

# Using Geometric Algebra to Study Optical Aberrations

Jack Hanlon and Hans Ziock  
Los Alamos National Laboratory

RECEIVED

MAY 05 1997

OSTI

This paper uses Geometric Algebra (GA) to study vector aberrations in optical systems with square and round pupils. GA is a new way to produce the classical optical aberration spot diagrams on the Gaussian image plane and surfaces near the Gaussian image plane. Spot diagrams of the third, fifth and seventh order aberrations for square and round pupils are developed to illustrate the theory.

## 1.0 Introduction

A new vector based mathematical structure called geometric algebra and calculus (GAC) is being developed by David Hestenes and others to clarify and extend many topics in classical mechanics, special relativity, quantum mechanics and mathematics [1-7]. Manthey [8] suggests that GAC provides a framework for understanding and developing new nanocomputer architectures, and there have even been recent applications of GAC to optics [9-10]. Many have noticed that GAC offers a near ideal mixture of algebra and geometric interpretation [7], and some even suggest that GAC should replace vector and tensor analysis as the basic tool for physics and engineering because GAC uses vector concepts without notation changes; it is more efficient than tensor analysis; it completes the connection between vector analysis and complex variable theory; it encourages geometric as well as algebraic thinking, and it often leads from analysis to diagrams that aid understanding. We use geometric algebra (GA) to clarify vector aberration theory and to provide another example of the application of the new algebra to a traditional optical engineering topic.

Optical aberration theory is an old but still useful and interesting optical engineering topic because it provides insight into optical system performance and methodology for optical design. In the early 1980's, Roland Shack and others developed a unique vector aberration theory [11-14] that addressed aberrations of all orders in centered optical systems and optical systems with tilted components and displaced pupils. Their theory is especially notable for its geometric insight and compact formulation, but although the vector algebra they used is straightforward, it does not use standard notation for some vector operations, and we thought we could improve their theory by rewriting it in the language of geometric algebra.

This paper has several sections. Section two is a brief review of the necessary GA basics, section three uses GA to calculate spot diagrams on the Gaussian image plane for 3<sup>rd</sup>, 5<sup>th</sup>, and 7<sup>th</sup> order aberrations for both round and square pupils. We show the aberration types for square pupils because square beams are used in a major current application [15]. Spot diagram shapes of aberrations on surfaces other than the Gaussian image plane are analyzed in section four. All the results are summarized in the appendices, which include a table of spherical aberration like and coma like spot diagram shapes that are cataloged according to a terminology outlined long ago by Cruickshank and Hills [16]. Appendix 6.2 outlines the connection between the approach used in this paper and traditional wave aberration theory.

## 2.0 Geometric Algebra Introduction

Many of the basis ideas of GA are contained in a new type of vector product called the geometric product. The geometric product of two vectors contains all the geometric and algebraic information about the two vectors. It is written as the sum of two different types of objects as:

$$ab = a \cdot b + a \wedge b$$

MASTER<sup>(21)</sup>

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

*ln*

## DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

**DISCLAIMER**

**Portions of this document may be illegible  
in electronic image products. Images are  
produced from the best available original  
document.**

The left hand side of (2.1) is called the geometric product of vectors  $\mathbf{a}$  and  $\mathbf{b}$ . The first term on the right side of (2.1), the scalar dot product of vector algebra, is well known as the basis of an expression for the projection of one vector onto another. The second term on the right side of (2.1), is called the outer product of vectors  $\mathbf{a}$  and  $\mathbf{b}$ , and as we will see, it is the basis of an expression for finding the rejection, *i.e.*, the component of one vector perpendicular to the other.

The outer product of two vectors is also called a bivector, which is related to, but not the same as, the vector cross product. In fact, the outer and cross products of two vectors are dual concepts; the vector cross product produces a vector perpendicular to the area defined by the outer product. Bivectors are algebraic representations of planes, and like vectors, bivectors have direction, magnitude and sign. The outer product of vectors  $\mathbf{a}$  and  $\mathbf{b}$  represents the direction, magnitude and sign of a segment of a plane defined by  $\mathbf{a}$  and  $\mathbf{b}$ . All bivectors in a plane are scalar multiples of a unit bivector that defines the direction of that plane, just as all vectors on a line are scalar multiples of a unit vector that defines the direction of that line. The bivector boldface symbol  $\mathbf{i}$  is the direction of a unit plane, and the bivector magnitude is the size of the area in the  $\mathbf{i}$  plane. Although the area can take any shape, it is often convenient to think of it as a parallelogram formed by two vectors in the  $\mathbf{i}$  plane; then, if the two vectors are  $\mathbf{a}$  and  $\mathbf{b}$  with an angle  $\theta$  between them, the bivector  $\mathbf{a} \wedge \mathbf{b} = \mathbf{i}|\mathbf{a}||\mathbf{b}|\sin\theta = \mathbf{i}\text{Area}$ . The bivector sign associates a circulation direction (orientation) for areas in the  $\mathbf{i}$  plane. The circulation direction is defined by attaching the tail of the second vector in a bivector product to the head of the first vector and following the direction of circular flow. According to convention, counterclockwise circulation defines a positive area, clockwise circulation a negative area.

Changing the order of vectors in the bivector product changes the circulation direction; consequently, it follows that

$$\mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a}. \quad (2.2)$$

Because of the sign change in (2.2), vectors in a geometric product like (2.1) do not normally commute. Vectors in geometric products commute only if they are collinear, because then  $\mathbf{a} \wedge \mathbf{b} = 0$ , and  $\mathbf{a}\mathbf{b} = \mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} = \mathbf{b}\mathbf{a}$ . Similarly, vectors in geometric products only anticommute if they are orthogonal, because then,  $\mathbf{a} \cdot \mathbf{b} = 0$  and  $\mathbf{a}\mathbf{b} = \mathbf{a} \wedge \mathbf{b} = -\mathbf{b} \wedge \mathbf{a} = -\mathbf{b}\mathbf{a}$ .

With (2.1) and (2.2) we can express the dot and outer products of two vectors in terms of either the geometric product or the angle  $\theta$  between the vectors.

$$\mathbf{a} \cdot \mathbf{b} = 1/2 (\mathbf{a}\mathbf{b} + \mathbf{b}\mathbf{a}) = |\mathbf{a}||\mathbf{b}|\cos\theta \quad (2.3)$$

$$\mathbf{a} \wedge \mathbf{b} = 1/2 (\mathbf{a}\mathbf{b} - \mathbf{b}\mathbf{a}) = \mathbf{i}|\mathbf{a}||\mathbf{b}|\sin\theta \quad (2.4)$$

The geometric and outer products broaden the geometric scope of vector analysis in a fundamental way. For example, bivectors represent area by an area concept and an area direction that is represented by the algebraic symbol  $\mathbf{i}$ . Furthermore, GA provides tools to manipulate areas as well as lines. Bivectors also provide a unification of vector and complex number concepts, and bivectors add significantly more geometric content to traditional vector analysis that is much more than mere cosmetics. In addition, bivectors are the simplest example of higher dimensional vector-like objects called multivectors and a higher dimensional noncommuting algebra.

Geometric products with more than two vectors are defined to follow the distributive and associative rules of arithmetic and algebra, but since vectors in geometric products do not normally commute, the order of vectors in the products must be preserved, *e.g.*,

$$(\mathbf{a}\mathbf{b})\mathbf{c} = \mathbf{a}(\mathbf{b}\mathbf{c}) \quad (2.5)$$

$$\mathbf{a}(\mathbf{b} + \mathbf{c}) = \mathbf{a}\mathbf{b} + \mathbf{a}\mathbf{c} \quad (2.6)$$

$$(\mathbf{b} + \mathbf{c})\mathbf{a} = \mathbf{b}\mathbf{a} + \mathbf{c}\mathbf{a} \quad (2.7)$$

Other geometric algebra rules evolve from (2.5) to (2.7) and definitions (2.1) to (2.4), and as in conventional vector analysis, the derived rules become identities, which are convenient forms used for analysis.

Some useful GA operations for two dimensional (2D) operations are listed below:

1.) An operation not defined in vector analysis is the reverse operation  $(\mathbf{M})^\dagger$ , defined by Hestenes<sup>1</sup>. In the reverse operation the elements of a product are arranged in reverse order, and then the elements themselves are reversed; *e.g.*,

$$(\mathbf{B}\mathbf{a})^\dagger = \mathbf{a}^\dagger \mathbf{B}^\dagger, \quad (2.8)$$

where  $\mathbf{B}$  is a general bivector. The reverse operation applies to vectors and bivectors as follows:

a.) By definition, the reverse of a vector reproduces the same vector;

$$\mathbf{v}^\dagger = \mathbf{v} \quad (2.9)$$

b.) Reversing a bivector changes the bivector sign and consequently, the circulation direction.

$$\mathbf{B}^\dagger = (\mathbf{a} \wedge \mathbf{b})^\dagger = \mathbf{b}^\dagger \wedge \mathbf{a}^\dagger = -\mathbf{a} \wedge \mathbf{b} = -\mathbf{B} \quad (2.10)$$

As a corollary, since  $\mathbf{i}$  is a bivector,  $\mathbf{i}^\dagger = -\mathbf{i}$ .

2.) The geometric product and the dot and outer products of a vector  $\mathbf{a}$  and a bivector  $\mathbf{B}$  are defined by relationships (2.11) and (2.12), which are similar in form to (2.2) and (2.3).

$$\mathbf{a}\mathbf{B} = \mathbf{a} \cdot \mathbf{B} + \mathbf{a} \wedge \mathbf{B} \quad \text{and} \quad \mathbf{B}\mathbf{a} = \mathbf{B} \cdot \mathbf{a} + \mathbf{B} \wedge \mathbf{a} \quad (2.11)$$

$$\mathbf{a} \cdot \mathbf{B} = 1/2 (\mathbf{a}\mathbf{B} - \mathbf{B}\mathbf{a}) = -\mathbf{B} \cdot \mathbf{a} \quad \text{and} \quad \mathbf{a} \wedge \mathbf{B} = 1/2 (\mathbf{a}\mathbf{B} + \mathbf{B}\mathbf{a}) = \mathbf{B} \wedge \mathbf{a} \quad (2.12)$$

In Appendix 1, we show that  $\mathbf{a} \cdot \mathbf{B}$  is a vector, which is an object one dimension lower than  $\mathbf{B}$ . The operation  $\mathbf{a} \wedge \mathbf{B}$  creates a trivector, which is an object one dimension higher than  $\mathbf{B}$ . For our analysis in this paper we set  $\mathbf{a} \wedge \mathbf{B} = 0$  because trivectors are not needed to solve two dimensional problems. Consequently, for 2D problems the product  $\mathbf{a}\mathbf{i} = \mathbf{a} \cdot \mathbf{i}$  is a vector, and using (2.12) or (2.8), (2.9) and (2.10), we get

$$\mathbf{a}\mathbf{i} = -\mathbf{i}\mathbf{a}. \quad (2.13)$$

The relationship  $\mathbf{a} \wedge \mathbf{B} = 0$  is also the condition that vector  $\mathbf{a}$  is in plane  $\mathbf{B}$ .

3.) The following general vector identities (see Appendix 6.1 for proofs), are often useful:

$$(\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c} = \mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) \quad (2.14)$$

$$\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b})\mathbf{c} - (\mathbf{a} \cdot \mathbf{c})\mathbf{b} \quad (2.15)$$

4.) Another useful identity is the expansion of vector  $\mathbf{a}$  relative to vector  $\mathbf{b}$  and  $\mathbf{b}^{-1} = \hat{\mathbf{b}} / \mathbf{b}$ :

$$\mathbf{a}\mathbf{b}\mathbf{b}^{-1} = (\mathbf{a} \cdot \mathbf{b})\mathbf{b}^{-1} + (\mathbf{a} \wedge \mathbf{b})\mathbf{b}^{-1} = \mathbf{a}. \quad (2.16)$$

Equation (2.16) expresses vector  $\mathbf{a}$  in components along  $\mathbf{b}$  and perpendicular to  $\mathbf{b}$ . The first term is the vector component of  $\mathbf{a}$  in the direction of  $\mathbf{b}$  and is called the projection of  $\mathbf{a}$  onto  $\mathbf{b}$ , while the second term,  $(\mathbf{a} \wedge \mathbf{b})\mathbf{b}^{-1}$ , is the vector component of  $\mathbf{a}$  perpendicular to  $\mathbf{b}$ , which is called the rejection of vector  $\mathbf{a}$  from vector  $\mathbf{b}$ .

