CHAPTER 1
OPTIMIZATION

For the first 40 years of the twentieth century, optical design was done using a mixture of Seidel theory, a little ray tracing, and a great deal of experimental work. All of the computations were done using log tables. The usual method was to design a prototype lens using Seidel theory, then to manufacture this prototype and measure its aberrations. The designer would then modify the target values for the Seidel aberrations and produce a new design. Ray tracing was relatively little used; in most companies skew rays were almost never traced until about 1940. This seems strange to us nowadays, but it must have been sensible at the time; ray tracing, after all, is only a simulation technique, and if it is cheaper and quicker to do the real experiment, there is little point in carrying out a simulation.

Mechanical calculators were little used until about 1940. Again, this is now difficult to comprehend, but the reason appears to be that techniques for the use of log tables were developed to the point that an experienced "computer," usually a young lady, could actually trace rays faster than with calculators.

Computers were first used in the UK for optical design in 1949, when C.G. Wynne had some ray tracing done at Manchester University. Unfortunately, this pioneering work took so long that the lens was made and delivered to the customer before the ray tracing results were ready! The use of computers for ray tracing did not really become practical until about 1957, when Taylor Hobson acquired an Elliott computer for this purpose, but for several years it was the only British company with a computer dedicated to optical design.

However, rapid ray tracing, in itself, was not in most cases sufficient to enable the design of more advanced lenses, except where lens manufacturers were building very complex systems. For example, the author was fortunate to be employed in the early 1960s by Rank Taylor Hobson, where G.H. Cook and P.A. Merigold were designing complex zoom lenses.[1] For these systems, the use of a computer for ray tracing did enable the design of lenses that would have been impractical otherwise, but this was an exceptional situation.

For most lens designers, the major impact of computers came when optimization techniques began to be used in lens design. While early workers in this field suggested various methods, the dominant method has become that of damped least squares (DLS), proposed by Wynne[2] in the UK and Girard[3] in France, and now known to computer scientists as the Levenberg-Marquardt method.[4,5] Many optical designers have developed variations on this basic technique.[6–12] Some designers have also used other methods with success, such as the adaptive method of Glatzel[13] and Rayces.[14]

1.1 Special characteristics of lens design as an optimization problem

In lens design, we seek to optimize the parameters that determine performance while at the same time producing a design that is manufacturable and, preferably, can be made at a reasonable cost. The optical performance can be most conveniently described, for optimization purposes, in terms of the sum of the squares of some appropriate aberrations, and optimizing these aberrations is indeed the main problem.

1.2 The nature of the merit function

From the beginning, it has generally been implicitly assumed that in lens design a useful merit function can be defined by the sum of squares of the aberrations of a lens. This assumption follows from recognizing that the effect of a positive aberration is the same as that of a negative aberration, at least in the sense that both are equally harmful to image quality. However, there are good theoretical reasons for using the sum of the squares of the aberrations as a merit function, as we will now show.

1.2.1 The Strehl ratio

One criterion of image quality is the Strehl ratio, defined as the ratio of the intensity at the maximum of the actual point spread function (the image of a point object, such as a star or pinhole) to the maximum of the aberration-free point spread function. It is well known that for small aberrations, the Strehl ratio depends on the variance of the wavefront aberration.[15–17] Although this is not the same as the sum of the squares of the wavefront aberrations, as they are normally defined, it is still essentially dependent on the squares of the aberrations and not on the absolute values of the aberrations.

1.2.2 MTF optimization

A more realistic criterion for many lenses that image extended objects with larger residual aberrations is the modulation transfer function (MTF). It has been shown that the MTF depends on the squares of a set of terms that are closely related to the aberrations. To be specific, Gostick[18] and Kidger[19] showed that the following could approximate the geometrical optics approximation to the MTF:

$$L_r(f) \propto \sum [1 - 2\sin^2(\pi \cdot f \cdot \delta\xi')], \tag{1.1}$$

where $L_r(f)$ is the real part of the sagittal geometric MTF at a spatial frequency of f and $\delta\xi'$ is the x-component of transverse aberration. Naturally, a similar expression applies for the tangential component of the MTF.

