and (2.4.1) that
\[ a = \frac{1}{\mu} g^2 (x_0^2 + y_0^2) - 1, b = \frac{g}{n_0 \mu} (x_0 \xi_0 + y_0 \eta_0) \] (2.4.14)

For the higher order coefficients \( \chi_i, i=1,2,3, \ldots \), a closer examination of the origin of various types of terms containing \( \mu^{i+1} \) reveals that these coefficients must be of the form
\[ \chi_i(t) = N_4 \chi_{i1}(t) \]
\[ \chi_2(t) = N_4 \chi_{21}(t) + N_6 \chi_{22}(t) \]
\[ \chi_3(t) = N_4 \chi_{31}(t) + N_4 N_6 \chi_{32}(t) + N_6 \chi_{33}(t) \] (2.4.15)

We consider in our ray-tracing formulae the effects of terms up to the eighth power of the radius in the refractive index distribution (2.3.6) i.e. \( i=1,2,3 \). After separating the various types of terms, we find that the set of equations for \( \chi_{ij} \) is given by
\[ \ddot{\chi}_{ij} + 4 \dot{\chi}_{ij} = f_{ij}(t) \] (2.4.16)
where we have used the abbreviations
The equations (2.4.16-17) can be solved one after the other. For i=0 the corresponding coefficient is given by Eqs. (2.4.13-14). It can be seen from Eqs. (2.4.17) that for each order i>0 and all values of j≤i, the term \( f_{rij} \) contains only coefficients of preceding orders (i.e. those having lower values of i). If the coefficients of preceding orders have already been determined as functions of \( t \), then the equation for \( \chi_{ij} \) is a second order linear differential equation with a \( t \)-dependent inhomogeneous term on the right side.

This type of differential equations can be conveniently solved by means of the Laplace transform /9/. The Laplace transform of an arbitrary function \( f(t) \) (satisfying certain requirements of transformability) is defined as

\[
F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t) dt
\]

We now take the Laplace transform of both sides of Eqs. (2.4.16). According to the general differentiation property of the Laplace transform we have

\[
\mathcal{L}\{\chi_{ij}\} = s^2 \mathcal{L}\{\chi_{ij}\} - s \chi_{ij,0} - \chi_{ij,0} = s^2 \mathcal{L}\{\chi_{ij}\}
\]  

(2.4.18)

because, as follows from Eq. (2.4.11), we have for the initial values \( \chi_{ij,0} = \chi_{ij,0} = 0 \).

Consequently, we obtain

\[
s^2 \mathcal{L}\{\chi_{ij}\} + 4 \mathcal{L}\{\chi_{ij}\} = \mathcal{L}\{f_{rij}\}
\]

and hence

\[
\mathcal{L}\{\chi_{ij}\} = \frac{1}{s^2 + 4} \mathcal{L}\{f_{rij}\}
\]  

(2.4.19)

From the tables of Laplace transforms (see e.g. /9/) we find that

\[
\frac{1}{s^2 + 4} = \mathcal{L}\left\{\frac{1}{2} \sin 2t\right\}
\]  

(2.4.20)

Since the product of the Laplace transforms of two functions is equal to the Laplace transform of the convolution of the functions, we rewrite the right hand side of Eq. (2.4.19) as

\[
\mathcal{L}\{\chi_{ij}\} = \mathcal{L}\left\{\frac{1}{2} \int_0^t \sin 2(t-t') f_{rij}(t') dt'\right\}
\]  

(2.4.21)

Finally, by back-transforming Eq. (2.4.21) and expanding \( \sin 2(t-t') \) according to the addition theorem for sines, we arrive at the solution

\[
f_{r11} = 6 \chi_0^2
\]

\[
f_{r21} = 12 \chi_0 \chi_{11}
\]

\[
f_{r22} = 8 \chi_0^3
\]

\[
f_{r31} = 6 \chi_{11}^2 + 12 \chi_0 \chi_{21}
\]

\[
f_{r32} = 24 \chi_0^2 \chi_{11} + 12 \chi_0 \chi_{22}
\]

\[
f_{r33} = 10 \chi_0^4
\]  

(2.4.17)
Thus, using Eq. (2.4.22), the equations (2.4.16-17) can be successively solved. For \( i=1 \) we obtain

\[
\chi_{ij}(t) = -\frac{1}{4} (a^2 - b^2) \cos 4t - \frac{1}{2} ab \sin 4t - 3 bt \cos 2t - \frac{1}{2} (a^2 + 2b^2 + 3) \cos 2t + 3at \sin 2t + \frac{1}{4} (2ab + 3b) \sin 2t + \frac{3}{4} (a^2 + b^2 + 2)
\]

(2.4.23)

Since for higher order coefficients the length of the formulae increases rapidly with \( i \) we do not reproduce them here. However, applying recently developed computer algebra software (e.g. DERIVE™ / 62/ or MATHEMATICA™ / 60/) makes the evaluation of these coefficients an easy matter. The results have the following structure

\[
\chi_{ij}(t) = T_{rij} + \sum_{k=1}^{i+1} \sum_{l=1}^{k} \left( C_{rijkl} t^{k-l} \cos 2lt + S_{rijkl} t^{k-l} \sin 2lt \right)
\]

(2.4.24)

where \( T_{rij} \), \( C_{ijkl} \), and \( S_{ijkl} \) are polynomials of degree \( i+1 \) in \( a \) and \( b \) with rational coefficients. These coefficients are used in what follows for solving Eqs. (2.4.2) and for determining the optical path length according to Eq. (2.4.3).

We arrive now at the second step of the ray path determination. As noted before, after substituting the solution \( \chi(t) \) into Eqs. (2.4.2), these equations become two similar linear differential equations with variable coefficients. The solution for both \( x \) and \( y \) can be written as a linear combination of the two fundamental solutions, denoted by \( \phi_c(t) \) and \( \phi_s(t) \). The corresponding optical direction cosines are then a linear combination of the derivatives of the fundamental solutions. The coefficients are found from the initial conditions as shown previously for Eqs. (2.3.22).

\[
x(t) = x_0 \phi_c(t) + \frac{\xi_0}{n_0} \phi_s(t)
\]

\[
y(t) = y_0 \phi_c(t) + \frac{\eta_0}{n_0} \phi_s(t)
\]

(2.4.25)

We seek both \( \phi_c(t) \) and \( \phi_s(t) \) as series expansions of the form

\[
\phi(t) = \phi_0(t) + \mu \phi_1(t) + \mu^2 \phi_2(t) + \mu^3 \phi_3(t) + ... 
\]

(2.4.26)

As in the case of \( \chi(t) \) we define
\[
\phi(t) = N_1 \phi_1(t) \\
\phi_2(t) = N_2 \phi_2(t) + N_6 \phi_{22}(t) \\
\phi_3(t) = N_3 \phi_3(t) + N_4 N_6 \phi_2(t) + N_6 \phi_{33}(t)
\] (2.4.27)

and substitute Eqs.(2.4.26-27), (2.4.10) and (2.4.15) into Eqs. (2.4.2) . Equating the terms with equal powers of \( \mu \) we obtain the set of equations

\[
\ddot{\phi}_0(t) + \phi_0(t) = 0
\] (2.4.28)

and

\[
\ddot{\phi}_0(t) + \phi_0(t) = f_{xy}(t)
\] (2.4.29)

with the abbreviations

\[
\begin{align*}
 f_{s11} &= 2 \chi_0 \phi_0 \\
 f_{s21} &= 2 \chi_1 \phi_0 + 2 \chi_0 \phi_1 \\
 f_{s22} &= 3 \chi_0 \phi_0 \\
 f_{s31} &= 2 \chi_2 \phi_0 + 2 \chi_1 \phi_1 + 2 \chi_0 \phi_2 \\
 f_{s32} &= 6 \chi_0 \chi_1 \phi_0 + 2 \chi_2 \phi_0 + 3 \chi_0 \phi_1 + 2 \chi_0 \phi_2 \\
 f_{s33} &= 4 \chi_0 \phi_0
\end{align*}
\] (2.4.30)

The two fundamental solutions \( \phi_0(t) \) and \( \phi_s(t) \) are then obtained by substituting the two linearly independent solutions of Eq.(2.4.28)

\[
\phi_0(t) = \cos t \\
\phi_0(t) = \sin t
\] (2.4.31)

as \( \phi_0 \) in Eq.(2.4.30) and using

\[
\phi_0(t) = \sin \int f_{xy}(t') \cos t' dt' - \cos \int f_{xy}(t') \sin t' dt'
\] (2.4.32)

for solving successively Eqs. (2.4.29). The derivation of Eq. (2.4.32) is the same as for Eq. (2.4.22).

These calculations have been performed symbolically on the computer and the final formulæ have been translated automatically into FORTRAN. Because of their length, we list the results separately in Appendix A. We give them here only for \( \iota=1 \)

\[
\begin{align*}
\phi_{s11}(t) &= -\frac{1}{8} \left( a \cos 3t - \frac{1}{8} \right) b \sin 3t - \frac{1}{2} b t \cos t + \frac{1}{2} (a + 2) t \sin t + \frac{7}{8} b \sin t \\
\phi_{s21}(t) &= -\frac{3}{8} b \cos 3t + \frac{3}{8} a \sin 3t + \frac{1}{2} (a + 2) t \cos t + \frac{1}{2} b t \sin t + \frac{1}{8} (3a + 8) \sin t \\
\phi_{s22}(t) &= \frac{1}{8} b \cos 3t - \frac{1}{8} a \sin 3t + \frac{1}{2} (a - 2) t \cos t - \frac{1}{8} b t \sin t - \frac{1}{8} (a - 8) \sin t \\
\phi_{s31}(t) &= -\frac{3}{8} b \cos 3t - \frac{3}{8} a \sin 3t + \frac{1}{2} b t \cos t + \frac{3}{8} a \cos t - \frac{1}{2} (a - 2) t \sin t + \frac{5}{8} b \sin t
\end{align*}
\] (2.4.33)
In Chapter 3 of this study, Eqs. (2.4.33) will be used for the derivation of the transfer contributions of radial GRIN lenses to the Seidel aberration coefficients of the system.

Finally, the ray path inside the radial GRIN lens is obtained by substituting Eqs. (2.4.31) and the results for $\phi_{cij}$ and $\phi_{sij}$ into Eqs. (2.4.25-27). Recall that $\mu$ and $t$ are given by Eqs.(2.3.12) and (2.3.8).

As noted in the previous section, the optical path length can be computed directly from the solution $\chi(t)$. Substituting into the left hand side of the second of Eqs. (2.4.3)

$$\Phi(t) = t + \mu\Phi_0(t) + \mu^2\Phi_1(t) + \mu^3\Phi_2(t) + \mu^4\Phi_3(t) + ...$$  \hspace{1cm} (2.4.34)

and

$$\Phi_1(t) = N_4\Phi_{11}(t)$$
$$\Phi_2(t) = N_4^2\Phi_{21}(t) + N_6\Phi_{22}(t)$$
$$\Phi_3(t) = N_4^2\Phi_{31}(t) + N_4^4\Phi_{32}(t) + N_6\Phi_{33}(t)$$  \hspace{1cm} (2.4.35)

and in the right hand side Eqs.(2.4.10) and (2.4.15) we find

$$\Phi_0 = \int (\chi_0) dt'$$
$$\Phi_{11} = \int (\chi_0^2 - \chi_{11}) dt'$$
$$\Phi_{21} = \int (2\chi_0\chi_{11} - \chi_{21}) dt'$$
$$\Phi_{22} = \int (\chi_0^3 - \chi_{22}) dt'$$
$$\Phi_{31} = \int (\chi_{11}^2 + 2\chi_0\chi_{21} - \chi_{31}) dt'$$
$$\Phi_{32} = \int (3\chi_0^2\chi_{11} + 2\chi_0\chi_{32} - \chi_{32}) dt'$$
$$\Phi_{33} = \int (\chi_0^4 - \chi_{33}) dt'$$  \hspace{1cm} (2.4.36)

The results, which enable the determination of the optical path length of arbitrary skew rays in the radial GRIN medium, are also listed in Appendix A.

In the special case of a purely parabolic refractive index distribution, i.e. when in Eq. (2.3.6) we have $N_4=N_6=N_8=...=0$, we obtain $\phi_c(t)=\phi_{c0}(t)$ and $\phi_s(t)=\phi_{s0}(t)$ and the ray path is given in a closed form by

$$x(t) = x_0 \cos t + \frac{\xi_0}{n_0g} \sin t$$
$$y(t) = y_0 \cos t + \frac{\eta_0}{n_0g} \sin t$$
$$\xi(t) = -n_0gx_0 \sin t + \xi_0 \cos t$$
$$\eta(t) = -n_0gy_0 \sin t + \eta_0 \cos t$$  \hspace{1cm} (2.4.37)
In can be seen that in the paraxial approximation Eqs. (2.4.37) are reduced to Eqs. (2.3.22).

This approximate method for determining the ray path has a close resemblance to what is called in mathematical physics a perturbation method. For a purely parabolic refractive index distribution, the ray path is given by an exact solution. Deviations from the parabolic shape lead to nonzero values for the higher order coefficients in Eq. (2.3.6) and also, according to Eqs. (2.4.26-27), to corrections to the ray path.

For positive radial GRIN media (i.e. \(k>0\)) the two quantities \(a\) and \(b\) given by Eqs. (2.4.14) are of the order of magnitude of one. (This can be shown easily for the parabolic case.) In this case, for a fixed value of \(t\), the convergence of the series expansions for \(\chi, \phi, \phi_s, \Phi\) is determined by \(\mu\) and \(N_4, N_6, N_8, \ldots\). The smaller the values of the dimensionless coefficients \(N_4, N_6, N_8, \ldots\) are, the larger is the domain of values of \(\mu\) for which a reasonable convergence of these series can be expected. Not too large values of these coefficients correspond to a refractive index distribution which is not very different from a parabola (especially in the near-axis domain). Therefore, the present method is especially useful for tracing rays in gradients with so-called parabolic-like profiles. However, previous experience with radial GRIN lenses has shown that precisely this type of profiles is most useful for aberration correction /34/. A typical example showing the convergence of these series will be given in Sec. 2.4.4.

For \(k<0\) and for certain values of the initial conditions, \(\mu\) can be zero and according to Eqs. (2.4.14) \(a\) and \(b\) become infinite. This problem can be avoided by taking advantage of the fact that each term of \(\chi, \phi, \phi_s, \Phi\) is a polynomial in \(a\) and \(b\) of the same degree as the corresponding power of \(\mu\), and by defining \(a = \mu \hat{a}, b = \mu \hat{b}\). Thus \(\chi, \phi, \phi_s, \Phi\) become power series expansions in \(\hat{a}, \hat{b}, \mu\).

2.4.3 Curved end surfaces

For practical applications, the ray-tracing formulae give the ray position and direction at the surface following the radial GRIN medium as functions of the ray parameters at the surface preceding it. The formulae developed in the previous section enable the determination of position and direction as functions of \(z\), which is the distance of propagation of the ray in the radial GRIN medium measured along the optical axis. Thus, if the end surfaces are plane (Wood lens) and the distance between them is denoted by \(d\), these formulae enable the tracing of rays by simply substituting \(z\) by \(d\).
Let now the radial GRIN medium be situated between two curved surfaces. In this case the value of \( z \) must be determined for each ray separately from the two ray-surface intersection points. If the medium before the first surface is homogeneous, the intersection point of the ray with this surface can be determined as usual /23/, /59/. Therefore, we discuss in what follows the case when the ray reaches the second surface.

Prior to our study /6/, the accurate computation of the ray-surface intersection point in the case of radial GRIN lenses with curved end surfaces was possible only with methods relying on iterative solutions of the ray equations /52/. In this section we show how this computation can be performed starting from the analytic formulae of the previous section.

We first determine the ray position and direction at the tangent plane at the vertex using the formulae of the previous section i.e. as for a plane end surface and regard these values as new initial conditions. To avoid confusion, we denote in this section all quantities related to the vertex of the end surface by a \(^{v} \) above the symbol.

The coordinates of the intersection point of the ray with a rotationally symmetric end surface satisfy an equation of the form

\[
\bar{z} = F(r^2) \quad .
\]  

(2.4.38)

As can be easily seen from Fig. (2.1.2), for a spherical surface of radius \( R = 1/\rho \) with the vertex at the origin of the coordinate system we have

\[
F(r^2) = R \left( 1 - \frac{1 - r^2}{R^2} \right) = R \left[ \frac{1}{2} \frac{r^2}{R^2} + \frac{1}{2 \cdot 4} \left( \frac{r^2}{R^2} \right)^2 + \frac{3}{2 \cdot 4 \cdot 6} \left( \frac{r^2}{R^2} \right)^3 + \frac{3 \cdot 5}{2 \cdot 4 \cdot 6 \cdot 8} \left( \frac{r^2}{R^2} \right)^4 + \ldots \right]
\]  

(2.4.39)

As follows from Eqs. (2.4.1), (2.3.8) and (2.4.38) we have the system of equations

\[
\begin{align*}
\frac{1}{g} \chi(\bar{r}) &= r^2 \\
\frac{\bar{r}}{\xi_0} &= \frac{n_0 g}{\xi_0} F(r^2)
\end{align*}
\]  

(2.4.40)

From the solution of Eqs. (2.4.40), the \( z \)-coordinate of the intersection point is given by

\[
\bar{z} = \frac{\bar{r}}{n_0 g} \bar{r}
\]  

(2.4.41)

We first eliminate \( r^2 \) by substituting the first of Eqs. (2.4.40) into the second one. Since \( \bar{r} \) is small we can use Taylor series expansions for the trigonometric functions in Eqs.(2.4.13) and (2.4.24) for the evaluation of \( \chi \). Since according to Eq.(2.4.10) \( \chi \) is a
power series expansion with respect to $\mu$, we seek the solution for $r$ also as a series of
the form

$$r = \mu^1 t_1 + \mu^2 t_2 + \mu^3 t_3 + \mu^4 t_4 + \mu^5 t_5 + \mu^6 t_6 + ...$$

(2.4.42)

Introducing Eq.(2.4.42) into Eq. (2.4.40) and equating (on the computer) the terms with
equal powers of $\mu$ we arrive at a set of equations for $t_i$ where, as in the case of $\chi$ and $\phi$,
each coefficient $t_i$ can be calculated from the coefficients of preceding orders. We found
that, in order to consider the effects of terms up to the eighth power of the radius in the
refractive index distribution (2.3.6) we had to compute six terms in Eq.(2.4.42). We note
at this point that the procedure presented above works only due to the fact that the series
expansion (2.4.10) for $\chi$ does not contain a zeroth power of the expansion variable $\mu$. The
results are listed in Appendix A. Here, we note only that the lowest order contribution to $\tilde{z}$
is precisely the same as for homogeneous media

$$\tilde{z} = \frac{1}{2} \rho (x_0^2 + y_0^2) + ...$$

(2.4.43)

Up to this point we have used in the ray-tracing equations a notation that seems to be the
most convenient for solving differential equations. However, for deriving aberration co-
efficients (as it will be shown in Chapter 3.) a slight change of notation is appropriate.
For a given finite ray consider the planes $A$ and $A'$ perpendicular to the symmetry axis
$Oz$ and passing through the intersection points of the ray with the first and second sur-
face. Let $\delta z$ and $\delta z'$ be the distances between the corresponding planes and surface vertices. As shown above, $\delta z'$ is given by Eqs.(2.4.41-42) (where it was denoted by $\tilde{z}$). Thus, the distance along $Oz$ for the propagation of the given ray in the radial GRIN medium of axial thickness $d$ is

$$z = d + \delta z' - \delta z$$

(2.4.44)

The ray coordinates $(x,y)$, the corresponding optical direction cosines $(\xi,\eta)$ and any
functions of them are now denoted as follows: before transfer (i.e. the values previously
denoted by the index 0) by unprimed symbols and after transfer (i.e. the values for $z$ given
by Eq.(2.4.44)) by primed ones. Thus, Eqs.(2.4.25) become

$$x' = x_\phi (t) + \frac{\xi}{n_0 g} \phi$$(t)

$$y' = y_\phi (t) + \frac{\eta}{n_0 g} \phi$$ (t)  

$$\xi' = n_0 g x_\phi (t) + \xi \phi$$(t)

$$\eta' = n_0 g y_\phi (t) + \eta \phi$$ (t)  

(2.4.45)
2.4.4 Test of accuracy

In order to verify the correctness of the analytic ray-tracing formulae described in the previous sections, we have made three types of comparisons with the results generated by these formulae.

i) Comparison with earlier results from the literature

ii) Comparison for a special case where an exact solution is known for the ray path

iii) Comparisons with iterative solutions

In this section we describe only the test of the formulae giving the ray position, direction and optical path length as functions of distance in positive radial gradients. A more complete test, for lenses with curved end surfaces and considering both positive and negative gradients, will be given in Sec. (A.3) of Appendix A.

i) We first compare our results with earlier analytic results of Marchand on the basis of a numerical example described in Ref. /28/. Marchand's analytic formulae are of a lower order than ours, considering the sixth order term in Eq.(2.3.6) for ray position and direction and fourth order term for optical path length. The following gradient profile was assumed:

\[ n_0 = 1.5, \ g = 2\pi/67, \ N_4 = 2/3, \ N_6 = -17/45, \ N_8 = 0 \]

For a typical ray and distance,

\[ x_0 = y_0 = 0.1, \ \xi_0 = 0.12, \ \eta_0 = 0.13, \ z = 10 \]

we compare in Table (2.4.1) the following four types of results for ray position, direction and optical path length:

1. Marchand's analytical results
2. Numerical results used by Marchand to test his formulae
3. Results with our formulae using the same order of approximation as Marchand (i.e. 6-th for \( x,y,\xi,\eta \) and 4-th for \( L \))
4. Our eighth order results

For clarity the relevant digits have been emphasized by underlining in the rows 1 and 3 and by boldfacing in the rows 2 and 4.

<table>
<thead>
<tr>
<th></th>
<th>( x )</th>
<th>( y )</th>
<th>( \xi )</th>
<th>( \eta )</th>
<th>( L )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.750554325</td>
<td>0.808204373</td>
<td>0.0594095553</td>
<td>0.0653051455</td>
<td>15.0364008</td>
</tr>
<tr>
<td>2</td>
<td>0.750554318</td>
<td>0.808204314</td>
<td>0.0594095443</td>
<td>0.0653051336</td>
<td>15.0364002</td>
</tr>
<tr>
<td>3</td>
<td>0.75055432505</td>
<td>0.80820432323</td>
<td>0.059409555294</td>
<td>0.065305145521</td>
<td>15.036400805</td>
</tr>
<tr>
<td>4</td>
<td>0.75055431614</td>
<td>0.80820431369</td>
<td>0.059409544136</td>
<td>0.065305133540</td>
<td>15.036400222</td>
</tr>
</tbody>
</table>

Tab. 2.4.1 Comparison of our results with the results of Marchand (see text)
It can be seen that our eighth order results are in full agreement with Marchand's iterative solution. However, at the same order of approximation, our results are also nearly the same as the analytical results of Ref. /28/ despite the fact that Marchand's formulae have a seemingly different form than ours.

ii) As shown in Sec. (2.3.3), if the refractive index distribution is given by Eq. (2.3.36), i.e. $N_4=2/3$, $N_6=-17/45$, $N_8=62/315$, the ray path of meridional rays can be expressed in closed form by Eqs. (2.3.41-42). We choose as in the previous case $n_0=1.5$, $g=2\pi/67$, $z=10$

and consider a meridional ray having $x_0=0.1$, $\xi_0=0.12$, $\eta_0=y_0=0$

In Table (2.4.2) we compare for the ray position and direction:

1. Exact results
2. Numerical results by iterative method
3. Eighth order analytic results

We observe that all three types of results are in an excellent agreement.

<table>
<thead>
<tr>
<th></th>
<th>$x$</th>
<th>$\xi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.748605255725</td>
<td>0.059540277329</td>
</tr>
<tr>
<td>2</td>
<td>0.748605254740</td>
<td>0.059540277345</td>
</tr>
<tr>
<td>3</td>
<td>0.748605254736</td>
<td>0.059540277335</td>
</tr>
</tbody>
</table>

Tab. 2.4.2 Comparison with exact solution

iii) We have made extensive comparisons of our approximate analytic solutions with the solutions generated by a code using the iterative method given in Ref. /50/. In Appendix A we show a comparison of the analytic results with the results obtained using CODE V /15/. For both positive and negative radial GRIN media the agreement was excellent in all cases.

From all these comparisons we find that for parabolic-like refractive index profiles, as long as the propagation distance $z$ is not very large, the accuracy of these various analytic and iterative methods is nearly the same and is given mainly by the number of terms considered in the refractive index distribution. However, from numerical tests we have found that the accuracy of the approximate analytic formulae decreases slowly with $z$. Our method for determining the ray path and optical path length is very accurate as long as $z$ is not larger than a few periods of the trigonometric functions appearing in these formulae. In most applications this condition is satisfied.
In order to illustrate the accuracy improvement with increasing order of approximation we give now a simple example. Consider the collimation of an axial point source situated directly in front of a quarter-pitch rod lens having the refractive index distribution (2.3.36) (See Fig. 2.3.1b) As shown in Sec.2.3.3, the exact value of the optical path length inside the lens is then independent of the initial ray direction and is equal to its value along the optical axis \( n_0 d \). (Eq. 2.3.48) In this example the parameter \( \mu \) which determines the convergence of the solution series is a function only of the numerical aperture. We have computed the optical path length with our method in four different orders of approximation by successively considering the effects of terms up to the 2-nd, 4-th, 6-th and 8-th order. By comparing these results with the exact value we determine the computation error of the wave aberration in the four cases as function of the numerical aperture as shown in Fig.(2.4.1)

\[
\begin{array}{c}
\text{numerical aperture} \\
0 & 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 \\
1,00E+00 & 1,00E+01 & 1,00E+02 & 1,00E+03 \\
\end{array}
\]

Fig. 2.4.1 Computation error of the wave aberration in four different orders of approximation

It can be seen that the number of terms that have to be computed in order to achieve a prescribed accuracy increases with the numerical aperture. For example, with eighth order terms the error at N.A. = 0.5 is less than 1/100 of a wavelength. For not too high values of the N.A. the accuracy improves with each additional order by a factor between 10 and 100, faster at low N.A. and slower at higher N.A. We have found the same features in the general skew-ray case for all series described in this work. In this case the convergence is determined by the size of both aperture and field.

As can be seen in Appendix A, the length of the analytic formulae increases rapidly with each additional order. For moderate values of aperture and field, low-order analytic formulae are as accurate as iterative methods but are computationally much faster, as
shown in Ref. /28/. With increasing aperture and field, high-order contributions become necessary and the speed advantage is progressively lost. However, from our comparisons we have found that our eighth-order formulae are still faster than iterative methods, especially if very accurate ray-tracing results are required.
2.5 Paraxial approximation for rotationally-symmetric optical systems

2.5.1 Positions of object and image

Earlier in this chapter we have derived the necessary linear transformation formulae for determining the position and direction of any paraxial ray after refraction at a surface or transfer through axial or radial GRIN media. We have also shown that each of these linear transformations can be written by means of Gauss matrices. In this section we show that, for the entire optical system, the Gauss-matrix formalism is a very effective tool for determining the relationships between system data and the positions and sizes of object and image.

Consider a rotationally symmetric optical system consisting of k surfaces and k-1 homogeneous or gradient-index media between the surfaces. The effect of all refractions and transfers within the system upon any paraxial ray can be expressed in a condensed form by means of the Gauss-matrix of the whole system. This matrix relates the paraxial ray parameters at the last surface (after refraction) to the parameters at the first surface (prior to refraction) and is given by the product in the reverse order of ray propagation of all refraction and transfer matrices within the system, starting with the refraction matrix of the last surface and ending with the refraction matrix of the first surface

$$G = B_k T_{k-1}...B_2 T_1 B_1$$

According to the type of the medium, $T$ denotes $G_z$ given by Eq. (2.2.23) for axial gradients, $G_r$ (Eq. (2.3.25)) for positive or negative radial gradients or $U$ (Eq. (2.3.30)) for homogeneous or shallow gradient media. In all cases, $B$ is given by Eq. (2.1.29).

For an optical system described by a Gauss-matrix, the familiar paraxial properties of homogeneous optical systems are retained. Thus, for any object plane, there exists a unique conjugate image plane such that, in the paraxial approximation, every point of the object plane is imaged sharply at a corresponding point of the image plane. Also, the paraxial transverse magnification is the same for all points of the given object plane. For a given value of the transverse magnification $\beta$, the position of these two planes can be determined from the coefficients of the Gauss-matrix of the system. The derivation of the corresponding formulae has been given by Professor Kross in Ref. /22/, here we just reproduce the final results.
We denote the object plane by \( P \) and the paraxial image plane by \( Q \). The media between these planes and the end surfaces of the system are considered to be homogeneous, the refractive indices being denoted by \( n_\text{P} \) and \( n_\text{Q} \). The distances from these planes to the corresponding end surfaces are then given by

\[
\beta = d_\text{P} = \frac{n_\text{P}}{G_{12}} (G_{11} - \beta^{-1}), \quad \beta = d_\text{Q} = \frac{n_\text{Q}}{G_{12}} (G_{22} - \beta) \tag{2.5.2}
\]

Let us now show that the two planes whose positions are given by Eqs. (2.5.2) are indeed conjugated to each other. Consider a paraxial ray starting from an object of size \( l_\text{P} \) with an inclination \( \omega_\text{P} \). The equations (2.5.2) then give the "axial thicknesses" that must be substituted in the homogeneous transfer matrix (2.3.30) in order to obtain the matrices \( U_\text{P} \) and \( U_\text{Q} \) describing the ray transfer between the end surfaces of the system and the planes \( P \) and \( Q \). By denoting at the latter plane the ray height and inclination by \( l_\text{Q} \) and \( \omega_\text{Q} \) we have

\[
\begin{pmatrix} n_\text{Q} \omega_\text{Q} \\ l_\text{Q} \end{pmatrix} = U_\text{Q} U_\text{P} \begin{pmatrix} n_\text{P} \omega_\text{P} \\ l_\text{P} \end{pmatrix}
\]

By performing the matrix multiplications we find

\[
U_\text{Q} U_\text{P} = \begin{pmatrix} \beta^{-1} & G_{12} \\ 0 & \beta \end{pmatrix} \tag{2.5.3}
\]

It follows immediately that

\[
\begin{align*}
  n_\text{Q} \omega_\text{Q} &= \beta^{-1} n_\text{P} \omega_\text{P} + G_{12} l_\text{P} \\
  l_\text{Q} &= \beta l_\text{P} \tag{2.5.4}
\end{align*}
\]

The second of Eqs. (2.5.4) shows that the ray height at the plane \( Q \) does not depend on \( \omega_\text{P} \), i.e. all paraxial rays from the object point pass through the corresponding image point. Also, the transverse magnification is the same for all points of the plane \( P \). Thus, for each plane \( P \) there exists a conjugate plane \( Q \) at the position given by Eqs. (2.5.2).

The first of Eqs. (2.5.4) can be written also as

\[
\beta n_\text{Q} \omega_\text{Q} = n_\text{P} \omega_\text{P} + G_{12} l_\text{Q}
\]

Consider now the special cases \( \beta=0 \) (object at infinity) and \( \beta=\infty \) (image at infinity). It follows from the first of Eqs. (2.5.4) and its above equivalent that we have

\[
\frac{1}{G_{12}} = \frac{l_\text{Q}}{n_\text{P} \omega_\text{P}} \tag{2.5.5.a}
\]

in the first case and

\[
\frac{1}{G_{12}} = \frac{l_\text{P}}{n_\text{Q} \omega_\text{Q}} \tag{2.5.5.b}
\]

in the second case. The equations (2.5.5.a-b) are definitions for the focal length of the system. (See /22/.) Thus, the power and focal length of the system are given simply by
\[ \phi = f^{-1} = G_{12} \]  

\[ \text{(2.5.6)} \]

### 2.5.2 The nature of the paraxial approximation

For rays propagating through the optical system close to the optical axis, i.e. when \( x, y, \xi, \eta \) are small, the paraxial approximation holds. As shown previously, at each refraction or transfer, only the linear terms have to be considered in this case in the expressions for finite rays giving the final values of the ray data as functions of the initial ones. However, for a ray undergoing a succession of refractions and transfers, as in the case of ray-tracing throughout the system, a more detailed analysis of the paraxial approximation for the quantities \( x, y, \xi, \eta \) is necessary, because for deriving aberration coefficients a careful distinction between terms of various orders must be made. (Recall that we denote the paraxial approximation for \( x, y, \xi, \eta \) and for any function of these variables by a tilde \( \sim \) over the corresponding quantity.)

We denote the marginal and chief ray heights at any surface by \( h \) and \( m \) and the corresponding marginal and chief ray inclinations by \( u \) and \( w \) (notation in use at the Technical University of Berlin / 22/, 23/), as shown in Fig. 2.5.1.

Consider now an arbitrary ray in an rotationally-symmetric optical system consisting of homogeneous and GRIN lenses. Let \((\tau_x, \tau_y)\) be the normalized Cartesian ray coordinates in the object plane and \((\sigma_x, \sigma_y)\) those of the intersection point of the ray with the entrance pupil plane. These coordinates are normalized to unity such that the ray coordinates of the marginal ray are \( \sigma_x = 1, \sigma_y = \tau_x = \tau_y = 0 \) and those of the chief ray are \( \tau_x = 1, \sigma_x = \sigma_y = \tau_y = 0 \). Thus, if \( r_{EP} \) is the radius of the entrance pupil and \( r_p \) is the maximum object height, the Cartesian coordinates at the object plane are given by

\[ x_p = r_p \tau_x, \quad y_p = r_p \tau_y \]  

\[ \text{(2.5.7)} \]

and at the entrance pupil plane by

\[ x_{EP} = r_{EP} \sigma_x, \quad y_{EP} = r_{EP} \sigma_y \]  

\[ \text{(2.5.8)} \]
If linearization is performed also in the opening formulae where the ray position and direction are given through \( (\sigma_x, \sigma_y) \) and \( (\tau_x, \tau_y) \), then \( \bar{x}, \bar{y}, \bar{z}, \bar{n} \) are given at each surface of the system by linear combinations of the aperture and field coordinates. Let us now show that the coefficients are the height and slope of the marginal and chief ray at that surface

\[
\begin{align*}
\bar{x} &= m \tau_x + h \sigma_x, \quad \bar{z} = -n_0 w \tau_x - n_0 u \sigma_x, \\
\bar{y} &= m \tau_y + h \sigma_y, \quad \bar{n} = -n_0 w \tau_y - n_0 u \sigma_y.
\end{align*}
\]  

(2.5.9)

For radial gradients, \( n_0 \) is the refractive index on the optical axis. For axial gradients, \( n_0 \) is the refractive index at the surface vertex. For homogeneous media, the index 0 can be dropped. Recall that the sign convention adopted for the angles \( u \) and \( w \) is that they are of opposite sign to the corresponding direction cosines.

First, we show that Eqs.(2.5.9) hold everywhere in the medium between object and the first surface of the system. As known from analytical geometry, the equation of a straight line in the plane \( xz \) passing through the points \( P_0(x_0,z_0) \) and \( P_1(x_1,z_1) \) is

\[
\frac{x-x_0}{z-z_0} = \frac{x_1-x_0}{z_1-z_0} = \frac{a_x}{a_z}
\]

where \( a_x \) and \( a_z \) are the corresponding direction cosines. Let \( P_0 \) and \( P_1 \) be the intersection points of the ray with the object plane and entrance pupil plane and let the distance between these two planes be denoted by \( d_{EP} \). Using Eqs. (2.5.7-8) and noting that \( \tilde{a}_z = 1 \) we obtain

\[
\begin{align*}
\bar{x} - r_p \tau_x &= \frac{r_{EP} \sigma_x - r_p \tau_x}{d_{EP}} \\
n_p \tilde{a}_z &= n_p \frac{r_{EP} \sigma_x - r_p \tau_x}{d_{EP}}
\end{align*}
\]

and after rearrangement
\begin{align}
\tilde{\chi} &= \sigma_x \frac{z r_{EP}}{d_{EP}} + \tau_x r_{EP} \left(1 - \frac{z}{d_{EP}}\right) \\
\tilde{\zeta} &= \sigma_x \frac{n_p r_{EP}}{d_{EP}} + \tau_x \frac{-n_p r_{EP}}{d_{EP}}.
\end{align}

(2.5.10)

By replacing \( x \) by \( y \) and \( \xi \) by \( \eta \) similar relations are obtained for the ray projection in the \( yz \) plane.

For the marginal and chief rays we have at the object plane \( h=0 \) and \( m=r_p \). At the entrance pupil plane, using the transfer equations yields

\[ h = r_{EP} = -u d_{EP}, \quad m = 0 = r_p - w d_{EP} \]

and therefrom we find that between the object plane and the first surface of the system we have

\[ u = -\frac{r_{EP}}{d_{EP}}, \quad w = \frac{r_p}{d_{EP}}, \quad h = \frac{z r_{EP}}{d_{EP}}, \quad m = r_p \left(1 - \frac{z}{d_{EP}}\right). \]

(2.5.11)

Inserting Eqs. (2.5.11) into Eqs. (2.5.10) gives Eqs. (2.5.9), which thus hold at the first surface of the system, prior to refraction. Now, since all refractions and transfers through the system are described in the paraxial approximation by linear transformations, it follows by mathematical induction that Eqs. (2.5.9) hold also after each of these transformations, i.e. that these equations are valid throughout the optical system. By denoting symbolically by \( G_X \) any of the Gauss matrices for refraction or transfer derived previously (see Eqs. (2.1.29), (2.2.23), (2.3.25), (2.3.30)), we have

\[ \begin{pmatrix} -\frac{\tilde{\eta}}{\tilde{\chi}} \\ \tilde{\zeta} \end{pmatrix} = G_X \begin{pmatrix} -\frac{\tilde{\eta}}{\tilde{\chi}} \\ \tilde{\zeta} \end{pmatrix} = G_X \left[ \sigma_x \left( \begin{array}{c} n_0 u \\ h \end{array} \right) + \tau_x \left( \begin{array}{c} n_0 w \\ m \end{array} \right) \right] = \sigma_x \left( n_0' u' \\ h' \right) + \tau_x \left( n_0' w' \\ m' \right) \]

i.e. the validity of Eqs. (2.5.9) prior to the transformation implies validity also after the transformation.

We note here that \( x = \tilde{x} + O(3) \). The same holds for the other ray parameters. For any integer \( p \), in this study we denote by \( O(p) \) any quantity containing only terms of total order \( \geq p \) in \( (\sigma_x, \sigma_y) \) and \( (\tau_x, \tau_y) \).

The equations (2.5.9) show that for tracing any paraxial ray throughout the system, it suffices to trace only two rays, the marginal ray and the chief ray. By using the matrix formalism introduced previously, tracing these two rays reduces to a succession of matrix multiplications. Paraxial ray tracing gives a good approximation to the actual ray path of a given ray if for that ray \( (\sigma_x, \sigma_y) \) and \( (\tau_x, \tau_y) \) are small.
As has been pointed out by H.H. Hopkins /18/, Eqs. (2.5.9) enable to remove a widespread ambiguity about the nature of the paraxial approximation. The marginal and chief rays are commonly traced using paraxial relations, despite the fact that their height and inclination are not infinitesimal, but finite quantities. Actually, by expanding $x,y,\xi,\eta$ everywhere within the system with respect to $(\sigma_x,\sigma_y)$ and $(\tau_x,\tau_y)$, it can be seen from Eqs. (2.5.9) that the ray data of the paraxially traced marginal and chief rays are the coefficients of the linear terms of the expansions. Thus the customary paraxial refraction and transfer formulae are exact relations between these coefficients and are independent on the magnitude of the latter.

