5. Immobilized molecules

5.1. Introduction, Signal-to-noise ratio

If the single molecules are immobilized in a solid or in a highly viscous sample, or at the surface of a solid, it becomes possible to study the same molecule for long periods of time. In FCS correlation experiments on liquid solutions, the statistics over a large number of molecules always makes it possible to find a signal, even in difficult conditions with strong background and noise. For molecules immobilized in solids or on surfaces, however, it is of crucial importance to be able to « recognize » the molecule, i.e. to distinguish its signal from background and noise.

The following discussion of the signal/noise ratio applies to microscopic images as well as to spectra of single molecules, which can be considered as images in the spectral domain (see the second part of this course on cryogenic experiments). Let us consider (see Fig. 5.1) the background B and the signal S of the molecule as functions of a scanning parameter (which can be one position coordinate of the sample, or the frequency of the exciting laser). If t is the acquisition time per channel or pixel, the number of background counts is Bt. Assuming shot noise to be dominant (which is usually the case for weak signals), the noise from the background is $\sqrt{Bt}$, which has to be compared to the signal St. The most favorable case to detect the molecule is when its signal occupies about one pixel or one channel width (as a broad structure is more difficult to distinguish from noise than a sharp one, we bin all data in one width to reduce statistical fluctuations). The molecule will be detectable if and only if the signal-to-noise ratio is significantly larger than unity. Let us choose a factor of, say, 3: $St > 3\sqrt{Bt}$, which gives $S > 3\sqrt{B/t}$, or $t > 10 \times B/S^2$.

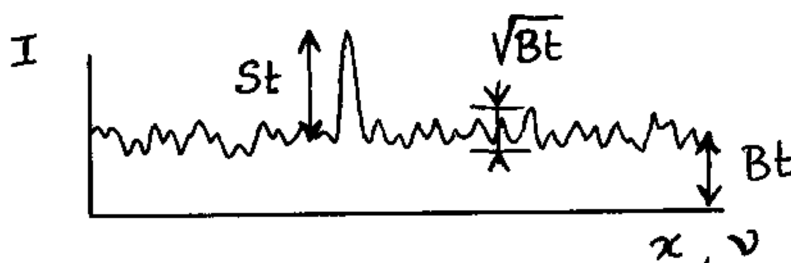

Figure 5.1 : Illustration of the spatial or spectral signal of a single molecule, to be detected against noise due to background fluctuations.

For a typical background of 100 cps arising from the dark counts of an avalanche detector, and for an acquisition time of 1 s, it follows that the signal has to be larger than 30 cps to be clearly distinguished from noise. This condition becomes only a *lower bound* if the molecule's image is spread over several pixels, or if noise from other sources adds to photon noise.

5.2. Sample preparation

In fluorescence experiments, the main source of background is fluorescence from residual impurities or from optical elements. Raman scattering is often negligible. To observe single molecules, it is therefore crucial to reduce contamination by fluorescent impurities, and to work with very clean optical parts. The excited volume should be as small as possible, and the substrate should be non-fluorescent. Fused silica, or very pure glass are often used as substrates and coverslips. Hereafter, we review a few usual methods to prepare samples for single-molecule fluorescence.
i) spin-coating: quick spinning of a flat substrate with a thin layer of polymer solution leads to a uniform film, draining slower and slower as time goes (see Fig. 5.2). At the same time, evaporation decreases the thickness too, and increases the concentration and the viscosity. In a first phase, viscous draining is dominant in reducing the thickness (see Exc. 5.1), then evaporation takes over and becomes dominant. The balance between viscous draining and evaporation leads to films of tunable thickness.

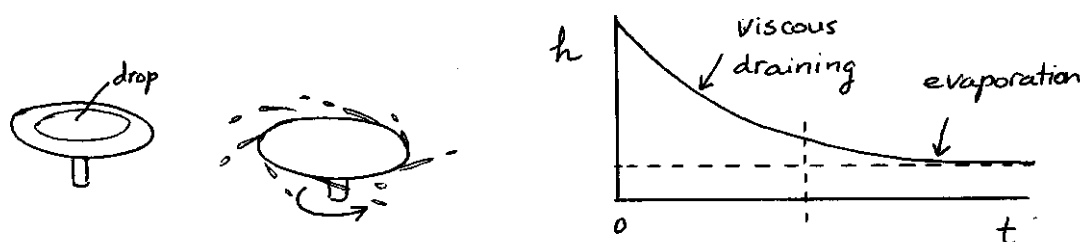

Figure 5.2 : Spin coating of a polymer solution. A drop is deposited on the flat substrate, which is then spun at high speed, resulting in draining of excess liquid and gradual thinning of the film. The thinning process is driven by viscous flow at the beginning of the operation, then by evaporation of the solvent in its last stage.

