Diffraction-unlimited optical microscopy

Optical microscopy, and fluorescence microscopy in particular, has emerged as one of the most powerful and convenient microscopic tools available today. This power does come at a price, however, in terms of a limited spatial resolution: traditionally fluorescence microscopy has been limited by diffraction to a resolution of a few hundred nanometers, far too large to discern nanostructuring in biological or material samples. Recent conceptual advances have emerged that challenge this once-thought ‘unbreakable’ barrier, and fluorescence microscopy with nanometer resolution is now within reach. In this review we highlight some of the approaches that have made this paradigm shift possible.

Peter Dedecker, Johan Hofkens, and Jun-ichi Hotta
Department of Chemistry, Katholieke Universiteit Leuven, Celestijnenlaan 200F, B-3001 Heverlee, Belgium
E-mail: junichi.hotta@chem.kuleuven.be

One of the great paradoxes of life is that in order to truly understand the everyday issues and concepts that affect and fascinate us, one invariably has to focus on the microscopic world around us. This idea, while no longer particularly surprising, was realized as soon as the first lenses and microscopes were developed, and has only become stronger as our ability to observe smaller and smaller details improved. While this idea is usually brought up in the context of the life sciences, it is no different for the materials of today and tomorrow. Designing new materials tailored to specific needs and properties requires both study and manipulation of their micro- and nanoscale structuring, and microscopes are invaluable tools in modern-day research.

Not all approaches to visualize small details are equally useful for a given sample or problem, however, and a wide range of microscopy techniques have been developed. Each of these has its share of strengths and weaknesses, but one of the most convenient is optical microscopy, which, in an ideal case, requires only the illumination of the sample. Because of this, optical microscopy tends to be relatively noninvasive, nondestructive, and convenient to use. Fluorescence microscopy in particular has proven to be one of the most powerful microscopy tools available today, by combining the advantages of optical microscopy with the selectivity and extreme sensitivity of fluorescence emission. Indeed, fluorescence microscopy allows for measurements down to the level of individual molecules, and has been successfully applied in the study of, for example, polymer dynamics, catalysts, and dendrimers.

Nevertheless, revealing nanoscale structuring and ordering remains one of the great challenges of science. Indeed, one aspect that is true for every microscopy technique is that there is a lower limit to the sizes of the details that can discerned, that is, the spatial resolution is
Fig. 1 Point spreading: after imaging, an object of negligible size (such as a single molecule) will show up as a distribution of much bigger dimensions. The centroids of these distributions are indicated by the green and blue markers. If the emitters are spaced close together, as is the case for the molecules in the blue circle, then they cannot be distinguished and the structural information is lost. (Adapted from 83. © Wiley-VCH Verlag GmbH & Co. KGaA. Reproduced with permission.)

limited. Intuitively, if the sample structure becomes too small, then it cannot be revealed. There are several possible reasons why this is so, ranging from technical aspects, such as instrument imperfections, vibrations, and measurement noise, to limitations imposed by the fundamental laws of nature. While the former limitations are, to some extent, specific to a particular instrument or experiment, the latter present limiting barriers that are hard to overcome.

In the case of optical microscopy, this fundamental limitation stems directly from the wave-like character of light: no matter what kind of optical components we use, we cannot focus the light waves to an infinitely small point, only to a finite region. This might seem surprising at first: when we collect the fluorescence emission from a single, isolated molecule and image it through a lens system (such as a microscope), the resulting image will not be a point but rather a three-dimensional shape of much larger dimensions (hundreds of nanometers; see Fig. 1). The same holds true for the excitation light: even if the light source is infinitely small, we can only focus the light to a diffraction-limited region and cannot confine it further. This is known as point spreading, and is an immutable law of nature.

Point spreading is problematic if we want to image a sample that contains fluorescent molecules in close proximity, such as in a finely structured sample. In this case the emission distributions from closely spaced fluorophores overlap in the resulting fluorescence image, making it impossible to distinguish between molecules that are close together (Fig. 1). Hence point spreading imposes a fundamental limitation on the spatial resolution that can be attained. (We will not go into further detail here; more extended, yet accessible, discussions have been published elsewhere[17–19].)