Consider now Eqs. (2.5.9) at the exit pupil plane and paraxial image plane. Denoting the radius of the paraxial image of the entrance pupil, i.e. the radius of the exit pupil, if pupil aberrations are neglected, by $r_{AP}$ and the maximum paraxial image height by $r_Q$, we have at the exit pupil plane $m=0$ and $h=r_{AP}$, i.e.

$$\tilde{x}_{AP} = r_{AP} \sigma_x, \tilde{y}_{AP} = r_{AP} \sigma_y$$

(2.5.12)

and at the image plane $h=0$ and $m=r_Q$, i.e.

$$\tilde{x}_Q = r_Q \tau_x, \tilde{y}_Q = r_Q \tau_y$$

(2.5.13)

The normalized ray coordinates which have been defined in object space (at the object plane and entrance pupil plane) can be defined equally well in image space i.e. at the exit pupil plane and paraxial image plane. By comparing Eqs. (2.5.12-13) with Eqs. (2.5.7-8) we find that for paraxial rays the two definitions lead to identical values of the normalized ray coordinates. For finite rays the difference between normalized coordinates in object and image space is therefore a quantity of $O(3)$.

In Chapter 3, the Seidel and chromatic paraxial aberration coefficients will be found to be the coefficients of the lowest order terms in the expansions of monochromatic and chromatic aberrations with respect to the normalized ray coordinates. Therefore, for deriving these coefficients for gradient index lenses, the difference between the two definitions of normalized ray coordinates is irrelevant. (Differences appear only in the case of higher order aberrations.) In what follows we will use the notations $(\sigma_x,\sigma_y)$ and $(\tau_x,\tau_y)$ in both cases. Sometimes, at the entrance or exit pupils, spherical surfaces are used instead of planes to define $\sigma_x,\sigma_y$, see e.g. /18/. If only lowest order aberrations are considered, the resulting difference can also be neglected.
3. Aberrations of gradient-index lenses

In this chapter we derive short and accurate analytic formulae for the Seidel coefficients and chromatic paraxial aberration coefficients (i.e., the primary aberration coefficients) of gradient-index lenses with arbitrary axial or radial refractive index distributions. Starting from the analytic ray-tracing formulae of the previous chapter, we develop a technique for decomposing the two components of the transverse aberration of an arbitrary skew ray in surface and inhomogeneous transfer contributions. For axial gradients, this technique leads directly to the primary aberration coefficients. For radial gradients, an additional technique for shortening the large expressions for the terms resulting from the transfer contributions is developed. Unlike previously known derivation methods, applying our method delivers simple algebraic expressions for all primary coefficients of radial GRIN lenses.

In the case of homogeneous optical systems, primary aberration coefficients are known for a long time to be a valuable tool for lens design. Even if the primary aberrations yield a description of the image quality which is strictly valid only for modest aperture and angular coverage, the primary coefficients are useful to analyse the effects of the change of various lens parameters on the aberrations of the system and to locate adequate starting values of the system data for subsequent optimization with ray-tracing. Only optical systems where both Seidel and chromatic aberrations can be reduced to acceptable values are capable of producing an image of good quality. Thus, controlling these aberrations is a necessary (but not sufficient) condition for obtaining successful optical designs.

For gradient-index lenses, attempts to develop convenient methods for the calculation of the Seidel coefficients have been made ever since the potential of the gradients to improve the performances of optical systems has been recognized /46/, /47/, /36/, /30/, /58/. Sands has shown for instance that, for general rotationally symmetric gradients, in addition to the ordinary surface contributions, the Seidel coefficients of gradient-index lenses consist of inhomogeneous surface contributions at both end faces and of inhomogeneous transfer contributions /46/. For axial gradients, the expressions of the Seidel coefficients have been obtained by the same author /47/. For radial gradients however, while the inhomogeneous surface contributions are given by simple aspheric-like terms (see Eqs. (3.3.35)), the methods for obtaining numerical values for the transfer contributions described in these references are highly complex, a simple expression being published only for the Petzval curvature /36/.
In Ref. /7/ we have developed a method to derive Seidel aberration coefficients for radial GRIN lenses directly from ray-tracing formulae and have obtained simple formulae for the remaining transfer contributions as well. This derivation method is discussed in detail in this chapter. To illustrate how it works, we first derive the Seidel coefficients in the more simple cases of the surface contributions (axial and radial gradients) and of the transfer contributions of axial gradients. In these cases, we obtain the same results as Sands, thus providing an independent confirmation of these formulae which have considerable practical utility.

Expressions for the chromatic paraxial aberration coefficients of axial and radial GRIN lenses have been published also by Sands /48/, /49/. While in the axial case we can confirm Sand's results, in the radial case we have to correct a computation error of Sands which, although mathematically trivial, prejudices the practical usefulness of the results (see Sec 3.6.4).

Previously, an elegant Lagrangian method suggested by Buchdahl /10/ has been employed for the derivation of aberration coefficients of GRIN lenses /46/, /48/, /13/. However, it is generally known that, for deriving aberration coefficients, a large amount of symbolic calculation is necessary and Buchdahl's method suffers from the practical drawback that eventual computation errors during the calculation (as that mentioned above) are extremely difficult to be detected and localized precisely. Therefore, we have preferred to develop a new derivation method based on ray-tracing, because it enables a direct control of the intermediate steps of the calculation by comparison with numerical ray-tracing data and thus provides additional safety. We have been encouraged in this attempt by the fact that, for homogeneous lenses, one of the existing methods for deriving aberration coefficients is also based on ray-tracing /12/.

After a general discussion in Sec. 3.1 of the properties of primary aberrations, the rest of this chapter is devoted to the derivation of aberration coefficients. Relying on the algebraic formulae for tracing rays in GRIN lenses developed in the previous chapter, we first describe a technique enabling the decomposition of the two components of the monochromatic transverse aberration of an arbitrary skew ray in contributions from refraction at each surface and transfer through each inhomogeneous medium of the system (Sec.3.2.). From the resulting algebraic expressions the Seidel coefficients are obtained by collecting the coefficients of all terms having third degree in aperture and field coordinates. In the case of the surface contributions (axial and radial gradients, Sec. 3.3) and of the transfer contributions of axial gradients (Sec 3.4) this technique leads directly to the Seidel coefficients. For radial gradients, however, at this stage the Seidel coeffi-
cients are given by very large expressions. We show that these expressions can be substantially shortened by using a heuristic symmetrization technique such that finally simple closed formulae are obtained for all Seidel coefficients of the radial GRIN lenses, as in the case of homogeneous and axial GRIN lenses (Sec.3.5). To prove the correctness of the Seidel formulae for radial GRIN lenses, an independent test by comparison with ray-tracing is performed. A similar method with that for the derivation of Seidel coefficients is then developed for deriving the chromatic paraxial aberration coefficients of axial and radial GRIN lenses (Sec. 3.6).

3.1 Primary aberrations of rotationally- symmetric optical systems

Before starting the derivation of the Seidel and chromatic coefficients of GRIN lenses, let us first recall some general properties of the primary aberrations of rotationally symmetric systems, which are well known from the study of homogeneous lenses, but which are valid as well for GRIN lenses. Since some of these properties are often derived using methods valid only for homogeneous media, we derive them below in the general context.

3.1.1 Types of primary aberrations

The types of aberrations which can occur in an optical system are determined by the symmetry of the system. A general treatment of this subject can be done using Hamiltonian methods /11/. However, for discussing only the lowest order monochromatic and chromatic aberrations, we can start from the more familiar concept of wave aberration. In what follows, we will derive aberration coefficients from the decomposition of the lowest order monochromatic and chromatic transverse aberrations in surface and transfer contributions. In this section, the types of terms which can occur in these decompositions and the physical significance of the terms will be determined using the relationship between the wave aberration and the transverse ray aberration.

Consider a finite ray through the rotationally - symmetric system. The components of the transverse aberration vector of the ray at the paraxial image plane are defined, as usual, by

\[ \Xi_x = x_0 - \bar{x}_0, \Xi_y = y_0 - \bar{y}_0 \]

(3.1.1)

Consider now the wavefront through the centre of the exit pupil, associated with the image-forming pencil and consider also a reference sphere, having is centre at the paraxial image and passing also through the centre of the exit pupil. The optical path length W
along the segment of the given finite ray situated between the wavefront and the reference sphere is called wave aberration. The two components of the transverse aberration vector of the ray can be expressed in terms of the wave aberration as

$$\Xi_x = -\frac{R_{AP}}{n_O} \frac{\partial W}{\partial x_{AP}}, \quad \Xi_y = -\frac{R_{AP}}{n_O} \frac{\partial W}{\partial y_{AP}},$$

(3.1.2)

where $R_{AP}$ is the radius of the reference sphere and $x_{AP}, y_{AP}$ are the Cartesian coordinates in the exit pupil plane. (For more details see e.g. Refs. /24/, /59/,/8/)

If only the lowest order aberrations are considered, $R_{AP}$ can be replaced in Eq.(3.1.2) by the distance $d_{AP}$ between the exit pupil plane and the paraxial image plane and the Cartesian coordinates can be replaced by their paraxial approximation given by Eq. (2.5.12). Thus we have

$$W(x_{AP}, y_{AP}, \sigma_x, \sigma_y, \tau_x, \tau_y) = -\frac{1}{n_O d_{AP}} \frac{\partial W}{\partial \sigma_x} \sigma_x - \frac{1}{n_O d_{AP}} \frac{\partial W}{\partial \sigma_y} \sigma_y - \frac{1}{n_O d_{AP}} \frac{\partial W}{\partial \tau_x} \tau_x - \frac{1}{n_O d_{AP}} \frac{\partial W}{\partial \tau_y} \tau_y$$

(3.1.3)

The transfer equation for the marginal ray between pupil and image plane yields

$$r_{AP} - u_0 d_{AP} = 0$$

(3.1.4)

Consequently, Eqs. (3.1.2) can be written as

$$\Xi_x = -\frac{1}{n_O u_0} \frac{\partial W}{\partial \sigma_x}, \quad \Xi_y = -\frac{1}{n_O u_0} \frac{\partial W}{\partial \sigma_y}$$

(3.1.5)

Consider now the series expansion of $W(\sigma_x, \sigma_y, \tau_x, \tau_y)$ with respect to the normalized aperture and field coordinates. For determining what kinds of terms can appear in this expansion we take advantage of the rotational symmetry of the system. Thus, $W$ must be left unchanged by any rigid rotation of the ray about the symmetry axis and therefore any terms of the expansion containing combinations of aperture and field coordinates which are not invariant at rotation must vanish.

Note first that, when the ray rotates, the length of the vectors $s = (\sigma_x, \sigma_y)$ and $t = (\tau_x, \tau_y)$ and the angle between them remain unchanged. Thus, for each ray, we have three rotational invariants: $\sigma^2$, $\tau^2$ and the scalar product $\sigma \tau$. Consequently, in the series expansion for $W$ the aperture and field coordinates can appear only through these three invariants. Up to the fourth order, we have

$$W(s,t) = W_0 + a_3 s^2 + a_5 s^4 + a_7 s^6 + a_3 t^2 +$$

$$+ b_3 s^4 + b_5 s^2 t + b_3 (st)^2 + b_5 s^4 t^2 + b_3 s t^4 + b_6 t^4 + O(6)$$

(3.1.6)

Consider first the case of monochromatic imaging. From the definition of the wave aberration, where we have assumed that both the wavefront and the reference sphere are chosen to pass through the centre of the exit pupil, it follows that for $\sigma = 0$ (i.e. for rays passing through the centre of the exit pupil) we must have $W = 0$, i.e.
irrespective of the value of $\tau$. Therefore we must have

$$W_0 = a_3 = b_6 = 0.$$  \hspace{1cm} (3.1.7)

As can be seen from Eqs. (3.1.5), the two remaining second-order terms would produce contributions to the transverse aberration vector which are linear in aperture and field coordinates. However, since the aberrations are measured taking as reference the paraxial image point, any linear contribution must vanish, i.e.

$$a_i = 0$$  \hspace{1cm} (3.1.8)

Thus, the five fourth order terms with coefficients $b_1$ to $b_5$ are the lowest order monochromatic aberration terms in the wave aberration expansion. For practical purposes, a different notation for these coefficients is more convenient. We rewrite the fourth-order contribution to the wave aberration as

$$W_4(\sigma_x, \sigma_y, \tau_x, \tau_y) = \frac{1}{8} \Gamma_1(\sigma_x^2 + \sigma_y^2) + \frac{1}{2} \Gamma_2(\sigma_x^2 + \sigma_y^2)(\sigma_x \tau_x + \sigma_y \tau_y) +$$

$$+ \frac{1}{2} \Gamma_3(\sigma_x \tau_x + \sigma_y \tau_y)^2 + \frac{1}{4} (\Gamma_3 + P)(\sigma_x^2 + \sigma_y^2)(\tau_x^2 + \tau_y^2) + \frac{1}{2} \Gamma_4(\sigma_x \tau_x + \sigma_y \tau_y)(\tau_x^2 + \tau_y^2)$$  \hspace{1cm} (3.1.9)

The coefficients $\Gamma_i$, $i=1,2,3,4$ and $P$ are the total Seidel aberration coefficients of the system. The physical significance of the coefficients for spherical aberration, $\Gamma_1$, coma, $\Gamma_2$, astigmatism, $\Gamma_3$, Petzval curvature, $P$, and distortion, $\Gamma_4$, can be found e.g. in Ref. /8/.

Substituting Eq. (3.1.9) in Eq. (3.1.5) we find that the two components $\Xi_{x,3}$, $\Xi_{y,3}$ of the third order transverse aberration of the ray are related to the total Seidel aberration coefficients of the system by

$$-2n_0 u_0 \Xi_{x,3} = \left[ \Gamma_1(\sigma_x^2 + \sigma_y^2) + 2\Gamma_2(\sigma_x \tau_x + \sigma_y \tau_y) + (\Gamma_3 + P)(\tau_x^2 + \tau_y^2) \right] \sigma_x +$$

$$+ \left[ \Gamma_2(\sigma_x^2 + \sigma_y^2) + 2\Gamma_3(\sigma_x \tau_x + \sigma_y \tau_y) + \Gamma_4(\tau_x^2 + \tau_y^2) \right] \tau_x$$

$$-2n_0 u_0 \Xi_{y,3} = \left[ \Gamma_1(\sigma_x^2 + \sigma_y^2) + 2\Gamma_2(\sigma_x \tau_x + \sigma_y \tau_y) + (\Gamma_3 + P)(\tau_x^2 + \tau_y^2) \right] \sigma_y +$$

$$+ \left[ \Gamma_2(\sigma_x^2 + \sigma_y^2) + 2\Gamma_3(\sigma_x \tau_x + \sigma_y \tau_y) + \Gamma_4(\tau_x^2 + \tau_y^2) \right] \tau_y$$  \hspace{1cm} (3.1.10)

As has been first shown by Sands /46/, for an optical system consisting of homogeneous and GRIN lenses, the total Seidel coefficients are obtained by summing up the corresponding ordinary contributions $S_p$, $P_S$ and inhomogeneous surface contributions $S_p^*$ over all surfaces and the inhomogeneous transfer contributions $T_p$, $P_T$ over all gradient media.

$$\Gamma_p = \sum_{\text{surfaces}} (S_p + S_p^*) + \sum_{\text{GRIN/media}} T_p, \ p = 1,2,3,4$$

$$P = \sum_{\text{surfaces}} P_S + \sum_{\text{GRIN/media}} P_T$$  \hspace{1cm} (3.1.11)
It will be seen later (see Eq. (3.2.21)) that the decompositions (3.1.11) follow directly from the properties of the so-called quasi-invariant, which will be defined in Secs. 3.2.1-2. The expressions for the various surface and transfer contributions will be derived in Secs. 3.3, 3.4 and 3.5.

Let us now consider the dispersion effects in an optical system designed for use over a given range of wavelengths. Because of the variation with wavelength of the refractive index parameters determining the paraxial properties of the system, the exact position (and size) of the paraxial image also depends on the wavelength. Thus, even if the quadratic terms in Eq. (3.1.6) can be annulled for a given reference wavelength $\lambda_0$ by properly shifting the image plane, they do not vanish over the entire wavelength range. Consequently, the lowest order terms of the chromatic wave aberration are those of second order, and, according to Eq. (3.1.5), the resulting terms in the transverse aberrations are linear in aperture and field coordinates, i.e., they are of paraxial nature.

By defining the chromatic paraxial aberration vector of a ray at the wavelength $\lambda$ as

$$\Xi_{\lambda_0,1} = \tilde{x}_Q(\lambda) - \tilde{x}_Q(\lambda_0), \Xi_{\lambda_0,1} = \tilde{y}_Q(\lambda) - \tilde{y}_Q(\lambda_0)$$

where $Q$ is the image plane for $\lambda_0$, we can write

$$-n_0 u_0 \Xi_{\lambda_0,1} = \Gamma_1 \sigma_s + \Gamma_2 \tau_s, -n_0 u_0 \Xi_{\lambda_0,1} = \Gamma_1 \sigma_y + \Gamma_2 \tau_y.$$  

The chromatic paraxial coefficients of the system $\Gamma_{\lambda_1}$ and $\Gamma_{\lambda_2}$ are called the total axial color and lateral color coefficients. Their name comes from the fact that they are related to the axial displacement of focus and to the chromatic change in image height. However, in Eq. (3.1.13), they both contribute to the transverse chromatic aberration. As in the case of the total Seidel coefficients, the total chromatic coefficients can be written as sums over surfaces and GRIN media

$$\Gamma_{\lambda_p} = \sum_{\text{surfaces}} S_{\lambda_p} + \sum_{\text{GRIN media}} T_{\lambda_p}, \ p = 1,2$$

The expressions for the surface and transfer contributions of GRIN lenses to the total chromatic coefficients will be derived in Sec. 3.6.

For practical applications, the two chromatic coefficients are usually considered together with the five Seidel coefficients as constituting a group of primary aberration coefficients. All expressions for primary aberration coefficients will be summarized in Appendix B.

Finally, let us consider the special case when the optical system, in addition to the rotational symmetry, is also symmetric with respect to the stop and is designed for use at $\beta=-1$. We now show that the additional symmetry produces a further limitation of the possible aberration types. In this case, for any ray, the lowest order monochromatic and
chromatic terms in Eq. (3.1.6) must be left unchanged if the object and paraxial image plane are interchanged.

Let us first note that, because of symmetry, the transverse magnification between the entrance and exit pupil plane is equal to unity. This is because this quantity is the product of the two mutually reciprocal transverse magnifications corresponding to the imaging of the stop through the identical halves of the system. Consequently the two pupil planes are identical with the principal planes.

Consider now the case when the object and image plane are interchanged. If the differences of O(3) between normalized coordinates in object and image space are neglected (as it can be done if the lowest order aberrations are considered only) then the lengths of the vectors $\sigma$ and $\tau$ remain unchanged. However, as can be seen from Fig. 3.1.1 the vector $\sigma$ keeps its direction, while $\tau$ has reversed its direction. Thus, Eq. (3.1.6) can contain only terms in $\sigma^2$ and $\tau^2$, i.e. the coefficients $a_2$, $b_2$ and $b_5$ of the terms containing $\sigma \tau$ must vanish. Consequently, we have

$$\Gamma_{42} = \Gamma_2 = \Gamma_4 = 0$$

i.e. the symmetric system is free of lateral color, coma and distortion.

Thus, the symmetrical system has for $\beta = -1$ only four primary aberrations: spherical aberration, astigmatism, Petzval curvature and axial color. For each of these aberrations, the corresponding coefficient is twice the value for each half system. The half system values can be determined, for instance, by considering the second half of the system as an independent optical system having the object at infinity.
3.1.2 Stop-shift formulae

The formulae indicating how the various primary aberration coefficients depend on the stop position (i.e. the stop-shift formulae) are of considerable importance in practical optical design. In this section we show that these formulae can be derived from general requirements, irrespective of the concrete nature of the system. Thus, for optical systems containing GRIN lenses, these formulae are precisely the same as for homogeneous optical systems.

Suppose that the stop position is changed such that the H invariant (object height \( x \) numerical aperture) is left unchanged. For keeping the numerical aperture constant, the diameter of the stop must also be changed at stop shift such that the marginal ray path remains unchanged through the optical system.

In order to find the changes in the primary coefficients due to a stop shift, we first note that the Cartesian ray coordinates at any arbitrary surface within the system are obviously left unchanged by the stop shift. This holds also for their paraxial approximation given by Eq. (2.5.9)

\[
\begin{align*}
\hat{x} &= m \tau_x + h \sigma_x \\
\hat{y} &= m \tau_y + h \sigma_y 
\end{align*}
\]

In these equations, \( h \) and \( \tau \) are not affected by the stop shift. Let us first see how the normalized aperture coordinates are affected by this shift. We take the position of the stop prior to the shift as reference and denote the variation of a given quantity due to the stop shift by \( \delta \) before that quantity. After and prior to the shift, we have

\[
h(s + \delta h) + (m + \delta m)t = hs + mt
\]

and hence

\[
\delta h = -t \delta m
\]  \hspace{1cm} (3.1.16)

where the abbreviation

\[
\delta \theta = \frac{\delta m}{h}
\]  \hspace{1cm} (3.1.17)

has been used. Thus, Eqs. (3.1.16-17) give the relation between the change in the aperture coordinates and the change in the chief ray height at an arbitrary surface when the stop position is changed. It follows from the above derivation that the value of \( \delta m/h \) is the same at each plane perpendicular to the symmetry axis.

The stop shift equations can be now obtained from the condition that the total monochromatic and chromatic transverse aberrations of the system given by Eqs. (3.1.10) and (3.1.13), which are expressed through aperture and field coordinates and primary aber-
ration coefficients, are left unchanged by the stop shift. Consider first the chromatic aberrations (3.1.13). We have
\[
(\Gamma_{a1} + \delta \Gamma_{a1})(s + \delta s) + (\Gamma_{a2} + \delta \Gamma_{a2})t = \Gamma_{a1}s + \Gamma_{a2}t \tag{3.1.18}
\]
or, after substituting Eq. (3.1.16) and rearranging the terms,
\[
\delta \Gamma_{a1}s + (-\Gamma_{a1}\vartheta - \delta \Gamma_{a1}\vartheta + \delta \Gamma_{a2})t = 0 \tag{3.1.19}
\]
Since Eq. (3.1.19) holds for arbitrary vectors \( \sigma \) and \( \tau \), we obtain the stop-shift equations for chromatic paraxial coefficients as
\[
\delta \Gamma_{a1} = 0 \tag{3.1.20}
\]
\[
\delta \Gamma_{a2} = \vartheta \Gamma_{a1}
\]

The derivation of the stop-shift equations for the Seidel coefficients can be done using the same principle. For simplifying the derivation, consider a meridional ray having \( \sigma_y = \tau_y = 0 \). Thus, the first of Eqs. (3.1.10) becomes
\[
-2n\vartheta u_0 \Xi_{x3} = \Gamma_{x3} + 2\Gamma_{x2}\sigma_x^2 \tau_x + (3\Gamma_3 + P)\sigma_x \tau_x^2 + \Gamma_x \tau_x^3 \tag{3.1.21}
\]
It will be shown in Secs. 3.3, 3.4 and 3.5 that the surface and transfer contributions to the Petzval curvature \( P \) depend only on the system data and are therefore independent of the stop position. (See Eqs. (3.3.40), (3.4.14), (3.5.16).) Thus we have
\[
\delta P = 0 \tag{3.1.22}
\]
As in the case of Eq. (3.1.18), requiring that Eq. (3.1.21) should be left unchanged by the stop shift, we can write an equation containing the four unknowns \( \delta \Gamma_p, p=1...4 \). Since the algebraic manipulations are simple, but lengthy, the author has performed the computations using computer algebra (Mathematica). After using Eq. (3.1.16) and rearranging the terms, the requirement that the coefficients of \( \sigma_x^3, \sigma_x^2 \tau_x, \sigma_x \tau_x^2 \) and \( \tau_x^3 \) vanish leads to four linear equations in the four unknowns \( \delta \Gamma_p, p=1...4 \). The solution was found to be
\[
\delta \Gamma_1 = 0
\]
\[
\delta \Gamma_2 = \vartheta \Gamma_1
\]
\[
\delta \Gamma_3 = 2\vartheta \Gamma_1 + \vartheta^2 \Gamma_1
\]
\[
\delta \Gamma_4 = \vartheta(3\Gamma_3 + P) + 3\vartheta^2 \Gamma_2 + \vartheta^3 \Gamma_1 \tag{3.1.23}
\]
The equations (3.1.20), (3.1.23) give the change of the primary aberration coefficients of a rotationally symmetric system as functions of the stop shift parameter \( \vartheta \) given by Eq.(3.1.17).

In initial stages of lens design, Eqs.(3.1.20), (3.1.23) are used for finding the value of \( \vartheta \) necessary for correcting certain primary aberrations. Let us now determine the relation between \( \vartheta \) and the corresponding value of the axial displacement of the stop position \( \delta z_s \). Therefore, we have to evaluate \( \delta m/h \) at the plane of the stop prior to the shift as shown in Fig. 3.1.2.
We denote the stop radius prior to the shift by $r_s$, the refractive index in the stop space (assumed to be constant) by $n_s$ and the inclination of the marginal ray in this space by $u_s$. Prior to the shift, we have at the stop plane $AE$ $m=0$ and $h=r_s$. After the shift we have at the old stop plane the height of $MF$, the new chief ray, $ME = \delta m$. At the new stop plane $BF$, the height of the marginal ray $AB$ is

$$BF = r_s - u_s \delta z_s,$$

while the height of the new chief ray is

$$\delta m - (w + \delta v) \delta z_s = 0 \quad (3.1.24)$$

Since the stop shift keeps $H$ unchanged, we have at the plane $BF$

$$H = -n_s (r_s - u_s \delta z_s) (w + \delta v). \quad (3.1.25)$$

From Eqs. (3.1.17) and (3.1.24) we have

$$\vartheta = \frac{\delta m}{h} = \frac{(w + \delta v) \delta z_s}{r_s}$$

Using (3.1.25), we finally obtain

$$\vartheta = -\frac{H \delta z_s}{n_s r_s (r_s - u_s \delta z_s)} \quad (3.1.26)$$

Let us finally note that in the general case of rotationally symmetric optical systems, the stop-shift equations (3.1.20), (3.1.22), (3.1.23) and (3.1.26) are precisely the same as in the special case of homogeneous media. (See e.g. Ref. /59/ where these equations have been derived using special properties of the homogeneous media.)
3.2 Decomposition of the transverse aberration in surface and transfer contributions

3.2.1 The quasi-invariant

As a first step in the determination of Seidel coefficients, we derive a technique which, given the ray position and direction at each surface, enables the decomposition of the transverse aberration vector of an arbitrary finite ray (Eqs. (3.1.1)) in contributions from refraction at the surfaces and from transfer through the inhomogeneous media of the system. This technique is based on the idea of the quasi-invariant, first introduced by Buchdahl /10/. Since the direct use of ray-tracing formulae requires several changes to Buchdahl's original definition, the idea of the quasi-invariant will be discussed in detail below.

Consider a finite ray through the system. In the paraxial approximation, a system invariant similar to $H$ (given by Eq. (2.2.24)) can be formed with the marginal ray and the projection on the $xz$ plane of the chosen finite ray. At any surface of the system, this invariant can be written as

$$\bar{\xi} = n_p \bar{\xi} + h \bar{\xi}. \quad (3.2.1)$$

(The plus sign appears in Eq. (3.2.1) because of the sign convention used. See also Eqs. (2.1.27).) Recall that we denote the object plane by $P$ and the paraxial image plane by $Q$ and that the media at these two planes are considered to be homogeneous. We now denote

$$\Lambda_{px} = n_p u_p x_p, \Lambda_{xQ} = n_Q u_Q x_Q \quad (3.2.2)$$

Due to the invariance of Eq. (3.2.1) at the object and image plane we have

$$\bar{\Lambda}_{xQ} = \bar{\Lambda}_{xp} = \Lambda_{xp} \quad (3.2.3)$$

Therefore we can write

$$n_Q u_Q (x_Q - \bar{x}_Q) = n_Q u_Q x_Q - n_Q u_Q \bar{x}_Q = n_Q u_Q x_Q - n_p u_p x_p \quad (3.2.4)$$

and

$$\Xi_x = \frac{\Lambda_{xQ} - \Lambda_{xp}}{n_Q u_Q} \quad (3.2.5)$$

The required decomposition of Eq. (3.2.5) can be accomplished by defining a quantity $\Lambda_x$ having Eq. (3.2.1) as its paraxial approximation. Since the paraxial approximation of $\Lambda_x$ is an invariant, $\Lambda_x$ itself will be called, following Buchdahl, a quasi-invariant. We seek $\Lambda_x$ as

$$\Lambda_x = F(n_q u) x + f(h) f(\xi) \quad (3.2.6)$$

where the unknown functions $F$ and $f$ have still to be determined. These functions will depend also on the coordinates of the finite ray, such that, according to Eq. (3.2.1),

$$\bar{F}(n_q u) = n_q u, \bar{F}(h) = h, \bar{f}(\xi) = \bar{\xi} \quad (3.2.7)$$
Consider now the variation of $\Lambda_x$ given by Eq.(3.2.6) at each refraction or transfer in the system starting from the object plane and ending with the image plane and sum up all the resulting terms. We have

$$\Lambda_{xO} - \Lambda_{xp} = \sum \Delta \Lambda_x$$  \hspace{1cm} (3.2.8)

As known from aberration theory, transfer through homogeneous media does not contribute to the aberration coefficients. Therefore the unknown functions $F$ and $f$ can be determined by requesting that $\Delta \Lambda_x$ should vanish at transfer through a homogeneous medium.

Consider first the case of the transfer between two planes. It follows from Eqs. (3.2.2) that at the object and paraxial image plane we must have $F(nu)=nu$, so we require that

$$F(nu) = n_0 u, F(h) = h$$  \hspace{1cm} (3.2.9)

at any plane surface. It can be easily verified that the transfer contributions vanish for

$$f(\xi) = \frac{n}{\zeta} \xi$$  \hspace{1cm} (3.2.10)

where $n$ is the refractive index of the homogeneous medium. If the planes are separated by the distance $z$, we have

$$\Delta x = \frac{\xi}{\zeta} z, \Delta h = -uz$$  \hspace{1cm} (3.2.11)

and therefore

$$\Delta \Lambda_x = nu \Delta x + \frac{n_0 \xi}{\zeta} \Delta h = 0$$  \hspace{1cm} (3.2.12)

The generalization of Eq.(3.2.10) for inhomogeneous media is not unique. However, it seems natural to simply replace the constant refractive index by a varying one. For rotationally-symmetric GRIN media we thus have

$$f(\xi) = \frac{n(r^2, z)}{\zeta} \xi$$  \hspace{1cm} (3.2.13)

For radial GRIN media, $n(r^2)$ is given by Eq.(2.3.6).
3.2.2 Curved end surfaces

Let now a homogeneous or GRIN medium be situated between two curved lens surfaces. For a given finite ray consider the planes A and A' perpendicular to the symmetry axis Oz and passing through the intersection points of the ray with the first and second surface. Consider the points \( \bar{M} \) and \( \bar{M}' \) situated at the intersection of the marginal ray OM in the corresponding medium (or of its prolongation) with the planes A and A' defined by the finite ray EF. (See Fig.3.2.1 for the definition of \( \bar{M} \) at the first surface of the system.)

Fig. 3.2.1: Definition of the barred marginal ray data. In this figure the medium after the first surface of the system was assumed to be homogeneous.

Let the corresponding values of the inclination and height of the marginal ray at \( \bar{M} \) and \( \bar{M}' \) be denoted by \( \bar{u}, \bar{h} \) and \( \bar{u}'', \bar{h}' \). Both for homogeneous and GRIN media the barred marginal ray data can be obtained from the unbarred ones by replacing in the paraxial transfer equations for the corresponding medium, Eqs. (2.2.22), (2.3.24) and (2.3.30), \( d \) by \( \delta z \) at the first surface and by \( \delta z' \) at the second one. Recall that \( \delta z \) and \( \delta z' \) are the distances between the planes A and A' and surface vertices. At the first surface \( \delta z \) is given in the lowest order of approximation by

\[
\delta z = \frac{1}{2} \rho (x^2 + y^2) + O(4)
\]  

(3.2.14)

and a similar relation holds at the second surface. (See e.g. Eq. (2.4.43)) For instance, in the homogeneous case we obtain

\[
\bar{u} = u, \bar{h} = h - u\delta z, \bar{u}' = u', \bar{h}' = h' - u'\delta z'
\]

(3.2.15)

Since according to Eq. (3.2.14) \( \delta z \) and \( \delta z' \) are quantities of \( O(2) \), the barred quantities differ from the unbarred also by quantities of \( O(2) \).
Note that the transfer equations for the barred marginal ray data are the same as those for the unbarred, where \( d \) is replaced by the distance \( z \) between the planes \( A \) and \( A' \) given by

\[
z = d + \delta z' - \delta z
\]  

(3.2.16)

For homogeneous media, by replacing the unbarred marginal ray data by barred ones, Eq.(3.2.12) also obviously holds for the transfer between the planes \( A \) and \( A' \). We now define the quantity \( \bar{n} \) as the paraxial approximation \( \bar{n}(\bar{M}) \) for the refractive index at the point \( \bar{M} \) i.e.

\[
\bar{n} = n_0 \text{ for radial gradients} \\
\bar{n} = n(\delta \bar{z}) \text{ for axial gradients}
\]

(3.2.17)

Therefore, by choosing

\[
F(n_0 \mu) = \bar{n} \bar{u}, F(h) = \bar{h}
\]

we obtain at each surface of the system, both for homogeneous and GRIN media, a definition of the quasi-invariant having all the necessary properties for the derivation of aberration coefficients

\[
\Lambda_x = \bar{n} \bar{u} x + \frac{\bar{h} n(r^2, \delta \bar{z}) \bar{z}}{\zeta}.
\]

(3.2.18)

Because of the rotational symmetry, \( x \) and \( \bar{\zeta} \) are odd functions in aperture and field coordinates while \( n(r^2, z), \bar{\zeta}, \bar{n}, \bar{u} \) and \( \bar{h} \) are even functions. Therefore Eq.(3.2.18) is an odd function. Since at each refraction or transfer the first order terms vanish because of \( \Delta \bar{\Lambda}_x = 0 \), the lowest order terms in \( \Delta \Lambda_x \) are those of third order, as required by the decomposition of \( \Xi_x \).

With the above definition of \( \Lambda_x \), the sum in Eq.(3.2.8) has to be considered only over surfaces and GRIN media

\[
\Lambda_{sQ} - \Lambda_{sP} = \sum_{\text{surfaces}} \Delta \Lambda_x + \sum_{\text{GRIN media}} \Delta \Lambda_x
\]

(3.2.19)

Substituting Eq. (3.2.19) into (3.2.5) yields

\[
n_0 \mu_0 \Xi_x = \sum_{\text{surfaces}} \Delta \Lambda_x + \sum_{\text{GRIN media}} \Delta \Lambda_x.
\]

(3.2.20)

Assume now that the surface and transfer contributions in Eq. (3.2.20) are expressed in aperture and field coordinates. At each refraction or transfer, denote the sum of the third order terms by \( (\Delta \Lambda_x)_3 \). Compare Eq.(3.2.20) with the first of Eqs. (3.1.10). We obtain
It can be seen that Eq. (3.2.21) leads to the decomposition of the total Seidel coefficients of the system in surface and transfer contributions. (See Eq. (3.1.11).) In Secs. 3.3, 3.4 and 3.5, the various Seidel coefficients for refraction and transfer appearing in Eq. (3.1.11) will be determined from the refraction and transfer contributions to \((\Delta \Lambda_x)_3\).

Eqs. (3.2.18-19) and (3.2.5) provide for any finite ray the decomposition of \(\Xi_x\) in surface and inhomogeneous transfer contributions. By changing \(x\) and \(\xi\) to \(y\) and \(\eta\), similar formulae can be obtained for \(\Xi_y\).

Besides providing a way to determine Seidel coefficients, the decomposition of \(\Xi_x\) and \(\Xi_y\) can also be used to analyse higher order aberration contributions of surfaces and GRIN media to the total aberrations of given finite rays. It can be easily seen that the numerical results of the decomposition of \(\Xi_x\) and \(\Xi_y\) do not actually depend on the initial inclination of the "marginal" ray. This is because in Eq. (3.2.18) both \(\widetilde{\Pi}\) and \(\widetilde{H}\) (at any surface) and \(n_n u_n\) are proportional to \(n_P u_P\). Consequently, both sides of Eq. (3.2.20) are proportional to \(n_P u_P\), which thus cancels. Therefore in this case any paraxially calculated ray starting on the optical axis from the basis of the object could be equally well employed as "marginal" ray. Since we have

\[
\frac{n_P u_P}{n_Q u_Q} = \beta
\]

a convenient value for the initial inclination of the marginal ray is \(n_P u_P = \beta\), which leads to \(n_Q u_Q = 1\). Thus, tracing paraxially this ray together with the given finite ray through the system and evaluating \(\Delta \Lambda_x\) and \(\Delta \Lambda_y\) according to Eq. (3.2.18) at each refraction or inhomogeneous transfer yields the required decomposition of the transverse aberration vector of the given finite ray. For homogeneous media a similar method is known as the Aldis Theorem /59/, /12/.
We now determine the surface contributions of GRIN lenses with arbitrary axial and radial refractive index distributions to the Seidel aberration coefficients of the optical system. We show that these contributions consist of two parts: the ordinary surface contributions, which are precisely the same as for homogeneous lenses, and the inhomogeneous surface contributions, which, for GRIN lenses with spherical end surfaces, are mathematically equivalent to the contributions of an aspheric surface.

### 3.3.1 Derivation of the Seidel coefficients

Since the following calculations are somewhat lengthy, let us first briefly describe, as a guideline for their understanding, the main ideas of the derivation of the Seidel coefficients. The derivation procedure for the transfer contributions (Sec. 3.4 for axial gradients and Sec. 3.5 for radial gradients) is essentially the same as that for deriving the surface contributions.

Consider the variation of $\Lambda_\chi$ at refraction at a surface. As mentioned before, the lowest order terms of $\Delta \Lambda_\chi$ in aperture and field coordinates are those of third order. We denote their sum by $(\Delta \Lambda_\chi)_3$. For finding the Seidel coefficients we have to determine $(\Delta \Lambda_\chi)_3$. For that we start by separating in $\Lambda_\chi$ and $\Lambda_\chi'$ the terms of first and third order in the ray parameters $x,y,\xi,\eta$ (i.e. the ray parameters prior to refraction) from the higher order terms. It will turn out that the variation of the first order terms vanishes. Thus, the lowest order terms of $\Delta \Lambda_\chi$ in $x,y,\xi,\eta$ are the third order terms.

Since $x,y,\xi,\eta$ differ from their paraxial approximations only by third order quantities in aperture and field coordinates, $(\Delta \Lambda_\chi)_3$ is obtained by replacing in the third order terms of $\Delta \Lambda_\chi$ in $x,y,\xi,\eta$ the ray parameters $x,y,\xi,\eta$ by their paraxial approximations and then express the paraxial approximations for $x,y,\xi,\eta$ through aperture and field coordinates using Eqs. (2.5.9).