Very importantly, because of a particular solution of this hydrodynamic problem, the thickness of the resulting film is constant over the whole area of the substrate. This is the main reason why such films are used as photoresists in the semiconductor

industry. Typical spin-coating parameters (1000 rpm, 1 % weight of polymer in such solvents as water for hydrophilic polymers, or toluene for hydrophobic ones) lead to thicknesses between 10 nm and a few microns. Spin-coating is a very simple and efficient way to reduce the illuminated volume, and therefore the background from impurities, via a reduction of the film thickness.

ii) Langmuir-Blodgett films: even thinner, monomolecular films are obtained by the Langmuir-Blodgett technique, i.e. the deposition onto a substrate from a monomolecular layer at the air-water interface. The dye molecules have to be introduced at low concentration in the spreading solution of the amphiphilic molecules (Fig. 5.3).

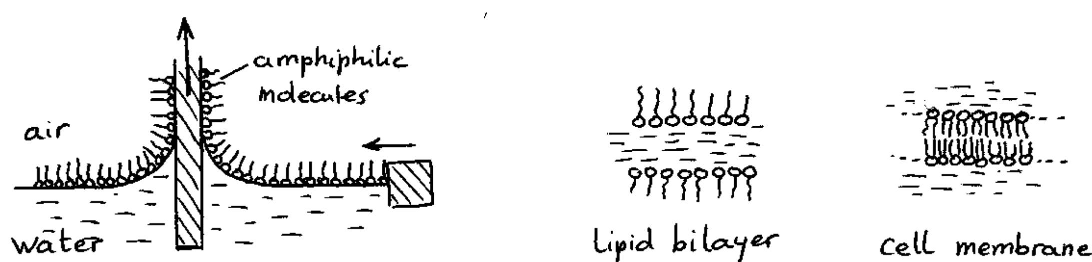

Figure 5.3 : Deposition of a Langmuir-Blodgett film onto a solid substrate. The scheme shows deposition on a hydrophilic substrate. Deposition on a hydrophobic substrate is performed by dipping it into the air-water interface. Schematic structures of bilayers that can be used as matrices or anchors for single molecules.

iii) bilayers, membranes or black lipid films are particular cases of a pair of monomolecular layers, which can be either deposited on a substrate, or suspended across a hole in a thin substrate. Cell membranes are particularly important (Fig. 5.3) examples of bilayers in which single molecules can be imaged and tracked.

iv) other preparation methods: molecules can be deposited directly on a substrate (in that case they are more sensitive to oxygen and water), included in nanocrystals and other nanoparticles, or simply selected or imaged within a 3D sample, such as a cell, for example. In the latter case, a thin sample should be preferred to limit background.

5.3. Microscopy images

Much information can sometimes be obtained from images with single-molecule spots. Counting them can provide concentrations, their intensity provides the stoichiometry of complexes, etc. For example, the number of labelled ligands in

complexes can be deduced from the intensity of the fluorescence signal (Schmidt et al., *Anal. Chem.* **68** (1996) 4397). Another important application is co-localization : two biomolecules can be labeled with two different fluorophores. If the two labels occur at the same points in an image, it means that the biomolecules are probably associated and/or interacting: they are colocalized (G. J. Schütz et al., *Biophys. J.* **74** (1998) 2223). Such evidence for colocalization of two molecules is very similar to the one we mentioned for the cross-correlation FCCS, and is very useful in molecular biology.

5.4. Orientation

The orientation of the *in-plane component* of the molecular transition moment can be obtained easily by polarization measurements. Two different methods are used dominantly :

i) polarization analysis in the detection: the fluorescence beam is sent to a polarizing beam splitter and the two beams are sent to different detectors. The intensity ratio gives two possible solutions for the in-plane orientation. Discriminating between those requires a third measurement. In combination with polarized excitation, polarized analysis gives the fluorescence anisotropy, related to the angular mobility and rotational diffusion of the fluorophore during fluorescence.