5.) Equations (2.3) and (2.4) provide another way to write the geometric product of unit vectors  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$ . If the unit bivector for the plane which contains  $\hat{\mathbf{a}}$  and  $\hat{\mathbf{b}}$  is  $\mathbf{i}$ , the geometric product can be written in several ways using well known identities.

$$\hat{\mathbf{a}}\hat{\mathbf{b}} = \hat{\mathbf{a}} \cdot \hat{\mathbf{b}} + \hat{\mathbf{a}} \wedge \hat{\mathbf{b}} = \cos\theta + \mathbf{i}\sin\theta = e^{i\theta} \quad (2.17)$$

Equation (2.17) defines the exponential  $e^{i\theta}$ , which is also called a two dimensional rotor. The rotor concept gives geometric significance to the exponential, and geometric content, as well as consistent vector

<sup>1</sup> Vectors are denoted by lower case boldface letters, unit vectors have a caret (^), and bivectors and trivectors are denoted by capital boldface letters. The exceptions are the unit bivector, which is denoted by a boldface  $\mathbf{i}$ , and the unit trivector which is denoted by an italic  $i$ .

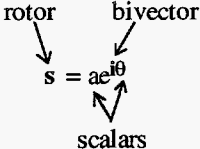
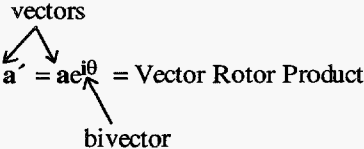
algebra content, to vectors in two dimensional complex variable theory. When a rotor multiplies a vector  $\mathbf{a}$ , the result is a new vector  $\mathbf{a}'$ , which is the old vector  $\mathbf{a}$  rotated by the angle  $\theta$  in the  $\mathbf{i}$  plane;

$$\mathbf{a}' = \mathbf{a} e^{i\theta} = \mathbf{a} \cos\theta + \mathbf{a} \mathbf{i} \sin\theta \quad (2.18)$$

According to (2.18), if  $\theta = 90^\circ$ , the vector bivector product  $\mathbf{a} \mathbf{i}$  is the vector  $\mathbf{a}$  rotated by  $90^\circ$  in the  $\mathbf{i}$  plane:

$$\mathbf{a}' = \mathbf{a} e^{i\pi/2} = \mathbf{a} \mathbf{i}. \quad (2.19)$$

The important result is that it takes the product of a rotor,  $e^{i\theta}$ , and a vector in the  $\mathbf{i}$  plane, to create a consistent two dimensional vector algebra expression for rotating a vector<sup>2</sup>. These ideas, which consistently link GA, vector analysis and complex numbers, are summarized in Table 2.1.

<u>2-D Rotor</u>	<u>Rotated Vector</u>
 <p style="text-align: center;"><math>s = \mathbf{a} e^{i\theta}</math></p>	 <p style="text-align: center;"><math>\mathbf{a}' = \mathbf{a} e^{i\theta} = \text{Vector Rotor Product}</math></p>
$s = \mathbf{a} \cos\theta + \mathbf{a} \mathbf{i} \sin\theta$	$\mathbf{a}' = \mathbf{a} \cos\theta + \mathbf{a}_0 \mathbf{i} \sin\theta$
rotor = scalar + bivector	vector = vector + vector

**Table 2.1.** Properties of a two dimensional rotor and a rotated vector.

A summary of the GA properties of the common elements of 2D geometric algebra are given in (2.20).

$$\mathbf{a} \mathbf{i} = -\mathbf{i} \mathbf{a} \quad \mathbf{i} \mathbf{i} = -1 \quad \mathbf{a} e^{i\theta} = (\mathbf{a} e^{i\theta})^\dagger = e^{-i\theta} \mathbf{a} \quad (2.20)$$

Using rules (2.20), geometric products of vectors in a plane reduce to standard rotated vector forms. In two dimensions, with vectors  $\mathbf{a}$  and  $\mathbf{b}$  written in rotor form as  $\mathbf{a} = a \hat{\sigma} e^{i\alpha}$  and  $\mathbf{b} = b \hat{\sigma} e^{i\beta}$ , with  $\hat{\sigma}$  being a unit vector in a reference direction in the  $\mathbf{i}$  plane, and with the angles  $\alpha$  and  $\beta$  being measured relative to  $\hat{\sigma}$ , we can use rules (2.20) to reduce a compound vector geometric product to a simple polar form, e.g.,

$$\mathbf{b} \mathbf{a} \mathbf{b} \mathbf{a} = a^2 b^3 \hat{\sigma} e^{i(3\beta - 2\alpha)} \quad (2.21)$$

## 3.0 Aberrations

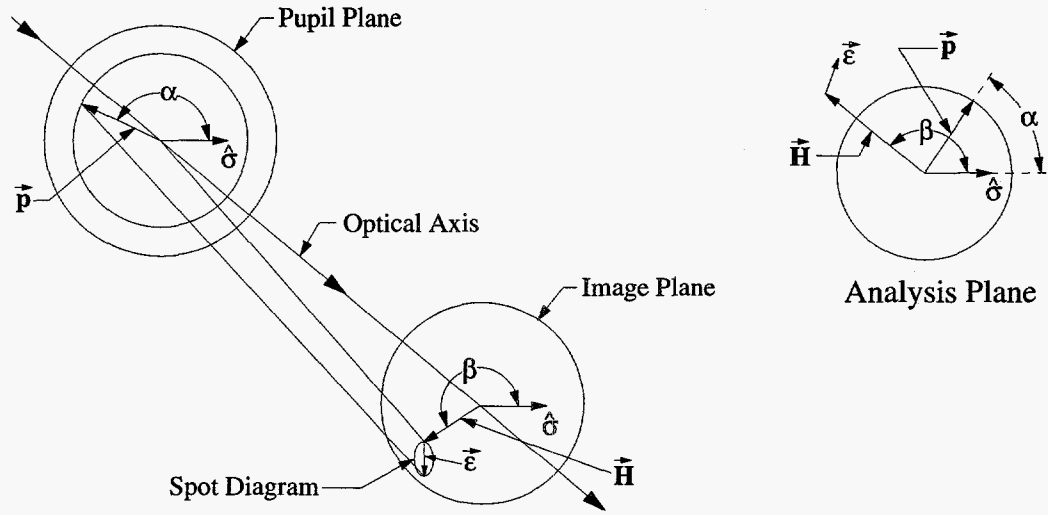
### 3.1 Introduction

Light and particle beam optical theories are usually studied in several parts: 1.) creating system layouts; 2.) understanding aberration types; 3.) learning what aberrations are created by different optical components and systems, and 4.) creating designs with prescribed first order properties that have satisfactory aberration characteristics. This paper is only concerned with understanding aberration types.

Ray aberration theory is based on a simple concept: when the wavefront leaving the pupil is an ideal spherical surface, all wavefront normals (rays) pierce the image plane at the ideal image point, but when the wavefront leaving the pupil is aberrated, the wavefront normals point in a direction slightly different from the ideal direction, and the normals pierce the image plane a vector distance  $\vec{\epsilon}$  from the ideal image point. The vector  $\vec{\epsilon}$  is called the ray aberration. The pattern that results from plotting all the  $\vec{\epsilon}$ 's at an image point for each point on the aberrated wavefront that is centered in the exit pupil is called a spot diagram. The geometry for the analysis is shown in Figure 3.1. The vector  $\mathbf{H} = h \hat{\sigma} e^{i\beta}$  is a normalized image plane

<sup>2</sup>Rotors and phasors, which are often used in engineering analysis, are easily confused. In phasor notation a rotated vector is written as  $\mathbf{a} = \mathbf{a} e^{i\theta}$ . This is not a legitimate equality that can be manipulated by the rules of vector algebra, because the left hand side is a vector, but the right hand side is not.

vector, and the vector  $\mathbf{p} = p\hat{\sigma}e^{i\alpha}$  is a normalized pupil plane vector. For round pupils the scalars  $p$  and  $h$  take any value between 0 and 1, and the angles  $\alpha$  and  $\beta$  are measured from the reference direction  $\hat{\sigma}$ .



**Figure 3.1.** a.) The pattern that results from plotting all the  $\vec{\epsilon}$ 's at an image point, for each point on the aberrated wavefront in the exit pupil, is a spot diagram. b.) The analysis plane.

A spot diagram is located in the image plane by  $\mathbf{H}$ , or equivalently  $h$  and  $\beta$ , and the ray aberrations are generated at  $\mathbf{H}$  by letting  $\mathbf{p}$  scan a set of points in the pupil, usually circles or squares.

The most general expansion for the ray aberration  $\vec{\epsilon}$  is given by (3.1).

$$\vec{\epsilon} = \sum_{c=1}^{\infty} \sum_{m=0}^c W_{m(c-m)} (\mathbf{H}^m \mathbf{p}^{c-m}) \quad (3.1)$$

Since the ray aberration  $\vec{\epsilon}$  must be a vector, and  $W_{m(c-m)}$  is a scalar, the sum of the powers of  $\mathbf{H}$  and  $\mathbf{p}$ , which is equal to  $c$ , must be an odd integer. If instead  $c$  is an even integer, (3.1) reduces not to a vector, but rather to a sum of scalars and bivectors. The even order expansions are identical to the traditional wave aberration polynomials as shown in Appendix 6.2. The first two odd power terms from (3.1) are (3.2) and (3.3).

$$\vec{\epsilon} = W_{01}\mathbf{p} + W_{10}\mathbf{H} \quad (3.2)$$

$$\vec{\epsilon} = W_{03}\mathbf{p}^3 + W_{12}(\mathbf{H}\mathbf{p}^2) + W_{21}(\mathbf{H}^2\mathbf{p}) + W_{30}\mathbf{H}^3. \quad (3.3)$$

Since  $\mathbf{H}$  and  $\mathbf{p}$  do not necessarily commute, we define the symbol  $(\mathbf{H}^m \mathbf{p}^{c-m})$  in (3.1) and (3.3) to represent the sum of different permutations of the vectors  $\mathbf{H}$  and  $\mathbf{p}$  in the product. This sum of permutations is necessary because  $\mathbf{H}\mathbf{p}^2$  can also be written as  $\mathbf{p}\mathbf{H}\mathbf{p}$  or as  $\mathbf{p}^2\mathbf{H}$ , but since  $\mathbf{H}$  and  $\mathbf{p}$  do not commute the different forms are not equivalent. We create spot diagrams by writing the vectors in polar form and reducing the results using the 2D geometric algebra results summarized by (2.20) and (2.21). With this interpretation (3.1) produces all the classical aberration types, which for third order (3.3) are spherical aberration, coma, astigmatism and distortion. The goal of the next section is to develop the geometric algebra of ray aberrations and produce spot diagrams.

## 3.2 Spot Diagrams

### 3.2.1 First Order Aberrations

The simplest ray aberration is the second vector in (3.2),  $W_{10}\mathbf{H} = W_{10}h\hat{\sigma}e^{i\beta}$ . It points along the direction of the field vector  $\mathbf{H}$  and changes the position of the image point. The first vector in (3.2),  $W_{01}\mathbf{p} =$

$W_{01} p \hat{\sigma} e^{i\alpha}$ , creates a scaled circle or square as the vector  $\mathbf{p}$  scans either a circle or square in the exit pupil. The spot diagrams on the ideal image plane correspond to out of focus image points. The amount of focus shift is determined by either  $W_{10}$  or  $W_{01}$  and  $L$ , the distance from the pupil to the ideal image plane. As shown in Figure 3.2, simple geometry leads to an equation for the defocus distance  $\Delta z$ , which is the axial distance between a flat image plane and an observation point. The focus shift is calculated by comparing similar triangles, which leads to two equations for the defocus distance  $\Delta z$ .

$$\Delta z = W_{01} L = W_{10} L. \quad (3.4)$$

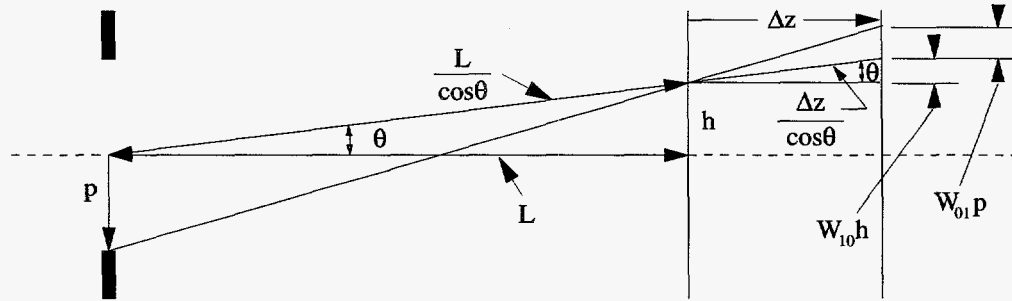


Figure 3.2. The relationship between defocus and 1<sup>st</sup> order aberrations.

