

Facts and artifacts in near-field optical microscopy

B. Hecht,^{a)} H. Bielefeldt, Y. Inouye,^{b)} and D. W. Pohl^{c)}
IBM Research Division, Zurich Research Laboratory, CH-8803 Rüschlikon, Switzerland

L. Novotny
Pacific Northwest National Laboratory, P.O. Box 999, Mail Stop K2-14, Richland, Washington 99352

(Received 8 October 1996; accepted for publication 18 November 1996)

Near-field optical (NFO) microscopes with an auxiliary gap width regulation (shear force, tunneling) may produce images that represent the path of the probe rather than optical properties of the sample. Experimental and theoretical evidence leads us to the conclusion that many NFO results reported in the past might have been affected or even dominated by the resulting artifact. The specifications derived from such results for the different types of NFO microscopes used therefore warrant reexamination. We show that the resolving power of aperture NFO microscopes, 30–50 nm, is of genuine NFO origin but can be heavily obscured by the artifact. © 1997 American Institute of Physics. [S0021-8979(97)00305-8]

I. INTRODUCTION

Scanning near-field optical (NFO) microscopy has become a field of considerable activity in recent years.^{1,2} In the present article, we draw attention to an artifact that frequently occurs in NFO imaging and that can lead to serious image misinterpretations. Estimates of resolving power based on such images can be extremely misleading.

NFO microscopes are usually operated with an auxiliary (“non-NFO”) gap-width-sensing mechanism such as the onset of electrical (tunneling)^{3,4} or mechanical (friction or normal force)^{5–7} contact. These mechanisms allow the probe to be moved along the sample surface reliably at very small distances. In addition to NFO images, such microscopes simultaneously provide topographic images, which greatly facilitate the identification of recorded object features.

Widely overlooked, however, is the fact that the motion of the tip normal to the sample surface (“*z* motion”) caused by the gap-width control can and usually *does* give rise to undesired crosstalk in the NFO images. This was noticed some time ago already^{8–11} but unfortunately mentioned only in passing.

The *z*-motion artifact generates features in the optical image that are highly correlated with the structures in the topographic image. Thus the presence of this artifact is not obvious because the sample topography is part of its optical structure. It is expected to be made visible in the NFO microscope, as it is in a conventional microscope. Thus an optical image showing only topographic features can be a genuine NFO picture as well as a pure *z*-motion artifact image, or a superposition of both.

We began to suspect the possible existence of severe artifacts in NFO imaging when noticing that a one-to-one correspondence frequently existed in the structures recorded in NFO and in topographic images. In particular, the same resolution was seen in both images.¹² Why should NFO and

auxiliary interaction mechanisms, in spite of their different physical origin, result in identical resolving power? Furthermore, the apertures inspected with a scanning electron microscope (SEM) were quite often too large for the “NFO” resolution observed.

A survey of the NFO literature revealed the same one-to-one correspondence in many published images. This is particularly true for images that show very small features. Such images are frequently used as a proof of the achieved resolution capability. In Ref. 1, for instance, about 50% of the NFO/topographical image pairs show such a high correlation and have the same resolution.

This is alarming news. *At worst*, it means that

- (i) many “NFO” images may demonstrate NFO-mediated recording of the probe *z*-motion rather than optical superresolution;
- (ii) the worldwide NFO development efforts have been misguided by this overlooked effect; and
- (iii) the various forms of NFO microscopy should be reexamined regarding their superresolution capability.

The experimental evidence we collected unfortunately supports this pessimistic view. We will demonstrate that NFO images can be completely dominated by nonoptical features, that the origin of the *z*-motion artifact can be readily understood but scarcely eliminated, and that its significance increases with decreasing object size. But we also present criteria for artifact-free imaging and show examples of images that indicate excellent, true NFO resolution capability of the *aperture scanning near-field optical microscope* (a-SNOM).

II. EXAMPLES OF Z-MOTION ARTIFACTS

To illustrate the significance of artifacts created by non-optical control, Figs. 1 and 2 show images of a frequently used test structure, the so-called latex sphere projection pattern¹³. The pattern consists of nearly triangular aluminum patches of 15 nm thickness on a thin glass slide, arranged in a hexagonal pattern with a nearest-neighbor distance of 110 nm. The lateral extension (triangle height) of the individual patches is about 50 nm. A tunnel near-field optical micro-

^{a)}Present address: ETH Zürich, Institut für physikalische Chemie, CH-8092 Zürich, Switzerland.

^{b)}Permanent address: Osaka University, Suita, Osaka 565, Japan.