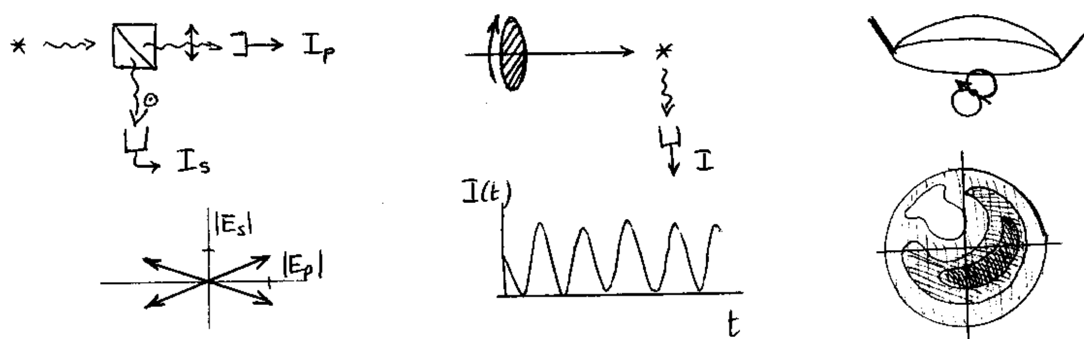

Figure 5.4 : Polarization measurements to determine the orientation of the transition dipole moment of single molecules. Left : detection by two orthogonally polarized channels. A third measurement is required to eliminate one of the two possible orientations. Center : excitation with a rotating linearly polarized wave. The phase of the sinusoidal intensity trace provides the molecular orientation. The contrast depends on the rotational diffusion of the fluorophore. Right : the angular distribution of emitted photons also gives the dipole orientation. This intensity distribution can be obtained in the focal plane by slightly defocusing the molecule's image on the CCD or detector.

ii) polarization modulation: the polarization of the *exciting* beam can be modulated in time, for example with a rotating polarizer. The phase and amplitude of the sinusoidal variation of the detected intensity gives information about the orientation and slow rotational diffusion of the fluorophore.
With the latter method, various modulation depths of the intensity can be found (see work by S. Weiss' lab). For a fixed molecule, the modulation is sinusoidal with contrast unity. For a fast isotropic diffusion, no modulation is found. For cases in between, Weiss et al. have proposed models of hindered rotators. For example, rotational diffusion can be isotropic within a cone with a given orientation and half-angle. With these two parameters, different contrasts and phases of the intensity oscillations can be fitted.
The rotational mobility of fluorophores was used to probe nano-environments close to the glass transition (Deschenes and Vanden Bout, *J. Phys. Chem. B* **106** (2002) 11438). These authors found that single molecules diffuse as Brownian particles on short timescales, but that different molecules diffuse at different rates, and that diffusion can change in time. This inhomogeneity gives rise to the complex (non-exponential) behavior of the glass just above the glass transition. Single-molecule experiments confirm the spatial inhomogeneity of glass dynamics.

As discussed in the section on optical microscopy, the full 3D orientation of the dipole moment can be obtained with a widely opened excitation beam, or by measuring the distribution of fluorescence photons as a function of the angles of emission. Alternatively, the image can be slightly defocused, providing a characteristic pattern depending on the 3D orientation of each molecule. These latter methods require a matrix of detectors for fast determination. The pattern can then be compared to the emission pattern of a dipole (Fig. 5.4, right).

5.5. Blinking

When the fluorescence intensity of a nano-object is recorded as a function of time (in a so-called fluorescence intensity trace), random variations of the average emission intensity are often seen. The variations can be progressive (such fluctuations are often called *flickering*), but often present sudden jumps between bright and dark states. We discuss only the latter behavior here, which is commonly known as *blinking* or

intermittency. Blinking is a characteristic feature of the emission of single nano-objects and has been observed for dye molecules, polymer molecules, color centers, defects in insulators or semiconductors, quantum dots, nanocrystals, etc.. In large populations of these objects, it is almost always hidden because the fluctuations of individual objects are usually not synchronized. In most experiments on large ensembles, an external control parameter is required to synchronize all blinking objects. An obvious example is the dynamics following changes in excitation intensity: either switching the exciting laser on, or illuminating the sample by short excitation pulses. In those cases, blinking will appear as a transient variation of the average signal. It is one of the most powerful features of single-molecule methods that no synchronization is needed to observe fluctuations directly. Blinking, as many other complex dynamical processes, can be observed on single objects under continuous-wave excitation.