### 3.2.2 Third Order Aberrations

The third order aberrations, which are defined by  $c = 3$  in (3.1) and separated out as (3.3), are:  $W_{03} \mathbf{p}^3$ ,  $W_{12} (\mathbf{H} \mathbf{p}^2)$ ,  $W_{21} (\mathbf{H}^2 \mathbf{p})$ , and  $W_{30} \mathbf{H}^3$ .

The third order spherical aberration term,  $\vec{\epsilon} = W_{03} \mathbf{p}^3 = W_{03} \mathbf{p}^2 \mathbf{p} = W_{03} p^3 \hat{\sigma} e^{i\alpha}$ , is independent of the field vector  $\mathbf{H}$ ; consequently, for round pupils, spherical aberration produces the same size blur spot at every point in the image field. At each point the maximum blur circle radius occurs when  $p = 1$ , and the maximum radius is equal to  $W_{03}$ . For square pupils the  $p^3$  dependence, perhaps surprisingly, changes the focal spot shape dramatically, as shown in Figure 3.3. For all the figures we use square and round pupils with the same area<sup>3</sup>, and we set all the  $W_{ij} = 1$  so the size of the spot diagram shapes can be directly compared. The real aberration sizes are of course much smaller because the  $W_{ij}$  will be much smaller than one, in fact, hopefully close to zero for a good design.

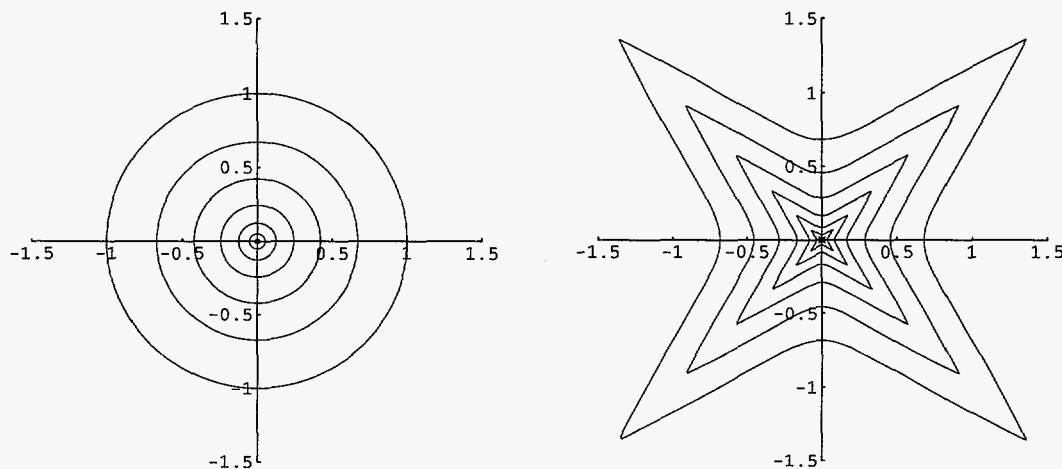


Figure 3.3 Third order spherical aberration for round and square pupils.

<sup>3</sup> Since we consider equal areas for round and square pupils, for a given radius  $p$  for a round pupil, the corresponding square pupil has a half-side length value of  $p_{sq} = p\sqrt{\pi/2}$ . Plotted are spot diagrams for round pupil radii of  $0/8, 1/8, 2/8, \dots, 8/8$ , and the corresponding half-sidelengths for square pupils of the same area.



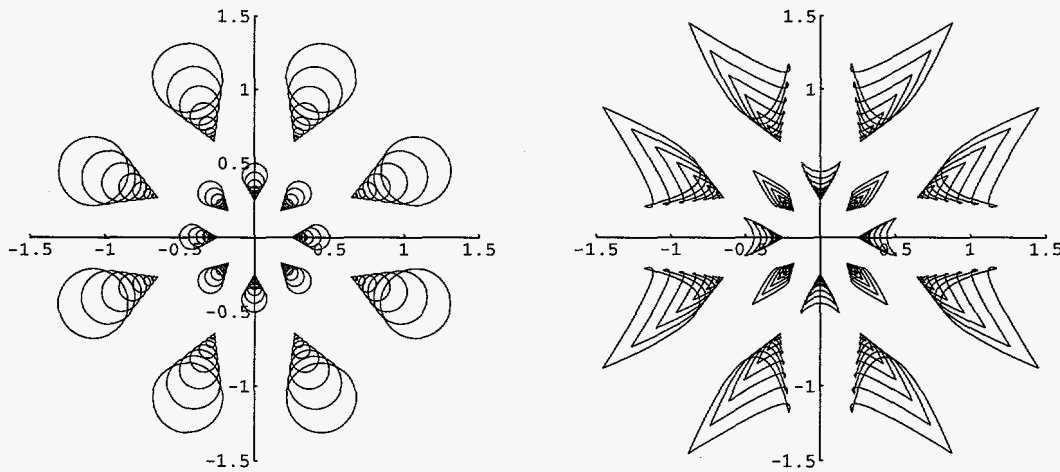
Third order coma,  $\vec{\epsilon} = W_{12}(\mathbf{H}\mathbf{p}^2)$ , is linear in  $\mathbf{H}$  and quadratic in  $\mathbf{p}$  and is sometimes called linear coma. This is the first term where the commutation properties of vectors  $\mathbf{H}$  and  $\mathbf{p}$  are needed. The  $3!/2!$  different permutations of  $(\mathbf{H}\mathbf{p}^2)$  are  $\mathbf{p}^2\mathbf{H}$ ,  $\mathbf{H}\mathbf{p}^2$  and  $\mathbf{p}\mathbf{H}\mathbf{p}$ . Since  $\mathbf{p}^2 = \mathbf{p} \cdot \mathbf{p} = p^2$  is a scalar, the first two forms are the same. We collect like forms and normalize the result by dividing by the total number of permutations to produce the ray aberration vector for coma given by

$$\vec{\epsilon} = 1/3 W_{12}(2p^2\mathbf{H} + \mathbf{p}\mathbf{H}\mathbf{p}), \quad (3.5)$$

which simplifies when we write it in polar form using results like (2.21) to

$$\vec{\epsilon} = 2/3 W_{12}h^2p (\hat{\sigma} + 1/2 \hat{\sigma} e^{2i(\alpha-\beta)}) e^{i\beta}. \quad (3.6)$$

For round pupils (3.6) produces the familiar coma spot diagrams. Probably the easiest way to sketch the geometric figure corresponding to (3.6) for round pupils is to draw the figure for  $\beta = 0$ , since the only effect of  $\beta$  is to rotate the figure to a new field position. With  $\beta = 0$ , the first vector in (3.6),  $\hat{\sigma}$ , is a displacement along the field vector direction. The second vector,  $(1/2 \hat{\sigma} e^{2i\alpha})$ , traces a circle twice in the image plane as  $\alpha$  scans the pupil once between 0 and  $2\pi$ . Both the displacement and circle radius depend on the coefficient  $2/3 W_{12}h^2p$ . The coma patterns point toward the origin if  $W_{12}$  is positive and away from the origin if  $W_{12}$  is negative. These results perfectly match the classical descriptions of third order coma for round pupils. For square pupils the quadratic dependence of coma on  $p$  effects the shape and field dependence of the coma spot diagrams in an interesting way. For square pupils the shape and the size vary as a function of field position and are difficult to sketch; computer generated results are shown in Figure 3.4 and Appendix 6.5.



**Figure 3.4** Third order coma for round and square pupils.

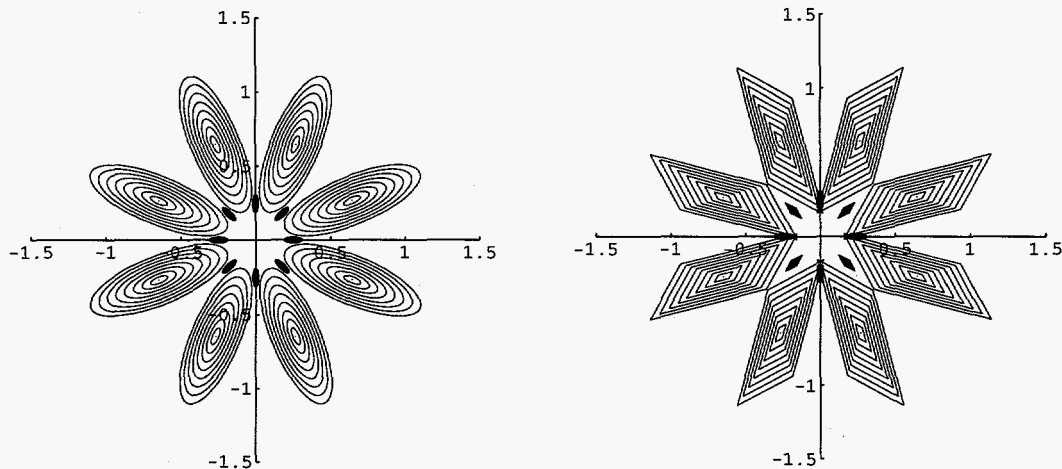
Third order astigmatism,  $\vec{\epsilon} = W_{21}(\mathbf{H}^2\mathbf{p})$ , is second order in  $\mathbf{H}$  and linear in  $\mathbf{p}$ . In a useful classification scheme developed by Cruickshank and Hills[16], this is called linear astigmatism. This term has the same form as (3.6) but with the roles of  $\mathbf{H}$  and  $\mathbf{p}$  interchanged. For third order astigmatism,

$$\vec{\epsilon} = W_{21}(\mathbf{H}^2\mathbf{p}) = 1/3 W_{21}(2h^2\mathbf{p} + \mathbf{H}\mathbf{p}\mathbf{H}), \quad (3.7)$$

which in polar form is

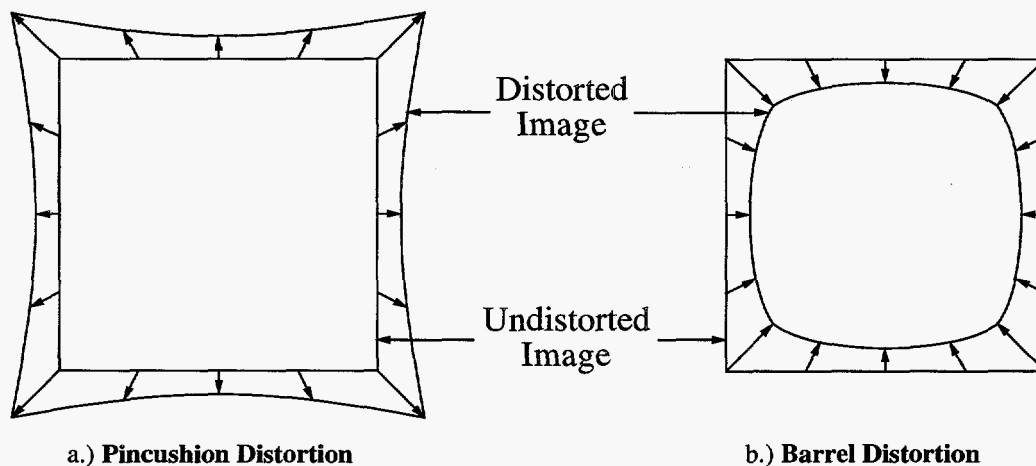
$$\vec{\epsilon} = 2/3 W_{21}h^2p (\hat{\sigma} e^{i(\alpha-\beta)} + 1/2 \hat{\sigma} e^{-i(\alpha-\beta)}) e^{i\beta}. \quad (3.8)$$

The interpretation of (3.8) is that of unequal length vectors spinning in opposite directions. The vector sum creates an ellipse centered on the field point in the image plane as  $\mathbf{p}$  traces a circle in the pupil plane. These elliptical patterns, with the major axis oriented along the field vector direction, are the characteristic patterns for astigmatism on the Gaussian image plane. The characteristic patterns for square pupils are rectangles, but the square pupil skews the pattern in the Gaussian image plane in the directions of the corners of the pupil. The spot diagrams for third order astigmatism for round and square pupils are shown in Figure 3.5.



