Characterization of acceptance angles of small circular apertures

Ying Min Wang,1* Guoan Zheng,2 and Changhuei Yang1,2

1Department of Bioengineering (MC138-78), California Institute of Technology, Pasadena, CA 91125, U.S.A.
2Department of Electrical Engineering (MC136-93), California Institute of Technology, Pasadena, CA 91125, U.S.A.
*yingmin@caltech.edu

Abstract: We characterize the acceptance angles of small circular apertures for light collection by simulations and experimental measurements. By examining the full width half maximum acceptance angle as a function of the aperture size, we show that the acceptance angle of a circular aperture reaches a minimum of 67 deg before rebounding around the transition between single mode and multimode transmission (approximately 400 nm). This behavior can be explained by the change of mode-coupling efficiency during the transition from single mode to multimode propagation regime. This work in understanding of the behavior of light transmission through subwavelength apertures will guide the design of better aperture based imaging devices where apertures are used as light collection units.

©2009 Optical Society of America

OCIS codes: (110.1220) Apertures; (130.0130) Integrated optics

References and links

1. Introduction

Light transmission through subwavelength apertures have yet to be fully characterized and understood. There have been many reports of fascinating diffraction and transmission characteristics of subwavelength circular apertures [1,2], subwavelength apertures with designed shapes [3–5], subwavelength apertures arrays [6–10] and subwavelength apertures surrounded by specially designed surface structures [11]. Furthering the understanding of transmission phenomena in the subwavelength regime will benefit applications such as biomolecular detection and analysis [12], and aperture-based imaging [13], where subwavelength apertures are utilized. In this manuscript, we theoretically and experimentally characterize the variation of numerical apertures with diameters of small circular apertures from the subwavelength regime to the wavelength scale regime.

The numerical aperture (NA) of a conventional microscope system is defined as:

\[ NA = n \sin \theta \]  

where \( \theta \) is the half angle of the largest cone of light that can be collected by the collection optics of the system (e.g. an objective lens). The NA is of particular interest in imaging systems that utilize widely-spaced submicron apertures for high resolution imaging. The optofluidic microscope system (OFM), a high-resolution lenless microscope system that deliver specimens microfluidically across slanted arrays of micrometer-size apertures fabricated on a metal-coated CMOS sensor to produce projection images, is one example of such imaging systems [13,14]. In aperture-based light collection systems such as the OFM, we define \( \theta \) as the incidence of a plane wave where its transmission is half that of the maximum. The full width half maximum (FWHM) angle is thus \( 2\theta \).

While it is intuitive that the NAs of such systems are related to the size of the apertures, the exact relationship cannot be assumed especially in the subwavelength regime. Since the NA of an imaging system is closely related to the resolving power of the system, it is important to characterize the relationship of the NA with diameter of the circular apertures. This knowledge will guide the choice of the sizes of circle apertures used in lensless, aperture-based imaging systems.

The optical geometry of the system we are discussing here deserves further explanation and justification as it is closely related to the definition of the NA of the system, and is fundamental to the work reported by this paper. Heng, et. al. termed imaging systems where apertures are as collection units type-II aperture-based imaging systems (ABID) [15]. The OFM is an example of a type-II ABID. In order to determine the numerical aperture of a type-II ABID, the geometry as illustrated in Fig. 1 is most straightforward: A plane wave illuminates a circular aperture in a thin metal film at varying incident angles to its normal. A detection unit below the aperture collects the transmitted light.