^{c)}Electronic mail: dp@zurich.ibm.com

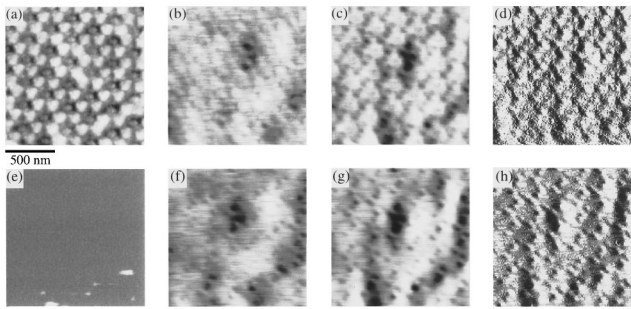


FIG. 1. TNO images of a latex sphere projection pattern, patch diameter ≈ 50 nm (prepared by U. Ch. Fischer, Univ. of Münster, Germany), $\lambda = 633$ nm, good NFO probe tip. Upper row (a)–(d): constant gap-width operation (CGM), lower row (e)–(h): constant height (CHM). (a) and (e) Shear-force (topographic) images; (b) and (f) allowed NFO images; (c) and (g) forbidden NFO images; (d) and (h) differentiated forbidden images.

scope (TNO)^{14,15} provided shear force (SF) images^{5,6} (left) simultaneously with pairs of NFO images (center). The latter represent the radiation transmitted at angles smaller (“allowed”) and larger (“forbidden”) than the critical angle of the glass substrate. The images on the right-hand side were obtained by differentiation, i.e., high-pass filtering of the forbidden light signals along the scan direction. This improves the visibility of small structures because the slow variations in light intensity are suppressed.

The upper row in Fig. 1 depicts a set of images obtained with the TNO operated in constant gap-width mode (CGM) using SF control. The SF image, Fig. 1(a), clearly shows the hexagonal pattern of metal patches. They appear slightly larger than in reality, indicating that the NFO probe⁵ formed a fairly blunt SF tip. The allowed and forbidden NFO images, Figs. 1(b)–(d), reveal patterns similar to that in Fig. 1(a), although the hexagonal structures cannot be recognized as easily. The resolution is in the 40–70 nm range. In the forbidden image, the elevations are dark as to be expected for an absorbing object. In the allowed image, they are bright, a finding which might lead one to speculate about contrast reversal in the imaging of small absorbing objects.

The lower row in Fig. 1, however, gives a distinctly different impression of the test object. It shows the same area of the test object as the upper one but was recorded in constant height mode (CHM), i.e., without z motion of the probe. The NFO probe was retracted from the surface of the object just far enough to avoid contact except at the highest

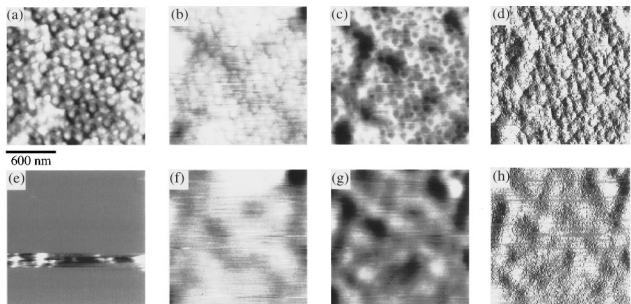


FIG. 2. TNO images as in Fig. 1, bad NFO probe tip.

elevations, where SF control started to operate.¹⁶ For the rest of the topographic image, the SF signal was constant, resulting in the uniformly gray part of Fig. 1(e). Thus the average distance to the object surface was only slightly larger than with SF control.

The hexagonal patterns appear quite differently in Figs. 1(f) and (g) as compared to 1(b) and (c). In the differentiated picture, Fig. 1(h), they are clearest. The patches are dark in both CHM images (hence there is *no* contrast reversal!) and smaller than before, agreeing now with the calculated size. The resolution is obviously better than with SF gap-width regulation, say ≤ 30 nm: The z -motion artifact had masked the excellent resolution of the NFO probe.

The second set of images, Fig. 2, was obtained with the same procedure, but using a different NFO probe. The hexagonal structure is obscured in the topographic image, Fig. 2(a), by superposition of three or four ghost images. Such perturbations are well known in scanning probe microscopy. They are caused by the formation of several minitips at the end of the scanning probe. The optical images depict the same superposition. This clearly suggests a nonoptical origin because it is unlikely that the minitips would act as NFO probes as well.

The CHM images, Figs. 2(f)–(h), indeed show only dull, slowly varying structures typical of a useless, wide-open aperture. The z -motion artifact obviously feigns all the fine features in Figs. 2(b)–(d)!

III. PHENOMENOLOGICAL DESCRIPTION

In the following sections, mechanisms will be discussed by which the auxiliary gap-width control can induce variations in the recorded NFO signals. For simplicity, the latter is assumed to be SF. Other non-NFO gap-width control mechanisms such as normal force (*atomic force microscopy*) or tunneling (*scanning tunneling microscopy*) introduce the same problems. The popular and relatively well-understood a-SNOM operated in emission mode^{3,4,17,18} has been chosen as a basis for discussion and illustration; we emphasize, however, that all related techniques such as (*photon scanning tunneling (optical) STOM, PSTM*),^{19–21} “*apertureless*,”^{22–26} and *reflection*^{27,28} near-field microscopies are subject to the same complications when used with CGM.

