Chapter 2

Fundamental Optics for DVD Pickups

2.1 Introduction to basic optics

The theory of the geometrical aberration and diffraction limits are introduced for estimating the focused laser beam spot of a DVD pickup. The concept and formula for evaluating the image quality of an optical system are also reviewed. Then the DVD specifications are presented. Finally, the optimization procedure is described.

2.2 Geometrical optics

To obtain a sufficiently small spot on the disk, the geometrical aberrations need to be well-controlled. The performance of focused spot can be evaluated with transverse ray aberration, or wavefront aberration. The wavefront aberration is a measurement of the differences in optical path length between the exact wavefront and the reference wavefront, as shown in Fig. 2- 1. From the shape of the wavefront, the nature of the aberrations can be visualized more easily.

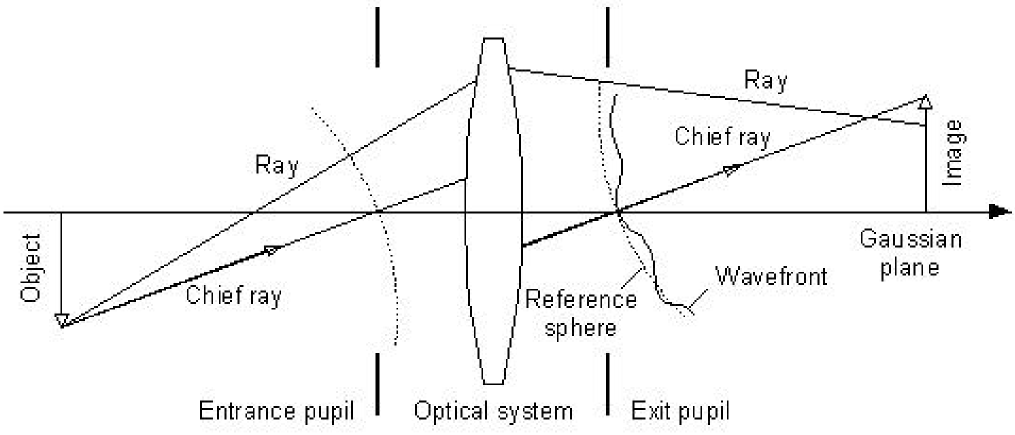

Fig. 2- 1 Ray and wavefront aberration for a general optical system

The wavefront and ray aberrations are related as follows [1]. Referring to Fig. 2-1, the rays emitted from the ideal point source, such as the laser diode, enter the pupil, pass through the optical system, leave the exit pupil, and then converge onto the image plane. If the system is aberration free, the wavefront in image space is spherical in shape and the wave associated with this wavefront converges to paraxial image point. In most system, the point where the ray passes through the pupil plane is represented by the fractional coordinate (x, y). The ray displacements (? ?, ? ?) measured from the chief ray (0, ?) is given by [1]

$$(\Delta x, \Delta h) = -\frac{\overline{R}}{n\, h_E}\left(\frac{\partial W}{\partial x}, \frac{\partial W}{\partial y}\right)$$

Where $\overline{R}$ is the radius of the reference sphere,

n is the refractive index of the image space, and

h_E is the exit pupil height.

Apart from a proportional factor $\overline{R}/n\, h_E$, the transverse aberration is the derivative of the wavefront aberration with respect to the coordinates of the ray at the exit pupil. For most cases, the wavefront shape of a given object point (?, ?) is a function of the pupil position (x, y), and therefore the wavefront aberration can be defined as $W(x, h; x, y)$.

Because the optical system in pickup head can be treated as a cascade of rotationally symmetrical subsystems, we will consider the optical system consisting of a series of refracting surfaces. An object point is displaced along y-axis (See Fig. 2- 2) by a distance ?, and the pupil position is expressed in polar coordinate (x, y) = (? sinf , ? cosf). Then the wavefront aberration can be expressed as a polynomial of parameters ?, ?, and f in terms of wavefront aberration coefficients ${}_kW_{lm}$

$$W(h; r, f) = \sum_{k,l,m} {}_kW_{lm} h^k r^l \cos^m f \qquad \text{(Eq. 2-1)}$$

Where k is the power of ?,

l is the power of ?, and

m is the power of cos f .

