12.1 Confocal microscopy

Confocal microscopy ~~is~~ can render depth-resolved slices through a 3D object by rejecting a much of the out of focus light via a pinhole. ~~In contrast with the microscopy methods discussed so far~~

~~...~~ In confocal microscopy the image is ~~formed~~ reconstructed serially, i.e. point by point, using a single photodetector, rather than in parallel (e.g., live cell, bright field, phase contrast, etc), where ~~at~~ a 2D image is recorded via a detector array, such as a camera. Thus, the illumination light is focused down to a small spot at the sample plane, while the detector records the light originating from the same spot, hence the name confocal. This method has been described in detail in many references ~~such as~~ [Wilson - Sheppard, Pawley]. In the following we describe the basic principles of confocal microscopy.

12.1.1. Principle

The principle of confocal microscopy was described prior to the invention of lasers [Ref. Minsky Patent 1957]. However, today most confocal imaging systems use lasers for ~~as~~ illumination. The confocal imaging system can operate either in transmission ~~or~~ reflection, as shown ~~in Fig. 1~~ in Fig. 1.

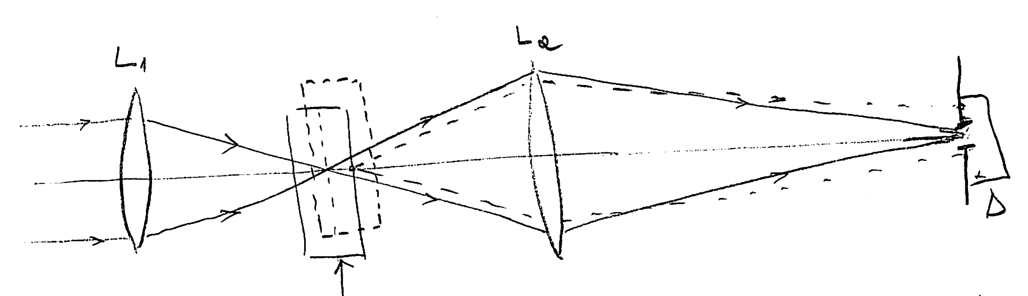

Fig. 1 a. Confocal imaging in transmission: $L_{1,2}$ lenses, S specimen, D detector; arrow indicates the plane of focus.

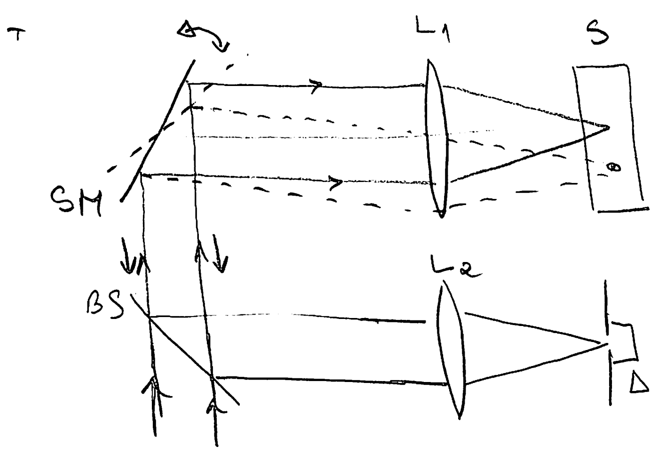

Figure 1b. Confocal imaging in reflection (epi-illumination): $L_{1,2}$ lenses, S pecimen, BS beam splitter, SM scanning mirror.

In both cases, the out of focus light is rejected by the pinhole in front of the detector, which is placed at a plane conjugate to the illumination plane.

The image reconstruction is performed either by scanning the sample or the specimen. Note that the transmission geometry (Fig. 1a) requires that the specimen is translated. Otherwise, i.e. scanning the illumination beam requires that the pinhole translates synchronously, which is impractical.

By contrast, in reflection, although the illuminating spot is translated within the sample by the scanning mirror (SM), the light is focused at the detector into a fixed point; It is said that the light is automatically "descanned". For imaging the specimen in the second transverse direction, a second scanning mirror is necessary.

Clearly, the "epi" geometry (reflection, Fig. 1b) is the main choice for fluorescence confocal microscopy, because the excitation light can be filtered-out more efficiently than in transmission.

(x) Of course, scanning the beam can be made much faster by using galvo-mirrors, for example, while translating the specimen is limited by ~~the~~ inertia.

