

# Choosing Mesh Spacings and Mesh Dimensions for Wave Optics Simulation

Steve Coy\*

MZA Associates Corporation 2021 Girard Blvd. SE, Suite 150, Albuquerque, NM 87106-3140

## ABSTRACT

One of the most common and important tasks in wave optics simulation is choosing what mesh spacings and mesh dimensions<sup>1</sup> to use for a given problem. To obtain correct results, it is crucial that the mesh spacings are sufficiently small and the mesh dimensions sufficiently large, but if one makes the spacings too small, or the dimensions too large, that can greatly increase the simulation run time, and that may be unaffordable. It is therefore important to understand exactly what the applicable constraints are, so that one may choose mesh spacings and dimensions that will yield correct results without being over-conservative. However this problem can be nontrivial, especially when modeling propagation through aberrating media, or when there is potentially useful a priori information available which might allow us to relax the modeling constraints. For example, if the light source is known to be well-collimated, we know that all of the light to be modeled will be concentrated along one axis, allowing us to use smaller meshes than we would if the light were uncollimated. Similarly, if the receiver has a limited field of view, we need not model any light incident upon it from angles outside its field of view. In this paper we present a simple general method to determine what mesh spacings and dimensions will work for any given wave optics propagation problem, including problems involving propagation through aberrating media and/or a priori information about the source and/or receiver.

## 1.0 INTRODUCTION

In wave optics simulation optical wavefronts are modeled using complex meshes where the complex amplitude represents the amplitude of the optical field, while the complex phase represents the difference between the phase of optical field and that of some planar or spherical *reference wave*. These wavefronts are propagated from plane to plane the optical axis using algorithms based on the Discrete Fourier Transform, or DFT, and optical components and effects are modeled by operating on the meshes in between propagations. Using these techniques it is possible to obtain highly accurate simulation results for many different kinds of optical systems, but only if the mesh spacings and mesh dimensions are chosen correctly. If the mesh spacings are too large or the mesh dimensions too small, the results will be incorrect. On the other hand, if the mesh spacings are too small, or the mesh dimensions made too large, the simulation run time can increase dramatically, even to the point where it becomes impractical to carry out the calculation. In this paper we present a general method for determining exactly what the constraints are that govern the choice of mesh parameters for any given wave optics modeling problem, taking into account any available a priori information concerning the nature of the source, the receiver, and the intervening optical path. The method can be used to determine the choice of mesh parameters that will minimize the mesh dimension, and thus the simulation execution time, while still obtaining correct<sup>2</sup> results. User-defined requirements can also be folded in, such as requiring that the initial and final mesh spacings be the same, or that they should be assigned “convenient” values, e.g. integer multiples of 1 cm. The method is amenable to computer automation, and we plan to incorporate it in WaveTrain<sup>TM</sup>, a tool for wave optics modeling of advanced optical systems. For the sake of brevity, we will present our results with some explanatory discussion, but without complete derivations.

The method we have developed relies upon the use of *complex rays* and *limiting apertures* to define both the set of rays to be modeled in the given problem, which we shall call the *ray region to be modeled*, and the set of rays that can be correctly modeled for a given set of mesh parameters, which we shall call the *region of validity*. Once the problem has

---

\* [coy@mza.com](mailto:coy@mza.com); phone 1 505 245-9970 ext. 115; fax 1 505 245-9971; mza.com

<sup>1</sup> By “mesh dimension” we mean the number of mesh points along one side; to refer to the physical size of the mesh we will use the term “mesh extent”.

<sup>2</sup> By “correct” we mean without errors related to wraparound or sampling not satisfying the Nyquist criterion; there will still be some approximation error due to discrete sampling of the continuous wavefront.

been cast in this form, the condition for correct modeling can be stated simply: the region to be modeled must be a subset of region of validity. The use of rays in the context of wave optics may seem surprising, since rays are more often associated with the traditional *ray optics*, which is less physically realistic than wave optics, but when one associates a complex amplitude with each ray, as opposed to a real-valued intensity, it can be shown that this ray-based description is mathematically equivalent to the more usual wave-based description of scalar diffraction theory. The advantage of using complex rays is that, because each ray crosses all planes, it makes it much easier to take into account geometric constraints present at different planes. In particular, we can define both the region to be modeled and the region of validity in terms of two sets of limiting apertures. In the case of the region to be modeled, these may be actual physical apertures and/or *effective apertures* used to take into account a priori information tending to limit the set of rays that must be modeled. For example, if the source is known to be a laser with some given beam quality passing through a beam waist at some given plane, we know that essentially all of its energy will pass through an effective aperture at that plane that is somewhat larger than the waist size. Similarly, if the light is impinging on a camera with a restricted field of view, the boundary of the field of view can be represented by another effective aperture at the camera's image plane. In the case of the region of validity, the relevant apertures correspond to the mesh extents at the initial and final planes, and also, when the propagation is carried out in multiple steps, at any and all intermediate planes.

We have devised two types of diagrams we find useful for analyzing these problems: *propagation diagrams* and *ray region diagrams*. Both types of diagrams deal with only one transverse dimension at a time, which can be thought of as either  $x$  or  $y$ , where  $z$  is the nominal optical axis; for our purposes the two transverse dimensions are interchangeable and independent. In propagation diagrams, we show a cross-sectional view of the optical system, indicating the limiting apertures for the region to be modeled, the rays connecting them, the meshes at the initial, final, and any intermediate planes, the phasefronts of the reference wave(s) at each mesh plane, and normals to those phasefronts. These diagrams are useful for understanding how the geometry of the problem relates to the constraints governing the choice of mesh parameters, and in some relatively simple cases those constraints can almost be read directly off the diagrams. In ray region diagrams, we plot the set of complex rays passing through a given set of limiting apertures as a contiguous region in slope-intercept space, where the ray intercepts are those for some particular plane along the optical axis, chosen for convenience. (The ray slopes are independent of the plane chosen.) In these diagrams both the ray region to be modeled for the given problem and the region of validity for a given choice of mesh parameters appear as polygons. The first is fixed, defined by the problem statement, but the second is subject to our control. By varying the mesh parameters we can adjust the size and shape of the region of validity so as to ensure that it contains the entire ray region to be modeled; that, in a nutshell, is our method.