It follows that an optimization program that minimizes $\sum \sin^2(\pi \cdot f \cdot \delta\xi')$ will maximize the MTF, according to the geometrical optics approximation. Naturally a similar but more complex expression can be derived for the accurate diffraction-based MTF value, but this will not be followed up here, as we are only concerned with showing that the sum of the squares of a set of terms is a logical form—and perhaps the only logical form—for a merit function for lens design optimization problems.

1.2.3 General comments

There are, of course, exceptions to the rule that the sign of an aberration does not have to be considered in a lens design optimization program, but they are few. One class of systems that does introduce some exceptions is almost any lens that is designed for visual use. In most visual lenses, such as eyepieces, it is effectively impossible to correct field curvature. However, in visual lenses the importance of the defocus introduced by field curvature is dependent on the sign of the aberration, because the normal observer is capable of accommodating to an image that is formed in front of, but not behind, the eye. Most lens design programs deal with this by setting a nonzero target for the appropriate aberration; the merit function is still considered to be the sum of squares of these modified “aberrations.” It should be pointed out that, apart from lens design, not all optimization problems have this characteristic that the merit function is naturally formed by the sum of squares of a set of terms.

1.2.4 Comparison with the optical thin-film design problem

To take one example, but still in an optical context, optimization programs are used in the design of thin-film filters (e.g., antireflection coatings) where a given reflectance or transmittance is required. Of course this reflectance or transmittance is dependent on wavelength, angle of incidence, and polarization, but this is not relevant to the present discussion.

In the case of an antireflection coating, reflectances should be zero, so in that sense they can be likened to aberrations in lens design; but it is not logical to define a thin-film merit function as the sum of the squares of the reflectances. It is much more logical to define the merit function as the sum of the reflectances to the first power, because the mean reflectance, integrated over the whole of the relevant waveband, is proportional to the weighted sum of the reflectances and not at all on the sum of the squares.

It can be argued that the intensity is itself the square of the amplitude, and indeed a thin-film optimization program could use a merit function consisting of the sum of the squares of the real and imaginary components of the amplitude; but a general-purpose optical thin-film optimization program is often required to optimize for nonzero reflectance, as in the case of beamsplitters, where a reflectance of 50% may be desired. In this case, it is not appropriate to define a merit function in terms of the squares of its components. However, we will not consider any further the optical thin-film design problem; we merely wish to emphasize that many optimization problems, apart from lens design, do not lend themselves automatically to a merit function that is built up as a sum of squares.

1.2.5 Nonlinearity of the aberrations

If the aberrations were linear functions of the design parameters (curvatures, separations, refractive indices, etc.), lens design would be relatively simple. It would be possible to easily construct a set of linear equations that could be solved to find a set of parameters that minimize the merit function. Apart from relatively minor problems such as the control of boundary conditions, and the detailed definition of the merit function, an optimization program would locate the minimum of the merit function very rapidly.

However, apart from a few special and nontypical cases, in lens design the aberrations are nonlinear functions of the design parameters. This has a great impact on the methods needed in lens design optimization, and on the effectiveness of lens design optimization programs.

Consider one very simple example. The Seidel spherical aberration of a single surface, S_1, is given by

$$S_1 = -A^2 h\left(\frac{u'}{n'} - \frac{u}{n}\right), \tag{1.2}$$

where $A = ni,$

- i is the angle of incidence of the paraxial marginal ray,
- h is the height of the paraxial marginal ray,
- u is the angle of the paraxial marginal ray,
- n is the refractive index,

and primes indicate quantities after refraction or reflection at a surface.

Since i is a linear function of the curvature, c, and $(u'/n' - u/n)$ is also a linear function of c, clearly the spherical aberration is a cubic function of c. The expressions for several other Seidel aberrations also demonstrate that they must be cubic functions of c. As we shall see in Chapter 2, the relationship between fifth-order spherical aberration and curvature is much more complex, but it certainly depends on at least the third power of c. In addition, all high-order aberrations are affected by induced aberrations; in other words, the aberration introduced by a surface is affected by the aberrations of previous surfaces.

It should be clear from the above discussion that in general there are no simple relationships between the aberrations of a lens and its design parameters. Admittedly, a few aberrations are simple. The Petzval sum of a surface is exactly linearly proportional to the curvature, as are first-order axial color and lateral color, but the effectiveness of an optimization procedure is determined by the most difficult cases, not by the most simple ones, and the most difficult aberrations to optimize are the highly nonlinear aberrations.

