Manuscript submitted to **Biophysical** *Journal*
Article

Variational calculus approach to Zernike polynomials with application to fluorescence correlation spectroscopy

Ivan Gligonov[1] and Jörg Enderlein[1,2,*]

[1]Third Institute of Physics – Biophysics, Georg August University, 37077 Göttingen, Germany
[2]Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC), Universitätsmedizin Göttingen, 37075 Göttingen, Germany
[*]Correspondence: jenderl.gwdg.de

ABSTRACT Zernike polynomials are a sequence of orthogonal polynomials that play a crucial role in optics and in particular in modeling microscopy systems. Introduced by Frits Zernike in 1934, they are particularly useful in expressing wavefront aberrations and thus imperfections of imaging systems. However, their origin and properties are rarely discussed and proven. Here, we present a novel approach to Zernike polynomials using variational calculus, and apply them to describe aberrations in fluorescence microscopy. In particular, we model the impact of various optical aberrations on the performance of one-photon and two-photon excitation fluorescence microscopy.

SIGNIFICANCE This manuscript explores the mathematical derivation of Zernike polynomials and highlights their critical role in describing optical wavefronts and aberrations, particularly in the domain of optical microscopy. Special emphasis is placed on their utility in simulating the effects of aberrations on one-photon and two-photon excitation fluorescence correlation spectroscopy.

INTRODUCTION

This paper is dedicated to the work and achievements of Watt W. Webb. Among his many important contributions to spectroscopy and microscopy are the development of Fluorescence Correlation Spectroscopy (FCS), together with Douglas Magde and Elliott Elson, (1–3), of two-photon excited fluorescence microscopy (4), together with Winfried Denk and James H. Strickler, and, the propagation of two-photon excitation FCS (5, 6), together with Petra Schwille and colleagues, which was first developed by Berland et al. (7). In all these methods, optical aberrations play an important role, because they lead to a change in the detection volume of a confocal microscope. In an FCS experiment, this results in a delayed decay of the correlation functions. Likewise, they will reduce the optical resolution of a two-photon microscope. The canonical way to describe optical aberrations caused by wavefront deformations in an optical system is to expand the wavefront deformations into an orthogonal set of functions over the unit disk, usually the Zernike polynomials. Frits Zernike, the inventor of the phase-contrast microscope (8) (which earned him the Nobel Prize in Physics in 1953), introduced these polynomials in 1934 (9). Their orthogonality makes them ideal for decomposing the wavefront error in optical systems into simpler, constituent aberrations. This decomposition enables an efficient understanding and correction of optical aberrations, which are critical for enhancing the quality of images produced by optical microscopy (10). Furthermore, Zernike polynomials play a crucial role in the development of adaptive optics systems for microscopy. These systems, which measure and actively compensate for wavefront errors, leverage devices like deformable mirrors or liquid crystal spatial light modulators. These devices are controlled based on the Zernike polynomial representation of aberrations, allowing for correction of distortions and the achievement of higher resolution and contrast images, especially in deep tissue imaging (11).

Although Zernike polynomials are well defined in many textbooks and research papers (12–14) and even on Wikipedia (15), the reasons behind their specific forms and the possibility of alternative function sets over the unit disk are less commonly known. This paper presents a derivation of Zernike polynomials based on variational calculus, providing insights into why they represent the simplest polynomial orthogonal function set over the unit disk. Along the way, we derive Bessel functions from a similar variational approach, noting that they are polynomial functions of infinite order.

In the second part of this paper, we will study the impact of different aberrations described by distinct Zernike functions on one-photon and two-photon excitation experiments. Although there have been extensive studies on the impact of various optical system imperfections on FCS experiments (see, e.g., Refs. (16, 17)), we are not aware of a comparative study between one-photon and two-photon FCS.

ORTHOGONAL FUNCTIONS OVER THE UNIT DISK: BESSEL FUNCTIONS

An ideal microscope transforms an incoming spherical wavefront from a point source into an outgoing spherical wavefront that converges into a focus point in the image plane. However, real optical systems are never perfect and exhibit various deviations from this ideal behavior, collectively called aberrations. The most significant aberrations are due to an additional position-dependent *phase* of the wavefront. Although the field amplitude across the wavefront can also be affected by an optical system (through attenuation/apodization), this effect is usually less important because it does not alter the shape of the resulting point spread function as severely as phase variations do. Therefore, we will focus here solely on wavefront deformations. Conveniently, this phase is described by a system of orthonormal functions over the unit disk, where the unit disk maps the spherical wavefront of the non-aberrated system. Therefore, we aim to identify a set of orthogonal functions defined on the unit disk that is most suitable for describing aberrations.

Our goal is to find a linear differential equation whose eigenfunctions form a set of orthogonal functions over the unit disk. The canonical approach is to define a functional over the unit disk, where its minimization leads to the desired differential equation. Naturally, such a functional will be constructed using derivatives of the scalar function f, for which the differential equation is sought. The lowest possible derivative, which is also independent of the underlying coordinate system, is the gradient of f. However, the gradient returns a vector field, whereas the functional must be a scalar. Therefore, the simplest possible functional is constructed from the square (scalar product) of the gradient, i.e., it is given by the integral

$$\int |\mathbf{grad} f|^2 \, dA \tag{1}$$

where dA represents the area element of the disk. In other words, we essentially target functions that show minimal deviation from being constant. Additionally, these functions should be normalized, which means the integral

$$\int f^2 dA \tag{2}$$

should equal one. By employing a Lagrange multiplier λ^2, we can merge this normalization condition with the initial minimal bending energy requirement into seeking functions that minimize the functional

$$S = \int_A dA \left(|\mathbf{grad} f|^2 - \lambda^2 f^2 \right) = \int_0^1 dr\, r \int_0^{2\pi} d\phi \, \left(|\mathbf{grad} f|^2 - \lambda^2 f^2 \right) . \tag{3}$$

The simplest function meeting our criteria is a constant function. However, due to the orthogonality constraint, each subsequent function in our series must exhibit increasing complexity compared to all previously identified functions. This approach enables the construction of a sequence of functions where each new addition not only minimizes the mean-square gradient but also remains orthogonal to every preceding function, facilitating a systematic and unique expansion of the function space over the unit disk.

The variation of the above functional will lead to a linear partial differential equation for f in the variables r and ϕ, which can be analyzed and solved using the standard method of separation of variables. However, it is more convenient to perform the separation of variables before minimizing the functional (3) and to focus on the resulting differential equation for the radial part of f. Given that the functions f must exhibit 2π periodicity in the angular variable ϕ, we decompose f into $f = R(r)\cos m\phi$ or $f = R(r)\sin m\phi$, where $R(r)$ are real functions and m is an integer number. This decomposition leads, up to a constant pre-factor, to the following functional

$$S = \int_0^1 dr\, r \left[\left(\frac{dR}{dr} \right)^2 + \left(\frac{m^2}{r^2} - \lambda^2 \right) R^2 \right] \tag{4}$$

Variation of this functional yields

$$\delta S = 2 \int_0^1 dr\, r \left[\frac{d(\delta R)}{dr} \frac{dR}{dr} + \left(\frac{m^2}{r^2} - \lambda^2 \right) \delta R R \right] \tag{5}$$

which, upon integration by parts, results in

$$\delta S = 2 \left(\delta R \frac{dR}{dr} \right)_{r=1} - 2 \int_0^1 dr\, r\, \delta R \left[\frac{1}{r} \frac{d}{dr} r \frac{dR}{dr} + \left(\lambda^2 - \frac{m^2}{r^2} \right) R \right] \tag{6}$$

For the variation to disappear, the radial function $R(r)$ must satisfy the differential equation

$$\frac{1}{r}\frac{d}{dr}r\frac{dR}{dr}+\left(\lambda^2-\frac{m^2}{r^2}\right)R=0 \tag{7}$$

with the boundary condition $dR/dr=0$ at $r=1$ ensuring that the partially integrated part in (6) disappears. Equation (7) is the Bessel differential equation, and its real and finite solutions $J_m(\lambda r)$ are the Bessel functions of the first kind J_m.

Thus, the looked-for orthonormal function set on the unit disk is given by

$$f_n^m(r,\phi)=c_n^m J_m(\lambda_n^m r)\begin{Bmatrix}\cos m\phi\\ \sin m\phi\end{Bmatrix} \tag{8}$$

where the λ_n^m are the roots of the equation

$$\left(\frac{dJ_m(\lambda_n^m r)}{dr}\right)_{r=1}=0. \tag{9}$$

and c_n^m is a normalizing constant.