Several possible mechanisms can lead to blinking. Most of them involve a 'dark' state of the emitter, i.e. a state which does not fluoresce, either because it does not absorb efficiently, or because its fluorescence yield is too low. Hereafter, we examine a few possible sources of blinking :

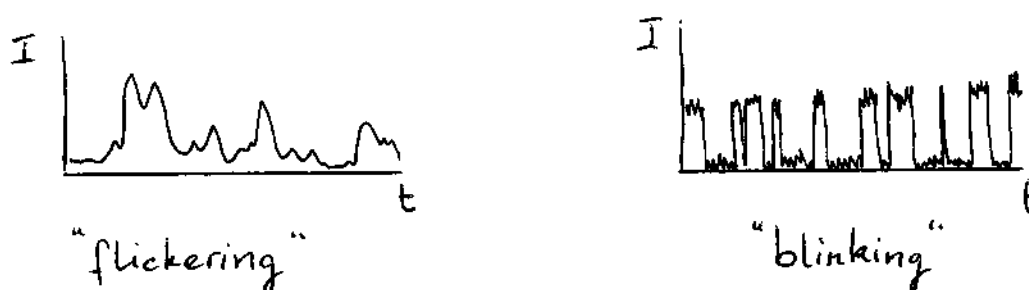

Figure 5.5 : Fluctuations of the fluorescence intensity as a function of time. The first case (left) is that of a flickering molecule, with continuous fluctuations. In the second case (right), sudden intensity jumps give the trace the look of a random telegraph signal. This is blinking.

i) triplet state: once in its triplet state, a molecule can in theory absorb and emit photons. Usually, however, the fluorescence yield of the triplet-to-triplet fluorescence is low, or its spectrum is shifted away from the singlet-to-singlet fluorescence, and the laser wavelength is not adapted to excite those triplet-to-triplet transitions. This means that the triplet state is usually dark. The fluorescence trace presents dark periods lasting for the triplet lifetime on average. Depending on the molecule, this time can be as short as microseconds and as long as tens of milliseconds for usual dyes.

Atmospheric oxygen has a strong influence on triplet lifetime. Since O_2 has a triplet ground state (it is paramagnetic), it can exchange electrons with a molecular triplet, yielding two singlet states, one being the ground state of the dye, the other one being singlet oxygen, a very reactive and phosphorescent (at 1270 nm or about 1 eV) species. Singlet oxygen can itself further react with the dye molecule and lead to its chemical degradation to a non-fluorescent product (thereby bleaching the dye, see below).

ii) electron transfer: the excited molecule may accept an electron from its environment, or give one electron to it. The radical-ion left usually has shifted spectra and/or very low fluorescence yield. The charge-transfer state lives for as long as the electron is away, giving rise to dark states with lifetimes longer than seconds if the electron goes far away. This process has been characterized in our group for rhodamine 6G in polyvinylalcohol (PVA), and appears to be rather general. More recent work in solution by the groups of Sauer and Tinnefeld has shown that the charge transfer rates can be controlled by adding redox species to the solution. This process, known as ROXS method (Vogelsang et al. Angew. Chem. Intl. Ed. 2008 47, 5465), can reduce the lifetime of the triplet state and shorten the lifetime of the charged dark states. The reduced lifetime of metastable states has the added benefit that bleaching (see next paragraph) is also significantly reduced. Controlling bright and dark times in bleaching enables the d-STORM scheme of super-resolution (see below).

iii) other reversible photochemical reactions: The excited molecule may change conformation. If the change is reversible, and if the two forms have different optical fluorescence properties, blinking will follow. An example of intramolecular reactions of this kind is cis-trans isomerization, which is common in cyanine dyes. The excited molecule may also react with its environment, for example by abstracting a proton from a nearby molecule or from an acid matrix. Several dyes, for example methylene blue, present protonated "leuco" forms, in which an additional proton shortens the conjugation path and shifts the absorption spectrum to the blue. The main feature of these photochemical reactions involving large rearrangements of atoms is that they often require large activation energies and are therefore frozen at low enough temperatures.

iv) other sources of blinking may exist in principle. For example, the orientation of the molecule could switch between different orientations with respect to the excitation

polarization. The absorption spectrum of the molecule could also shift between different spectral positions. This spectral diffusion is a prominent effect in low-temperature experiments (see later chapters of the course), and has also been observed at room temperature in polymers (Lu and Xie, *Nature* **385** (1997) 143). However, it is probable that in most cases, such « physical » processes are not the ones leading to significant blinking, i.e., blinking presenting large intensity fluctuations.