The Seidel contributions at refraction are then obtained from $(\Delta \Lambda_\chi)_3$ from the coefficients of the various third-order terms in aperture and field coordinates. At refraction, it turns out that $(\Delta \Lambda_\chi)_3$ can be written in a form which is completely symmetric in data prior and after refraction. This feature considerably simplifies the finding of a short form for the expressions of Seidel coefficients.
After the same derivation steps as those named above, in Sec. 3.4 a symmetric form for \((\Delta \Lambda_x)_{ij}\) will be found also in the case of transfer through axial gradients. In the case of radial gradients, however, such a symmetric form cannot be obtained directly. Therefore, for obtaining short expressions for the Seidel coefficients, an a posteriori symmetrization of the raw expressions for the coefficients must be performed. This additional derivation step produces simple expressions for the Seidel coefficients also in the case of radial GRIN lenses.

### 3.3.2 Third-order terms at refraction

Consider the refraction of a finite ray at a spherical surface and assume for generality that the media on both sides of the surface are gradient media. As described above, we start by separating in all quantities that appear in \(\Lambda_x\) and \(\Lambda_x'\) the terms of various orders in the ray parameters \(x, y, \xi, \eta\), up to the third order. (Recall that quantities prior to refraction are denoted by unprimed symbols and quantities after refraction by primed ones.) We denote the sum of the terms of a given order in \(x, y, \xi, \eta\) of a quantity by a corresponding subscript to the symbol of that quantity.

Both for axial and for radial gradients, we write the refractive index at the intersection point of the finite ray with the surface, prior to refraction, as

\[
n = n_0 + n_2 + O(4)
\]

(3.3.1)

For radial gradients, we obtain from Eq. (2.3.6)

\[
n = n_0 \left(1 - \frac{1}{2} kr^2\right) + O(4)
\]

(3.3.2)

and therefrom

\[
n_2 = -\frac{1}{2} n_0 k (x^2 + y^2)
\]

(3.3.3)

For axial gradients, we assume that in the neighbourhood of the surface vertex the refractive index distribution \(n=n(z)\) can be expanded into a power series with respect to \(z\). By denoting

\[
N_z = \frac{dn}{dz}\bigg|_{z=0}
\]

(3.3.4)

we can write generally

\[
n(\delta z) = n_0 + N_z \delta\xi + O(4)
\]

(3.3.5)

where \(\delta z\) is given by Eq. (3.2.14)

\[
\delta\xi = \frac{1}{2} \rho (x^2 + y^2) + O(4)
\]

Thus,
For the barred marginal ray data we have
\[ \overline{\pi} = n_0 \mu + \frac{d}{dz} \bigg|_{z=0} \frac{\partial \varepsilon}{\partial \mu} + O(4) = n_0 \mu + (\overline{\pi})_2 + O(4) \]  
\[ \overline{h} = h + \frac{dh}{dz} \bigg|_{z=0} \frac{\partial \varepsilon}{\partial \mu} + O(4) = h + \overline{h}_2 + O(4) \]  

For radial gradients, we obtain from Eqs. (2.3.24)
\[ \frac{d}{dz} \bigg|_{z=0} (\pi \mu) = - n_0 g \sin g z + n_0 g^2 h \cos g z \bigg|_{z=0} = n_0 g^2 h = n_0 kh \]
\[ \frac{dh}{dz} \bigg|_{z=0} = - g h \sin g z - g \frac{n_0 \mu}{n_0 g} \cos g z \bigg|_{z=0} = -u \]
and consequently
\[ (\overline{\pi})_2 = \frac{1}{2} \rho n_0 kh \left( x^2 + y^2 \right) \]
\[ \overline{h}_2 = - \frac{1}{2} \rho u \left( x^2 + y^2 \right) \]  

For axial gradients, it follows from Eqs. (2.2.22) that
\[ (\overline{\pi})_2 = 0 \]
\[ \overline{h}_2 = - \frac{1}{2} \rho u \left( x^2 + y^2 \right) \]  

Note that the first of Eqs. (3.3.9) is a consequence of the invariance at transfer through axial gradients of the product \( n u \). It can be seen that the second of the Eqs. (3.3.9) is identical with the second of the Eqs. (3.3.8) for radial gradients. Eqs. (3.3.9) hold also for homogeneous media.

Similarly, from Eq. (2.1.10) we obtain
\[ \frac{n}{\zeta} = \frac{n}{\sqrt{n^2 - \zeta^2 - \eta^2}} = 1 + \frac{1}{2n_0} \left( \zeta^2 + \eta^2 \right) + O(4) = 1 + \left( \frac{n}{\zeta} \right)_2 + O(4) \]  

Thus, prior to refraction, the quasi-invariant \( \Lambda_x \) can be written as
\[ \Lambda_x = \overline{\pi} \mu x + \overline{h} \left( \frac{n}{\zeta} \right) \xi = (n_0 \mu + (\overline{\pi})_2) x + (h + \overline{h}_2) \left( 1 + \left( \frac{n}{\zeta} \right)_2 \right) \xi + O(5) = \]
\[ = n_0 \mu x + h \xi + (\overline{\pi})_2 x + \overline{h}_2 \xi + h \left( \frac{n}{\zeta} \right)_2 \xi + O(5) \]  

\[ \text{(3.3.11)} \]
Similar equations can be written for the quantities after refraction. In addition, the ray parameters after refraction must also be expanded as power series with respect to the parameters prior to refraction:

\[
x' = x' + O(5), y' = y' + O(5), \xi' = \xi' + O(5), \eta' = \eta' + O(5).
\] (3.3.12)

In order to determine the quantities appearing in Eq. (3.3.12), we start from the refraction equations for a spherical surface (2.1.25-26). Let us denote for the moment:

\[
\eta = n_0 x + \xi J_0 - \rho x \xi
\] (3.3.13)

\[
J = \sqrt{n_0^2 - \xi^2 - \eta^2}
\] (3.3.14)

Thus, the x-component of the vectorial refraction equation (2.1.25) now reads:

\[
\xi' = \xi - \rho x J
\] (3.3.15)

A similar equation is valid for the y-component. Recall that at refraction we have \(x'=x\) and \(y'=y\).

For \(\theta\) we can write:

\[
\theta = \theta_0 + \theta_2 + O(4)
\] (3.3.16)

where \(\theta_2\) splits into a homogeneous and a gradient part:

\[
\theta_2 = \theta_{2h} + \theta_{2g}
\] (3.3.17)

After some elementary calculations, it follows from Eq. (3.3.13) that:

\[
\theta_0 = n_0
\]

\[
\theta_{2g} = n_2
\] (3.3.18)

\[
\theta_{2h} = -\frac{1}{2n_0} \left[ (n_0 \rho x + \xi)^2 + (n_0 \rho y + \eta)^2 \right]
\]

For \(J\) we also have:

\[
J = J_0 + J_2 + O(4)
\] (3.3.19)

and from Eqs. (3.3.14) and (3.3.16-18) we find after some algebra that:

\[
J_0 = \Delta n_0
\]

\[
J_2 = \Delta n_2 - \theta_{2h} \frac{\Delta n_0}{n_0}
\] (3.3.20)

From Eq. (3.3.15) and (3.3.19) it can be seen that:

\[
\xi' = \xi - \rho x J_0
\]

\[
\xi' = -\rho x J_2
\] (3.3.21)

Thus, after refraction, the quasi-invariant \(\Lambda'_x\) can be written as
\[ \Lambda' = \pi' \pi x + \bar{h} \left( \frac{n'}{\xi} \right)' \xi' = (n_o' u' + (\pi' \pi') x + (h + \bar{h} \xi) \left( 1 + \left( \frac{n'}{\xi} \right) \right) \xi' + \xi + O(5) = \]
\[ = n_o' u' x + h \xi' + (\pi' \pi') x + \bar{h} \xi + h \xi + O(5) \]  

(3.3.22)

The variation at refraction of \( \Lambda_x \) is obtained by subtracting Eq. (3.3.11) from Eq. (3.3.22). First, we note that the variation of the first order terms of \( \Lambda_x \) vanish. Indeed, using Eqs. (2.1.28) and (3.3.20-21) we find that

\[ n_o' u' x + h \xi' = x \Delta(n_o, u) + h(\xi' - \xi) = x h \rho \Delta n_0 + h(\xi' - \xi) = 0 \]  

(3.3.23)

Thus, the lowest order terms in \( \Delta \Lambda_x \) in \( x, y, \xi, \eta \) are those of third order

\[ \Lambda' - \Lambda_x = x \Delta(n \pi) + \bar{h} \xi + h \xi + O(5) \]  

(3.3.24)

Since Eq.(3.3.24) is already of third order in \( x, y, \xi, \eta \), \( \Delta \Lambda_x \) is obtained as follows: In each quantity depending on \( x, y, \xi, \eta \), we replace \( x, y, \xi, \eta \) by their paraxial approximations which, according to Eqs. (2.5.9), , are linear combinations of aperture and field coordinates. Substituting the various quantities appearing in Eq. (3.3.24) according to the formulae derived in this section, we find

\[ \Delta \Lambda_x = \frac{1}{2} \rho (\bar{x}^2 + \bar{y}^2) \Delta u \bar{\xi} - \frac{1}{2} h \Delta \left( \frac{\bar{\xi}(\xi'^2 + \eta^2)}{n_o^2} \right) - h \rho \bar{x} \left( \partial n_2 - \partial \Delta n_0 \right)n_0 \]  

(3.3.25)

We now separate in \( \Delta \Lambda_x \) the terms containing gradient contributions from the purely homogeneous ones.

\[ \Delta \Lambda_x = \frac{1}{2} \rho (\bar{x}^2 + \bar{y}^2) \Delta u \bar{\xi} + \frac{1}{2} h \Delta \left( \frac{\bar{\xi}(\xi'^2 + \eta^2)}{n_o^2} \right) + \frac{1}{2} h \rho \bar{x} \left( n_0 \rho \bar{\bar{\xi}} + \bar{\xi} \right) + \left( n_0 \rho \bar{\bar{\xi}} + \bar{\xi} \right)^2 \]  

(3.3.26)

Using Eqs. (3.3.18) it follows that the homogeneous part of \( \Delta \Lambda_x \) is

\[ \Delta \Lambda_x = \frac{1}{2} \rho (\bar{x}^2 + \bar{y}^2) \Delta u \bar{\xi} + \frac{1}{2} h \Delta \left( \frac{\bar{\xi}(\xi'^2 + \eta^2)}{n_o^2} \right) + \frac{1}{2} h \rho \bar{x} \left( n_0 \rho \bar{\bar{\xi}} + \bar{\xi} \right) + \left( n_0 \rho \bar{\bar{\xi}} + \bar{\xi} \right)^2 \]  

(3.3.27)

The gradient part of \( \Delta \Lambda_x \) is

\[ \Delta \Lambda_x = \frac{1}{2} \rho (\bar{x}^2 + \bar{y}^2) \Delta u \bar{\xi} + \frac{1}{2} h \Delta \left( \frac{\bar{\xi}(\xi'^2 + \eta^2)}{n_o^2} \right) + \frac{1}{2} h \rho \bar{x} \left( n_0 \rho \bar{\bar{\xi}} + \bar{\xi} \right) + \left( n_0 \rho \bar{\bar{\xi}} + \bar{\xi} \right)^2 \]  

(3.3.28)

For radial gradients, substituting Eqs. (3.3.3) and (3.3.8) yields

\[ \Delta \Lambda_x = \frac{1}{2} \rho (\bar{x}^2 + \bar{y}^2) \Delta u \bar{\xi} - \rho \bar{x} \Delta n_2 \]  

(3.3.29)

For axial gradients, it follows from Eqs. (3.3.6) and (3.3.9) that

\[ \Delta \Lambda_x = -\frac{1}{2} \rho (\bar{x}^2 + \bar{y}^2) \rho^2 h \Delta N_2 \]  

(3.3.30)

Eq. (3.3.27) will lead in the next section to the ordinary surface contributions, while Eqs. (3.3.29-30) will lead to the inhomogeneous surface contributions of axial and radial GRIN lenses to the Seidel coefficients. We note that in Eq.(3.3.27), in addition to \( \bar{x}, \bar{y}, h, \)
the expressions \( n_0 \rho \tilde{\eta} + \zeta \) and \( n_0 \rho \tilde{\eta} + \tilde{\eta} \) are also quantities which do not change at refraction. This can be seen by inserting the first of Eqs. (3.3.20) into the paraxial approximation for Eq. (3.3.15). Thus, in Eqs. (3.3.27) and (3.3.29-30) we have a complete symmetry between quantities before and after refraction. This symmetry considerably simplifies the derivation of short formulae for the Seidel coefficients.

### 3.3.3 Ordinary and inhomogeneous surface contributions

We now replace in Eqs. (3.3.27) and (3.3.29-30) the paraxial ray parameters \( \tilde{x}, \tilde{y}, \tilde{\xi}, \tilde{\eta} \) by their expressions in aperture and field coordinates \( \sigma_x, \sigma_y, \tau_x, \tau_y \) using Eqs. (2.5.9). It can be seen from these formulae that the result for the homogeneous part must be of the form

\[
(\Delta \Lambda_x)_{h} = \left[ c_1 \left( \sigma_x^2 + \sigma_y^2 \right) + c_2 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + c_3 \left( \tau_x^2 + \tau_y^2 \right) \right] \sigma_x + \\
+ \left[ c_4 \left( \sigma_x^2 + \sigma_y^2 \right) + c_5 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + c_6 \left( \tau_x^2 + \tau_y^2 \right) \right] \tau_x
\]

(3.3.31)

It will be shown explicitly below that the gradient part also has the same structure.

It was noted in Sec. 3.2.1 that the definition of \( \Lambda_x \) is not unique. All definitions can lead to the Seidel aberration coefficients, but the derivation is more direct if the decomposition of \( \Lambda_x \) is such that all third-order contributions at refraction and transfer have the same structure as their sum given by Eq.(3.2.21). Inserting Eqs. (3.1.11) into Eq. (3.2.21) we find that for the ordinary surface contributions this condition can be written as

\[
-2(\Delta \Lambda_x)_{h} = \left[ S_1 \left( \sigma_x^2 + \sigma_y^2 \right) + 2S_2 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + S_3 + P_3 \left( \tau_x^2 + \tau_y^2 \right) \right] \sigma_x + \\
+ \left[ S_2 \left( \sigma_x^2 + \sigma_y^2 \right) + 2S_3 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + S_4 \left( \tau_x^2 + \tau_y^2 \right) \right] \tau_x
\]

(3.3.32)

and for the inhomogeneous surface contributions as

\[
-2(\Delta \Lambda_x)_{g} = \left[ S'_1 \left( \sigma_x^2 + \sigma_y^2 \right) + 2S'_2 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + S'_3 \left( \tau_x^2 + \tau_y^2 \right) \right] \sigma_x + \\
+ \left[ S'_2 \left( \sigma_x^2 + \sigma_y^2 \right) + 2S'_3 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + S'_4 \left( \tau_x^2 + \tau_y^2 \right) \right] \tau_x
\]

(3.3.33)

This is not the case for the original definition of \( \Lambda_x \) given in Ref. / 10/. Therefore, in Ref. / 46/ terms in the surface contributions of \( \Delta \Lambda_x \) that deviate from Eqs.(3.3.32-33) are cancelled by corresponding terms in the transfer contributions.

With our definition of the quasi-invariant (Eq. (3.2.18)), the conditions (3.3.32-33) are satisfied. Consider first the gradient parts of \( (\Delta \Lambda_x)_3 \) given by Eqs.(3.3.29-30). In both cases, the dependence on aperture and field coordinates is given by
By comparing Eqs. (3.3.33) and (3.3.34), it follows from Eqs. (3.3.29-30) that the inhomogeneous surface contributions to the Seidel aberration coefficients are

\[
\begin{align*}
S_1' &= -2h^3 \rho \Delta (n_k) \\
S_2' &= -2h^3 m \rho \Delta (n_k) \\
S_3' &= -2h^2 m^2 \rho \Delta (n_k) \\
S_4' &= -2h m^3 \rho \Delta (n_k)
\end{align*}
\]  

(3.3.35)

for radial gradients and

\[
\begin{align*}
S_1' &= h^3 \rho^2 \Delta N_z \\
S_2' &= h m \rho^2 \Delta N_z \\
S_3' &= h^2 m^2 \rho^2 \Delta N_z \\
S_4' &= h m^3 \rho^2 \Delta N_z
\end{align*}
\]  

(3.3.36)

for axial gradients. Eqs (3.3.35-36) have been first obtained by Sands /46/. It can be seen from these equations that, both for axial and for radial gradients, the dependence of the inhomogeneous surface contributions on the heights of the marginal and chief rays is precisely the same as for an homogeneous aspherical surface.

Consider now the ordinary surface contributions. If condition (3.3.32) is satisfied, then only five of the six coefficients of Eq.(3.3.31) are linearly independent. Otherwise, the deviating terms would cause all six coefficients to be independent. By explicitly determining the six coefficients as described above, we have found that \(c_2=2c_4\). Thus Eq. (3.3.32) is satisfied.

Raw expressions for the ordinary surface contributions to the Seidel aberration coefficients result from comparing the coefficients of Eqs. (3.3.31) and (3.3.32)

\[
\begin{align*}
S_1 &= -2c_1, S_2 = -2c_4 = -c_2, S_3 = -c_5, P_S = -2c_1 + c_5, S_4 = -2c_6
\end{align*}
\]  

(3.3.37)

With the notation introduced in Sec. 3.1.1, \(S_1\) stands for the spherical aberration, \(S_2\) for coma, \(S_3\) for astigmatism, \(P_S\) for Petzval curvature, and \(S_4\) for distortion. The starred symbols in Eqs.(3.3.35-36) have the same meaning.

Since by its definition Eq. (3.3.27) does not contain gradient contributions, we expect that the expressions (3.3.37) are precisely the same as those derived for homogeneous lenses. By performing the calculations in detail we find that this is indeed the case.
Since the expressions of the ordinary surface contributions are well known, we give below only the derivation of the expression for spherical aberration. The coefficient $c_1$ can be simply obtained by putting in Eqs. (2.5.9) $\sigma_x=1, \sigma_y=\tau_x=\tau_y=0$, i.e. by substituting in Eq. (3.3.27) $\vec{x}=h\vec{\xi}=-n_0u$. We obtain

$$c_1 = -\frac{1}{2} \rho h^2 \Delta (-n_0u^2) + \frac{1}{2} h \Delta (-n_0u^2) + \frac{1}{2} h^2 \rho (n_0 \rho h - n_0u)^2 \Delta \left( \frac{1}{n_0} \right)$$

(3.3.38)

Since according to Eq. (2.1.30) the quantity $n_0 \rho h - n_0 u = n_0i$ is a paraxial refraction invariant, Eq. (3.3.38) can be rearranged as

$$c_1 = \frac{1}{2} h \Delta (\rho h n_0 u^2 - n_0 u^3 + \rho h n_0 i^2) = \frac{1}{2} h n_0 i \Delta (u^2 + \rho i) = \frac{1}{2} h n_0 i [\Delta (-u i) + h \rho \Delta (u + i)] = -\frac{1}{2} h (n_0 i)^2 \Delta \left( \frac{u}{n_0} \right)$$

Writing for the chief ray the paraxial refraction invariant as

$$n_0 i = n_0 m \rho - n_0 w$$

(3.3.39)

the remaining coefficients can be rearranged similarly. (We have used therefore computer algebra software.) Finally, substituting the results into Eq. (3.3.37) we find

$$S_1 = (n_0 i)^2 h \Delta (u/n_0)$$

$$S_2 = n_0 i n_0 j h \Delta (u/n_0)$$

$$S_3 = (n_0 j)^2 h \Delta (u/n_0)$$

$$P_8 = -\rho H^2 \Delta (1/n_0)$$

$$S_4 = (n_0 j)^2 m \Delta (u/n_0) + n_0 j H \Delta (w/n_0)$$

(3.3.40)

where $H$ is given by Eq. (2.2.24).

$$H = mn_0 u - h n_0 w$$

The incidence angles for the marginal and chief rays $i$ and $j$ appearing in the refraction invariants are shown in Fig. (2.5.1). (See Sec.2.5.2.)

Thus, the ordinary and inhomogeneous surface contributions of GRIN lenses to the Seidel coefficients can be calculated from the paraxial marginal and chief ray data at the lens surfaces and the paraxial refraction invariants. In Sec.3.5, for radial GRIN lenses, the validity of all Seidel formulae, including Eqs. (3.3.35) and (3.3.40) will be verified on the basis of an independent numerical test.
3.4 Transfer contributions of axial gradient-index lenses to the Seidel coefficients

3.4.1 Third-order terms at transfer

For deriving the transfer contributions of GRIN lenses to the Seidel coefficients, we follow, as in the case of the surface contributions, the guideline described in Sec 3.3.1. In this section we collect from the variation of $\Lambda_\chi$ at transfer through an axial or radial GRIN medium the third order terms in $(\sigma_x, \sigma_y)$ and $(\tau_x, \tau_y)$, denoting their sum by $(\Delta \Lambda_\chi)_3$

Let us first separate in the expressions of the quantities appearing in the quasi-invariant (Eq. (3.2.18)) the terms of first and third order in $x, y, \xi, \eta$ from the higher order terms. For $x', y', \xi', \eta'$, i.e. the ray parameters at the second surface after transfer, we write, as in Sec. 3.3.2

$$x' = x_0' + O(5), y' = y_0' + O(5), \xi' = \xi_0' + O(5), \eta' = \eta_0' + O(5).$$

The explicit form of the various first and third-order quantities will be determined in the next sections from the ray-tracing formulae derived in Sec. 2.2 for axial gradients and in Sec. 2.4 for radial gradients. At this point we note only that in the ray-tracing formulae, $x', y', \xi', \eta'$ are functions of $x, y, \xi, \eta$ and of the propagation distance inside the gradient, $z$, given by Eq. (3.2.16)

It will be seen later that, in order to simplify the derivation of $(\Delta \Lambda_\chi)_3$, in the following calculations we can separate the terms of various orders in a way that differs from an usual power series expansion in $x, y, \xi, \eta$ in two respects:

1. We disregard the dependence of the position of the planes A and A' on $x, y, \xi, \eta$ as given by Eq. (3.2.16) and keep therefore $z$ and the barred quantities $\overline{nu}, \overline{nh}$ constant.

2. For radial gradients, we regard $\mu$ (see Eq. (2.3.15)) as a quantity containing only second order terms in $x, y, \xi, \eta$ and neglect the remaining terms of $O(4)$.

Writing prior to transfer

$$\frac{n}{\xi} = \left(\frac{n}{\xi}\right)_0 + \left(\frac{n}{\xi}\right)_2 + O(4)$$

we obtain as in the case of Eq. (3.3.10)

$$\frac{n}{\xi} = \frac{n}{\sqrt{n^2 - \xi^2 - \eta^2}} = \frac{1}{\sqrt{1 - \frac{1}{n_0^2} (\xi_0^2 + \eta_0^2)}} = 1 + \frac{1}{2n_0} (\xi_0^2 + \eta_0^2) + O(4)$$

Recall that for radial gradients $n_0$ is the quantity appearing in the refractive index distribution (2.3.6), i.e. $n_0 = n_0'$, while for axial gradients we have $n_0 = n(0)$ and $n_0' = n(d)$. 
Before transfer the series expansion up to the third order in \(x, y, \xi, \eta\) of Eq.(3.2.18) is given by

\[
\Delta_x = \overline{u}x + \overline{h} \xi + \frac{1}{2n_0} \overline{h} \xi \left(\xi^2 + \eta^2\right) + O(5)
\]  

(3.4.1)

and after transfer by

\[
\Delta'_x = \overline{\nu}'u'(x'_1 + x'_j) + \overline{h}' \left(\left(\frac{n'}{\xi}\right)_0 + \left(\frac{n'}{\xi}\right)_2\right)(\xi_{j'} + \xi'_j) + O(5) = \\
= \overline{\nu}'u'x'_1 + \overline{h}' \xi'_1 + \overline{\nu}'u'x'_j + \overline{h}' \xi'_j + \frac{1}{2n_0'} \overline{h}' \xi' \left(\xi'^2 + \eta'^2\right) + O(5).
\]  

(3.4.2)

Subtract now Eq. (3.4.1) from Eq. (3.4.2). As in Sec. 3.3.2. the first order contributions vanish because we have

\[
\overline{\nu}u + \overline{h} \xi = \overline{\nu}u'x'_1 + \overline{h}' \xi'_1
\]  

(3.4.3)

Eq. (3.4.3) can be verified both for axial and for radial gradients by direct substitution of the transfer equations for the primed quantities. (See Secs. 3.4.2 and 3.5.1 for the equations for \(x', \xi', \eta', \xi_1', \eta_1'\).) However, a closer analysis shows directly that Eq.(3.4.3) holds for any rotationally symmetric gradient. Indeed, it follows from the definitions of the various quantities appearing in Eq. (3.4.3) that \(-\xi'_1\) and \(x'_1\) are related to \(-\xi_1\) and \(x\) by precisely the same linear transformation as that relating \(\overline{h}u', \overline{h}'\) to \(\overline{\nu}u, \overline{h}\). This transformation is described by the Gauss matrix for the transfer through the corresponding medium, where \(d\) is replaced by \(z\) given by Eq.(3.2.16). Thus, Eq. (3.4.3) has the same origin as the invariant \(H\) (see Sec. 2.2.3). Consequently, we have

\[
\Delta'_x - \Delta_x = \overline{\nu}'u'x'_3 + \overline{h}' \xi'_3 + \frac{1}{2n_0'} \overline{h}' \xi'_3 \left(\xi'^2 + \eta'^2\right) - \frac{1}{2n_0} \overline{h} \xi \left(\xi^2 + \eta^2\right) + O(5)
\]  

(3.4.4)

Since Eq.(3.4.4) is already of third order in \(x, y, \xi, \eta\), \((\Delta \Delta_x)_3\) is obtained as follows: In each quantity depending on \(x, y, \xi, \eta\) we keep only the lowest order terms in these variables and then replace \(x, y, \xi, \eta\) by their paraxial approximation (2.5.9). (Recall that a quantity thus modified is denoted by a tilde over it.) Thus, the barred marginal ray data are replaced by the unbarred, \(\overline{\nu}\) by \(n_0\), and \(x, y, \xi, \eta\) by \(z\). (This is why we could keep \(z\) and the barred quantities constant in the above derivation. The \(O(2)\) contributions stemming from these quantities cancel out because of Eq. (3.4.3).) Thus, we obtain from Eq. (3.4.4)

\[
(\Delta \Delta_x)_3 = n'_0u'x'_3 + h' \xi'_3 + \Delta \left(\frac{1}{2n_0} h \xi \left(\xi^2 + \eta^2\right)\right)
\]  

(3.4.5)
3.4.2 Axial transfer contributions

We now determine the various quantities appearing in Eq. (3.4.5) in the case of axial gradients. The invariance at transfer of the optical direction cosines $\xi$ and $\eta$ (see Eqs. (2.1.16)) leads to

$$\xi' = \xi, \eta' = \eta$$

and therefore we have

$$\xi' = \xi, \quad \eta' = \eta, \quad \xi' = 0.$$  \hspace{1cm} (3.4.6)

Similarly, for the marginal and chief rays we have the transfer invariants

$$n'u' = n_0u, \quad n'_w = n_0w.$$  \hspace{1cm} (3.4.7)

From Eqs. (2.2.4) and (2.2.8-9) we find

$$x' = x + \xi Z_1(z) + \frac{1}{2} (\xi^2 + \eta^2)Z_3(z) + O(5)$$  \hspace{1cm} (3.4.8)

which leads to

$$x'_\xi = x + Z_1(z)\xi, \quad x'_\eta = \frac{1}{2}Z_3(z)(\xi^2 + \eta^2).$$  \hspace{1cm} (3.4.9)

In Eq. (3.4.5) the paraxial approximation for the second of Eqs. (3.4.9) is required, where $z$ is replaced by $d$, as mentioned above. As for $Z_1$ (see Eqs. (2.2.21)), for $Z_3$ we define

$$Z_3(d) = \left\langle n^{-3} \right\rangle d^3n_0^{-3} \left( n_0u\xi d + \Delta \left( \frac{h}{n_0^2} \right) \right).$$  \hspace{1cm} (3.4.10)

Thus, for axial gradients Eq. (3.4.5) reads

$$\Delta \Lambda_x = \frac{1}{2} \xi \left( \xi^2 + \eta^2 \right) + \left(n_0u\xi d + \Delta \left( \frac{h}{n_0^2} \right) \right).$$  \hspace{1cm} (3.4.11)

Note that in Eq. (3.4.11) we have a complete symmetry between quantities before and after transfer.

We now express Eq. (3.4.11) through aperture and field coordinates. Using Eqs. (2.5.9) we obtain

$$\xi (\xi^2 + \eta^2) = -(n_0u\sigma_x + n_0w\tau_x)\left[ (n_0u\sigma_x + n_0w\tau_x)^2 + \left( n_0u\sigma_x + n_0w\tau_y \right)^2 \right] \xi = -[n_0^2 u^2 \sigma^2 + n_0 w^2 \tau_x + n_0 w \sigma_x + n_0 w \tau_y]$$  \hspace{1cm} (3.4.12)

Proceeding as in Sec. 3.3.3, it follows from Eqs. (3.4.11-12) that for transfer through axial gradients we have
\[-2(\Delta \Lambda )_3 = \left[ T_1 \left( \sigma_x^2 + \sigma_y^2 \right) + 2T_2 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + \left( T_3 + P_T \right) \left( \tau_x^2 + \tau_y^2 \right) \right] \sigma_x + \left[ T_2 \left( \sigma_x^2 + \sigma_y^2 \right) + 2T_3 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + T_4 \left( \tau_x^2 + \tau_y^2 \right) \right] \tau_x \quad (3.4.13)\]

By comparing the coefficients, we finally obtain
\[
T_1 = n_0^2 u^2 \left( n_{ud}/n_0 \right) \left( 1 + \Delta \left( \frac{h}{n_0^2} \right) \right)
\]
\[
T_2 = n_0^2 u^2 n_0 w \left( n_{ud}/n_0 \right) \left( 1 + \Delta \left( \frac{h}{n_0^2} \right) \right)
\]
\[
T_3 = n_0 w n_0^2 \left( n_{ud}/n_0 \right) \left( 1 + \Delta \left( \frac{h}{n_0^2} \right) \right)
\]
\[
T_4 = n_0^3 w^3 \left( n_{ud}/n_0 \right) \left( 1 + \Delta \left( \frac{h}{n_0^2} \right) \right)
\]
\[
P_T = 0 \quad (3.4.14)
\]

The equations (3.4.14) and (3.4.10) give the contributions of the transfer through an arbitrary axial GRIN medium to the Seidel aberration coefficients of the system. With the notations introduced in Sec. 3.1.1., \( T_1 \) stands for spherical aberration, \( T_2 \) for coma, \( T_3 \) for astigmatism, \( T_4 \) for distortion, and \( P_T \) for the Petzval curvature. The change at transfer of the height of the marginal ray is given by Eq. (2.2.22).

By taking into account the differences in notation, the equations (3.4.14) are precisely the same as the expressions for the Seidel coefficients obtained by Sands /47/. As far as we know, the above derivation is the first independent confirmation of Sands' results. Since the expressions of Seidel aberration coefficients are of a considerable practical importance, such an independent confirmation, which is based on a derivation method differing from that of Sands, was necessary to ensure the correctness of Sands' Seidel formulae.

It can be seen from Eqs.(3.4.14) that the axial gradient medium does not contribute to the Petzval curvature. For the remaining coefficients, we have
\[
\frac{T_4}{T_3} = \frac{T_3}{T_2} = \frac{T_2}{T_1} = \frac{n_0 w}{n_0 u} \quad . \quad (3.4.15)
\]

Consider now the paraxial transfer equations (2.2.21-22) together with the Seidel surface contributions given by Eqs.(3.3.36) and (3.3.4), and the Seidel transfer contributions given by Eqs. (3.4.14)and (3.4.10). As can be seen from these equations, the arbitrary axial refractive index distribution \( n=n(z) \) determines the Seidel coefficients of a GRIN lens through the following parameters:

1. the refractive indices \( n_0 \) at the vertices of the two end surfaces,
2. the derivatives $N_z$ of the refractive index at the vertices of the two end surfaces,
3. the average values over the axial thickness of $1/n$ and $1/n^3$ i.e. $<n^{-1}>$ and $<n^{-3}>$.

As can be seen from Eqs.(2.2.21) and (3.4.10), the numerical values for $<n^{-1}>$ and $<n^{-3}>$ must be found by quadrature. For general axial GRIN media, the quadrature must be carried out numerically.

### 3.4.3 Special cases

If $n(z)$ has a simple analytical form, $<n^{-1}>$ and $<n^{-3}>$ can be found without the numerical quadrature. For instance, for a linear axial distribution
\[ n(z) = n_0(1 + \alpha z) \]
(3.4.16)
it follows by putting $q=1$ in Eqs. (2.2.16-18) that
\[ \begin{align*}
\langle n^{-1} \rangle &= \frac{1}{n_0 \alpha d} \ln(1 + \alpha d), \\
\langle n^{-3} \rangle &= -\frac{1}{2n_0 \alpha d} [(1 + \alpha d)^{-2} - 1]. 
\end{align*} \]
(3.4.17)

In this case we have
\[ n'_0 = n_0(1 + \alpha d), h' = h - \frac{u}{\alpha} \ln(1 + \alpha d), N_z = n_0 \alpha \]

Assuming that the media outside the lens are homogeneous, the sum of the two inhomogeneous surface contributions to the spherical aberration at the end faces is obtained from Eqs.(3.3.36) as
\[ S_i^r = n_0 \alpha \left( h^4 \rho_1^2 - h'^4 \rho_2^2 \right) \]
(3.4.18)
where $\rho_1$ and $\rho_2$ are the two surface curvatures. Similar formulae are valid for coma, astigmatism and distortion. Note that, because at the end surfaces $\Delta N_z$ at refraction has opposite signs, the two inhomogeneous surface contributions in Eq. (3.4.18) tend to cancel each other out, irrespective of the signs of the surface curvatures. Thus, the effectiveness of this type of axial gradient for aberration correcting is considerably reduced. Additional comments on the use of axial gradients for correcting the Seidel aberrations of the system will be made in Chapter 4.

The transfer contribution to the spherical aberration can be written as
\[ T_i = n_0 \alpha \left\{ -\frac{u}{2n_0 \alpha} \left[ (1 + \alpha d)^{-2} - 1 \right] + \frac{h - \frac{u}{\alpha} \ln(1 + \alpha d)}{n_0^2 (1 + \alpha d)^3} - \frac{h}{n_0^2} \left\{ \left( h - \frac{u}{2\alpha} \right) \left[ (1 + \alpha d)^{-2} - 1 \right] - \frac{u \ln(1 + \alpha d)}{\alpha (1 + \alpha d)^2} \right\} \right\} \]
(3.4.19)
The remaining transfer contributions follow immediately from Eqs. (3.4.15).

The axial refractive index distribution $n=n(z)$ must not be a continuous function of the variable $z$. For example, two distinct homogeneous media separated by a vertical plane surface can also be regarded as forming an axial gradient medium with a refractive index
distribution given by a step function. This gives us the possibility to verify in this special case the validity of the formulae for the axial transfer contributions to the Seidel coefficients.

Consider a doublet lens having a plane cemented surface which separates two homogeneous media having the indices \( n_A \) and \( n_B \). (See Fig. 3.4.1.) The Seidel coefficients of this lens can be computed in two ways:
1. the contributions of the end surfaces plus the contribution of the cemented surface,
2. the contributions of the end surfaces plus the contribution of the equivalent "axial" gradient.

![Fig. 3.4.1 Verification of Eqs. (3.4.14) in the special case when n(z) is a step-function](see text)

At the end faces we have \( N_z = 0 \) and therefore the "inhomogeneous" surface contributions vanish. Thus, the coefficients \( T_p \), \( p = 1..4 \) of the equivalent "axial" gradient given by Eqs.(3.4.14) must be equal to the coefficients \( S_p \) calculated for the cemented surface by means of the expressions for the ordinary surface contributions (3.3.40).

Let \( d_A \) and \( d_B \) be the axial thicknesses of the two homogeneous media. Eq.(3.4.10) thus reads

\[
\frac{1}{n^{-2}} = \frac{1}{d_A + d_B} \left( \frac{d_A}{n_A^2} + \frac{d_B}{n_B^2} \right)
\]

(3.4.20)

By denoting the marginal ray heights at the two end surfaces and at the cemented surface by \( h_A \), \( h_B \) and \( h_{AB} \), and the marginal ray inclinations in the two media by \( u_A \) and \( u_B \), the "axial transfer" contribution to the spherical aberration is given by
\[ T_1 = (n_A u_A)^3 \left[ n_A u_A \left( \frac{d_A}{n_A^3} + \frac{d_B}{n_B^3} \right) + \frac{h_B}{n_B^2} - \frac{h_A}{n_A^2} \right] \tag{3.4.21} \]

At the transfer through the two homogeneous media we have
\[ h_A = h_{AB} + u_A d_A \]
\[ h_B = h_{AB} - u_B d_B \tag{3.4.22} \]

Because the cemented surface has \( \rho = 0 \), we obtain from Eqs. (2.1.30) and (3.3.39)
\[ n_A u_A = -n_A i_A = -n_B i_B = n_B u_B \]
\[ n_A w_A = -n_A j_A = -n_B j_B = n_B w_B \tag{3.4.23} \]

Substituting Eqs. (3.4.22) into Eq. (3.4.21) we obtain
\[ T_1 = (n_A u_A)^3 \left[ \frac{u_A d_A}{n_A^2} + \frac{u_B d_B}{n_B^2} - \frac{h_{AB} - u_B d_B}{n_B^2} + \frac{u_A d_A}{n_A^2} \right] = (n_A u_A)^3 h_{AB} \left( \frac{1}{n_B^2} - \frac{1}{n_A^2} \right) = (n_A u_A)^3 h_{AB} \left( \frac{u_B}{n_B} - \frac{u_A}{n_A} \right) \tag{3.4.24} \]

Using Eqs. (3.4.23) then yields
\[ T_1 = (n_A i_A)^2 h_{AB} \left( \frac{u_B}{n_B} - \frac{u_A}{n_A} \right) = S_1 \tag{3.4.25} \]

For coma and astigmatism we have
\[ \frac{S_2}{S_1} = \frac{S_3}{S_2} = \frac{n_A j_A}{n_A i_A} = \frac{n_A w_A}{n_A u_A} \tag{3.4.26} \]

By comparing Eqs. (3.4.26) and (3.4.15) it follows that also for coma and astigmatism we have
\[ T_2 = S_2, T_3 = S_3 \tag{3.4.27} \]

Finally, for distortion we have
\[ S_4 = (n_A j_A)^2 m_{AB} \left( \frac{u_B}{n_B} - \frac{u_A}{n_A} \right) + n_A j_A H \left( \frac{w_B}{n_B} - \frac{w_A}{n_A} \right) = \]
\[ = \left( \frac{1}{n_B^2} - \frac{1}{n_A^2} \right) \left[ (n_A w_A)^2 m_{AB} n_A u_A - (n_A w_A)^2 (m_{AB} n_A u_A - h_{AB} n_A w_A) \right] = (n_A w_A)^2 h_{AB} \left( \frac{1}{n_B^2} - \frac{1}{n_A^2} \right) = T_4 \tag{3.4.28} \]

Thus, the general expressions (3.4.14) for the axial transfer contributions to the Seidel coefficients have produced in this special case the expected results and turned out to be correct.