Two pairs of signal functions $S_{\text{NFO}}(x, y, z), S_{\text{SF}}(x, y, z)$ and $R_{\text{NFO}}(x, y), R_{\text{SF}}(x, y)$ plus the NFO probe path function $z_{\text{scan}}(x, y)$ will be introduced to put the discussion on a sound basis. The coordinate system shall be fixed with regard to the object, with x pointing in the scan line direction and z in the direction normal to the average object surface. The values of x, y, z shall refer to the center of the aperture at the apex of the glass tip, cf. Fig. 3. The metal film surrounding the aperture is assumed to have a distinct protrusion on one side. In reality this might be the grain of aluminum sticking farthest out from the coating at the tip apex.¹⁶

A. System signal functions

The *system signal functions* $S_{\text{NFO}}(x, y, z)$ and $S_{\text{SF}}(x, y, z)$ are the (electric) signals that optical and SF detector, respectively, can provide as a function of the position

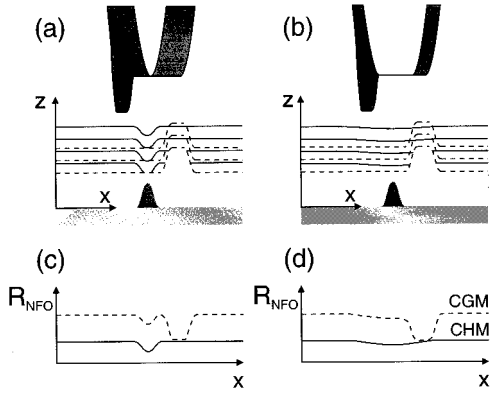


FIG. 3. (a) “Good” and (b) “bad” NFO probes, schematic; contour lines of S_{NFO} (solid lines) and S_{SF} (dashed line), plotted as a function of position of the aperture center. (c) and (d) Scan lines for the different modes of operation, derived from (a) and (b).

of the reference point (x, y, z) . The system signal functions are defined for the entire half space above the sample accessible to the probe. They include all effects caused by illumination, probe, sample, and detection electronics.

Depending on the type of NFO microscope used, S_{NFO} may represent transmitted or reflected light, its polarized or depolarized components, or the response of a fluorescent medium to irradiation from the NFO probe. S_{NFO} can also be the amplitude of modulation if differential techniques such as probe dithering are employed.

The optical properties of the sample are encoded in S_{NFO} . This comprises contributions from the (complex) dielectric constant as well as from topography, such as the surface curvature.²⁹ Owing to the rapid decay of evanescent waves next to the NFO probe, the degree of correspondence between the *optical structure* of the object surface and the lateral variations of S_{NFO} increases for decreasing gap width. The finest resolvable structures are approximately equal to the aperture radius in a-SNOM.

S_{NFO} and S_{SF} are distinctly different functions of x, y, z because of their different interaction and detection mechanisms and also because the size and position of NFO and SF interaction zones on the NFO probe are not necessarily the same. The relations between S_{NFO} and S_{SF} are illustrated schematically in Fig. 3 for a planar sample with a single metal bump on its surface and two different aperture probes.

S_{NFO} and S_{SF} are sketched in the form of contour plots in Fig. 3. Close to the sample surface, the contours of S_{SF} follow more or less the object topography convoluted with the NFO probe geometry; they are also influenced by variations of the tribological properties of the sample.^{10,30,31} The lowest $S_{\text{SF}} = \text{const.}$ curve appears at a distance from the sample surface which corresponds to the height of the SF sensing protrusion. This line defines the minimum distance between aperture center and sample surface and thus limits the range of definition of the signal functions.

The SF response to the bump is shifted to the right in Fig. 3 because the protrusion is on the left-hand side of the reference point. Such a displacement is natural to aperture probes and in fact can help identify nonoptical features in an

SNOM image (cf. Section V). In contrast, the centers of the interaction zones coincide in general for the tips used in STOM/PSTM and other “apertureless” techniques.

The S_{NFO} response to the bump is influenced by the aperture size of the two probes. The small aperture of Fig. 3(a) can provide high-resolution NFO images; accordingly, S_{NFO} is modulated over a narrow range above the bump on the sample surface only. The wide aperture of Fig. 3(b), on the other hand, results in a very gentle effect of the bump on S_{NFO} .

B. Recorded signals

The signals *available* during raster scanning are $z_{\text{scan}}(x, y) \propto U_z$, where U_z is the voltage applied to the “z piezo,” and $R_{\text{NFO}}(x, y)$ and $R_{\text{SF}}(x, y)$. The latter are derived from $S_{\text{NFO}}(x, y, z)$ and $S_{\text{SF}}(x, y, z)$. They depend on the path that the NFO probe actually takes, i.e., the mode of operation.

1. Constant intensity mode (CIM)

The probe follows a contour line $S_{\text{NFO}} = \text{const.}$, given by the set value R_{set} of the controller:

$$R_{\text{NFO}}(x, y) = S_{\text{NFO}}[x, y, z_{\text{scan}}(x, y)] \approx R_{\text{set}}, \quad (1)$$

$$z_{\text{scan}} = \bar{z} + \delta z^{\text{CIM}}(x, y). \quad (2)$$

The signal providing the NFO image is $\delta z^{\text{CIM}}(x, y)$, the variation of z_{scan} around its average value \bar{z} . It represents true optical features only. Nevertheless CIM is rarely used because S_{NFO} does not strictly follow the sample topography nor is it necessarily monotonous in z . (An example will be discussed in Section IV B 1.) Thus control of a very small gap width by CIM is not feasible in general.