Let us check that two functions $J_m(\lambda_a r)$ and $J_m(\lambda_b r)$ with $\lambda_a\neq\lambda_b$ are indeed orthogonal to each other. We will use this method later also for Zernike polynomials. We first introduce the differential operator $\hat{\mathbf{O}}$ defined by

$$\hat{\mathbf{O}}J_m=\frac{1}{r}\frac{d}{dr}r\frac{dJ_m}{dr}. \tag{10}$$

With this operator, the Bessel differential equation (7) can be rewritten as

$$\hat{\mathbf{O}}J_m(\lambda r)=\left(\frac{m^2}{r^2}-\lambda^2\right)J_m(\lambda r). \tag{11}$$

Application of this operator to function $J_m(\lambda_a r)$ with left multiplication by $J_m(\lambda_b r)$ yields

$$J_m(\lambda_b r)\hat{\mathbf{O}}J_m(\lambda_a r)=\left(\frac{m^2}{r^2}-\lambda_a^2\right)J_m(\lambda_b r)J_m(\lambda_a r). \tag{12}$$

Subtraction of $J_m(\lambda_a r)\hat{\mathbf{O}}J_m(\lambda_b r)$ and subsequent integration results in

$$\int_0^1 drr[J_m(\lambda_b r)\hat{\mathbf{O}}J_m(\lambda_a r)-J_m(\lambda_a r)\hat{\mathbf{O}}J_m(\lambda_b r)]=(\lambda_b^2-\lambda_a^2)\int_0^1 drrJ_m(\lambda_b r)J_m(\lambda_a r). \tag{13}$$

Integration by parts in the left hand side of the above equation leads to

$$(\lambda_b^2-\lambda_a^2)\int_0^1 drrJ_m(\lambda_b r)J_m(\lambda_a r)=\left[rJ_m(\lambda_b r)\frac{dJ_m(\lambda_a r)}{dr}-rJ_m(\lambda_a r)\frac{dJ_m(\lambda_b r)}{dr}\right]_0^1=0 \tag{14}$$

where we have used the fact that both λ_a and λ_b shall be roots of the equation $[dJ_m(\lambda r)/dr]_{r=1}=0$.

For our derivation of Zernike polynomials later, it is useful to recall the power series method for finding explicit solutions of differential equations such as the Bessel equation (7). We expand the function $J_m(r)$ into a power series

$$J_m(r)=\sum_{j=0}^{\infty}a_j r^j \tag{15}$$

insert this into the differential equation (7) with $\lambda=1$, and compare the coefficients of equal power. This leads to the recursive relation

$$a_{j+2}=-\frac{1}{(j+2)^2-m^2}a_j=-\frac{1}{(j+2-m)(j+2+m)}a_j \tag{16}$$

for the coefficients a_j. This shows that the first non-vanishing coefficient must be a_m (to avoid infinite terms in the series of coefficients) and we find, after setting $a_m=1$,

$$J_m(r)\propto r^m-\frac{r^{m+2}}{2(2m+2)}+\frac{r^{m+4}}{2\cdot 4(2m+2)(2m+4)}-\ldots=\sum_{j=0}^{\infty}\frac{(-1)^j m!}{2^{2j}j!(m+j)!}r^{m+2j}. \tag{17}$$

This leads, after canceling the common factor $2^m m!$, to the standard definition of the Bessel function $J_m(r)$

$$J_m(r) = \sum_{j=0}^{\infty} \frac{(-1)^j}{j!(m+j)!} \left(\frac{r}{2}\right)^{m+2j}. \tag{18}$$

For finding its normalization, we consider the expression

$$\frac{1}{2}\frac{d}{dr}\left(r\frac{dJ_m(\lambda r)}{dr}\right)^2 = \left(r\frac{dJ_m(\lambda r)}{dr}\right)\left(\frac{d}{dr}r\frac{dJ_m(\lambda r)}{dr}\right). \tag{19}$$

The second bracket on the r.h.s. can be transformed with Bessel's differential equation (7) to yield

$$\frac{1}{2}\frac{d}{dr}\left(r\frac{dJ_m(\lambda r)}{dr}\right)^2 = \left(r\frac{dJ_m(\lambda r)}{dr}\right)\left(\frac{m^2}{r} - \lambda^2 r\right)J_m(\lambda r) = \frac{1}{2}\left(m^2 - \lambda^2 r^2\right)\frac{d[J_m^2(\lambda r)]}{dr}. \tag{20}$$

Thus, after integrating the l.h.s. and r.h.s of the last equation from $r = 0$ to $r = 1$, one obtains

$$\frac{1}{2}\left[\left(r\frac{dJ_m(\lambda r)}{dr}\right)^2\right]_0^1 = \frac{1}{2}\left[m^2 J_m^2(\lambda r) - \lambda^2 r^2 J_m^2(\lambda r)\right]_0^1 + \lambda^2 \int_0^1 dr r J_m^2(\lambda r). \tag{21}$$

where we have employed once again a partial integration on the r.h.s. Consequently, when choosing for λ one of the roots λ_n^m of $[dJ_m(\lambda r)/dr = 0]_{r=1}$, and taking into account that $mJ_m(0) = 0$ for all integer m, see (18), we find the normalization

$$\int_0^1 dr r J_m^2(\lambda_n^m r) = \frac{1}{2}\left(1 - \frac{m^2}{(\lambda_n^m)^2}\right)J_m^2(\lambda_n^m). \tag{22}$$

Thus, the normalizing constant c_n^m in (8) is

$$c_n^m = \frac{\lambda_n^m}{\sqrt{\pi\left[(\lambda_n^m)^2 - m^2\right]}J_m(\lambda_n^m)}. \tag{23}$$

Figure 1 shows density plots of the first 12 modes $J_m(\lambda_n^m r)\cos m\phi$ and $J_m(\lambda_n^m r)\sin m\phi$ which are similar to the first 12 Zernike functions that describe the lowest order aberrations, see Figure 2 in the next section.

ZERNIKE POLYNOMIALS

As we have seen, Bessel functions can be used to construct an orthonormal set of functions over the unit disk, but their use has several drawbacks: Bessel functions are transcendental functions that have to be computed numerically, and even finding the roots of $[dJ_m(\lambda r)/dr = 0]_{r=1}$ is a nontrivial task. Although in the current era of powerful computers this is not a principal obstacle anymore, it was much more so in the times of Frits Zernike, who wanted to come up with a much simpler set of functions where the radial part can be represented by finite-order polynomials. From the power series method used to derive the series expansion of Bessel functions, it is evident that the Bessel differential equation cannot have finite-order polynomial solutions. This is because the differential operator reduces any power of r by two, while the term with λ^2 does not. As a result, the recurrence relation for the series expansion coefficients can never terminate at a finite power of r. To rectify this situation, one may introduce into the functional S another differential operator which does *not* lower the order of a polynomial but still is *rotation-invariant*. Such an operator is $\mathrm{div}(\mathbf{r}f)$, so that we may use the modified functional

$$S = \int_A dA\left(|\mathbf{grad}\, f|^2 + \alpha\,[\mathrm{div}(\mathbf{r}f)]^2 - \lambda^2 f^2\right) \tag{24}$$

with some constant α as the starting point. As can be checked later, the alternative choice of $\mathbf{r}\cdot\mathbf{grad}\, f$ instead of $\mathrm{div}(\mathbf{r}f)$ gives the same final result for the differential equation. Although it is not guaranteed that the modified functional (24) will lead to a differential equation of f, we again use the Ansatz $f = R(r)\cos m\phi$ or $f = R(r)\sin m\phi$ which now leads to

$$S = \pi\int_0^1 dr r\left[\left(\frac{dR}{dr}\right)^2 + \alpha\left(\frac{1}{r}\frac{d(r^2R)}{dr}\right)^2 + \left(\frac{m^2}{r^2} - \lambda^2\right)R^2\right] \tag{25}$$

Variation and partial integration of this functional leads to

$$\delta S = 2\pi \left[\delta R \left(\frac{dR}{dr} + \alpha \frac{d(r^2 R)}{dr}\right)\right]_{r=1} - 2\pi \int_0^1 dr r \delta R \left[\frac{1}{r}\frac{d}{dr} r \frac{dR}{dr} + \alpha r \frac{d}{dr}\frac{1}{r}\frac{d(r^2 R)}{dr} + \left(\lambda^2 - \frac{m^2}{r^2}\right) R\right] \tag{26}$$

To make the partially integrated term at $r = 1$ independent on the derivative of R, one sets $\alpha = -1$, so that the above equation simplifies to

$$\delta S = -4\pi \left(\delta R R\right)_{r=1} - 2\pi \int_0^1 dr r \delta R \left[\frac{1}{r}\frac{d}{dr} r \frac{dR}{dr} - r \frac{d}{dr}\frac{1}{r}\frac{d(r^2 R)}{dr} + \left(\lambda^2 - \frac{m^2}{r^2}\right) R\right] \tag{27}$$

which leads to the ordinary differential equation

$$\frac{1}{r}\frac{d}{dr} r \frac{dR}{dr} - \frac{m^2}{r^2} R - r \frac{d}{dr}\left(2R + r\frac{dR}{dr}\right) + \lambda^2 R = 0. \tag{28}$$

for $R(r)$. When we use a finite-order polynomial Ansatz for $R(r)$ with highest order n, we see that this can indeed work if $\lambda^2 = n(n+2)$, and if the coefficients a_j for $j < n$ obey the recursive relation

$$a_j = -\frac{(j+2)^2 - m^2}{n(n+2) - j(j+2)} a_{j+2} = -\frac{(j+2+m)(j+2-m)}{(n+j+2)(n-j)} a_{j+2}. \tag{29}$$