While imperfections in the imaging system worsen the point spreading beyond the theoretical minimum, the optics in today’s microscopes and objectives have advanced to the point where the level of aberrations is so low that their contribution is negligible. Therefore little gain in resolution can be expected from further improvements in the quality of the microscope optics. While it is possible to modify the optical system in such a way that the point spreading is strongly reduced, e.g. by including a second objective[20,21], these approaches tend to be limited by the fact that the instrument design and alignment become increasingly complex.

It thus appears that the spatial resolution in optical microscopy is limited to a few hundred nanometers, and that details at smaller distances are fundamentally out of reach. Fortunately it turns out that it is possible to obtain an imaging resolution that is fundamentally unlimited by diffraction, and a variety of techniques that aim to achieve this have recently appeared. Intuitively these techniques work by making fluorescence emission into a rare event: if neither the excitation illumination nor the fluorescence emission can be confined further, then we have to make sure that only a small fraction of the molecules emits fluorescence even though all of them may experience a similar intensity excitation. Alternatively, if we process the images mathematically, then we can invert the problem, and try to minimize the number of molecules that are not emitting beyond what is normally allowed by diffraction. It turns out that the answer to this problem lies in the properties of the fluorescent molecules themselves, and in the fact that fluorescence is a dynamic phenomenon[22–24].

Nonlinearity

When we think about fluorescence emission, we tend to think about it as a fairly well-behaved, constant phenomenon. If we take a fluorescent molecule and irradiate it with excitation light, we expect to see a continuous, and predictable, fluorescence emission. Moreover,
Fig. 2 The emission of a single molecule of a fluorescent protein, clearly showing complex fluorescence dynamics, and the resulting nonlinearity in the fluorescence emission. The inset shows an expansion of the emission trace.

if we double the excitation intensity then we expect to see a twofold increase in fluorescence emission, and so on. This seems intuitive, because it agrees with the observation of emission from a dye-loaded solution in a fluorescence cuvette.

This apparent proportionality, however, only holds at the ensemble (averaged) level. One of the most fascinating findings from single-molecule experiments is that fluorescence emission is in fact a very dynamic phenomenon: individual molecules are seen to rapidly cycle between emissive and nonemissive states, displaying a highly complex and fascinating dynamic behavior [25-29] (Fig. 2). The origin of this on-off ‘blinking’ behavior is not always clear, though it can sometimes be attributed to, for example, the formation of triplet states, photoisomerization, or electron transfer. It is also known to be virtually ubiquitous, being observed in a wide variety of systems, including organic molecules [23-27,29], semiconductor nanocrystals [28], and fluorescent proteins [30].

Apart from transient nonfluorescent or ‘dark state’ formation, interruptions in the fluorescence emission can arise for other reasons as well, including saturation (the fact that there is an upper limit for the emission rate due to the finite excited-state lifetime), and phenomena such as photoswitching or photoactivation [31-41].

The net result is that the fluorescence emission is no longer strictly proportional to the excitation intensity. For a given sample and experimental setup the emission can depend on time, on the excitation intensity or intensities, or, by applying a spatial intensity distribution, on the position of the molecule with respect to the microscope’s focus.

Every approach to achieve a diffraction-unlimited resolution makes use of this fact in some way. It is possible to roughly divide these approaches based on whether they make use of nonlinearities in space (that is, try to exploit the nonlinearity to increase the spatial confinement of the fluorescence emission) or time (that is, try to separate the emission of the different labels in time). In what follows we will discuss each of these approaches, and mention some possible advantages and disadvantages of each.

**Nonlinearity in space**

For a more detailed example of how nonlinearities can be put to work, let us look at the process of fluorescence emission itself. The initial absorption of an excitation photon causes the molecule to enter an electronically excited state, where it remains for (usually) a few nanoseconds, before returning to the ground state through the emission of a fluorescence photon. The central issue here is that there is always some nonzero delay between the absorption and emission events. In other words, the emission rate can only increase up to the point where the molecule is constantly in the excited state, when it reaches a plateau. At this point, increasing the excitation intensity does not increase the emission rate any further, and the fluorescence emission is said to be saturated (and strongly nonlinear).

