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 the 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 (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 (a broad structure is more difficult to distinguish from noise than a sharp one). 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:

\[
St > 3\sqrt{Bt}, \quad \text{which gives} \quad S > 3\sqrt{B/t}, \quad \text{or} \quad t > 10 \times B/S^2.
\]

![Figure 5.1: Illustration of the spatial or spectral signal of a single molecule, to be detected against noise due to background fluctuations.](image)

For a typical background of 100 cps arising from the dark counts of the avalanche detector, and for an acquisition time of 1 s, it follows that the signal has to be larger than...
than 30 cps to be clearly distinguished from noise. This condition becomes only a lower bound if the molecule’s image is spread over more than one pixel.

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 must be as small as possible, and the substrate should be non-fluorescent (fused silica, or very pure glass). Hereafter, we cite a few possible methods to prepare samples.

i) spin-coating: quick spinning of a flat substrate with a thin layer of 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. The balance between viscous draining and evaporation leads to films of tunable thickness, constant over the whole area of the substrate. Such films are used as photoresists in the semiconductor industry. Typical spin-coating parameters (1000 rpm, 1 % weight of polymer in a solvent like water) lead to thicknesses between 10 nm and a few microns. Spin-coating is very interesting to reduce the illuminated volume, and therefore the background, via the 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, then by evaporation of the solvent.

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. Schematic structures of current bilayers used as matrices for single molecules.](image)

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 to a hole. Cell membranes are particularly important (Fig. 5.3).

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, etc., 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, 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 associated and interact: they are colocalized (G. J. Schütz et al., Biophys. J. 74 (1998) 2223). This argument is very similar to the one we mentioned for the cross-correlation, 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.

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 diffusion of the fluorophore.

With the latter method, various modulation depths of the intensity were found (Weiss and coll.). For a fixed molecule, the modulation is sinusoidal with contrast 1. 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 (L. A. Deschenes and D. A. Vanden Bout, J. Phys. Chem. B 106
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 (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. In large populations, it is almost always hidden because the fluctuations of individual objects are not synchronized. In most experiments on large ensembles, an external control parameter is required to synchronize all blinking objects (an obvious example is switching the exciting laser on or applying a short excitation pulse). 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, for example in blinking.

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:

i) triplet state: the molecule in the triplet state can in theory absorb and emit photons, but usually, the fluorescence yield is low, the spectrum is shifted from the singlet-
singlet fluorescence, and the laser wavelength is not adapted to excited triplet-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\textsubscript{2} has a triplet ground state (it is paramagnetic), it can exchange electrons with a molecular triplet, yielding two singlet states, one the ground state of the dye, the other one singlet oxygen, a very reactive and phosphorescent (1260 nm) species. Singlet oxygen can itself react with the dye molecule and lead to its chemical degradation to non-fluorescent products (bleaching, see below).

![Figure 5.5](image)

*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.*

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 very low fluorescence yield. The charge-transfer state lives 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, 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.

iii) other reversible photochemical reactions: The excited molecule may change conformation. If the change is reversible, and if the two forms have different optical
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 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 temperatures.

iv) other sources of blinking may exist in principle. For example, the orientation of 
the molecule could switch between parallel and perpendicular 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 second part 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, with 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 goes 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. Therefore, many single-molecule fluorescence studies are done in the presence of reducing agents (oxygen scavengers, mercapto-ethanol, trolox(R), etc.). 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 a lower rate. Because photobleaching involves chemical reactions, low temperatures considerably decrease 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 substitution.

5.7. Superresolution

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 accuracy much higher 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

\[
\frac{\Delta x}{\sqrt{N}},
\]

where \( \Delta x \) 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 \( \sqrt{2N\gamma} \), 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 superresolved image (van Oijen et al., 2000). Because many photons can be accumulated for each single molecule, one can reach an accuracy of a few nanometers in locating the molecules. This has been done recently on a sample of millions of molecules in a crystal (Naumov et al., Angew. Chem. Intl. Ed. 2009 48, 9747).
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 they were proposed with different photochemical reactions and 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 superresolution 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 an auxiliary laser. The concentration of these activated switches is 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 more recently (Sauer and Tinnefeld, method known as direct STORM or dSTORM), which makes use of the blinking of single molecules and 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 superresolution 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 the surface or object to be imaged. Only the immobile molecules give rise to a measurable spot, whose center determines with superresolution the source's location. This variant is called PAINT.

A completely different manner to obtain superresolution is based on nonlinear optics and on saturation. It was invented and demonstrated by S. Hell in the mid-1990’s, and has recently been implemented in a commercial device. 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 in the fluorescence spectrum). By increasing the intensity of the STED beam, one makes the remaining fluorescent spot (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 superresolution imaging.

In summary, we have two main different routes towards superresolution. Both approaches to superresolution 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 molecules. 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:

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

In the case of the single-molecule imaging methods, the factor $I / I_{\text{sat}}$ 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 $N$. Although this process is obtained at very low excitation intensity for each individual photon, 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 superresolution microscopy methods.

**Exercise 5.1:** Consider spin coating of a Newtonian liquid (density $\rho$, viscosity $\eta$, rotation velocity $\omega$) on a flat, infinite rotating substrate and neglect evaporation of the solvent. Equate the shear viscosity force and the centrifugal force (neglect Coriolis force for a slow flow) to obtain the following relation in polar coordinates for the velocity $v$ of the flow:

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

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

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

Use the conservation equation to deduce the differential equation for the thickness:

$$\frac{\partial h}{\partial t} = -K \frac{1}{r} \frac{\partial}{\partial r} \left( r^2 h^3 \right) ,$$

with $K = \frac{\rho \omega^2}{3\eta}$.

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.

**Exercise 5.2:** 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 5.3:** 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. Deduce from this that the
distribution of the average of $N$ variables has a width $\sqrt{N}$ times smaller than that of each individual variable.