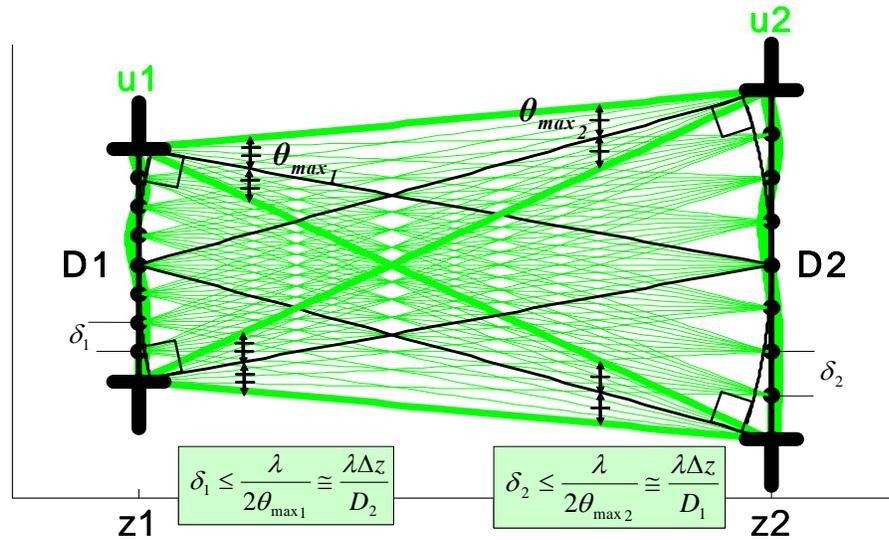
Strictly speaking, a set of complex rays remains the same only within a contiguous volume filled with vacuum or a uniform dielectric medium; as soon as the light is perturbed in any way, e.g. by an aperture or a phase screen, the set of rays changes. Therefore, when modeling propagation through an optical system or through an aberrating medium, one must take into account the resulting series of changes in the ray region to be modeled. For an ideal optical system all the effects of the lenses and mirrors may be removed by simply picking one plane to work in, e.g. the source plane, and then replacing the actual receiver, apertures, aberrations, and so forth with their images as seen through the intervening optics. Any non-ideal optical effects, e.g. aberrations in the lenses, must be retained in the re-imaged model. Limiting apertures, when seen through nearer apertures, must be enlarged to take into account blurring due to diffraction at the edges of the nearer apertures. (Blurring can be thought of as bending of the complex rays.) Similarly, limiting apertures seen through random aberrating media, such as the earth's turbulent atmosphere, must be enlarged to take into account blurring due to those random aberrations. With these relatively minor adjustments, we can now apply the same method, including both propagation diagrams and ray region diagrams, to any wave optics problem.

Using this method one can reliably choose mesh spacings and mesh dimensions appropriate for any given wave optics propagation problem – however we hasten to point out that that does not suffice to guarantee correct results. Our method implicitly assumes that the initial representation of the light leaving the source will be correctly constructed on the chosen source plane mesh, and containing only that portion of the light leaving the source that will make it through all the previously determined limiting apertures (taking into account diffraction effects) before arriving at the source. If there is more light leaving the source, the extra light must be removed. Also, there are numerical artifacts associated with DFT propagation and modeling propagation through aberrating media that can affect simulation results unless

appropriate precautions are taken. A full treatment of these issues would be beyond the scope of the present paper, but we will discuss them briefly in Section 5, Related Issues.

## 2.0 PROPAGATION DIAGRAMS AND ONE-STEP AND TWO-STEP DFT PROPAGATION

Let us consider a relatively simple wave optics problem in which there are two limiting apertures,  $D_1$  and  $D_2$ , we know the optical field  $u_1$  at the plane of the first aperture, and we wish to obtain the corresponding propagated field,  $u_2$ , at the plane of the second aperture, where the space between the two apertures is assumed to be filled with vacuum or an ideal dielectric medium.  $u_1$  is assumed to be zero everywhere outside the aperture  $D_1$ . The propagated field  $u_2$  is not assumed to be zero outside  $D_2$ , but we do assume that  $D_2$  defines the region over which we wish to determine the propagated field. To obtain the propagated field we will use a one-step DFT propagation. The propagation diagram for this problem is shown in Figure 1:



**Figure 1. Propagation diagram for a simple wave optics problem.**

In a propagation diagram, as shown in Figure 1, the nominal optical axis ( $z$ ) is horizontal, while the transverse dimension ( $x$  or  $y$ ) is vertical. Each limiting aperture is indicated by a pair of “T” shapes, one upright, one inverted, while the two complex meshes at the initial and final plane are indicated by vertical dotted lines. In this case the two meshes are shown exactly filling the two apertures. In a one-step DFT propagation the reference wave for the mesh at the initial plane ( $z_1$ ) is a spherical wave centered on the mesh at the final plane ( $z_2$ ), and vice versa; these reflect the two quadratic phase factors that appear in the Fresnel diffraction integral. The reference wave phasefronts for each mesh are indicated by thin black arcs, but the one at the plane  $z_1$  is hard to see because it overlaps the dotted line for the mesh itself. Also shown in black are the normals to reference wave at the edges of the mesh; these are important because the maximum allowable mesh spacing for each mesh is defined by the maximum angle formed between its reference normal and any ray that is to be modeled. A representative subset of the rays to be modeled is indicated by thin green lines connecting points in the two apertures. Four rays, connecting the extreme edges of the aperture are shown as thicker green lines. These rays form the largest angles with the reference normals and thus control the maximum allowable mesh spacing. The constraint follows directly from the Nyquist criterion: the mesh spacing  $\delta$  must be small

enough so that the change in phase between adjacent meshpoints will be less than or equal to half a wave for the highest angle light to be represented:

$$\delta \cdot \theta_{\max} \leq \frac{\lambda}{2}, \quad \delta \leq \frac{\lambda}{2\theta_{\max}} \quad (1)$$