Thus, we find for the radial polynomials $R_n^m(r)$ the explicit expressions

$$\begin{aligned} R_n^m(r) &= a_n \left[r^n - \frac{(n+m)(n-m)}{(2n)2} r^{n-2} + \frac{(n+m)(n+m-2)(n-m)(n-m-2)}{(2n)(2n-2)2\cdot 4} r^{n-4} - \dots\right] \\ &= a_n \sum_{j=0}^{\frac{n-|m|}{2}} (-1)^j \frac{(n+m)!!(n-m)!!}{(n+m-2j)!!(n-m-2j)!!} \frac{(2n-2j)!!}{(2n)!!(2j)!!} r^{n-2j} \end{aligned} \tag{30}$$

where $x!!$ denote the double factorial $x!! = x(x-2)(x-4)\cdots$. The index m in R_n^m describes the periodicity of the final function $R_n^m(r)\cos(m\phi)$ or $R_n^m(r)\sin(m\phi)$ in the angular variable ϕ, but n is now the polynomial order of the radial function. Importantly, we see that $n - m$ must be even so that no negative powers of r occur in the above expansion. When taking into account that for even integer numbers k one has $k!! = 2^{k/2}(k/2)!$, one finds

$$R_n^m(r) = \sum_{j=0}^{\frac{n-|m|}{2}} (-1)^j \frac{(n-j)!}{j!\left(\frac{n-m}{2} - j\right)!\left(\frac{n+m}{2} - j\right)!} r^{n-2j} = \sum_{j=0}^{\frac{n-|m|}{2}} (-1)^j \binom{\frac{n-m}{2}}{j} \binom{n-j}{\frac{n-m}{2}} r^{n-2j} \tag{31}$$

where we dropped the highest coefficient a_n and all factors that do not depend on the summation variable j.

Next, we want to calculate the value of R_n^m at $r = 1$. For doing this, consider the qth power $(x-y)^q$ of the difference between two auxiliary variables x and y and expand it into a binomial series

$$(x-y)^q = \sum_{j=0}^{q} (-1)^j \binom{q}{j} x^{q-j} y^j \tag{32}$$

When multiplying both sides of this equation with x^{n-q}, differentiating q time with respect to x, and dividing by $q!$, we find

$$\frac{1}{q!}\frac{d^q}{dx^q}\left[x^{n-q}(x-y)^q\right] = \sum_{j=0}^{q} (-1)^j \binom{q}{j}\binom{n-j}{q} x^{n-j-q} y^j \tag{33}$$

However, when using the binomial formula for expanding the qth derivative on the l.h.s., we find

$$\frac{1}{q!}\frac{d^q}{dx^q}\left[x^{n-q}(x-y)^q\right] = \frac{1}{q!}\sum_{j=0}^{q}\binom{q}{j}\frac{d^j(x^{n-q})}{dx^j}\frac{d^{q-j}\left[(x-y)^q\right]}{dx^{q-j}} = \frac{1}{q!}\sum_{j=0}^{q}\binom{q}{j}\frac{(n-q)!}{(n-q-j)!} x^{n-q-j}\frac{q!}{j!}(x-y)^j \tag{34}$$

When setting in both previous equations $x = y = 1$, this yields

$$\sum_{j=0}^{q}(-1)^j\binom{q}{j}\binom{n-j}{q} = 1 \tag{35}$$

which shows that $R_n^m(r=1) = 1$. Thus, all radial functions R_n^m are constant and equal to one along the circumference of the unit disk ($r = 1$), so that we have indeed $\delta R|_{r=1} = 0$ in the variation (27).

To establish the orthogonality of the Zernike polynomials R_n^m, we start by acknowledging that the orthogonality of $R_n^m(r)\cos m\phi$ and $R_n^m(r)\sin m\phi$ for differing m values is secured by the intrinsic orthogonality of the trigonometric functions $\cos m\phi$ and $\sin m\phi$. Our focus now shifts to demonstrating the orthogonality of $R_n^m(r)$ across different n values while maintaining the same m. For doing that we abbreviate the differential operator in (28) by $\hat{\mathbf{O}}$ as in the previous section, i.e.

$$\hat{\mathbf{O}}R = \frac{1}{r}\frac{d}{dr}r\frac{dR}{dr} - r\frac{d}{dr}\left(2R + r\frac{dR}{dr}\right) \tag{36}$$

so that, when using the differential equation (28), we have

$$R_{n'}^m\hat{\mathbf{O}}R_n^m - R_n^m\hat{\mathbf{O}}R_{n'}^m = \left[n'(n'+2) - n(n+2)\right]R_{n'}^m R_n^m \tag{37}$$

On the other hand, we obtain after partial integration,

$$\begin{aligned}\int_0^1 drrR_{n'}^m\hat{\mathbf{O}}R_n^m &= \int_0^1 dr\left[R_{n'}^m\frac{d}{dr}r\frac{dR_n^m}{dr} - r^2R_{n'}^m\frac{d}{dr}\left(2R_n^m + r\frac{dR_n^m}{dr}\right)\right]\\ &= -2\left[R_{n'}^mR_n^m\right]_{r=1} - \int_0^1 drr\left[\frac{dR_{n'}^m}{dr}\frac{dR_n^m}{dr} - \left(2R_n^m + r\frac{dR_n^m}{dr}\right)\left(2R_n^m + r\frac{dR_n^m}{dr}\right)\right]\end{aligned} \tag{38}$$

so that

$$\int_0^1 drr\left[R_{n'}^m\hat{\mathbf{O}}R_n^m - R_n^m\hat{\mathbf{O}}R_{n'}^m\right] = 0 = \left[n'(n'+2) - n(n+2)\right]\int_0^1 drrR_{n'}^mR_n^m \tag{39}$$

which proves the orthogonality of the R_n^m for different values of n but same m.

Finally, we want to find the normalization of the Zernike polynomials, i.e. we want to calculate the integral $\int_0^1 drr\left[R_n^m(r)\right]^2$. Here, we can use the fact that the R_n^m are n^{th} order polynomials that are orthogonal to each other. Thus, the part of R_n^m that contains all the powers of r smaller than n can be expressed by an expansion into polynomials $R_{n'}^m$ with $n' < n$. Due to the orthogonality of the Zernike polynomials and using the explicit expression (31), we thus find

$$\int_0^1 drr\left[R_n^m(r)\right]^2 = \int_0^1 drr\binom{n}{\frac{n-m}{2}}r^nR_n^m(r) = \binom{n}{\frac{n-m}{2}}\sum_{j=0}^{\frac{n-m}{2}}(-1)^j\binom{\frac{n-m}{2}}{j}\binom{n-j}{\frac{n-m}{2}}\frac{1}{2(n-j+1)} \tag{40}$$

To simplify the r.h.s. of this equation, we consider the integral

$$\int_0^y dx\frac{x^q}{q!}\frac{d^q}{dx^q}\left[x^{n-q}(x-y)^q\right] = \sum_{j=0}^{q}(-1)^j\binom{q}{j}\binom{n-j}{q}\frac{y^{n+1}}{n-j+1} \tag{41}$$

where the r.h.s. was found by expanding $(x-y)^q$ into the sum $\sum_{j=1}^{q}(-1)^j\binom{q}{j}x^{q-j}y^j$ and then differentiating and integrating each term of this sum separately. However, we can also perform q times a partial integration on the l.h.s., which leads to the alternative result

$$\int_0^y dx\frac{x^q}{q!}\frac{d^q}{dx^q}\left[x^{n-q}(x-y)^q\right] = (-1)^q\int_0^y dx\left[x^{n-q}(x-y)^q\right] = \binom{n}{q}^{-1}\int_0^y dxx^n = \binom{n}{q}^{-1}\frac{y^{n+1}}{n+1} \tag{42}$$

where we have, for coming from the second to the third term, applied q times a partial integration to the powers of x. By equating the r.h.s. of (41) and (42), we have

$$\binom{n}{q}\sum_{j=0}^{q}(-1)^j\binom{q}{j}\binom{n-j}{q}\frac{1}{n-j+1} = \frac{1}{n+1} \tag{43}$$

#	n	m	function name	$R_n^m(r)\begin{Bmatrix}\cos m\phi\\ \sin m\phi\end{Bmatrix}$	name
1	1	1	Z_1^1	$2r\cos\phi$	horizontal tilt
2	1	−1	Z_1^{-1}	$2r\sin\phi$	vertical tilt
3	2	0	Z_2^0	$\sqrt{3}(2r^2-1)$	defocus
4	2	2	Z_2^2	$\sqrt{6}r^2\cos 2\phi$	vertical astigmatism
5	2	−2	Z_2^{-2}	$\sqrt{6}r^2\sin 2\phi$	oblique astigmatism
6	3	1	Z_3^1	$\sqrt{8}(3r^2-2)r\cos\phi$	horizontal coma
7	3	−1	Z_3^{-1}	$\sqrt{8}(3r^2-2)r\sin\phi$	vertical coma
8	3	3	Z_3^3	$\sqrt{8}r^3\cos 3\phi$	oblique trefoil
9	3	−3	Z_3^{-3}	$\sqrt{8}r^3\sin 3\phi$	vertical trefoil
10	4	0	Z_4^0	$\sqrt{5}(6r^4-6r^2+1)$	primary spherical
11	4	2	Z_4^2	$\sqrt{10}(4r^2-3)r^2\cos 2\phi$	vertical secondary astigmatism
12	4	−2	Z_4^{-2}	$\sqrt{10}(4r^2-3)r^2\sin 2\phi$	oblique secondary astigmatism

