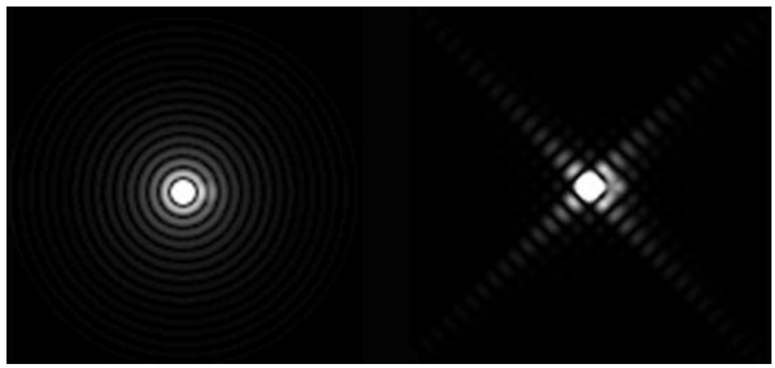
APODIZATION FOR OBSTRUCTED APERTURES

Harold R. Suiter

Apodization is a complicated word for the simple process of shading part of the aperture, usually the outside. "Apod" mean "without feet" and, judging from the name, the purpose of apodization is to reduce the intensity of the diffraction rings. However, it has come to represent any form of modification to the diffraction structure to emphasize certain desirable features of imaging at the expense of others.

One of the simplest forms of apodization is to impose a square shape on the aperture (or a diamond if we tilt our heads) to allow resolution of unequal doubles stars where the separation is just beyond the first diffraction ring's radius. The diffraction structure compresses into two bars emerging from the image and the little star at 45 degrees doesn't have to contend with the rings at all, However, the diagonal of a square is larger than its side, so we have to give up resolution. We have, in effect, stopped down the aperture to change the shape of the diffraction pattern.



A "diamond" one-dimensional apodizer. A weak companion next to a bright primary is difficult to see, particularly in the presence of turbulence. Putting a square mask over the entire aperture and rotating it until the companion is in the direction of one of the points forces the bright parts of the diffraction structure into two non-interfering directions. The dim star is likelier to be seen.

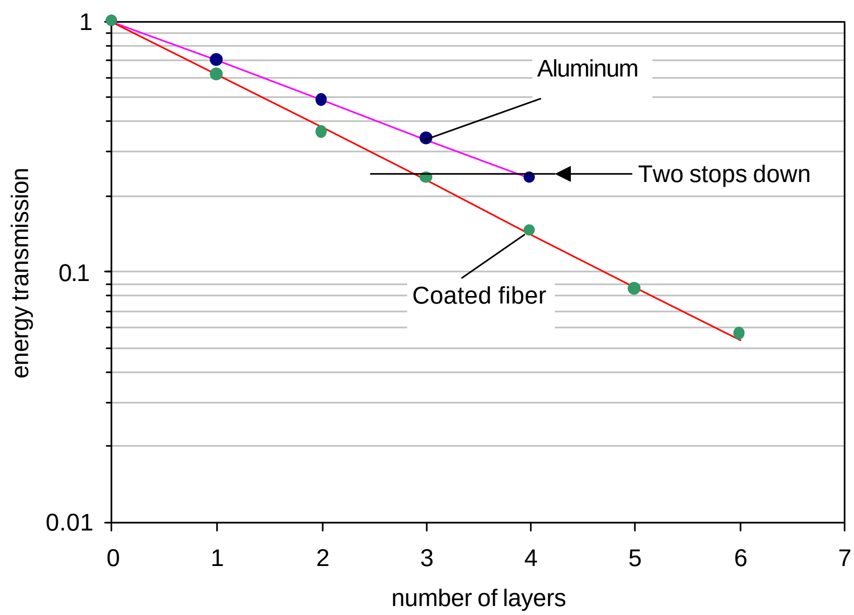
Apodization is "rediscovered" every ten years or so, and seems to be currently enjoying a renaissance after years of neglect. The window-screen apodizer was first mentioned in an article by Arthur S. Leonard in a letter sent to the telescope-making column of the June 1954 *Scientific American.* Prior to Mr. Leonard's suggestion, apodization schemes depended on depositing opaque films on glass or coating only partially transmissive mirrors. Naturally, few if any such elaborate and expensive plans were actually carried through to completion. Mr. Leonard's suggestion was easy and, more important, reversible. The gradual decrease of transmission is achieved be cutting holes of ever increasing radius in each layer of window screen. Each layer is carefully centered and laid on the previous layer at such an angle that the opportunity for nasty moiré effects are minimized.