For the example illustrated, the maximum angles for the two meshes can be read directly off the diagram; they are just the angular subtenses of the two apertures, as seen by one another. This gives us the constraints on the mesh spacings for the initial and final meshes:

$$\delta_1 \leq \frac{\lambda}{2\theta_{\max 1}} \cong \frac{\lambda \Delta z}{D_2}, \quad \delta_2 \leq \frac{\lambda}{2\theta_{\max 2}} \cong \frac{\lambda \Delta z}{D_1} \quad (2)$$

Finally, we can obtain the constraint on the mesh dimension  $N$ , by applying the requirement that all meshes must be large enough to capture all the rays to be modeled. In this example, there are just two meshes, located at the same planes as the two apertures, so we need only ensure that the two meshes are larger than the two apertures:

$$N \geq \max\left(\frac{D_1}{\delta_1}, \frac{D_2}{\delta_2}\right) \geq \frac{D_1 D_2}{\lambda \Delta z} \quad (3)$$

Of course by definition  $N$  must be an integer, and if we wish to compute the DFT using the Fast Fourier Transform (FFT),  $N$  must be a power of two, or a product of powers of two and other low primes, depending on what FFT implementation we use. Finally, when using one-step DFT propagation we cannot choose the initial and final mesh spacings independently; instead they must be chosen so as to satisfy the following equation:  $\delta_1 \delta_2 = \lambda \Delta z / N$ . This limitation is one reason why it is sometimes preferable to use two-step DFT propagation, where the initial and final mesh spacings *can* be chosen independently.

A two-step propagation consists of two consecutive propagations, first from the initial plane to some intermediate plane, and then from that intermediate plane to some final plane. By adjusting the position of the intermediate plane ( $z_{im}$ ), we can control the propagation magnification,  $M$ , the ratio of the final mesh spacings to the initial mesh spacing:

$$\delta_2 = \frac{\lambda |z_2 - z_{im}|}{N \delta_{im}} = \frac{\lambda |z_2 - z_{im}|}{N \frac{\lambda |z_{im} - z_1|}{N \delta_1}} = \frac{|z_2 - z_{im}|}{|z_{im} - z_1|} \delta_1 \quad (4)$$

$$M = \frac{\delta_2}{\delta_1} = \frac{|z_2 - z_{im}|}{|z_{im} - z_1|} \quad (5)$$

Next, we can solve the above to determine the location of the intermediate plane:

Letting

$$\Delta z = z_2 - z_1, \quad \Delta z_1 = z_{im} - z_1, \quad \Delta z_2 = z_2 - z_{im} \quad (6)$$

We obtain

$$\frac{\Delta z_1}{\Delta z_2} = m, \quad \frac{\Delta z_1}{\Delta z} = \frac{1}{1 \pm m}, \quad \frac{\Delta z_2}{\Delta z} = \frac{m}{1 \pm m} \quad (7)$$

This implies that for any given value of  $M$  there are two possible choices for  $z_{itm}$ , one in between  $z_1$  and  $z_2$ , and one outside that interval, corresponding to the planes of the two points that project the two meshes onto on another. In the first case the two one-step propagations are in the same direction, in the second they are in opposite directions. In both cases the reference wave for both the initial and final mesh is a spherical wave centered on the projection point at the intermediate plane. There is one special case that deserves mention: in the limit where the intermediate plane goes to either plus or minus infinity,  $M \rightarrow 1.0$ ,  $\delta_2 \rightarrow \delta_1$ , the intermediate mesh becomes infinite in size but remains finite in angular extent, and the reference wave is planar; this is sometimes called *frequency domain propagation*. For all other cases the intermediate plane is at some finite distance and the intermediate mesh is finite in extent, which is convenient when drawing propagation diagrams.

Figure 2 contrasts two-step propagation with one-step propagation for the same case considered earlier, showing both of the available choices for the intermediate plane. For the sake of simplicity we show only the rays connecting the extreme edges of the two apertures; as before, these are the rays that drive the sampling requirements. For the case of two-step propagation using the intermediate outside the interval  $[z_1, z_2]$  we have extended these rays, shown dashed, to the intermediate plane. For both of the two-step propagations we have chosen to show the mesh at the intermediate plane as being just wide enough to capture all the rays connecting the two apertures; this is the minimum size we can use and still obtain correct results.