Besides providing a way to verify the validity of the Seidel transfer formulae, Eqs. (3.4.24-28) will enable in Chapter 4 a physical interpretation of the thin-lens approximation for axial GRIN lenses.
3.5 Transfer contributions of radial gradient-index lenses to the Seidel coefficients

3.5.1 Third-order terms at transfer through radial gradients

The first stage of the derivation of Seidel transfer contributions of radial GRIN lenses is the same as for axial gradients. At transfer through radial gradients, Eq. (3.4.5) reads

\[ (\Delta \lambda^*_s) = n_0 n_0' \xi^* + h^* n_0' + \frac{1}{2n_0^2} \Delta \left( h^* \xi^* + n_0^* \right) \]  

(3.5.1)

Bearing in mind the observations made in Sec. 3.4.1, let us now determine \( \xi' \) and \( \xi'' \) from the analytic ray tracing formulae developed in Sec. 2.4. By substituting in Eqs. (2.4.45)

\[ \phi_s = \phi_s + \phi_s' + O(4), \phi_s' = \phi_s' + O(4) \]  

(3.5.2)

we obtain

\[ \xi_p' = x \phi_{s,p} + \frac{\xi}{n_0 g} \phi_{s,p} \]  

(3.5.3)

Similar expressions can also be written for \( \eta' \) and \( \xi' \). From Eqs. (2.4.26-27) we find

\[ \phi_s = \phi_s + \mu N_s \phi_{s,11} + ... \]  

(3.5.4)

The equations (2.4.31) and (2.4.33) for the coefficients and their derivatives can be re-written as

\[ \phi_{s,0} = \cos \theta, \phi_{s,0} = -\sin \theta \]  

(3.5.5)

\[ \phi_{s,11} = -\frac{1}{8} a \cos t - \frac{1}{8} b \sin t - \frac{1}{2} b t \cos t + \frac{1}{2} a \cos t + \frac{1}{2} (a + 2) \sin t + \frac{7}{8} b \sin t \]

\[ \phi_{s,11} = \frac{3}{8} b \cos t + \frac{3}{8} a \sin t + \frac{1}{2} (a + 2) \cos t + \frac{3}{8} b \cos t + \frac{1}{2} b t \sin t + \frac{1}{8} (3a + 8) \sin t \]  

(3.5.6)

where \( a, b \) and \( \mu \) are given by Eqs. (2.4.14) and (2.3.12) and \( t \) by Eq. (2.3.8).

For determining the quantities appearing in Eq. (3.5.2) note first that \( t \) is also depending on \( x, y, \xi, \eta \). Therefore, it follows from Eq. (2.3.12) that

\[ t = \frac{n_0}{\xi_0} g = \frac{g z}{\sqrt{1 - 2 \mu}} = g z (1 + \mu) + O(4) \]  

(3.5.7)

In Eqs. (3.5.5) we insert the Taylor series expansions
\[
\cos t = \cos gz - \mu gz \sin gz + O(4)
\]
\[
\sin t = \sin gz + \mu gz \cos gz + O(4)
\]

Consequently, it follows that
\[
\phi_{c,0} = \cos gz, \phi_{s,0} = \sin gz, \dot{\phi}_{c,0} = -\sin gz, \dot{\phi}_{s,0} = \cos gz
\]
\[
\phi_{c,2} = -\mu gz \sin gz + \mu N_4 \phi_{c,11}, \dot{\phi}_{c,2} = -\mu gz \cos gz + \mu N_4 \dot{\phi}_{c,11}
\]
\[
\phi_{s,2} = \mu gz \cos gz + \mu N_4 \phi_{s,11}, \dot{\phi}_{s,2} = -\mu gz \sin gz + \mu N_4 \dot{\phi}_{s,11}
\]

Thus, Eqs. (3.5.3) read for \( p=1 \)
\[
x'_1 = x \cos gz + \frac{\xi}{n_0 g} \sin gz
\]
\[
\xi'_1 = -n_0 g x \sin gz + \frac{\zeta}{n_0 g} \cos gz
\]

and, for \( p=3 \), after some algebra,
\[
x'_3 = \mu \frac{\xi'_1}{n_0} + \mu N_4 \left( x \phi_{c,11} + \frac{\zeta}{n_0 g} \phi_{s,11} \right)
\]
\[
\xi'_3 = -\mu n_0 g^2 z x'_1 + \mu N_4 \left( n_0 g x \phi_{c,11} + \zeta \phi_{s,11} \right)
\]

Finally, the paraxial approximation for Eqs. (3.5.11) is
\[
\tilde{x}'_3 = \tilde{\mu} \frac{\tilde{\xi}'_1}{n_0} + \tilde{\mu} N_4 \left( \tilde{x} \phi_{c,11} + \frac{\tilde{\zeta}}{n_0 g} \phi_{s,11} \right)
\]
\[
\tilde{\xi}'_3 = -\tilde{\mu} n_0 g^2 \tilde{d} x' + \tilde{\mu} N_4 \left( \tilde{n}_0 g \tilde{x} \phi_{c,11} + \tilde{\zeta} \phi_{s,11} \right)
\]

We now substitute into Eqs.(3.5.1), (3.5.6) and (3.5.12) the following equations: (2.4.14) for \( a \) and \( b \), (2.3.15) for \( \mu \), (2.3.22) for \( \tilde{x}' \) and \( \tilde{\xi}' \), and (2.3.24) for \( u' \) and \( h' \) and also replace everywhere \( t \) by \( gd \). In the expression for \( (\Delta \Lambda_x)_3 \) thus obtained, expressing \( \tilde{x}, \tilde{y}, \tilde{\xi}, \tilde{\eta} \) through \( (\sigma_x, \tau_x) \) and \( (\tau, \sigma_x) \) (see Eqs (2.5.9)) yields, as in the case of the surface contributions (Eq. (3.3.31)), an expression of the form
\[
(\Delta \Lambda_x)_3 = \left[ c_1 (\sigma_x^2 + \tau_x^2) + c_2 (\sigma_x \tau_x + \pi_x \pi_x) + c_3 (\tau_x^2 + \tau_x^2) \right] \sigma_x + \left[ c_4 (\sigma_x^2 + \sigma_x^2) + c_5 (\pi_x \pi_x + \sigma_x \tau_x) + c_6 (\tau_x^2 + \tau_x^2) \right] \tau_x
\]

Because of the large amount of symbolic calculation, this substitution and the subsequent processing of the results can be most conveniently done by means of computer algebra software. (We have used DERIVE and MATHEMATICA.) The expressions for the coefficients \( c_1-c_6 \) are of a considerable length and therefore we do not reproduce them here. As can be expected from Eqs.(3.5.6),(2.3.22) and (2.3.24), these expressions contain sums of sines and cosines. The coefficients of the trigonometric functions are quantities of total order four in the (unprimed) marginal and chief ray data.

As in the case of the surface contributions (see Sec. 3.3.3 for details ) comparing the expressions obtained for \( c_2 \) and \( c_4 \), we have found that \( c_2 = 2c_4 \).

Thus, we have
Raw expressions for the transfer contributions of the radial GRIN medium to the Seidel aberration coefficients result from comparing the coefficients of Eqs. (3.5.13) and (3.5.14)

\[
-2(\Delta \lambda_x) = \left[ T_1 \left( \sigma_x^2 + \sigma_y^2 \right) + 2T_2 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + (T_3 + P_T) \left( \tau_x^2 + \tau_y^2 \right) \right] \sigma_x
+ \left[ T_5 \left( \sigma_x^2 + \sigma_y^2 \right) + 2T_6 \left( \sigma_x \tau_x + \sigma_y \tau_y \right) + T_4 \left( \tau_x^2 + \tau_y^2 \right) \right] \tau_x
\]  

(3.5.14)


where, as in the case of axial gradients, \( T_1 \) stands for the spherical aberration, \( T_2 \) for coma, \( T_3 \) for astigmatism, \( P_T \) for Petzval curvature, and \( T_4 \) for distortion.

A simple expression is obtained at this stage only for the Petzval curvature

\[
P_T = kd H^2 / n_0
\]

(3.5.16)

as first obtained by Moore and Sands / 36/. For the remaining Seidel coefficients, the technique described above produces large algebraic expressions which seem similar to those given earlier in Ref. / 36/. For shortening these expressions we use a heuristic method based on symmetry requirements. This method turns out to be very powerful. As in the case of the surface contributions and axial transfer contributions, simple closed expressions will be obtained in the next section for all radial transfer contributions.

### 3.5.2 Transfer contributions of positive and negative gradients

In order to find simple expressions for the remaining transfer coefficients, we apply a heuristic technique. As noted before, at this stage of the derivation the coefficients \( c_1-c_6 \) containing large sums of trigonometric functions are all expressed through marginal and chief ray data before transfer. However, this particular form of the expressions resulting from derivation described above is not necessarily the most convenient one.

The technique for simplifying these expressions is suggested by an analogy with the case of the ordinary surface contributions, Eqs. (3.3.40), which are completely symmetric in primed and unprimed quantities. It can be easily verified that, if the various invariants and the ray data after refraction would be replaced in these coefficients through ray data prior to refraction, the length of the expressions would increase considerably. Similarly, in the case of transfer contributions, since we have no reason to prefer ray data before transfer to those after transfer, we expect that a suitable combination of terms which are completely symmetric in primed and unprimed quantities can lead to shorter expressions for the coefficients \( c_1-c_6 \). By expressing the terms also through marginal and chief ray data after transfer, we try to eliminate in Eqs. (3.5.15) the explicit appearance of the trigonometric functions, which are responsible for the considerable length of these coeffi-
ficients. Therefore we set up several requirements and seek all possible terms satisfying all of them.

1. We assume that the radial transfer contributions can be expressed by sums of terms in which the ray data appear either through invariants or through differences between quantities after and prior to transfer. No trigonometric functions should appear.

2. All terms must be of total order four in $u, h, w, m, u', h', w', m'$. Here we note that all invariants, $H$ given by Eq. (2.3.31) and $e_1, e_2, e_3$ given by Eqs. (2.3.32-34), are of second order.

3. All terms must have the dimension of length. This is because both $\Xi_x$ and $\Lambda_x$ have that dimension and because $(\sigma_x, \sigma_y)$ and $(\tau_x, \tau_y)$ are dimensionless. The elementary quantities having the dimension of length are $d, H, h, m, h', m'$, the dimensionless ones are $e_1, e_2, e_3, u, w, m, u', w'$, and $k = g^2$ has the dimension of length$^{-2}$.

4. All terms must vanish when $d$ tends to zero.

5. No term should contain a factor $d$ at a power larger than one. It can be observed that in Eqs. (3.5.6)(with $t$ replaced by $gd$) $d$ appears either in the argument of trigonometric functions or as a linear factor - no higher powers occur. Therefore, after the insertion of Eqs. (3.5.6), $(\Delta \Lambda_x)_j$ also contains $d$ only as a linear factor.

6. All terms must be real for $k < 0$ if the same formulae shall be usable both for positive and negative gradients.

We have found only a limited number of independent terms satisfying all the above requirements. These terms can be divided into three groups:

i) terms containing a product of two invariants:

$$kdH^2, de_pe_q, p,q=1,2,3$$

ii) terms containing an invariant and the difference of a second order ray data product:

$$H\Delta(u^2), H\Delta(uw), H\Delta(w^2),
\begin{align*}
ed_p\Delta(hu), & e_p\Delta(mu), e_p\Delta(mw), p=1,2,3
\end{align*}$$

Because of the invariance of $H$, $e_1, e_2, e_3$, other terms of this type can be expressed through the above six terms. We have

$$e_p\Delta(hw) = e_p\Delta(mu), Hk\Delta(h^2) = -H\Delta(u^2), Hk\Delta(hm) = -H\Delta(uw), Hk\Delta(m^2) = -H\Delta(w^2)$$

iii) terms containing the difference of a fourth order ray data product:

$$\Delta(hu^3), \Delta(hu^2w), \Delta(huw^2), \Delta(hw^3),
\begin{align*}
\Delta(mu^3), & \Delta(mu^2w), \Delta(muw^2), \Delta(mw^3)
\end{align*}$$

7. Because of Eqs. (3.1.10-11) each of the transfer coefficients (3.5.15) must be of the same order in marginal ray data (MRD) $h, u, h', u'$ and chief ray data (CRD) $m, w, m', w'$ as the corresponding surface coefficient in Eqs. (3.3.35) or (3.3.40). Therefore, the transfer coefficients are of the form:

$$T_1: \text{MRD}^4, T_2: \text{MRD}^3\text{CRD}, T_3: \text{MRD}^2\text{CRD}^2, T_4: \text{MRD CRD}^3$$
where the superscript stands for the corresponding total order.

For each transfer coefficient we select all possible terms having the required MRD and CRD orders and assume that all transfer coefficients can be expressed as linear combinations of these terms. We assume e.g. \( T_1 \) to be of the form

\[
T_1 = n_0 \left( C_1 d e_1^2 + C_2 e \Delta(h u) + C_3 \Delta(h u^3) \right)
\]  

(3.5.17)

Inserting Eqs.(2.3.24) and (2.3.32-34) into (3.5.17) and then Eq.(3.5.17) into the first of the Eqs. (3.5.15) we have found

\[
\begin{align*}
-n_0 (C_3 - N_4 - 1) & \left[ \frac{1}{4} h u (g^2 h^2 - u^2) \cos 4 g d + \frac{1}{16 g} (g^4 h^4 - 6 g^2 h^2 u^2 + u^4) \sin 4 g d \right] + \\
+ n_0 (2 C_2 + C_3 + 4 N_4 - 1) & \left[ \frac{1}{4} h u (g^2 h^2 + u^2) \cos 2 g d + \frac{1}{8 g} (g^4 h^4 - u^4) \sin 2 g d \right] + \\
+ \frac{n_0 d}{4} (2 C_1 + 3 N_4 - 2) (g^2 h^2 + u^2) - \frac{n_0}{4} (2 C_2 + 5 N_4) g^2 h^3 u - \frac{n_0}{4} (2 C_2 + 2 C_3 + 3 N_4 - 2) h u^3 = 0
\end{align*}
\]  

(3.5.18)

The assumption (3.5.17) means that all brackets containing unknown coefficients must simultaneously vanish.

\[
\begin{align*}
C_3 - N_4 - 1 &= 0 \\
2 C_2 + C_3 + 4 N_4 - 1 &= 0 \\
2 C_1 + 3 N_4 - 2 &= 0 \\
2 C_2 + 5 N_4 &= 0 \\
2 C_1 + 2 C_3 + 3 N_4 - 2 &= 0
\end{align*}
\]  

(3.5.19)

The system of linear equations has the solutions

\[
C_1 = 1 - 3 N_4 / 2, C_2 = -5 N_4 / 2, C_3 = 1 + N_4
\]  

(3.5.20)

and assumption (3.5.17) turns out to be correct. Short expressions for \( T_2, T_3 \) and \( T_4 \) are obtained by the same technique. Finally, together with Eq.(3.5.16) , the transfer contributions to the Seidel coefficients read

\[
\begin{align*}
T_1 &= n_0 d e_1^2 (1 - 3 N_4 / 2) + n_0 (1 + N_4) \Delta(h u^3) - 5 n_0 N_4 e_1 \Delta(h u) / 2 \\
T_2 &= n_0 d e_1 e_2 (1 - 3 N_4 / 2) + n_0 (1 + N_4) \Delta(h u^2 w) - 5 n_0 N_4 e_1 \Delta(h u) / 2 - N_4 H \Delta(u^2) \\
T_3 &= n_0 d e_2^2 (1 - 3 N_4 / 2) + n_0 (1 + N_4) \Delta(h u v^2) - 5 n_0 N_4 e_2 \Delta(h u) / 2 - 2 N_4 H \Delta(u v) - N_4 P_T / 2 \\
T_4 &= n_0 d e_2 e_3 (1 - 3 N_4 / 2) + n_0 (1 + N_4) \Delta(h v^3) - 5 n_0 N_4 e_2 \Delta(m u) / 2 - N_4 H \Delta(w^2) / 2
\end{align*}
\]  

(3.5.21)

It can be easily verified that if \( k \) tends to zero for any finite value of \( N_4 \), i.e. in the case of a homogeneous medium, all transfer contributions vanish.
3.5.3 Transfer contributions of shallow gradients

If the series expansion for the refractive index does not contain a quadratic term (shallow gradient) then Eq.(2.3.6) must be replaced by

\[ n^2(r^2) = n_0^2(1 + \varepsilon r^4) + O(6) \]  \hspace{1cm} (3.5.22)

In this case all inhomogeneous surface contributions (3.3.35) vanish and the paraxial transfer equations for the marginal and chief rays are given by Eq.(2.3.30). However, the transfer contributions to the Seidel coefficients can be obtained directly from Eqs. (3.5.16) and (3.5.21) by letting \( k \) tend to zero and \( N_4 \) tend to infinity such that \( \varepsilon = k^2 N_4 \) is kept constant. This derivation can also be conveniently performed by means of computer algebra. In this case after applying a symmetrization technique too, we obtain

\[ P_T = 0 \]  \hspace{1cm} (3.5.23)

and

\[ T_1 = -\frac{4}{5} d n_o \varepsilon \left( h^4 + h^3 h' + h^2 h'^2 + h h'^3 + h'^4 \right) \]

\[ T_2 = -\frac{1}{5} d n_o \varepsilon \left( 4 h^3 m + 3 h^2 h' m + 2 h h'^2 m + h'^3 m + 2 h^2 h' m' + 3 h h'^2 m' + 4 h'^3 m' \right) \]

\[ T_3 = -\frac{2}{15} d n_o \varepsilon \left( 6 h^2 m^2 + 3 h h' m^2 + h'^2 m^2 + 3 h^2 m m' + 4 h h' m m' + 3 h'^2 m m' + 3 h' m'^2 + 3 h h'^2 + 6 h'^2 m'^2 \right) \]

\[ T_4 = -\frac{1}{5} d n_o \varepsilon \left( 4 h m^3 + h' m^3 + 3 h m^2 m' + 2 h m^2 m' + 3 h m m'^2 + 2 h m m'^2 + 3 h m'^2 + h m'^3 + 4 h' m'^3 \right) \]  \hspace{1cm} (3.5.24)

3.5.4 Test of the results

The validity of all Seidel aberration formulae for radial GRIN lenses derived in this chapter can be verified by means of an independent numerical test based on ray-tracing. For the transverse magnification \( \beta = -1 \) we have found with the above formulae several cases of very thick single radial GRIN lenses where all five Seidel coefficients are zero , irrespective of the stop position, e.g. for

\[ n_0 = 1.637, \; k = 0.108796, \; N_4 = 0.694658, \; \rho_1 = -\rho_2 = -0.783032, \; d = 9.169290 \]

The distance from the object and the paraxial image to the corresponding end surfaces is 0.11487. All data are normalized such that the effective focal length is equal to unity. (The way the above lens data have been obtained will be described in Sec. 4.1.3.)

We use this example to test the correctness of the Seidel formulae. Consider in an optical system a ray propagating close to the optical axis but otherwise arbitrary. Because for such a ray the aberration series converges rapidly, the only significant contributions to \( \Xi_x \) and \( \Xi_y \) are the ones of the lowest order that do not vanish and all higher order contributions can be neglected. Therefore, by increasing all initial ray data by a factor \( \alpha \), \( \Xi_x \) and \( \Xi_y \) will increase
nearly by the factor $\zeta = \alpha \gamma$, where $\gamma$ stands for the lowest order of the nonvanishing contributions. Set $\alpha = 2$. If the Seidel formulae are correct, then the lowest order contributions are in this case the fifth order contributions, and we expect $\zeta = 2^5 = 32$. Otherwise, the lens would not really have all Seidel coefficients zero and we would obtain $\zeta = 2^3 = 8$.

Since we need for this test accurate transverse aberration values that vary over several orders of magnitude, an analytical ray-tracing method is more adequate for this purpose than a numerical one because it does not involve step size changes and numerical error evaluations. Therefore, to determine $\Xi_x$ and $\Xi_y$ we have used the analytical method described in Sec.2.4. As shown previously, for not too large values of aperture and field, by comparison with numerical solutions and with the exact solutions known in some special cases, these ray-tracing formulae have been found to be highly accurate.

We first trace a skew ray having, for simplicity, at the object plane $x_p = \eta_p = 0$ and $y_p = \xi_p = 0.01$. We then double the initial ray data in four steps (to ensure that the results are not accidental) and determine for both aberration components the ratio $\zeta$ between the current value and the value at the previous step. As shown in Table 3.5.1, in all cases the result is in the vicinity of $2^5 = 32$, as required for proving the correctness of the Seidel formulae. The same result is obtained for different rays. A similar test has been successfully performed also for arbitrary lenses by subtracting from the total aberrations the third-order contributions using Eqs.(3.1.10). The resulting values of $\zeta$ were in all cases close to 32, too.

For small values of $k$ and large values of $N_4$ (see Eq. (2.3.6)) the numerical results obtained with the general Seidel formulae have been found to approach those obtained with the shallow gradient formulae thus providing a successful test for the latter formulae, too.

<table>
<thead>
<tr>
<th>$y_p = \xi_p$</th>
<th>$\Xi_x$</th>
<th>$\zeta_x$</th>
<th>$\Xi_y$</th>
<th>$\zeta_y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0.01</td>
<td>9.1920 $10^{-11}$</td>
<td>5.2739 $10^{-11}$</td>
<td>33.00</td>
<td></td>
</tr>
<tr>
<td>1 0.02</td>
<td>2.9681 $10^{-9}$</td>
<td>32.29</td>
<td>1.7406 $10^{-9}$</td>
<td>32.00</td>
</tr>
<tr>
<td>2 0.04</td>
<td>9.5351 $10^{-8}$</td>
<td>32.12</td>
<td>5.6207 $10^{-8}$</td>
<td>32.29</td>
</tr>
<tr>
<td>3 0.08</td>
<td>3.0740 $10^{-6}$</td>
<td>32.23</td>
<td>1.8134 $10^{-6}$</td>
<td>32.26</td>
</tr>
<tr>
<td>4 0.16</td>
<td>1.0106 $10^{-4}$</td>
<td>32.87</td>
<td>5.9527 $10^{-5}$</td>
<td>32.82</td>
</tr>
</tbody>
</table>

Table 3.5.1. Confirmation of the correctness of the Seidel aberration formulae by ray-tracing (see text)

3.5.5 Special case
In the case of axial imaging by a quarter-pitch rod lens described in Sec. 2.3.3, the Seidel aberration formulae become particularly simple. Assume that the axial object point is situated at infinity. As can be seen from Fig. 2.3.1.a in Sec. 2.3.3, the entrance pupil plane coincides with the first surface.

Because the end surfaces are plane and the marginal ray in the object space is parallel to the symmetry axis, we have at the first surface \( i = - u = 0 \) and \( h = r_{EP} \). Since according to Eq. (2.3.47) we have \( g_d = \pi/2 \), it follows from Eqs. (2.3.24) that at the second surface we have \( h' = 0 \) and \( u' = g r_{EP} \). At the image plane (which is situated at the second end surface) we have \( n_0 u_Q = n_0 g r_{EP} \).

The only one nonvanishing Seidel coefficient is that of spherical aberration. It can be observed from Eqs. (3.3.40) and (3.3.35) that at the end surfaces both the ordinary and the inhomogeneous surface contributions to the spherical aberration vanish, at the first surface because of \( i = - u = 0 \) and \( \rho = 0 \) and at the second surface because of \( h' = 0 \). For the same reasons, in the transfer contribution given by Eq. (3.5.21) we have \( h u = h u' = h u^3 = 0 \). Thus the total Seidel coefficient for spherical aberration is

\[
\Gamma_1 = T_1 = n_0 d e_1^2 \left( 1 - \frac{3}{2} N_4 \right)
\]

(3.5.25)

where, according to Eq. (2.3.32) we have \( e_1 = g^2 r_{EP}^2 + u^2 = g^2 r_{EP}^2 \). Substituting the expressions for \( d \) and \( e_1 \) into Eq. (3.5.25) we find

\[
\Gamma_1 = \frac{\pi}{2} n_0 g^3 r_{EP}^4 \left( 1 - \frac{3}{2} N_4 \right)
\]

(3.5.26)

As can be seen from Eq. (3.5.26), the spherical aberration vanishes for \( N_4 = 2/3 \). This is precisely the value of \( N_4 \) in the refractive index distribution (2.3.36) for which we have proven that this type of axial imaging is free of aberrations in all orders.

In this special case, the total third-order transverse aberration of any ray can also be determined analytically. Considering a ray with \( \sigma_y = 0 \), substituting Eq. (3.5.26) into the first of Eqs. (3.1.10) yields

\[
\Xi_{x3} = -\frac{1}{2 n_0 u_Q} \Gamma_1 \sigma_x^3 = -\frac{\pi}{4} g^2 r_{EP}^3 \left( 1 - \frac{3}{2} N_4 \right) \sigma_x^3.
\]

By denoting the ray distance to the optical axis at the first surface by \( r_0 = r_{EP} \sigma_x \) we obtain

\[
\Xi_{x3} = -\frac{\pi}{4} g^2 r_0^3 \left( 1 - \frac{3}{2} N_4 \right) = -\frac{\pi^3 r_0^3}{16} d^2 \left( 1 - \frac{3}{2} N_4 \right)
\]

(3.5.27)
Finally, for rays propagating close to the optical axis, we compare the numerical results produced by Eq. (3.5.27) with the transverse aberration obtained numerically using the finite ray-tracing method described in Sec. 2.4. If we choose \( g=0.1 \), the corresponding thickness according to Eq. (2.3.47) is \( d=15.707963 \). For \( n_0=1.5 \), \( N_4=2 \), \( N_6=N_8=0 \), \( r_0=0.1 \), \( \rho_1=\rho_2=\beta=10^{-10} \), Eq. (3.5.27) gives

\[
\Xi_{x,3} = 1570710^{-5}
\]

while finite ray tracing yields

\[
\Xi_{x,3} = 1571210^{-5}.
\]

The very good agreement between these two results provides an additional check of the aberration formalism derived in this chapter.
3.6 Primary chromatic aberrations of gradient-index lenses

Expressions for the primary chromatic aberration coefficients of axial and radial gradient-index lenses have been first published by Sands /48/, /49/. Since for radial gradients we have found a computation error in Ref. /49/ (Sec. 3.6.3), we give our independent derivation of all expressions for the surface and transfer contributions of axial and radial gradients in detail. For axial gradients, we obtain precisely the same results as Sands (Sec. 3.6.2).

3.6.1 The chromatic quasi-invariant

For deriving the primary chromatic aberration coefficients of gradient-index lenses, we first introduce a quantity which is similar to the monochromatic quasi-invariant defined in Sec. 3.2. As has been shown in Sec. 3.1.1, the primary chromatic aberrations are of paraxial nature. Therefore, it suffices to consider in what follows all quantities in the paraxial approximation.

Consider a ray through the system at an arbitrary wavelength \( \lambda \). For the convenience of notation, we denote each quantity at this wavelength by the subscript \( \lambda \) at the symbol of the quantity. All quantities without this subscript are considered at the reference wavelength \( \lambda_0 \). Thus, Eqs. (3.1.2) for the components of the chromatic paraxial aberration vector of the ray now read

\[
\Xi_{\lambda,1} - \Xi_{\lambda,1} = \Lambda_{\lambda,1} - \Lambda_{\lambda,1} = \Lambda_{\lambda,1} - \Lambda_{\lambda,1} = \Xi_{\lambda,1} \tag{3.6.1}
\]

Recall that the image plane \( Q \) is that for the reference wavelength \( \lambda_0 \). We now denote

\[
\Lambda_{\lambda,1} = n_\lambda u_\lambda x_{\lambda,1} = \Lambda_{\lambda,1} = n_\lambda u_\lambda x_{\lambda,1} \tag{3.6.2}
\]

From Eqs. (3.6.2) and (3.2.2) we obtain

\[
n_\lambda u_\lambda \left( x_{\lambda,1} - x_{\lambda,1} \right) = \Lambda_{\lambda,1} - \Lambda_{\lambda,1} \tag{3.6.3}
\]

However, since at the object plane \( P \) we have \( x_{\lambda,1} = x_P \), it follows that

\[
\Lambda_{\lambda,1} = \Lambda_{\lambda,1} = \Lambda_{\lambda,1} \tag{3.6.4}
\]

We thus have

\[
n_\lambda u_\lambda \Xi_{\lambda,1} = \Lambda_{\lambda,1} - \Lambda_{\lambda,1} = \Lambda_{\lambda,1} - \Lambda_{\lambda,1} \tag{3.6.5}
\]

By defining at each surface of the system a quantity \( \Lambda_{\lambda,1} \) and by summing up its variation over all refractions and transfers in the system, Eq. (3.6.5) can be rewritten as

\[
n_\lambda u_\lambda \Xi_{\lambda,1} = \sum \Delta \Lambda_{\lambda,1} \tag{3.6.6}
\]
As in the case of the monochromatic quasi-invariant, we determine the form of $\tilde{\Lambda}_{\lambda}$ by requesting that its variation at transfer through an homogeneous medium vanishes. Since in this case we have

$$\Delta\tilde{\gamma}_\lambda = \frac{\tilde{z}_\lambda d}{n_\lambda}, \Delta h = -ud$$

it can be easily seen that $\tilde{\Lambda}_{\lambda}$ written as

$$\tilde{\Lambda}_{\lambda} = n_0 u\tilde{\gamma}_\lambda + \frac{n_0}{n_0 \lambda} h\tilde{\gamma}_\lambda$$

satisfies the required condition.

The equations (3.6.6-7) give the decomposition of the transverse chromatic paraxial aberration of an arbitrary ray in surface and inhomogeneous transfer contributions. Because this decomposition is similar with the decomposition (3.2.20) in the monochromatic case, the quantity $\tilde{\Lambda}_{\lambda}$ will be called chromatic quasi-invariant. Substituting Eq. (3.6.6) into the first of Eqs. (3.1.13) yields

$$-\sum \Delta\tilde{\Lambda}_{\lambda} = \Gamma_{\lambda_1}\sigma_x + \Gamma_{\lambda_1}\tau_x$$

It can be seen from Eq. (3.6.8) that the two total chromatic coefficients of the system can be written as sums over surfaces and GRIN media, as anticipated in Eq. (3.1.14).

As in the monochromatic case, the definition of $\tilde{\Lambda}_{\lambda}$ is not unique. It turns out, however, that the definition (3.6.7) is the paraxial approximation for the chromatic quasi-invariant first introduced by Buchdahl /10/.

The chromatic paraxial aberrations occur because of the variation with wavelength of the refractive index parameters determining the paraxial properties of the system. For the change with wavelength of the refractive index at the surface vertex we write

$$\delta_\lambda n_0 = n_{0\lambda} - n_0$$

In addition to Eq. (3.6.9), for transfer through GRIN media we must consider the quantities

$$\delta_\lambda \langle n^{-1} \rangle = \langle n_{\lambda^{-1}} \rangle - \langle n^{-1} \rangle$$

for axial gradients and

$$\delta_\lambda k = k_{\lambda} - k$$

for radial gradients. In the following derivations, we will use the operator $\delta_\lambda$ also for denoting the chromatic difference of arbitrary functions of the above refractive index parameters.

Consider for the moment $\lambda - \lambda_0$ as a small quantity. (In chapter 4 we come back to this point. For practical applications, $\lambda - \lambda_0$ is usually taken as the full wavelength range over
which the system is to be used.) In order to determine the surface and transfer contributions to the chromatic coefficients, we consider in the next three sections the variation at refraction and transfer of \( \tilde{\Delta}_d \) and keep in the resulting expressions only the first order terms in all \( \delta_\lambda \)'s. Expressing the results through aperture and field coordinates will then lead to the primary chromatic aberration coefficients.

### 3.6.2 Surface contributions

Since from Eq. (3.6.9) we obtain

\[
\frac{n_\lambda}{n_{0,\lambda}} = \frac{n_0}{n_{0,\lambda}} - \delta_\lambda \frac{n_0}{n_{0,\lambda}} = 1 - \delta_\lambda \frac{n_0}{n_{0,\lambda}}, \tag{3.6.12}
\]

the variation of \( \tilde{\Delta}_d \) at refraction can be written as

\[
\Delta \tilde{\Delta}_d = \tilde{x}_d \Delta(n_0 u) + h \Delta \left( \frac{n_0}{n_{0,\lambda}} \tilde{\xi}_d \right) = \tilde{x}_d \Delta(n_0 u) + h \Delta \tilde{\xi}_d - h \Delta \left( \frac{\delta_\lambda n_0}{n_{0,\lambda}} \tilde{\xi}_d \right) = \tilde{x}_d \Delta(n_0 u) + h \Delta \tilde{\xi}_d - h \Delta \left( \frac{\delta_\lambda n_0}{n_{0,\lambda}} \tilde{\xi}_d \right) \tag{3.6.13}
\]

Since we have

\[
\Delta \tilde{\xi}_d = -\rho \tilde{x}_d \Delta(n_0 u), \Delta(n_0 u) = h \rho \Delta n_0 \tag{3.6.14}
\]

and

\[
\Delta n_{0,\lambda} - \Delta n_0 = n'_{0,\lambda} - n_0 - n_0 - n_0 = (n'_{0,\lambda} - n_0') - (n_{0,\lambda} - n_0) = \Delta(\delta_\lambda n_0), \tag{3.6.15}
\]

it follows that

\[
\Delta \tilde{\Delta}_d = \tilde{x}_d h \rho \Delta n_0 - h \rho \tilde{x}_d \Delta n_0 - h \Delta \left( \frac{\delta_\lambda n_0}{n_{0,\lambda}} \tilde{\xi}_d \right) = -\tilde{x}_d h \rho \Delta n_0 - h \Delta \left( \frac{\delta_\lambda n_0}{n_{0,\lambda}} \tilde{\xi}_d \right) = -\tilde{x}_d h \rho \Delta \left( \delta_\lambda n_0 \right) - h \Delta \left( \frac{\delta_\lambda n_0}{n_{0,\lambda}} \tilde{\xi}_d \right)
\]

We finally obtain

\[
\Delta \tilde{\Delta}_d = -h \left( n_{0,\lambda} \tilde{x}_d \rho + \tilde{\xi}_d \right) \Delta \left( \frac{\delta_\lambda n_0}{n_{0,\lambda}} \right) \tag{3.6.16}
\]

where we took advantage of the fact that, according to the first of Eqs. (3.6.14), the quantity in the first bracket is invariant at refraction.

Since Eq. (3.6.16) is already of first order in \( \delta_\lambda n_0 \), we can replace in this equation all quantities at wavelength \( \lambda \) with the corresponding quantities at \( \lambda_0 \), i.e. we can drop the index \( \lambda \). From Eqs. (2.5.9), (2.1.30) and (3.3.39) we obtain

\[
n_0 \tilde{x}_d \rho + \tilde{\xi} = n_0 i \sigma_x + n_0 j \tau_x \tag{3.6.17}
\]

Thus, the variation at refraction of the chromatic quasi-invariant expressed in aperture and field coordinates is given by

\[
\Delta \tilde{\Delta}_d = -h \left( n_0 i \sigma_x + n_0 j \tau_x \right) \Delta \left( \frac{\delta_\lambda n_0}{n_0} \right) \tag{3.6.18}
\]
Inserting Eqs. (3.1.14) into Eq. (3.6.8) and comparing the result with Eq. (3.6.18) we fin-
nally obtain the expressions for the surface contributions to the primary chromatic aber-
ration coefficients

\[
S_{\lambda} = h n_{\lambda} \Delta \left( \frac{\delta \lambda n_{0}}{n_{0}} \right)
\]

(3.6.19)

where \( S_{\lambda} \) stands for axial color and \( S_{\lambda} \) for lateral color. Since the paraxial refraction equations do not contain any gradient contributions, the surface contributions to the primary chromatic aberrations are for GRIN lenses the same as for the homogeneous ones.

### 3.6.3 Transfer contributions of axial gradients

Using Eq. (3.6.12), the variation of \( \tilde{\Lambda}_{\lambda} \) at transfer through axial gradients can be written as

\[
\Delta \tilde{\Lambda}_{\lambda} = n_{\lambda} u \Delta \tilde{\lambda} + \tilde{\xi}_{\lambda} \Delta \left( \frac{n_{0} h}{n_{0 \lambda}} \right) = n_{\lambda} u \Delta \tilde{\lambda} + \tilde{\xi}_{\lambda} \Delta \left( \frac{\delta \lambda n_{0}}{n_{0 \lambda}} h \right)
\]

(3.6.20)

Since we have (see Sec. 2.2.3)

\[
\Delta \tilde{\lambda} = \tilde{\xi}_{\lambda} d \frac{n_{\lambda}}{\lambda} \Delta h = -d \Delta \left( \frac{\lambda n}{n_{0 \lambda}} \right)
\]

(3.6.21)

Eq. (3.6.20) can be written as

\[
\Delta \tilde{\Lambda}_{\lambda} = n_{\lambda} u \Delta \tilde{\lambda} \left[ -\tilde{\xi}_{\lambda} d \frac{n_{\lambda}}{\lambda} \Delta \left( \frac{\lambda n}{n_{0 \lambda}} \right) \right] = n_{\lambda} u \Delta \tilde{\lambda} \left[ n_{\lambda} \Delta \left( \frac{\lambda n}{n_{0 \lambda}} \right) - d \Delta \left( \frac{\lambda n}{n_{0 \lambda}} \right) \right]
\]

(3.6.22)

or, using Eq. (3.6.10), as

\[
\Delta \tilde{\Lambda}_{\lambda} = \tilde{\xi}_{\lambda} n_{\lambda} u \Delta \tilde{\lambda} \left[ -\Delta \left( \frac{\lambda n}{n_{0 \lambda}} \right) \right]
\]

(3.6.23)

Since Eq. (3.6.22) is already of first order in \( \delta \lambda \), we can replace in this equation all quantities at wavelength \( \lambda \) with the corresponding quantities at \( \lambda_0 \). Inserting into Eq. (3.6.22)

\[
\tilde{\xi} = -n_{\lambda} u \sigma_s - n_{\lambda} w \tau_s
\]

yields

\[
\Delta \tilde{\Lambda}_{\lambda} = \left( n_{\lambda} u \sigma_s + n_{\lambda} w \tau_s \right) \left[ n_{\lambda} u \Delta \tilde{\lambda} \right] \left( \frac{n_{\lambda}}{n_{0 \lambda}} \Delta \left( \frac{\lambda n}{n_{0 \lambda}} \right) \right)
\]

(3.6.23)

It follows immediately that the transfer contributions to the primary chromatic aberrations are
where \( T_{x1} \) stands for axial color and \( T_{x2} \) for lateral color. In these equations, an equivalent form for \( \delta_x n^{-1} \) is

\[
\delta_x n^{-1} = \delta_x \left( \frac{1}{d} \int \frac{dz}{n(z)} \right) = \frac{1}{d} \int \delta_x \left( \frac{1}{n(z)} \right) dz = -\frac{1}{d} \int \frac{\delta_x n(z)}{n^2(z)} dz
\]

The equations (3.6.24-25) for the contributions of the transfer through an axial gradient to the primary chromatic aberrations of the system are equivalent with the results of Sands /49/. (Note, however, that in Refs. /48/, /49/ all chromatic coefficients are defined with the opposite sign as compared to our definition.)