1.2.6 Changes needed to reduce high-order aberrations

Another factor also determines the strategy of lens design, even before we start to think about using computers and optimization programs. It has been well understood for many decades, as a generalization, that high-order aberrations change relatively slowly, while third-order terms change quickly with changes in the lens parameters. In the precomputer period of lens design, this fact was used quite consciously. As already mentioned, it was a common procedure for designers to set up a prototype lens design with given Seidel aberrations, and then either to measure the aberrations of an actual prototype lens, or to compute the aberrations by ray tracing. In both cases, it was possible to determine a new set of Seidel aberrations, which would improve the lens performance, on the assumption that the high-order aberrations would not change much.

All of these factors, when applied to the lens design optimization problem, mean that in interesting cases we expect that an optimization program should be able to reduce the high-order aberrations. However, to do this, large parameter changes are often required, and the nonlinearity of the Seidel aberrations therefore implies that Seidel aberrations will usually change rapidly, and nonlinearly, while the high-order aberrations are gradually being reduced.

1.2.7 A method of visualizing the problem of optimization in lens design

It is convenient to think of the simplest case, a hypothetical two-dimensional problem. Given that large parameter changes are normally required to reduce high-order aberrations, we can predict that the distance from the starting point to the minimum is, in general, large, and therefore contours of the merit function will be elongated ellipses. This method of visualization is in good agreement with the general experience that in a complex system it is easy to correct the Seidel aberrations, but that small changes from the optimum design parameters will degrade the performance significantly.

If we visualize the contours of the merit function as representing a valley, the bottom of the valley represents the region where the Seidel aberrations are corrected. As we travel downhill along the valley floor, the high-order aberrations are slowly reducing, and of course the valley is not straight, because the Seidel aberrations are nonlinear functions of the design parameters.

The basic problem in quickly finding the minimum is therefore to find the lowest point at the bottom of this curved valley; of course, in important cases the number of variables is quite large, typically between about 20 and 50, and the two-

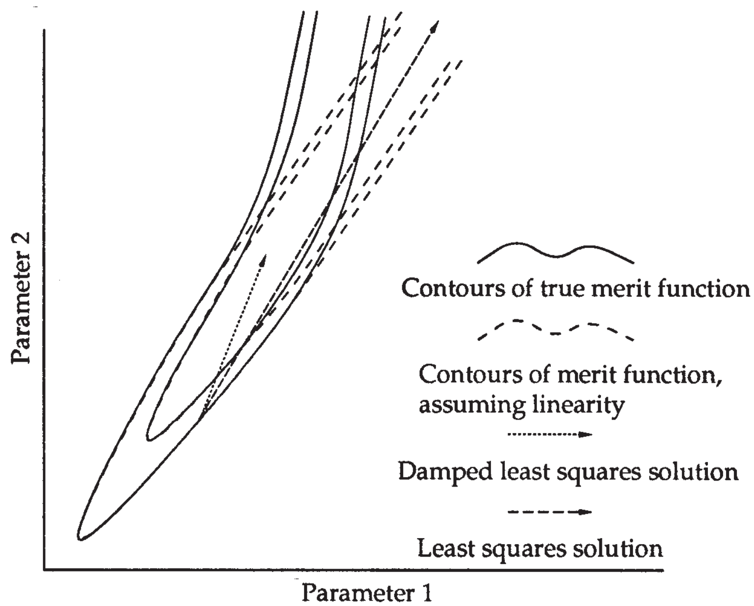

Figure 1.1 An illustration of the merit function.

dimensional picture in Fig. 1.1 is highly simplified, but it nevertheless gives a useful model for visualizing the optimization process.

1.3 Theory of damped least squares (Levenberg-Marquardt)

Suppose that the aberrations of an optical system can be described by a vector **g**. The components of this vector, g_i, are the aberrations of the lens, and typically there might be about 100 of these aberrations, although in extreme cases there can be many hundreds of aberrations. In practice these aberrations are often multiplied by a weighting factor, and they are also sometimes related to a nonzero target value, but this does not affect the following discussion.

The derivatives of the aberrations with respect to the variable design parameters are written as $a_{ij} = \partial g_i / \partial x_j$, and together they are written as a matrix **A**. The individual parameter changes (i.e., changes to curves and separations, typically) are written as x_j, and the set of parameter changes is written as a vector **x**. Typically, the number of variable parameters is a few tens.