2. Constant height mode (CHM)

In CHM, the probe is scanned in a plane parallel to the average object surface, resulting in

$$z_{\text{scan}} = z_{\text{set}}, \quad (3)$$

$$R_{\text{NFO}}(x, y) = \bar{S}_{\text{NFO}}(z_{\text{set}}) + \delta S_{\text{NFO}}(x, y, z_{\text{set}}). \quad (4)$$

Any structure visible in the scan image corresponds to a lateral variation of S_{NFO} and hence to the *optical* properties of the sample surface. The information obtained with CIM and CHM is practically equivalent.³² CHM, however, allows the close-distance operation necessary for high-resolution imaging (cf. Figs. 1 and 2).

3. Constant gap-width mode (CGM)

Itinerant mechanical contact forces the NFO probe to follow a path of (nearly) *constant SF*:

$$R_{\text{SF}}(x, y) = S_{\text{SF}}(x, y; z_{\text{scan}}) \approx R_{\text{set}}, \quad (5)$$

$$z_{\text{scan}} = \bar{z} + \delta z^{\text{CGM}}(x, y), \quad (6)$$

$$R_{\text{NFO}}(x, y) = \bar{S}_{\text{NFO}}(\bar{z}) + \delta S_{\text{NFO}}(x, y, \bar{z}) + \left. \frac{\partial S_{\text{NFO}}}{\partial z} \right|_{\bar{z}} \cdot \delta z^{\text{CGM}}. \quad (7)$$

In Eq. (5) the \approx symbol indicates possible deviations caused by technical limitations of the electromechanical feedback circuit. Such deviations can become significant when the topography undergoes rapid changes and/or the scan speed is very high. It should be emphasized that the following considerations are valid for any path that the probe may take, no matter whether it follows the topography exactly or not.

The signal $R_{\text{NFO}}(x, y)$ in Eq. (7) is developed into a power series of δz^{CGM} .³³ The first two terms on the right-hand side represent the same purely optical information as in the case of CHM operation. The third term couples the z motion of the probe with the optical signal. For the optical properties to dominate the variations of light intensity in a scan image, the δ -terms in Eq. (7) have to satisfy

$$\delta S_{\text{NFO}}(x, y; \bar{z}) \gg \frac{\partial S_{\text{NFO}}}{\partial z} \bigg|_{\bar{z}} \cdot \delta z^{\text{CGM}}. \quad (8)$$

The z -motion artifact is represented by the right-hand side of Eq. (8). It is proportional to the *relative* variation of the NFO signal with z :

$$\Gamma_a \equiv \frac{1}{S_{\text{NFO}}(x, y, \bar{z})} \cdot \frac{\partial S_{\text{NFO}}}{\partial z}. \quad (9)$$

The CHM and CGM signals that would be recorded with the probes of Figs. 3(a) and (b) are plotted schematically in Figs. 3(c) and (d). Only the small aperture probe of Fig. 3(a), operated in CHM, provides a good image of the *optical* structure. The aperture in Fig. 3(b) is far too wide to generate a high-resolution NFO image, and in CGM operation, the scan lines are dominated by the SF response to the passage over the bump. In Fig. 3(c) the true NFO signal can still be recognized but in Fig. 3(d) the CGM trace gives a completely misleading impression about the imaging capability of the probe.

IV. NFO SIGNAL VERSUS PROBE HEIGHT

A. Physical effects

The coupling factor Γ_a introduced in Eq. (9) depends on the properties of the SNOM as well as on those of the object. Γ_a is a superposition of several components in general. The following list can give a feeling for the size of the coupling $(\partial \bar{S}_{\text{NFO}} / \partial z) / S_{\text{NFO}}$ between the z motion of the probe and the artifact it can generate in the NFO image.

(1) *Proximity fields*: The electric fields next to very small structures ($\ll \lambda$), such as SNOM apertures, decay over distances comparable to their lateral size, giving rise to large gradients in S_{NFO} . However, the contribution is significant only for near-field detection^{34,35} because these fields are mainly nonradiative.

(2) *Evanescent fields*: These are inhomogeneous, guided waves of the type created by total internal reflection. They are exploited in STOM/PSTM, in the ‘‘forbidden’’ channel of the TNOM as well as in some ‘‘apertureless’’ schemes. These fields decay exponentially with gap width. The decay length is typically of the order $\lambda/2\pi \approx 100$ nm, which yields $\Gamma_a^{\text{evan}} \approx -1\%$ /nm.

(3) *Standing waves*: In arrangements with light propagation in z direction, reflection/scattering from the sample surface and from the NFO probe can produce interference undulations in the approach curve. The phase angle at the sample surface depends on the dielectric properties of the object. Γ_a^{interf} can vary between positive and negative values of up to 1%/nm.

(4) *Inhomogeneous illumination*: In some types of NFO microscopes, the probe is illuminated by a sharply focused laser beam whose orientation is either collinear to the axis of the NFO probe^{23,36,37} or tilted at a large angle.^{25,38} On the axis, the intensity of illumination varies as $1/[1+(as)^2]$, where s is the distance from the focal plane and $a = (k_0/2) \times [\arcsin(\text{NA}/n)]^2$ with $k_0 = 2\pi/\lambda$, $n =$ index of refraction, and NA = numerical aperture of the focusing objective. The exact position of the NFO-active part of the probe with respect to the beam waist has a strong influence on S_{NFO} . For an NA=1.3 oil immersion objective, it is sufficient that the NFO probe be out of focus by as little as ± 100 nm to generate $\Gamma_a^{\text{illum}} = \pm 0.5 \dots 0.7\%$ /nm.