**Figure 3.5** Third order astigmatism on the Gaussian image plane for round and square pupils.

The last third order term,  $W_{30} \mathbf{H}^3$ , is third order distortion. Distortion terms do not depend on  $\mathbf{p}$ , and they do not blur the image spot; they just displace the image point along the field vector direction by an amount that nonlinearly depends on the field position. If the object is square the image is either pincushion or barrel shaped depending on the sign of  $W_{30}$ ; this is summarized in Figure 3.6; a result that is independent of the shape of the pupil.



**Figure 3.6** Third order distortion of a square object for a.) positive and b.) negative distortion coefficients.

### 3.2.3 Fifth Order Aberrations

There are six, fifth order aberration terms produced when  $c = 5$  in (3.1). They are  $W_{05} \mathbf{p}^5$ ,  $W_{14}(\mathbf{H} \mathbf{p}^4)$ ,  $W_{23}(\mathbf{H}^2 \mathbf{p}^3)$ ,  $W_{32}(\mathbf{H}^3 \mathbf{p}^2)$ ,  $W_{41}(\mathbf{H}^4 \mathbf{p})$ , and  $W_{50} \mathbf{H}^5$ . The first and last terms are fifth order spherical aberration and fifth order distortion respectively. The geometric interpretation for these terms is the same as for the corresponding third order theory except of course for the higher order  $\mathbf{p}$  and  $\mathbf{h}$  dependence. For square pupils, the fifth order dependence of spherical aberration on  $\mathbf{p}$  produces dramatically spiked spot diagram shapes as shown in Appendix 6.5.

The fifth order coma term linear in  $\mathbf{H}$  is  $W_{14}(\mathbf{H} \mathbf{p}^4)$ . There are  $5!/4!$  different permutations of  $\mathbf{H} \mathbf{p}^4$  to reduce using geometric algebra, but the 5 permutations reduce to just two simple forms. It is time consuming to

count and reduce the number of different permutations by hand, so we developed a short Mathematica<sup>4</sup> program, which we called *reduce* (see appendix), to carry out the process. As an example, for  $\mathbf{Hp}^4$ , *reduce*[[H,p,p,p,p]] produces the following results:

1.) A listing of the permutations of the list {H,p,p,p,p}, which are:

$$\{\{\mathbf{H},\mathbf{p},\mathbf{p},\mathbf{p},\mathbf{p}\},\{\mathbf{p},\mathbf{H},\mathbf{p},\mathbf{p},\mathbf{p}\},\{\mathbf{p},\mathbf{p},\mathbf{H},\mathbf{p},\mathbf{p}\},\{\mathbf{p},\mathbf{p},\mathbf{p},\mathbf{H},\mathbf{p}\},\{\mathbf{p},\mathbf{p},\mathbf{p},\mathbf{p},\mathbf{H}\}\};$$

2.) A listing of the reduced forms of each sublist;

$$\{\{\mathbf{H}\},\{\mathbf{p},\mathbf{H},\mathbf{p}\},\{\mathbf{H}\},\{\mathbf{p},\mathbf{H},\mathbf{p}\},\{\mathbf{H}\}\};$$

3.) A listing of the separate forms;

$$\{\{\mathbf{H}\},\{\mathbf{p},\mathbf{H},\mathbf{p}\}\}, \text{ and}$$

4.) A listing of the coefficients for each form. These coefficients are simply the number of times an element of list 3 appears in list 2.

For this example there are 3 H's and 2 pHp's in list 2, so the coefficient list is, {3,2}

With this information the expression for the 5<sup>th</sup> order ray aberration is easily developed.

$$\vec{\epsilon} = W_{14}(\mathbf{Hp}^4) = 1/5 W_{14}(3 \mathbf{p}^4 \mathbf{H} + 2 \mathbf{p}^2 \mathbf{pHp}); \quad (3.9)$$

$$\vec{\epsilon} = 3/5 W_{14} h p^4 (\hat{\delta} + 2/3 \hat{\delta} e^{2i(\alpha-\beta)}) e^{i\beta}. \quad (3.10)$$

Fifth and third order coma spot diagrams have similar shapes. The only difference is the coefficient in front of the second term, which gives the fifth order coma flare angle a value of  $2 \sin^{-1}(2/3) = 83.6^\circ$ . For third order coma, the flare angle is only  $2 \sin^{-1}(1/2) = 60^\circ$ . The 5<sup>th</sup> order coma patterns for square and round pupils are shown in Appendix 6.5.

The ray aberration  $W_{23}(\mathbf{H}^2 \mathbf{p}^3)$  is fifth order astigmatism (2<sup>nd</sup> order in  $\mathbf{H}$ ). The  $5!/(2!3!)$  permutations reduce to a sum of 3 simple forms:

$$\vec{\epsilon} = 1/10 W_{23}(6 h^2 p^2 \mathbf{p} + 3 p^2 \mathbf{HpH} + \mathbf{pHpHp}); \quad (3.11)$$

$$\vec{\epsilon} = 3/5 W_{23} h^2 p^3 (\hat{\delta} e^{i(\alpha-\beta)} + 1/2 \hat{\delta} e^{-i(\alpha-\beta)} + 1/6 \hat{\delta} e^{3i(\alpha-\beta)}) e^{i\beta}. \quad (3.12)$$

The sum of these 3 rotating vectors create a peanut shaped spot diagram centered on the field point as  $\mathbf{p}$  scans a round pupil. This aberration is sometimes called fifth order oblique spherical aberration because it is similar to spherical aberration in that the pattern is centered on the image point. It is also called cubic astigmatism.

The next 5<sup>th</sup> order term  $W_{32}(\mathbf{H}^3 \mathbf{p}^2)$  is called 5<sup>th</sup> order elliptical coma. It has the same form as (3.11), but with  $\mathbf{H}$  and  $\mathbf{p}$  interchanged.

$$\vec{\epsilon} = 1/10 W_{32}(6 h^2 p^2 \mathbf{H} + 3 h^2 \mathbf{pHp} + \mathbf{HpHpH}); \quad (3.13)$$

$$\vec{\epsilon} = 3/5 W_{32} h^3 p^2 (\hat{\delta} + 1/2 \hat{\delta} e^{2i(\alpha-\beta)} + 1/6 \hat{\delta} e^{-2i(\alpha-\beta)}) e^{i\beta}. \quad (3.14)$$

The first term is a displacement along the field vector, and the sum of the next two vectors create an ellipse. The major axis is along the field vector direction, and the ratio of major to minor axis is 3/1. This pattern has many of the characteristics of coma patterns and is also called cubic coma.

The last 5<sup>th</sup> order aberration is  $W_{41}(\mathbf{H}^4 \mathbf{p})$ ,

$$\vec{\epsilon} = W_{41}(\mathbf{H}^4 \mathbf{p}) = 1/5 W_{41}(3 h^4 \mathbf{p} + 2 h^2 \mathbf{HpH}); \quad (3.15)$$

$$\vec{\epsilon} = 3/5 W_{41} h^4 p (\hat{\delta} e^{i(\alpha-\beta)} + 2/3 \hat{\delta} e^{-i(\alpha-\beta)}) e^{i\beta}. \quad (3.16)$$

<sup>4</sup> Mathematica is a commercially available symbolic computing environment created by Wolfram Research.

The spot diagram for this 5<sup>th</sup> order aberration is created by ellipses centered on the field point. The major axis of the ellipse is oriented along the field vector  $\mathbf{H}$ ; the major to minor axis ratio is 5/1. This aberration has the same shape as 3<sup>rd</sup> order astigmatism but a different field dependence.

This analysis of spot diagrams is easily continued. A summary chart of ray aberration vectors and spot diagrams for 3<sup>rd</sup>, 5<sup>th</sup>, 7<sup>th</sup> order aberrations are summarized in the appendix. The basic shapes are constructed from sums of rotating vectors and vectors displaced along the field vector direction. The aberrations, which separate into two basic classes were usefully classified by Cruickshank and Hills[16]. Spot diagrams for aberrations that are even order in  $\mathbf{H}$  are all centered on the field point  $\mathbf{H}$ , and just resized as the magnitude of  $\mathbf{p}$  varies between 0 and 1. These patterns are similar to spherical aberration. Spot diagrams for aberrations which are odd order in  $\mathbf{H}$  are created by resizing the basic shape and displacing the resized basic shapes from the ideal image point along the direction of the field vector as the magnitude of  $\mathbf{p}$  varies between 0 and 1; these patterns are similar to coma.

## 4.0 Combining Aberrations

In the following sections we combine the first and third order aberrations to reproduce more classical results using geometric algebra.

### 4.1 Defocus + 3<sup>rd</sup> Order Spherical Aberration

When defocus and 3<sup>rd</sup> order spherical aberration are added together the ray aberration vector  $\vec{\epsilon}$  is given by (4.1).

$$\vec{\epsilon} = W_{10}\mathbf{H} + W_{01}\mathbf{p} + W_{03}\mathbf{p}^3 = W_{10}\mathbf{H} + (W_{01}\mathbf{p} + W_{03}\mathbf{p}^3) \hat{\delta} e^{i\alpha} \quad (4.1)$$

For round pupils this  $\vec{\epsilon}$  describes blur circles at the field point with a radius which depends on the amount of defocus, which is given by  $W_{10}$  or  $W_{01}$ . At the paraxial focus, *i.e.*, the Gaussian image plane;  $W_{10} = W_{01} = 0$ , and the blur circle radius is simply  $W_{03}\mathbf{p}^3$ , which has a maximum value of  $W_{03}$  when  $\mathbf{p} = 1$ . The marginal focus (really defocus) position is defined by the position of an image plane where the marginal rays ( $\mathbf{p} = 1$  rays) produce a zero radius blur circle. The amount of defocus required to reach the marginal focal plane is found by setting the coefficient of the second term, the blur radius, equal to 0 when  $\mathbf{p} = 1$ , *i.e.*,  $W_{10} + W_{03} = 0$  which from Figure (3.2) gives  $W_{10} = \Delta z / R = -W_{03}$ ; consequently, the distance from the paraxial image plane to the marginal image plane is  $\Delta z = -RW_{03}$ . On the marginal image plane the radius of the blur is given by (4.2),

$$r = |W_{10}\mathbf{p} + W_{03}\mathbf{p}^3| = |-W_{03}\mathbf{p} + W_{03}\mathbf{p}^3|, \quad (4.2)$$

which has a maximum when  $\mathbf{p} = 1/\sqrt{3}$ . At this maximum  $r = |2W_{03} / (3\sqrt{3})|$ . See Figure 4.1.

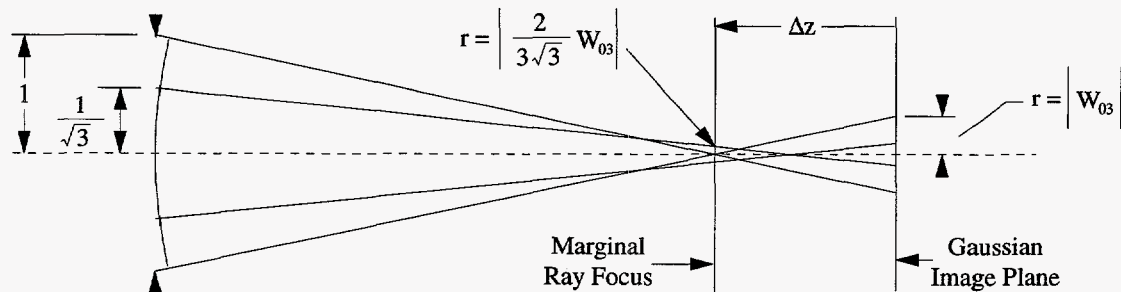


Figure 4.1 Defocus + 3<sup>rd</sup> Order Spherical Aberration

### 4.2 Defocus + 3<sup>rd</sup> Order Coma

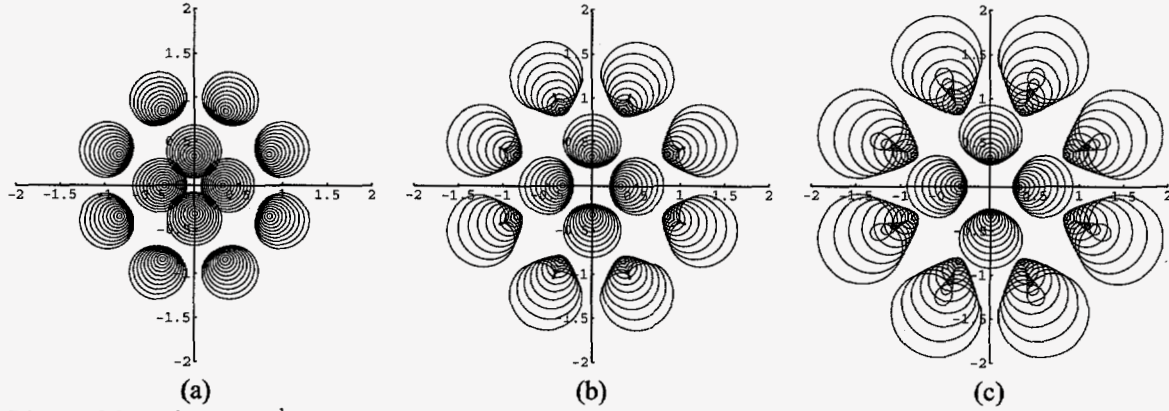
When defocus and 3<sup>rd</sup> order coma are added together the ray aberration vector  $\vec{\epsilon}$  is given by (4.3).