Consider now, as in Sec. 3.4.2, the special case when the "axial" gradient medium consists of two homogeneous media separated by a plane surface perpendicular to the optical axis, i.e. \( n(z) \) is a step function. (See Fig. 3.4.1) We expect that in this case the transfer contributions (3.6.24) yield the same results as the surface contributions (3.6.19) for the plane surface separating the two homogeneous media. As in Sec. 3.4.2, we use this property both to verify the validity of the formulae for the transfer contributions and to enable a physical interpretation of the thin-lens approximation for axial GRIN lenses (see Chapter 4).

In this case we have

\[
\delta_x n^{-1} = \frac{1}{d_A + d_B} \delta_x \left( \frac{d_A}{n_A} + \frac{d_B}{n_B} \right)
\]

Substituting Eqs. (3.6.26) and (3.4.22) into the first of Eqs. (3.6.24) yields

\[
T_{x1} = n_A u_A \left[ n_A u_A \delta_x \left( \frac{d_A}{n_A} + \frac{d_B}{n_B} \right) - \left( \frac{\delta_x n_B}{n_B} h_B - \frac{\delta_x n_A}{n_A} h_A \right) \right] =
\]

\[
= n_A u_A \left[ u_A d_A n_A \delta_x \left( \frac{1}{n_A} \right) + u_A d_B n_B \delta_x \left( \frac{1}{n_B} \right) - \left( h_{AB} - u_A d_A \right) \frac{\delta_x n_B}{n_B} - (h_{AB} + u_A d_A) \frac{\delta_x n_A}{n_A} \right]
\]

Since we have

\[
n_A \delta_x \left( \frac{1}{n_A} \right) = n_A \left( \frac{1}{n_A} \right) = \frac{n_A}{n_A} - 1 = \frac{n_A - n_{A\lambda}}{n_{A\lambda}} = -\frac{\delta_x n_A}{n_A}
\]

and since a similar relation can be obtained by replacing the index A with B, cancelling terms and using Eqs. (3.4.23) yields
\[ T_{\lambda 1} = -n_{\lambda} u_{\lambda} h_{AB} \Delta \left( \frac{\delta_{\lambda} n}{n} \right) = n_{\lambda} i_{\lambda} h_{AB} \Delta \left( \frac{\delta_{\lambda} n}{n} \right) = S_{\lambda 1} \]  

(3.6.28)

Since in this special case we have
\[ \frac{T_{\lambda 2}}{T_{\lambda 1}} = \frac{n_0 w}{n_0 i} = S_{\lambda 2} \]

it follows that
\[ T_{\lambda 2} = S_{\lambda 2} \]

(3.6.29)

i.e. for this special case, the validity of the formulae (3.6.24) has been confirmed.

### 3.6.4 Transfer contributions of radial gradients

At transfer through radial GRIN media, the variation of the chromatic quasi-invariant (3.6.7) is given by
\[ \Delta \bar{\lambda}_{\lambda \delta} = n_0 \Delta (\alpha \lambda_{\lambda}) + \frac{n_0}{n_{\lambda \delta}} \Delta \left( h \bar{g}_{\lambda} \right) \]

(3.6.30)

The derivation of the expressions for the transfer contributions to axial color and lateral color is very similar with that described in the previous section for axial gradients. However, since the algebraic calculations are somewhat lengthy, we have preferred to perform them by means of computer algebra software (MATHEMATICA). We give below all steps of the derivation, but without reproducing the intermediate results.

Writing the paraxial transfer equations (2.3.22) for quantities at wavelength \( \lambda \) yields
\[ x_{\lambda}' = x_{\lambda} \cos g_{\lambda} d + \frac{\xi_{\lambda}}{n_{0 \lambda} g_{\lambda}} \sin g_{\lambda} d \]

(3.6.31)

\[ \xi_{\lambda}' = -n_{0 \lambda} g_{\lambda} x_{\lambda} \sin g_{\lambda} d + \xi_{\lambda} \cos g_{\lambda} d \]

From Eqs. (2.3.9) and (3.6.11) it follows that
\[ g_{\lambda} = g \left( 1 + \frac{1}{2} \frac{\delta_{\lambda} k}{k} + \ldots \right) \]

(3.6.32)

where only the linear term in \( \delta_{\lambda} k \) is kept in the series expansion of the square root. We first write
\[ \bar{x}_{\lambda} = \bar{x} + \delta_{\lambda} \bar{\xi}, \bar{\xi}_{\lambda} = \bar{\xi} + \delta_{\lambda} \bar{\xi}, n_{0 \lambda} = n_0 + \delta_{\lambda} n_0 \]

(3.6.33)

and substitute these equations together with Eq. (3.6.32) into Eq. (3.6.31). Then, we expand Eqs. (3.6.31) as a series in the operator \( \delta_{\lambda} \) and keep only the first order term. The result and the transfer equations for the marginal ray (2.3.24) are then inserted in Eq. (3.6.30). Finally, \( \bar{x} \) and \( \bar{\xi} \) are replaced with their expressions in aperture and field coordinates given by Eq. (2.5.9). Raw expressions for the coefficients for axial color and lateral color are then immediately obtained by using Eq. (3.6.8). It turns out that these
first order expressions in $\delta_k$ do not contain any of the $\delta_{\lambda}x, \delta_{\lambda}x$, introduced by Eq. (3.6.33), but only $\delta_k$.

The raw expressions for the two chromatic coefficients contain sums of sines and cosines and are considerably more complex than the corresponding expressions for the surface contributions or axial transfer contributions. We have found that these raw expressions can be put into a much shorter form by using the same heuristic symmetrization technique as for the Seidel coefficients (see Sec.3.5.2). The final results read

$$T_{\lambda 1} = n_0 \frac{\delta \lambda k}{2} [de_1 + \Delta hu]$$

$$T_{\lambda 2} = n_0 \frac{\delta \lambda k}{2} [de_2 + \Delta hw]$$

(3.6.34)

where $e_1$ and $e_2$ are given by Eqs. (2.3.32-33). The transfer equations for the marginal ray are Eqs. (2.3.24). It can be observed that in Eqs. (3.6.34) we have a complete symmetry between quantities prior to transfer and after transfer. This symmetry requirement turned out to be essential for finding a short form for the expressions of the two chromatic coefficients.

Expressions for the transfer contributions of radial gradients to the paraxial chromatic aberration coefficients of the system have been already published by Sands /49/. However, these older results contain a computation error. We have detected the error by comparing our raw expressions prior to symmetrization with the final results of Ref. /49/. The discrepancy was found to be caused by a coefficient 1/2 which was lost by Sands when he calculated the second of Eqs. (21) in his paper according to Eq. (20). Thus, the correct result following from Eq. (20) is

$$g_1 = \frac{1}{\alpha} \int \cos(\alpha x) \sin(\alpha x) dx = \frac{1}{2\alpha^2} \sin^2(\alpha x)$$

This is of course only a minor omission occurring in a derivation which is in principle correct, but the effect of this omission on the results is very harmful. This is because the lost coefficient does not appear e.g. as a common factor in the final results but it affects only one term of the sum in each chromatic contribution.

After correcting the expression of $g_1$, Sands' results turn out to be equivalent with our raw expressions. It is interesting to note that the expressions containing the error cannot be put into a short symmetric form, while applying our symmetrization technique upon the corrected expressions produces Eqs. (3.6.34).
4. Aberration correction in optical systems with GRIN lenses

For designing an optical system meeting given specifications, the constructional parameters of the system are varied until the aberrations are reduced to acceptable values. For each coefficient in the aberration series to be controlled, one degree of freedom is required. Therefore, the higher the required degree of correction is, the larger the number of design parameters must be.

For gradient index lenses, if the required values are in the producible range, the parameters of the refractive index distribution are additional degrees of freedom for aberration correction. Therefore, the use of gradients in optical systems enables a considerable reduction of the number of elements as compared to homogeneous designs of same performance.

A necessary condition for obtaining a successful design is the control of the primary aberrations. For homogeneous lenses, it is known, however, that the control of these aberrations is subjected to certain basic limitations. In this chapter we show how the additional design parameters of GRIN lenses provided by the refractive index distribution can be used for correcting the primary aberrations of the system and investigate to what extent the design limitations for homogeneous lenses apply also for GRIN lenses.

For homogeneous optical systems, an essential tool for revealing the basic limitations of design possibilities is known to be the thin lens approximation. This approximation provides a much better insight than the exact formalism into the influence of the lens parameters on the paraxial properties and on the primary aberration coefficients of the system. Thus, in early design stages the thin lens theory enables the designer to see if a proposed system layout is capable of yielding the required degree of aberration correction.

At present, no satisfactory thin lens theory for GRIN lenses is known from the literature. For radial gradients, thin-lens expressions are known only for lens power, Petzval curvature and axial color /2/. For the first time, the generalization of the homogeneous thin-lens approximation for all primary aberration coefficients both for radial and for axial gradients is described in what follows (Secs. 4.1 and 4.2). For axial gradients the generalization is straightforward and we find that, within the range of validity of the thin lens approximation, an axial GRIN lens with spherical end surfaces is equivalent to a pair of homogeneous aspherical lenses in contact, having a common plane surface. For radial
gradients, however, considering lenses with zero thickness, as in the case of the standard thin lens approximation, is not sufficient for providing the required insight into the gradient properties. For radial gradients, we show that it is possible to define a so-called extended thin lens approximation which enables a very good description of the effects of the refractive index parameters on the power and primary aberrations of the lens.

Finally, the thin lens formulae are used for describing specific features of aberration correction in the case of simple systems with radial and axial GRIN lenses (Sec. 4.3).

4.1 Extended thin-lens approximation for radial gradient-index lenses

For homogeneous optical systems, it is known that the lens thicknesses are generally a much less effective degree of freedom for controlling aberrations than the curvatures of the surfaces and the air spaces between the lenses. The aberration coefficients of a homogeneous lens having a finite but not too large thickness are of the same order of size and show approximately the same variation with the other lens parameters as in the case of the corresponding lens considered to have zero thickness /17/. On the other hand, substituting $d=0$ into the exact paraxial and aberration formulae enables a considerable simplification of the corresponding expressions for the entire lens. Thus, even if the results of the thin-lens approximation differ to some extent in absolute magnitude from the exact results, the thin-lens approximation yields an useful qualitative insight into the general properties of the lens aberrations.

For radial GRIN lenses, however, the transfer through the GRIN medium can have a significant contribution to the power and to the primary aberrations of the lens. Thus, for the aberration control of radial GRIN lenses, the lens thickness must be regarded as a design parameter as significant as the others. Since all transfer contributions vanish for $d=0$, the standard thin lens approximation yields only a poor insight into the correction possibilities of radial GRIN lenses.

We will see below that if we keep in the transfer contributions to power and primary aberration coefficients the lowest order non vanishing terms in the lens thickness $d$, i.e. the linear terms, it is possible to derive a set of approximate formulae for power and primary aberrations for the entire lens which have the same structure as the homogeneous thin-lens formulae (Secs. 4.1.1 and 4.1.2). Thus, even if the lenses considered in this new approximation are not necessarily "thin" in the standard sense, this approximation will be
The extended thin lens approximation indicates which parameters of the radial GRIN lens are effective for controlling a given primary aberration (Sec. 4.1.3). By comparison with the results produced by the exact formulae, it will be shown that the extended thin lens approximation can be used to determine the shape of generalized bending curves for various aberrations (i.e. curves giving the change of the aberrations if power is transferred from the surfaces to the medium).

The limitations of aberration correction of radial GRIN lenses are also investigated by means of the extended thin lens formalism. For example, it will be found that, unless the lens becomes very thick or the gradient very strong, a single aplanatic radial GRIN lens cannot be corrected also for astigmatism, i.e. we have the same basic limitation as in the homogeneous case. On the other hand, as already known from the literature, it is in principle possible to control simultaneously Petzval curvature and axial color of a single lens.

### 4.1.1 Paraxial approximation

In this section we define the "extended thin lens approximation" and derive the expression for the gradient contribution to the power of a thin radial GRIN lens.

Consider first a homogeneous lens in air, having a thickness which is small in comparison to the radii of the two end surfaces. In this case, the change inside the lens of the height of a paraxial ray $\Delta h = -ud$ is negligibly small. If the stop is situated at the first surface of the lens, this condition can be written as $\left|ud\right| \ll r_{EP} \ . \ (4.1.1)$

In this case we can insert $d=0$ into the coefficient $G_{21}$ of the Gauss matrix for transfer (Eq. (2.3.30)) which now becomes a unity matrix. In the thin lens approximation, the Gauss matrix of the lens (Eq. (2.5.1)) is given then by

$G = B_2 U B_1 = B_2 B_1 = \begin{pmatrix} 1 & \varphi_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \varphi_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \varphi_1 + \varphi_2 \\ 0 & 1 \end{pmatrix} \quad (4.1.2)$

where $\varphi_1$ and $\varphi_2$ are the powers of the two end surfaces. Thus, the power of the homogeneous thin lens is simply the sum of the two surface powers

$\varphi = G_{12} = \varphi_1 + \varphi_2 \quad (4.1.3)$

We replace now the homogeneous medium inside the lens with a radial gradient one. In this case the Gauss matrix for transfer is given by Eqs. (2.3.28-29) and we have

$\Delta h = -dE_2 (kd^2) \mu + \left(E_1 (kd^2) - 1\right) h \quad (4.1.4)$
If the quadratic refractive index coefficient $k$ is not too large, such that
$$|k|d^2 << 1,$$
we can replace $kd^2$ by 0 in Eqs. (2.3.29) for $E_1$ and $E_2$. We obtain
$$G_{11} = G_{22} = E_1(0) = E_2(0) = 1.$$  
As in the homogeneous case, because of Eq. (4.1.1), we can replace $d$ by 0 into the coefficient $G_{21}$. Thus, the Gauss matrix for transfer (2.3.28) then reads
$$G_r = \begin{pmatrix} 1 & n_0kd \\ 0 & 1 \end{pmatrix}$$
and Eq. (4.1.4) vanishes, i.e., as in the homogeneous case, the ray height inside the lens does not change. By writing
$$\varphi_1 = (n_0 - 1)\rho_1, \varphi_2 = -(n_0 - 1)\rho_2, \quad \varphi_h = \varphi_1 + \varphi_2$$
and by defining
$$\varphi_g = n_0kd,$$
the Gauss matrix of the lens is now given by
$$G = B_2 G_r B_1 = \begin{pmatrix} 1 & \varphi_2 & 1 & \varphi_1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & \varphi_g & 1 & \varphi_h \end{pmatrix} = \begin{pmatrix} 1 & \varphi_1 + \varphi_2 + \varphi_g \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & \varphi_g & 1 & \varphi_h \end{pmatrix}$$
and the total power by
$$\varphi = \sum \varphi_i$$
The quantities $\varphi_h$ and $\varphi_g$ will be called homogeneous and gradient contributions to the total power. The equation (4.1.11) can be easily generalized for a system of thin lenses in contact, situated in air
$$\varphi = \sum \varphi_i$$
where the index $i$ denotes the various homogeneous and gradient contributions of the components to the power of the lens system.

The approximation defined by the conditions (4.1.1) and (4.1.5) will be called extended thin lens approximation, in order to distinguish it from the standard thin lens approximation (i.e. $d=0$ in all formulae). In what follows, the word "thin" will be used in the sense of the extended thin lens approximation. In this approximation, the gradient contribution to the power of a thin radial GRIN lens is obtained by keeping in the coefficient $G_{112}$ of the transfer matrix only the linear term in the lens thickness $d$. It will be seen in the next section that the extended thin lens approximation can be applied for the primary aberration coefficients also by keeping only the linear terms in $d$ in the corresponding transfer formulae.

The limits of validity for practical purposes of the formulae derived in this approximation, such as Eq. (4.1.11) or the formulae for the primary aberration coefficients, which will be derived in the next section, can be determined only by applying these formulae in
concrete cases. As can be seen from Eq. (4.1.9), $\phi_g$ is proportional to the quadratic refractive index coefficient $k$. Thus, if the thickness is kept reasonably small, only somewhat stronger gradients can have a significant contribution to the total power. Note that, even if $kd^2 \ll 1$, the product $kd$ can still have non-negligible values. For example, in the case of an aplanatic lens for an optical disk system having $n_0=1.637$, $\sqrt{k} = g = 0.1475 \text{ mm}^{-1}$, $d=1.45 \text{ mm}$ and the surface radii $R_1=2.8 \text{ mm}$, $R_2=\infty / 21/$, we obtain
\[
kd^2 = 0.0457 << 1
\]
\[
G_{r12} = n_0gsingd = 0.05124 \text{ mm}^{-1}
\]
\[
\phi_g = n_0kd = 0.05164 \text{ mm}^{-1}
\]
\[
\phi_1 = \left( n_0 -1 \right)/ R_1 = 0.2275 \text{ mm}^{-1}
\]
\[
\phi_2 = 0
\]
The exact value of the power is
\[
\phi_{exact} = 0.27356 \text{ mm}^{-1}
\]
while the approximate value according to Eq. (4.1.11) is
\[
\phi_{approx} = \phi_1 + \phi_g = 0.2791 \text{ mm}^{-1}.
\]
If the gradient contribution to power would be neglected, the total power would be $\phi_1 = 0.2275 \text{ mm}^{-1}$, i.e. a much poorer approximation of the exact value.

As in the homogeneous case, in the extended thin lens approximation, the Gauss matrix, and therefore all paraxial properties of the lens, are determined only by the total power. For a thin lens producing a real image, the equations (2.5.2) giving the position of object and paraxial image plane for a given value of $\beta$ now read
\[
d_p = \frac{1}{\phi} \left( 1 - \beta^{-1} \right), d_Q = \frac{1}{\phi} \left( 1 - \beta \right)
\]  
(4.1.13)
The same equations are valid also for a system of thin lenses in contact. For general imaging (i.e. real or virtual), Eqs.(4.1.13) can be easily rewritten by substituting
\[
d_p = -s, d_Q = s'
\]  
(4.1.14)
such that $s$ ($s'$) is positive if the object (image) is situated to the right of the lens and negative otherwise. Eliminating $\beta$ yields the well known relation between position of object and image of a thin lens
\[
\frac{1}{s'} = \frac{1}{f} + \frac{1}{s}
\]  
(4.1.15)
For optical systems consisting of thin lenses separated by air spaces, Eq.(4.1.15) can be used to determine the position of the intermediate image (real or virtual) produced by each lens. For the given value of the total transverse magnification, the position of the object with respect of the first lens can be determined using Eqs. (2.5.2). Then, the positions of the intermediate images are determined successively using Eq. (4.1.15) by taking for each lens as object the intermediate image of the previous lens.
4.1.2 Primary aberration coefficients

The expressions derived in Chapter 3 for the primary aberrations of GRIN lenses are expressions for the contribution of refraction at each surface and transfer through each inhomogeneous medium to the corresponding aberrations. These expressions are adequate for obtaining numerical values for the various aberration coefficients, but offer only a limited insight into the correction possibilities that can be expected for the entire lens or for the entire optical system.

In this section, simplified expressions for the primary aberrations of the entire radial GRIN lens are derived on the basis of the extended thin lens approximation. By using adequate lens variables we obtain expressions for the sum of the various contributions to each aberration coefficient of the entire lens which have the same structure as their homogeneous counterparts. These expressions are then generalized for systems of thin lenses in air. Finally, from the resulting formulae the correction possibilities of the Seidel and chromatic aberrations as well as their limitations are described.

We first consider the transfer contributions of radial gradients to the primary aberrations. As in the paraxial case, where we keep only the lowest order transfer contribution to power, for the transfer contributions to the primary aberrations we consider now the power series expansions of these expressions with respect to the lens thickness $d$ and also keep only the lowest order terms in $d$. These calculations can be conveniently performed by means of computer algebra. We give below only the result.

Consider first the case of positive or negative gradients (i.e. nonzero $k$). After substituting the paraxial transfer equations (2.3.34) into Eqs. (3.5.21), keeping only the first term in the power series expansion with respect to $d$ and using Eq. (4.1.9) yields

$$T_1 = \varphi g \left[ h^4 k (1 - 4 N_4) + 5 h^2 u^2 \right]$$

$$T_2 = \varphi g \left[ h^4 m k (1 - 4 N_4) + 5 h^2 uw + 2 hu \frac{H}{n_0} \right]$$

$$T_3 = \varphi g \left[ h^2 m^2 k (1 - 4 N_4) + 5 h^2 w^2 + 4 hw \frac{H}{n_0} \right]$$

$$T_4 = \varphi g \left[ h m^3 k (1 - 4 N_4) + 5 hw^2 + mw \frac{H}{n_0} \right]$$

Similarly, from Eqs. (3.6.34) we obtain
\[ T_{\lambda 1} = \varphi_g \frac{\delta_k^2 k h^2}{k} \]  
\[ T_{\lambda 2} = \varphi_g \frac{\delta k}{h} \]  

Using Eq. (4.1.9), the Petzval curvature for transfer (Eq. (3.5.16)) can be immediately rewritten as

\[ P_T = \varphi_g \frac{H^2}{n_0^2} \]  

Note that all expressions in (4.1.16-18) are proportional to the gradient power \( \varphi_g \), i.e. to the product \( k d \).

For deriving expressions in the extended thin lens approximation for the various primary aberrations, it is convenient to consider first the case when the aperture stop is at the lens (i.e. at the both surfaces, since the lens is thin ) and then to allow for arbitrary stop positions by means of the stop-shift formulae derived in Sec. 3.1.2.

Consider now the various surface and transfer contributions if the stop is at the lens. In all formulae we have \( m=0 \) and the invariants (3.3.39) and (2.2.24) now become

\[ n_0 j = n_0 m \rho - n_0 w = - n_0 w \]  
\[ H = m n_0 u - h n_0 w = - h n_0 w \]  

Thus, the angles \( j \) and \( w \) can be written as

\[ n_0 j = \frac{H}{h} = - n_0 w \]  
i.e. all quantities related to the chief ray are removed from the aberration expressions.

Recalling that in this approximation we have \( h' = h \), the homogeneous surface contributions (3.3.40) and (3.6.19) become

\[ S_1 = (n_0 j)^2 h \Delta (u/n_0) \]  
\[ S_2 = n_0 j h \Delta (u/n_0) \]  
\[ S_3 = \frac{H^2}{h} \Delta (u/n_0) \]  
\[ P_S = - \rho H^2 \Delta (1/n_0) \]  
\[ S_4 = n_0 j H n_0 w \Delta (1/n_0^2) = - \frac{H^3}{h^3} \Delta (1/n_0^2) \]  

and

\[ S_{\lambda 1} = h n_0 j \Delta \left( \frac{\delta_{\lambda 1} n_0}{n_0} \right) \]  
\[ S_{\lambda 2} = h \Delta \left( \frac{\delta_{\lambda 2} n_0}{n_0} \right) \]  

\[ (4.1.22) \]
For the inhomogeneous surface contributions (3.3.35), we have
\[ S'_1 = -2h^4 \rho \Delta(n_j, k) \]
\[ S'_2 = S'_3 = S'_4 = 0 \]  \hspace{1cm} (4.1.23)
i.e. because of \( m=0 \), all contributions vanish, excepting the spherical aberration.

Finally, the transfer contributions (4.1.16-17) now read
\[ T_1 = \varphi_g [h^4 k (1 - 4N_4) + 5h^2 u^2] \]
\[ T_2 = \varphi_g \left[ 5h^2 uw + 2h u \frac{H}{n_0} \right] = -3 \varphi_g hu \frac{H}{n_0} \]  \hspace{1cm} (4.1.24)
\[ T_3 = \varphi_g \left[ 5h^2 w^2 + 4hw \frac{H}{n_0} \right] = \varphi_g \left( \frac{H}{n_0} \right)^2 \]
\[ T_4 = 0 \]
and
\[ T_{\Delta_1} = \varphi_g \frac{\delta_j}{k} h^2 \]  \hspace{1cm} (4.1.25)
\[ T_{\Delta_2} = 0 \]

As in the case of homogeneous thin lens theory, before summing up the various contributions for each aberration coefficient we must introduce two thin lens variables for the following purposes:

1. In the above formulae, the marginal ray inclination \( u \) must be expressed in a convenient way.
2. For a given value of the homogeneous power \( \varphi_h \) (Eq.(4.1.8)), a bending variable must be defined for expressing the relationship between the two surface curvatures \( \rho_1 \) and \( \rho_2 \).

We define the two thin lens variables such that in the limiting case of a homogeneous medium, the extended thin lens formulae reduce to the homogeneous thin lens formulae first derived by Argentieri / 1/, / 53/, / 23/ . For homogeneous media, other definitions of the thin lens variables are also possible, e.g. the Coddington variables / 59/, / 17/. However, for radial gradients, the Coddington variables seem to be less appropriate, because one of them becomes indefinite in the special case when both end faces are plane (Wood lens). We define
\[ \psi = \frac{1 + \beta \varphi_h \sigma}{1 - \beta \sigma} = \rho_1 + \rho_2 - \psi \]  \hspace{1cm} (4.1.26)
In the subsequent calculations, the indices 1 and 2 denote several quantities related to the two end surfaces. For the marginal ray inclination, primed and unprimed quantities denote values after and prior to refraction at the given surface.

Using the first of Eqs. (4.1.26) we now express the marginal ray inclination in all aberration formulae through the variable $\psi$. Observing that

$$su_i = h = s'u_2'$$  \hspace{1cm} (4.1.27)

it follows from Eq. (4.1.15) that the total change of $u$ produced by the thin lens is given by

$$u' - u_i = h\varphi = h(\varphi_h + \varphi_g)$$  \hspace{1cm} (4.1.28)

Since we have

$$u_i = \beta u_2'$$  \hspace{1cm} (4.1.29)

we can also write

$$\psi = \frac{1 + \beta}{1 - \beta} \varphi_h = \frac{u_2' + \beta u_1'}{u_2' - \beta u_1'} \varphi_h = \frac{u_2' + u_1}{u_2' - u_1} \varphi_h = \frac{u_2' + u_1}{h(\varphi_g + \varphi_h)} \varphi_h$$  \hspace{1cm} (4.1.30)

or using Eq. (4.1.27)

$$\psi = \left(1 - \frac{1}{s'}\right) \frac{\varphi_h}{\varphi_g + \varphi_h}$$  \hspace{1cm} (4.1.31)

From Eq. (4.1.30) we obtain

$$u_2' + u_i = h\psi \left(1 + \frac{\varphi_g}{\varphi_h}\right)$$  \hspace{1cm} (4.1.32)

which, together with Eq. (4.1.28) can be regarded as forming a linear system of equations for $u_1$ and $u_2'$. The solution reads

$$u_2' = \frac{1}{2} h \left[ \psi \left(1 + \frac{\varphi_h}{\varphi_g}\right) + \varphi_h + \varphi_g \right]$$

$$u_i = \frac{1}{2} h \left[ \psi \left(1 + \frac{\varphi_g}{\varphi_h}\right) - \varphi_h - \varphi_g \right]$$  \hspace{1cm} (4.1.33)

Similarly, from Eqs. (4.1.8) and (4.1.26) we obtain the linear system of equations for $\rho_1$ and $\rho_2$

$$\rho_1 - \rho_2 = \frac{\varphi_h}{n_0 - 1}$$  \hspace{1cm} (4.1.34)

$$\rho_1 + \rho_2 = \sigma + \psi$$

Thus, the surface curvatures are expressed through the Argentieri variables as
The equations (4.1.33) give the marginal ray inclinations at the two end surfaces, outside the thin lens. The values at the end surfaces, but inside the lens (i.e. after refraction at the first surface and prior to refraction at the second surface) follow immediately from Eq.(2.1.28) as

\[ u'_1 = \frac{1}{n_0} [u + (n_0 - 1)h \rho_1], \quad u'_2 = \frac{1}{n_0} [u'_2 + (n_0 - 1)h \rho_2] \]  

(4.1.36)

We must now decide what value should be assigned to the quantity \( u \) in Eqs. (4.1.24). In the derivation of these equations, \( u \) was assumed to be the marginal ray inclination in the gradient medium prior to transfer, i.e. the value after refraction at the first surface \( u = u'_1 \). We do not have, however, any reason to prefer the value prior to the transfer to the value after transfer. Thus, we will use in Eqs. (4.1.24)

\[ u = \frac{u'_1 + u'_2}{2} \]  

(4.1.37)

Note that according to Eq. (4.1.7) we have

\[ k \Delta u = n_0 u'_1 + n_0 kd \]  

(4.1.38)

i.e. \( u_2 \) differs from \( u'_1 \) by a linear term in \( d \). Thus, the replacement (4.1.37) is consistent with the approximation used for deriving Eqs. (4.1.24) - that only lowest order non vanishing terms are kept - because, as has been noted before, these equations are already linear in \( d \).

Consider now the quantities \( n_0 j \) and \( \Delta(u/n_0) \) appearing in the surface contributions (4.1.21-22). At the two surfaces we have

\[ (n_0 j) = h \rho_1 - u'_1 (n_0 j) \]  

(4.1.39)

\[ \Delta(u/n_0)_1 = u'_1 / n_0 - u'_1, \Delta(u/n_0)_2 = u'_2 / n_0 \]  

(4.1.40)

Finally, we sum up for each aberration coefficient the surface contributions at the two end faces and the transfer contributions (Eqs.(4.1.21-25)). Note first that the quantities

\[ \Delta(n_0 k), \Delta(1/n_0), \Delta(1/n_0^2) \]  

and \( \Delta \left( \frac{\delta n_0}{n_0} \right) \)

have at the two end surfaces equal magnitudes but opposite signs. Thus, for distortion and lateral color the sum vanishes because the two surface contributions cancel each other out. For the remaining aberration coefficients, substituting Eqs.(4.1.33), (4.1.35-37) and (4.1.39-40) into
Eqs. (4.1.21-25) gives the surface contributions at the two end faces and the transfer contributions expressed through the Argentieri variables. In the resulting sums, only the linear terms in $\phi_g$ (i.e. the linear terms in $d$) are kept. These calculations have been done using computer algebra. The results for the total primary aberration coefficients of the thin lens with the stop at the lens are listed below:

**Spherical aberration:**

\[
\Gamma_1 = S_{1,1} + S_{1,2} + S_{1,3} + S_{1,4} + T_1 = A = A_h + A_g + A_0^* + A_4^*
\]

\[
A_h = \frac{\varphi_g h^4}{4} \left[ \frac{n_0 + 2}{n_0} \sigma^2 - 2 \sigma \psi + \left( \frac{n_0}{n_0 - 1} \right)^2 \varphi_g^2 \right]
\]

\[
A_g = \varphi_g h^4 \left[ \frac{\sigma^2 (n_0 - 1)(3n_0 - 1)}{2n_0^2} + \frac{\sigma \psi (n_0 - 3)}{n_0} + \frac{7\psi^2}{4} + \frac{(3n_0 + 1)\varphi_g^2}{4(n_0 - 1)} \right]
\]

\[
A_0^* = -2\varphi_n h^4 \frac{n_0k}{n_0 - 1}
\]

\[
A_4^* = \varphi_g h^4 k(1 - 4N_4)
\]

**Coma:**

\[
\Gamma_2 = S_{2,1} + S_{2,2} + T_2 = B = B_h + B_g
\]

\[
B_h = \frac{\varphi_h H h^2}{2} \left( \frac{n_0 + 1}{n_0} \sigma - \psi \right)
\]

\[
B_g = \frac{\varphi_g H h^2}{2} \left( \frac{n_0 - 1}{n_0} \sigma - 3 \frac{n_0 + 1}{n_0} \psi \right)
\]

**Astigmatism and Petzval curvature:**

\[
\Gamma_3 = H^2(\varphi_h + \varphi_g) = H^2 \varphi
\]

\[
P = H^2 \left( \frac{\varphi_h}{n_0} + \frac{\varphi_g}{n_0^2} \right)
\]

**Distortion:**

\[
\Gamma_4 = 0
\]

**Axial and lateral color:**

\[
\Gamma_{41} = h^2 \left[ \varphi_h \frac{\delta_0 n_0}{n_0 - 1} + \varphi_g \left( \frac{\delta_0 n_0}{n_0} + \frac{\delta_0 k}{k} \right) \right] = h^2 \left[ \varphi_h \frac{\delta_0 n_0}{n_0 - 1} + \varphi_g \frac{\delta_0(n_0k)}{n_0k} \right]
\]

\[
\Gamma_{42} = 0
\]

As in the homogeneous case, the coefficient for axial color (4.1.52) is proportional to the chromatic change of the total power of the thin radial GRIN lens. As has been first shown by Mc. Laughlin et al. / 29/ for the two wavelengths $\lambda_0$ and $\lambda$ we have
\[
\varphi_a - \varphi = \delta_a \varphi = \delta_a \varphi_n + \delta_a \varphi_g = (\rho_1 - \rho_2) \delta_a n_0 + d \delta_a (n_0 k) = \varphi_n \frac{\delta_a n_0}{n_0 - 1} + \varphi_g \frac{\delta_a (n_0 k)}{n_0 k} \quad (4.1.54)
\]

Thus, if the axial color vanishes, the thin radial GRIN lens has the same power at the two wavelengths $\lambda_0$ and $\lambda$ (achromatic condition).

In Sec. 3.6, for deriving the expressions of the primary chromatic aberrations we assumed that the wavelength difference $\lambda - \lambda_0$ is a small quantity. The equation (4.1.54), however, is valid for arbitrarily large values of $\lambda - \lambda_0$. Thus, in view of the relationship between axial color and chromatic change of power we will consider in Eq. (4.1.52) $\lambda - \lambda_0$ to be the full wavelength range over which the system is to be used. It is thus possible to define an Abbe number for the gradient material in analogy with the Abbe number of conventional glasses. In the existing literature on gradient-index optics, the wavelengths conventionally used are those for the helium d line (587.6 nm - yellow light) and for the hydrogen C and F lines (656.3 nm and 486.1 nm - red and blue light). The homogeneous Abbe number is

\[
V_h = \frac{n_{6d} - 1}{n_{6f} - n_{6C}}. \quad (4.1.55)
\]

By defining the gradient Abbe number as

\[
V_g = \frac{(n_0 k)_d}{(n_0 k)_f - (n_0 k)_C} \quad (4.1.56)
\]

the expression (4.1.52) for axial color now reads

\[
\Gamma_{\lambda 1} = h^2 \left( \frac{\varphi_n}{V_h} + \frac{\varphi_g}{V_g} \right) \quad (4.1.57)
\]

We note at this point that in the literature, the index of refraction of a radial gradient-index glass is often given by

\[
N(r) = N_{00} + N_{10} r^2 + N_{20} r^4 + \ldots \quad (4.1.58)
\]

Up to the fourth order, the coefficients of the power series (4.1.58) are related to the coefficients of Eq. (2.3.6) by

\[
N_{00} = n_0, N_{10} = -k n_0 / 2, N_{20} = n_0 k^2 (4N_4 - 1) / 8, \ldots \quad (4.1.59)
\]

or conversely

\[
n_0 = N_{00}, k = -2 N_{10} / N_{00}, \quad N_4 = \left( N_{10}^2 + 2 N_{00} N_{20} \right) / \left( 4 N_{10}^2 \right), \ldots \quad (4.1.60)
\]

In the literature, the gradient Abbe number is usually defined as

\[
V_g = \frac{N_{10d}}{N_{10f} - N_{10C}} \quad (4.1.61)
\]

However, it can be seen from the second of the Eqs. (4.1.59) that the two definitions of the gradient Abbe number (4.1.56) and (4.1.61) produce the same numerical results.
Studies of the chromatic properties of gradient-index glasses have shown that the gradient Abbe number is a design parameter with remarkable properties. Investigations of various glasses and ion exchange pairs (see Chapter 5 for details about fabrication methods) have shown that a large range of gradient Abbe numbers from -2000 to +5000 is obtainable, particularly if multielement exchange processes are utilised. Thus, even negative gradient Abbe numbers are possible. Moreover, unlike the discrete Abbe number of conventional materials, the gradient Abbe number can be regarded as continuously varying /41/.

We now consider the case when the stop is situated at a certain distance from the lens. The expressions for the primary aberrations can be immediately derived from Eqs. (4.1.41-53) by means of the stop-shift formulae of Sec. 3.1.2. From Eqs.(3.1.23) and (3.1.20) we obtain

\[
\begin{align*}
\Gamma_1 &= A \\
\Gamma_2 &= B + \vartheta A \\
\Gamma_3 &= H^2 \varphi + 2\vartheta B + \vartheta^2 A \\
P &= H^2 \left( \frac{\varphi_h}{n_0} + \frac{\varphi_g}{n_0} \right) \\
\Gamma_4 &= 3\vartheta H^2 \varphi + \vartheta P + 3\vartheta^2 B + \vartheta^3 A \\
\Gamma_{\lambda_1} &= h^2 \left( \frac{\varphi_h}{V_h} + \frac{\varphi_g}{V_g} \right) \\
\Gamma_{\lambda_2} &= \vartheta \Gamma_{\lambda_1}
\end{align*}
\]

By evaluating the stop-shift parameter (3.1.17) at the lens plane, we find, since \(m=0\) when the stop is at the lens, that \(\Delta m=m\), i.e.

\[
\vartheta = \frac{m}{h}
\]

(4.1.63)

Let \(Z\) and \(Z'\) be the positions of the entrance and exit pupils, measured from the lens. In Eq. (3.1.26) we thus have \(n_s=1\) and \(r_s=h\). In the object space we have \(u_s = u_1\) and \(\delta z_s = Z\) and in the image space \(u_s = u'_2\) and \(\delta z_s = Z'\). Using Eq.(4.1.27) yields

\[
\vartheta = \frac{H}{h^2 \left( \frac{1}{s} - \frac{1}{Z} \right)} = \frac{H}{h^2 \left( \frac{1}{s^2} - \frac{1}{Z'} \right)}
\]

(4.1.64)

For positive and negative gradients, the quantities \(A\) and \(B\) appearing in Eqs. (4.2.62) are given by Eqs (4.1.41-48). In the special case of the shallow gradients (i.e. \(k=0\)) , a similar derivation leads also to the formulae (4.1.62) in which we must substitute \(k=0\) in all expressions, excepting that for \(A^*_1\). By substituting \(h'=h\) in Eqs. (3.5.24) we find

\[
A^*_1 = -4dn_0 \epsilon h^4
\]

(4.1.65)
All other gradient contributions vanish

\[ A_g = A_0^g = B_g = 0 \]  \hspace{1cm} (4.1.66)

The expressions for Petzval sum and axial color become the same as for homogeneous media. The expressions for \( A_h \) and \( B_h \) remain the same (Eqs.(4.1.42) and (4.1.47)).

For optical systems consisting of thin lenses separated by air spaces, the primary aberrations are given simply by summing up for all components the individual contributions given by Eqs. (4.1.62). However, for determining the individual contributions, the Argentieri variables (4.1.26) and (4.1.31) and the quantity \( m/h \) given by Eqs. (4.1.63-64) must be evaluated for each lens. The quantities \( s,s',Z,Z' \) appearing in these formulae are the values corresponding to the given lens. As noted in the previous section, this succession of object and stop images is obtained by applying Eq. (4.1.15) for each thin lens in the system.