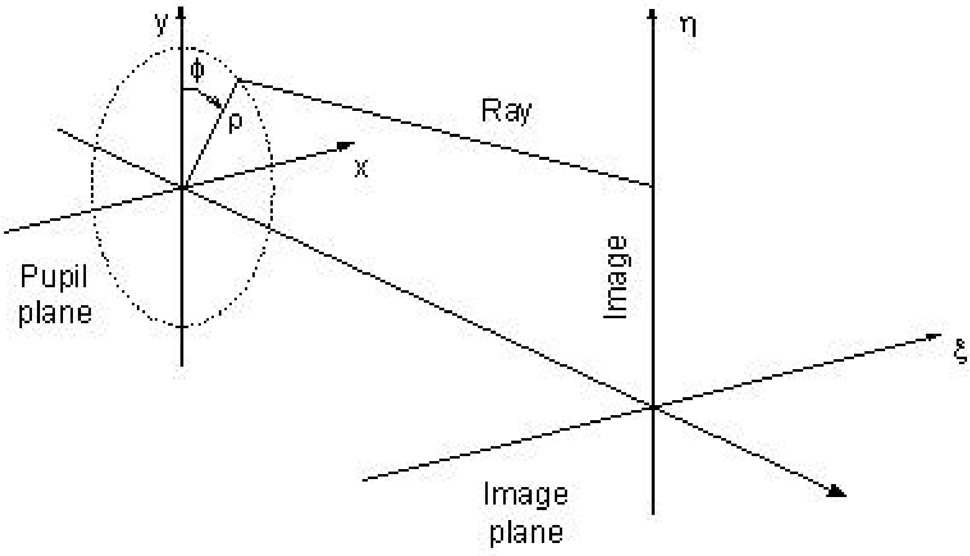

Fig. 2- 2 Coordinate of the pupil and the image plane

The coefficients ${}_kW_{lm}$ have numerical values dependent on the construction parameters, including the curvatures of the surfaces, the refractive indices of the materials, and the thickness between the elements.

As a consequence of the rotational symmetry, the net power ($k+l$) of the length factors (h and r) can only be 0, 2, 4, 6, etc which determines the order of the aberration. These terms are referred to as the second-, fourth-, sixth-order wavefront aberrations and so on. Because the ray aberration is given by the derivative of the wavefront function, these terms correspond to the first-, third-, fifth-order of ray aberrations and so on, and are, traditionally, named as paraxial, primary, secondary, tertiary, aberrations etc. The following paragraph summarizes the wavefront polynomial of the first three orders.

(1) Constant term ($k+l=0$):

$$W(h; r, f) = {}_0W_{00}, \qquad \text{(Eq. 2-2)}$$

(2) Paraxial (first-order) terms ($k + l = 2$):

$$W(h; r, f) = W_{200}h^2 + W_{111}h\, r \cos f + W_{020} r^2 , \qquad \text{(Eq. 2-3)}$$

(3) Primary (third-order) terms ($k + l = 4$):

$$W(h; r, f) = W_{040} r^4 + W_{131}h\, r^3 \cos f + W_{222}h^2 r^2 \cos^2 f + W_{220}h^2 r^2 + W_{311}h^3 r \cos f \qquad \text{(Eq. 2-4)}$$

In practice, the aberration coefficients can be computed through the finite ray tracing [2, 3, 4]. First, we determine the wavefront shape through finite ray tracing at several points on the exit pupil, and then subtract a sphere from that shape, giving the wavefront aberration. Interpolation of a two-dimensional polynomial provides an analytical expression. However, the calculation of the exact aberrations generally is a tedious procedure. In the next paragraph, an elegant method for calculating primary-aberration is presented.