Fig. 3 shows an example of this. In this figure a cross-section of the fluorescence emission rate along the focus of a confocal fluorescence microscope is shown, at different excitation powers. The shape of the diffraction-limited excitation intensity distribution is shown by the red curve in the figure, and does not change throughout the experiment.
However, as the excitation power is increased, the saturation of the emission drastically alters the resulting emission distribution. The most crucial aspect of Fig. 3 is that it demonstrates that saturation enables us to create emission distributions with sharp edges and strongly confined non emissive regions, even though these were not allowed in linear systems by diffraction.

This simple example shows how using fluorescence nonlinearities can allow us to bypass the limitations of diffraction.

Saturation of the emission rate is not the only possibility, however, as any transition to a state with different emissive properties can be used as long as it is saturable (that is, its light-induced formation is much faster than its recovery) \[22\]. Approaches that make use of this principle to create 'sharp' distributions in emission are known as RESOLFT microscopy, standing for 'reversible, saturable optically linear fluorescence transitions' \[22\].

By itself, the example shown in Fig. 3 is not directly useful for attaining a higher imaging resolution, since the fluorescence confinement is not improved but, rather, worsened compared to the 'standard' diffraction-limited case. Hence we must either modify the imaging to use the saturation in a different way or process the images mathematically. A number of techniques have been developed that attempt both approaches, and we will expand on these in what follows.

**STED microscopy**

While the focusing of the excitation light will always lead to some spatial distribution, the exact shape of this distribution can be modified into different 'modes' by manipulating the relative phases of the input laser beam. Especially interesting are 'donut modes', which are characterized by a sharp intensity zero at the maximum, surrounded by a region of intense irradiation \[45-48\] (Fig. 4). Only at the exact center of the microscope focus is the intensity truly — mathematically — zero, while there is some irradiation intensity everywhere else in the focal volume.

Let us now assume that we have shaped the output from a laser into a donut mode, and that the wavelength and duration of the light is chosen so that it can deactivate the dye molecules. This deactivation could be because the light induces the formation of a nonfluorescent state (e.g. triplet or charge-transfer states, fluorescence photoswitching), or because it induces stimulated emission. Moreover, due to the intensity distribution of the donut shape, this deactivation occurs everywhere in the focal volume except at the center of the focus. This approach becomes even more interesting when it is combined with saturation (Fig. 5). In this case, it is clear that the confinement of the molecules in the fluorescent state can be made arbitrarily small, limited only by the power and duration of the donut-mode beam.

At present the most successful technique that uses this approach is 'stimulated-emission depletion' (STED) microscopy \[45-50\]. The principle underlying this approach is shown in Fig. 6, and operates as follows: in the first step of the acquisition, the sample molecules in the diffraction-limited spot are excited by a normal mode (Airy disk) laser pulse with a duration of picoseconds or less. This is followed immediately (within a few hundred picoseconds) by a second donut-mode laser pulse irradiating the same region, which induces stimulated
emission in the majority of the dye molecules, except those at the center of the focus (due to the intensity distribution of the donut mode).

The sample molecules that are still in the excited state are then left to fluoresce normally, which is detected. Because of the depletion, however, the fluorescence collection is effectively confined to a spot with sub-diffraction dimensions (Fig. 5b), leading to an increased resolution. This measurement cycle is then repeated at the next point on the sample, until a complete image is obtained. In this way, a spatial resolution down to a few tens of nanometers has been reported. As mentioned above, any saturable transition to a nonfluorescent state can be used, including triplet-state formation and fluorescence photoswitching, the sequence of the pulses likely has to be interchanged, and an additional laser pulse may be necessary.

The most attractive features of donut-based microscopy are that the effective resolution increase is completely dictated by the experimental setup and the irradiation powers used. Moreover, the imaging is altered directly, and obtaining a subdiffraction resolution does not require any additional processing. In addition, if stimulated emission is used, then the image acquisition times can be as fast as any laser-scanning microscope.