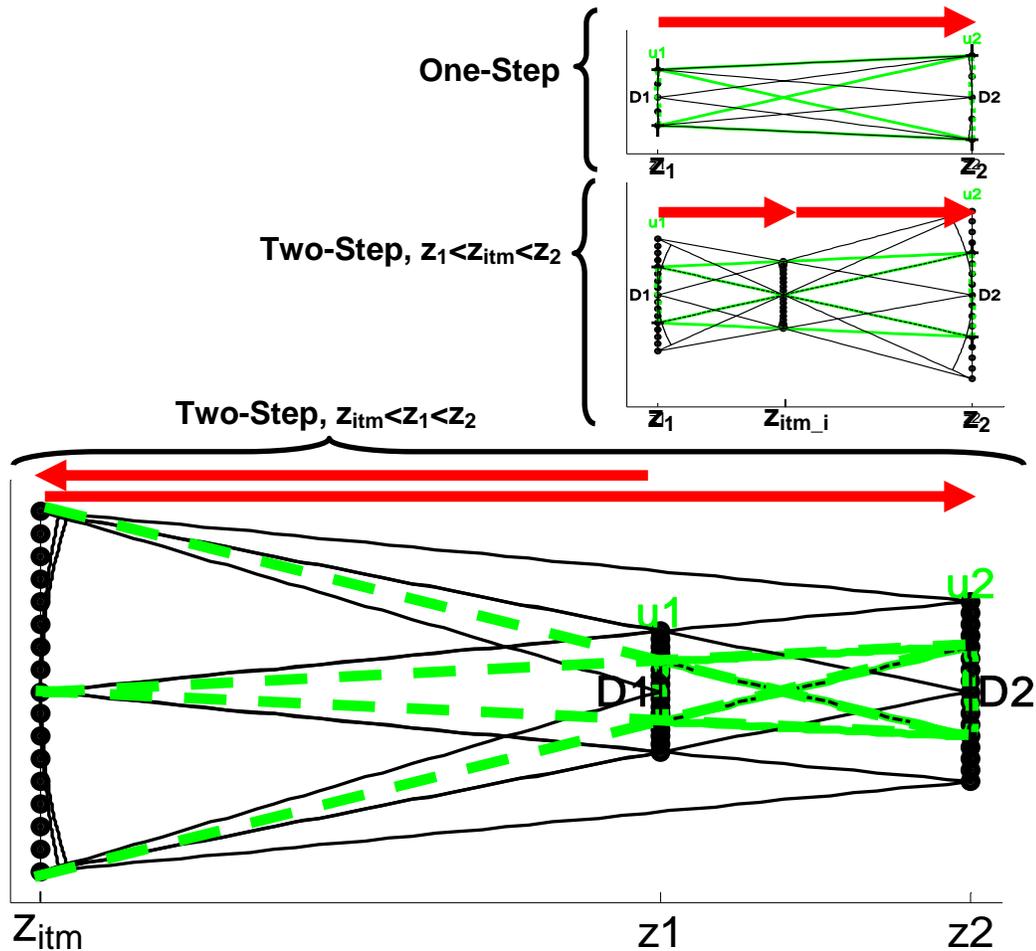
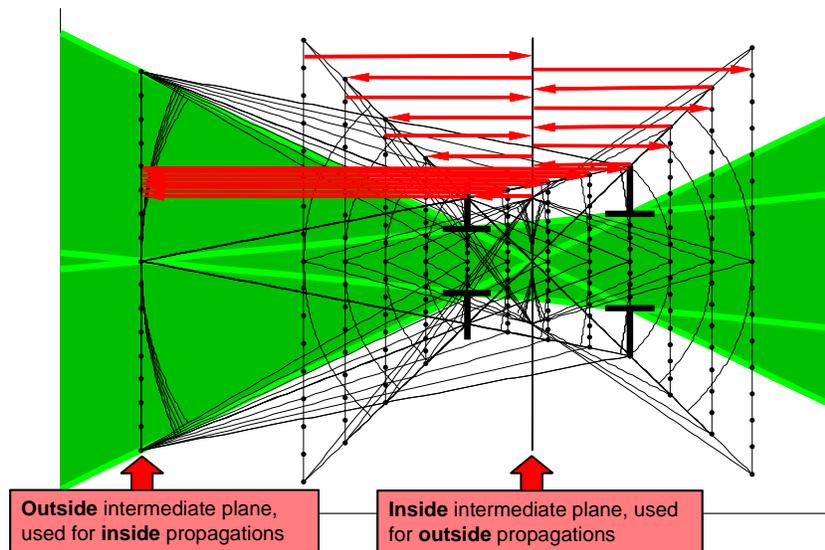


Figure 2. Propagation diagrams for one-step and two-step DFT propagations.

We now have three distinct procedures for propagating an optical wavefront from one plane to another. The first, one-step DFT propagation, is obviously the simplest, and in some cases, including the example described above, it is much more efficient. – more than 32 times more efficient in the case of the example. This raises an obvious question: why would one ever use two-step propagation? One reason we have already mentioned: it makes it possible to choose the mesh spacing at the initial and final planes independently. But there is another, more important, reason: using two-step propagation we can adjust the curvature of the reference wave to match, at least approximately, the overall curvature of the wavefront to be propagated. This reduces the maximum angle formed between the rays to be propagated and the reference normal, allowing us to use coarser meshes, reducing the number of mesh points required, and thus reducing the simulation execution time. (As compared to a two-step propagation where the curvature is less well-matched.) Thus if the light to be modeled is collimated, converging, or diverging, the reference wave should generally be chosen to be collimated, converging, or diverging, respectively. However, this can only be done when we have enough a priori information available to allow us to infer the approximate curvature of the wavefronts to be modeled. For example, consider the case where we have just two limiting apertures, as before, but now consider propagations between *any* two planes along the optical axis, not just between one aperture plane and the other. The set of rays to be modeled is the same at all planes, but the average curvature of the wavefronts they form changes from converging, to somewhat collimated, to diverging, as the light first approaches, then passes through, and then recedes from the pair of apertures.

Because of this, the optimum<sup>3</sup> reference wave for all propagations between any two planes both *inside* the interval between the apertures is a spherical wave centered on the point *outside* that interval that projects the two apertures onto one another. Similarly, the optimum reference wave for all propagations between any two planes both *outside* that interval is a spherical wave centered on the point *inside* the interval that projects the two apertures onto one another. For propagations between the two aperture planes either of the two intermediate planes may be used. For propagations from outside the interval to inside, or vice versa, if we wish to use the minimum possible mesh dimension it is necessary to first propagate to from the initial plane to an aperture plane, switch intermediate planes, and then propagate to the final plane. The mesh dimension is the same at all planes, and the mesh extent at each plane is twice the cross-sectional



**Figure 3. Propagation diagram for the case of two limiting apertures with propagations between multiple planes.**

<sup>3</sup> For the purposes of this paper the “optimum” mesh spacings, mesh dimension(s), reference wave, and/or location for the intermediate plane used in two-step DFT propagation shall be defined as those which minimize the mesh dimension  $N$  while still allowing us to obtain correct results for the given wave optics problem.

extent of the ray region to be modeled at that plane. All of this is illustrated in Figure 3.