This classic approach was developed by Seidel (1856), Hopkins (1950) and Welford (1986) [1, 5, 6], stating that the primary wavefront aberrations (the first five terms in the aberration polynomial) can be calculated through the Seidel sums S_I , S_{II} , S_{III} , S_{IV} , and S_V

$$W = \frac{1}{8} S_I r^4 + \frac{1}{2} S_{II} h r^3 \cos f + \frac{1}{2} S_{III} h^2 r^2 \cos^2 f + \frac{1}{4}(S_{III} + S_{IV}) h^2 r^2 + \frac{1}{2} S_V h^3 r \cos f \qquad \text{(Eq. 2-5)}$$

where the formulae of the Seidel sums are given by

$$S_I = -\sum_n A_n^{\,2} h_n \Delta\left(\frac{u}{n}\right)_n , \qquad \text{(Eq. 2-6)}$$

$$S_{II} = -\sum_{n} A_n \bar{A}_n h_n \Delta\left(\frac{u}{n}\right)_n, \quad \text{(Eq. 2-7)}$$

$$S_{III} = -\sum_{n} \bar{A}_n^{\,2} h_n \Delta\left(\frac{u}{n}\right)_v, \quad \text{(Eq. 2-8)}$$

$$S_{IV} = -\sum_{n} H^2 c_n \Delta\left(\frac{1}{n}\right)_n, \quad \text{(Eq. 2-9)}$$

$$S_V = -\sum_{n} \frac{\bar{A}_n^{\,3}}{A_n} h_n \Delta\left(\frac{u}{n}\right)_n + \frac{\bar{A}_n}{A_n} H^2 c_n \Delta\left(\frac{1}{n}\right)_n. \quad \text{(Eq. 2-10)}$$

In the formulae, the summation is carried over all the surfaces in the optical system; c_n is the curvature of the n th surface indicated; n_n is the refractive index of the n th medium; $\Delta(u/n)_n = (u/n)_{n+1} - (u/n)_n$, $\Delta(1/n)_n = (1/n)_{n+1} - (1/n)_n$; the paraxial Lagrange invariants are $A_n = n_n(h_n c_n + u_n)$ and $\bar{A}_n = n_n(\bar{h}_n c_n + \bar{u}_n)$ for the marginal rays and chief rays, respectively. All the quantities in the formulas are only related to the marginal and chief rays.

The relation between the Seidel sums and geometrical aberration is tabulated in Tab. 2- 1. The common geometrical aberrations are spherical aberration, coma, astigmatism, field curvature, and distortion.

Tab. 2- 1 Geometrical aberrations and Seidal sums

Name	Wavefront Aberrations
Spheical	$S_I/8$
Coma	$S_{II}/4$
Astigmatism	$3S_{III}/4$
Petzval	$S_{IV}/4$
Distortion	$S_V/2$

In our design, the monochromatic laser diode is used. The spherical aberration is the primary concern among the monochromatic aberrations tabulated in Tab. 2- 1.

2.2.1 Spherical aberration

Spherical aberration is defined as the variation of focus with aperture. A convex lens is taken as an example (Fig. 2- 3). As the ray height at the lens increases, the position of the ray intersection with the optical axis move further away from the paraxial focus. The distance from the paraxial focus to the ray intersection on the optical axis is called longitudinal spherical aberration.

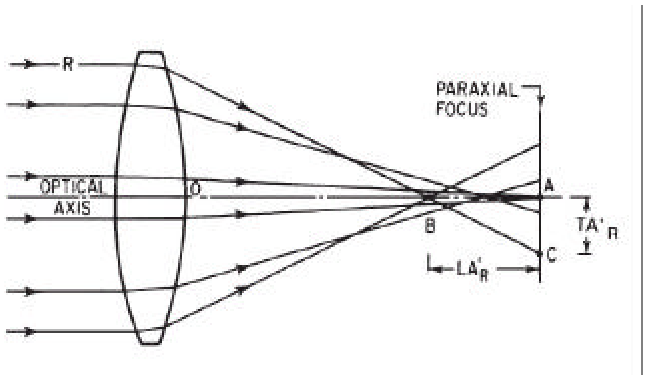

Fig. 2- 3 A simple converging lens with undercorrected spherical aberration

The aspheric surface is a powerful design approach for correcting primary aberrations especially spherical aberration. There are various forms of aspheric surface. The most common form of an aspheric surface is

$$z = \frac{(cv)r^2}{1+\sqrt{1-(cv)(cc+1)r^2}} + (ad)r^4 + (ae)r^6 + (af)r^8 + (ag)r^{10} \qquad \text{(Eq. 2-11)}$$

Where	cv	is the spherical radius of the surface,
	cc	is the conic constant of a surface,
	ad, ae, af, ag	are the coefficient for higher order terms

The common fabrication approaches for aspheric components are injection-molding, diamond-turning, and glasses molding [6].