As indicated, the effective resolution increase that can be obtained

---

**Fig. 5** The effect of saturation using a donut-mode beam. (a) By saturating the transition to a nonfluorescent state, the spot of molecules that remain in the bright state can be made arbitrarily small (the relative excitation powers are 1, 5, 20, 100, and 1000, respectively). (b) The resulting apparent spot size.
depends on the number of irradiation photons in the dump beam; more irradiation leads to a greater depletion and tighter confinement, and hence an improvement in resolution. While this is a conceptual advantage, as the diffraction-limit is effectively broken, it also means that rather high levels of irradiation are required, especially when the depletion has to occur within a very short time window (as is the case when using stimulated emission). This can be mitigated somewhat by making use of longer-lived nonfluorescent states\textsuperscript{51-53}, but photobleaching or photodestruction is likely to remain a limiting factor in techniques that rely on saturation to create sharp emission distributions. Nevertheless, a wide range of fluorescent dyes, including fluorescent proteins\textsuperscript{57,58} and ‘atto’ dyes\textsuperscript{58}, have been found to be stable enough to produce impressive results\textsuperscript{59,60}.

**Nonlinear structured illumination**

As already discussed, saturation of the emission rate can be used directly to achieve a diffraction-unlimited resolution, though it requires processing of the raw images to discern the underlying sample structure. One way to do this is by nonlinear structured illumination microscopy\textsuperscript{61}.

Structured illumination relies on projecting the excitation light as a standing wave along the focal plane of the microscope, with a sufficiently high intensity to saturate the emission (though other saturable transitions can also be used). A cross-section of the fluorescence emission along the standing wave pattern looks similar to Fig. 5a, with ‘fluorescence deactivation’ replaced by ‘fluorescence emission’. By recording a set of images of the same sample, but with shifted orientations of the pattern, a diffraction-unlimited picture can be reconstructed (Fig. 7). Moreover, it is possible to extend the underlying principle to three dimensions, as has recently been demonstrated\textsuperscript{61-64}.

Nonlinear structured illumination is fairly similar to STED microscopy in that the measurement time is theoretically independent of the labeling density and the photostability of the sample determines the performance. They differ, however, in that STED is based on confocal microscopy instead of widefield microscopy, and that image
Fig. 7 Examples of STED and nonlinear structured illumination. (Left) A comparison between a standard confocal image and a STED image of 20 nm fluorescent beads. (Right) A widefield image of 50 nm fluorescent beads. (a, b) Conventional and filtered image; (c) linear structured illumination; (d) nonlinear structured illumination. (Structured illumination data reproduced with permission from[60]. © National Academy of Sciences, USA.)

processing is essential for structured illumination but not required for STED.

The exact mechanism by which the resolution is increased, and the required imaging processing is usually expressed in terms of Fourier components and spatial frequencies. A detailed discussion of this is beyond the scope of this paper, but has been published elsewhere[60].

In a sense, the two approaches that we have introduced here both make use of steady-state saturation: the high-resolution information is only obtained after the system has reached saturated conditions. While this makes for an intuitively clear picture, it is not essential. We will not consider ‘dynamic saturation optical microscopy’ in detail here, but two possible approaches have been discussed elsewhere[63,64].

**Nonlinearity in time**

The problem of limited spatial resolution can be approached in a

![Diagram](image_url)

Fig. 8 The principle behind PALM imaging: by repeatedly activating, localizing, and bleaching, individual emitters a high-resolution picture of the sample can be reconstructed. (a) A simulated conventional image of the sample. (b) The PALM imaging. (c) The reconstructed picture.
different way: let us assume that we somehow know that only a single emitting molecule is present in a given diffraction-limited spot. We can easily pinpoint the exact position of this molecule with nanometer precision, simply by looking at the centroid of the emission distribution\textsuperscript{65,66} (indicated by the markers in Fig. 1). If the density of the labeling molecules is sufficiently low, then this approximation is likely valid, though in that case it would be impossible to obtain any detailed structural information on the sample.