The dark-field image is comprised of rainbow spikes around a planet or other small bright object, centered around a small, dark region of ordinary-looking images. As long as the periods and spacing of the screen is maintained, there is little blurry light diffracted into that region of exclusion in the center.

Using 1/18th-inch window screen, the first-order rainbow appears at 58 arcseconds for blue wavelengths and even farther for other colors. Jupiter fits easily inside 58 arcseconds as viewed from anywhere on the orbit of the Earth. Finer screens divert spectra farther. Coarser screens divert the spectra into lower angles that may eventually intrude into the region of interest. That is a general theme with screen-type apodizing filters. They can only be used to image small, bright, objects on dark backgrounds, but small objects are where we are most concerned with the diffraction pattern and the night sky is naturally dark. Screen-type apodizing filters are of no use on the moon, sun, and extended objects.

Building such an apodizer is fairly straightforward (see article on "making"). What is more difficult is designing one. It depends on many factors, not the least of which is the size of the central obstruction. As we shall see below, effective apodization can only be achieved with obstructions about 20% and below. However, we may achieve some lessening of the first ring up to obstruction ratios of about 30%.

Another important factor in screen design is the layer-by-layer transmission of the screen. If you fabricate an apodizer using only word-of-mouth recipes without measuring the transmission of your particular screen, you will achieve only a crude approximation of the effects you wish to achieve. I measured the transmission of several layers of aluminum and coated-fiber screens with a professional radiometer and saw the behavior in the figure. I followed the coated-fiber screen down through an absurd number of layers to see if the relation ever broke down, but was pleased to find that it did not. As can be seen in the graph, plastic-coated fiber or plastic screens tend to intercept more incoming light than aluminum screens. But please remember, these are only two examples of the many screens available. *You must measure your screening material yourself.*



The energy transmitted by various thicknesses of ordinary window-screen materials. The transmission of two layers is just the transmission of one layer multiplied by itself, i.e., T^2. Three layers is T^3 and so forth.

My two screens are best fit by a one-screen transmission coefficient of T=0.70 for the aluminum and T=0.61 for the coated fiber screen. Of course, professional radiometers are not common, but perhaps the reader or a friend possesses a high-quality radiometer in the form of a manual SLR camera metering system. Just put the camera on a tripod, point it at a well-lit uniform wall, and put more and more layers of screen in front of the lens until it meters an integer number of f-stops away.

For example, I was gratified to find that my Olympus OM-1 metered four layers of the aluminum and three layers of the coated fiber screens at two stops from the unimpeded case. The radiometer values of about 0.24 in both cases is very close to the two-stops value of 1/4 or 0.25. If I didn't have the radiometer, I would have calculated these transmissions at $0.25^{1/4} = 0.71$ and $0.25^{1/3} = 0.63$ respectively.