Table 1: The first 12 Zernike functions.

and then comparing the l.h.s. of this expression with (40), we finally find

$$\int_0^1 drrR_{n'}^m(r)R_n^m(r) = \frac{\delta_{n'n}}{2(n+1)}. \tag{44}$$

The complete formulation of Zernike functions is concisely captured by the equation:

$$Z_n^{\pm m} = c_n^m R_n^m(r)\begin{Bmatrix}\cos m\phi\\ \sin m\phi\end{Bmatrix}, \tag{45}$$

wherein the superscript $\pm m$ distinguishes between the cosine function (for a positive m) and the sine function (for a negative m). The normalization coefficient c_n^m is selected to ensure that the integral over the unit disk is $\int_0^{2\pi} d\phi \int_0^1 drr(Z_n^{\pm m})^2 = \pi$. Using the normalization detailed in (44), this leads to the explicit definition of c_n^m as:

$$c_n^m = \sqrt{\frac{2(n+1)}{1+\delta_{m,0}}}, \tag{46}$$

which standardizes the Zernike polynomials across the unit disk, facilitating their application in optical systems for aberration correction and wavefront analysis.

ABERRATIONS, FOCUS SHAPE, AND FLUORESCENCE CORRELATION SPECTROSCOPY

In this section, we consider the impact of aberrations as described by Zernike functions on the focusing of light through an objective, and the consequences for fluorescence correlation spectroscopy using a confocal microscope. In the first subsection, we will consider focusing of a plane wave, in the second subsection, the detection efficiency distribution of confocal imaging, and in the third subsection, how aberrations affect the outcome of an FCS measurement.

Focusing

When focusing a linearly polarized plane wave (polarization direction along the x-direction in Figure 3) with field amplitude E_0 through an objective into a sample with refractive index n_r, the electric field distribution close to the focus position

$\mathbf{r} = 0 = \{0, 0, 0\}$ is given by (18, 19)

$$\mathbf{E}_{\text{exc}}(\mathbf{r}) = E_0 \int_0^{2\pi} d\phi \int_0^{\Theta} d\theta \frac{\sin\theta}{\sqrt{\cos\theta}} \left(\cos\phi \hat{\mathbf{e}}_\perp - \sin\phi \hat{\mathbf{e}}_\parallel\right) \exp\left[i\mathbf{k}\cdot\mathbf{r} + i\Psi(\theta,\phi)\right]. \tag{47}$$

which represents a superposition of plane waves with wave vectors $\mathbf{k}$. The meaning of the involved vectors $\mathbf{k}, \hat{\mathbf{e}}_\perp$ and $\hat{\mathbf{e}}_\parallel$ is visually explained in Figure 3(a). The factor $1/\sqrt{\cos\theta}$ ensures conservation of energy flux along the optical axis while the light is focused through the objective. Abbe's sine condition relates the distance r from a point on the incident planar wavefront to the angle θ at a corresponding point on the focused spherical wavefront through the relationship $r = f \sin\theta$. Here, f denotes the focal length of the lens focusing the wavefront. The upper integration limit Θ is the half-angle of light collection of the focusing objective, which is related to its numerical aperture via $\text{NA} = n \sin\Theta$.

The function $\Psi(r, \phi)$ describes wave front deformations of the incoming plane wave and/or ones induced by an imperfect focusing optics. These wavefront deformations result in aberrations of focus shape and alter the light intensity distribution across the focus. The function $\Psi(r, \phi)$ is efficiently represented by a series expansion in terms of Zernike functions:

$$\Psi(\theta,\phi) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} a_n^m Z_n^m \left(\frac{\sin\theta}{\sin\Theta}, \phi\right). \tag{48}$$

The coefficients a_n^m are the expansion coefficients that quantify the contribution of each Zernike function to the overall wavefront distortion. We now have everything in place to delve into the effects of various aberrations on the shape and size of the laser light focus. The lowest non-trivial Zernike functions, specifically $aZ_1^{\pm 1}$ with an amplitude factor a, represent linear gradients along the x- or y-axis. Incorporating these linear expressions into the exponent on the r.h.s. of (47) reveals that they induce a lateral displacement of the focal position by $2a/(k \sin\Theta)$, without modifying its shape. In a similar vein, the aberration denoted by aZ_2^0 leads to a first-order shift of the focus along the optical axis, quantified by $4\sqrt{3}a/\sin^2\Theta$. Hence, these initial Zernike functions effectively capture simple translational shifts of the focus, illustrating their superior utility in optical system descriptions as opposed to the orthogonal function system that involves Bessel functions. The latter, depicted in Figure 1, which parallel the Zernike expansion modes presented in Figure 2, convey a considerably more intricate physical interpretation. For instance, the primary three Bessel modes, corresponding to straightforward focus shifts in the Zernike framework, also impact the focus' shape, as they cannot be succinctly represented by simple first or second-order polynomials in the radial coordinate. In Figure 4, we examine the influence of the lowest significant Zernike modes on the shape of the focus. Specifically, we investigate modes that characterize first and second order astigmatism, coma, primary spherical aberration, and trefoil. In each instance, we have set the amplitude of the corresponding Zernike functions to unity. Astigmatism is typically caused by non-planar surfaces of reflective elements such as beam splitters, dichroic mirrors, or plane mirrors. Coma is most often introduced by later misalignment of lens elements with respect to the optical axis. Primary spherical aberration is primarily caused by the classical spherical aberration of spherical (non-parabolic) lenses, although high-quality optical systems are quite effective at compensating for these effects. Finally, trefoil aberrations can be induced by diffraction from obstructions within the optical pathway, such as the three-armed holders used for optical elements in mirror telescopes.

Confocal detection

In an epi-fluorescence confocal microscope, as typically used in conventional FCS, the fluorescence excited within the focused laser light is collected by the same objective and then focused through a confocal aperture before being detected by a point light detector. This is done to ensure efficient rejection of out-of-focus light and thus to generate a detection volume with its lateral site determined by the diameter of the focused laser light, and its axial size by the confocal detection.

Basically all common fluorescent dyes, proteins or quantum dots behave as ideal electric dipole emitters. Thus, for modeling the detection efficiency distribution of a confocal microscope, we consider the imaging of an electric dipole emitter in sample space through the objective into the plane of the confocal pinhole. The electric and magnetic field distributions at lateral position $\boldsymbol{\rho}$ in that plane generated by an electric dipole emitter (fluorescent molecule) with dipole moment amplitude $\mathbf{p}$ at position $\mathbf{r}$ in sample space are given by (see for example (20))

$$\begin{Bmatrix} \mathbf{E}_{\text{pinhole}}(\boldsymbol{\rho},\mathbf{r},\mathbf{p}) \\ \mathbf{B}_{\text{pinhole}}(\boldsymbol{\rho},\mathbf{r},\mathbf{p}) \end{Bmatrix} = \int_0^{2\pi} d\phi \int_0^{\Theta} d\theta \frac{\sin\theta}{w(\theta)} \sqrt{\frac{\cos\theta}{\cos\theta'}} \begin{Bmatrix} \hat{\mathbf{e}}'_\perp(\hat{\mathbf{e}}_\perp\cdot\mathbf{p}) + \hat{\mathbf{e}}_\parallel(\hat{\mathbf{e}}_\parallel\cdot\mathbf{p}) \\ -\hat{\mathbf{e}}_\parallel(\hat{\mathbf{e}}_\perp\cdot\mathbf{p}) + \hat{\mathbf{e}}'_\perp(\hat{\mathbf{e}}_\parallel\cdot\mathbf{p}) \end{Bmatrix} \exp\left(i\mathbf{k}'\cdot\boldsymbol{\rho} + i\mathbf{k}\cdot\mathbf{r}\right) \tag{49}$$

where the meaning of all vectors $\mathbf{k}, \mathbf{k}', \hat{\mathbf{e}}_\perp, \hat{\mathbf{e}}'_\perp$ and $\hat{\mathbf{e}}_\parallel$ is visually explained in Figure 3(b). The factor $\sqrt{\cos\theta/\cos\theta'}$ in (49) takes again care of conservation of the energy flux along the optical axis, and $w(\theta) = nk_0\cos\theta$ is the axial component of the wave

vector $\mathbf{k}$ in sample space. Next, the detection efficiency through the confocal pinhole for such an emitter is given by the integral of the time-averaged axial component of the Poynting vector, $\hat{\mathbf{e}}_z \cdot \mathbf{P} = (c/8\pi)\hat{\mathbf{e}}_z \cdot (\mathbf{E}_{\text{pinhole}} \times \mathbf{B}_{\text{pinhole}})$, over the pinhole area,

$$U_{\text{det}}(\mathbf{r}, \mathbf{p}) = \frac{c}{8\pi} \int_{\text{pinhole}} d^2\boldsymbol{\rho}\ \hat{\mathbf{e}}_z \cdot \left[\mathbf{E}_{\text{pinhole}}(\boldsymbol{\rho}, \mathbf{r}, \mathbf{p}) \times \mathbf{B}_{\text{pinhole}}(\boldsymbol{\rho}, \mathbf{r}, \mathbf{p})\right]. \tag{50}$$

Finally, for an isotropic distribution of molecule orientations, the detection efficiency is obtained by averaging the above expression over all orientations of the electric dipole vector $\mathbf{p}$, i.e. $U_{\text{det}}(\mathbf{r}) = \langle U_{\text{det}}(\mathbf{r}, \mathbf{p})\rangle_{\hat{\mathbf{p}}}$. For the sake of simplicity, when modeling an FCS measurement in the next subsection we assume that aberrations are only present in the focusing of the fluorescence excitation light, but not in the imaging from sample to image space. However, this is no general restriction–imaging aberrations could be treated by including a similar phase function $\Phi(\theta, \phi)$ into the exponent in (49) as was done in (47) when considering aberrations in excitation light focusing.