This configuration is realistic and relevant to the context of the type-II ABID and can be used to characterize the NA of a type-II ABID. A wave generated by any source can be considered a superposition of plane waves. A sample modulates a normally incident uniform source of illumination mainly by scattering and/or absorbing the light. The light exiting the sample with varying angular direction and magnitude, also essentially a superposition of plane waves, is passed through a circular aperture before being collected and detected by the detection unit. Light incident at different angles are collected by the circular aperture with varying collection efficiencies. This can be characterized by measuring the transmitted power of light incident on the circular apertures at different angles, using the optical geometry in Fig. 1. The values of transmitted power can be plotted against the corresponding angles of incidence. The FWHM of the plot obtained is taken to be the acceptance angle of the aperture. This acceptance angle is directly related to the NA of the system as stated in Eq. (1).
The primary objective of this work is to characterize the FWHM acceptance angles of single circular apertures with diameters ranging from the subwavelength regime to the wavelength scale regime by theoretically modeling and experimentally measuring transmission characteristics of a plane wave incident onto a single circular aperture at varying incident angles. This work will further understanding of the behavior of light transmission through subwavelength apertures and guide the design of better type-II ABIDs where apertures are used as light collection units. We will elaborate on the significance of this work on the design of type-II ABIDs in Section 2. We will describe the experimental methods employed to characterize the FWHM angles of the circular apertures in Section 3. The electromagnetic simulation method used to model this optical geometry will be reported in Section 4. In Section 5, we will discuss the goodness of fit between the experimental data and the simulation results. In Section 6, we will discuss and address the following: 1) the non-intuitive trend of the variation of FWHM acceptance angles with aperture sizes. 2) the theoretical FWHM acceptance limit of a type-II ABID as determined by the diameter of its circular apertures.

2. Significance of work on type-II ABID design

In order to fully appreciate the significance and motivation of this work, we will presently discuss the effect of the NA of a standard microscope and a type-II ABID on their depth of field and resolution.

The depth of field of the system is related to its NA and is an important parameter in any imaging system. A large NA and thus a short depth of field are desirable in standard microscopes used for thin specimens or in sectioning microscopes. The sharpness of the features of the imaged object above and below the image plane quickly degrades. In other words, the NAs of such systems are related to its resolution in the z-direction. A small NA and thus a long depth of field is desirable for a dissection microscope, an optical projection tomography system and, more pertinent to the purpose of our discussion, an OFM system. Firstly, since an OFM image is essentially the projection image of an object, a greater depth of field is more forgiving to the varying vertical (z-axial) positions with respect to the collection aperture a specimen may assume as it is flowing through the OFM channels. More importantly, a greater depth of field reduces the volume of object contributing to a light spot collected by the aperture (see Fig. 2). This results in improved resolution in an ABID system. Thus, it is established that a type-II ABID like the OFM will benefit from a small NA.

However, the exact relationship of the diameter of a circular aperture with its NA is unknown. As we will see in Section 6, the NAs of circular apertures can be analytically predicted at the large aperture and small aperture limits. In the intermediate regime, such analytical solution is non-trivial. Through the characterization of the relation of a circular aperture’s collection angle to its diameter, our work that will be described in the following sections aims to answer two questions: 1) Is there an aperture diameter at which the NA of the
aperture is at a minimum? 2) If so, at what diameter does a circular aperture have the smallest NA?

![Diagram](image)

(a) Smaller depth of field  
(b) Greater depth of field

Fig. 2. Illustration of effect of depth of field on imaged spot size and thus resolution.

3. Experimental method

The geometry shown in Fig. 1 is used for the characterization of the NA of a circular aperture. We define the NA of the aperture as in Eq. (1) where $2\theta$ is FWHM angle of the transmission through an aperture from a plane wave.

The experimental setup is illustrated in Fig. 3(b)–3(d). A diode pumped solid state laser ($\lambda = 532$ nm, 100 mW) is expanded and collimated to act as the plane wave source. The circular apertures were milled in a titanium film ($t = 200$ nm) coated on a thin glass cover slip by high resolution focused ion beam milling (FEI Nova 200 dual-beam FIB). The aperture diameters ranged from 100 nm to 900 nm. The aperture sizes were verified by using a scanning electron microscope (FEI Nova 200 dual-beam FIB). The image of an aperture 100nm in diameter is shown in Fig. 3(a). For each aperture diameter value, multiple holes were milled on the titanium film with more than 50µm spacing between apertures in order to avoid crosstalk induced by the adjacent apertures. The light transmitted through the aperture is collected by an oil immersion objective lens (DIN 100x, NA = 1.25, Edmund Optics) and detected by a monochrome CCD camera (DMK 31BU03, the Imaging Source) attached to the objective via a 160mm lens tube. Immersion oil with a refractive index of 1.518 is used. In order to vary the incident angle of the laser beam precisely, the titanium coat glass slide with milled apertures and the detection setup are assembled on a motorized rotation stage. This is illustrated in Fig. 3(c) and 3(d). The experiment is repeated on at least three apertures of each diameter. Each set of experimental data is Gaussian fitted. The average FWHM angle and the corresponding standard deviation are obtained for circular apertures of diameters: 100 nm, 200 nm, 300 nm, 400 nm, 500 nm, 600 nm and 900 nm.
To obtain the true collection angle of the apertures, the detector unit should have $\theta = 90$ deg. Our experimental setup, however, has a detector setup with finite NA. This deviation is caused by the practical limits of a real light collection and detection system. The consequence of this deviation will be addressed by simulations (Section 3) and discussed in Section 4.