2.3 Diffraction limit

The laser beam is focused by objective lens onto the optical disk to perform the function of read/write/erase for optical recording. From the viewpoint of geometrical optics, the light can be focused into an infinitesimal point if no aberration presented. However, the fundamental limitation for minimum spot size is the wave behavior of light. When the light passes through a lens with a finite aperture sizes, it will be diffracted and form a diffraction pattern which is called Airy disk with roughly 84% encircled energy in the center ring. Therefore, the minimum spot size (full-width at $1/e^2$) can not be reduced without any limitation.

The diffraction limited spot size is determined by laser wavelength and NA of objective lens depicted in Fig. 2- 4.

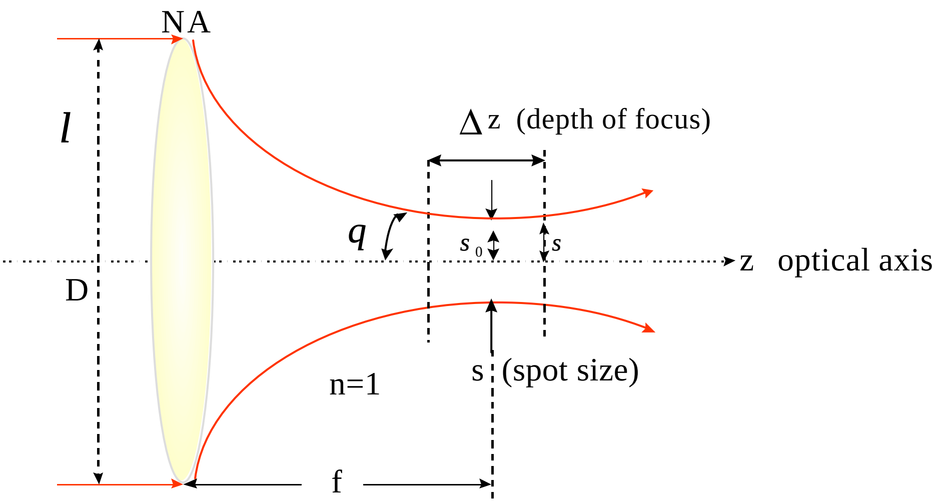

Fig. 2- 4 The diffraction limits of forming a focused Gaussian beam diagram

The total power P and the variance of Gaussian beam s^2 are used for representing the intensity distribution along radial direction. [7]

$$I(r) = \frac{P}{2p s^2} e^{-\frac{r^2}{2s^2}} \qquad \text{(Eq. 2-12)}$$

The variance s^2 is expressed as:

$$s^2 \cong \frac{\left(1+e^{-\frac{1}{2k}}\right)}{8kp^2\left(1-e^{-\frac{1}{2k}}\right)}\left(\frac{l}{NA}\right)^2 \qquad \text{(Eq. 2-13)}$$

Where, ? is the wavelength of laser diode.