$$\vec{\epsilon} = W_{10}\mathbf{H} + W_{01}\mathbf{p} + 2/3 W_{12}(\mathbf{H}\mathbf{p}^2 + 1/2 \mathbf{p}\mathbf{H}\mathbf{p}) \quad (4.3)$$

Collecting and reducing terms leads to (4.4)

$$\begin{aligned} \vec{\epsilon} &= (W_{10} + 2/3 W_{12}p^2)\mathbf{H} + (W_{01}\mathbf{p} + 1/3 W_{12}\mathbf{p}\mathbf{H}\mathbf{p}) \\ &= (W_{10} + 2/3 W_{12}p^2) h \hat{\sigma} e^{i\beta} + (W_{01}p \hat{\sigma} e^{i\alpha} + 1/3 W_{12}p^2 h \hat{\sigma} e^{2i(\alpha-\beta)} e^{i\beta}) \\ &= (W_{10} + 2/3 W_{12}p^2) h \hat{\sigma} e^{i\beta} + (W_{01}p \hat{\sigma} e^{i(\alpha-\beta)} + 1/3 W_{12}p^2 h \hat{\sigma} e^{2i(\alpha-\beta)}) e^{i\beta} \\ &= W_{10} \{ 1 + [2W_{12}p^2 / (3W_{01})] \} h \hat{\sigma} e^{i\beta} + W_{10}p \{ \hat{\sigma} e^{i(\alpha-\beta)} + [W_{12}ph / (3W_{01})] \hat{\sigma} e^{2i(\alpha-\beta)} \} e^{i\beta} \end{aligned} \quad (4.4)$$

The first term in (4.4) is a displacement along the field vector, and the second term produces a Limascon of Pascal as shown in Figure 4.2. The size and orientation of the patterns depend on the amount and sign of the defocus and coma coefficients and the field position  $h$ .



**Figure 4.2** Defocus + 3<sup>rd</sup> Order Coma, for the same defocus but different amounts of astigmatism:  $W_{10} = W_{01} = 0.3$  and a)  $W_{12}/W_{01} = 3/4$ , b)  $W_{12}/W_{01} = 5/2$ , c)  $W_{12}/W_{01} = 4$ .

### 4.3 Defocus + 3<sup>rd</sup> Order Astigmatism

The results for 3<sup>rd</sup> order astigmatism on the Gaussian image plane are summarized by equation (3.8). When defocus is added to 3<sup>rd</sup> order astigmatism the ray aberration vector  $\vec{\epsilon}$  is given by (4.5).

$$\begin{aligned} \vec{\epsilon} &= W_{10}\mathbf{H} + W_{01}\mathbf{p} + 2/3 W_{21}(h^2\mathbf{p} + 1/2 \mathbf{H}\mathbf{p}\mathbf{H}) \\ &= W_{10}h \hat{\sigma} e^{i\beta} + (W_{01}p + 2/3 W_{21}h^2p) \hat{\sigma} e^{i\alpha} + 1/3 W_{21}h^2p \hat{\sigma} e^{-i(\alpha-\beta)} e^{i\beta} \\ &= W_{10}h \hat{\sigma} e^{i\beta} + \{ (W_{01}p + 2/3 W_{21}h^2p) \hat{\sigma} e^{i(\alpha-\beta)} + 1/3 W_{21}h^2p \hat{\sigma} e^{-i(\alpha-\beta)} \} e^{i\beta} \end{aligned} \quad (4.5)$$

Equation (4.5) has the same form as (3.8) except that the coefficients for the clockwise and counterclockwise spinning vectors are different in (4.5) than in (3.8); consequently, in general, we expect astigmatism to produce elliptical patterns whose size increases with  $h^2$ . There are however, some well known special cases.

#### Special Case I

We use the defocus parameter  $W_{10}$  to search for surfaces with special characteristic shapes. For example, a special case will result when the coefficients of the clockwise and counterclockwise spinning vectors are equal to each other. For this case, from (4.5),

$$W_{01} = -1/3 W_{21} h^2, \quad (4.6)$$

and thus, using (3.4) and (4.6), the defocus is given by,

$$\Delta z = W_{01} L = -1/3 L W_{21} h^2, \quad (4.7)$$

which describes a parabolic surface in  $\Delta z$  and  $h$  that is called the sagittal surface.

On the sagittal surface we easily determine that the characteristic shape of the aberrations are straight lines pointing along the field direction. Since the coefficients in (4.5) are equal because of (4.6), the length of the lines is given by  $2/3 W_{21} h^2$ . This is show in Figure 4.3.

### Special Case II

By adjusting the defocus parameter  $W_{01}$ , we can set the coefficient of the first spinning vector in (4.5) equal to zero; then, as  $\mathbf{p}$  scans the pupil, the characteristic patterns are circles with radius  $|1/3 W_{21} h^2 p|$ . The condition for the first coefficient in (4.5) to equal zero is given by (4.8), and the amount of defocus necessary to maintain this condition is given by (4.9)

$$W_{01} = -2/3 W_{21} h^2 \quad (4.8)$$

$$\Delta z = W_{01} L = -2/3 L W_{21} h^2 \quad (4.9)$$

Equation (4.9) shows that the circles will be produced on a parabolic surface called the medial surface.

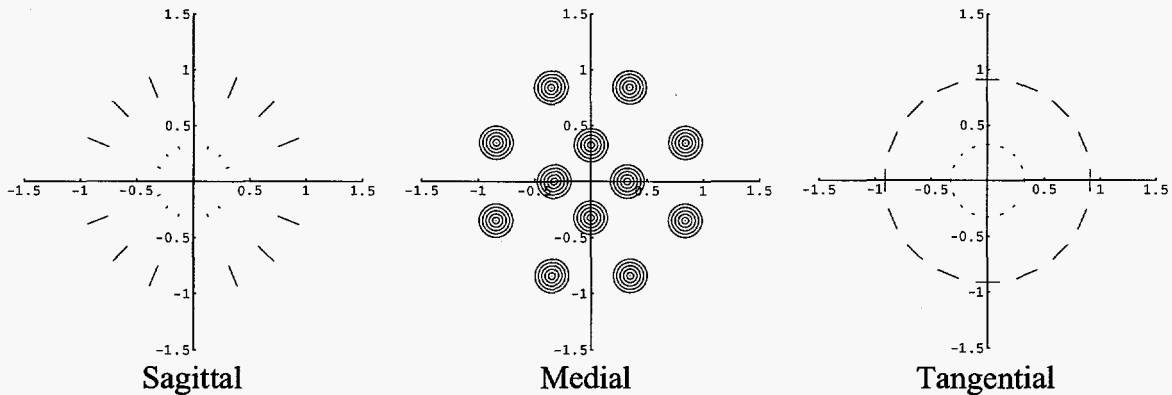
### Special Case III

When the defocus parameter  $W_{01}$  is adjusted so the coefficients of the spinning vectors are the negative of each other,

$$W_{01} = - W_{21} h^2 \quad (4.10)$$

$$\Delta z = W_{01} L = -L W_{21} h^2 \quad (4.11)$$

Again, the condition (4.10) for a particular shape is maintained on a parabolic surface, which in this case is called the tangential surface. When the spinning vectors are added they produce lines that are perpendicular  $\mathbf{H}$ . On both the sagittal and the tangential surfaces the line length is  $|2/3 W_{21} h^2|$ .

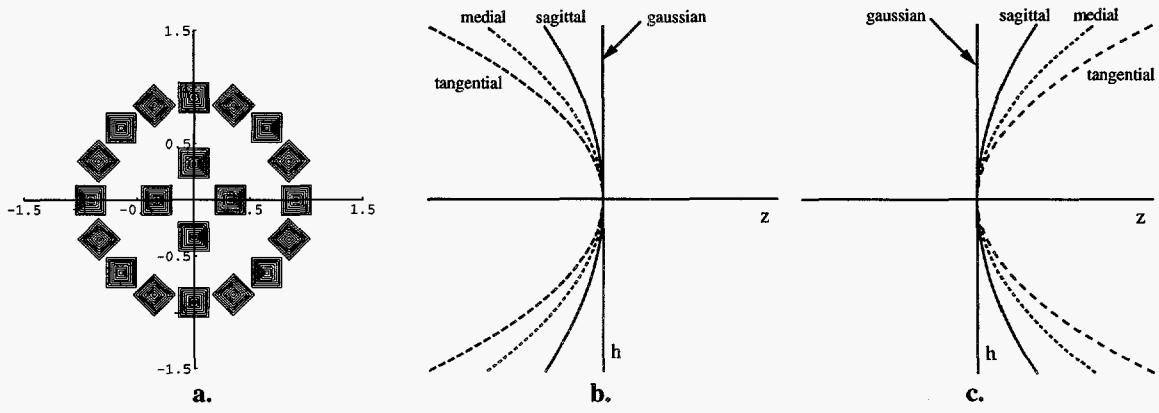


**Figure 4.3** Defocus and Astigmatism on the sagittal, medial and tangential spherical surfaces. The Gaussian image plane is always closest to the sagittal surface.

At a particular field height  $h$ , the distance between the Gaussian and sagittal surfaces, the sagittal and medial surfaces and the medial and tangential surfaces are equal. Starting at the Gaussian surface, the surfaces always occur in the order: Gaussian, sagittal, medial, tangential; however, the direction that the paraboloids open up depends on the sign of  $W_{21}$ .

Usually  $\Delta z$  is small and the paraboloids can be approximated by a section of a sphere. In this case we can treat  $\Delta z$  as the sag of a spherical surface away from a plane surface. The radius of the spherical surface,  $R$ , is related to the sag by (4.12). When we use (4.12) with (4.7), (4.9) and (4.11) we find the radius of each of the spherical surfaces that approximate the paraboloids near the optical axis.

$$\Delta z = h^2/(2R) \quad (4.12)$$



**Figure 4.4** a.) The astigmatism characteristic patterns on the medial surface when the pupil is a square. b.) and c.) schematically depict the order of the Gaussian, sagittal and tangential surfaces for both signs of the astigmatism coefficient. A defocused plane image surface will show characteristic shapes of all the special surfaces because it cuts through the parabolic surfaces.

The reasoning for square pupils is exactly the same and the characteristic shapes on the sagittal and tangential surfaces are again lines along and perpendicular to the direction to the field point. On the medial surface the characteristic shapes are squares, and the orientation of the squares on the spherical medial surface depends on the position in the field, as shown in Figure 4.4.

### 5.0 Conclusion

Optical aberration theory provides an interesting way to learn GA in an easy two dimensional setting that is readily understandable and easy to visualize; in return, GA provides an interesting and concise description of classical optical aberration theory. The noncommuting nature of GA is necessary to obtain all the classical well known results. We developed a theory for classical aberration shapes for square and round pupils could be extended to learn about the aberrations for other pupil shapes and aberration in optical systems with displaced pupils and tilted components following the theory already developed by Shack and Thompson. There also seems to be much in common between this approach to optical aberration theory and the aberration theory developed for particle beam optics using Lie algebra [17]. Given the fast PC's and elegant design codes currently available for designing optical systems, our results are primarily useful for learning about aberration types, developing intuition about design code results, exercising GA as a new analysis tool, and having some fun learning about old things in a new way.