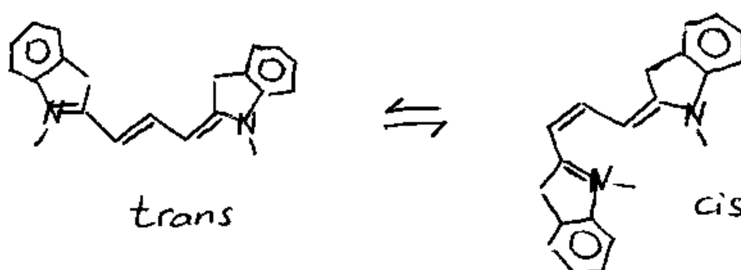

Figure 5.6 : Cis-trans isomerization of a cyanine molecule. Fluorescence is much stronger in the trans conformation. The chemical equilibrium between the two forms leads to blinking.

5.6. Bleaching

Many photochemical reactions are irreversible. In that case, the molecule stays in the dark state the first time it gets there. Fluorescence is lost for good, and a new single molecule must be found. All fluorescent organic molecules photobleach sooner or later at room temperature. Much effort has been spent to screen, adapt and design dyes able to resist photobleaching as far as possible. This has been crucial in lasers, but also for fluorescent labelling in biology. In spite of several decades of efforts, little is known about photobleaching and how to reduce it. Atmospheric oxygen and small reactive molecules such as water obviously open efficient channels for photobleaching. An efficient strategy is to lower the HOMO-LUMO gap of the molecule, so that its triplet state doesn't excite singlet oxygen any more. This is the proposed mechanism by which terrylene-di-imide resists bleaching very efficiently even under atmospheric oxygen. Many single-molecule fluorescence studies are done in the presence of reducing agents such as mercapto-ethanol, Trolox®, and other oxygen scavengers. The ROXS scheme (see above), by shortening the lifetimes of metastable and reactive states, also improves the resistance of dyes to bleaching. However, even in inert and dry atmosphere, or under high vacuum, photobleaching still occurs, albeit at much lower rates. Because photobleaching involves chemical

reactions, low temperatures considerably delay and reduce the process. Temperature studies of rhodamine in PVA have shown that, indeed, bleaching is reduced by several orders of magnitude at low temperatures for some molecules, but not for all of them. Recent articles report dyes of the perylene- or terrylene-diimide family, which present excellent resistance to bleaching, and still can be made water soluble by convenient chemical functionalization.

5.7. Superresolution

5.7.1 Single molecules (PALM, STORM, PAINT,…)

We start with a brief historical overview. It was realized from the very beginning of single-molecule observations that the center of the imaging spot of a single molecule can be found with much higher accuracy than the width of the point-spread function. If N photons are detected, *and if we know that they all arise from the same single molecule,* the position of their center of gravity is determined down to roughly $\Delta x/\sqrt{N}$, where Δx is the spatial width of the point spread function. To see why, look at the probability distribution of the sum of N variables distributed as normal Gaussian variables. This sum is distributed according to a convolution of N Gaussians with width $\Gamma=\sqrt{N\gamma^2}$, therefore the width of the probability distribution of the average is $\frac{\Gamma}{N}=\frac{\gamma}{\sqrt{N}}$.

In low-temperature experiments (see part II of this course), single molecules can be discriminated by their resonance frequency, and the selection of different molecules by the laser frequency may give rise to a super-resolved image (van Oijen et al., 1998). Because many photons can be accumulated for each single molecule, one can reach an accuracy of a few nanometers in locating the molecules, exciting each molecule in turn by changing the laser frequency. This can been done on a sample of millions of molecules in a crystal (Naumov et al., *Angew. Chem. Intl. Ed.* **48** (2009) 9747). This spectral selection does not work at room temperature because the absorption bands of the molecules are very broad and overlap one another.