Although the problem illustrated in Figure 3 is a special case, in that it involves only two limiting apertures, this case has a more general significance than might be apparent, because even in the most general case, with any number of limiting apertures, it is always possible to find a pair of apertures at any two planes such that all rays that pass through the full set of apertures also pass through those two. This yields a propagation geometry similar to that shown in Figure 3 that will work for all propagations between any two planes along the optical axis. By adjusting the positions of the two planes and the sizes of the two apertures to fit the ray region to be modeled reasonably closely, we can ensure that the mesh dimension will be reasonably close to the minimum consistent with obtaining correct simulation results. We conjecture that if we were to *optimize* the two plane positions and aperture diameters so as to minimize the mesh dimension, that would be identically equal to the minimum mesh dimension obtainable taking the full set of limiting apertures into account directly, but that hypothesis remains to be proven.

Incidentally, there is no comparable solution that works for propagation between any two planes using one-step DFT propagation, because in that case the mesh dimension varies inversely with the propagation distance, and in the limit where the propagation distance goes to zero the mesh dimension goes to infinity. Of course, one can get around this problem by propagating from the specified initial plane to some distant plane, and then propagating from there to the specified final plane, but that is just two-step DFT propagation.

### 3.0 RAY REGION DIAGRAMS AND MULTIPLE LIMITING APERTURES

As we have seen, propagation diagrams can swiftly become very complicated even for relatively simple cases such as vacuum propagation with two limiting apertures. They become even more complicated when one treats the mesh spacings, mesh dimensions, intermediate plane locations as free parameters rather than fixed, or when there are more limiting apertures, not necessarily aligned with the source and receiver apertures, or when one needs to take into account the bending of rays that occurs when propagating through aberrating media, or all of the above. To handle such cases we found it necessary to devise another, more compact, means of representing the same information: ray region diagrams.

In a ray region diagram, we plot the set of complex rays passing through a given set of limiting apertures as a contiguous region in slope-intercept space, where the ray intercepts are those for some particular plane along the optical axis, chosen for convenience. By convention we plot the ray-intercept along the x-axis and the ray slope along the y-axis; this is convenient when comparing ray region diagrams for different planes along the optical axis because when we plot them side-by-side corresponding rays have the same ordinate value, because a ray's slope is independent of  $z$ , whereas its intercept is not. Plotted this way, the set of rays passing through a single limiting aperture forms a straight strip bounded by a pair of parallel lines, its horizontal width equal to the aperture diameter, and its slope inversely proportional to the distance between the aperture and the plane used for the diagram. The set of rays passing through two limiting apertures in different planes forms a parallelogram, and the set of rays passing through  $n$  limiting apertures forms a polygon with between 3 and  $2n$  sides.

For any given wave optics problem the ray region to be modeled can always be described in terms of a set of two or more limiting apertures at different planes – there have to be at least two, or else it would be impossible to simulate the problem, because the meshes would have to be infinitely large. Often there are just two, e.g. the source and receiver apertures, as in the example we used in the previous section. But if the source is known to be collimated, that gives us an additional effective aperture at the plane of the beam waist, and if the receiver is known to have a finite field of view, that gives us yet another. In principle there can be any number of effective limiting apertures, related to vignetting within the optical systems at the source and/or receiver, but in a well-designed optical system there should be little or no vignetting, so for most problems two to four limiting apertures will generally suffice to define the ray region to be modeled. Even when four or more potentially limiting apertures are present, depending on their relative sizes and positions, some may be dominated by others, so that the first have no effect on the ray region to be modeled, and therefore may be disregarded. Figure 4 shows the ray region diagram for the ray region to be modeled for the case of two limiting apertures discussed in the previous section. We have drawn the diagram for the plane  $z=z_1$ ; as a result the set of rays passing through the aperture  $D_1$  at that same plane forms a vertical strip.

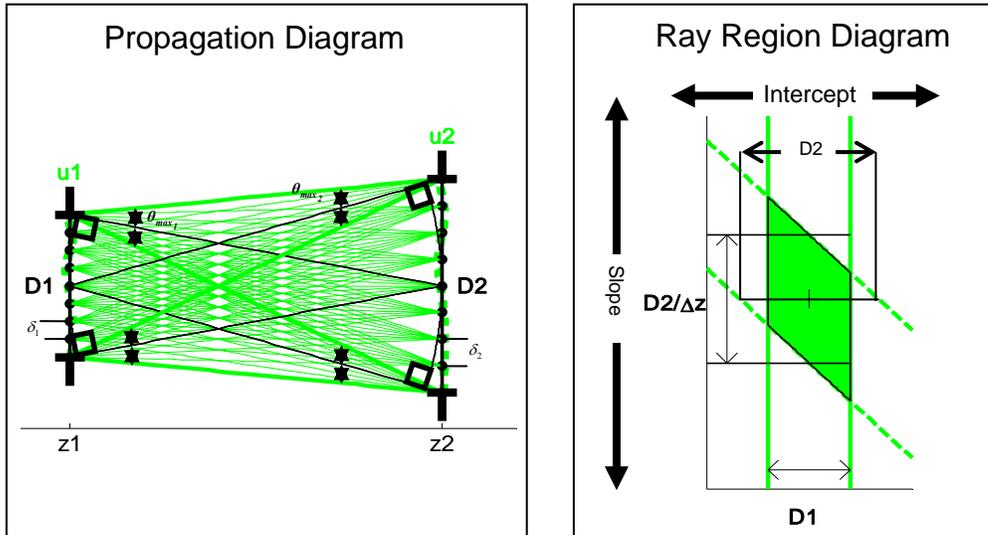


Figure 4. A comparison between the propagation diagram and the corresponding ray region diagram for the same simple wave optics problem shown in Figure 1. The diagram has been drawn for the plane  $z=z_1$ .

When there are only two limiting apertures, the ray region to be modeled is a parallelogram centered on the line connecting the two apertures, which may or may not be the nominal optical axis. When there are three or more limiting apertures, they may not lie along the same line, further complicating the geometry, as illustrated in Figure 5.