(5) *Confocal detection*: In order to discriminate stray light, the NFO probe is frequently imaged onto a small pinhole. Any deviation from perfect alignment caused by the z motion of the probe results in a variation of detected light intensity similar to the previous case. To reduce this effect, the position of the NFO probe may be kept fixed and the sample position be adjusted instead. Variations of the object position still cause defocusing $\propto (n-1) \cdot z$, where n is the refractive index of the sample substrate. Hence $\Gamma_a^{\text{confoc}} \approx \Gamma_a^{\text{illum}}$.

(6) *Topographic influences*: Besides these general features, Γ_a also depends on the specific optical and topographic structures of the sample. The influence of these factors is significant but difficult to predict. An example of such behavior will be presented in the next section.

B. Experimental illustration

1. Determination of Γ_a

Figure 4 shows the approach characteristics for a glass grating measured with our TNOM.^{14,15} The grating consists of nearly rectangular elevations 8 nm high and 190 nm wide, interleaved with flat valleys of the same width [Fig. 4(a)]. The NFO probe is scanned in the x, z plane, which is oriented perpendicular to the surface and the grating stripes. Each scan line represents one retraction curve starting from the point of SF contact.

The characteristic features of allowed and forbidden light, viz., interferences fringes and exponential decay,¹⁵ can be readily seen [Figs. 4(b) and (c)]. Near contact, one recognizes additional structure in the allowed light, which is obviously related to the shape of the grating. Figure 4(d) depicts two pairs of allowed and forbidden approach curves obtained by averaging 12 individual curves from top and bottom positions, respectively. The latter curves appear shifted ≈ 8 nm to the right with respect to the first ones in Fig. 4(c). The shift equals the variation of the starting point for top and valley positions. When corrected for the 8-nm

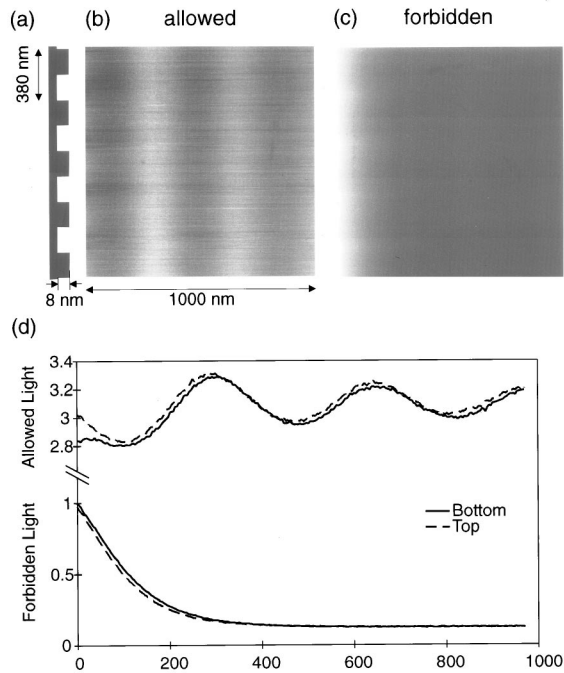


FIG. 4. (a) Profile of the test sample “glass grating” (prepared by B. Curtis, PSI Zurich); (b) and (c) allowed and forbidden x/z scan images; (d) approach curves for top and valley positions of the NFO probe (averaged over 12 scan lines each).

displacement, the two forbidden curves coincide exactly, and so do the allowed ones except very close to contact.

The Γ_a values near contact are $-1\%/nm$ for forbidden light, independent of position. For allowed light, Γ_a varies between $+0.02$ (bottom position) and $-0.2\%/nm$ (top). This behavior demonstrates that the weight of the artifact can vary considerably across an object surface and can change even from positive to negative contrast.

2. Cross-check: Artificial artifacts

In another control experiment, the attenuated output voltage from a function generator was added to the SF signal while the NFO probe was scanned over a plane glass plate. The gap-width control circuit of the SNOM retracted the NFO probe by 1 nm when the voltage was turned on. This generated the grating structure seen in Fig. 5(a). Without z -motion crosstalk, the NFO signal would have remained unaffected because there is no optical structure on the sample

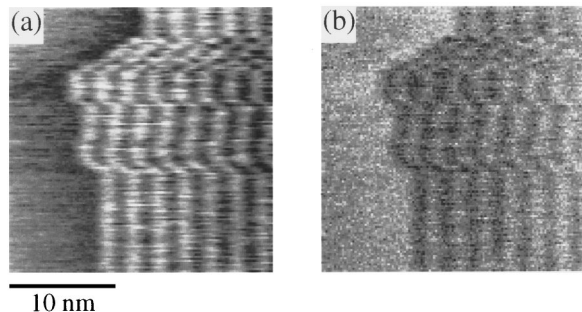


FIG. 5. False high-resolution (a) SF and (b) forbidden NFO images demonstrating the influence of a 1-nm gap-width variation.