Representative results from apertures with diameters 100 nm, 400 nm, 600 nm and 900 nm are shown in Fig. 4. We note that the FWHM angles of the apertures decreases as the aperture diameters decrease from 900 nm to 400 nm, while the FWHM angles of the 400 nm aperture is smaller than that of the 100 nm aperture. The second observation may appear non-intuitive, but will be further supported by more experimental data, simulation results and theoretical explanation in the following sections.
Fig. 4. Representative experimental data for apertures of diameters (a) 100 nm, (b) 400 nm, (c) 600 nm and (d) 900 nm.

4. Simulation method

In order to verify our experimental findings, we employed a numerical simulation scheme using the time-domain solver in CST Microwave Studio.

A plane wave ($\lambda = 532$ nm) is used as the excitation source. A thin square perfect electrical conductor (PEC) sheet ($t = 50$ nm, $w = 10\lambda$) with an aperture of diameter $d$ is used to approximate the experimental setup (see Fig. 5). The background material is vacuum. The thin PEC sheet is a fair approximation of the Titanium layer of finite thickness since Titanium has a very shallow skin depth of 25 nm at $\lambda = 532$ nm. The power transmitted is integrated over a square detector ($a = 6\lambda$) placed at a distance $h = \lambda$ below the aperture. The integration area was not extended to the entire PEC sheet to avoid integrating numerical errors at material boundaries into the results.

Fig. 5. Illustration of geometry for simulation with CST.

As mentioned in Section 3, the experimental setup deviated from the ideal because of the finite NA of the detector. To investigate the effect of collection optics with limited NA,
simulation results are also obtained by integration of transmitted power over a smaller square area, \( a = 2.9\lambda \). This approximates the light cone that can be collected with the experimental detection unit consisting of microscope objective with NA = 1.25 utilizing standard immersion oil for microscopy \((n = 1.518)\). The acceptance angle of the objective and with immersion oil is calculated to be 110.9 deg using Eq. (1). The circular cross section of a light cone with this acceptance angle at \( h = \lambda \) below the aperture has a radius of 1.45\(\lambda\). This is approximated in our simulation by a 2.9\(\lambda\) by 2.9\(\lambda\) square area. This method of simulating a detector with limited NA is justified by assuming that rays of light in the far field does not deviate from its path before reaching the detector optics.

Figure 6 shows the examples of 2D plots of power flow obtained with the simulation. Figure 7 shows representative plots of the simulation results. For each simulation with the same aperture size, smaller values of FWHM angles are obtained from a simulation with a smaller detector than with a larger detector. The difference is, however, insignificant when the FWHM angle is small compared to the collection angle of the detector [see Fig. 7(b), \( d = 400 \text{ nm} \)]. These observations will be further discussed in Section 5.

Fig. 6. Examples of simulation results showing power flow due to plane wave incident at oblique (−45 deg) and normal incidence (0 deg) at three representative aperture sizes \((d = 100 \text{ nm}, 400 \text{ nm} \text{ and } 900 \text{ nm})\).
Fig. 7. Plots of the total detected power versus incident angle of the plane wave obtained from simulation results with aperture diameters (a) 100 nm, (b) 400 nm, (c) 600 nm and (d) 900 nm.

5. Comparison of simulation results to experimental results

In this section, we will compare the experimental results with the simulation results for the purpose of cross validation. The FWHM angles of the transmission profiles obtained by simulations (with large detector and with limited NA small detector) and by experiments are plotted in Fig. 8. The standard deviation of the experimental FWHM measurements at each aperture diameter is shown by the error bars.