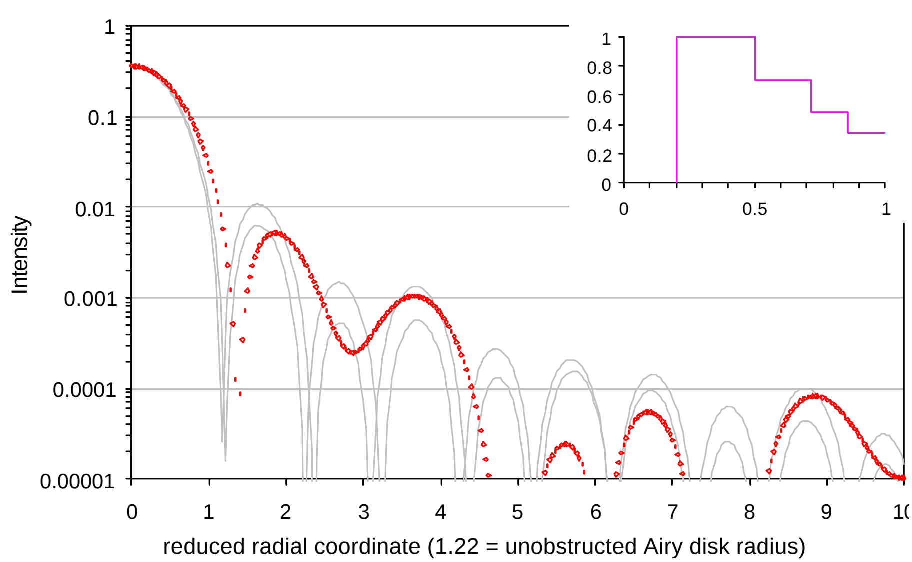
THE DESIGN

Okay, we have all the input constants, the radius of obstruction and the transmission coefficients of all of the successive layers of screen. We need to choose the radii where the first, second, and third layers start. Moving from the inside out, call these radii r_1, r_2, and r_3 respectively. I have chosen, somewhat arbitrarily, to limit the apodizer to three radii.

I have chosen as my quality criterion the *encircled energy ratio* at the angular radius of the perfect Airy disk. Let me explain just what the encircled energy ratio is. A circle at the first minimum of the Airy pattern encloses 83.77 percent of the total energy in the pattern. That is, the Airy disk contains this fraction of the total energy collected by the aperture during a fixed exposure. The rest of the energy is out in the diffraction rings. Similarly, if we were to account for all of the energy collected by our apodized aperture and ask what fraction has been collected together at the same radius, we will get some other value. If it is more than 83.77% of the total energy in the apodized pattern, then the encircled energy ratio at the radius of the Airy disk is more than 1 and the apodized aperture "wins." If that ratio is less than 1, the apodizer cannot pack a greater fraction of the energy inside the Airy disk than the unobstructed case, but it can perhaps beat the obstructed case. Note that this is not the only quality criterion possible but it is the one chosen here. I might well have decided to minimize the relative brightness of the first ring, or used some other obscure criterion.

Naturally, having all these windows screens in front of the aperture darkens the image. The raw value of energy collected cannot be as great as the Airy disk of an unobstructed aperture, but its *fraction* can. In using this criterion, I separate the *shape* of the image from mere *darkening*. It is this shape and not the raw energy transmitted that most concerns the person doing the apodization. We are assuming that the image target has more than enough brightness to overcome the darkening. Planetary observers frequently diminish the surface brightness of the planet by using a neutral density filter. The dimming caused by the apodizer often is not a disadvantage but a benefit.

Our basic design tool will be an Excel spreadsheet (on website) Instructions are with the spreadsheet). The best fit for a 20-percent obstruction using three layers of aluminum window screen is displayed in the graph. The general procedure is to start the process with the apodizer radii at almost the whole aperture. (The apodizer starts with just a fuzzy edge.) First the inner radius r_1 is pulled inward until the quality criterion is maximized. Then the same is done with r_2, and r_3 in turn. Each time the improvement is less, but it does creep upward. Then, the whole sequence is repeated again and again until



Best fit for the aluminum screen, obstruction = 20% (quality criterion = 1.0123) Upper gray curve: unapodized obstructed. Lower gray curve: perfect aperture. Red diamonds: apodized obstructed aperture. Inset contains profile of transmission.

no more improvement is possible. After three or four rounds of adjusting the radii, we find the radii where the quality criterion cannot be increased. The figure inset depicts the transmission coefficient profile.