12.1.2 Resolution

Since confocal microscopy renders 3D images, we must define both transverse and longitudinal resolution (Fig. 2), which establishes the closest two points that can be "resolved" perpendicular to and along the optical axis, respectively.

x, z, S, δx, δz, Ob

Fig. 2. Transverse (δx) and longitudinal (δz) resolutions in confocal microscopy: Ob, objective, S, specimen.

As with any type of imaging, what we consider resolved is subject to convention. Still, no matter the convention, the impulse response, or its Green's function for the confocal system, is given by the 3D distribution of the field in the vicinity of focus. In order to calculate this field distribution, we need to describe the input field. We distinguish here three cases, although a similar discussion can be expanded to include further cases. Figure 3 shows the three cases: a) plane wave truncated by the lens aperture, b) Gaussian beam truncated by lens aperture, and c) Gaussian beam unrestricted by lens aperture.

R

a) b) c)

Figure 3. Illumination beams at the lens aperture in confocal microscopy: a) truncated plane wave; b) truncated Gaussian beam; c) infinite Gaussian beam.

In all cases, the transmission function of the lens can be approximated by $e^{-i k_0 \frac{(x^2+y^2)}{2f}} \cdot P(x,y)$.

(x) $$t(x,y) = e^{-i\frac{k_0(x^2+y^2)}{2f}} \cdot P(x,y). \qquad (1)$$

In Eq. 1, f is the focal distance of the lens, $k_0 = \frac{\omega}{c}$, and P is the aperture function, which for a circular aperture of radius R has the form

$$P(x,y) = \begin{cases} 1, & x^2+y^2 \leq R^2 \\ 0, & \text{rest} \end{cases}$$

$$= \Pi\left(\frac{\sqrt{x^2+y^2}}{2R}\right). \qquad (2).$$

In 1956, Linfoot and Wolf provided the field near focus for case a) [Linfoot and Wolf, 1956], i.e., plane wave incident. The field propagating behind the lens, U, is the convolution of $t(x,y)$ and the spherical wavelet, i.e.

$$U(x,y,z) = t(x,y) \circledv \frac{e^{ik_0\sqrt{x^2+y^2+z^2}}}{\sqrt{x^2+y^2+z^2}}. \quad (3)$$

Recall from chapter 4 that the spherical wave can be expressed in the plane wave representation, the (k_x, k_y, z) representation, referred to as the Weyl's formula [Weyl, 1909]

$$\frac{e^{ik_0 r}}{r} \xrightarrow{\mathcal{F}_{x,y}} \frac{e^{iqz}}{q} \quad (4a)$$

$$r = \sqrt{x^2+y^2+z^2} \quad (4b)$$

$$q = \sqrt{k_0^2-k_x^2-k_y^2} \quad (4c)$$

Thus, Fourier transforming Eq. 3 with respect to variables (x,y), we obtain

$$\tilde{U}(k_x,k_y,z) = \frac{e^{iqz}}{q} \cdot \tilde{t}(k_x,k_y)$$

$$= \frac{e^{iqz}}{q} \cdot \left[e^{-i\frac{f(k_x^2+k_y^2)}{2k_0}} \circledv \frac{J_1\left(\sqrt{k_x^2+k_y^2}\cdot R\right)}{2\sqrt{k_x^2+k_y^2}\cdot R} \right]. \quad (5)$$

In Eq. 5, J_1 is the Bessel function of first order (and first kind). Evaluating the right hand side of Eq. 5 requires numerical integration, as Linfoot and Wolf performed more than 50 years ago. Figure 4 shows the amplitude and phase of field $U(x,y,z)$.

According to Abbe's criterion, the transverse resolution (δx) is the radius of the first dark ring (Airy ring) in the plane of focus, $z = f$. Similarly the axial resolution (δz), is the distance from $z=0$ to the first zero on the z-axis (Fig. 4). Evaluating Eq. 5, δx and δz are

$$\delta x = 0.61 \frac{\lambda}{NA} \qquad (6a)$$

$$\delta z = \frac{2 n \lambda}{(NA)^2}, \qquad (6b)$$

where λ is the wavelength in vacuum and n is the average refractive index of the object. It is important to note that unlike transverse resolution, axial resolution has a stronger dependence on NA, is proportional to $1/NA^2$.