## 6.0 Appendix

### 6.1 Proofs of Several Geometric Algebra Identities

To prove that  $\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) = (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - (\mathbf{a} \cdot \mathbf{c}) \mathbf{b}$ , one starts with  $\mathbf{a} \mathbf{b} \mathbf{c} = \mathbf{a} \mathbf{b} \mathbf{c}$ , which one rewrites as  $(\mathbf{a} \mathbf{b}) \mathbf{c} = \mathbf{a}(\mathbf{b} \mathbf{c})$ , and then expands using (2.1) to give

$$(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \mathbf{c} = \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) + \mathbf{a} (\mathbf{b} \wedge \mathbf{c}) \quad (6.1)$$

$$(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c} = \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) + \mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) + \mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) \quad (6.2)$$

Equating the trivector parts of the above equation gives the desired result

$$(\mathbf{a} \wedge \mathbf{b}) \wedge \mathbf{c} = \mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}).$$

To prove (2.15), one equates the remaining parts of equation (6.2) which gives

$$(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) + \mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) \quad (6.3)$$

One also expands  $\mathbf{a} \mathbf{b} \mathbf{c}$  using (2.1) and rewrites the result using (2.3) to give

$$\begin{aligned} (\mathbf{a} \mathbf{b}) \mathbf{c} &= (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} + (\mathbf{a} \wedge \mathbf{b}) \mathbf{c} = (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} + 1/2 (\mathbf{a} \mathbf{b} - \mathbf{b} \mathbf{a}) \mathbf{c} \\ (\mathbf{a} \mathbf{b}) \mathbf{c} &= 2(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - \mathbf{b} \mathbf{a} \mathbf{c} \end{aligned} \quad (6.4)$$

Starting instead with  $\mathbf{b} \mathbf{a} \mathbf{c}$ , one can likewise show that

$$\mathbf{b} (\mathbf{a} \mathbf{c}) = 2\mathbf{b} (\mathbf{a} \cdot \mathbf{c}) - \mathbf{b} \mathbf{c} \mathbf{a}$$

Using that result in equation (6.4) gives

$$\begin{aligned} (\mathbf{a} \mathbf{b}) \mathbf{c} &= 2(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - 2\mathbf{b} (\mathbf{a} \cdot \mathbf{c}) + \mathbf{b} \mathbf{c} \mathbf{a} \\ \mathbf{a} \mathbf{b} \mathbf{c} - \mathbf{b} \mathbf{c} \mathbf{a} &= 2(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - 2\mathbf{b} (\mathbf{a} \cdot \mathbf{c}) = 2(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - 2(\mathbf{a} \cdot \mathbf{c}) \mathbf{b}. \end{aligned} \quad (6.5)$$

The left hand side is then expanded using (2.1), (2.12), and recalling that the outer product of two vectors is a bivector to yield

$$\begin{aligned} \mathbf{a} \mathbf{b} \mathbf{c} - \mathbf{b} \mathbf{c} \mathbf{a} &= \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) + \mathbf{a} (\mathbf{b} \wedge \mathbf{c}) - (\mathbf{b} \cdot \mathbf{c}) \mathbf{a} - (\mathbf{b} \wedge \mathbf{c}) \mathbf{a} \\ &= \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) + \mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) + \mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) - (\mathbf{b} \cdot \mathbf{c}) \mathbf{a} - (\mathbf{b} \wedge \mathbf{c}) \cdot \mathbf{a} - (\mathbf{b} \wedge \mathbf{c}) \wedge \mathbf{a} \\ &= \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) - (\mathbf{b} \cdot \mathbf{c}) \mathbf{a} + \mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) - (\mathbf{b} \wedge \mathbf{c}) \cdot \mathbf{a} + \mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) - (\mathbf{b} \wedge \mathbf{c}) \wedge \mathbf{a} \end{aligned} \quad (6.6)$$

Equating the right hand sides of (6.5) and (6.6) and noting the trivector parts must be equivalent gives

$$\mathbf{a} \wedge (\mathbf{b} \wedge \mathbf{c}) - (\mathbf{b} \wedge \mathbf{c}) \wedge \mathbf{a} = 0. \quad (6.7)$$

Similarly, noting that the remaining parts must be equivalent gives

$$2(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - 2(\mathbf{a} \cdot \mathbf{c}) \mathbf{b} = \mathbf{a} (\mathbf{b} \cdot \mathbf{c}) - (\mathbf{b} \cdot \mathbf{c}) \mathbf{a} + \mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}) - (\mathbf{b} \wedge \mathbf{c}) \cdot \mathbf{a}. \quad (6.8)$$

Recalling that  $(\mathbf{b} \cdot \mathbf{c})$  is a scalar, and a scalar commutes with a vector [*i.e.*  $\mathbf{a} (\mathbf{b} \cdot \mathbf{c}) = (\mathbf{b} \cdot \mathbf{c}) \mathbf{a}$ ], and using (2.11), we can rewrite (6.8) as

$$2(\mathbf{a} \cdot \mathbf{b}) \mathbf{c} - 2(\mathbf{a} \cdot \mathbf{c}) \mathbf{b} = 2\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c}), \quad (6.9)$$

which is equivalent to the identity (2.15). This also shows that  $\mathbf{a} \cdot (\mathbf{b} \wedge \mathbf{c})$  and thus  $\mathbf{a} \cdot \mathbf{B}$  is a vector.

### 6.2 Wave Aberration Expansion

Wave aberration is an alternate way to describe the aberrations. In the wave aberration theory the ideal wavefront leaving the pupil is a perfect spherical surface. Aberrations add an additional phase to the wavefront in the exit pupil that depends on  $\mathbf{p}$  and  $\mathbf{H}$ . The additional phase shift caused by aberrations must be a scalar, and the most general scalar is given by



$$W = \sum_{c=1}^{\infty} \sum_{m=0}^c W_{m(c-m)} (\mathbf{H}^m \mathbf{p}^{c-m})$$

where the sum of the powers of  $\mathbf{H}$  and  $\mathbf{p}$  must be even. Just as for ray aberrations, the term  $(\mathbf{H}^m \mathbf{p}^{c-m})$ , represents the sum of the different ways that the vectors  $\mathbf{H}$  and  $\mathbf{p}$  can be rearranged because  $\mathbf{H}$  and  $\mathbf{p}$  do not commute. When  $c = 4$  the terms are  $(\mathbf{H}\mathbf{H}\mathbf{H}\mathbf{H})$ ,  $(\mathbf{H}\mathbf{H}\mathbf{H}\mathbf{p})$ ,  $(\mathbf{H}\mathbf{H}\mathbf{p}\mathbf{p})$ ,  $(\mathbf{H}\mathbf{p}\mathbf{p}\mathbf{p})$ , and  $(\mathbf{p}\mathbf{p}\mathbf{p}\mathbf{p})$ . When we use our Mathematica program *reduce* on each of these terms to find all the rearrangements, just as we did for ray aberrations, we get the following results, which exactly match classical wave aberration theory[18]:

$(\mathbf{H}\mathbf{H}\mathbf{H}\mathbf{H})$  reduces to  $h^4$ , which is constant over the entire pupil and is called a piston error;

$(\mathbf{p}\mathbf{p}\mathbf{p}\mathbf{p})$  reduces to  $p^4$ , which is spherical aberration;

$(\mathbf{H}\mathbf{p}\mathbf{p}\mathbf{p})$  reduces to  $2h p^3 \{e^{i(\alpha-\beta)} + e^{-i(\alpha-\beta)}\} = 4h p^3 \cos(\alpha-\beta)$ , which is coma;

$(\mathbf{H}\mathbf{H}\mathbf{p}\mathbf{p})$  reduces to  $4h^2 p^2 + h^2 p^2 \{e^{2i(\alpha-\beta)} + e^{-2i(\alpha-\beta)}\} = 2h^2 p^2 + 4h^2 p^2 \cos^2(\alpha-\beta)$ , which is a field curvature term plus astigmatism, and

$(\mathbf{H}\mathbf{H}\mathbf{H}\mathbf{p})$  reduces to  $2h^3 p \{e^{i(\alpha-\beta)} + e^{-i(\alpha-\beta)}\} = 4h^3 p \cos(\alpha-\beta)$ , which is distortion.

### 6.3 The Mathematica Code to Reduce Aberration Forms

The Mathematica code to reduce the forms for the aberrations is listed below. The program *newperm* takes a list to be reduced, *i.e.*,  $\{\mathbf{H}, \mathbf{p}, \mathbf{p}, \mathbf{p}\}$ , and compares the elements in pairs, starting with the first pair, *i.e.*, element 1,  $\mathbf{H}$ , and element 2,  $\mathbf{p}$ . Since these elements are not the same the comparison shifts one element to the right and compares element 2 and element 3, which are both  $\mathbf{p}$ 's; both  $\mathbf{p}$ 's are dropped from the list and the process starts over at the beginning and proceeds until the list pair has been compared.

```
newperm[list_List]:= →
Module[{i,list1},
list1=list /. {H → 1,p → 2};
i=1;
While[i ≤ Length[list1]-1,
If[list1[[i]] = list1[[i+1]],
list1=Drop[list1,{i,(i+1)}]; i=1,
i++]];
list1/.1 → H, 2 → p}]
```

The next section of code called *reduce* applies *newperm* to each permutation of the list, sorts out the list of different permutations, and counts the number of times each separate reduced permutation occurs in the list of reduced permutations.

```
reduce[list_List]:=
Module[{perm,list2,list3,coef},
perm=Permutations[list];
list2=Map[newperm,perm];
list3=Union[list2];
coef=Table[Count[list2,list3[[i]]],{i,1,Length[list3]}];
Print[{perm,list2,list3,coef}]
]
```

After typing the code into Mathematica, the results are calculated by typing `reduce[{list}]` followed by shift enter, where `{list}` is the list of symbols to be reduced.  $\{\mathbf{H}, \mathbf{H}, \mathbf{p}, \mathbf{p}, \mathbf{p}\}$  is an example of an acceptable list.

## 6.4 Aberration Summary

### Third Order Aberration Summary

1.) Third Order Spherical

$$W_{03} \mathbf{p}^3 = W_{03} p^3 \hat{\sigma} e^{i\alpha}$$

2.) Third Order Linear Coma

$$\begin{aligned} W_{12}(\mathbf{H} \mathbf{p}^2) &= 2/3 W_{12}(\mathbf{H} p^2 + 1/2 \mathbf{p} \mathbf{H} \mathbf{p}); \\ &= 2/3 W_{12} h p^2 (\hat{\sigma} + 1/2 \hat{\sigma} e^{2i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

3.) Third Order Linear Astigmatism

$$\begin{aligned} W_{21}(\mathbf{H}^2 \mathbf{p}) &= 2/3 W_{21}(h^2 \mathbf{p} + 1/2 \mathbf{H} \mathbf{p} \mathbf{H}); \\ &= 2/3 W_{21} h^2 p (\hat{\sigma} e^{i(\alpha-\beta)} + 1/2 \hat{\sigma} e^{-i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

4.) Third Order Distortion

$$W_{30} \mathbf{H}^3 = W_{30} h^3 \hat{\sigma} e^{i\beta}$$

### Fifth Order Aberration Summary

1.) Fifth Order Spherical Aberration

$$W_{05} \mathbf{p}^5 = W_{05} p^5 \hat{\sigma} e^{i\alpha}$$

2.) Fifth Order Linear Coma

$$\begin{aligned} W_{14}(\mathbf{H} \mathbf{p}^4) &= 3/5 W_{14}(\mathbf{H} p^4 + 2/3 p^2 \mathbf{p} \mathbf{H} \mathbf{p}); \\ &= 3/5 W_{14} h p^4 (\hat{\sigma} + 2/3 \hat{\sigma} e^{2i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

3.) Fifth Order Cubic Astigmatism (Oblique Spherical Aberration)

$$\begin{aligned} W_{23}(\mathbf{H}^2 \mathbf{p}^3) &= 3/5 W_{23}(h^2 p^2 \mathbf{p} + 1/2 p^2 \mathbf{H} \mathbf{p} \mathbf{H} + 1/6 \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p}); \\ &= 3/5 W_{23} h^2 p^3 (\hat{\sigma} e^{i(\alpha-\beta)} + 1/2 \hat{\sigma} e^{-i(\alpha-\beta)} + 1/6 \hat{\sigma} e^{3i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

4.) Fifth Order Cubic Coma (Elliptical Coma)

$$\begin{aligned} W_{32}(\mathbf{H}^3 \mathbf{p}^2) &= 3/5 W_{32}(h^2 p^2 \mathbf{H} + 1/2 h^2 \mathbf{p} \mathbf{H} \mathbf{p} + 1/6 \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H}); \\ &= 3/5 W_{32} h^3 p^2 (\hat{\sigma} + 1/2 \hat{\sigma} e^{2i(\alpha-\beta)} + 1/6 \hat{\sigma} e^{-2i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