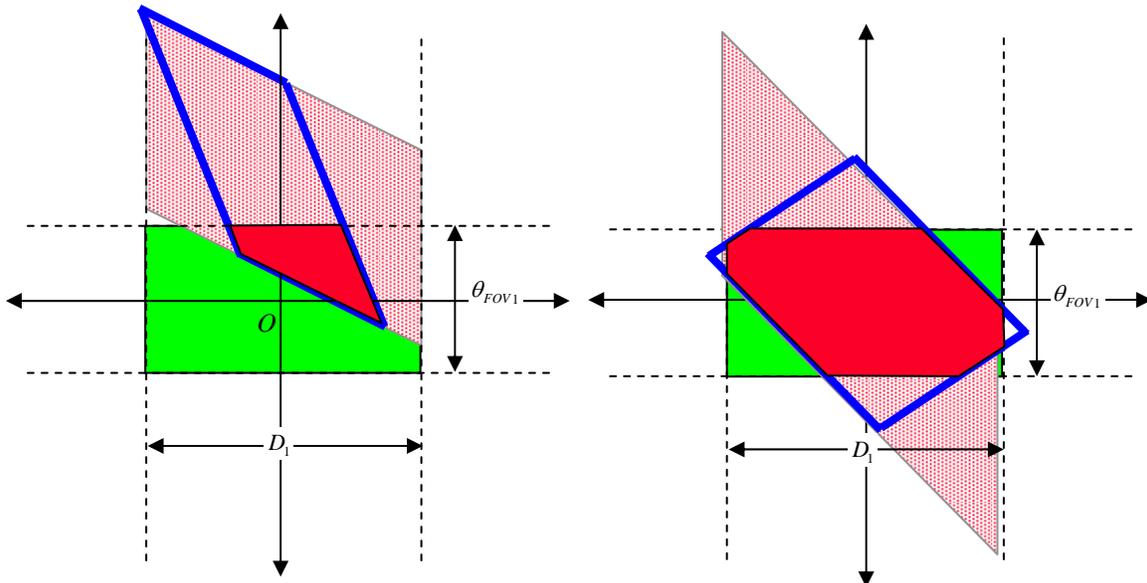
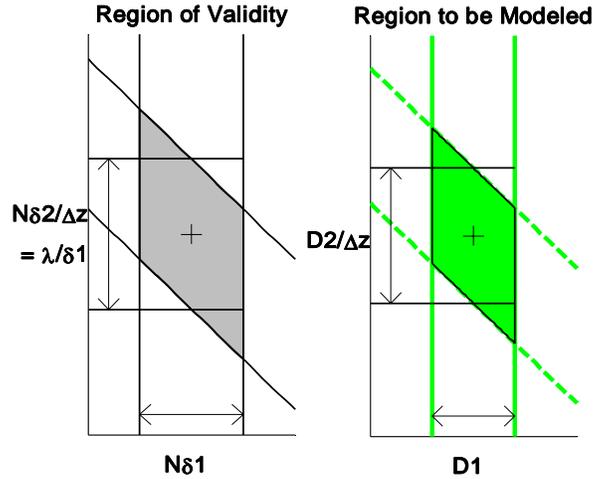


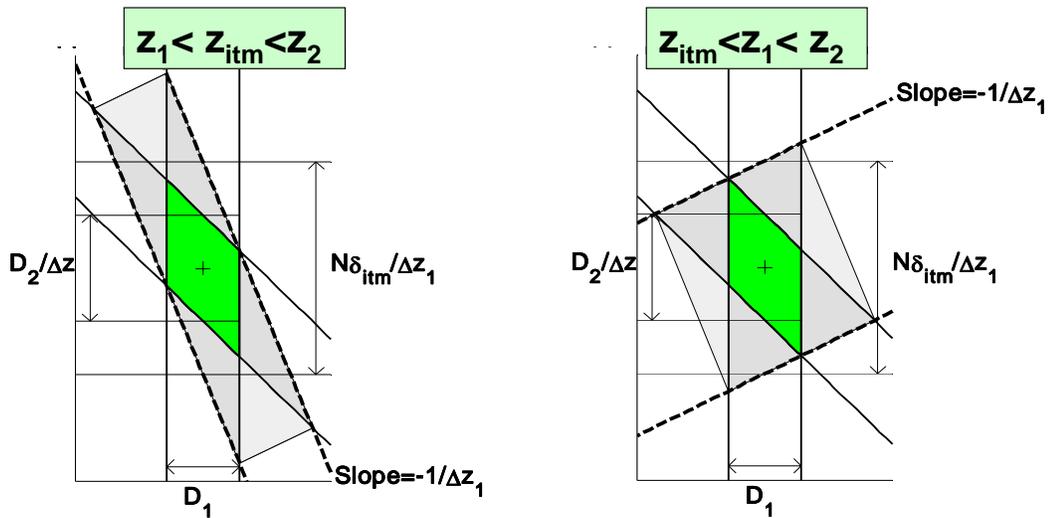
Figure 5. Two of the many possible geometries for the ray region bounding the light to be modeled in an optical propagation problem when a priori information about the source collimation and/or the receiver field of view is available. The green rectangle is the region defined by the source collimation, the red dotted parallelogram is the region defined by the source-receiver aperture geometry, and the blue parallelogram is the region defined by the receiver field of view.

The region of validity for any given DFT propagation can be defined in terms of two limiting apertures, corresponding to a parallelogram in a ray region diagram. For a one-step DFT propagation, shown Figure 6, the limiting apertures are just the mesh extents at the initial and final planes; we can adjust their sizes, but not their locations. For a two-step DFT

propagation, shown in Figure 7, we are free to choose both the sizes and the locations of the limiting apertures independent of the initial and final planes for the propagation; this allows us to better fit the region of validity to the ray region to be modeled, and we can then use those same limiting apertures for all propagations, between any and all planes of interest. However this added flexibility carries a price: the mesh dimension for a two-step DFT propagations is four times that for a one-step DFT with the same limiting apertures; at the planes of the limiting apertures the mesh spacings are *halved* and the mesh extents are *doubled*. *Outside* the two aperture planes the mesh spacings vary linearly with distance from the *inside* intermediate plane (described in the previous section), and vice versa.