The best radii for three-radius screens is given in the table below as they vary with obstruction ratio. The one-layer transmission coefficient is the one I measured for aluminum of 0.70. We see the "after" quality criterion dip below 1.0 at an obstruction ratio slightly beyond 20%. Nevertheless, the criterion can still be improved beyond what it was for the unfiltered case even at 30%, but packing a bigger fraction of the energy inside the classical Airy radius is becoming very difficult at obstructions above 20 to 25%.

A quick look at the generalized Strehl ratios and quality criteria for both the before and after cases are a sharp lesson that no one-number criterion tells the whole story of optical quality.

		Obstruction	0	0.15	0.175	0.2	0.25	0.3
before								
	Strehl ratio=	EER(0)	1	0.978	0.969	0.96	0.938	0.91
	quality criterion=	EER(1.22)	1	0.949	0.932	0.912	0.867	0.817
after								
	Strehl ratio=	EER(0)	0.856	0.842	0.837	0.832	0.822	0.806
	quality criterion=	EER(1.22)	1.141	1.066	1.041	1.012	0.948	0.878
		Energy transmitted	0.432	0.422	0.427	0.424	0.427	0.434
layers								
1: 70.0%		r_1	0.475	0.485	0.495	0.505	0.525	0.555
2: 49.0%		r_2	0.680	0.695	0.705	0.715	0.735	0.770
3: 34.3%		r_3	0.825	0.840	0.850	0.860	0.885	0.910

Radii as a fraction of the whole radius are given for three-layer apodizing screens made of the aluminum window screen measured earlier. Obstruction ratios are along the top. The generalized Strehl ratio and the encircled-energy quality criterion are presented both before and after filtering. "

DOES IT WORK?

Although I have expended much effort the past 15 years in coming up with ways of designing, building, and using apodizing screens, I have yet to be convinced that any improvements I see are due to the lowering of the rings and not some other effect. This is why I have written this article. If enough people start experimenting with these filters (and do the design properly), perhaps we can come up with a consensus. I have observed that apodizers sometimes improve the image of bright planets, but what is acting to improve the image? Clearly, the apodizers reduce the diffraction rings in the case of small obstruction, and clearly the images are better, but ring reduction may not be the mechanism of improvement. Three alternatives, any one of which could be responsible for the improvement as well as decreasing the diffraction rings, occur.

1) The image is darker.

Darker images help dazzling images. Darker images detected with the same eye are better to the extent that the "trash" diffraction rings and other scattering surrounding the stellar image is lessened in intensity until the eye no longer pays any attention to it. It is below threshold. Look at the apparent diameters of various brightness stars the next time you observe. Even though you know that the shape of all stellar images is the same as determined by diffraction, dimmer stars appear smaller. More of the trash surrounding the image drops below threshold. Whether this "trash" explanation helps extended images is less certain, but filtering certainly helps to reduce the dazzling intensity of bright planets.

HOW THESE CALCULATIONS WERE DONE

I adapted an expression found in Daniel Schroeder's *Astronomical Optics* that is used there to calculate the diffraction pattern of an obstructed aperture. The adapted expression calculates the more general case of several partially transmitting rings. Each annular region contributes the following term to the field.

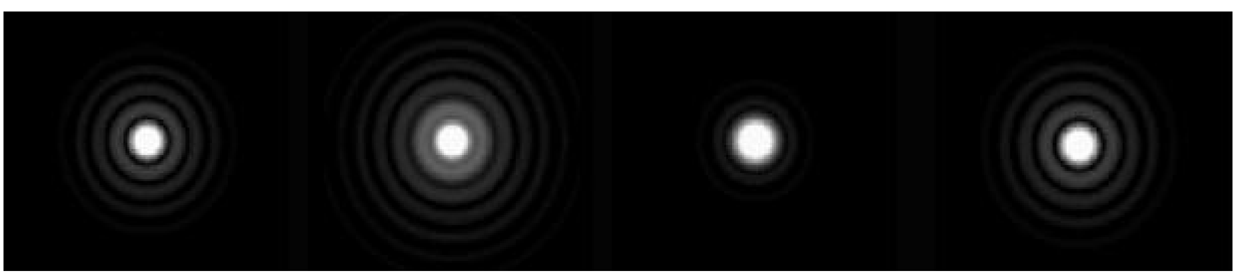
$$2T_{region}\left(r_{outer}\frac{J_1(r_{outer}\pi\rho)}{\pi\rho} - r_{inner}\frac{J_1(r_{inner}\pi\rho)}{\pi\rho} \right)$$