The definition of spot size s is full width at half maximum (FWHM) of Gaussian distribution.

$$s = 2\sqrt{2\ln 2}s \cong 2.35s \qquad \text{(Eq. 2-14)}$$

For example, when apodization factor $\frac{1}{k}$ equals 2.5, then

$$s \cong 0.25\frac{l}{NA} \qquad \text{(Eq. 2-15)}$$

Substituting the Eq. 3-16 into Eq. 3-15, the spot size

$$\mathrm{s} \cong 0.59\frac{l}{\mathrm{NA}} = 0.59\frac{l_0}{n_1 \times NA} \qquad \text{(Eq. 2-16)}$$

And, the NA is,

$$NA = n\,sinq \sim \frac{n \cdot D}{2f} \quad ; \quad n = 1\ (for\ air) \qquad \text{(Eq. 2-17)}$$

n is the refractive index of the media in image space and D is the diameter of lens. The f represents the focal length of objective, and $?_0$ is the wavelength in the air. According to Eq. 2-16, the spot size is proportional to the wavelength and inversely proportional to NA. Data density on the disk can be increased by using shorter wavelength laser diodes or increasing the numerical aperture of objective.

After focusing, the Gaussian beam will propagate with Gaussian distribution with the beam size varying along the optical axis, and the total power keeping constant. The depth of focus is given by the distance of change within which the spot size will not alter too much. The depth of focus has significant influence on the mechanical tolerance in the longitudinal direction. The longer the depth of focus, the larger error tolerance of the position of the servo.

From Marechal's criteria [8]:

Strehl ration: $\dfrac{I_0(z)}{I_0(0)} > 0.8$ (Eq. 2-18)

$I_0(z)$ is the beam intensity at a distance z from the beam waist. $I_0(0)$ is the beam intensity at the beam waist.

$$I_0(z) \propto \frac{\sigma_0^2}{\sigma^2} \qquad \text{(Eq. 2-19)}$$

And σ_0 is the root mean square radius at the waist.

$$\frac{\sigma}{\sigma_0} = \left\{1 + \frac{\lambda^2 z^2}{16\pi^2 \sigma_0^4}\right\}^{\frac{1}{2}} \qquad \text{(Eq. 2-20)}$$

Substituting into Eq. 2-18,

$$z \cong \pm 6.3 \frac{\sigma_0^2}{\lambda} \qquad \text{(Eq. 2-21)}$$

Replace the σ_0 with σ in Eq. 2-20

$$\Delta z \cong 0.8 \frac{\lambda}{NA^2} \qquad \text{(Eq. 2-22)}$$

According to Eq. 2-22, the depth of focus is in inversely proportional to the square of numerical aperture. Although increasing numerical aperture can reduce the spot size, the difficulty in focusing servo rises due to the significant reduction of depth of focus.

A briefly summary is as follows.

1. The diffraction limited full width at half maxima spot size gives

   $$s \cong 0.59 \frac{\lambda}{\mathrm{NA}}$$

   For DVD, wavelength = 650 nm and spot size = 1.08 μm, NA is 0.6.

2. The depth of focus, Δ z, gives

   $$\Delta z \cong \pm 0.8 \frac{\lambda}{NA^2}$$

   For DVD, wavelength = 650 nm and NA = 0.6, the depth of focus is 1.44 μm which is within the vibration of disk when rotating.

From the derivation above, NA is 0.6 for DVD application. For aplantic system with well corrected geometrical aberrations like the objective lens in DVD pickup, NA is approximately equal to D/f, where D and f are the diameter and the focal length of the objective lens, respectively. From the approximation, the ratio of lens diameter and focal length can be determined.

2.4 DVD specification

The spot size is the key specification related to NA of objective lens. Fig. 2- 5 can be used to define the required spot size for CD, DVD, and high density digital versatile disk (HDDVD). The optical demands of DVD pickup are tabulated as Tab. 2- 2. Usually the NA of objective lens is measured by using collimated plane wave. However, in our system, the spot is practically a focused Gaussian beam emerging from the edge emitting laser. The NA of objective lens derived from spot size in the previous section is slightly different from the NA of DVD specification listed below.