If we are interested in subdiffraction microscopy, then our sample should contain many fluorophores within nanometers of each other. Hence, the only way to apply this method would be to separate the emission of the molecules in time, so that at any given moment only a single molecule within a given diffraction-limited spot would emit. Then, by locating and mapping out the positions of all of the molecules in time, a high-resolution image can be constructed, effectively bypassing the diffraction limit.

Such a separation would be possible if one could precisely control the distribution of the molecules between the fluorescent and nonfluorescent states over an extended period. Until recently this task would have been considered virtually impossible, especially for large numbers of dye molecules. Now, however, a number of new fluorescent dyes have become available that display highly efficient fluorescence photoswitching or photoactivation\textsuperscript{31-41}. Photoactivation means that the molecule starts out in a thermally stable nonfluorescent state but can be converted to a fluorescent state through irradiation with light of an appropriate wavelength, where it remains until photobleaching. Photoswitching, on the other hand, means that the molecule can exist in both a thermally stable fluorescent state and a nonfluorescent state, and can reversibly interconvert between these states through irradiation with light of an appropriate wavelength. By using these properties, precise control over emission of individual fluorescent labels is now possible. However, other processes can be used as well, such as diffusion and docking\textsuperscript{67}.

Several related techniques have appeared that make use of photoswitching or photoactivation in combination with localization, and are commonly known as 'photoactivation localization microscopy' (PALM)\textsuperscript{68}, 'fluorescence photoactivation localization microscopy' (FPALM)\textsuperscript{69}, or 'stochastic optical reconstruction microscopy' (STORM)\textsuperscript{70}. For convenience, we will use these techniques as PALM microscopy. Briefly, the principle behind PALM, shown in Fig. 8, is as follows: a given sample is densely labeled with photoactivatable or photoswitchable dyes, which are initially in the nonfluorescent state. The sample is then weakly irradiated with a small amount of activation light, so that only a very small fraction of the molecules are activated. These activated molecules are then imaged until they deactivate again, either through photoswitching or photobleaching, and their locations are precisely determined. Next, a new set of molecules is activated, and the cycle is repeated until the position of (nearly) all molecules has been determined. Because the activation of the fluorescence is a stochastic process, different molecules are activated in each cycle, and the probability that two activated molecules will be present within a given diffraction-limited spot can be neglected. Hence, by mapping out all the localized positions, a high-resolution image of the sample can be reconstructed (Fig. 9).

PALM is a fascinating technique in that it does not attempt to modify the optical capabilities of the instrument in any way. Instead, it sidesteps the issue completely by ensuring that situations where there are two emitters within the same diffraction-limited spot are avoided. By this technique, localization precision to within a few tens of nanometers has been demonstrated\textsuperscript{68,71}. The advantages of this technique are that it is relatively simple, as no special equipment is required, and fairly gentle, in the sense that only low irradiation intensities are required. Also, as some of the most promising photoswitchable and photoactivatable fluorescent markers are genetically encodable fluorescent proteins, this approach is especially useful for biological samples.

However, the technique is not without its disadvantages: at present, the resolution along the optical axis is still fairly limited, and the requirement for a large number of activation–localization–bleaching cycles means that a fairly long measurement time (several minutes or more) is required. Nevertheless, attempts to extend the imaging to three dimensions\textsuperscript{71,72} and to use different spectrally shifted labels simultaneously\textsuperscript{73,74}, as well as to increase the speed of the acquisition\textsuperscript{75-77}, are currently underway.
Outlook and conclusions

It does not happen often that a barrier that has stood for over a hundred years is suddenly broken, but that is precisely what is happening today with fluorescence microscopy. Over the course of just a few years, several techniques have appeared that tackle or avoid this limitation in different ways, thereby allowing fluorescence microscopy to cross a familiar frontier into a previously inaccessible world.