5.) Fifth Order Linear Astigmatism

$$\begin{aligned} W_{41}(\mathbf{H}^4 \mathbf{p}) &= 3/5 W_{41}(h^4 \mathbf{p} + 2/3 h^2 \mathbf{H} \mathbf{p} \mathbf{H}); \\ &= 3/5 W_{41} h^4 p (\hat{\sigma} e^{i(\alpha-\beta)} + 2/3 \hat{\sigma} e^{-i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

6.) Fifth Order Distortion

$$W_{50} \mathbf{H}^5 = W_{50} h^5 \hat{\sigma} e^{i\beta}$$

### Seventh Order Aberration Summary

1.) Spherical

$$W_{07} \mathbf{p}^7 = W_{07} \mathbf{p}^7 \hat{\delta} e^{i\alpha}$$

2.) Linear Coma

$$\begin{aligned} W_{16}(\mathbf{H} \mathbf{p}^6) &= 4/7 W_{16}(\mathbf{H} \mathbf{p}^6 + 3/4 \mathbf{p}^4 \mathbf{p} \mathbf{H} \mathbf{p}); \\ &= 4/7 W_{16} h \mathbf{p}^6 (\hat{\delta} + 3/4 \hat{\delta} e^{2i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

3.) Quintic Astigmatism

$$\begin{aligned} W_{25}(\mathbf{H}^2 \mathbf{p}^5) &= 12/21 W_{25}(h^2 \mathbf{p}^4 \mathbf{p} + 1/2 \mathbf{p}^4 \mathbf{H} \mathbf{p} \mathbf{H} + 1/4 \mathbf{p}^2 \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p}); \\ &= 4/7 W_{25} h^2 \mathbf{p}^5 (\hat{\delta} e^{i(\alpha-\beta)} + 1/2 \hat{\delta} e^{-i(\alpha-\beta)} + 1/4 \hat{\delta} e^{3i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

4.) Cubic Coma

$$\begin{aligned} W_{34}(\mathbf{H}^3 \mathbf{p}^4) &= 18/35 W_{34} (h^2 \mathbf{p}^4 \mathbf{H} + 2/3 h^2 \mathbf{p}^2 \mathbf{p} \mathbf{H} \mathbf{p} + 2/9 \mathbf{p}^2 \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} + 1/18 \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p}); \\ &= 18/35 W_{34} h^3 \mathbf{p}^4 (\hat{\delta} + 2/3 \hat{\delta} e^{2i(\alpha-\beta)} + 2/9 \hat{\delta} e^{-2i(\alpha-\beta)} + \hat{\delta} e^{4i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

5.) Cubic Astigmatism

$$\begin{aligned} W_{43}(\mathbf{H}^4 \mathbf{p}^3) &= 18/35 W_{43} (h^4 \mathbf{p}^2 \mathbf{p} + 2/3 \mathbf{p}^2 h^2 \mathbf{H} \mathbf{p} \mathbf{H} + 2/9 h^2 \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} + 1/18 \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H}); \\ &= 18/35 W_{43} h^4 \mathbf{p}^3 (\hat{\delta} e^{i(\alpha-\beta)} + 2/3 \hat{\delta} e^{-i(\alpha-\beta)} + 2/9 \hat{\delta} e^{3i(\alpha-\beta)} + \hat{\delta} e^{-3i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

6.) Quintic Coma

$$\begin{aligned} W_{52}(\mathbf{H}^5 \mathbf{p}^2) &= 12/21 W_{52} (h^4 \mathbf{p}^2 \mathbf{H} + 1/2 h^4 \mathbf{p} \mathbf{H} \mathbf{p} + 1/4 h^2 \mathbf{H} \mathbf{p} \mathbf{H} \mathbf{p} \mathbf{H}); \\ &= 4/7 W_{52} h^5 \mathbf{p}^2 (\hat{\delta} + 1/2 \hat{\delta} e^{2i(\alpha-\beta)} + 1/4 \hat{\delta} e^{-2i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

7.) Linear Astigmatism

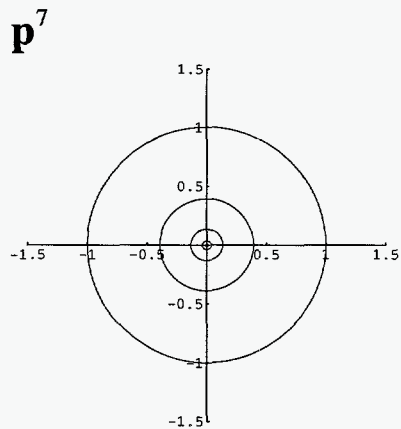
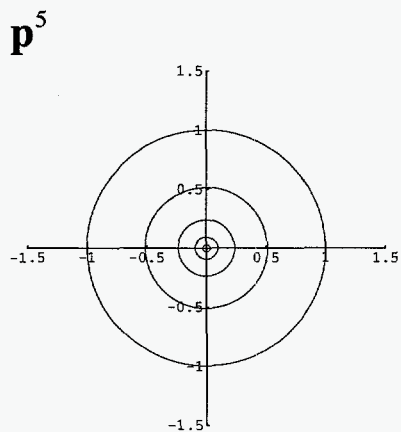
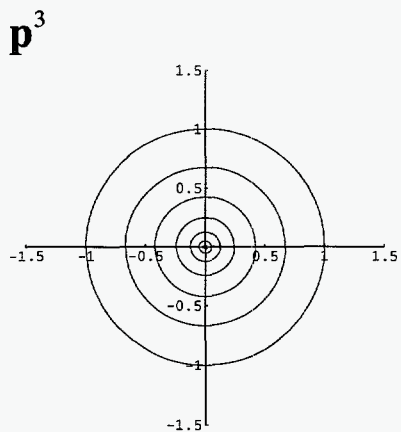
$$\begin{aligned} W_{61}(\mathbf{H}^6 \mathbf{p}) &= 4/7 W_{61} (h^6 \mathbf{p} + 3/4 h^4 \mathbf{H} \mathbf{p} \mathbf{H}); \\ &= 4/7 W_{61} h^6 \mathbf{p} (\hat{\delta} e^{i(\alpha-\beta)} + 3/4 \hat{\delta} e^{-i(\alpha-\beta)}) e^{i\beta} \end{aligned}$$