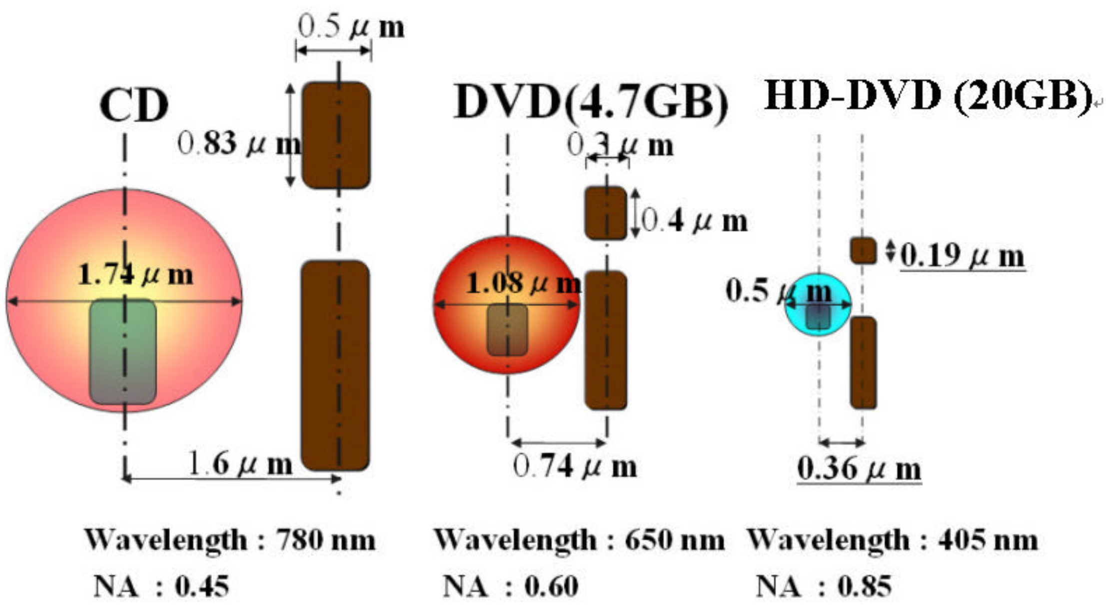

Fig. 2- 5 The optical specification for CD, DVD, and HDDVD

Tab. 2- 2 The data formats and optical demands of DVD pickup [9]

Name	DVD (4.7 GB)	HDDVD (18 GB)
Disk		
Layer in front of recording layer (mm)	0.6 (substrate)	0.1 (cover layer)
Track Pitch (μm)	0.74	0.35 ~ 0.4
Shortest Pit Length (μm)	0.4	~2
Optical Stylus		
Wavelength (nm)	650	405
Planer Wave NA	0.6	0.85
Spot Size (μm)	1.08	~0.5

To design an objective lens yield a nearly diffraction limited performance, the spot size must be smaller than the Airy disk and the optical path difference must meet Rayleigh criteria [6].

2.5 Optimization procedures

With ray-tracing program, the optical system performance can be examined to meet the given design specification. In optimization, the performance can be characterized by a single function, called the merit function F, which can be expressed in the form of a weighted sum of the squares of the defect functions:

$$\Phi = \sum_{i=1}^{m} w_i f_i^2(<x>) \qquad \text{(Eq. 2-23)}$$

Where the vector <x> denotes the optimization variables, and

$$<x> = <x_1, x_2, x_3, ..., x_n>$$

Where fi (i=1, 2..., m) are called operands. The fi(<x>) denotes the deviations from the target values, which in general are nonlinear functions of construction parameters x, and the desired attribute, such as the paraxial constants, aberration coefficients, and exact ray displacements. The surface parameters xj (j=1, 2, 3… n) are the variables to be adjusted, including the surface curvature, arrangement of elements, refractive indices of materials, etc. The target of optimization is to search for the suitable xj to make the merit function as small as possible. Minimization of the merit function is based on a piece-wise linear model of operand dependencies on the variables. If a small change of operands has been made, the new distributions, described by a vector f (x+? x), can be expanded in the vicinity of the initial value x as

$$f_i(x+\Delta x) = f_i(x) + A\Delta x \qquad \text{(Eq. 2-27)}$$

where $<\Delta x> = <\Delta x_1, \Delta x_2, \Delta x_3, ..., \Delta x_n>$ denotes the step vector and A is the derivative of the matrix of each of the operands with respect to each of the variables,

$$\mathbf{A} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \qquad \text{(Eq. 2-24)}$$

The merit function is at its minimum value when all the operands are zero. The merit function is written as

$$\Psi = f^T f \qquad \text{(Eq. 2-25)}$$

Where

$$f = < f_1, f_2, f_3, ..., f_m > \qquad \text{(Eq. 2-26)}$$

According to the optimization theory, if the step vector Δx is the solution of the least-squares equation,

$$(A^T A)\Delta x = -A^T f(x) \qquad \text{(Eq. 2-27)}$$

the merit function is optimal.