Here, T is the transmission coefficient, the r's are the normalized annular radii in the text, ρ is the normalized radius (1.22 = edge of Airy disk), and the J's are the usual Bessel functions. The spreadsheet simply adds up all the regions and squares the sum. If r_{outer} and T is 1 and r_{inner} is zero, we get the familiar Airy expression. The results of this expression have been independently confirmed by a routine that works another way.

As a test device, it is good to make an apodizer that does not apodize. This so-called "null screen" is a uniform filter (perhaps comprised of two screening layers or one of a less transmissive screen) completely across the front of the telescope, or with just big enough a hole to let the secondary-mirror post jut through. You can use it to test the performance of the true apodizer by putting it on when others observe through the instrument. The reason for doing it this way, rather than merely using an eyepiece neutral density filter, is that it too has side-order spectra and the guest observers don't know they're being fooled. Otherwise, they'll politely tell you what they think you want to hear.

2) Aberrations are obscured.

The next figure shows the effect of a severe rolled-edge aberration (a full wavelength starting at 9/10 the radius) on an unobstructed aperture. Even at the best focus for the unfiltered aperture, the aberration has blurred the Airy disk into the first ring. However, when the apodizer is added and the aperture refocused yet again, the diffraction pattern makes a rather spectacular recovery. This is a particularly good example of the conflict between aperture and aberration. Given a certain amount of aberration on the aperture, there is perhaps a subset of the aperture that can be used to achieve superior resolution. In other words, we go through



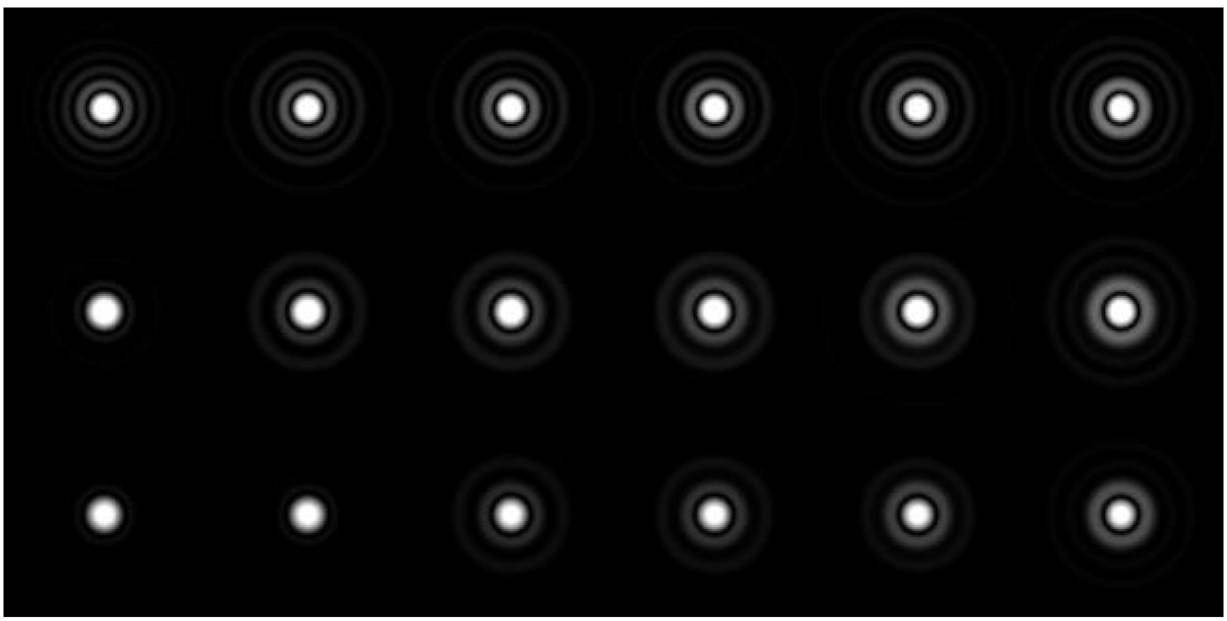
From left to right: perfect unobstructed image, image with severe rolled edge, apodized severe rolled edge, image as it would be if we merely stopped down below the edge difficulty.