8.) Distortion

$$W_{70} \mathbf{H}^7 = W_{70} h^7 \hat{\delta} e^{i\beta}$$

## 6.5 Spot Diagrams for Round and Square Pupils

### Spherical Aberrations S Type Round Pupil



# Coma Aberrations

## C Type

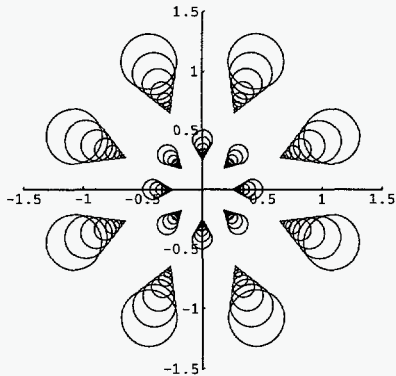
### Round Pupil

**Linear Coma**

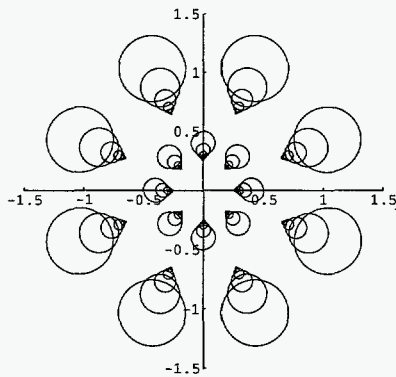
**Cubic Coma**

**Quintic Coma**

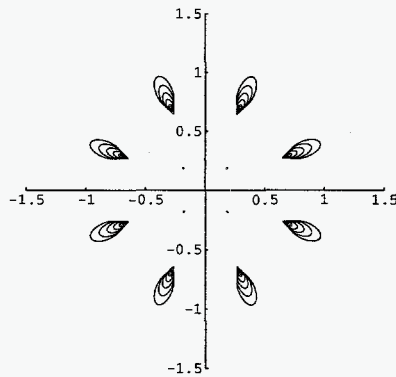
**$Hp^2$**



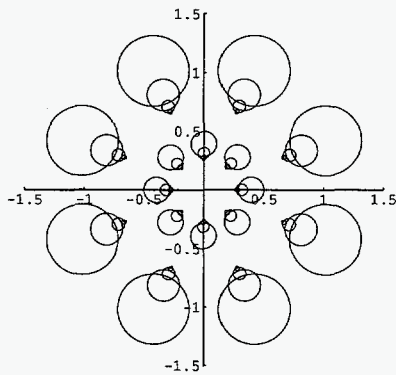
**$Hp^4$**



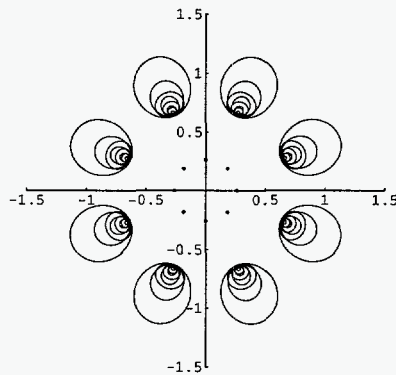
**$H^3p^2$**



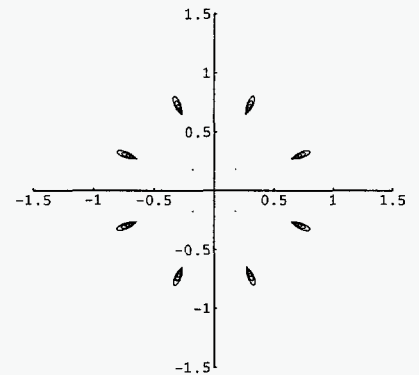
**$Hp^6$**



**$H^3p^4$**



**$H^5p^2$**



# Astigmatism Aberrations

## S Type

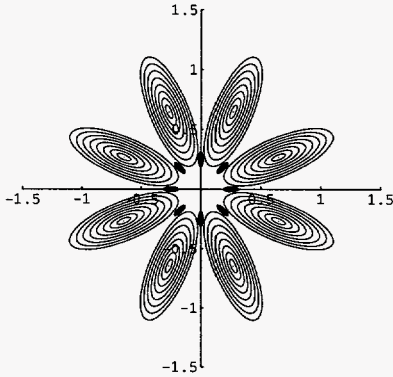
### Round Pupil

**Linear Astigmatism**

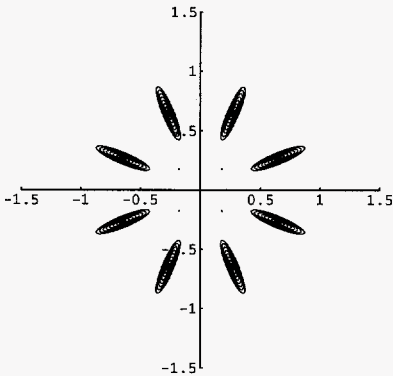
**Cubic Astigmatism**

**Quintic Astigmatism**

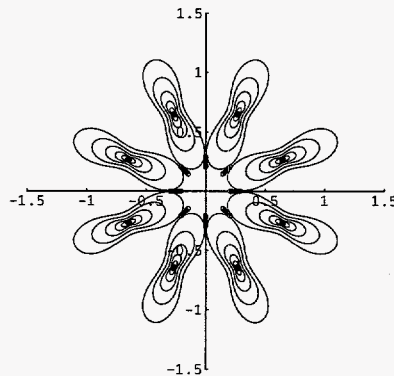
**$H^2p$**



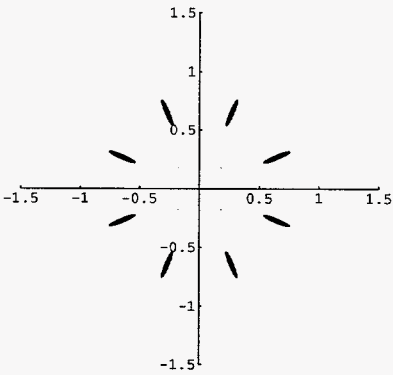
**$H^4p$**



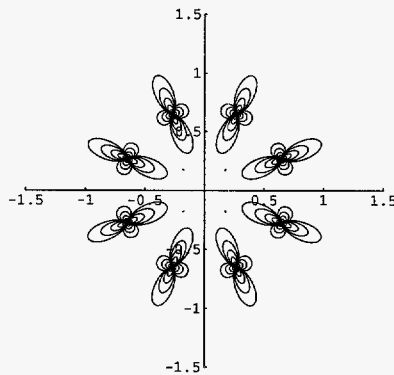
**$H^2p^3$**



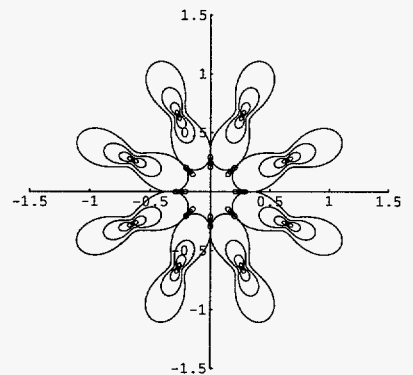
**$H^6p$**



**$H^4p^3$**



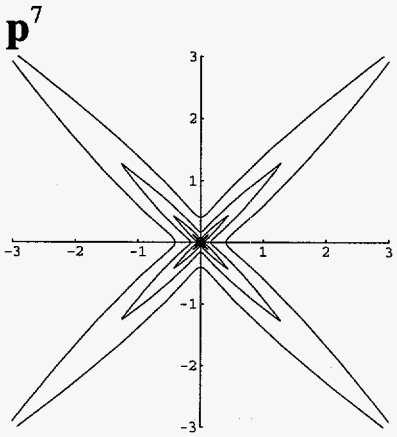
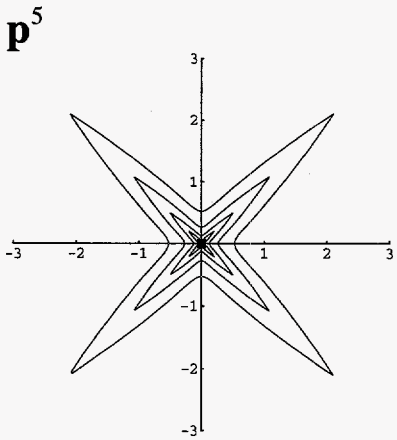
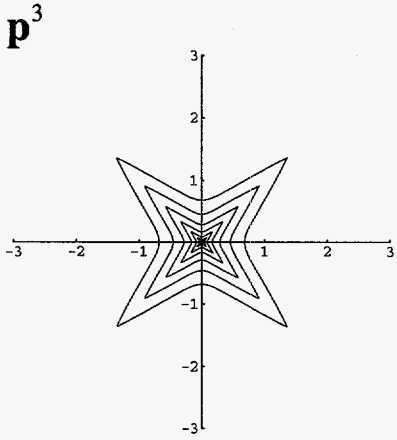
**$H^2p^5$**



# Spherical Aberrations

## S Type

### Square Pupil



# Coma Aberrations

## C Type

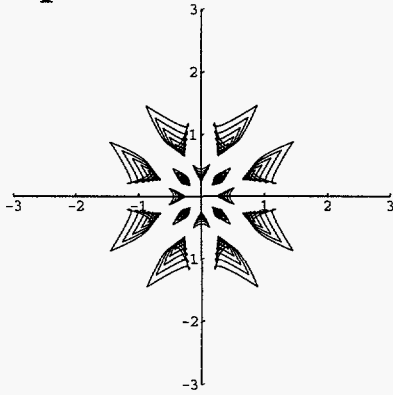
### Square Pupil

**Linear Coma**

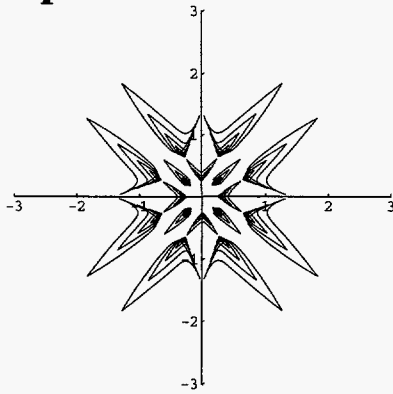
**Cubic Coma**

**Quintic Coma**

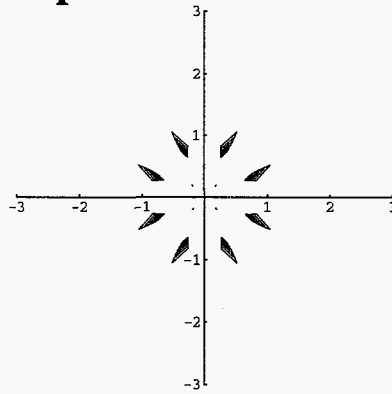
$Hp^2$



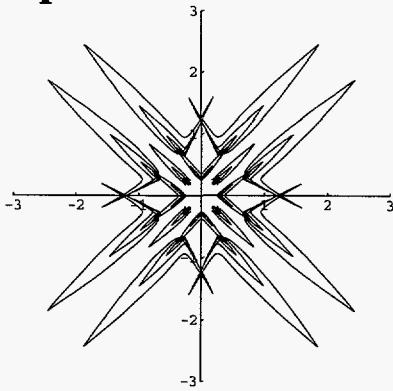
$Hp^4$



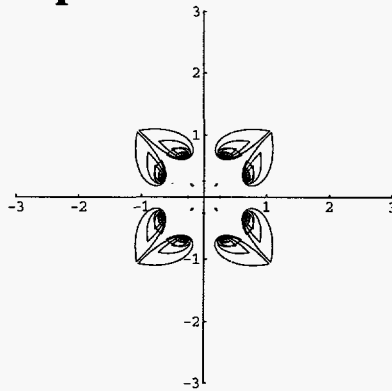
$H^3p^2$



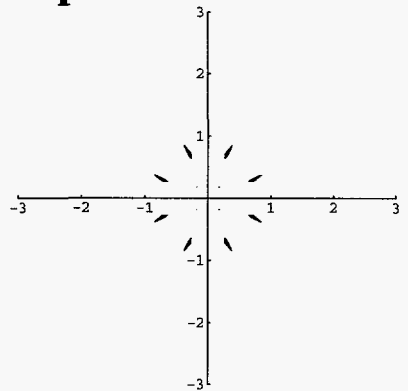
$Hp^6$



$H^3p^4$



$H^5p^2$





# Astigmatism Aberrations

## S Type

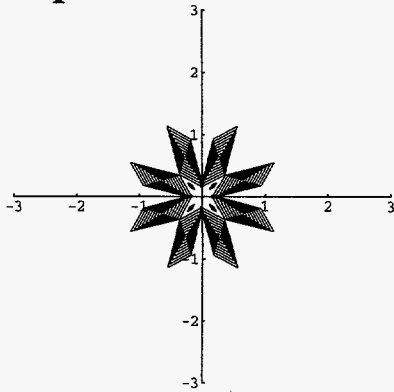
### Square Pupil

**Linear Astigmatism**

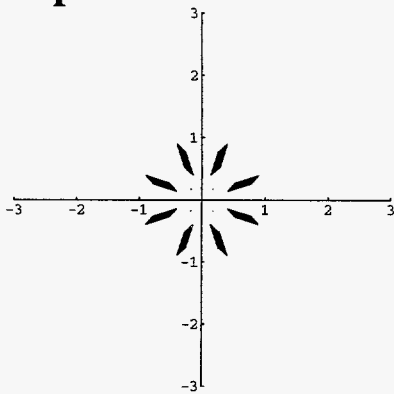
**Cubic Astigmatism**

**Quintic Astigmatism**

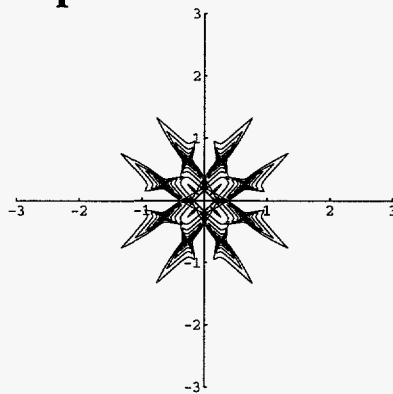
$H^2p$



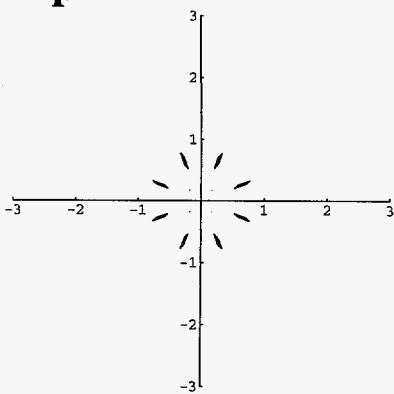
$H^4p$



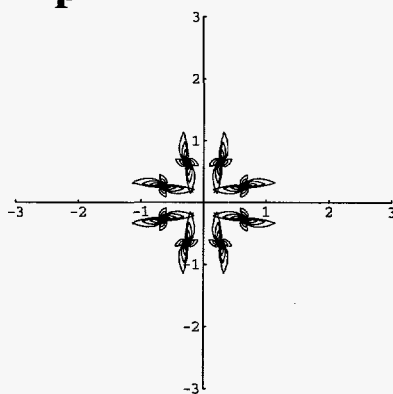
$H^2p^3$



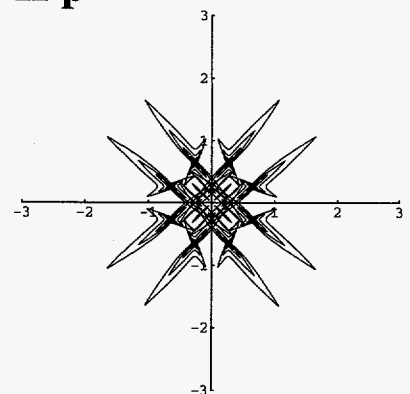
$H^6p$



$H^4p^3$



$H^2p^5$



## 7.0. References

1. Hestenes, David, "New Foundation for Classical Mechanics", (Reidel, 1986-87)
2. Jancewicz, Bernard, "Multivectors and Clifford Algebra in Electrodynamics", (World Scientific Publishing Co., Pet. Ltd., 1988)
3. Gull, S. , Lasenby, A., Doran, C., "Imaginary Numbers Are Not Real-The Geometric Algebra of Space Time", Foundations of Physics, Vol. 23, No. 9, 1993, 1175-1201.
4. Baylis, W.E., Huschilt, J., Wei, Jiansu, "Why i?", Am. J. Phys, 60 (9), Sept. 92, 788-797.
5. Vold, T.G., "An introduction to geometric algebra with application in rigid body mechanics", Am. J. Phys. 61, (6), June 93, 491-504.
6. Gough, W., "Mixing Scalars and Vectors—an Elegant View of Physics", Eur. J. Phys., 11 (1990) 326-333.
7. Baylis, William E., Editor, "Clifford (Geometric) Algebras, with Applications in Physics, Mathematics, and Engineering", (Birkhauser, Boston , 1996).
8. Manthey, M., "Toward and Information Mechanics", 1994 IEEE.
9. Baylis, W.E., Bonenfant, J., Derbyshire, J. and Huschilt, J., "Light Polarization: A geometric algebra approach ", Am. J. Phys. 61 (6), June 93, 534-545.
10. Hanlon, J., and Ziocck, H., "Using Geometric Algebra to Understand Pattern Rotations in Multiple Mirror Optical Systems". To be Published.
11. Shack, R.V., "Lectures in Optical Design", Optical Science Center, University of Arizona (1977).
12. Bucroeder, R.A., "Tilted Component Optical Systems", PhD Dissertation, Optical Sciences Center, University of Arizona (1976).
13. Thompson, K.P., "Aberration Fields in Tilted and Decentered Optical Systems", PhD Dissertation, Optical sciences Center, University of Arizona (1980).
14. Sassian, Jose M., "How to Approach the Design of a Bilateral Symmetric Optical System", Optical Engineering, June 1994, Vol. 33, No. 6, 2045
15. Lawrence Livermore National Laboratory, "National Ignition Facility Conceptual Design Report", NIF-LLNL-94, (L-17026-1), Livermore, CA., April 1994.
16. Cruickshank, F.D. and Hills, G.A., "Use of Optical Aberration Coefficients in Optical Design", Journal of the Optical Society of America, April 1960, Vol.50,(4),379-387.
17. Dragt, Alex J. and Forest, E., "Lie Algebraic Theory of Charged-Particle Optics and Electron Microscopes", Advances in Electronics and Electron Physics, Vol. 67, Academic Press, Inc. 1986.
18. Rogers, Handbook of Microwave and Optical components, Editor Chang pg 125