In addition, to assure the convergence of the equation above, a dumping factor μ was proposed by Levenberg, Wynne, and Girard [2, 3, 4]. A damping term μ is added to the merit function so that the following function is to be minimized.

$$\Psi = f^T f + m\Delta x^T \Delta x \qquad \text{(Eq. 2-28)}$$

Therefore, we derive the damped least-square equation

$$(A^T A)\Delta x + mI\Delta x = -A^T f(x), \qquad \text{(Eq. 2-29)}$$

where I is an identity matrix. In carrying out the optimization procedure, one takes the x+? x as a new initial point, and aberration of their derivation is re-computed. The program will iterate the process until a minimal merit function is found.

To design the aspheric profile of objective lens, the curvature, conic constant, and higher order coefficients are taken as optimization variables, <x>, shown in Fig. 2- 6. Then, the merit function, f(<x>), is sum of edge thickness and spot size of the objective lens. To optimize the lens profile, the target of merit functions is spot diagram smaller than the Airy disk and positive thickness. Next, the variables are optimized and the aspheric profiles are obtained.

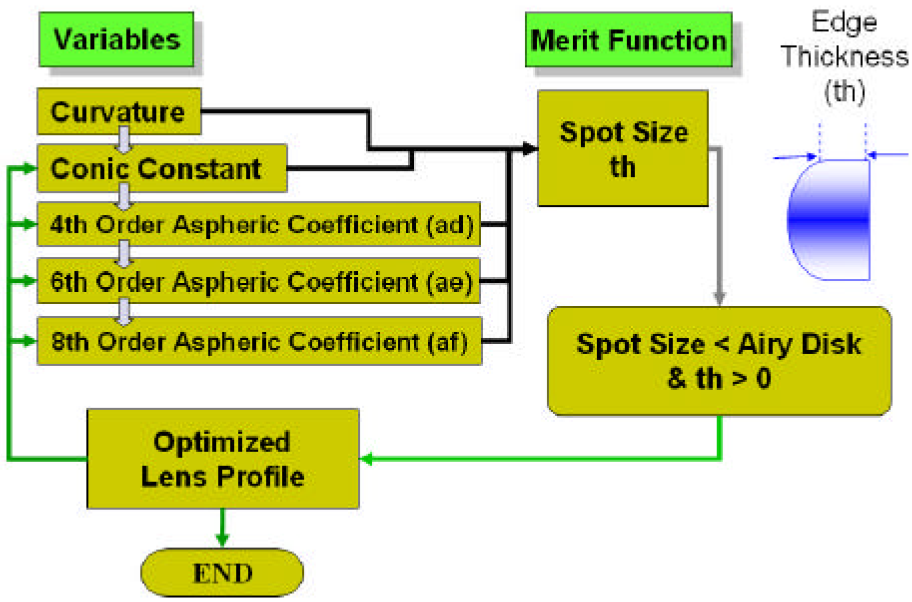

Fig. 2- 6 Optimization flow chart for aspheric profiles

2.6 Summary

To design an objective lens for focusing the spot size smaller then 1.08 µm in the disk is the main objective of this work. First, the theories related to the image evaluation were reviewed, which includes the aberration function. The aberration theory presents the image qualities dependent on constructional parameters. To eliminate the geometrical aberrations, the main aberration should be considered in designing the objective lens is spherical aberration. Next, the minimum spot size restricted by the wave characteristic of light is the diffraction limit. Based on the diffraction theory, the diffraction limits and depth of focus were discussed. Then, the DVD specification is illustrated. Finally, we described the merit function and numerical methods for searching the optimized optical systems.