In 2006, three important papers appeared almost simultaneously, demonstrating superresolution based on photochemical switching of single molecules. These methods are actually one and the same, but as they were demonstrated with different

photochemical reactions, they were given different names. The original references are: PALM by E. Betzig's group [Betzig et al. *Science* **313** (2006) 1642], STORM by X.-W. Zhuang [Rust et al. *Nat. Meth.* **3** (2006) 793], and a third paper [ST Hess et al., *Biophys. J.* **91** (2006) 4258] presents the same idea under the acronym FPALM. The central idea of super-resolution imaging with photo-switches is to randomly select a small sub-population of all fluorescent molecules present in the sample by switching them to a bright state, for example by a short illumination with an auxiliary laser. The concentration of these activated switches is kept low enough that one can extract a maximum number of photons of every one of them *individually*, until they go to a dark state or are irreversibly bleached. Because the molecules are still well separated from each other, one can find the center of each individual spot very accurately. Therefore, the number of switched molecules after every photo-activation cycle has to be small enough, so that their centers can be located with a much better precision than the resolution of the microscope, down to a few nm for the brightest molecules, which provide many photons. A similar scheme has been proposed by Sauer and Tinnefeld, and is known as direct STORM or dSTORM. It makes use of the random blinking of single molecules under controlled redox conditions by means of the ROXS scheme to tune the on- and off-times. There, the transitions between dark and bright states are not induced by an auxiliary laser, they occur spontaneously upon irradiation by the excitation laser. The transition rates can be tuned by properly choosing the concentration of oxidizing and reducing agents. All photons detected during a bright period are assigned to the same molecule, which is then located with super-resolution accuracy. The on- and off-times of the fluorophores and the density of labels can be tuned to optimize resolution.

Another method uses the transient binding of fluorescent molecules to a surface or to the object to be imaged. Only the transiently immobilized molecules give rise to measurable spots, whose centers determine with super-resolution the sources' locations. The diffusing molecules only contribute background. This super-resolution method based on transient binding of fluorophores to fixed targets, for example complementary short DNA strands, is called PAINT (Point Accumulation for Imaging in Nanoscale Topography).

Not all super-resolution methods are based on the detection and identification of single molecules. For example, SOFI (Super-resolution Optical Fluctuation Imaging, proposed by J. Enderlein) is based on spatial maps of the contrast of a correlation

function. With correlation functions of N-th order, these maps exhibit a resolution improvement by $\sqrt{N}$.

5.7.2 Point-spread function engineering (STED, RESOLFT)

A completely different manner to obtain super-resolution is based on nonlinear optics and on saturation. It was invented in the mid-1990's and demonstrated by S. Hell, and has recently been implemented in commercial devices. It makes use of the strong nonlinear character of saturation and of its behavior in the vicinity of a zero of the optical intensity. A good example of such a process is stimulated emission depletion (STED), which brings excited molecules back to the ground state when they are illuminated at the proper wavelength, usually a strong line or band resonant with one of the transitions in the dye's fluorescence spectrum. The STED beam is engineered in such a way that it presents a node (zero of electric field) at the center of its PSF, and that this dark spot is completely surrounded by bright areas. The PSF of the STED beam has a donut shape. By increasing the intensity of the STED beam, one makes the remaining fluorescent spot at the center of the donut (in which the molecules have not been brought back to their ground state by the STED beam) smaller and smaller, in principle down to arbitrarily small sizes, in practice easily down to a few tens of nm in size. This latter method does not rely on any single-molecule identification, only on photophysics and on the nonlinearity of the saturation process. Other processes than STED, for example triplet saturation, or photochemistry, and in particular photochemical switching, can also be used for this sort of super-resolution imaging, which has been called RESOLFT by Hell and his group.

5.7.3 Combination of single-molecule and STED illumination, MINFLUX

The two methods for super-resolution described above are completely independent. Therefore, there is no problem to apply them independently. In this way, a narrow STED PSF can be still improved further if we exploit the knowldege that N fluorescence photons all stem from the same single molecule. In this way, localization accuracies as high as 1 nm can be reached. S. Hell has even designed a method, called MINFLUX (Balzarotti et al. *Science* **355** (2017) 606), to minimize the number of photons required to locate each molecule. This scheme makes use of the information

provided by the *absence of detection* of a molecule, once the molecule's presence within the donut has been ascertained.