Gaussian beams are most relevant for laser scanning confocal systems. Note that an infinite Gaussian beam, i.e. not truncated by an aperture (Fig. 3c), does not generate zeros of intensity around the focal point of the lens. While this situation cannot be attained fully in practice, experiments with a Gaussian beam with its waist much smaller than the lens aperture can be approximated well by such a beam. In this case, a good measure for axial resolution is the Rayleigh range, z_0 (Fig. 4). As discussed in Chapter 4, z_0 is the distance over which the beam waist, w, increases to $\sqrt{2}$ of its minimum, w_0

$$z_0 = \frac{\pi w_0^2}{\lambda} \qquad (7a)$$

$$w^2 = w_0^2 \left(1 + \frac{z}{z_0}\right) \qquad (7b)$$

$$w_0 = \frac{\lambda}{\pi \theta} \qquad (7c)$$

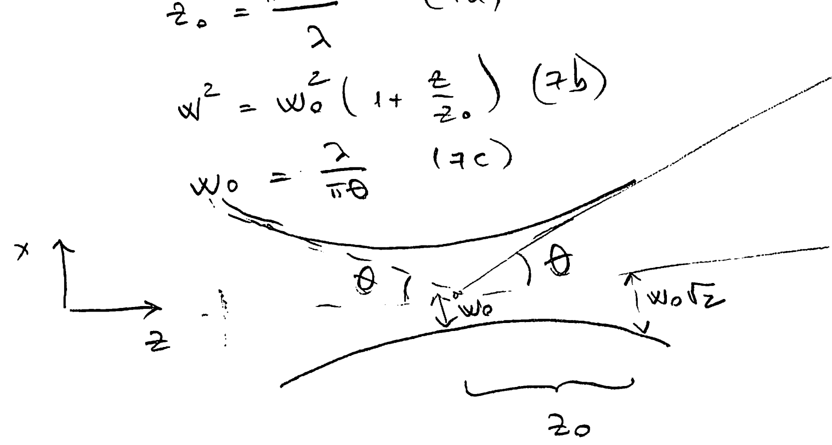

Fig. 4. Diffraction of a Gaussian beam: w_0 minimum waist, z_0 Rayleigh range, θ diffraction angle.

In Eq. 7c, θ is the diffraction angle, which is the analog of the NA in Eq. 6a. Similarly, combining Eqs. 7a and 7c, we obtain z_0, which is a measure of axial resolution,

$$z_0 = \frac{\lambda}{\pi\theta^2}, \quad (8)$$

where we recover the dependence on angle squared, encountered in Eq. 6b.
Thus, Equations 7b and 8 can be used as measures of transverse and longitudinal resolution, respectively, for the idealized case of infinitely broad Gaussian beams. Of course, other conventions for resolution can be used, but the dependence on wavelength and numerical aperture of the lens will remain the same.

Clearly, aberrations, both chromatic and geometrical, have the net effect of lowering the resolving power of the confocal microscope. Aberration correction is a critical issue especially when operating at very high NA [Ref aberr.]. However, even with an instrument that is capable in principle of producing high resolution images, the amount of details ultimately distinguishable in the final image depends also on how different the various regions of instrument are, i.e. on contrast.

12.1.3 Contrast

Note that confocal microscopy is intrinsically an intensity-based method. Therefore, we anticipate low contrast when operating in reflectance mode, as the refractive index variation within tissues is not very high. Further, the residual out of focus light that makes it to the detector has the effects of blurring the image, lowering both the contrast and resolution.

Thus, there is always a compromise between the size of the pinhole and the contrast of the image. A small pinhole generates higher contrast at the expense of lower signal detected.

Most commonly, confocal microscopy is used with fluorescence, which provides significantly higher contrast. Theoretically, the contrast in fluorescence image is infinite, i.e. untagged (background) structures give zero signal. However, practical issues related to dark signals in the detector, limited dynamic range, saturation, and out of focus light, lower the contrast. Still, fluorescence confocal microscopy offers a great tool for biological studies, especially cell biology, as illustrated in the section 12.1.5.

12.1.4. Further developments

Confocal microscopy has important advantages over full-field methods, including the ability to section through optically thick specimens in 3D, field of view restricted only by the scanning range, and enhanced resolution.

Past and current research in confocal microscopy deals mainly with achieving higher acquisition rates and deeper penetration. The "spinning disk" confocal microscope was built upon an idea due to Nipkow (Nipkow, 1884). It uses a rotating disk with perforated pinholes arranged in a spiral, such that the entire field of view is scanned upon one rotation of the disk, which allows fast acquisition rates [Ref. disk]

The penetration depth has been improved significantly once confocal microscopy was combined with nonlinear optics, in particular two-photon fluorescence [Ref.]. This will be discussed in more detail in chapter 16.

12.1.5. Biological applications