For a single thin radial GRIN lens, the most important limitation of the possibilities of aberration correction can be seen from Eqs. (4.1.62). Within the domain of validity of the extended thin lens approximation ( lenses with a thickness which is small as compared to the surface radii and gradients which are not very strong, \( k|d^2| << 1 \) for a given value of the total power, the five Seidel coefficients are expressed only through four independent functions of the lens parameters: \( A,B,P \) and \( \vartheta \). Thus, it is not possible to correct simultaneously all five Seidel aberrations. For example, for any thin lens with nonzero power, if the spherical aberration and coma are corrected ( \( A=B=0 \) ) or if the stop is at the lens ( \( \vartheta=0 \) ), the astigmatism is proportional with the lens power and is therefore uncorrected. The same limitation occurs also for a group of thin lenses in contact. This limitation, which is well known from the homogeneous case, cannot be removed by the use of gradients.

On the other hand, if the refractive index parameters \( k, N_4 \) and \( \nu_g \) are regarded as variable and the required values are in the producible range, the larger number of design parameters creates design possibilities which are not available for homogeneous lenses. (Note that the lens thickness \( d \) is also more effective for controlling aberrations than in the homogeneous case, since it determines the gradient power (4.1.9).) As will be shown in detail in Sec.4.3.1, aplanatic correction is possible in a variety of ways. (This can already be seen by inspecting Eqs. (4.1.41-48).)

An important advantage of radial GRIN lenses is the possibility of correcting simultaneously the Petzval curvature and axial color of a single lens (plan-achromat condition). It can be seen from Eqs.(4.1.62) that if we have
the Petzval curvature vanishes. The value of the homogeneous power can then be adjusted such that the total power has the required value. For given values of $\varphi_g, \varphi_h$ and $\nu_h$, choosing

$$\nu_g = -\nu_h \frac{\varphi_g}{\varphi_h}$$  \hspace{1cm} (4.1.68)

anullates the axial color. GRIN lenses satisfying the plan-achromat condition (Eqs. (4.1.67-68)) have lead to excellent design results. A microscope objective consisting of two GRIN lenses satisfying both of them this condition which was designed at Olympus Optical Co. has nearly the same performance as a homogeneous lens design having seven lens elements / 57/. A photographic objective designed by Atkinson et al. / 2/ will be discussed in detail in Sec. 4.3.2.

In the thin lens approximation, the lateral color is automatically corrected if the axial color is corrected. This is not generally the case, however, if the exact formalism is used. In the latter case, some remaining value of the lateral color must be taken into account.

### 4.1.3 Effects of the change of the lens parameters on the Seidel aberrations

In the extended thin lens approximation, the presence of a radial GRIN medium inside the lens produces two additional types of contributions to the Seidel aberrations of the lens in comparison with the homogeneous case:

i) Inhomogeneous surface contributions (Eqs. (4.1.44) and (4.1.65) )which do not depend on the lens thickness and exist also in the limiting case $d=0$ (i.e. they are described also by the standard thin lens approximation). For lenses having a moderate thickness, these contributions are in many cases the largest contributions due to the gradient medium.

ii) Transfer contributions proportional to the gradient power $\varphi_g = n_g k d$. As noted before, these terms have a significant influence on the total aberrations of the lens (and therefore on the capabilities of the lens for correcting aberrations) only if the lens thickness $d$ has a certain nonzero value. On the other hand, with increasing thickness, the accuracy of even the standard thin lens approximation is gradually lost. In order to see whether under these seemingly contradictory circumstances the approximate expressions for the transfer contributions derived in the previous section are of practical usefulness, it is necessary to compare numerically for realistic lens data the results produced by the extended thin lens approximation with the exact results.
For homogeneous lenses, a large amount of design experience has shown that the thin lens approximation has a practical value even for lenses having finite thickness. Unless the thickness becomes very large, the qualitative conclusions that can be drawn from a thin-lens analysis upon the possibilities for aberration correction of a given layout for an optical system are in many cases confirmed by the exact calculations. In this section, we show that even if the radial GRIN lens has a finite thickness, the extended thin lens approximation can provide, as its homogeneous counterpart, a considerable insight into the dependence of the aberrations of the lens on the lens parameters.

It must be emphasized that for numerical comparison between approximate thin lens formulae and exact aberration formulae only realistic lens data are relevant. As an example in the homogeneous case, a lens having the stop at the first surface and \( n=1.518 \), \( d=0.1 \) and the surface radii \( R_1=-0.285 \), \( R_2=-0.206 \) (Lens data normalized to \( f=1 \)) has at \( \beta=-1 \) zero astigmatism despite the fact that thin lens theory forbids this. This surprising result is possible because the second surface of this lens produces very large contributions to spherical aberration and coma. Thus, zero astigmatism is a consequence of the stop shift effect produced by the small, but finite thickness \( d \). However, it is very unlikely that such a lens could be an element in an optical system of reasonable quality, because it is known that, even if they are corrected by other components in the system, large surface contributions to Seidel aberrations tend to produce very large higher order aberrations /4/.

The numerical values given in what follows for the Seidel aberration coefficients are subjected to the following two normalizations

1. \( f=1 \),
2. \( H=1 \).

The first normalization means simply that the unity of length for all system data is the effective focal length. (Obviously, this normalization can be used only for systems that are not afocal.)

We consider now the implications of the second normalization. We have seen previously (Eqs. 2.5.9) that in the paraxial approximation the ray position and direction are given everywhere in the system by linear combinations of suitably defined aperture and field coordinates, and that the coefficients of the linear combinations result from the paraxial ray tracing of two rays. Up to this point, these two rays have been chosen to be the paraxially traced marginal and chief rays. Consequently, as has been shown in Chapter 3, all primary aberration coefficients can be computed from the ray data of these two paraxially traced rays. However, any pair of rays, where one of the rays starts at the axial
object point and the other passes through the centre of the stop, can be used as well for that purpose.

For a general analysis of the properties of aberration coefficients, without reference to specific values of aperture and field, other choices of the two paraxial rays may be appropriate. Thus, if the stop is at the first surface of the lens, we choose for the initial ray data of the fictitious "marginal" and "chief" rays at the object plane

$$n_d u = n_p / d_p, h = 0, m = 1/(n_d u) n_0 w = 1$$  \hspace{1cm} (4.1.69)

where $d_p$ is given by Eq. (2.5.2). With these starting values, the "marginal" ray height at the first surface of the lens is $h_{EP}=-1$, as required by the condition $H=-h_{EP} n_0 w=1$. If the stop is shifted, $h_{EP}$ must be modified such that the path of the "marginal" ray remains unchanged and the initial value of $n_0 w$ is chosen as

$$n_0 w = -1/h_{EP}$$  \hspace{1cm} (4.1.70)

Note that Eqs. (4.1.69-70) require a definition of the aperture and field coordinates of a ray which differs from Eqs. (2.5.7-8). Inserting Eqs. (4.1.69-70) into Eqs. (2.5.9) yields the new relations between normalized and Cartesian ray coordinates. At the object plane we have

$$x_p = (d_p / n_p) r_x, y_p = (d_p / n_p) r_y$$  \hspace{1cm} (4.1.71)

and at the entrance pupil plane we have

$$x_{EP} = h_{EP} \sigma_x, y_{EP} = h_{EP} \sigma_y$$  \hspace{1cm} (4.1.72)

The relations (3.1.10) and (3.1.13), expressing the two components of the monochromatic third-order transverse aberration and of the chromatic paraxial aberration through the primary aberration coefficients and normalized ray coordinates, remain unchanged.

For comparing the approximate and exact Seidel formulae, we choose an aplanatic plano-convex radial GRIN lens designed as objective for a compact disk system, which can be currently fabricated /38/. The optimized lens parameters given in Table 1 of Ref./38/ lead to an effective focal length of $f=4.53$mm. Normalized to $f=1$, these parameters are given in Tab. 4.1.1

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_1$=0.718</td>
<td>$d=0.416$</td>
<td>$k=0.3334$</td>
</tr>
<tr>
<td>$R_2=\infty$</td>
<td>$n_0=1.58$</td>
<td>$N_4=-0.863$</td>
</tr>
</tbody>
</table>

Tab. 4.1.1 Normalized lens parameters of a radial GRIN lens designed as objective for a CD system /38/. The object is at infinity and the stop is situated at the first surface.

For the above lens parameters, Tab. 4.1.2 shows the ordinary and inhomogeneous surface contributions to the Seidel coefficients at the two end faces (denoted by 1 and 2), the
contributions of the gradient medium and the total Seidel aberrations. (The objective produces the image on the back side of the optical disk having a thickness of 1.2 mm which acts as an additional plane parallel plate. The effect of this plane parallel plate will not be considered below.)

For spherical aberration, note that the largest contribution is the inhomogeneous contribution of the first surface and that the contribution of the medium has the same order of magnitude as the two ordinary surface contributions. The astigmatism, which according to the thin lens theory, with the present normalization should be equal to unity, has the smaller value 0.79 because the lens thickness is not negligibly small as compared to the radius of the first surface.

<table>
<thead>
<tr>
<th></th>
<th>spherical aberration</th>
<th>coma</th>
<th>astigmatism</th>
<th>Petzval curvature</th>
<th>distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ordinary 1</td>
<td>0.627</td>
<td>0.45</td>
<td>0.323</td>
<td>0.511</td>
<td>0.599</td>
</tr>
<tr>
<td>inhomog. 1</td>
<td>-1.463</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>medium</td>
<td>0.486</td>
<td>-0.141</td>
<td>0.037</td>
<td>0.088</td>
<td>0.037</td>
</tr>
<tr>
<td>ordinary 2</td>
<td>0.457</td>
<td>-0.444</td>
<td>0.431</td>
<td>0</td>
<td>-0.418</td>
</tr>
<tr>
<td>inhomog. 2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>total</td>
<td>0.107</td>
<td>-0.135</td>
<td>0.791</td>
<td>0.599</td>
<td>0.218</td>
</tr>
</tbody>
</table>

Tab. 4.1.2 Seidel aberrations corresponding to the parameters of Tab. 4.1.1

The thin-lens equations (4.1.41-53) enable a qualitative description of the way the aberrations of the lens change if various lens parameters are modified. In what follows, the results produced by the approximate expressions for spherical aberration, coma and astigmatism are compared with the exact results when several lens parameters are varied. For each value of the variable parameter, the curvature of the first surface is adjusted such that the total lens power remains equal to unity.

For homogeneous lenses, it is known that, if the shape of a lens is changed while keeping its power constant, the aberrations of the lens are changed markedly. Thus, lens "bending" is one of the basic tools of the optical design /54/. For radial gradients, Eq. (4.1.11) shows that the total power of the lens consists of three terms: the contributions of the two end surfaces and the contribution of the medium. Thus, for a given value of the total power, two types of bending are possible:

i) homogeneous bending - the transfer of power from a surface to the other
ii) gradient bending - the transfer of power from a surface to the medium
In the case of the homogenous bending of a thin radial GRIN lens, $\varpi$ (Eq. (4.1.26)) is the only thin lens variable which is changing in Eqs.(4.1.41-53). As for homogeneous lenses, the thin-lens spherical aberration is quadratic in $\varpi$ (Eqs. (4.1.41-43)) and thin-lens coma is linear (Eqs(4.1.46-48)). The same conclusions apply if the curvature of the second surface is chosen as bending variable, since according to Eq. (4.1.35) $\rho_2$ is linear in $\varpi$.

As can be seen from Fig. 4.1.1, these qualitative features occur also for the exact aberrations. Within the domain of variation of $\rho_2$ considered here, the spherical aberration has an extremum while coma changes monotonically. Note that for $\rho_2=0$, i.e. the value in the optimized aplanatic design given in Tab.4.1.1, the spherical aberration reaches its minimum, the corresponding value being close to 0. In order to indicate the value of the varying lens parameter which corresponds to the optimized design, the position of the vertical axis in all figures of this section is such that it cuts the horizontal axis precisely at the optimized value given in Tab. 4.1.1.

In Fig. 4.1.2 the change by homogeneous bending of the exact spherical aberration is compared to that given by the formulae of the extended thin-lens approximation and of the standard thin lens approximation (i.e. $d=0$ in all formulae). In addition, $\beta$ was chosen to be -1 in order to show also the effect of change of the transverse magnification as compared to Fig. 4.1.1. (In all other figures we have $\beta=0$.) It can be seen that the extended thin lens approximation gives a much better quantitative description of the exact results than the standard approximation. The difference between the curves for the two types of approximations is due to the transfer contributions proportional to the gradient power (Eqs (4.1.43) and (4.1.45)).

We consider now the case of gradient bending. Keeping the second surface plane, the variation of the quadratic refractive index coefficient $k$ gives rise to large changes of the distribution of power between the first surface and the medium. For $k=0$, the lens is homogeneous. As can be seen from Fig. 4.1.3, when $k$ reaches approximately the value 1.6, the power of the first surface vanishes, i.e. this surface is also plane. Thus, for this value of $k$ the power has been transferred entirely to the gradient medium and the lens is now a Wood lens.

The change of spherical aberration by gradient bending is shown in Figs. 4.1.4-5. and that of coma in Fig. 4.1.6. It can be seen that the extended thin lens approximation gives a good qualitative description of the exact results, while the standard thin lens approximation does not. For spherical aberration, it can be seen from Fig. 4.1.5 that the differ-
ence between the exact values and the extended approximation is for all values of $k$ nearly the same as for $k=0$. Thus, as for the corresponding homogeneous lens when $k=0$, this difference comes mainly from the fact that the lens thickness is more than 40% of the focal length. Note that for the maximal value of $k$ considered here, $k=1.8$, we have $kd^2=0.31$.

Fig. 4.1.1. Change of exact spherical aberration and coma by homogeneous bending ($\beta=0$)

Fig. 4.1.2 Change of spherical aberration by homogeneous bending ($\beta=-1$)
Fig. 4.1.3 Gradient bending: By increasing the quadratic refractive index coefficient \( k \), the power of the first surface (decreasing curve) is transferred to the gradient.

Fig. 4.1.4 Change of spherical aberration by gradient bending.
Fig. 4.1.5. Same as Fig. 4.1.4, but smaller domain of variation for $k$

Fig. 4.1.6 Change of coma by gradient bending
Fig. 4.1.7 Exact astigmatism is left almost unchanged by gradient bending

Fig.4.1.8 Exact spherical aberration, coma and astigmatism as functions of $N_4$
Fig. 4.1.9. Spherical aberration as function of $n_0$

Fig. 4.1.10. Coma as function of $n_0$
According to Eq. (4.1.49) the astigmatism of a thin lens with a fixed value of the power, having the stop at the lens, cannot be affected by changes of any lens parameters. In our example of gradient bending, figure 4.1.7 confirms this prediction. It can be observed that, even if during the gradient bending the distribution of power between first surface and medium changes significantly, the exact value of astigmatism is left almost unchanged at about 0.8. As for coma (Fig. 4.1.6), for astigmatism the standard thin lens approximation produces a curve which has no resemblance at all with the exact curve.

Changes of the fourth order refractive index coefficient \(N_4\) have no influence on the power. In the exact Seidel formulae for transfer (3.5.21) \(N_4\) enters linearly in all expressions excepting the Petzval curvature. From the extended thin lens approximation it follows, however, that, if the stop is at the lens, only the spherical aberration is affected by changes of \(N_4\). This conclusion is confirmed by Fig. 4.1.8. Coma changes very slowly with \(N_4\), while astigmatism is practically constant. As can be seen from Eq. (4.1.45), the effect of \(N_4\) on the spherical aberration is proportional with the gradient power \(\varphi_g\). It will be shown in Sec. 4.3.1 that for radial GRIN lenses having the second surface plane, as in the case of the aplanatic CD objective considered here, coma correction at \(\beta=0\) requires small values of \(\varphi_g\). In our case, \(\varphi_g\) is approximately 20% of the total power. Thus, for lenses where the gradient power has a larger contribution, the same change of \(N_4\) would produce a larger effect on the spherical aberration as that shown in Fig. 4.1.8.

Finally, let us observe that in Eqs. (4.1.41-48) the dependence of spherical aberration and coma upon the refractive index on the optical axis \(n_0\) is rather complex. Figures 4.1.8-9 show the exact and approximate values for these dependences in the case of the lens given in Tab. 4.1.1 As before, \(\rho_1\) is adjusted to keep the power constant. It can be seen that within the available range of refractive indices for glass the change of spherical aberration and coma is relatively small. The approximate curves have nearly the same shape as the exact ones, but are shifted with respect to them.

As mentioned before, if
\[
\rho_1d << 1, \rho_2d << 1, k|d|^2 << 1
\]
then it follows from Eqs. (4.1.62) that, if a lens with nonzero power is corrected for spherical aberration and coma, it cannot be corrected also for astigmatism. In what follows, examples will be given when all three aberrations mentioned above are simultaneously corrected. It will be shown, however, that in these cases the conditions of validity of the extended thin lens approximation are severely violated.
Several examples of single radial GRIN lenses where spherical aberration, coma and astigmatism have been simultaneously corrected are given in a patent of Olympus Optical Co., /63/. We consider the lens given in Tab. 52 of Ref. /63/. The lens data and the corresponding Seidel coefficients are shown in Tabs. 4.1.3-4. We note that the total values of the three corrected Seidel aberrations are not exactly zero. This is because in optimized designs as this one the residual Seidel aberrations are balanced against higher order aberrations. As in the case of the lens given in Tab. 4.1.1, for spherical aberration the largest contribution is also in this case the inhomogeneous contribution of the first surface. Note also that, since this lens was optimized for a numerical aperture of 0.5 where higher order aberrations become important, the distribution of the Seidel aberrations between surface and medium is such that large individual contributions, which could produce large higher order aberrations, are avoided.

In this example, correction of astigmatism is possible because the lens thickness is large. As can be seen from Tab. 4.1.3, d has the same order of magnitude as R₁ and R₂ and kd²= 0.827 has the order of magnitude of the unity.

The condition $|k|d^2 \ll 1$ is violated also if for moderate values of the thickness the absolute value of k is large. Note, however, that k cannot be arbitrarily large, because it is related to the maximal value of the radial change of the refractive index $\delta n$. Since from Eq. (3.3.2) we have

$$-\frac{1}{2}kr^2 \approx \frac{n(r^2) - n_0}{n_0}$$

(4.1.73)

the aperture of the lens $r_{max}$ is related to $\delta n$ by

$$|k|\frac{r_{max}^2}{n_0^2} = \frac{2\delta n}{n_0}$$

(4.1.74)

Thus, for a given value of $\delta n$, very large absolute values of k would severely limit the lens aperture.

The first example in the literature for a radial GRIN lens where third-order spherical aberration, coma and astigmatism are simultaneously corrected has been given by Moore in Ref. /31/ (see Tabs. 4.1.5-6). In this example, distortion can also be corrected, namely by stop-shift because Petzval curvature is uncorrected. This lens is not optimized for higher order aberrations.

Correction of astigmatism is possible because d has the same order of magnitude as R₂ and kd²=-0.395 is not negligibly small. However, in this case the coefficient k=-12.9
$R_1 = 1.697 \quad d = 1.560 \quad k = 0.340 \quad N.A. = 0.5 \quad \beta = 0$

$R_2 = -1.356 \quad n_0 = 1.50 \quad N_4 = 0.208 \quad N_6 = 1.640 \quad f = 1$

Tab. 4.1.3 Lens parameters of an objective with a single radial GRIN lens, which is corrected for spherical aberration, coma and astigmatism /63/. The stop is situated at the first surface.

<table>
<thead>
<tr>
<th></th>
<th>spherical aberration</th>
<th>coma</th>
<th>astigmatism</th>
<th>Petzval curvature</th>
<th>distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ordinary 1</td>
<td>0.045</td>
<td>0.077</td>
<td>0.131</td>
<td>0.196</td>
<td>0.556</td>
</tr>
<tr>
<td>inhomog. 1</td>
<td>-0.601</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>medium</td>
<td>0.339</td>
<td>-0.182</td>
<td>-0.063</td>
<td>0.353</td>
<td>0.247</td>
</tr>
<tr>
<td>ordinary 2</td>
<td>0.337</td>
<td>0.103</td>
<td>0.031</td>
<td>0.246</td>
<td>0.085</td>
</tr>
<tr>
<td>inhomog. 2</td>
<td>-0.011</td>
<td>-0.029</td>
<td>-0.074</td>
<td>0</td>
<td>-0.192</td>
</tr>
<tr>
<td>total</td>
<td>0.109</td>
<td>-0.03</td>
<td>0.025</td>
<td>0.796</td>
<td>0.695</td>
</tr>
</tbody>
</table>

Tab. 4.1.4 Seidel aberrations corresponding to the parameters of Tab. 4.1.3

$R_1 = 0.333 \quad d = 0.175 \quad k = -12.920 \quad \beta = 0$

$R_2 = -0.190 \quad n_0 = 1.522 \quad N_4 = 2.788 \quad f = 1$

Tab. 4.1.5 Normalized lens parameters of a radial GRIN lens which is corrected for spherical aberration, coma and astigmatism /31/.

<table>
<thead>
<tr>
<th></th>
<th>spherical aberration</th>
<th>coma</th>
<th>astigmatism</th>
<th>Petzval curvature</th>
<th>distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ordinary 1</td>
<td>6.084</td>
<td>0.817</td>
<td>0.11</td>
<td>1.029</td>
<td>0.153</td>
</tr>
<tr>
<td>inhomog. 1</td>
<td>117.985</td>
<td>-23.479</td>
<td>4.672</td>
<td>0</td>
<td>-0.93</td>
</tr>
<tr>
<td>medium</td>
<td>-413.86</td>
<td>56.423</td>
<td>-10.03</td>
<td>-1.486</td>
<td>2.22</td>
</tr>
<tr>
<td>ordinary 2</td>
<td>71.964</td>
<td>-16.816</td>
<td>3.93</td>
<td>1.808</td>
<td>-1.341</td>
</tr>
<tr>
<td>inhomog. 2</td>
<td>217.725</td>
<td>-16.933</td>
<td>1.317</td>
<td>0</td>
<td>-0.102</td>
</tr>
<tr>
<td>total</td>
<td>-0.101</td>
<td>0.012</td>
<td>-0.001</td>
<td>1.352</td>
<td>0</td>
</tr>
</tbody>
</table>

Tab. 4.1.6 Seidel aberrations corresponding to the parameters of Tab. 4.1.5
has a large absolute value. Therefore, the inhomogeneous surface contributions become very large (see Eqs. (3.3.35)) and thus require large values from the other contributions in order to be compensated. This can be seen by comparing the distribution of the total aberrations in Tab. 4.1.6 with that in Tab 4.1.4. Recall that in both cases the same normalization for the Seidel coefficients has been used.

As a final example in this section, it will be now shown that in some cases the astigmatism can behave as predicted by the thin lens theory even for lenses where the conditions of validity of the extended thin lens approximation are severely violated.

Consider at $\beta=-1$ the simultaneous correction of spherical aberration, coma and Petzval curvature when the stop is situated at the first surface of the lens. Annulling these three aberrations and keeping $f=1$ yields a nonlinear system of four algebraic equations. This nonlinear system has been solved numerically by means of a program using the FORTRAN subroutine NLEQ1E developed at the Konrad Zuse Zentrum für Informations-technik Berlin / 39/. As unknowns, $\rho_1, \rho_2, k$ and $N_4$ are chosen. For a fixed of $n_0=1.637$, we obtain a family of solutions determined by the value of the remaining free parameter, the lens thickness.

Figure 4.1.11 shows the change of astigmatism as function of $d$. It can be seen that as long as $d < 3f$, the astigmatism is nearly constant and has approximately the value 1 predicted by the thin lens theory. For larger values of $d$, it decreases slowly and becomes 0 for $d=9.16f$. The exact lens parameters in this case have been already given in Sec. 3.5.4. In this case, in addition to astigmatism, the distortion also becomes zero. Thus, this lens is corrected for all Seidel coefficients. Moreover, this correction state is now independent of the stop position. Several lenses have been found at $\beta=-1$ where all five Seidel coefficients are zero. All of them are very thick.

From the examples given in this section, it can be concluded that, even if the extended thin lens approximation for radial GRIN lenses does not provide accurate values for the absolute magnitude of the aberrations, it gives nevertheless, for lenses which are not too thick, a good qualitative description of the variation of these aberrations when the lens parameters are changed. Unlike the exact expressions, in the extended thin lens approximation the total aberrations of a lens can be directly related to the lens parameters by relatively simple formulae. Thus, at the expense of quantitative accuracy, additional insight in the correction properties and their limitations has been obtained.
Fig. 4.1.11 Astigmatism as function of the lens thickness if spherical aberration, coma and Petzval curvature are zero
4.2 Thin-lens approximation for axial gradient-index lenses

In this section we develop the thin lens approximation for axial GRIN lenses. From an analysis of the paraxial transfer equations and of the transfer contributions to the primary aberration coefficients we find that, within the domain of validity of this approximation, the primary aberrations of an axial GRIN lens with spherical end surfaces are equivalent to those for a pair of homogeneous aspherical lenses in contact, having a common plane surface.

As has been first shown by Sands /46/, the inhomogeneous surface contributions of axial GRIN lenses to the Seidel coefficients (Eqs. (3.3.36)) are equivalent to those of aspherical surfaces. (However, as will be seen later in this section -Eqs. (4.2.10-11), while the aspheric surface contributions are curvature independent, the inhomogeneous contributions are proportional to the square of the surface curvature.) Thus, it is possible to redesign lens systems with aspherical lenses by replacing the aspherical surfaces by axial gradient materials /32/.

The transfer contributions, however, have not yet been discussed adequately in the literature. As long as the total axial refractive index change $\delta n$ is small, as usually in the case of that produced by ion exchange (see Chapter 5) the effect of the transfer contributions is in most cases considerably smaller than that of the inhomogeneous surface contributions. Therefore, for qualitative discussions of the effect of the axial gradient, the transfer contributions have simply been neglected. However, recently developed fabrication methods of axial gradients by controlled fusion of thin layers of glass can produce large index changes. (Values up to $\delta n \sim 0.5$ have been reported. /5/) In what follows, we examine the effect of the transfer through an axial GRIN medium on the power and on the primary aberration coefficients within the frame of the thin lens approximation.

We have seen in Chapter 3 that for calculating the Gauss matrix for transfer and the transfer contributions to the primary aberrations, three quadratures of the refractive index data must be performed (Eqs. (2.2.21),(3.4.10),(3.6.25)). We first evaluate the order of magnitude of these three quantities. Since all these quantities can be regarded as average values over the lens thickness of given functions of gradient material properties, their order of magnitude is the same as that for the corresponding function at any of the lens surfaces. By denoting the values of $n_0$ at the two end surfaces by $n_{01}$ and $n_{02}$, we have for instance
In the paraxial approximation, the change of the ray height inside the lens (see Eq. (2.2.22)) is of the order of 
\[ \delta h \approx \frac{\delta n(z)}{n^2(z)} \, dz \sim \frac{\delta n_0}{n_0^2} \quad (4.2.1) \]

In the paraxial approximation, the change of the ray height inside the lens (see Eq. (2.2.22)) is of the order of 
\[ \Delta h = -d \int n^{-1}(n_0 u \sim -ud) \quad (4.2.2) \]

where we recall that \( n_0 u = n_0 u' \). As in the homogeneous case, in the thin lens approximation for axial gradients we assume \( ud<<h \). We neglect the change of the ray height inside the lens by setting \( d=0 \) in Eqs.(2.2.22) and also in the Gauss matrix for transfer given by Eq.(2.2.23). Thus, in the thin lens approximation, the Gauss matrix for transfer becomes the unity matrix as in the case of homogeneous media, i.e. the transfer through an axial gradient does not contribute to the lens power.

The net effect of the gradient in the paraxial approximation is simply that \( n_0 \neq n_0' \). The lens power is 
\[ \varphi = \varphi_1 + \varphi_2, \varphi_1 = (n_0 - 1)\rho_1, \varphi_2 = -(n_0' - 1)\rho_2 \quad (4.2.3) \]

We now consider the transfer contributions to the primary aberrations. For spherical aberration, we have according to Eq.(3.4.14)
\[
T_i = n_0^3 u^3 \left[ n_{01}^{-3} \left\langle u d + \frac{h - \int n^{-1}(n_0 u d}{h} \right\rangle \frac{n_0 u d}{n_0 n_0'} - h \right] = n_0^3 u^3 \left[ n_0^{-1} \left( \frac{1}{n_0^2} - \frac{1}{n_0' n_0} \right) - u d \left( \frac{1}{n_0^2} \right) n_0^{-1} \left( n_0 - n_0' \right) n_0^{-3} \right] \quad (4.2.4)
\]

According to Eq. (4.2.1), the second term in the square bracket of Eq.(4.2.4) has the order of magnitude
\[ -u d \left( \frac{1}{n_0^2} \right) n_0^{-1} \left( n_0 - n_0' \right) n_0^{-3} \approx -u d \left( \frac{1}{n_0^2} \right) n_0^{-1} \left( n_0 - n_0' \right) n_0^{-3} \quad (4.2.5) \]

and therefore the ratio between the second and the first term in the square bracket is \( \sim -u d/h \). Thus, setting \( d=0 \) also in the formulae for the Seidel aberrations is consistent with the usual thin lens approximation. We find
\[ T_i = n_0^3 u^3 h \left( \frac{1}{n_0^2} - \frac{1}{n_0'} \right) \quad (4.2.6) \]

The remaining Seidel coefficients can be derived directly from Eq (3.4.15). Similarly, from Eq.(3.6.24) we find for axial color
\[ T_{\lambda 1} = n_{0_1} u \left[ n_{0_1} \delta_\lambda \right] n^{-1} u d - \left( \frac{\delta_\lambda n_{0_2}}{n_{0_2}} \left( \delta_\lambda n_{0_2} \frac{n^{-1}}{n_{0_1}} \right) - \frac{\delta_\lambda n_{0_1}}{n_{0_1}} h \right] = \]

\[ = -n_{0_1} u \left[ h \left( \frac{\delta_\lambda n_{0_2}}{n_{0_2}} - \frac{\delta_\lambda n_{0_1}}{n_{0_1}} \right) - u d \left( \frac{\delta_\lambda n_{0_2}}{n_{0_2}} n^{-1} \left( n_{0_1} + n_{0_1} \delta_\lambda \right) n^{-1} \right) \right] \]  

(4.2.7)

Since

\[ -u d \left( \frac{\delta_\lambda n_{0_2}}{n_{0_2}} n^{-1} \left( n_{0_1} + n_{0_1} \delta_\lambda \right) n^{-1} \right) \sim -u d \left( \frac{\delta_\lambda n_{0_2}}{n_{0_2}} - \frac{\delta_\lambda n_{0_1}}{n_{0_1}} \right) \]  

(4.2.8)

we can again neglect this term by setting \( d=0 \) in Eq. (4.2.7) and obtain thus

\[ T_{\lambda 1} = -n_{0_1} u h \left( \frac{\delta_\lambda n_{0_2}}{n_{0_2}} - \frac{\delta_\lambda n_{0_1}}{n_{0_1}} \right) \]  

(4.2.9)

It can be observed from the above formulae that, in the thin lens approximation, only gradient material properties at the two end surfaces appear in the expressions of the transfer contributions to the primary aberration coefficients. Thus, the effect of transfer through the axial GRIN medium upon the aberrations of the system does not depend on the specific form of the \( n=n(z) \) inside the medium.

Consequently, the effects of the transfer terms are for an arbitrary refractive index distribution precisely the same as in the case of an "axial" gradient consisting of two homogeneous media, separated by a plane surface (i.e. \( n=n(z) \) is a step function) or, in other words, the same as the effects of a plane surface separating two homogeneous media. The exact expressions for the primary aberration coefficients in the latter case have been derived in Secs. 3.4.3 and 3.6.3. It can be verified immediately that by replacing the indices \( A \) and \( B \) by \( 0_1 \) and \( 0_2 \) and by setting \( h_{AB}=h \), the equations (3.4.24) and (3.6.28) become Eqs.(4.2.6) and (4.2.9).

Consider now all contributions to the primary aberrations of an axial thin lens. By writing the aspherical contribution to the spherical aberration of the first surface of an homogeneous asphere as

\[ S_{\theta} = 8 \kappa h^4 (n-1) \]  

(4.2.10)

we conclude that, within the frame of the thin lens approximation, the primary aberrations of an axial GRIN lens are equivalent to those of a pair of homogeneous aspherical lenses in contact, having a common plane surface and
\[ n_1 = n_{01}, \kappa_1 = \frac{\rho^2 N_{z,1}}{8(n_1 - 1)} \]  
\[ n_2 = n_{02}, \kappa_2 = \frac{\rho^2 N_{z,2}}{8(n_2 - 1)} \]  

(4.2.11)

and having Abbe numbers \( \nu_1 \) and \( \nu_2 \) identical with the values corresponding to \( n_{01} \) and \( n_{02} \).

Thus, the design possibilities and limitations for thin axial GRIN lenses are precisely the same as for the equivalent aspherical pair. Thin-lens aberration expressions for each of the two equivalent homogeneous aspheres can be obtained by inserting \( k=0 \) in the corresponding expressions for radial GRIN lenses derived in Sec. 4.1.2 and considering as an additional term for spherical aberration \( A^* = S_1^* \) given by Eq.(4.2.10-11). Results for the total primary aberrations of the equivalent aspherical pair can then be obtained as shown there for systems of thin lenses. For instance, if the stop is at the axial GRIN lens, the astigmatism is as in the homogeneous and radial gradient cases

\[ \Gamma_3 = H^2(\phi_1 + \phi_2) = H^2 \phi \]  

(4.2.12)

Note also that, if the Abbe numbers \( \nu_1 \) and \( \nu_2 \) at the end surfaces differ significantly, then an achromatic axial-gradient singlet can be obtained. The powers of the end surfaces (Eqs.(4.2.3)) then satisfy the condition

\[ \frac{\phi_1}{\nu_1} + \frac{\phi_2}{\nu_2} = 0 \]  

(4.2.13)
4.3 Simple optical systems with gradient-index lenses

4.3.1 Aplanatic correction of a single radial gradient-index lens

In this section, two simple examples of aplanatic correction of single radial GRIN lenses are given where the thin lens formulae derived in Sec. 4.1.2. are used to explain features of solutions optimized by finite ray tracing.

Consider a radial GRIN lens having the stop at the first surface. If the lens is to be used for imaging at a finite aperture, but small field radius, then the aberration types that must be corrected are spherical aberration and coma. If all parameters of the refractive index distribution given by Eq. (2.3.6) can be regarded as independent variables, then correction of spherical aberration can be easily accomplished in all orders. Thus, for $p=4,6,8\ldots$, the parameter $N_p$ in the refractive index distribution can be used to correct the spherical wave aberration of order $p$ irrespective of the values of the other lens parameters. As can be seen from Eqs. (4.1.46-48), for a fixed value of the total lens power, third-order coma can be corrected either by homogeneous bending (the variable $\varpi$) or by gradient bending (i.e. by changing the ratio between $\varphi_h$ and $\varphi_g$). See also Figs. 4.1.1 and 4.1.6 in Sec. 4.1.3.

As a first example of aplanatic correction, in the case $\beta=-1$ thin lens theory, exact Seidel aberrations and optimization with finite ray tracing are successively applied.

Let us first see what relationships between lens parameters follow from the thin lens theory. For $\beta=-1$, the thin lens equations become particularly simple. It follows from Eqs. (4.1.26) that

$$\varpi = 0, \quad \vartheta = \rho_1 + \rho_2$$

(4.3.1)

Consider first coma correction. As can be seen from Eqs. (4.1.46-48), coma vanishes either if $\varpi=0$, irrespective of the values of $\varphi_h$ and $\varphi_g$, or if $\varphi_h$ and $\varphi_g$ are such that $B_g=-B_h$, irrespective of the value of $\varpi$. The first type of solutions have $\rho_1=-\rho_2$, i.e. these are the symmetric solutions described in Sec. 3.1.1, where coma, distortion and lateral color vanish because of symmetry.

The solutions of second type are asymmetric and, for a fixed value of $\varphi$, expressions for the values for $\varphi_h$ and $\varphi_g$ can be easily found from the condition $B_g=-B_h$. Solutions of asymmetric type can be obtained also at the exact Seidel level, i.e. when the exact third-order spherical
aberration and coma are simultaneously anulled. However, at the level of finite ray-tracing, in a search for local minima performed with CODE V (described in more detail below), all solutions which have been found turned out to be of the symmetric type. This is probably because the asymmetric solutions seem to have very large higher order aberrations. For lens data corrected at the exact Seidel level, tracing several finite rays both in symmetric and in asymmetric solutions (with the analytic ray tracing method described in Sec 2.4) showed indeed that the same rays suffer in the asymmetric case much larger aberrations than in the symmetric case. In what follows, only the symmetric solutions will be further analysed.

Assuming that the focal length is kept fixed, \( f=1 \), we consider now the correction of third-order spherical aberration in the case of the symmetric solutions. As mentioned above, this can be done by adjusting the value of \( N_4 \). However, since the lens has more free parameters than conditions to be fulfilled, we obtain an infinite number of solutions. Let us now see how the value of \( N_4 \) which corrects spherical aberration depends on the values of the remaining lens parameters.

By substituting in Eqs. (4.1.42-45)
\[
\psi = 0, \varphi = 0, \varphi_h = n_0 kd, \varphi_h = 1 - n_0 kd
\]
and keeping only terms up to the first order in \( d \), we obtain

\[
A_h = \frac{h^4 n_0^2 (1 - 3 n_0 kd)}{4(n_0 - 1)^2}, \quad A_g = \frac{h^4 n_0 kd (3 n_0 + 1)}{4(n_0 - 1)}
\]

\[
A_0^* = -\frac{2 h^4 n_0 k (1 - n_0 kd)}{(n_0 - 1)} \quad A_1^* = h^4 n_0 d k^2 (1 - 4 N_4)
\]

Solving the equation

\[
A_h + A_g + A_0^* + A_1^* = 0
\]

with respect to \( N_4 \) yields after some algebra

\[
N_4 = \frac{3 n_0 - 1}{4(n_0 - 1)} \frac{k [d(2 n_0 + 1) + 8(n_0 - 1)] - n_0}{16 d k^2 (n_0 - 1)^2}
\]

The condition \( f=1 \) yields

\[
\rho_1 = -\rho_2 = \frac{1 - n_0 kd}{2(n_0 - 1)}
\]

Since the solutions (4.3.5-6) are functions of three lens parameters, for better insight we assign fixed values for two of them, e.g.

\[
n_0 = 1.5, \quad d = 0.4
\]
As a second step, the aplanatic correction has been investigated at the exact Seidel level. For the fixed values (4.3.7) and 0.2<k<2, the three nonlinear equations resulting from annulling spherical aberrations and coma and keeping the focal length unchanged have been solved numerically using the program already mentioned at the end of Sec. 4.1.3. Even if, with the stop at the first surface, the optical system is not perfectly symmetric because of the finite thickness of the lens, the solutions still have ρ₁=−ρ₂. This is because, with zero spherical aberration, coma does not depend on the stop position and has the same value as in the case of a symmetric stop position. Thus, for this type of solutions, correcting coma requires ρ₁=−ρ₂.

Finally, solutions corrected at the exact Seidel level have been used as starting configurations for optimization with finite ray tracing. Using CODE V, for the fixed values (4.3.7) four local minima have been found. Figure 4.3.1 shows one of the solutions. (The other three are similar.) The parameters of the four solutions are given in Tabs. 4.3.1-2. Even if the lens parameters (especially the refractive index distribution) differ considerably in the four cases, it can be observed that the error function computed by CODE V and the Seidel aberrations are nearly the same. As expected, spherical aberration and coma have small values, while astigmatism is ~0.7-0.8. (The values of Seidel aberrations are normalized to H=1.) Note that the relation ρ₁=−ρ₂ holds fairly well also for the optimized solutions.