this elaborate design procedure based on the assumption that our aperture is perfect, only to achieve our ends because we have inadvertently stopped down the instrument. The reputation of apodizers as "seeing filters" may somehow be related to this stop-down feature of the apodizer.

3) The energy is collected better

Lessening the diffraction rings is the nominal purpose of apodizers, However, if they work to improve the image, they may work not because the diffraction rings are lower, but because *they gather the distant energy in the image more efficiently*. Diminishing the intensity of diffraction rings may be only an unimportant side-effect of their action. Indeed, I set the quality criterion to follow energy collection. Only as a side-issue do I notice that it does very well at minimizing diffraction rings.

We can examine the next figure to obtain some insight into the performance of apodizers. The first thing that can be noted is that except for small obstruction ratios (0 to 20 percent), the action of apodizers is very subtle. If you have witnessed spectacular increases in performance by putting an apodizer on the typical 20 to 30-percent obstructed Newtonian, you have probably invoked one of the causes listed above rather than used your apodizer in its declared function of reducing diffraction rings.



0 15 17.5 20 25 30

The performance of apodizers is shown in these three rows. At the top, simple unfiltered apertures with obstruction ratios of 0, 15, 17.5, 20, 25, and 30 percent. The middle row shows the apodizers of the earlier table on the same apertures. The bottom row depicts the darkening of the these apodizers more faithfully.

The second thing to notice is that the effect of darkening appears to be more powerful than true apodization. The middle row has the theoretical apodized patterns unceremoniously jerked upward to appear equal in central brightness to the unfiltered case in the top row. The bottom row shows the appearance of the apodized apertures more nearly at the roughly 50% transmissions that they actually possess. It is clear that to correctly assess the performance of the apodizers as opposed to just intensity filters, you should make a "null" screen as a part of your kit.

....

So, I haven't answered the question yet. Does apodization work? I must confess that I really don't know. It is clear that something is happening. Nevertheless, I am skeptical about something that drifts in and out of fashion. I would think that if it offered any true benefit, such an inexpensive screen would have long ago become a standard accessory. I would be disappointed if the only lesson apodization teaches us is that darker images are better if the images are dazzling.

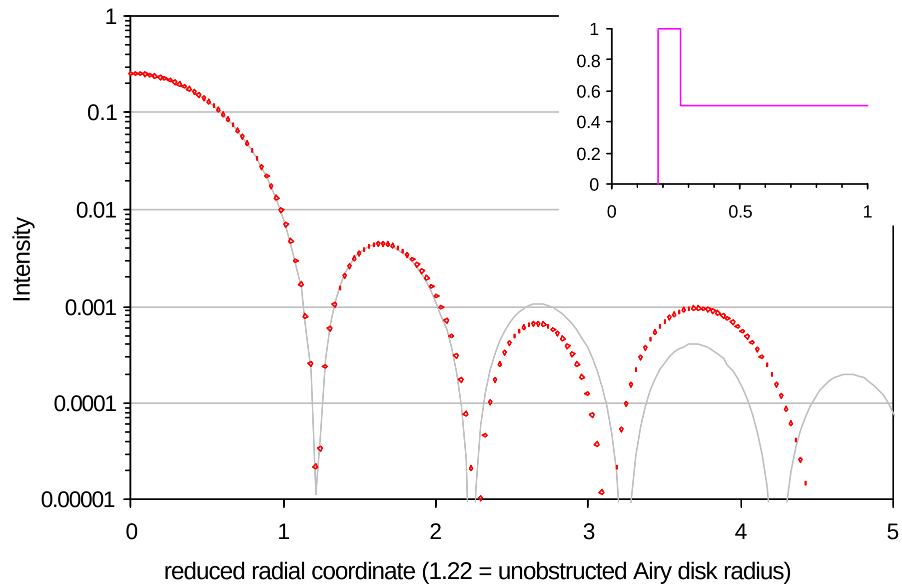
I must admit, though, it has only recently become possible for ordinary observers to undertake the diffraction calculations required to properly make a good screen. Also, it has only been in the last decade or so that obstruction ratios have drifted down as the scale of instruments has drifted up. During the past thirty years, with the effective dominance of the amateur market by the Schmidt-Cassegrainian telescope, the 33-percent obstruction would have offered little benefit to even those innovators inclined to try an apodizer.