Fluorescence correlation spectroscopy

FCS is a single-molecule sensitive technique that measures and evaluates the intensity fluctuations of the fluorescence that is excited within the minuscule detection volume of a confocal microscope (1–3, 21). The autocorrelation function calculated from these intensity fluctuations reveals insights into any physical processes that influence the fluorescence intensity fluctuations within this volume. A primary application of FCS is to study the translational diffusion of fluorescent molecules through the detection volume, providing data on the diffusion coefficients of these molecules. However, the diffusion-related decay time of the autocorrelation function is notably sensitive to the size and shape of this detection volume (16, 17). Consequently, various FCS adaptations have been developed to offer robustness against potential optical aberrations that might occur in a confocal microscope setup. Among these methods, notable examples include dual-focus FCS (2fFCS) (22, 23), scanning FCS (24, 25), and z-scan FCS (26, 27). Two-photon excitation FCS (5–7) also promises to minimize impact from optical aberrations, thanks to the quadratic intensity dependence of fluorescence excitation. This dependency selectively emphasizes the contributions from the focus's high-intensity central region over its lower-intensity surrounding areas. In this study, we utilize Zernike functions to model optical aberrations incurred while focusing a laser beam through an objective and assess how these aberrations influence one-photon and two-photon excitation FCS experiments.

The excitation intensity distributions for one- and two-photon excitation are determined by the second and fourth absolute power of the electric field distribution described by (47). In case of one-photon excitation FCS, the excitation intensity distribution has to be multiplied with the detection efficiency distribution $U_{\text{det}}(\mathbf{r})$ of (50), i.e. the three-dimensional function that describes how well an emitting molecule is detected through the pinhole of the confocal microscope, to obtain the total molecule excitation and detection probability distribution. In case of two-photon excitation FCS, which does not require confocal detection (due to its intrinsic optical sectioning capability), this excitation and detection probability distribution is given by the square of the excitation intensity distribution directly.

With the excitation and detection probability distributions determined, the diffusion-related component of the autocorrelation function, $g(\tau)$, is defined as the average of the product of three probabilities: the probability of exciting and detecting a fluorescence photon emitted by a molecule initially at position $\mathbf{r}_1$, the probability that this molecule diffuses to position $\mathbf{r}_2$ within the lag time τ, and the probability of subsequently exciting and detecting another photon from the molecule at its new position. This concept is mathematically represented by a six-dimensional integral

$$\begin{aligned} g_{1\text{h}\nu}(\tau) &\propto \int d^3\mathbf{r}_2 \int d^3\mathbf{r}_1 U_{\text{det}}(\mathbf{r}_2)\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_2)|^2\, G_D(\mathbf{r}_2 - \mathbf{r}_1, \tau) U_{\text{det}}(\mathbf{r}_1)\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_1)|^2 \\ g_{2\text{h}\nu}(\tau) &\propto \int d^3\mathbf{r}_2 \int d^3\mathbf{r}_1\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_2)|^4\, G_D(\mathbf{r}_2 - \mathbf{r}_1, \tau)\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_1)|^4 \end{aligned} \tag{51}$$

where $G_D(\mathbf{r}_2 - \mathbf{r}_1, \tau)$ is Green's function for the free diffusion equation,

$$G_D(\mathbf{r}_2 - \mathbf{r}_1, \tau) = \frac{1}{(4\pi D\tau)^{3/2}} \exp\left[-\frac{(\mathbf{r}_2 - \mathbf{r}_1)^2}{4D\tau}\right] \tag{52}$$

with D being the diffusion coefficient of the diffusing fluorescent molecules.

For the numerical calculation of this autocorrelation function, we first determined the electric field distribution $\mathbf{E}_{\text{exc}}(\mathbf{r})$ and the detection efficiency distribution $U_{\text{det}}(\mathbf{r})$ across a grid of finite elements, each measuring $10 \times 10 \times 20$ nm^3, within a volume of $4 \times 4 \times 8$ µm^3. The calculations were performed with the following parameter values: one-photon excitation wave length of 532 nm, two-photon excitation wavelength of 1064 nm, water immersion objective with an NA of 1.2, aqueous sample

medium with refractive index of 1.33, magnification of imaging 60 times, and pinhole diameter of 86 µm (corresponding to the position of the second Airy ring and assuring that 97% of light from an emitter at the objective's focus that hits the plane of the pinhole is transmitted through the pinhole). Next, we computed the convolution $\int d^3\mathbf{r}_1 G_D(\mathbf{r}_2 - \mathbf{r}_1, \tau) U_{\text{det}}(\mathbf{r}_1)\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_1)|^2$ for the one-photon excitation FCS and $\int d^3\mathbf{r}_1 G_D(\mathbf{r}_2 - \mathbf{r}_1, \tau)\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_1)|^4$ for the two-photon excitation FCS in (51), respectively, employing Fast Fourier Transforms (FFTs) (28). This approach leverages the principle that the convolution of two functions in real space is proportional to the inverse Fourier transform of the product of their individual Fourier transforms. Finally, we achieved the second integration in (51) by aggregating the products of the convolution and U across all finite elements. Finally, we multiplied the results of these convolutions with either $U_{\text{det}}(\mathbf{r}_2)\, |\mathbf{E}_{\text{exc}}(\mathbf{r}_2)|^2$ for the one-photon excitation case or with $|\mathbf{E}_{\text{exc}}(\mathbf{r}_2)|^4$ for the two-photon excitation case and integrated the result over the coordinate $\mathbf{r}_2$. This yields the value of the corresponding autocorrelation function at lag time τ.

Figure 5 shows both the explicit autocorrelation curves for the ideal and different aberrated detection volumes and the resulting diffusion times defined by the lag time when the autocorrelation function has fallen to half of its maximum value. One observes that different types of aberrations change the shape of the FCS autocorrelation curve in different ways. For example, secondary astigmatism changes the curve shape in such a way that it is sufficiently distinct from that of an aberration-free experiment, which would help to discern it as coming from a measurement with substantial aberration. However, trefoil mostly shifts the correlation curve to longer times without much changing its shape, so that such a measurement could be easily confused with an aberration-free measurement of a slower diffusing sample. This is a principal problem of conventional FCS when using it for precise measurements of diffusion coefficients: It is very sensitive to optical aberrations, and its is hard to detect this by just inspecting FCS curve shape and resulting fit quality (16, 17). Comparison of the diffusion times with that of an ideal aberration-free measurement shows that, except for astigmatism, two-photon excitation FCS is slightly more robust against optical aberrations than one-photon excitation FCS. However, the difference between the performance of one- and two-photon excitation FCS is not as striking as one may hope for, emphasizing again the importance of more accurate FCS variants mentioned in the beginning of this section (22–27) when one wants to measure diffusion coefficients with high accuracy.

Another crucial parameter in an FCS measurement is the maximum brightness of a molecule, defined as the brightness measured for a molecule positioned at the center of the focus and on the optical axis. Figure 6(a,b) shows this value compared to that observed for an ideal system without aberrations. It is evident that aberrations have a much more dramatic impact on maximum brightness than on diffusion time, and this effect is significantly worse for two-photon excitation experiments compared to one-photon excitation experiments. This observation correlates well with the impact of aberrations on the detection volume, which will be discussed next.