In Figs. 4.3.2-3 the values of ρ₁ and N₄ as functions of k are shown at the three levels of approximation described above. As could be expected for the moderate values of the chosen aperture and field, the four CODE V solutions are situated very close to the continuous Seidel curve. However, it can be observed that the shape of the exact Seidel curve is described remarkably well by the thin lens formulae (Eqs.(4.3.8-9)). The shift of the exact Seidel curves with respect to the thin-lens curves is due to the fact that d=0.4 while f=1, i.e. d is not very small.
Fig. 4.3.1 Aplanatic radial GRIN lenses obtained with CODE V (See Tabs. 4.3.1-2)

\[
d = 0.4 \quad f = 1 \quad \text{Numerical aperture } = 0.2
\]
\[
n_0 = 1.5 \quad \beta = -1 \quad \text{Maximal object height } = 0.1
\]

Tab. 4.3.1 Common parameters of the four solutions

<table>
<thead>
<tr>
<th>solution</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_1 )</td>
<td>1.090</td>
<td>1.117</td>
<td>1.318</td>
<td>1.673</td>
</tr>
<tr>
<td>( R_2 )</td>
<td>-1.088</td>
<td>-1.117</td>
<td>-1.327</td>
<td>-1.679</td>
</tr>
<tr>
<td>( k )</td>
<td>0.263</td>
<td>0.301</td>
<td>0.529</td>
<td>0.789</td>
</tr>
<tr>
<td>( N_4 )</td>
<td>5.433</td>
<td>2.946</td>
<td>-0.787</td>
<td>-0.756</td>
</tr>
<tr>
<td>( N_6 )</td>
<td>-21.60</td>
<td>-17.35</td>
<td>-6.444</td>
<td>-2.556</td>
</tr>
<tr>
<td>( N_8 )</td>
<td>1253.</td>
<td>464.8</td>
<td>-24.76</td>
<td>-2.555</td>
</tr>
<tr>
<td>S.A.</td>
<td>-0.186</td>
<td>-0.168</td>
<td>-0.114</td>
<td>-0.125</td>
</tr>
<tr>
<td>coma</td>
<td>-0.025</td>
<td>-0.022</td>
<td>-0.012</td>
<td>-0.015</td>
</tr>
<tr>
<td>astigmatism</td>
<td>0.690</td>
<td>0.698</td>
<td>0.740</td>
<td>0.782</td>
</tr>
<tr>
<td>Petzval curv.</td>
<td>0.682</td>
<td>0.677</td>
<td>0.645</td>
<td>0.608</td>
</tr>
<tr>
<td>distortion</td>
<td>0.366</td>
<td>0.367</td>
<td>0.376</td>
<td>0.386</td>
</tr>
<tr>
<td>error function</td>
<td>0.279</td>
<td>0.281</td>
<td>0.286</td>
<td>0.292</td>
</tr>
</tbody>
</table>

Tab. 4.3.2 Parameters and Seidel aberrations of the four solutions
Fig. 4.3.2 Aplanatic correction - thin lens theory, exact Seidel formulae and optimization with finite ray tracing: $\rho_1$ as function of $k$

Fig. 4.3.3 Aplanatic correction - thin lens theory, exact Seidel formulae and optimization with finite ray tracing: $N_4$ as function of $k$
For practical purposes, the thin lens formulae (in this example Eqs. (4.3.5-6)) could thus be of considerable value by providing insight in the approximate relationships between the parameters of the lens to be fabricated and its aberrations. Since the presently known manufacturing technologies impose certain restrictions on the shapes of the refractive index distributions that can be produced, thin lens theory could be used in various cases as a first step in investigating the feasibility of the lens.

Single aplanatic radial GRIN lenses have been designed as compact-disk objectives /\(^{38}\), /\(^{21}\), /\(^{63}\). For autofocussing and automatic track search in the CD system, the objective should have a weight as small as possible. Therefore, the single GRIN lens objectives are more convenient than equivalent homogeneous objectives consisting of three spherical lenses, being an alternative to the CD objectives with aspherical singlets.

The CD lenses are infinite conjugate systems (\(\beta=0\)). Often, one of the end surfaces is plane (usually the second one). The parameters and Seidel aberrations of an aplanatic GRIN lens of this type (fabricated by Nippon Sheet Glass Co. /\(^{38}\)) have already been given in Tabs. 4.1.1-2. Let us now explain some features of this lens by means of thin lens theory.

We consider first coma correction, assuming that the stop is at the lens. Since \(\beta=\rho_2=0\), it follows from Eqs. (4.1.26) and (4.1.8) that

\[
\psi = \varphi_h \cdot \varphi_g = \varphi_h \frac{2-n_0}{n_0-1}
\]  

(4.3.10)

Assuming \(\varphi=1\) and substituting \(\varphi_h=1-\varphi_g\), the condition \(B_h+B_g=0\) (see Eqs. (4.1.47-48)) yields after some algebra

\[
\varphi_g = \frac{\varphi_h(2-n_0)}{n_0-1}
\]  

(4.3.11)

Thus, in the thin lens approximation, the distribution of power between the first surface and the gradient medium necessary for coma correction is determined only by \(n_0\). One of the solutions of the equation

\[
n_0^2 - n_0 - 1 = 0
\]  

(4.3.12)

is

\[
n_0 = \frac{1}{2} (1 + \sqrt{5}) = 1.61803
\]  

(4.3.13)

Thus, for this value of \(n_0\), zero coma requires \(\varphi_g=0\). We observe at this point that the optimized lens has \(n_0=1.58\), i.e. a value close to 1.618. The equation (4.3.11) then gives
\( \varphi_g = 0.0366 \), a value which is negligible in comparison with unity. Considering now the correction of spherical aberration, it follows that the transfer contributions to the thin-lens spherical aberration \( A_0' \) and \( A_1' \), which are proportional to \( \varphi_g \), must therefore be small in comparison to the surface contributions \( A_h \) and \( A_0' \) (see Eqs. (4.1.41-45)) and can at this stage of analysis be neglected.

For simplicity, assume for the moment that \( n_0 \) is given by Eq. (4.3.13). Using Eqs. (4.3.10-13), correction of spherical aberration requires that

\[
A_h + A_0' = h^4 \left( 1 + \sqrt{5} \right) / 2 - h^4 k \left( 3 + \sqrt{5} \right) = 0
\]

and therefore

\[
k = \left( \sqrt{5} - 1 \right) / 4 = 0.309016
\]

Finally, the radius of the first surface is

\[
R_i = (n_0 - 1) / \varphi_g = n_0 - 1 = 0.618033
\]

If \( \varphi_g = 0 \), all transfer contributions of the medium to Seidel aberrations vanish. Thus, the above correction of spherical aberration is independent of \( N_4 \), because \( A_1' \), the only term containing \( N_4 \), vanishes in this case (see Eq.(4.1.45)). Note, however, that for nonzero \( k \) the condition \( \varphi_g = 0 \) requires \( d = 0 \). Since a real lens has nonzero thickness, the actual value of \( \varphi_g \) becomes also nonzero.

At this point, the limits of the above thin-lens analysis should be recalled. First, as shown in Sec. 4.3.1, for lenses with nonzero thickness, the thin-lens aberration curves have nearly the same shape as the exact Seidel curves, but are shifted with respect to them. Therefore, the values of the lens parameters for which thin lens aberrations are zero differ from the values for which the exact Seidel aberrations vanish. Secondly, in the optimized design, aberrations of various orders are balanced against each other and therefore, for optimal image quality, third-order spherical aberration and coma are not exactly zero.

For comparison, in Table 4.3.3 the thin-lens parameters (4.3.13) and (4.3.15-16) are shown together with the parameters of the optimized design (i.e. the values of Tab. 4.1.1). In the optimized design where \( d = 0.416 \) we have according to Eq. (4.1.9) \( \varphi_g = 0.219 \), a value which
Tab. 4.3.3 Aplanatic plano-convex radial GRIN lens as CD objective
a) thin lens design, b) optimized design / 38/

Fig. 4.3.4 Distributions of the Seidel aberrations between surfaces and medium corresponding to the parameters of Tab 4.3.3. In case b, the main difference is the transfer of spherical aberration from the first surface to the medium. (These contributions are marked by arrows)

is still small in comparison with unity, but which cannot be neglected any more. Consequently, the transfer through the medium also contributes to the Seidel aberrations.

The distributions of the Seidel aberrations between surfaces and medium in the thin lens design and in the optimized one are shown in Figs. 4.3.4a-b. In the second case, the main differences are the contribution of the medium to spherical aberration and a decrease of the corresponding contribution at the first surface. The remaining aberration contributions show in both cases nearly the same patterns. Thus, the fact already noted in Sec. 4.1.3 that for spherical aberration the largest contribution is the inhomogeneous contribution of the first surface, is described very well by thin lens theory.

The optimal values of $d$ and $N_4$ for reducing higher order aberrations can be determined by optimization with finite ray tracing. However, since the value of $d$ determines the distribution of power between the first surface and the medium, $d$ can be also used to achieve the right amount of sensitivity of the imaging quality at changes in the refractive index profile.
Thus, it follows from Eq. (4.1.45) that a too small value of $\varphi_g$ would cause $N_4$ to be less effective as a design parameter. (As noted above, for $\varphi_g = 0$ the Seidel aberrations are independent of $N_4$.) On the other hand, if $\varphi_g$ is too large, small changes of $N_4$ would have a large influence on the spherical aberration, thus leading to severe tolerances on the refractive index profile during the manufacturing process.

### 4.3.2 Correction of all Seidel aberrations with two gradient-index lenses

We have seen previously that with a single thin radial or axial GRIN lens it is not possible to correct all Seidel aberrations simultaneously. For instance, if the stop is at the lens, or if the lens is already corrected for spherical aberration and coma, then astigmatism cannot be corrected. However, as has been already shown in the literature, correction of all Seidel aberrations becomes possible with two GRIN lenses separated by an air space. Pairs of GRIN lenses with zero Seidel aberrations have lead to designs of high-quality photographic objectives which illustrate the possibility of reducing the number of elements as compared to homogeneous systems with the same specifications /2/, /3/.

Consider first, at $\beta = -1$, the case of two identical radial GRIN lenses arranged in a symmetric way with respect to the stop. In this case, both the Seidel aberrations and the paraxial chromatic aberrations can be simultaneously corrected. As shown in Sec. 3.1.1, a symmetrical optical system is automatically corrected for coma, distortion and lateral color and the remaining aberrations (spherical aberration, astigmatism, Petzval curvature and axial color) are also corrected if they are corrected for a half system. Thus, in what follows the aberration correction of the back half of the system at $\beta = 0$ will be discussed.

For correcting astigmatism, the stop must be situated at some distance from the back half lens. It will be immediately seen that the number of free parameters is larger than the numbers of conditions to be fulfilled. For instance, the lens thickness and the stop position can be kept fixed. The Petzval curvature and axial color can be corrected by choosing for $k$ (i.e. for $\varphi_g$) and $\nu_g$ the values that satisfy the plan-achromat condition (Eqs. (4.1.67-68)). As in the first example of the previous section, spherical aberration can be corrected with the coefficient $N_4$. The third of Eqs. (4.1.62) shows that correction of astigmatism can be achieved by controlling the back-half coma (the parameter $B$). As can be seen from Eqs. (4.1.46-48) the required value of $B$ for annulling astigmatism can be achieved by homogeneous bending (the parameter $\varpi$). Finally $\varphi_h$ can be adjusted in order to obtain for the whole system the
required focal length. Thus, all primary aberrations of the symmetric system have been corrected.

\[ R_1 = -0.417 \quad d = 0.5 \quad k = 3.387 \quad \beta = 0 \]
\[ R_2 = 1.012 \quad n_0 = 1.5 \quad N_4 = 1.036 \quad f = 1 \]

Tab. 4.3.4 Parameters of the back-half lens in a symmetric system corrected for all Seidel aberrations. The stop is situated at the distance -0.2 in front of the first surface.

<table>
<thead>
<tr>
<th></th>
<th>Spherical aberration</th>
<th>Coma</th>
<th>Astigmatism</th>
<th>Petzval curvature</th>
<th>Distortion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinary 1</td>
<td>-3.066</td>
<td>0.665</td>
<td>-0.144</td>
<td>-0.799</td>
<td>0.205</td>
</tr>
<tr>
<td>Inhomog. 1</td>
<td>24.367</td>
<td>4.873</td>
<td>0.975</td>
<td>0</td>
<td>0.195</td>
</tr>
<tr>
<td>Medium</td>
<td>-29.53</td>
<td>-10.543</td>
<td>-3.255</td>
<td>1.129</td>
<td>-0.426</td>
</tr>
<tr>
<td>Ordinary 2</td>
<td>0.001</td>
<td>-0.02</td>
<td>0.344</td>
<td>-0.329</td>
<td>1.047</td>
</tr>
<tr>
<td>Inhomog. 2</td>
<td>8.227</td>
<td>4.138</td>
<td>2.081</td>
<td>0</td>
<td>0.776</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
<td>-0.887</td>
<td>0</td>
<td>0</td>
<td>0.776</td>
</tr>
</tbody>
</table>

Tab. 4.3.5 Seidel aberrations corresponding to the parameters of Tab. 4.3.4

As an example, the parameters of the back-half lens in a symmetric system corrected for all Seidel aberrations are given in Tab. 4.3.4. For simplicity, the back half is normalized to \( f = 1 \). For fixed values of thickness and stop position, the values for the two radii and for the variables \( k \) and \( N_4 \) have been obtained by solving numerically the four nonlinear equations resulting from annulling spherical aberration, astigmatism and Petzval curvature and keeping \( f = 1 \). The corresponding Seidel aberrations are given in Tab. 4.3.5.

For the first time, symmetric correction with two radial GRIN lenses has been discussed in Ref. / 2/. A system thus corrected has been chosen there as starting point for the design of a photographic objective having the same specifications as a homogeneous design consisting of six elements. It turned out that, by giving up exact symmetry between the two lenses, correction of all Seidel aberrations is possible also at \( \beta = 0 \). However, in this case paraxial axial and lateral color cannot be corrected simultaneously and some compromise must be made. The third-order design thus obtained has been optimized with ray tracing. It has been found that the resulting GRIN photographic objective has nearly the same performances as the
standard homogeneous system. However, the required number of elements has been reduced from six to two.

Correction of all Seidel aberrations is possible also with two axial GRIN lenses. This can be easily shown by means of the thin lens formulae. Assume that the total axial change of the refractive index is small, such that the effect of transfer can be neglected. Then, each of the two axial GRIN lenses is equivalent with an homogeneous aspherical lens. The expressions of the Seidel aberrations for each thin GRIN lens can be obtained by considering in the expressions for radial GRIN lenses derived in Sec. 4.1.2 only the homogeneous parts (by inserting k=0 in these expressions) and by considering an additional aspheric-like term for spherical aberration. Assume that the gradient region is limited to the vicinity of one of the surfaces (i.e. the gradient is not so deep as to reach the other surface). If the gradient is at the first surface is follows from the first of Eqs. (3.3.36) that the aspheric-like term is

$$A^* = S_1^* = h^4 \rho_1^2 N_z$$

(4.3.17)

A similar expression is valid at the second surface, but with opposite sign. Recall that $N_z$ is the first derivative of $n(z)$ at the surface (see Eq. (3.3.4)).

Let $\phi_1 = -\phi_2$ and let the average index have the same value for both lenses. Thus, total Petzval curvature vanishes. If $d_{12}$ is the distance between the two lenses, then the power of the system is

$$\phi = \phi_1 + \phi_2 - d_{12} \phi_1 \phi_2 = d_{12} \phi_1^2$$

(4.3.18)

Annulling the remaining four Seidel aberrations is possible by means of the so called "antisymmetric correction" /23/. Therefore, choose for the two lenses equal values of B and values having equal magnitudes, but opposite signs for A and for the ratio m/h. In this case, it can be easily shown that spherical aberration and astigmatism vanish automatically. Requiring zero coma and distortion yields a linear system of two equations with unknowns A and B. The solutions are given in Ref. /23/. For each lens, the required value of A can be obtained by controlling $N_z$, while that of B can be obtained by bending. Thus, all Seidel aberrations can be corrected.

The correction of all Seidel aberrations in a system with two axial GRIN lenses has been first described in Ref. /3/. Such a system has been used as starting point for the design of a wide angle photographic objective. Choosing $\phi_1 < 0$ leads to a reverse telephoto design (back focal length greater than the effective focal length), which is convenient for wide angle objectives. At both lenses, the gradient has been placed at the surfaces having the larger curvature. (The effect of a given $N_z$ is proportional with the square of the curvature - see Eq.(4.3.17)) While
correction of all Seidel aberrations could be achieved with the two axial GRIN lenses, correction of higher order aberrations and of chromatic aberrations required a third homogeneous lens. After optimization, two different configurations of three- lens designs with two axial gradients showed almost the same performances as an homogeneous objective consisting of six lenses.
5. Fabrication of gradient-index media

As shown previously, if the parameters of the refractive index distribution can be used as free design parameters, then GRIN lens designs of imaging systems can be obtained which have a much smaller number of elements than equivalent homogeneous designs. However, the range of the producible values of these parameters is determined by the fabrication method used. At present, the application of gradients to imaging systems is limited by fabrication capabilities. Therefore, considerable effort is spent for the development of fabrication methods for achieving the variety of gradient shapes and depths encountered in the designs /37/. In what follows, three of the most important fabrication methods of GRIN lenses will be briefly described. A description of other methods can be found in Ref. /33/.

At present, the most widely used technique for fabricating GRIN glass is ion exchange in a molten salt. Therefore, the glass is immersed in a molten salt bath containing cations that differ from those of the glass. By a diffusion process, cations from the glass are exchanged with those of the salt bath. Diffusion in glass is most rapid when the temperature of the salt bath is in the range where the material makes a transition from the glassy state to that of a highly viscous state, but below the melting range. Keeping the glass and molten salt in contact for many hours, the cation exchange creates a concentration gradient of both types of cations within the glass which leads to a refractive index gradient. After the glass is withdrawn from the salt bath and cooled, the diffusion process is stopped and the resulting refractive index profile, extending a certain distance (depth) from the surface into the glass, becomes permanent /25/. The profile shape can then be further adjusted by post-annealing.

The parameters that affect the ion exchange process and the resulting refractive index distribution are the choice of the glass-salt system, effective diffusion time, temperature, glass geometry, use of an electric field and anneal cycle. At present, it is difficult to obtain changes in refractive index larger than 0.08 and gradient depths larger than 5 mm /45/ while the set of profile shapes is also predetermined. Thus, the limited gradient depth imposes severe restrictions on the diameter of radial GRIN materials.

More recently, an alternative fabrication method of GRIN materials based on the sol-gel process has been developed. Sol-gel processing is a low-temperature chemical synthesis process for the fabrication of glasses and other materials enabling a high degree of composi-
tional control. The hydrolysis of a solution containing the index-modifying ions as a dopant produces the sol, an even distribution of small clusters of molecules or colloids. Through loss of fluidity, the prepared sol changes to a porous jelly-like mass, the gel. The dopant is then partially leached out, e.g., by means of an acid (which happens very quickly because the gel is porous). Finally, the GRIN glass is obtained by drying and sintering the gel. Sol-gel processes for fabricating GRIN glasses can have more degrees of freedom than ion diffusion because there is potential to control either the diffusion, chemical reactions, or the interaction between the two. This process has advantages for the fabrication of radial GRIN materials of larger diameters.

Finally, for axial gradients a fabrication method by controlled fusion of thin layers of glass has been developed. Large axial changes of the refractive index up to ~0.5 have been obtained for a variety of lens sizes and shapes. The layers may be powder mixtures or solid plates of glasses with compatible chemistry that span the desired refractive index range. During fusion and subsequent heat treatment, the interfaces between layers are eliminated and a gradual transition in refractive index and other properties occurs.

Since the benefits of using GRIN materials are already well known, it is reasonable to expect that these materials will be used in an increasing number in optical imaging systems as soon as suitable materials become available.
6. Summary

In this study, new results related to the theoretical foundations underlying lens design with radial and axial gradient-index lenses have been obtained.

A new method to derive algebraic formulae for tracing skew rays in radial GRIN lenses is described. The ray position and direction, optical path length and the coordinates of the ray-surface intersection points are given by power series with coefficients determined by successive approximations. Thus, for the first time, the entire ray-tracing process can be performed analytically. The number of terms that have to be computed in these series in order to achieve a prescribed accuracy is determined by the sizes of aperture and field. In all formulae, expressions of the terms up to the eighth order have been obtained.

A new method for deriving aberration coefficients for rotationally symmetric gradients on the basis of analytical ray tracing has been developed. For radial GRIN lenses, for the first time short and accurate analytic formulae have been obtained for all Seidel and chromatic paraxial aberration coefficients. For deriving expressions of the Seidel aberration coefficients, the two components of the transverse aberration of an arbitrary skew ray are decomposed in surface and transfer contributions. The large expressions for the third order terms resulting from the transfer contributions are shortened by means of a heuristic symmetrization technique.

For radial and axial GRIN lenses, generalizations of the homogeneous thin lens approximation have been developed which provide additional insight into aberration correction with gradients. It is shown that, as in the homogeneous case, within the domain of validity of these approximations it is not possible for a single GRIN lens having nonzero power to correct all Seidel aberrations simultaneously. For radial GRIN lenses, thin-lens expressions for the primary aberrations having nearly the same structure as in the homogeneous case have been obtained. On the basis of these formulae, features of aplanatic lenses optimized by finite ray tracing are explained. For axial gradients it is found that the primary aberrations of a thin GRIN lens are equivalent to those of a pair of homogeneous aspherical lenses in contact, having a common plane surface and having refractive indices and Abbe numbers equal to the corresponding axial GRIN values at the two end surfaces.
Appendix A  Eighth-order analytic ray-tracing formulae for radial GRIN lenses

A.1 Formulae for ray position, direction and optical path length

We list here all necessary formulae for programming an eighth-order analytic ray-tracing routine giving the position, direction and optical path length of finite skew rays inside radial GRIN media. The formulae for curved end faces are given in Sec. (A.2).

A.1.1 Summary of previously introduced formulae

The formulae already introduced in Chapter 2. are given with references to the original numbering.

- Refractive index distribution:

  \[ n^2(r^2) = n_0^2 \left( 1 - kr^2 + N_4 r^4 + N_6 r^6 + N_8 r^8 + \ldots \right) \]  

  \[ (2.3.6) \]

- Independent variable:

  \[ t = \frac{n_0 g z}{\xi_0} \]  

  with \( \xi_0 = n_0^2 \left( x_0^2 + y_0^2 \right) \)\( - \eta_0^2 \)  

  \[ (2.3.8) \] \( (2.3.13) \] \( (2.3.9) \]

- Quantities appearing in the formulae for ray position, direction and optical path length:

  \[ \mu = \frac{1}{2} \left( 1 - \frac{\xi_0^2}{n_0^2} \right) \]  

  , \( a = \frac{1}{\mu} \frac{g^2 (x_0^2 + y_0^2)}{n_0^2 \mu} \)  

  \[ (2.3.12) \] \( (2.4.14) \]

- Ray path:

  \[ x(t) = x_0 \phi_x(t) + \frac{\xi_0}{n_0 g} \phi_y(t) \]  

  \[ y(t) = y_0 \phi_x(t) + \frac{\eta_0}{n_0 g} \phi_y(t) \]  

  \[ \xi(t) = n_0 g x_0 \phi_x(t) + \frac{\xi_0}{n_0 g} \dot{\phi}_y(t) \]  

  \[ \eta(t) = n_0 g y_0 \phi_x(t) + \eta_0 \dot{\phi}_y(t) \]  

  \[ (2.4.25) \]
Both $\phi_c(t)$ and $\phi_s(t)$ are series expansions of the form

$$
\phi(t) = \phi_0(t) + \mu \phi_1(t) + \mu^2 \phi_2(t) + \mu^3 \phi_3(t) + \ldots
$$

with

$$
\phi_1(t) = N_4 \phi_{11}(t) \\
\phi_2(t) = N_4^2 \phi_{21}(t) + N_6 \phi_{22}(t) \\
\phi_3(t) = N_4^3 \phi_{31}(t) + N_4 N_6 \phi_{32}(t) + N_8 \phi_{33}(t)
$$

- Optical path length:

$$
L = \frac{n_0}{g} \Phi(t)
$$

$$
\Phi(t) = t + \mu \Phi_0(t) + \mu^2 \Phi_1(t) + \mu^3 \Phi_2(t) + \mu^4 \Phi_3(t) + \ldots
$$

$$
\Phi_1(t) = N_4 \Phi_{11}(t) \\
\Phi_2(t) = N_4^2 \Phi_{21}(t) + N_6 \Phi_{22}(t) \\
\Phi_3(t) = N_4^3 \Phi_{31}(t) + N_4 N_6 \Phi_{32}(t) + N_8 \Phi_{33}(t)
$$

### A.1.2 Formulae derived automatically

The coefficients of the power series expansions for $\phi_c$ and $\phi_s$, their derivatives, and the coefficients for $\Phi$ have been obtained by computer algebra as described in the main text (See Sec. 2.4.2) We have performed the computation twice, using both DERIVE and MATHEMATICA, and have obtained identical results. These results have been automatically translated into FORTRAN and only minor changes have been made with the text editor. The resulting code has been tested as shown in Secs. (2.4.4) and (A.3) and has been proven to be free of errors.

Since these formulae have a simple structure but a considerable length, their readability in current mathematical notation would not be much better than in FORTRAN notation. Therefore, in order to avoid transcription errors we prefer to reproduce them here directly from the FORTRAN code. The symbols containing Greek letters have been coded as follows:

$$
\phi_{ij} \leftrightarrow x_{ij}, \phi_{ij} \leftrightarrow x_{ij}, \Phi_{ij} \leftrightarrow dx_{ij}, \Phi_{ij} \leftrightarrow dx_{ij}, \Phi_{ij} \leftrightarrow opl_{ij}
$$

where i and j are integers, e.g. $\phi_{22} \leftrightarrow dx_{22}$ . Powers of the three quantities $a$, $b$, and $t$ and trigonometric functions of multiples of $t$ have been evaluated separately and have been denoted

$$
a^i \leftrightarrow a_i, b^i \leftrightarrow b_i, t^i \leftrightarrow t_i, \sin(i) \leftrightarrow s(i), \cos(i) \leftrightarrow c(i)
$$

e.g. $a^3 \leftrightarrow a3$, $\cos(t) \leftrightarrow c(1), \cos(3t) \leftrightarrow c(3)$. The plus sign appearing at the beginning of lines denotes continuation lines.
xc0 = c(1)
xcl1 = -(a*c(3)+b*s(3)+(4*t*b-a)*c(1)-(4*t*(a+2)+7*b)*s(1)) / 8
xc21 = ((a2-b2)*c(5)+2*a*b*s(5)+2*(4*t*b*(a+4)+a2-11*a+8*b+6)*
+ c(3)-2*(2*t*(a+8*a-b2)+b*(9*a+17))*s(3)+(8*t2*(a2-6*a+b2-4)-8
+ t*b*(a-19)-3*a2+22*a-3*(5*b2+4))*c(1)-2*(24*t2*b-4*t*>(3*a2+7*
+ a+2*(2*b2+5))-b*(22*a+111))*s(1) / 64
xc22 = -((a2-b2)*c(5)+2*a*b*s(5)+3*(a2+4*a-b2))*c(3)+2*(3*b*(a+2)*
+ s(3)+2*)((12*t*b-a-3*a+b2)*c(1)-(6*t*(a2+2*a+b2)+2)+7*b*(a+3))
+ s(1)) / 32
xc31 = -(3*a*(a2-3*b2)*c(7)+3*b*(3*a2-b2))*s(7)+6*(2*t*b*(3*a2+28*
+ a-b2)+3*a2-23*a2+a*(25*b2+12)+35*b2)*c(5)-12*(t*(a3+14*a2-3*a*
+ b2-14*b2)+b*(5*a2+29*a-6*(b2+1)))*s(5)+6*(4*t2*(a3-18*a+2*a*(b2
+ -64)+18*b2)-2*t*b*(41*a2+240*a+29*b2+388)+60*a3-47*a2-2*a*(23*
+ b2-200)-545*b2-276)*c(3)+6*(4*t2*b*(a2-36*a+b2-64)+2*t*(31*a3+
+ 20*a2+a*(19*b2+292)-4*(47*b2+24))+b*(97*a+506*a+5*(3*b2+116))
+ )*s(3)+(32*t3*b*(a2+b2-28)+24*t2*(37*a3-6*a2+a*(29*b2+208)-2*(
+ 9*b2-28)+24*t*b*(89*a2+106*a+85*b2+754)-3*(127*a3-140*a2+2*a*(
+ 824-45*b2)-12*(85*b2+46)))^c(1)-(32*t3*a3+8*a2+a*(b2-28)+8*(
+ b2-1)-24*t2*b*(23*a2+12*a+b2+228)+24*t*(10*a3+87*a2+2*a*(
+ 23*b2+73)+193*b2+188)+3*b*(1063*a2+2456*a+763*b2+8080))*s(1))
+ / 1536
xc32 = (3*a*(a2-3*b2)*c(7)+3*b*(3*a2-b2))*s(7)+(4*t*b*(3*a2+28*a-
+ b2)+11*a3+4*a2+a*(35*b2+24)+20*b2)*c(5)-(4*t*(a3+14*a2-3*a*b2-
+ 14*b2)+b*(5*a2+16*a-19*b2-24))*s(5)+(12*t*b*(5*a2+36*a+5*b2+52
+ -83*a-12*a2-3*a*(b2+168)+4*(141*b2+52))*c(3)-(12*t*(5*a3+18*
+ a2+a*(5*b2+52)-18*b2)+b*(181*a2+688*a+85*b2+744))*s(3)-(96*t2*
+ (a3+a*(b2+10)+4)+8*t*b*(69*a+68*a+67*b2+438)-69*a3-8*a2+a*(23
+ *b2-480)+8*(73*b2+26))*c(1)-(96*t2*b*(a2+b2+10)-8*t*(21*a3+116
+ *a2+3*a*(17*b2+46)+4*(49*b2+38))*b*(985*a2+2144*a+661*b2+4992)
+ )*s(1)) / 256
xc33 = -(a*(a2-3*b2)*c(7)+3*b*(3*a2-b2))*s(7)+2*(a3+6*a2-3*a*b2-6*b2
+ )*c(5)+2*b*(3*a2+12*a-b2)*s(5)+18*(a3+2*a2+a*(a2+4)))*b2-
+ c(3)+18*b*(a2+4*a+b2+4)*s(3)+3*(24*t*b*(a2+b2+4)-7*a3-16*a2-3*
+ a*(b2+8)+16*b2)*c(1)-(24*t*(3*a3+12*a2+3*a*(b2+4)+4*(3*b2+2))+
+ b*(177*a2+336*a+109*b2+504))*s(1)) / 96
dxc0 = -s(1)
dxc11 = -(3*b*c(3)-3*a*s(3)-(4*t*a+2)+3*b)*c(1)-(4*t*b+3*a+8)*
\[ dxc_{21} = \frac{(10ab^5c - 5(a^2 - b^2)s + 4t(6a^2 + b^2 - 2a) + 8b)}{8} \]

\[ dxc_{22} = \frac{-10ab^5c - 5(a^2 - b^2)s + 4t(7a + 9)b}{32} \]

\[ dxc_{31} = \frac{-21b^7c - 21a^2b^2s + 12(5a^3 - 3b^2)a}{1536} \]

\[ dxc_{32} = \frac{21b^7c - 20a^2b^5 + 4t(3a^2 + 28a - b^2) + b(45a^2 + 68a + 43b^2 + 198)}{256} \]

\[ dxc_{33} = \frac{7b^7c - 10a^2b^5 + 54(a^2 + 4a + b^2 + 4)c - 3t(24t + b^2 + 194) + 3b(2^3 + 125b^2 + 1488)}{96} \]

\[ xs_0 = s(1) \]

\[ xs_11 = \frac{b^5c^3 - 4t(a^2 - b) + 8b}{s(1)} \]

\[ xs_{21} = \frac{-2ab^5c + (b^2 - a^2)s + 2(4t^2 + 8a - b^2 - a + 8)t}{} \]
\[
\begin{align*}
&+ c(3) + 2* (4t*b*(a-4) - 3a^2 + 15*a - 2*(2b+23)) * s(3) + 2*(24t^2b+4t b) + \\
&+ (7*a^2 - 11*a + 2*(4b+2b+15)) + b*(2*a+21)) * c(1) - (8t^2*a(2a+6+a+b-4) + \\
&+ 24*t*b*(a+5) + 37*a^2 - 30*a + 41(b+204) * s(1)) / 64
\end{align*}
\]

\[
xs_{22} = (2*a*b*c(5) + (b-a-2) * s(5) - 6*b*(a-2) * c(3) + 3*(a-2-a*b)) + \\
+ s(3) - 4*(6t*(a-2+a+b+2)-b*(a-3)) * c(1) + 4*(12t*b+5*a^2-a+7) + \\
+ b+12) * s(1) / 32
\]

\[
xs_{31} = (3*b*(3*a^2-b)^2 * c(7) - 3*a*(a^2-b)^2 * s(7) + 12*(t*(a^3-a^2*a^2-a) + \\
+ b+14) * b^2 - b*(a+23+a-a^2*(2b+2^3)) * c(5) + 6*(2t*b*(3^2-a^2-28*a-b^2) + \\
- 5*a^3 + 3^2*(9*a^2-a*(5b+2^4)-13*b^2)) * s(5) + 6*(4t^2*b*(a^2+36*a+b-2^2) + \\
+ 64+2*t^2*(47*a^2-92*b^2+a^2+67*b+404)*4*b+96)+b*(97*a+198*(a^2+79) + \\
+ b+844) * c(3)-6*(4t^2*(a+18*a^2+a*(b-2^4)-18*b^2)-2*t*b*(3^2-a^2) + \\
+ 64*a^5+53*b^2+500)+92*a^3-161*a^2+2*a*(29*b+2316)-311*b^2-324)* \\
+ s(3) + (32*t^3*(a-3-a*b+a*(b-2^2)-8*(b-1^2))-24*t^2*b*(31*a+36*a+a) + \\
+ 31*b^2+296)+24*t*(106*a+327*a^2+2*a*(19*b+269)-569*b^2-1012)-3^2*b^2 + \\
+ b^2*(193*a^2+264*a+173*b^2+1792)) * c(1) + (32*t^3*b*(a^2+b^2+2^4)+24*t^2*b + \\
+ 21*a^3-34*a^2+a^2*(29*b^2+298)+24*t*b*(73*a+166*a+ \\
+ 53*b^2+530)-3*(431*a^3-1804*a^2+a^2*(83*b+2008)-4*(751*b^2+1634)) + \\
+ s(1) / 1536
\]

\[
xs_{32} = -(3*b*(3*a^2-b)^2 * c(7) - 3*a*(a^2-b)^2 * s(7) + (4t^2*(a^3-a^2-a^2-a) + \\
+ b^2+14b^2-b^2*(19*a^2+32*a*(7*b^2+8)))*c(5) + (4t^2*b*(a^2^2-a-b^2)-5*a^3 + 4^2-a^2-a^2*(15*b^2+8)-2^8*b^2)*s(5) + 3*(4t^2*(5*a^3-a^2+a^2 + \\
+ 5*b^2+52)+18^2*b^2+b^2*(47*a^2+63*b^2+344)+c(3) - (12*t*b*(5^2+a^2 + \\
+ 36*a^2+b^2+52)-147*a^3+260*a^2-a^2*(211*b^2+792)+52*(b^2+4)) * s(3)+96 + \\
+ t^2*b^2*(a^2+18^2)+10^8*t^4*(77*a^3+368*a^2+3^2+a*(17*b+2118)-76*(5^5*b+6) + \\
+ b^2*(151*a^2-64*a+a^9*(19*b^2+112))*c(1)-96*t^2*(a^3+a^2*(b^2+10)-4)+ \\
+ 8*t^2*b^2*(69*a^2+92*a^2+59*b^2+366)-277*a^3+1544*a^2-a^2*(3^2+2000)+8^2* \\
+ 323*b^2+378)) * s(1) / 256
\]

\[
xs_{33} = (b*(3*a^2-b)^2 * c(7) - a*(a^2-b^2)^2 * s(7) - 2*b*(3*a^2-a^2-b^2)* \\
+ c(5^2+2*a^2-3^2+a^3+a^2+b^2)*s(5)+18*b*(a^2-a+b^2+4)^2*c(3) - 18^2*( \\
+ a^3-a^2+a^2*(b^2+4)+2^2)*s(3)+(24*t^2*(3*a^3-a^2+a^2*(b^2)+4^2)+3^2* \\
+ b^2+2^2)-b^2*(15*a^2-4^2+a^2+19*b)^2)*c(1)+3*(24*t^2*b*(a^2+b^2)+7*a^3 + \\
+ 80*a^2-a^2*b^2+8)+16*(7*b^2+4)) * s(1) / 96
\]

dx_{s0} = c(1)
\]

\[
dx_{s1} = -(3*a*c(3)+3*b^2*s(3)-(4t*b+3*a)*c(1)+(4t^2*(a-2)-5*b) + \\
+ s(1) / 8
\]

\[
dx_{s21} = (5*(a-b^2)^2*c(5)+10*a*b*s(5)-2*(12*t*b*(a-4)-7*a^2+29*a-2^2) + \\
+ 7*b^2+9)) * c(3)+2^2*(6*t*(a^2-b^2)-b*(13*a+47)) * s(3)+(8*t^2*(a^2+6 \\
+ 2^2)}
\]