In summary, we have two main different routes towards super-resolution. Both approaches to super-resolution rely on a photochemical switching mechanism. Both of them also rely on a strongly nonlinear process. This nonlinearity is obvious in the case of saturation methods (STED), but it is more subtle in the case of single-molecule methods. There, it relies on the *correlation* of all photons observed during a bright period. Because we assign them to the same molecule, we know they were issued from the same point in space. In the case of STED, one can show that the spot size is reduced compared to Abbe's limit by a factor depending on the ratio of the intensity to a saturation intensity of a nonlinear process:

$$\Delta x_{eff} = \frac{\lambda}{2 \times NA} \frac{1}{\sqrt{1 + I / I_{sat}}} \ .$$

In the case of the single-molecule imaging methods, the factor I / I_{sat} is replaced by the number of photons detected per molecule. The same type of relation can be generalized to the case of single molecules, because the detection of a large number, N, of photons from the same molecule amounts to a correlation (a nonlinear process) of order N. A similar correlation is achieved in the SOFI method on small ensembles of molecules. Although each individual photon in this process can be obtained at very low excitation intensity, the accumulation of all photons with the *knowledge that they all arise from the same molecule* at the same position in space (arising from the clever use of another, switching wavelength, or from additional knowledge about blinking) amounts to probing a high-order susceptibility. The 2014 Nobel Prize in Chemistry was awarded to E. Betzig, S. Hell and W. E. Moerner for the development of super-resolution microscopy methods.

Exercise 5.1: What volume of a mM *solution of arachidic acid ($C_{22}H_{45}COOH$) in chloroform should be spread on a water surface to create a monolayer of 300* cm^2 *in area? (The cross sectional area of a saturated chain is about 20* Å^2*).*

Exercise 5.2: Write the probability distribution of the sum of two variables distributed as normal Gaussian variables. This function is also a Gaussian whose width is the square root of the sum of square widths. From this relation, deduce that the distribution of the average of N variables, all with the same Gaussian distribution, has a Gaussian distribution with width $\sqrt{N}$ *times smaller than that of each individual variable.*

Exercise 5.3: A molecule is illuminated by an excitation laser with intensity I (absorption cross section σ*) and with a STED laser with intensity* I_S *(cross section for stimulated emission* Σ*). The fluorescence lifetime is* τ_f*. The STED laser is resonant with a transition from the excited state to a vibrational state, whose lifetime is supposed to be very short.*

i) Write a system of differential equations describing the population of the ground and excited state of the molecule under illumination by the two lasers.

ii) What is the excited state population in steady state for constant intensities?

iii) We assume that the excited state has been prepared at time zero and that a STED pulse of duration T *is applied at time zero. What is the probability of finding the molecule in the excited state after the STED pulse?*

Exercise 5.4: Consider spin coating of a Newtonian liquid (density ρ*, viscosity* η*, rotation velocity* ω*) on a flat, infinite rotating substrate and neglect evaporation of the solvent. Equate the shear viscosity force and the centrifugal force (neglecting the Coriolis force for a slow enough flow) to obtain the following relation in polar coordinates for the velocity* v *of the flow:*

$$-\eta \frac{\partial^2 v}{\partial z^2} = \rho\omega^2 r \ .$$

i) Solve this equation with the flow boundary conditions $\frac{\partial v}{\partial z}=0$ *for* $z=h$ *(the film thickness), and* $v\left(z=0\right)=0$. *Deduce the total amount of fluid flow* q *per unit length perpendicular to the flow by integration to find:*

$$q=\frac{\rho\omega^2 r h^3}{3\eta} \ .$$

ii) Use the conservation equation to deduce the following differential equation for the thickness:

$$\frac{\partial h}{\partial t}=-K\frac{1}{r}\frac{\partial}{\partial r}\left(r^2 h^3\right) \ , \text{ with } K=\frac{\rho\omega^2}{3\eta} \ .$$

iii) In the case of an initially uniform film, find the law of thinning as a function of time:

$$\frac{dh}{dt}=-2Kh^3 \ ,$$

and solve this equation to deduce the film thickness as a function of time.