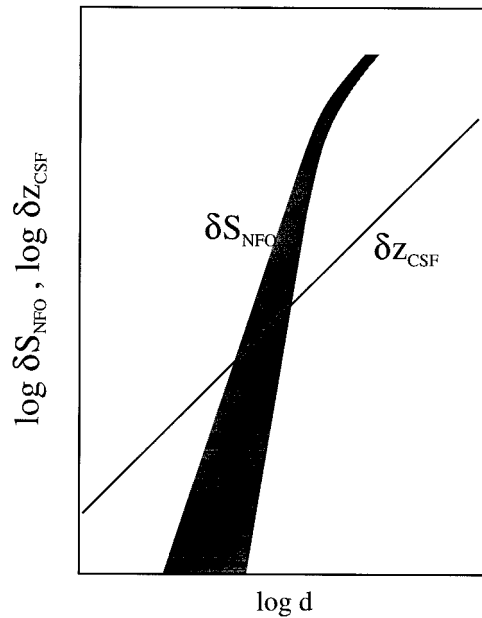


FIG. 6. Variation of NFO and SF signals with particle size, schematic.

surface. The very clear response of the forbidden NFO signal [Fig. 5(b)] proves that topographic motion can fake an optical structure which does not really exist. The modulation was visible in the NFO image down to an amplitude of 0.1 nm, i.e., down to atomic size variations in gap width!

C. Influence of object size

The discussion in the previous two sections leads to the conclusion that $|\Gamma_a|$ is of the order of 1%/nm in general. The second quantity required for an estimate of the different influences is the ratio $[\delta S_{\text{NFO}}(x, y; \bar{z}) / \bar{S}_{\text{NFO}}] / \delta z^{\text{CGM}}$. There is no simple way to determine this quantity. A general trend with regard to object size, however, might be deduced from the following consideration: A small object such as the bump in Fig. 3 may be considered a Mie scatterer. In the electrostatic limit (diameter $d \ll \lambda$), the electric field at a fixed distance from the center of the particle is proportional to its volume.^{39,40} If the size of the bump shrinks uniformly in all three dimensions and if the distance of the NFO zone from the sample is given by the height of the probe protrusion, then the variation of the NFO signal δS_{NFO} decreases with the third (sixth) power of the particle size for heterodyne (homodyne) detection. The topographic scan signal δz^{CGM} , on the other hand, depends linearly on the particle diameter:

$$\delta S_{\text{NFO}} \propto d^3 \dots^6, \quad \delta z^{\text{CGM}} \propto d. \quad (10)$$

This situation is depicted schematically in the double-log plot in Fig. 6. The reduced slope of δS_{NFO} for large d reminds us that these considerations are valid only for very small objects. The different dependence on d makes the z -motion artifact increasingly dominant with decreasing particle diameter. It should be noted that Γ_a does not scale with the particle size because the denominator in Eq. (9) is given by the properties of the bare sample surface.

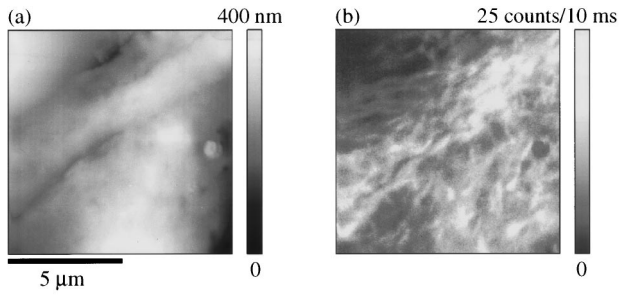


FIG. 7. (a) SF and (b) NFO fluorescence images, CGM control, almost free of z -motion artifacts. The image was made in collaboration with P. Descouts and M. Jobin, University of Geneva.

V. IDENTIFICATION OF TRUE NFO FEATURES

It was shown in Section III B that operation in *constant height* mode is the only way—except for the *constant intensity* mode not further considered here—to avoid the occurrence of the artifact completely. High resolution, on the other hand, requires that the probe be scanned very close to the object surface. Imaging of highly corrugated samples may therefore necessitate auxiliary feedback to be retained. In this case, the resulting scan records have to satisfy condition (8).

In order to obtain unperturbed NFO images in CGM mode, it has been suggested that the crosstalk term in Eq. (7) be determined and subtracted from R_{NFO} .²⁶ It is true that $\partial S_{\text{NFO}}(x, y, z)/\partial z$ can be measured locally by dithering the NFO probe or from approach (retraction) curves of the type shown in Fig. 4(d). The z -motion $\delta z^{\text{CGM}}(x, y)$ is the signal recorded anyway in the topographic image. Thus, a quasi-CHM image might be constructed from Eq. (7). The always existing uncertainty in the experimental data makes it difficult, however, to determine whether perfect compensation of the z -motion crosstalk was achieved.