The detection volume V_{det} is a crucial characteristic of an FCS experiment when estimating the concentration of diffusing molecules. Its product with the concentration is inversely proportional to the relative amplitude of an FCS curve, given by $\lim_{\tau\to 0} \frac{g(\tau)}{\bar{I}^2}$, where $\bar{I}$ is the average measured fluorescence intensity. Thus, by measuring the relative FCS amplitude and knowing the detection volume, one can calculate the absolute concentration of fluorescent molecules in a sample. Mathematically, V_{det} is defined by

$$V_{\text{det}} = \frac{\left(\int d\mathbf{r} U(\mathbf{r})\right)^2}{\int d\mathbf{r} U(\mathbf{r})^2} \tag{53}$$

where we have $U(\mathbf{r}) = U_{\text{det}}(\mathbf{r})\, |\mathbf{E}_{\text{exc}}(\mathbf{r})|^2$ for one-photon excitation FCS, and $U(\mathbf{r}) = |\mathbf{E}_{\text{exc}}(\mathbf{r})|^4$ for two-photon excitation FCS. The computational results for the detection volume are shown in Figure 6(c,d). As illustrated, the impact of aberrations on V_{det} is similarly severe to that on brightness, and it is surprisingly more pronounced for two-photon excitation than for one-photon excitation. This behavior was previously reported in Ref. (17), which also noted a much stronger impact of aberrations on FCS-based estimates of concentration compared to diffusion coefficients (for one-photon excitation FCS).

Finally, to characterize the similarity (or lack thereof) between autocorrelation curves measured with and without aberrations, we use the reduced logarithmic-time Pearson coefficient $1 - P_{\log\text{time}}$. Here, $P_{\log\text{time}}$ is defined as:

$$P_{\log\text{time}} = \frac{\int d(\ln\tau)\, g(\tau) g_{\text{ideal}}(\tau)}{\sqrt{\int d(\ln\tau)\, g(\tau)^2}\sqrt{\int d(\ln\tau)\, g_{\text{ideal}}(\tau)^2}}. \tag{54}$$

Calculating this coefficient with logarithmic time accounts for the rapid decay of an autocorrelation curve over time, as seen in the logarithmic plots in Figure 5(a,b), so most of its information is contained at short timescales. The reduced Pearson coefficient $1 - P_{\log\text{time}}$ will be exactly zero when comparing identical curves, and its absolute value increases with the difference between two curves. The computational results for this coefficient are shown in Figure 6(e,f). As illustrated, the difference

between aberrated and ideal autocorrelation curves, as measured by $1 - P_{\log \text{time}}$, is similar across all considered aberration types for one-photon excitation, with trefoil aberration showing the most significant impact. However, for two-photon excitation FCS, astigmatism has by far the largest impact, causing the strongest deviation of the autocorrelation curve from its ideal, aberration-free case. Please note the qualitatively similar trends in the impacts of aberrations on diffusion time and on the similarity between autocorrelation curves.

CONCLUSION

In this paper, we have presented a derivation of Zernike polynomials using a variational calculus approach and have derived the fundamental properties regarding orthogonality and normalization. On the way, we have elucidated the connection between Bessel functions and Zernike polynomials. Finally, we have applied Zernike functions for modeling the focusing of light through a high-aperture objective and then used this result for modeling the impact of different aberrations on one and two-photon excitation FCS measurements.

CODE AVAILABILITY

The MATLAB code for performing all model calculations described in this paper is freely available from the authors upon reasonable request.

AUTHOR CONTRIBUTIONS

Jörg Enderlein developed the general idea of the manuscript, wrote its core part, performed all numerical calculations and prepared all figures. Ivan Gligonov helped with editing the manuscript and checked all mathematical derivations and equations.

DECLARATION OF INTERESTS

The authors declare no competing interests.

ACKNOWLEDGMENTS

Jörg Enderlein acknowledges the financial support of the European Research Council (ERC) for financial support through project “smMIET” (grant agreement no. 884488) under the European Union’s Horizon 2020 research and innovation program, and financial support from the Deutsche Forschungsgemeinschaft through the German Excellence Strategy EXC 2067/1-390729940. Ivan Gligonov acknowledges funding by the International Max Planck Research School for Physics of Biological and Complex Systems and by the European Union via the HORIZON–MSCA–2022–DN “Improving BiomEdical diagnosis through LIGHT-based technologies and machine learning – BE-LIGHT” (GA n° 101119924 – BE-LIGHT).

REFERENCES

1. Magde, D., E. Elson, and W. W. Webb, 1972. Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. *Physical Review Letters* 29:705.

2. Magde, D., E. L. Elson, and W. W. Webb, 1974. Fluorescence correlation spectroscopy. II. An experimental realization. *Biopolymers: Original Research on Biomolecules* 13:29–61.

3. Magde, D., W. W. Webb, and E. L. Elson, 1978. Fluorescence correlation spectroscopy. III. Uniform translation and laminar flow. *Biopolymers: Original Research on Biomolecules* 17:361–376.

4. Denk, W., J. H. Strickler, and W. W. Webb, 1990. Two-photon laser scanning fluorescence microscopy. *Science* 248:73–76.

5. Schwille, P., U. Haupts, S. Maiti, and W. W. Webb, 1999. Molecular dynamics in living cells observed by fluorescence correlation spectroscopy with one-and two-photon excitation. *Biophysical Journal* 77:2251–2265.

6. Schwille, P., J. Korlach, and W. W. Webb, 1999. Fluorescence correlation spectroscopy with single-molecule sensitivity on cell and model membranes. *Cytometry: The Journal of the International Society for Analytical Cytology* 36:176–182.

7. Berland, K., P. So, and E. Gratton, 1995. Two-photon fluorescence correlation spectroscopy: method and application to the intracellular environment. *Biophysical Journal* 68:694–701.

8. Zernike, F., 1953. How I discovered phase contrast. *Science* 121:345–349.

9. Zernike, F., 1934. Beugungstheorie des Schneidenverfahrens und seiner verbesserten Form, der Phasenkontrastmethode. *Physica* 1:689–704.

10. Born, M., and E. Wolf, 1999. Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light. Cambridge University Press.

11. Booth, M. J., 2014. Adaptive optical microscopy: the ongoing quest for a perfect image. *Light: Science & Applications* 3:e165.

12. Goodwin, E. P., and J. C. Wyant, 2006. Field Guide to Interferometric Optical Testing. SPIE Press, Bellingham, WA.

13. Lakshminarayanan, V., and A. Fleck, 2011. Zernike polynomials: a guide. *Journal of Modern Optics* 58:545–561.

14. Tyson, R., and V. Lakshminarayanan, 2012. Adaptive optics. *Journal of Modern Optics* 59:1032–1033.

15. Wikipedia, 2023. Zernike polynomials. https://en.wikipedia.org/wiki/Zernike_polynomials.

16. Enderlein, J., I. Gregor, D. Patra, and J. Fitter, 2004. Art and Artefacts of Fluorescence Correlation Spectroscopy. *Current Pharmaceutical Biotechnology* 5:155–161.

17. Enderlein, J., I. Gregor, D. Patra, T. Dertinger, and U. B. Kaupp, 2005. Performance of Fluorescence Correlation Spectroscopy for Measuring Diffusion and Concentration. *ChemPhysChem* 6:2324–2336.

18. Wolf, E., 1959. Electromagnetic diffraction in optical systems-I. An integral representation of the image field. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 253:349–357.

19. B., R., W. E., and G. Dennis, 1959. Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 253:358–379.

20. Fazel, M., K. S. Grussmayer, B. Ferdman, A. Radenovic, Y. Shechtman, J. Enderlein, and S. Pressé, 2024. Fluorescence microscopy: a statistics-optics perspective. *Reviews of Modern Physics* 96:025003.

21. Webb, W. W., 1976. Applications of fluorescence correlation spectroscopy. *Quarterly Reviews of Biophysics* 9:49–68.

22. Dertinger, T., V. Pacheco, I. von der Hocht, R. Hartmann, I. Gregor, and J. Enderlein, 2007. Two-focus fluorescence correlation spectroscopy: A new tool for accurate and absolute diffusion measurements. *ChemPhysChem* 8:433–443.

23. Müller, C. B., A. Loman, V. Pacheco, F. Koberling, D. Willbold, W. Richtering, and J. Enderlein, 2008. Precise measurement of diffusion by multi-color dual-focus fluorescence correlation spectroscopy. *EPL (Europhysics Letters)* 83:46001.

24. Petrášek, Z., and P. Schwille, 2008. Precise measurement of diffusion coefficients using scanning fluorescence correlation spectroscopy. *Biophysical Journal* 94:1437–1448.

25. Petrášek, Z., S. Derenko, and P. Schwille, 2011. Circular scanning fluorescence correlation spectroscopy on membranes. *Optics Express* 19:25006–25021.

26. Benda, A., M. Beneš, V. Marecek, A. Lhotskỳ, W. T. Hermens, and M. Hof, 2003. How to determine diffusion coefficients in planar phospholipid systems by confocal fluorescence correlation spectroscopy. *Langmuir* 19:4120–4126.

27. Humpolíčková, J., E. Gielen, A. Benda, V. Fagulova, J. Vercammen, M. Hof, M. Ameloot, Y. Engelborghs, et al., 2006. Probing diffusion laws within cellular membranes by Z-scan fluorescence correlation spectroscopy. *Biophysical Journal* 91:L23–L25.

28. Press, W. H., W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, 2007. Numerical Recipes 3rd Edition: The Art of Scientific Computing. Cambridge University Press.