Maybe the time of the apodizer has come.

APPENDIX

THE HIGH RING SCREEN

We can use a trick of diffraction to almost precisely reproduce the behavior of the Airy disk and the first diffraction ring as long as our obstruction is not too large. In many current large Dobsonians the obstruction ratio is very small, so let's use 18 percent as the blockage. Going outside of this obstruction by about a factor of 1.5 (to 27 percent) and applying a 50 percent screen yields the following diffraction pattern:



The obstructed aperture with the High Ring Screen compared with a perfect Airy pattern (gray line)

Note that the extra height of the ring outside the obstruction would slightly overfill the hole if the excess were scraped there by a shovel and tamped! Except for the slightly different radii, we are simulating a complete unobstructed aperture by smoke and mirrors. We can't get away with this ruse forever, though. The pattern has a slightly lower second ring and a significantly higher third ring. Still, this may be a tempting alternative to an apodizer for those wanting to resolve very close double stars and who are unwilling to give up the resolution lost in using true apodization.

FAQS ABOUT APODIZERS

Q: Are there any apodizers that lead to better resolution?
A: Yes, but they do not simultaneously decrease diffraction rings. The crudest form of a resolution-increasing mask is an obstruction, which leads to a smaller diffraction disk but larger rings. There are gentler ways of doing this, but in general apodizing masks transmit less toward the outside of the pupil and resolution masks transmit less in the middle. Of course, defining resolution as the radius of the diffraction disk leaves something to be desired.

Q: Are there any bad effects from the small holes?
A: As long as the holes are evenly spaced, and there is a sufficient number of them, then there is no reason that there should be any light scattered within the rainbows. However, screen periods are measured in dozens, not hundreds, and screens have defects of non-periodicity (they are stretched, the ends are bent over, etc.), I wouldn't be surprised if the effect of putting on a screen, even though most of the energy is diverted beyond the rainbows, has an effect on the center of the image as bad as another spider. However, spiders have effects in different spatial-frequency regions than the effects that apodizers produce. Changing the diffraction structure may be a trade-off that the observer is willing to accept.

Q: You seem to make a distinction between what you call the "transmission coefficient" and energy transmission, yet you measured the energy transmission of the screen material to derive the transmission coefficient.
A: The transmission coefficient is a weight on the wave *field*, not the energy. Since energy is proportional to the field squared, I should take the square root of the energy. However, a screen is a special case; its transmission is either off or on. When I am imaging through a radiometer in the near-field I am collecting the zeroth-order light and the diffracted orders too. The net effect is that, for binary on-off screens, *and for no other case*, the energy transmitted close to the screen is the same as the transmission coefficient.

Q: So if I want to design an evaporated-aluminum apodizer, I want to take the square root of the energy transmitted?
A: Yes. And you have to worry about the phase also.

Q: The phase?
A: Yes, the phase. Putting down different thicknesses of media with a different speed of light implies there will be relative phase shifts when the wave exits. Perhaps they won't be very great but almost none can be tolerated. In fact, making an evaporated apodizer is something that seems like a good idea at the time, but is a can of worms.

Q: Is there any way of handling a phase change in the spreadsheet?
A: Yes, but only a half-wave phase shift. You put in the transmission coefficient as a negative number.