We note that the experimental FWHM versus aperture diameter trend closely follows the numerical prediction obtained from the simulation results obtained with a large detector. However, there appears to be a systematic deviation of the experimental FWHM from the FWHM values obtained from simulation. As mentioned in Section 3, these systematic errors are expected in the experimental data due to the finite NA of the collection optics in the detector unit. Due to the directionality in transmission (see Fig. 6), a fraction of the transmitted light will not be collected by the detector unit when plane wave is incident at an oblique angle. This is not an issue when the detector is large or when the FWHM collection angles of the apertures are sufficiently small such that the detector NA is not limiting. However, the experimental collection optics has limited NA. Thus, the FWHM angle measured will be smaller than that in the ideal scenario with no detector NA limitation. This explanation is validated by considering the simulation results obtained with a small detector that mimics the limited NA detector in the experimental setup. The systematic deviation between the simulation results and the experimental results decreased when a small detector was employed in the simulation (see Fig. 7 and Fig. 8). The source of the deviation of experimental data from the simulation results is thus validated. Furthermore, this result justifies the simulation model and allows the corroboration of the experimental data with the simulation results.
6. Discussion

In the previous section, we showed good agreement between the experimental data and the numerical prediction obtained from our simulation model. Both simulation and experimental results showed that the collection angle of an aperture decreases with increasing aperture diameter in the zero and single mode propagation regime. The trend was reversed in the multimode propagation regime. Most important to our purpose, we found that a circular aperture with a diameter of 400 nm had the smallest FWHM acceptance angle (67 deg) and thus the smallest NA.

This section is devoted to the discussion of the above-mentioned trend of FWHM angle versus aperture size. In order to facilitate further discussion, the results are plot in Fig. 8 along with indication of regions of possible propagation modes corresponding to the aperture sizes. For a PEC, the zero to single mode propagation transition occurs at 305 nm and the single to multimode propagation transition occurs at 406 nm [16].

For apertures with diameters ranging from 400 nm to 900 nm, it is not surprising that the FWHM collection angle of the aperture increases with aperture diameter. In this regime, as the aperture diameters increase, the collection efficiencies of incident light at oblique angles increase because multimodes can be excited. At the large aperture limit (\(d \gg \lambda\)), we expect the transmitted power of the incident plane wave will no longer be limited by the collection efficiency of the apertures but will become equivalent to the component of the power of the plane wave normal to the detector plane. Therefore, the transmitted power is expected to be the cosine of the incident angle. FWHM angle is expected to be equivalent to the FWHM of the cosine function, 120 deg. This theoretical limit is indicated in Fig. 8 along with the results obtained. In an experiment, the limit is imposed by the NA of the detector unit as have been discussed before.

While it is intuitive that FWHM acceptance angle of the aperture should increase as the diameter increases from 400 nm to 900 nm, the cause of the increasing FWHM acceptance angle with diameter as aperture size decreases from 400 nm to 100 nm (single and zero mode transmission regime) requires more reasoning.
In the single mode propagation regime, the first mode that can be excited in the aperture is the TE$_{11}$ mode. In the zero mode regime, no propagative mode exists. However, because 1) the TE$_{11}$ mode decays gently along the depth of the aperture, and 2) the aperture is in a thin film of metal and thus is shallow, the following reasoning applies to both the single and the zero mode propagation regimes.

The transverse electric field components of the TE$_{11}$ mode, the lowest mode in the circular PEC waveguide, are

\begin{align}
E_\rho &= \frac{-i\omega \mu}{\left(\omega^2 \mu \epsilon - k_z^2\right)} \rho J_1(k_\rho \rho) \cos \phi e^{i \zeta z} \\
E_\phi &= \frac{i\omega \mu}{\left(\omega^2 \mu \epsilon - k_z^2\right)} J'_1(k_\rho \rho) \sin \phi e^{i \zeta z}
\end{align}

where $k_z$ and $k_\rho$ are wave vector along ‘z’ and ‘\rho’ directions [see Fig. 9(a)]. As can be seen in Fig. 9(b), the phase difference across the x direction of the electric field profile is small.

At oblique incidence, a phase difference exists between the electric field vectors along the x-axis due to the differences in path length, $\delta$. The coupling efficiency of this mode increases with decreasing the incident angle. Because $\delta$ decreases with decreasing aperture radius, the coupling efficiency increases with decreasing diameter at each incident angle for apertures with diameters in the below the multimode cutoff.