We define a merit function as

$$\Psi = \mathbf{g}^2. \tag{1.3}$$

After a set of parameter changes has been made, the aberrations can be described by a new vector, $\mathbf{g}'$, where

$$\mathbf{g}' = \mathbf{g} + \mathbf{A} \cdot \mathbf{x} + \ldots \tag{1.4}$$

assuming that the parameter changes are linear. It may be shown[5] that the new merit function is minimized if the parameter changes **x** are the solution of

$$\mathbf{A}^{\mathbf{t}}\mathbf{A} \cdot \mathbf{x} = -\mathbf{A}^{\mathbf{t}}\mathbf{g}. \tag{1.5}$$

However, as explained above, the parameter changes **x** given by Eq. (1.5) are normally very large, as the high-order aberrations change slowly. In addition, the nonlinearities discussed above limit the range of validity of Eq. (1.4), and therefore we normally find that a lens with parameter changes given by Eq. (1.5) will not have a smaller merit function. For this reason, Levenberg, Wynne, and Girard suggested that the equations should be "damped," as follows:

Instead of minimizing Ψ, we minimize

$$\Psi' = (\mathbf{g} + \mathbf{A} \cdot \mathbf{x})^2 + p^2\mathbf{x}^2, \tag{1.6}$$

where we have added a term p^2 to the merit function. p is known as the *damping factor*. It is squared in Eq. (1.6) merely to illustrate that it must be positive. Again it may be shown[5] that the parameter changes that minimize this new merit function are given by

$$(\mathbf{A}^{\mathbf{t}}\mathbf{A} + p^2\mathbf{I}) \cdot \mathbf{x} = -\mathbf{A}^{\mathbf{t}}\mathbf{g}. \tag{1.7}$$

The effect of this is that the size of the parameter changes is reduced, compared with the simple least squares method, and by a suitable choice of p the parameter changes can be chosen so that they are within the region of validity of Eq. (1.4). In practice, however, there is no requirement to do exactly this; it is more useful simply to choose the value of the damping factor that minimizes the merit function, and this implies that some departure from Eq. (1.4) is accepted. This is the basis of the Levenberg-Marquardt optimization method, known to lens designers as damped least squares (DLS).

It is important to point out that one of the properties of DLS is that the changes computed by this method are the smallest that can give a specific reduction in the merit function. This assumes that we can define the size of the parameter changes, but this point is not important in practice, except that it leads us to one variation of the basic method, the use of multiplicative damping, which is described below.

One variant of the basic DLS method, proposed by Spencer[6, 7], is to add the condition that some equations must be exactly solved, while at the same time the merit function is to be minimized. If we add the condition that the following equation must be exactly satisfied:

$$\mathbf{B} \cdot \mathbf{x} = \mathbf{e}, \tag{1.8}$$

where **B** is a matrix representing a set of derivatives, b_{ij}, and **e** is a vector representing a a set of quantities, e_j, the equations to be solved are the following:

$$\begin{aligned} (\mathbf{A}^{\mathbf{t}}\mathbf{A} + p^2\mathbf{I}) + \mathbf{B}^{\mathbf{t}} \cdot \lambda &= -\mathbf{A}^{\mathbf{t}} \cdot \mathbf{g}, \\ \mathbf{B} \cdot \mathbf{x} &= \mathbf{e}, \end{aligned} \tag{1.9}$$

where λ is a vector representing a scalar set of multipliers. The solution of Eq. (1.9) consists of a vector combining λ and **x**, but only **x** is actually required. Several lens optimization programs have used equations such as this for the "exact" solution of boundary conditions, but this topic is discussed in more detail in section 1.4.9 below.

1.4 Some details of damped least squares as used in lens design

The merit function may comprise many different types of aberration, depending on the designer's preferences, experience, and the application for which the lens is intended.

1.4.1 Paraxial (first-order) properties

The most important paraxial properties are focal length, image position, and magnification. These are controlled with two "aberrations" and their targets and weights. The first aberration is the focal length or the magnification. In the case of lenses with an object at infinity, the first "aberration" is taken to be the focal length. For finite-conjugate lenses, the first aberration is the magnification. In the case of afocal systems, the first aberration is the angular magnification. If we wish to control the back-focus, we must use a second aberration, which is the back focus error.