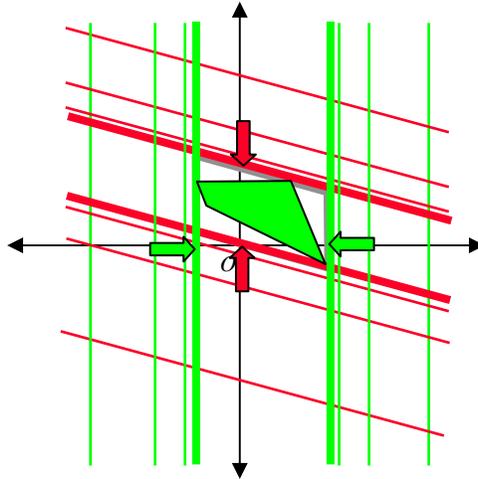
**Figure 6.** The region of validity for a one-step DFT propagation, compared with the ray region to be modeled for the case of two limiting apertures. As shown, the initial and final planes for the propagation are limiting apertures are assumed to coincide with the planes of the two limiting apertures for the region to be modeled. The diagram has been drawn for the plane  $z=z_1$ . Note that the area of the R.O.V. is proportional to  $N$ .



**Figure 7.** The region of validity for any two-step DFT propagation, shown for propagations using both the outer and inner intermediate planes, on the left and right respectively. Note that the region of validity is the same in both cases and matches the ray region defined by two limiting apertures,  $D_1$  and  $D_2$ , shown in green. The diagram has been drawn for the plane of the first limiting aperture,  $z=z_1$ . In the diagrams  $\Delta z_1 = z_{itm} - z_1$ .

#### 4.0 CHOOSING MESH SPACINGS AND MESH DIMENSIONS

For one-step DFT propagation, we can adjust the sizes of the two limiting apertures defining the region of validity, which controls the horizontal separations of the two line pairs in the ray region diagram, but we cannot adjust the locations of the limiting apertures along the z-axis, which determines the slopes of the two line pairs. We can however adjust the *transverse* positions of the two limiting apertures, which controls the transverse offsets of the line pairs with respect to the origin. Given a ray region to be modeled, we can always ensure that the region of validity will contain the region to be modeled by adjusting the sizes and transverse positions of the two line pairs, as illustrated in Figure 8. If we wish to minimize the mesh dimension,  $N$ , we should bring each of the four lines in until it just touches the ray region to be modeled, as illustrated in Figure 8. The mesh dimension is directly proportional the area of the region of validity; this can be verified by inspection of Figure 6.



**Figure 3-19. In the case of one-step propagation, fitting the bounding region for the light to be modeled involves adjusting the positions of two pairs of fixed-slope parallel lines.**

Applying the requirement that the region of validity must be a superset of the region to be modeled, and taking into account the DFT constraint relating the mesh spacings and the mesh dimension, we can use the ray region diagram to derive the same set of inequalities obtained in the previous section from the propagation diagram:

$$\delta_1 \leq \frac{\lambda}{2\theta_{\max 1}} \cong \frac{\lambda\Delta z}{D_2}, \quad \delta_2 \leq \frac{\lambda}{2\theta_{\max 2}} \cong \frac{\lambda\Delta z}{D_1} \quad (8)$$

$$N\delta_1 \geq D_1, \quad N\delta_2 \geq D_2 \quad (9)$$

$$N = \frac{Area_{ROV}}{\lambda} = \frac{\lambda\Delta z}{\delta_1\delta_2} \geq \frac{D_1D_2}{\lambda\Delta z} \quad (10)$$

For two-step DFT propagation, we can adjust not only the sizes of the two limiting apertures defining the region of validity, but also their locations of the limiting apertures along the z-axis; this means that we can adjust not only the horizontal separation of the two line pairs, but also their slopes. As in the case of one-step propagation, the mesh dimension is directly proportional to the area of the region of validity, but for two-step propagation the constant of proportionality is four times higher,  $4/\lambda$  instead of  $1/\lambda$ . If we wish to minimize the mesh dimension in this case it is

somewhat more complicated, because we have two more degrees of freedom, corresponding to the locations of the aperture planes along the optical axis. For any given choice of aperture locations it reduces to essentially the same problem as for the one-step case, but to determine which aperture locations will minimize the mesh dimensions we would need to use some type of optimization algorithm on that two-dimensional search space. That is not particularly difficult, but on the other hand it is generally not necessary to strictly minimize the mesh dimension, as long as we can keep it within acceptable bounds. Often it is more practical to choose the two limiting apertures for convenience, without worrying about whether they are exactly optimal or not; we can place them at any two planes along the optical axis, and we can adjust their sizes, so long as we make sure each contains the entire cross-sectional area of the ray region to be modeled as it cross that plane. (However, generally one would not wish to place the two planes too close together, or make the two apertures too much larger than necessary, as that could increase the mesh dimension required dramatically.) For a given pair of limiting apertures, it can be shown that the inequalities governing the choice of mesh spacings and dimensions for two-step propagation analogous to those previously obtained for one-step propagation:

$$\delta_1 D_2 + \delta_2 D_1 \leq \lambda \Delta z, \quad N = 4 \cdot \frac{Area_{ROV}}{\lambda} \geq \frac{D_1}{\delta_1} + \frac{D_2}{\delta_2} \quad (11)$$

Here  $D_1$  and  $D_2$  are the diameters of the two limiting apertures,  $\delta_1$  and  $\delta_2$  are the mesh spacings at the two aperture planes, and  $\Delta z$  is the distance between the two apertures. The mesh spacings at other planes vary as described in the previous section: *Outside* the two aperture planes the mesh spacings vary linearly with distance from the *inside* intermediate plane, and vice versa.