As we approach the small aperture limit ($d \ll \lambda$), the electric field of the plane wave incident onto the aperture is in phase. There is no mode coupling preference with respect to the incident angle. The power transmitted is equivalent to the component normal to the detector plane of the power of the plane wave incident on the aperture. Similar to the case in the large aperture limit, we expect the transmitted power to be the cosine of the incident angle. Bethe, in the 1940s, concluded the same by solving the Maxwell equation analytically for a circular aperture with a diameter that is small compared to a wavelength. The transmission coefficient of a circular aperture where $d \ll \lambda$, as derived from Eq. (57) in Bethe’s work [17] is

\begin{equation}
T = \frac{8}{27\pi^4} \left(kd\right)^4 \cos \theta
\end{equation}

This indicates that the transmission versus angle of incidence is a cosine profile and thus has a FWHM angle of 120°. This is in line with our prediction and reasoning that stemmed from our experimental and numerical observations.
In the multimode propagation regime, the coupling preference of the incident angle becomes less significant even when the electric field vectors across the hole are not in phase because multi-modes can be excited. Therefore, it is not surprising that in the zero and single mode regime, the FWHM acceptance angle of the apertures decreases with increasing diameter with a minimum observed just below the multimode cutoff; and with further increasing aperture diameter beyond the multimode cutoff, the FWHM acceptance angle of the apertures increases.

Thus, it has been shown that the collection angle of an aperture decreases with increasing aperture diameter in the zero and single mode propagation regime. The trend is reversed in the multimode propagation regime. Most importantly, a circular aperture with a diameter of 400nm has the smallest NA in our set of experiments. This set of findings holds significance in the optimization of the resolution of a type-II ABID system, like the OFM.

To illustrate the significance of our work to the optimization of the resolution of a type-II ABID system like the OFM, we can compare the projection spot size of an object at a given height above a 400 nm aperture and a 900 nm aperture. These two aperture sizes were respectively found to have the largest (89 deg) and the smallest (68 deg) FWHM acceptance angle in the experiment we performed. The resolution improvement when a 400nm aperture (minimum FWHM) is used, compared to a 900nm aperture (the largest aperture investigated in this study), can be calculated by comparing the tangents of the FWHM half angles ($\tan \theta_{400nm}/\tan \theta_{900nm}$) of the apertures. According to our experimental data, using an optimized aperture can decrease the projection spot size by 30%. This improvement can be even more significant, as shown by our simulation with a large detector unit, if the detector NA is more ideal than that used in our experiments. Furthermore, since the FWHM acceptance angle of the apertures increases as the aperture diameter decreases below 400 nm, our findings suggest that using an aperture smaller than 400 nm in diameter provides diminishing returns in terms of resolution.

7. Conclusion

In summary, the acceptance angles of subwavelength to wavelength scale circular apertures were characterized via both experiments and simulations. We performed a set of experiments with a collimated laser beam ($\lambda = 532$ nm) incident on a rotating set up consisting of apertures in a thin titanium film (diameters ranging from 100 nm to 900 nm) and a limited NA detector unit. Using the same geometry, a set of simulations was also performed using CST Microwave Studio. The simulation results verified our experimental findings: 1) For apertures with diameters in the multimode propagation regime, the FWHM acceptance angle of the apertures increase with increasing aperture diameter. 2) For apertures with diameters below the multimode cutoff, the FWHM acceptance angles of the apertures decrease with increasing aperture diameter. 3) The diameter of circular aperture for which the collection angle and thus NA is at a minimum was approximately 400 nm, the aperture diameter at which the single mode to multimode transition occurs. Based on our findings, it is possible to capitalize on the existence of a sweet spot in the relationship between NA and diameter of circular apertures for a good ABID design. The minimum NA occurred at an aperture diameter of 400 nm, near the single mode to multimode transition in our set of experiments.

Acknowledgements

The authors thank Dr Meng Cui, Dr Jigang Wu, Dr Xin Heng and Dr Xiquan Cui for their helpful discussions. We also thank the meticulous reviewers for kindly pointing out the formatting and typographical errors in the draft of this paper and suggesting revisions benefiting the clarity of this manuscript. Ying Min Wang acknowledges support provided by the National Science Scholarship from the Agency of Science, Technology and Research (A*STAR), Singapore. This project is funded by Wallace Coulter Foundation and NIH R21EB008867.