Other first-order properties that may be controlled include

- Object-to-image distance, in the finite conjugate case. Often this distance can be quite large (several times the focal length), so a very small weight might be appropriate.
- Total thickness of "glass" elements in the system, that is, for spaces in which the refractive index is not 1.0. The effect of controlling this quantity is that the weight of the lens and the cost of the glass are more or less controlled.
- Distance from the first surface of the lens to the image surface.
- Distance from the first surface of the lens to the last surface of the lens.

- Defocus. Normally we assume that the image surface is at the paraxial image position, but we can specify a defocus from the paraxial position. If the variable label for the image surface is set, the program will use the defocus as a variable parameter in optimization. However, we do not want an excessive amount of defocusing, so a weight on defocus may be used to prevent this.

1.4.2 Seidel and Buchdahl coefficients

As we shall see in Chapter 2, the aberration function can be expanded as a power series in aperture and field, and analytic expressions exist for the calculation of the coefficients. The third-order (Seidel) aberrations can be simply and quickly computed. Computation of the fifth-order aberration coefficients is also possible. Expressions for seventh-order aberrations also exist, but are rarely used.

Even the seventh-order expressions (which are taken to include the third- and fifth-order terms) are not sufficient to infallibly describe the aberrations of a typical photographic lens, for example, because their calculation is based on paraxial rays. It is therefore almost universal for optimization programs to use real ray-tracing results.

1.4.3 Transverse ray or wavefront aberrations

Exact ray tracing is, in principle, a simple and infallible process for calculating the aberrations of particular rays as they pass through a lens. In practice, great care is needed to ensure that the results are correct and accurate in all cases. A general-purpose lens design program must be capable of ray tracing through many different types of surface, but these details are of no concern to the present discussion.

Since no approximations are involved, ray tracing is almost universally used in optimization programs. The results of ray tracing are most often expressed as either transverse or wavefront aberrations. Transverse ray aberrations are the first derivatives of the wavefront aberrations, so that if the transverse aberrations are fully known, the wavefront aberrations can be computed, and vice versa.[20]

1.4.4 Aberration balancing and choice of weighting factors

During optimization, the program does not differentiate between primary and higher-order aberrations, but they are all included in the ray-trace aberrations. To a certain extent, the fact that a least-squares process is used will automatically produce some sort of aberration balance, but a sensible choice of weighting factors is essential if we are to achieve anything like the best possible performance. The program therefore needs to select default weighting factors, according to the following criteria:

1. In some cases it is better to have a sharp center to the spread function, with a large flare, than a smaller flare with a less sharp center. To attempt to achieve this, we normally give low-aperture rays a larger weighting factor than marginal rays.

2. In some cases, it is permissible for the performance at the edge of the field to be worse than at smaller field angles. Very often, of course, it is inevitable, but sometimes (as in photocopiers and microlithography) it is not at all desirable. It is therefore sensible, often, but not always, for weighting factors to be smaller for larger field positions.

3. When using the Conrady chromatic aberration formula (i.e., nearly always), one must remember that this aberration is a wavefront aberration, and is normally smaller than the transverse ray aberration. To see why, remember that

$$\delta\eta' = \frac{-1}{n'\sin\alpha'} \cdot \frac{\partial W}{\partial y} \tag{1.10}$$

and

$$\delta\xi' = \frac{-1}{n'\sin\alpha'} \cdot \frac{\partial W}{\partial x}. \tag{1.11}$$

Assuming that $W = W_{40}y^4$, $\partial W / \partial y = 4W_{40}y^3$, so at the edge of the aperture, when $y = 1$,

$$\partial W/\partial y = 4W_{40} = 4W. \tag{1.12}$$

Taking a numerical aperture of 0.1 as typical, we have $dh' = -10\partial W/\partial y = -10 \cdot 4W = -40W$. What this means is that for a given ray, if the system has third-order spherical aberration, the transverse ray aberration for the marginal ray will be 40 times the wavefront aberration. If, therefore, we have a merit function that includes transverse aberrations as well as wavefront aberrations, the wavefront aberrations should have a much larger weight.

4. If we are using the Coddington astigmatism formula, these are longitudinal aberrations and will normally be larger than transverse aberrations. They should therefore have a smaller weight.

5. It is often important that we should not attempt to control aberrations that are uncorrectable. Examples of this are distortion in eyepieces, chromatic aberration in Ramsden eyepieces, and astigmatism in doublets.