While the field of diffraction-unlimited fluorescence microscopy is still very young, it has evolved and continues to evolve at a very rapid pace. The validity of the different approaches presented here has been demonstrated, and the experimental focus has started to shift away from the ‘proof-of-concept’ stage to that of a valid and useful research tool. That is not to say, however, that there is a lack of challenges for subdiffraction imaging. For example, until recently, most experiments focused entirely on the imaging resolution along the focal plane of the microscope, using very thin samples, while effectively neglecting the resolution along the optical axis for the most part. By itself, this is not unreasonable, especially considering the recent nature of these approaches. However, in many experiments the ability to image ‘thick’ samples or to construct optical sections is essential. While considerable progress is being made in this direction,27-32, no conclusive approach has emerged so far.

A second complication is the time resolution of the acquisition, though this is currently mostly a limitation for the PALM-based approaches. As we have tried to highlight in the paper, the approaches that aim to exploit nonlinearities to increase the confinement (RESOLFT, nonlinear structured illumination) have the advantage that the duration to acquire an image is always the same, irrespective of the labeling density of the sample, and that it is possible to collect fewer photons per molecule, as they do not need to be individually localized. In contrast, in PALM-based microscopy, the rate at which molecules can be localized is essentially fixed, and is limited by the requirement that each activated molecule should emit enough photons to be accurately localized. As such, it remains to be seen whether localization approaches will prove capable of resolving fast dynamics (minutes or less) on meaningful samples.

But what exactly is a ‘meaningful sample’? To answer this question we have to look in closer detail at the imaging itself. Fluorescence microscopy is interesting in that it never observes the sample structure directly, but observes the labeling molecules that have been introduced artificially. The ‘structure’ that is observed thus depends on two factors: (i) which part of the sample has been labeled; and (ii) the density of the labeling. Intuitively: if we have an optical resolution of 50 nm but the average distance between the labels is 200 nm, then the resulting fluorescence image contains no meaningful information on features smaller than 200 nm. This is an aspect that has become very relevant with the development of high-resolution imaging.

Fig. 10 Applications in material science: a composite image showing the three-dimensional colloidal packing of 200 nm fluorescent latex spheres, acquired using STED microscopy. Both well organized areas and defects are clearly visible. (Reprinted with permission from Ref. 67. © 2007 American Chemical Society.)

This discussion leads to another issue. While diffraction-unlimited fluorescence microscopy has, to some extent, been made possible through technical advances, a perhaps even more crucial aspect has been the fluorescent label itself. As we have emphasized here, all of the techniques available today rely on special properties of the fluorescent label. This ranges from exceptional photostability to the presence of long-lived nonfluorescent states, or ‘special’ properties such as photoactivation or photoswitching. These processes are not yet fully understood, and the dyes that are available today all have their own drawbacks, directly limiting the imaging. As such, there is a strong requirement for new dyes with improved properties.

As for the applicability of these techniques to practical problems, in the experiments demonstrated so far the focus has mostly been on biological samples, and comparatively little has been done with respect to the material sciences81,82 (Fig. 10). In principle, however, these techniques can be applied to any sample that can be studied by conventional fluorescence microscopy. One possible limitation here might be the requirement for special probe molecules, though, with careful experimentation and the knowledge and labels available today, we feel that it should be possible to obtain a significant improvement in resolution for most kinds of samples.

In particular, research areas where these techniques are set to have a significant impact include nanometer-sized particles or structures such as plasmonic particles, phase transitions in colloidal
systems, single active-site catalysis, laser-induced lithography and micropolymerization. Similar possibilities exist in polymer dynamics, where the material properties depend strongly on the conformation of individual polymer chains, which is hard to obtain directly using diffraction-limited techniques. With time it should be possible to extend these techniques to most fields where nanoscopic particles or features are important.

There is no doubt that diffraction-unlimited fluorescence microscopy is here to stay, and will continue to evolve rapidly. All of the techniques that we have presented here have their own strengths and weaknesses, and hence will be more or less suitable for a particular experiment. At the moment it is impossible to say which, if any, of these techniques will eventually turn out to be superior, or if the way forward will be to use them in a complimentary fashion. Perhaps an as-yet undiscovered approach will take over — it is too early to tell. What can be said, however, is that subdiffraction fluorescence microscopy has opened up an entirely new playing field, and that many untold possibilities await.