In our experience, at least one of the following requirements has to be satisfied by an “NFO image” to be credible:

- (1) The NFO image is obtained in CHM or CIM.
- (2) The NFO image is obtained in CGM but:
 - (a) *Topographic and NFO images are highly uncorrelated.* This means that cross correlation must be minor for *any part* of the image because $\partial S_{\text{NFO}}(x, y, \bar{z})/\partial z$ is not necessarily constant across the scan area (see Fig. 4). Therefore, varying or even reversed correlation within one frame is not sufficient.
 - (b) *Correlated structures are displaced by a constant amount.* This indicates that optical and nonoptical interaction zones do not coincide, cf. Fig. 3.
 - (c) *The resolution of NFO and SF images are clearly different.*

Conditions 2(a) and (c) are satisfied, for example, in Fig. 7, depicting part of a T3 fibroblast cell, imaged with an a-SNOM adapted to fluorescence recording.¹⁶ The actin cytoskeleton is labeled with a fluorescent dye. The SF image shows the more or less smooth cell surface, whereas the NFO image clearly reveals the skeleton structure. The little bump on the right and the fold on the left, however, are so

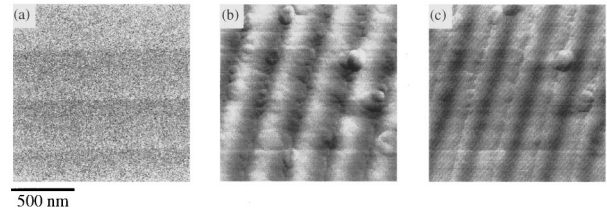


FIG. 8. High-resolution NFO imaging without z -motion artifacts: Grating of Fig. 4, imaged in CHM: (a) SF channel, (b) allowed, (c) forbidden light images. Details 10 nm in size can be recognized thanks to an exceptionally good aperture probe.

similar in both pictures that the z -motion artifact could well be the origin for their appearance in the NFO image.

Conditions 2(b) and (c) are satisfied by the “eye-bee” image published in Ref. 41. The surface roughness is visible in the SF image only; the displacement is obvious in the forbidden NFO image which displays, besides the true NFO features, the SF contours as a ghost image. The latter is invisible in the allowed NFO image because of the smaller Γ_a , cf. Fig. 4.⁴²

A similar situation apparently existed in our first NFO studies.^{3,4} The equipment available at that time allowed the registration of only a single image. Single scan spot checks with two-pen recorders, however, showed large offsets and different broadening of the recorded features in general.⁴³ This was not unexpected because the aperture NFO probes were fabricated in a way that generated extremely extended, very flat plateaus at the tip apex. The formation of the STM contact far off the aperture was therefore probable.

In our experience, the best way to obtain well-defined NFO images is to scan first in CGM and then in CHM. This allows one to scan very close to contact in CHM and provides sets of images similar to those depicted in Figs. 1 and 2. The resolution was excellent when the microscope was operated in this way with a good aperture probe. For example, Fig. 8 depicts the same grating as in Fig. 4. The fine structures reveal 10-nm details, which correspond to the pixel size. The image was obtained at the end of a long series of measurements with the same NFO probe, displaying considerably less sharp features. It is possible that the metal coating that usually protrudes from the aperture proper was gradually ground away, allowing the aperture to get closer and closer to the sample surface.

VI. APPEAL

For NFO to remain a credible area of research, it is indispensable that all images be carefully checked for the artifact described. Its elimination will more easily allow us to focus future R&D efforts on the true problems of NFO imaging, namely on reliability, simplicity, correct interpretation, and progress towards higher resolving power.

ACKNOWLEDGMENTS

The authors appreciate the permission of P. Descouts and M. Jobin to include Fig. 7 in the present article. They owe thanks for support and discussions to R. Carminati, D. Courjon, U. Dürig, C. Durkan, G. Eggers, U. Ch. Fischer, M.

Fujihira, N. Garcia, J.-J. Greffet, H.-J. Güntherodt, H. Heinzelmann, K. Karrai, O.J.F. Martin, F. Meixner, M. Ohtsu, V. Sandoghdar, and I. Shvets. This work was supported in part by the SPP *OPTIQUE* of the Board of the Swiss Federal Institutes of Technology and by the European program *Human Capital and Mobility*.