FIGURES

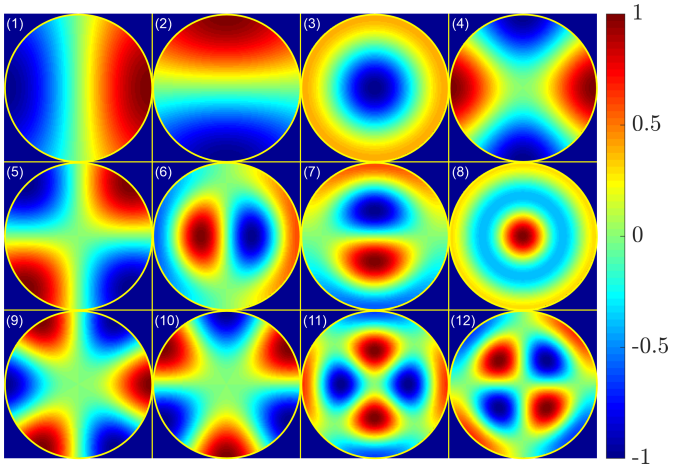

Figure 1: Density plots of the first low-order orthogonal functions $J_m(\lambda_n^m r)\exp(im\phi)$: (1) $J_1(\lambda_1^1 r)\cos\phi$; (2) $J_1(\lambda_1^1 r)\sin\phi$; (3) $J_0(\lambda_1^0 r)$; (4) $J_2(\lambda_1^2 r)\cos(2\phi)$; (5) $J_2(\lambda_1^2 r)\sin(2\phi)$; (6) $J_1(\lambda_2^2 r)\cos\phi$; (7) $J_1(\lambda_2^2 r)\cos\phi$; (8) $J_0(\lambda_2^2 r)$; (9) $J_3(\lambda_1^3 r)\cos(3\phi)$; (10) $J_3(\lambda_1^3 r)\sin(3\phi)$; (11) $J_2(\lambda_2^2 r)\cos(2\phi)$; and (12) $J_2(\lambda_2^2 r)\sin(2\phi)$. In the figures, all functions were normalized to their absolute maximum. The choice of the modes was done to make them most similar to the first canonical 12 Zernike aberrations, see Figure 2 and next section.

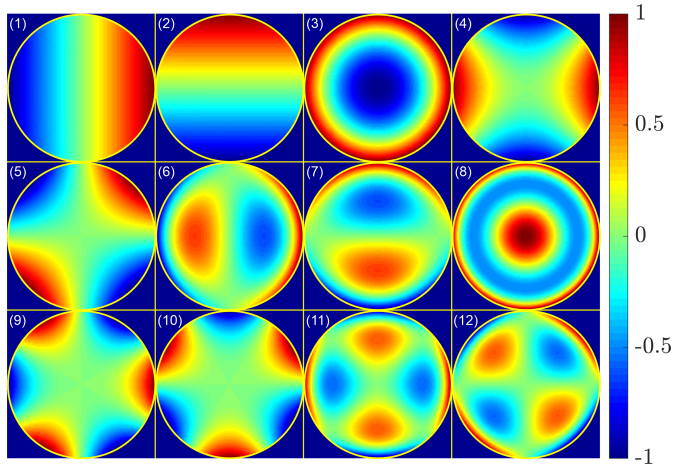

Figure 2: Density plots of the first twelve Zernike polynomials as presented in table 1: (1) horizontal or x tilt, $Z_1^1 = 2R_1^1(r)\cos\phi$; (2) vertical or y tilt, $Z_1^{-1} = 2R_1^1(r)\sin\phi$; (3) defocus, $Z_2^0 = \sqrt{3}R_2^0(r)$; (4) vertical astigmatism, $Z_2^2 = \sqrt{6}R_2^2(r)\cos(2\phi)$; (5) oblique astigmatism, $Z_2^{-2} = \sqrt{6}R_2^2(r)\sin(2\phi)$; (6) horizontal coma, $Z_3^1 = \sqrt{8}R_3^1(r)\cos\phi$; (7) vertical coma, $Z_3^{-1} = \sqrt{8}R_3^1(r)\sin\phi$; (8) primary spherical aberration, $Z_4 = \sqrt{5}R_4^0(r)$; (9) oblique trefoil, $Z_3^3 = \sqrt{8}R_3^3(r)\cos(3\phi)$; (10) vertical trefoil, $Z_3^{-3} = \sqrt{8}R_3^3(r)\sin(3\phi)$; (11) vertical secondary astigmatism, $Z_4^2 = \sqrt{10}R_4^2(r)\cos(2\phi)$; and (12) oblique secondary astigmatism, $Z_4^{-2} = \sqrt{10}R_4^2(r)\sin(2\phi)$.

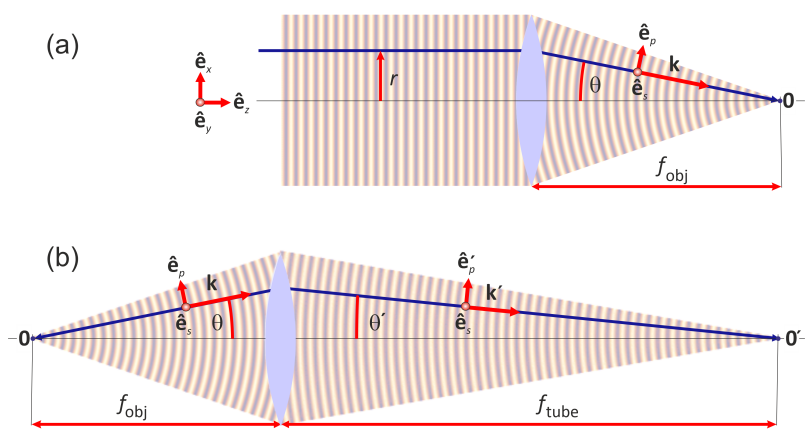

Figure 3: Excitation and detection in a FCS experiment. (a) Focusing of a linearly polarized plane wave (polarization along vector $\hat{\mathbf{e}}_x$) through an objective lens with focal length f_{obj}. The electric field close to the focal point 0 can be represented as a superposition of plane waves traveling along all possible wave vectors $\mathbf{k} = nk_0\{-\sin\theta\cos\phi, -\sin\theta\sin\phi, \cos\theta\}$, where n is the refractive index of the medium where the light is focused, and $k_0 = 2\pi/\lambda$ is the length of the vacuum wave vector at wavelength λ. The polarization of each plane wave can be divided into two components along the unit vectors $\hat{\mathbf{e}}_p = \{\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta\}$ and $\hat{\mathbf{e}}_s = \{-\sin\phi, \cos\phi, 0\}$. (b) Wave optics of imaging a point source (electric dipole emitter) in sample space (left hand side) into the plane of the confocal pinhole (right hand side). Each plane wave component with wave vector $\mathbf{k} = nk_0\{\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta\}$ in sample space is converted into a plane wave with wave vector $\mathbf{k}' = k_0\{-\sin\theta'\cos\phi, -\sin\theta'\sin\phi, \cos\theta'\}$ in image space. The relation between θ and θ' is given by Abbe's sine condition, $n\sin\theta = M\sin\theta'$, where $M = f_{\text{tube}}/f_{\text{obj}}$ is the imaging magnification. While the unit vector for the s-wave $\hat{\mathbf{e}}_s = \{-\sin\phi, \cos\phi, 0\}$ is the same in the sample and the image space, the unit vector for the p-wave in sample space is $\hat{\mathbf{e}}_p = \{\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta\}$ and in image space is $\hat{\mathbf{e}}'_p = \{\cos\theta'\cos\phi, \cos\theta'\sin\phi, \sin\theta'\}$.

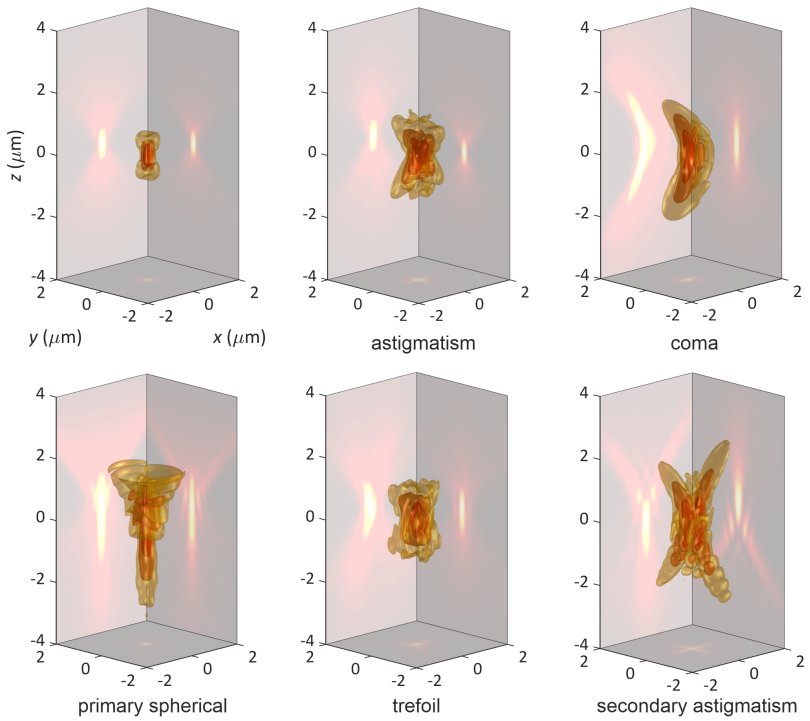

Figure 4: Iso-surface plots of $U(\mathbf{r})_{\text{exc}} = |\mathbf{E}(\mathbf{r})|^2$ (shown are the surfaces where the intensity U_{exc} has fallen to $1/e$, $1/e^2$, and $1/e^3$ of its maximum value in the center) and maximum intensity projections of U for different types of aberrations. Calculations were done for focusing an x-polarized plane wave with 532 nm wavelength through a water immersion objective with a numerical aperture of NA = 1.2. The top left panel shows the ideal aberration-free case. All other panels show the effect of different aberrations as indicated below each panel, and all axes values are in μm as shown in the first panel.