Once we have chosen a pair of limiting apertures, in order to minimize the mesh dimension  $N$  the mesh spacings and dimension should be chosen as follows:

$$\delta_1 = \frac{\lambda \Delta z}{2D_2}, \quad \delta_2 = \frac{\lambda \Delta z}{2D_1}, \quad N \geq \frac{4D_1 D_2}{\lambda \Delta z} \quad (12)$$

Another important special case is that of plane wave propagation, where we impose the additional requirement that the mesh spacings at the planes of the two limiting apertures and all planes in between should be kept equal. This results in the following set of inequalities:

$$\delta_1 = \delta_2 = \delta, \quad \delta \leq \frac{\lambda \Delta z}{D_1 + D_2}, \quad N \geq \frac{D_1 + D_2}{\delta} \quad (13)$$

## 5.0 RELATED ISSUES

In addition to choosing the mesh spacings and dimensions, there are a number of related issues that must be addressed if one is to obtain correct results using wave optics simulation. These include:

1. The task of constructing the correct representation of the wavefront leaving the source, retaining all the light falling within the ray region to be modeled, and discarding all light falling outside the region of validity. This can be done, for example, by directly evaluating of the Huygens-Fresnel integral at the receiver plane mesh points falling within the ray region to be modeled, and then propagating the resulting mesh back to the source plane using a two-step DFT propagation.
2. Spurious *Fresnel ringing* artifacts associated with the propagation mesh rather than any physical aperture that can arise in two-step propagation even when the mesh spacings and dimensions have been chosen correctly. These artifacts can be removed by appropriately filtering the wavefront mesh at the source, but our present filtering techniques work only for a single plane; at other planes there will still be spurious ringing artifacts.
3. To model propagation through an aberrating medium, one breaks up the propagation from the source to receiver into multiple steps, alternating vacuum propagations with *phase screens*, used to represent the integrated optical path difference over the corresponding interval. At each phase screen some of the incident light is bent toward the receiver, and some is bent away from it, thus changing the set of rays we need to

model. We can take this into account by increasing the diameter used for limiting apertures seen through aberrating media to reflect their increase in apparent (i.e. optical) size due to blurring by the media.

4. To model propagation through an optical system, we need to take into account gross changes in the direction the light propagating and the curvature of the wavefront caused by mirrors and lenses, as designed, and the somewhat subtler changes caused by aberrations, apodizations, vignetting, unwanted reflections, and so forth. Fortunately, all the gross changes in the direction and curvature of the wavefront can generally be taken into account by (a) defining the optical axis for wave optics simulation (i.e. the z-axis) to correspond to the nominal optical axis defined by the system designers, (b) removing the nominal effects of all mirrors and lenses by picking some one physical plane to work in, e.g. the source plane, and replacing all limiting apertures with their images as seen at that plane, through any intervening optics, (c) adding in any small deviations from the nominal effects, as small tilts and or curvatures, which can be represented in the reference wavefront as opposed to the complex mesh. Finally, the aberrations, apodizations, and so forth must still be modeled explicitly, by operating on the complex mesh at the appropriate planes.
5. When modeling problems that involve substantial changes in the ray region to be modeled as a function of position along the optical axis, for example propagation through strong turbulence, it is sometimes useful to break the propagation path into two or more pieces, and choose different pairs of limiting apertures for each.

## 6.0 SUMMARY AND CONCLUSIONS

We have presented a general method for determining exactly what the constraints are governing the choice of mesh parameters for any given wave optics modeling problem, taking into account any available a priori information concerning the nature of the source, the receiver, and the intervening optical path. In brief, the method involves using the whatever a priori information is available to identify two limiting apertures at two different planes such that all of the light we need to model (thought of as a set of complex rays) passes through both apertures. Generally it is desirable to choose the two apertures so that they together fit the set of rays to be modeled at least reasonably closely. These two limiting apertures can correspond to real physical apertures, or they can be imaginary apertures used to describe where the light of interest passes through a given plane. Once we have chosen those two limiting apertures, we can then choose the mesh spacings and mesh dimension according to the inequalities presented in Section 4:

$$\delta_1 D_2 + \delta_2 D_1 \leq \lambda \Delta z, \quad N \geq \frac{D_1}{\delta_1} + \frac{D_2}{\delta_2} \quad (14)$$

Here  $D_1$  and  $D_2$  are the diameters of the two limiting apertures,  $\delta_1$  and  $\delta_2$  are the mesh spacings at the two aperture planes, and  $\Delta z$  is the distance between the two apertures. Once we have chosen  $\delta_1$ ,  $\delta_2$ , and  $N$ , that gives us a single propagation geometry that will work for *all* propagations in the given problem, between *any* two planes, not just the two aperture planes. In between the two aperture planes the mesh spacing varies linearly from  $\delta_1$  to  $\delta_2$ . Outside that interval, the mesh spacing varies linearly with distance from the point in between the two apertures that projects each aperture onto the other, as illustrated in Figure 3.

When the available a priori information allows us to identify more than two limiting apertures, that may reduce the size of the ray region to be modeled, and that may allow us to reduce the mesh dimension and still obtain correct results. In a ray region diagram, the ray region to be modeled for any given set of limiting apertures appears as a polygon, while the region of validity for a two-step propagation geometry appears as a parallelogram. By adjusting the sizes and locations of the limiting aperture, we can control the shape of the parallelogram; if we wish to minimize  $N$ , we must seek to minimize the area of the parallelogram subject to the constraint that it must contain the entire ray region to be modeled – this is a relatively simple optimization problem. On the other hand, it is generally not necessary to strictly minimize  $N$ , especially in view of the fact that we are likely to have round it upward anyway, to make it a power of two, or at least product of powers of a few low primes, depending on the particular FFT algorithm we are using. The real importance of ray region diagrams is that they provide a systematic way take into account all available a priori information, whether it pertains to the source, the receiver, or the intervening optical path.