- ¹Near Field Optics, Vol. 61, Nos. 1–4 of *Ultramicroscopy*, edited by M. Paesler and N. van Hulst (Elsevier, Amsterdam, 1995); Proc. 3rd Int'l Conf. on Near-Field Optics (NFO III), Brno, Czech Republic, May 9–11, 1995.
- ²M. A. Paesler and P. J. Moyer, *Near Field Optics Theory, Instrumentation, and Applications* (Wiley, New York, 1996).
- ³D. W. Pohl, W. Denk, and M. Lanz, *Appl. Phys. Lett.* **44**, 651 (1984).
- ⁴U. Dürig, D. W. Pohl, and F. Rohner, *J. Appl. Phys.* **59**, 3318 (1986).
- ⁵E. Betzig, P. L. Finn, and S. J. Weiner, *Appl. Phys. Lett.* **60**, 2484 (1992).
- ⁶R. Toledo-Crow, P. C. Yang, Y. Chen, and M. Vaez-Iravani, *Appl. Phys. Lett.* **60**, 2957 (1992).
- ⁷N. F. van Hulst, M. H. P. Moers, O. F. J. Noordman, R. G. Tack, and F. B. Segerink, *Appl. Phys. Lett.* **62**, 461 (1993).
- ⁸E. Betzig, in *Near Field Optics*, Vol. 242, *NATO ASI Series E: Applied Sciences*, edited by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993), p. 7.
- ⁹G. A. Valaskovic, M. Holton, and G. H. Morrison, *Appl. Phys. A* **34**, 1215 (1995).
- ¹⁰C. Durkan and I. V. Shvets, *J. Appl. Phys.* **80**, 5659 (1996).
- ¹¹V. Sandoghdar, S. Wegscheider, G. Krausch, and J. Mlynek, *J. Appl. Phys.* **81**, 2499 (1997).
- ¹²Pictures of this type are reproduced in Ref. 15. In retrospect, it is clear that the “allowed” images represent true NFO features whereas the “forbidden” ones are dominated by the artifact.
- ¹³U. C. Fischer and H. P. Zingsheim, *J. Vac. Sci. Technol.* **19**, 881 (1981).
- ¹⁴B. Hecht, H. Heinzelmann, and D. W. Pohl, *Ultramicroscopy* **57**, 228 (1995).
- ¹⁵B. Hecht, D. W. Pohl, H. Heinzelmann, and L. Novotny, in *Photons and Local Probes*, Vol. 300, *NATO ASI Series E: Applied Sciences*, edited by O. Marti and R. Möller (Kluwer, Dordrecht, 1995), p. 93.
- ¹⁶B. Hecht, Ph.D. thesis, University of Basel, Switzerland, 1996.
- ¹⁷E. Betzig, A. Lewis, A. Harootunian, M. Isaacson, and E. Kratschmer, *Biophys. J.* **49**, 269 (1986).
- ¹⁸E. Betzig and J. Trautman, *Science* **257**, 189 (1992).
- ¹⁹P. C. Reddick, R. J. Warmack, and T. L. Ferrell, *Phys. Rev. B* **39**, 767 (1989).
- ²⁰D. Courjon, K. Sarayedine, and M. Spajer, *Opt. Commun.* **71**, 23 (1989).
- ²¹F. de Fornel, J. P. Goudonnet, L. Salomon, and E. Lesniewska, *Proceedings of the Society of Photooptical Instrument Engineering* (SPIE, Bellingham, 1994), Vol. 1139, p. 297.
- ²²U. C. Fischer, in *Scanning Tunneling Microscopy and Related Methods*, Vol. 184, *NATO ASI Series E: Applied Sciences*, edited by R. J. Behm, N. Garcia, and H. Rohrer (Kluwer, Dordrecht, 1990), p. 475.
- ²³F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **65**, 1623 (1994).
- ²⁴P. Gleyzes, A. C. Boccara, and R. Bachelot, *Ultramicroscopy* **57**, 318 (1995).
- ²⁵S. Kawata and Y. Inouye, *Ultramicroscopy* **57**, 313 (1995).
- ²⁶Y. Martin, F. Zenhausern, and K. Wickramasinghe, *Appl. Phys. Lett.* **68**, 2475 (1996).
- ²⁷M. Spajer and A. Jalocha, in *Near Field Optics*, Vol. 242, *NATO ASI Series E: Applied Sciences*, edited by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993), p. 87.
- ²⁸R. Bachelot, P. Gleyzes, and A. C. Boccara, *Opt. Lett.* **20**, 1924 (1995).
- ²⁹B. Hecht, D. P. Pohl, H. Heinzelmann, and L. Novotny, *Ultramicroscopy* **61**, 99 (1995).
- ³⁰Th. Lacoste, Th. Huser, H. Heinzelmann, and H.-J. Güntherodt, in *Photons and Local Probes*, Vol. 300, *NATO ASI Series E: Applied Sciences*, edited by O. Marti and R. Möller (Kluwer, Dordrecht, 1995), p. 123.
- ³¹A. Jalocha, M. H. P. Moers, A. G. T. Ruiters, and N. F. van Hulst, *Ultramicroscopy* **61**, 221 (1995).
- ³²R. Carminati, Ph.D. thesis, UPR 228, CNRS et Ecole Centrale Paris, France, 1996; R. Carminati and J. J. Greffet, *Opt. Lett.* **21**, 1208 (1996).
- ³³A similar approach is described in Ref. 32.
- ³⁴W. E. Moerner, T. Plakhotnik, T. Irngartinger, U. P. Wild, D. W. Pohl, and B. Hecht, *Phys. Rev. Lett.* **73**, 2764 (1994).
- ³⁵G. Kolb, K. Karrai, and G. Abstreiter, *Appl. Phys. Lett.* **65**, 2090 (1994).
- ³⁶R. Bachelot, P. Gleyzes, and A. C. Boccara, *Ultramicroscopy* **61**, 111 (1995).
- ³⁷F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, *Science* **269**, 1083 (1995).
- ³⁸J. Jersch and K. Dickmann, *Appl. Phys. Lett.* **68**, 868 (1996).
- ³⁹D. W. Pohl, in *Scanning Near-field Optical Microscopy (SNOM)*, Vol. 12, *Advances in Optical and Electron Microscopy* (Academic, London, 1991), p. 243.
- ⁴⁰M. Born and E. Wolf, *Principles of Optics* (Pergamon, London, 1959).
- ⁴¹D. W. Pohl, L. Novotny, B. Hecht, and H. Heinzelmann, *Thin Solid Films* **273**, 161 (1996).
- ⁴²Another good example of an SF-induced ghost image is depicted and correctly described in Fig. 5 of Ref. 8.
- ⁴³U. Dürig (private communication).