---
\[ dxs22 = \frac{-(5^2(a^2-b^2)^2)c(5)+10^2(a^2+b^2)s(5)-9^2(a^2-4^2a-b^2)c(3)-18^2b^2(a-2)\}s(3)-4\{12^2t^2b^2a^2+9^2a^2+b^2\}c(1)-4\{6^2t^2(a^2-2^2a+b^2+2^2b)\}b^2(a-15)\}s(1)}{64}\]

\[ dxs31 = \frac{-(21^2a^2-3^2b^2)c(7)+21^2b^2(a^2-2^2b^2)s(7)-6\{10^2t^2b^2(3^2a^2)\}s(1)}{1536}\]

\[ dxs32 = \frac{(21^2a^2-3^2b^2)c(7)+21^2b^2(a^2-2^2b^2)s(7)+3^2a^2+12^2a^2+a^2+12^2b^2)s(1)}{256}\]

\[ dxs33 = \frac{-(7^2a^2-3^2b^2)c(7)+7^2b^2(a^2-2^2b^2)s(7)-10^2(a^3-6^2a^2-3^2a^2+b^2)\{5^2b^2+2\}c(1)+(24^2t^2b^2+12^2b^2+4^2b^2)s(1)}{96}\]

\[ oplo = \frac{(b^2c^2-a^2s^2-2^2t^2-b^2)}{2}\]

\[ opll1 = \frac{-6^2a^2b^2c^4-3^2(a^2-b^2)^2s(4)-8\{3^2b^3a^2b^2+*(a^2+b^2)\}c^2-4\{6^2t^2\}b^2+a^2+2^2b^2+3^2s(2)+2^2\{2^2t^2(a^2+b^2+2^2b)\}b^2(a^2+b^2)\}s^2(4)}{16}\]

\[ opll2 = \frac{(19^2b^2(a^2-2^2b^2)^2s(6)-24\{18^2t^2(a^2-2^2b^2)\}b^2+2^2a^2+2^2b^2+3^2b^2)^2c(1)+(24^2t^2(3^2a^3-12^2a^2+3^2a^2b^2+4^2b^2)^2c^2-17^2a^3-16^2b^2)}{96}\]
\begin{align*}
+ \quad & b^2-12)-b*(45*a^2+256*a-3*(5*b^2-64))*c(2)*3*(288*t^2*a+24*t*b*(5 \\
+ \quad & a^2+8*a+5*b^2+26)-9*a^3+80*a^2+3*a*(25*b^2-24)+16*(22*b^2+15)) \\
+ \quad & s(2)+96*t^4*a^3+3*a-6*(b^2+1))-8*b*(15*a^2+36*a+10*b^2+99)) / \\
+ \quad & 384 \\
\text{opl}_{22} = & -(5*b*(3*a^2-b^2)c(6)-5*a*(a-3*b^2)*s(6)+120*a*b*c(4)-60*( \\
+ \quad & a^2-b^2)+s(4)-3*(24*t*a*(a^2+b^2+4)+b*(21*a^2+64*a+3*(3*b^2+16)))* \\
+ \quad & c(2)-3*(24*t*b*(a^2+b^2+4)-a^3+32*a^2+3*a*b^2+32*(2*b^2+1)))*s(2)+8*( \\
+ \quad & 6*t^4*(3*a^2+3*b^2+2)+b*(6*a^2+9*a+4*b^2+18))) / 96 \\
\text{opl}_{31} = & -(44*a*b^2+a^2+b^2)*c(8)-11*(a^4-6*a^2*b^2+b^4)*s(8)-8*(57*t*a^3( \\
+ \quad & a^2-b^2)+b^2*(117*a^2-57*a^3*(b^2+1)-58*b^2))*c(6)-4*(114*t*b^2*(3*a^2- \\
+ \quad & b^2+19*a^4-59*a^3+19*a^2*(5*b^2+3)+291*a^2-38*b^4-57*b^2)*s(6)-16*( \\
+ \quad & 432*t^2*a^2-3*a^2*(15*a^4+24*a^3+98*a^2-24*a^2+b^2)+5*b^2*(3*b^2+34) \\
+ \quad & )-2*b^2*(25*a^3+38*a^2-2*a^3*(4*b^2-87)-2*(56*b^2+57))))^c(4)+4*(864*t^2 \\
+ \quad & +(a^2-b^2)+24*t^2*b^2*(15*a^3+12*a^2+a^2*(15*b^2+134)-12*(2*b^2+3))-109*a^4 \\
+ \quad & +232*a^3+64*a^2*(30*b^2+97)+8*a^2*(188*b^2+87)+13*b^4+990*b^2+108*) \\
+ \quad & s(4)+8*(288*t^3*a+72*t^2*b^*(5*a^2+4*a+5*b^2+22)+3*t*(60*a^4-169*a^3+ \\
+ \quad & 4*a^2*(152*b^2+83)-a^2*(205*b^2+504)+4*(10*b^4+139*b^2+114)-b^2*(128*a^3 \\
+ \quad & +395*a^2+3*a^2*(19*b^2+403)-70*b^2+768)^c(2)+4*(576*t^3*b^2-144*t^2*(5 \\
+ \quad & +a^2-2*a^2*(5*b^2+16)-2*(2*b^2+3))-30*t*b^*(53*a^2+8*a^2*(b^2+7)+65) \\
+ \quad & +b^2+168)-59*a^4+259*a^3-3*a^2*(157*b^2+275)+a^2(1128-1115*b^2-9*(26^ \\
+ \quad & +b^4+439*b^2+216)^s(2)+4*(6*t^4*(25*a^4-64*a^3+2*a^2*(25*b^2+96)-192*a \\
+ \quad & +25*b^4+384*b^2+264)+b^2*(45*a^3+720*a^2+3*a^2*(25*b^2+304)+16*(40*b^2+ \\
+ \quad & 153)))/1024 \\
\text{opl}_{32} = & -(36*a*b^2*(a^2-b^2)*c(8)-9*(a^4-6*a^2*b^2+b^4)*s(8)-8*(15*t*a^3( \\
+ \quad & a^2-b^2)-b^2*(18*a^2+15*a^2*(b^2+1)-b^2))*c(6)-4*(30*t*b^2*(3*a^2+b^2)-5*a^4 \\
+ \quad & +17*a^3+5*a^2*(5*b^2+3)-21*a^2-5*b^2*(2*b^2+3)/s(6)-16*(3*t^2*(3*a^4 \\
+ \quad & +32*a^2-2*b^2*(3*b^2+23))+2*b^2*(13*a^3+11*a^2+10*a^2*(b^2+7)-22*b^2+21)) \\
+ \quad & c(4)+4*(24*t*a^2*(3*a^2+3*b^2+32)-37*a^4+88*a^3+4*a^2*(9*b^2+46)+8*a \\
+ \quad & +*(44*b^2+21)+b^2*(49*b^2+376))/s(4)-8*(72*t^2*b^*(a^2+b^2+4)+3^t*(12* \\
+ \quad & a^4-89*a^3+12*a^2*(b^2+7)-a^2*(109*b^2+160)+4*(2*b^4+27*b^2+20))-b^2*(64* \\
+ \quad & a^3+206*a^2*a^2*(49*b^2+433)+65*b^2+288))/c(2)+4*(144*t^2*a^2*(a^2+b^2+4) \\
+ \quad & +6^t*b^2*(121*a^2+8*a^2*(b^2+9)+133*b^2+256)+35*a^4-105*a^3+2*a^2*(207*b^2+ \\
+ \quad & 361)+a^2*(269*b^2-320)+122*b^4+1271*b^2+448))/s(2)-8*t^4*(31*a^4-80*a^3+ \\
+ \quad & 6*a^2*(13*b^2+40)-176*a^3+39*b^4+432*b^2+208)-4*b^2*(33*a^3+360*a^2+3*a^2* \\
+ \quad & (13*b^2+112)+8*(38*b^2+93))/256 \\
\text{opl}_{33} = & -(28*a*b^2*(a^2-b^2)*c(8)-7*(a^4-6*a^2*b^2+b^4)*s(8)+84*b^2*(3*a^2- \\
+ \quad & b^2)+c(6)-84*a^2*(a^2-b^2)*s(6)+176*a*b^2*(a^2+b^2+6)*c(4)-88*(a^4+6* \\
+ \quad & a^2-2*b^2*(b^2+6)))/s(4)-4*(120*t*a^3*(3*a^2+3*b^2+4)+b^2*(96*a^3+351*a^2+32
In this section the formulae for determining the z-coordinate of the intersection point of a ray with the second surface of a radial GRIN lens are listed. This quantity is measured from the vertex of the second surface. The surface is assumed to be spherical, its radius being denoted by R.

- z-coordinate:

\[
\bar{z} = \frac{\zeta}{n_0 g} \tag{2.4.41}
\]

\[
\bar{t} = \mu^1 t_1 + \mu^2 t_2 + \mu^3 t_3 + \mu^4 t_4 + \mu^5 t_5 + \mu^6 t_6 + \ldots \tag{2.4.42}
\]

The derivation method is described in Sec. (2.4.3). Instead of the coefficients \( t_i \), we list below the coefficients \( p_i \) related to \( t_i \) by

\[
(2.4.42)
\]

These coefficients have been determined using MATHEMATICA. Their correctness is ensured by the numerical test given in Sec. (A3). As in the previous section, we reproduce the results directly in FORTRAN notation. The formulae derived automatically contain the following abbreviations:

\[
c_4 = N_4 \left( \frac{n_0}{\zeta_0} \right)^2, c_6 = N_6 \left( \frac{n_0}{\zeta_0} \right)^4, c_8 = N_8 \left( \frac{n_0}{\zeta_0} \right)^6
\]

\[
d = \frac{a + 1}{2}, q = \frac{n_0}{\zeta_0} g R
\]

The remaining variables have the same meaning as in the previous section. The powers of the quantities \( d, b \) and \( q \), which have been evaluated separately, and the indexed quantities \( p_i \) and \( c_i \) have been coded in FORTRAN as

\[
p_i \leftrightarrow p(i), c_i \leftrightarrow c(i), d^i \leftrightarrow d(i), b^i \leftrightarrow b(i), q^i \leftrightarrow q(i)
\]

- Formulae derived automatically:

\[
p(1) = d \times q
\]

\[
p(2) = \frac{d2 \times q}{2} + b \times d \times q2
\]
\[ p(3) = d^3 \cdot q / 2 + 3 \cdot b \cdot d^2 \cdot q^2 / 2 + (b^2 \cdot d + d^2 - 2 \cdot d^3) \cdot q^3 \]
\[ p(4) = 5 \cdot d^4 / q / 8 + 5 \cdot b \cdot d^3 / 2 + (3 \cdot b^2 \cdot d + d^2 - 2 \cdot d^3 - 4) / 2 \cdot d^4 \cdot q^3 / 3 + 6 \cdot c^4 \cdot d^4 \]
\[ p(5) = 7 \cdot d^5 / q / 8 + 35 \cdot b \cdot d^4 / 2 + (3 \cdot b^3 \cdot d + 9 \cdot b^2 \cdot d + 2 \cdot d^2 - 2 \cdot b \cdot d^3) / 3 + 6 \cdot c^4 \cdot d^5 \]
\[ p(6) = 21 \cdot d^6 / q / 16 + 63 \cdot b \cdot d^5 / 8 + (3 \cdot b^4 \cdot d + 18 \cdot b^3 \cdot d + 6 \cdot d^3 - 44 \cdot b^2 \cdot d^3 - 25 \cdot d^4 + 26 \cdot d^5 + 36 \cdot c^4 \cdot d^5) / 3 + 22 \cdot b \cdot c^4 \cdot d^4 \cdot q^5 / 16 + 6 \cdot c^6 \cdot d^5 \cdot q^7 
\]
A.3 Comparison of the analytic ray-tracing results with CODE V

In order to test our entire ray-tracing formalism (refraction formulae, analytic formulae for transfer through radial GRIN media, analytic formulae for determining the ray-surface intersection points, paraxial formulae) we have compared for simple test problems the numerical results produced by these formulae with the results produced by CODE V.

Two cases are shown, corresponding to radial GRIN lenses having positive and negative refractive index distributions. The lens data and initial ray data are arbitrary. For the comparison, the refractive index data used by us (see Eq.(2.3.6)) have been transformed to the form used in CODE V.

\[ n(r) = n_0 + N_{10} r^2 + N_{20} r^4 + N_{30} r^6 + N_{40} r^8 + \ldots \]

In both cases, a ray is traced through a single lens, from the object plane to the paraxial image plane. The Cartesian coordinates of the intersection points of the ray with the two lens surfaces and with the object and image planes are compared for the two programs. (The z-coordinate is measured from the vertex of the corresponding surface.) The results are given in Tables (A.3.1-4). The CODE V results are parts of the output listings. The relevant numerical data have been emphasized by underlining. The agreement is very good, as shown below.

i) Single radial GRIN lens with positive refractive index distribution
- lens data
  \[ n_0=1.6, k=0.0225, d=0.897598, R_1=3, R_2=9.06435, N_4=2/3, N_6=-17/45, N_8=62/315 \]
  CODE V: \[ N_{10}=-1.8 \times 10^{-2}, N_{20}=1.6875 \times 10^{-4}, N_{30}=-1.544062 \times 10^{-6}, N_{40}=1.40857 \times 10^{-8} \]
  transverse magnification = -1
  distance from the object to the first surface = 11.7757
  distance from the second surface to the paraxial image = 10.9112
  effective focal length = 5.8066
- skew ray
  \[ y_0=1, \xi_0=0.1031157, x_0=\eta_0=0 \]

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>object</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>first surface</td>
<td>1.26976611</td>
<td>1.0</td>
</tr>
<tr>
<td>second surface</td>
<td>1.20541925</td>
<td>0.92003163</td>
</tr>
<tr>
<td>paraxial image</td>
<td>-0.06849063</td>
<td>-0.99548835</td>
</tr>
</tbody>
</table>

Tab. A.3.1 Results with analytic ray-tracing formulae
ii) Single radial GRIN lens with negative refractive index distribution
- lens data
  \[ n_0 = 1.6, \ k = -0.01, \ d = 1.5, \ R_1 = 3, \ R_2 = -6, \ N_4 = 2/3, \ N_6 = -17/45, \ N_8 = 62/315 \]
  \[
  \text{CODE V : } N_{10} = 8.0 \times 10^{-3}, \ N_{20} = 3.333 \times 10^{-5}, \ N_{30} = 1.355 \times 10^{-7}, \ N_{40} = 5.496 \times 10^{-10}
  \]
  transverse magnification = -1
  distance from the object to the first surface = 7.36039
  distance from the second surface to the paraxial image = 6.99912
  effective focal length = 3.83920
- skew ray
  \[ y_0 = 1, \ \xi_0 = 0.2, \ x_0 = \eta_0 = 0 \]

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>TANX</th>
<th>TANY</th>
<th>LENGTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBJ</td>
<td>0.00000</td>
<td>1.00000</td>
<td>0.00000</td>
<td>0.10367</td>
<td>0.00000</td>
</tr>
<tr>
<td>STO</td>
<td>1.26977</td>
<td>1.00000</td>
<td>0.47261</td>
<td>-0.10854</td>
<td>-0.13854</td>
</tr>
<tr>
<td>2</td>
<td>1.20542</td>
<td>0.92003</td>
<td>0.12774</td>
<td>-0.11814</td>
<td>-0.17763</td>
</tr>
<tr>
<td>IMG</td>
<td>-0.06849</td>
<td>-0.99549</td>
<td>0.00000</td>
<td>-0.11814</td>
<td>-0.17763</td>
</tr>
</tbody>
</table>

Tab. A.3.2 Results with CODE V
Appendix B  Summary of the formulae for Seidel and chromatic aberration coefficients

For the numerical computation of Seidel and chromatic aberration coefficients, it is convenient to have all necessary formulae collected together. For axial and radial GRIN lenses, the necessary formulae are summarized below with reference to the original numbering.

The Seidel coefficients of the optical system are sums over all surfaces and GRIN media of

i) ordinary surface contributions , $S_p , P_S$
ii) inhomogeneous surface contributions , $S^*_p$
iii) inhomogeneous transfer contributions , $T_p , P_T$

where $p=1$ stands for spherical aberration, $p=2$ for coma, $p=3$ for astigmatism and $p=4$ for distortion. $P_S$ and $P_T$ are the surface and transfer contributions to the Petzval curvature.

The chromatic paraxial coefficients are sums of

i) surface contributions , $S_\lambda p$
ii) inhomogeneous transfer contributions , $T_\lambda p$

with $p=1$ - axial color and $p=2$ - lateral color.

All these aberration coefficients are calculated using paraxial marginal and chief ray data at the lens surfaces and several paraxial invariants. The paraxial ray-tracing formulae are given below for the marginal ray. For the chief ray, the inclination $u$ and height $h$ must be replaced by $w$ and $m$. In what follows $\Delta ( )$ denotes the difference between the values after and prior to refraction or transfer of the quantity in the parentheses.
B.1 Ordinary surface contributions

\( n_0 \) : value of the refractive index
- at the surface vertex - for axial gradients
- on the optical axis - for radial gradients (same value at both end surfaces)

Paraxial refraction equation:
\[
n_0'u' = n_0u + h\rho\Delta n_0
\]  
(2.1.28)

Paraxial refraction invariants:
\[
n_{0,j} = n_0 h\rho - n_0 u
\]  
(2.1.30)
\[
n_{0,j} = n_0 m\rho - n_0 w
\]  
(3.3.39)
\[H = mn_0u - hn_0w\]  
(2.2.24)

Contributions to the Seidel coefficients:
\[
S_1 = (n_0)^2 h\Delta(u/n_0)
\]
\[
S_2 = n_0 im n_0 j h\Delta(u/n_0)
\]
\[
S_3 = (n_0 j)^2 h\Delta(u/n_0)
\]  
(3.3.40)
\[P_S = -\rho H^2 \Delta(1/n_0)\]
\[
S_4 = (n_0 j)^2 m\Delta(u/n_0) + n_0 j H\Delta(w/n_0)
\]

Contributions to the chromatic coefficients:
\[
S_{\lambda_1} = hn_{0,j}\Delta \left( \frac{\delta n_0}{n_0} \right)
\]  
(3.6.19)
\[
S_{\lambda_2} = hn_{0,j}\Delta \left( \frac{\delta n_0}{n_0} \right)
\]
\[
\delta n_0 = n_0 (\lambda_0 + \delta \lambda) - n_0 (\lambda_0)
\]  
(3.6.9)

The surface contributions \( S_p \), \( P_S \), \( S_{\lambda p} \) are identical for homogeneous lenses and GRIN lenses with arbitrary rotationally symmetric refractive index distributions.
B.2 Axial gradients

Parameters of refractive index distribution \( n = n(z) \) determining the primary aberrations:
- at the vertices of the two end surfaces: the refractive indices \( n_0 \) and the first derivatives
  \[
  N_z = \frac{dn}{dz} \bigg|_{z=0}
  \]
  (3.3.4)
- the average values over the axial thickness of \( 1/n \) and \( 1/n^3 \) (Eqs. (2.2.21) and (3.4.10))
  \[
  \langle n^{-1} \rangle = \frac{1}{d} \int_{z}^{z+d} \frac{dz}{n(z)}, \quad \langle n^{-3} \rangle = \frac{1}{d} \int_{z}^{z+d} \frac{dz}{n^3(z)}
  \]
- the chromatic changes of \( n_0 \) at the end surfaces and the chromatic change of \( <n^{-1}> \)
  \[
  \delta x \langle n^{-1} \rangle = \delta x \left( \frac{1}{d} \int_{z}^{z+d} \frac{dz}{n(z)} \right) = \frac{1}{d} \int_{z}^{z+d} \delta x \left( \frac{1}{n(z)} \right) dz = - \frac{1}{d} \int_{z}^{z+d} \frac{\delta x n(z)}{n^3(z)} dz
  \]
  (3.6.25)

Paraxial transfer equation:
  \[ h' = h - d \langle n^{-1} \rangle n_0 u \] (2.2.22)

Transfer invariants: \( n_0 u, n_0 w, H \)

Inhomogeneous surface contributions to the Seidel coefficients:

\[
S_4^* = h^4 \rho^2 \Delta N_z
S_2^* = h^3 \rho \Delta N_z
S_1^* = h^2 \rho^2 \Delta N_z
S_4^* = h \rho^3 \Delta N_z
\]
(3.3.36)

Transfer contributions to the Seidel coefficients:

\[
P_T = 0
\]

\[
T_i = n_0 u \left[ n_0 u d \langle n^{-3} \rangle \left( \Delta \left( \frac{h}{n_0^2} \right) \right) \right]
\]
(3.4.14)

\[
\frac{T_4}{T_3} = \frac{T_4}{T_2} = \frac{T_4}{T_1} = \frac{n_0 w}{n_0 u}
\]
(3.4.15)

Transfer contributions to the chromatic coefficients:

\[
T_{x1} = n_0 u \left[ n_0 u d \delta_x \langle n^{-3} \rangle \left( - \Delta \left( \frac{\delta x n_0}{n_0^2} h \right) \right) \right]
\]
(3.6.24)

\[
T_{x2} = n_0 w \left[ n_0 u d \delta_x \langle n^{-3} \rangle \left( - \Delta \left( \frac{\delta x n_0}{n_0^2} h \right) \right) \right]
\]
B.3 Positive and negative radial gradients

Refractive index distribution:

\[ n^2(r^2) = n_0^2 \left( 1 - kr^2 + N_4 k^2 r^4 + \ldots \right) \]  \hspace{1cm} (2.3.6)

Paraxial transfer equations (Eqs. (2.3.24),(2.3.26))

For \( k > 0 \): \( u' = uc \cos gd + hgs \sin gd \), \( h' = \frac{u}{g} \sin gd + h \cos gd \), \( g = \sqrt{k} \)

For \( k < 0 \): \( u' = uc \cosh \hat{g} d - h \hat{g} \sinh \hat{g} d \), \( h' = -\frac{u}{\hat{g}} \sinh \hat{g} d + h \cosh \hat{g} d \), \( \hat{g} = (-k)^{1/2} \)

For \( k = 0 \) see next section.

Transfer invariants: (Eqs. (2.3.31-34))

\[ e_1 = kh^2 + u^2, e_2 = km + uw, \quad e_3 = km^2 + w^2, H = mn_0 u - hn_0 w \]

Inhomogeneous surface contributions to the Seidel coefficients:

\[ S_1' = -2h^4 \rho \Delta (n_0 k) \]
\[ S_2' = -2h^3 m \rho \Delta (n_0 k) \]
\[ S_3' = -2h^2 m^2 \rho \Delta (n_0 k) \]
\[ S_4' = -2h m^3 \rho \Delta (n_0 k) \]  \hspace{1cm} (3.3.35)

Transfer contributions to the Seidel coefficients:

\[ T_1 = n_0 de_1^2 (1 - 3N_4 / 2) + n_0 (1 + N_4) \Delta (hu^3) - 5n_0 N_4 e_1 \Delta (hu) / 2 \]
\[ T_2 = n_0 de_1 e_2 (1 - 3N_4 / 2) + n_0 (1 + N_4) \Delta (hu^2 w) - 5n_0 N_4 e_2 \Delta (hu) / 2 - N_4 H \Delta (u^2) \]
\[ T_3 = n_0 de_2^2 (1 - 3N_4 / 2) + n_0 (1 + N_4) \Delta (huw^2) - 5n_0 N_4 e_2 \Delta (hu) / 2 - 2N_4 H \Delta (uw) - N_4 P_T / 2 \]
\[ T_4 = n_0 de_2 e_3 (1 - 3N_4 / 2) + n_0 (1 + N_4) \Delta (hw^3) - 5n_0 N_4 e_3 \Delta (mu) / 2 - N_4 H \Delta (w^2) / 2 \]  \hspace{1cm} (3.5.21)

\[ P_T = kd H^2 / n_0 \]  \hspace{1cm} (3.5.16)

Transfer contributions to the chromatic coefficients:

\[ T_{\lambda 1} = \frac{1}{2} \frac{\delta \lambda k}{k} [de_1 + \Delta (hu)] \]  \hspace{1cm} (3.6.34)
\[ T_{\lambda 2} = \frac{1}{2} \frac{\delta \lambda k}{k} [de_2 + \Delta (hw)] \]
\[ \delta \lambda k = k(\lambda_0 + \delta \lambda) - k(\lambda_0) \]  \hspace{1cm} (3.6.11)
B.4 Shallow radial gradients

Refractive index distribution:

\[ n^2(r^2) = n_0^2\left(1 + er^4\right) + O(6) \]  \hspace{1cm} (3.5.22)

Paraxial transfer equation: as for homogeneous media

\[ h' = h - ud \]

The inhomogeneous surface contributions to the Seidel coefficients and the transfer contributions to the chromatic coefficients vanish.

Transfer contributions to the Seidel coefficients:

\[ T_1 = -\frac{4}{5} dn_0 \varepsilon \left(h^4 + h^3 h' + h^2 h'^2 + hh'^3 + h'^4\right) \]
\[ T_2 = -\frac{1}{5} dn_0 \varepsilon \left(4h^3 m + 3h^2 h'm + 2hh'^2 m + h'^3 m + h'm'h' + 2h^2 h'm' + 3hh'^2 m' + 4h'^3 m'\right) \]
\[ T_3 = -\frac{2}{15} dn_0 \varepsilon \left(6h^2 m^2 + 3hh'm^2 + h'^2 m^2 + 3h^2 mm' + 4hh'mm' + 3h'^2 mm' + h^2 m'^2 + 3hh'm'^2 + 6h'^2 m'^2\right) \]
\[ T_4 = -\frac{1}{5} dn_0 \varepsilon \left(4hm^3 + h'm^3 + 3hm^2 m' + 2h'm^2 m' + 2hm'm'^2 + 3hm'm'^2 + hm'^3 + 4h'm'^3\right) \]  \hspace{1cm} (3.5.24)

\[ P_T = 0 \]  \hspace{1cm} (3.5.23)
Appendix C  Tables of symbols and conventions of notation

The meaning of symbols which are repeatedly used and the conventions of notation are collected in this appendix with references to the sections where they are defined. Some symbols used only at specific locations and symbols defined by formulae are not listed because their meaning is explained in the text or a reference to the definition formula is made whenever they appear. The conventions of notation are given for an arbitrary quantity denoted by $X$.

### C.1 Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning of symbol</th>
<th>Defined in</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a$</td>
<td>direction vector of the ray</td>
<td>2.1.1</td>
</tr>
<tr>
<td>$B$</td>
<td>Gauss matrix for refraction</td>
<td>2.1.3</td>
</tr>
<tr>
<td>$d$</td>
<td>axial thickness of the lens</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$d_{AP}$</td>
<td>distance between the exit pupil and the paraxial image</td>
<td>3.1.1</td>
</tr>
<tr>
<td>$d_{EP}$</td>
<td>distance between object and entrance pupil</td>
<td>2.5.2</td>
</tr>
<tr>
<td>$d_p$</td>
<td>distance between object and first surface of the optical system</td>
<td>2.5.1</td>
</tr>
<tr>
<td>$d_Q$</td>
<td>distance between last surface of the system and paraxial image</td>
<td>2.5.1</td>
</tr>
<tr>
<td>$f$</td>
<td>focal length</td>
<td>2.5.1</td>
</tr>
<tr>
<td>$G$</td>
<td>Gauss matrix of the optical system</td>
<td>2.5.1</td>
</tr>
<tr>
<td>$G_r$</td>
<td>Gauss matrix for transfer through a radial GRIN medium</td>
<td>2.3.2</td>
</tr>
<tr>
<td>$G_z$</td>
<td>Gauss matrix for transfer through an axial GRIN medium</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$h$</td>
<td>height of the marginal ray (or of a general paraxial ray)</td>
<td>2.1.3,2.5.2</td>
</tr>
<tr>
<td>$H$</td>
<td>paraxial invariant of the system</td>
<td>2.2.3</td>
</tr>
<tr>
<td>$i$</td>
<td>incidence angle of the marginal ray (or of a general paraxial ray)</td>
<td>2.1.3,3.3.3</td>
</tr>
<tr>
<td>$j$</td>
<td>incidence angle of the chief ray</td>
<td>3.3.3</td>
</tr>
<tr>
<td>$k$</td>
<td>coefficient of the 2d order term in the radial index distribution</td>
<td>2.3.1</td>
</tr>
<tr>
<td>$l_p$, $l_Q$</td>
<td>object and image size</td>
<td>2.5.1</td>
</tr>
<tr>
<td>$L$</td>
<td>optical path length</td>
<td>2.3.3</td>
</tr>
<tr>
<td>$m$</td>
<td>height of the chief ray (or of a general paraxial ray)</td>
<td>2.2.3,2.5.2</td>
</tr>
<tr>
<td>$n$</td>
<td>refractive index</td>
<td>2.1.1</td>
</tr>
<tr>
<td>$n_0$</td>
<td>for radial gradients: refractive index on the optical axis</td>
<td>2.3.1</td>
</tr>
<tr>
<td></td>
<td>for axial gradients: refractive index at the surface vertex</td>
<td>2.5.2</td>
</tr>
<tr>
<td>$n_{01}$, $n_{02}$</td>
<td>values of $n_0$ at the two end surfaces for axial gradients</td>
<td>4.2</td>
</tr>
</tbody>
</table>
\( \mathbf{N} \)  unit vector along the normal to the surface 2.1.3
\( N_4 \)  coefficient of the 4th order term in the radial index distribution 2.3.1
\( N_6 \)  coefficient of the 6th order term in the radial index distribution 2.3.1
\( N_8 \)  coefficient of the 8th order term in the radial index distribution 2.3.1
\( P \)  total Petzval curvature of the system 3.1.1
\( P_S , P_T \)  surface and transfer contributions to the Petzval curvature 3.1.1
\( r \)  distance to the optical axis in GRIN media 2.1.2
\( r_{AP}, r_{EP} \)  radius of the exit and entrance pupils 2.5.2
\( r_p, r_Q \)  maximum object height and maximum paraxial image heigth 2.5.2
\( r_s \)  stop radius 3.1.2
\( \mathbf{R} \)  position vector 2.1.1
\( R \)  radius of the surface 2.4.3
\( s \)  arc length along the ray path 2.1.1
\( s, s' \)  position of object and image of a thin lens 4.1.1
\( S_p \)  ordinary surface contributions to the Seidel coefficients for spherical aberration, coma, astigmatism and distortion 3.1.1
\( S_{\lambda p} \)  ordinary surface contributions to the chromatic coefficients for axial color and lateral color 3.1.1
\( S_p^* \)  inhomogeneous surface contributions to the Seidel coefficients 3.1.1
\( T_p \)  inhomogeneous transfer contributions to the Seidel coefficients 3.1.1
\( T_{\lambda p} \)  inhomogeneous transfer contributions to the chromatic coefficients for axial color and lateral color 3.1.1
\( u \)  angle with the optical axis of the marginal ray (or of a general paraxial ray)* 2.1.3,2.5.2
\( U \)  Gauss matrix for transfer through a homogeneous or shallow radial GRIN medium 2.3.2
\( w \)  angle with the optical axis of the chief ray (or of a general paraxial ray)* 2.2.3,2.5.2
\( W \)  wave aberration 3.1.1
\( x \)  Cartesian x-coordinate 2.1.1
\( y \)  Cartesian y-coordinate 2.1.1
\( z \)  Cartesian z-coordinate or propagation distance of the ray (along Oz) inside the medium 2.1.1
\( Z, Z' \)  positions of the entrance and exit pupils for a thin lens with remote stop 4.1.2
\( \alpha \)  optical ray vector 2.1.1
\( \beta \) transverse magnification 2.5.1
\( \Gamma_p \) total Seidel aberration coefficients of the system for spherical aberration, coma, astigmatism and distortion 3.1.1
\( p=1...4 \) \n\( \Gamma_{\lambda_p} \) total chromatic paraxial aberration coefficients of the system for axial color and lateral color 3.1.1
\( p=1,2 \) \n\( \delta z \) \( z \) - coordinate of a point on a surface measured from the vertex 2.4.3
\( \delta z_s \) axial displacement of the stop position ( in stop-shift equations ) 3.1.2
\( \delta n \) total radial or axial change of the refractive index 4.1.3
\( \varphi \) power 2.5.1
\( \varphi_h, \varphi_g \) homogeneous and gradient contributions to the power 4.1.1
\( \kappa \) aspherical constant 4.2
\( \lambda \) wavelength in vacuum 2.1.1
\( \lambda_0 \) reference wavelength for considering dispersion effects 3.1.1
\( \Lambda_x, \Lambda_y \) components of the quasi-invariant 3.2.1
\( \Phi \) phase 2.1.1
\( \nu_h, \nu_g \) homogeneous and gradient Abbe numbers 4.1.2
\( \rho \) curvature of the surface 2.1.3
\( (\sigma_x, \sigma_y) \) normalized Cartesian ray coordinates of the intersection point of the ray with the entrance pupil plane. 2.5.2
\( (\tau_x, \tau_y) \) normalized Cartesian ray coordinates in the object plane 2.5.2
\( \tau \) independent variable for determining the ray path 2.1.1
\( \xi \) x-component of the optical direction cosine 2.1.1
\( \Xi_{x,y} \) components of the monochromatic transverse aberration 3.1.1
\( \Xi_{\lambda x,y} \) components of the chromatic transverse aberration 3.1.1
\( \eta \) y-component of the optical direction cosine 2.1.1
\( \zeta \) z-component of the optical direction cosine 2.1.1

Note*: The sign convention adopted for the angles \( u \) and \( w \) is that they are of opposite sign to the corresponding direction cosines.
### C.2 Conventions of notation

<table>
<thead>
<tr>
<th>Notation</th>
<th>Meaning of notation</th>
<th>Defined in</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_{AP}$</td>
<td>quantity at the exit pupil plane</td>
<td>2.5.2</td>
</tr>
<tr>
<td>$X_{EP}$</td>
<td>quantity at the entrance pupil plane</td>
<td>2.5.2</td>
</tr>
<tr>
<td>$X_p$</td>
<td>quantity at the object plane</td>
<td>2.5.2</td>
</tr>
<tr>
<td>$X_Q$</td>
<td>quantity at the paraxial image plane</td>
<td>2.5.2</td>
</tr>
<tr>
<td>$X_s$</td>
<td>quantity at the stop plane</td>
<td>3.1.2</td>
</tr>
<tr>
<td>$X_0$</td>
<td>initial value after entering the GRIN medium</td>
<td>2.1.2</td>
</tr>
<tr>
<td>$X_i$</td>
<td>sum of the terms of order $i$ in $x, y, \xi, \eta$ of $X$</td>
<td>3.3.2</td>
</tr>
</tbody>
</table>
| $X_{\lambda}$ | quantity at the wavelength $\lambda$  
( the value at $\lambda_0$ is left without subscript) | 3.6.1      |
| $X'$     | value after refraction or transfer                                                  | 2.1.3, 2.2.3 |
| $\bar{X}$ | paraxial approximation                                                               | 2.1.3      |
| $<X>$    | average value over the lens thickness                                               | 2.2.3      |
| $\dot{X}$ | $\frac{dX}{dt}$ (used for radial gradients with $t$ given by Eq. (2.3.8))          | 2.3.1      |
| $\ddot{X}$ | quantity related to the vertex of the second surface of a GRIN lens                | 2.4.3      |
| $L\{X(t)\}$ | Laplace transform                                                                  | 2.4.2      |
| $O(X)$   | quantity containing only terms of total order $\geq X$ in $(\sigma_x, \sigma_y)$ and $(\tau_x, \tau_y)$ | 2.5.2      |
| $\delta_{\lambda}X$ | $X_{\lambda} - X$                                                                | 3.6.1      |
References

/ 1/ Argentieri, D. Ottica Industriale
     Hoepli, Milano, 1954 (in Italian)

/ 2/ Atkinson, L., Houde-Walter, S., Moore, D., Ryan, D.,
     Stagaman, J. Design of a gradient-index photographic objective
     Appl. Opt. 21 (1982), 993-998

/ 3/ Atkinson, L., Downie, J, Moore, D.
     Gradient-index wide-angle photographic objective
     Stagaman, J., Lyon Voci, L.
     design, Appl. Opt. 23 (1984), 1735-1741

/ 4/ Berek, M. Grundlagen der praktischen Optik
     de Gruyter, Berlin, 1970 (in German)

/ 5/ Blankenbecler, R. Gradient-Index Lens Blanks through
     Controlled Diffusion, Spectroscopy, 6 (1991), 40-41

/ 6/ Bociort , F., Kross , J. New ray-tracing method for radial gradient-index

/ 7/ Bociort, F. Seidel aberration coefficients for radial gradient-index
     (A summary with the main formulae has been published

/ 9/ Bronstein, I.N., Semendjajew, K.A. Taschenbuch der Mathematik (German translation from Russian), 22. Aufl., Teubner, Leipzig, 1985, Sec. 4.4.3


/ 11/ Buchdahl, H. A. An Introduction to Hamiltonian Optics Cambridge University Press, 1970


/ 14/ Fletcher, A., Murphy, T., Young, A. Solutions of two optical problems, Proc. Roy. Soc. (London), 223 (1953), 216-225 reprinted in Ref. / 37/


/ 16/ Haun, N., Moore, D. Control of radial gradient-index profile by leaching of a gel, Appl. Opt, 31 (1992), 5225-5229
/ 17/ Hopkins, H. H. Wave Theory of Aberrations
Clarendon Press, Oxford, 1950, Ch. IX

/ 18/ Hopkins, H. H. The nature of the paraxial approximation ,

/ 19/ Houde-Walter, S. Gradient-index optical imaging systems

/ 20/ Hovestadt, H.H Jena Glass and Its Scientific and Industrial
Applications, MacMillan, London, 1902,
reprinted in Ref. / 37/

/ 21/ Kitano, I Current status of aplanatic gradient-index lens systems
Appl. Opt. 29 (1990), 3992-3997

/ 22/ Kross, J. Optische Abbildung ( in German) in:
Bergmann-Schaefer, Lehrbuch der Experimentalphysik,

/ 23/ Kross, J. Technische Optik III, (Lecture notes in German)
TU Berlin, Berlin, 1978

/ 24/ Kross, J. Technische Optik IV, (Lecture notes in German)
TU Berlin, Berlin, 1979
Marchand, E. Gradient Index Optics, Academic Press, New York, 1978, Secs. 4.2.3-4, 5.1, 5.4, 11.1


Moore, D.T. Reality in Gradient Index Optics Optics & Photonics News (Dec. 1992) p.60


Nishi, H., Ichikawa, H., Toyama, M., Kitano, I. Gradient-index objective lens for the compact disk system, Appl. Opt. 25 (1986), 3340-3344


Sakamoto, T. GRIN lens profile measurement by ray trace analysis Appl. Opt. 22 (1983), 3064-3069


/ 45/ Samuels, J, Moore, D. Gradient-index profile control from mixed molten salt baths, Appl. Opt. 29 (1990), 4042-4050


/ 53/ Slevogt, H. Die Seidelschen Fehlerausdrücke für Systeme aus dünnten Linsen und Spiegelflächen, Optik 8 (1951), 369-375 (in German)

/ 54/ Smith, W.J. Modern Optical Engineering McGraw-Hill, New York, 1966, Sec.3.4

/ 55/ Sono, K., Yamasaki, T., Kishimoto, T Graded-index rod lenses, in a commercial presentation of the SELFOC micro lenses by Nippon Sheet Glass Co. Ltd. Japan


/ 59/ Welford, W.T. Aberrations of Optical Systems Adam Hilger, Bristol, 1986 Secs. 2.2, 4.3, 7.2, 8.2, 8.5, 9.2, 12.1
/ 60/ Wolfram, S. *Mathematica*, a system for doing mathematics by computer, Addison-Wesley, New York, 1991


/ 62/ *** DERIVE Handbook (Version 2)
Soft Warehouse, Honolulu, 1992

/ 63/ German Patent Aus einer Einzellinse bestehendes Objektiv
DE 35 37 883 C 2 Olympus Optical Co., Ltd., Tokyo, Japan
Erfinder: Takada, K., (1992)
Zusammenfassung

In dieser Arbeit wurden die rechnerischen Verfahren zur Ermittlung der Abbildungseigenschaften der Gradientenlinsen mit radialer und axialer Brechkraftverteilung weiterentwickelt.


Auf der Basis der analytischen Strahldurchrechnung wurde eine neue Methode zur Ableitung von Aberrationskoeffizienten der GRIN-Linsen entwickelt. Damit wurden für radiale GRIN-Linsen erstmals genaue und relativ einfache Ausdrücke sowohl für die Seidelschen Bildfehler als auch für die primären Farbfehler erhalten. Für die Ableitung der Seidelschen Bildfehler wurden die beiden Komponenten der Queraberration als Summen von Flächen- und inhomogenen Übergangsbeiträgen dargestellt. Die Ausdrücke für die Terme dritter Ordnung, die aus den Teilbeiträgen bestimmt wurden, waren von beträchtlicher Komplexität, besonders im Fall der Übergangsbeiträge. Eine neu entwickelte heuristische Methode ermöglichte dann die Gestaltung aller Ausdrücke der Seidelschen Bildfehler in einer einfachen Form.