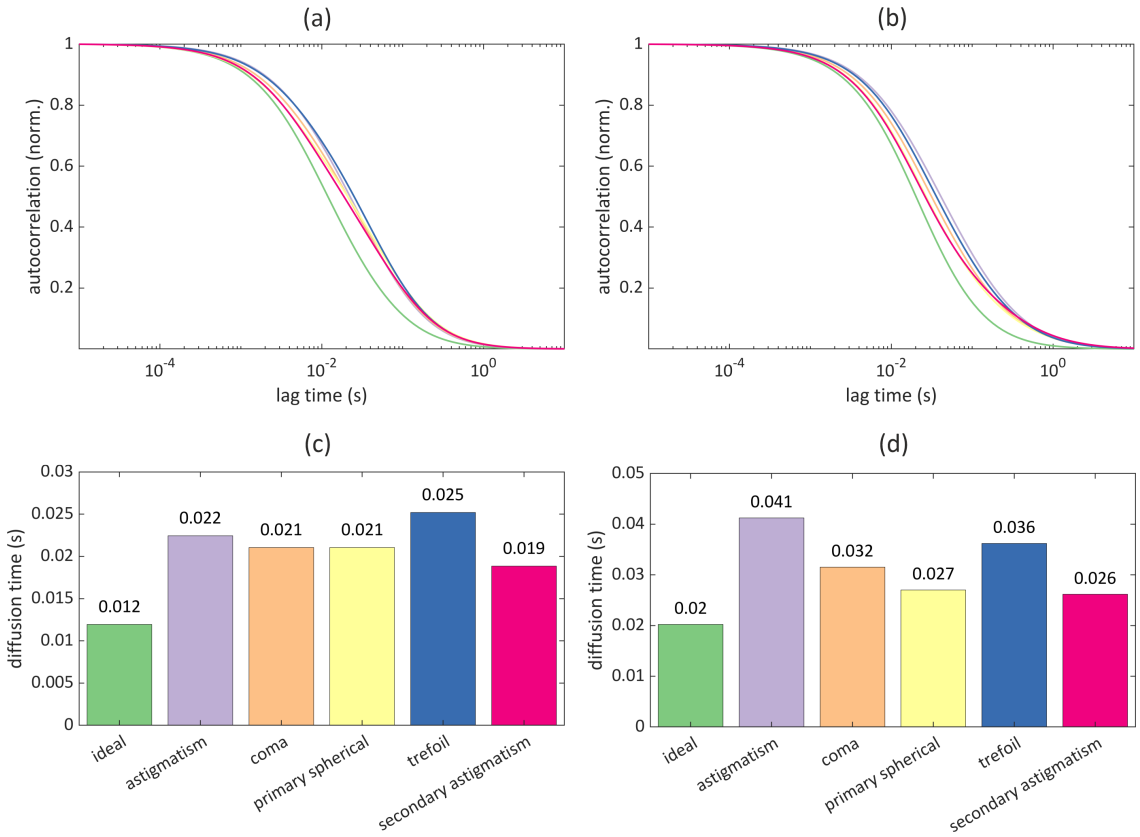

Figure 5: Fluorescence Correlation Spectroscopy (FCS) with aberrations. (a) Modeling of one-photon excitation FCS curves for pure diffusion through the ideal and the five aberrated detection volumes corresponding to the intensity distribution functions shown in Figure 4. The FCS curve for the ideal aberration-free detection volume is the leftmost curve. (b) Same as panel (a), but for two-photon excitation FCS. (c) Diffusion times (lag time when the autocorrelation has fallen off to half its maximum value) corresponding to the autocorrelation curves shown in panel (a). The colors of the bars correspond to the lines of the FCS curve plots, values above each bar denote the diffusion time in seconds. (d) Same as panel (c), but for two-photon excitation FCS.

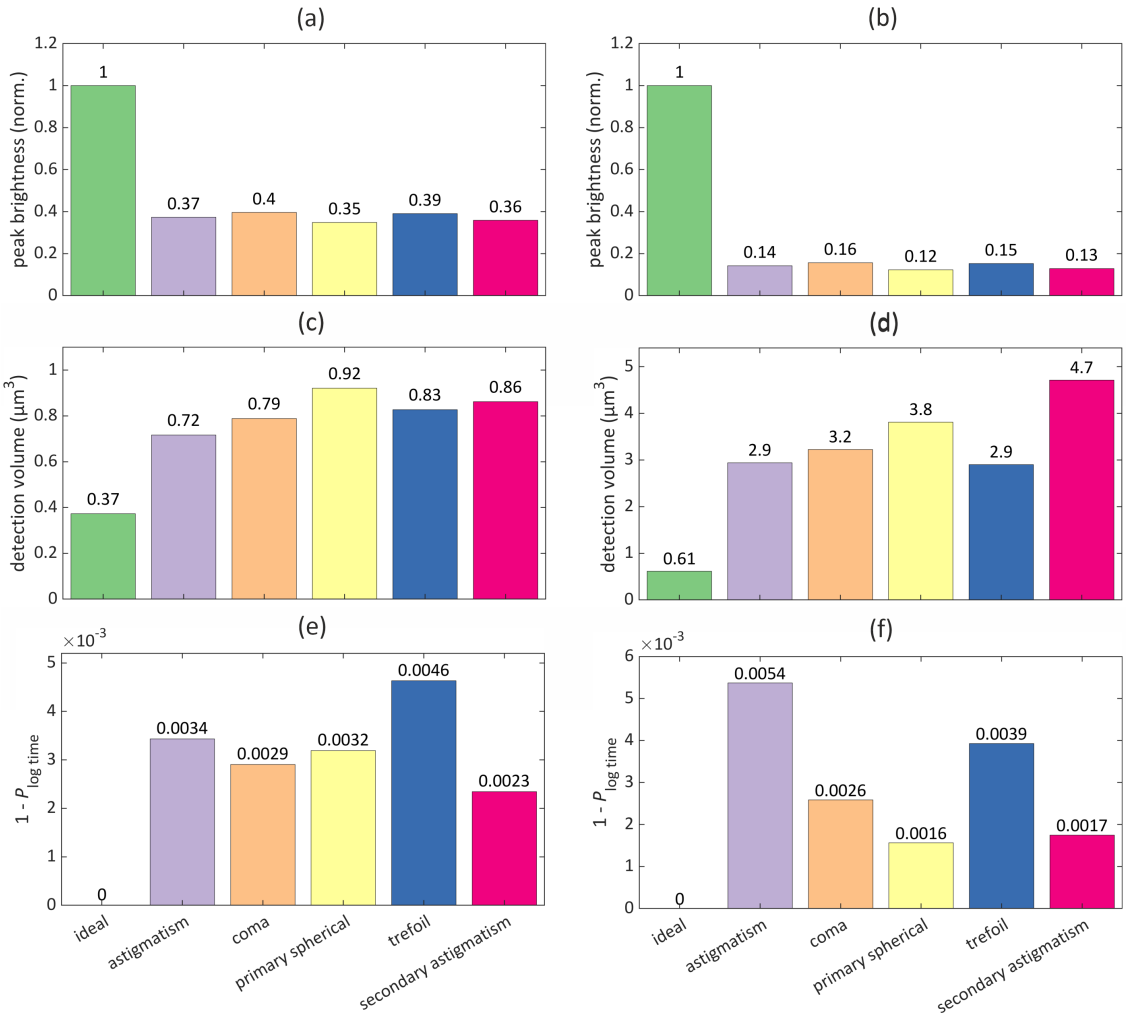

Figure 6: Brightness, detection volume, and curve similarity. Left panels show the result for 1-photon excitation FCS, right panels for two-photon excitation FCS. All model parameters are the same as used for Figure 5. (a) and (b) show the brightness that would be measured from a molecule in the focal plane and on the optical axis, as compared to a molecule a the same position in a measurement without any aberrations. (c) and (d) Effective detection volumes, see eq. (53). (e) and (f) Reduced logarithmic-time Pearson coefficients, see eq. (54), which are a measure of the similarity between autocorrelation curves measured with and without aberrations.