REVIEW ARTICLE

Near-field optics theories

To cite this article: Christian Girard and Alain Dereux 1996 *Rep. Prog. Phys.* **59** 657

View the article online for updates and enhancements.

Related content

- Near-field physics
  Christian Girard, Christian Joachim and
  Sébastien Gauthier

- Near fields in nanostructures
  Christian Girard

- Near field microscopy and near field optics
  D Courjon and C Bainier

Recent citations

- A multi-frequency iterative imaging method
  for discontinuous inverse medium problem
  Lei Zhang and Lixin Feng

- Extreme call amplitude from near-field
  acoustic wave coupling in the stridulating
  water insect Micronecta scholtzi
  (Micronectinae)
  Andrew Reid *et al*

- Carbon nanotube-based three-
  dimensional monolithic optoelectronic
  integrated system
  Yang Liu *et al*

REVIEW ARTICLE

Near-field optics theories

Christian Girard† and Alain Dereux‡

† Laboratoire de Physique Moléculaire URA CNRS 772, Université de Franche-Comté, F-25030 Besançon, France
‡ Laboratoire de Physique, Optique submicronique URA CNRS 1796, Université de Bourgogne, F-21004 Dijon, France

Abstract

The development of near-field optics theory is reviewed. We first recall that near-field optics is not limited to near-field microscopy. Broadly speaking, it concerns phenomena involving evanescent electromagnetic waves. The importance of such waves was ignored for a long time in optical and surface physics until the emergence of scanning near-field optical microscopes. Taking evanescent waves into account prevents the use of any simple approximation in the set of Maxwell's equations. The various theoretical approaches of near-field optics are discussed from the point of view of their ability to assess evanescent electromagnetic waves. We discuss the main results of the application of the various practical schemes which all rely on a numerical procedure.

This review was received in February 1996

Contents

1. Introduction

Some 15 years ago, the first optical topographic signals with subwavelength resolution were recorded independently by different groups in the world. These early works were encouraged and stimulated by speculations about the properties of evanescent and confined fields concentrated near the surface of materials. Since this exploratory period, a broad variety of scanning near field optical microscopes (SNOM) have been elaborated and continuously improved. An historical presentation of this pioneering period is detailed in the proceedings of the first near-field optics conference (Pohl and Courjon 1993).

The common feature of all SNOMs is the nanometre-sized detector able to collect or emit photons after coupling with a subwavelength size object deposited on a surface. Depending on the experimental design, this nano-detector can be used to transmit the collected light to an appropriate macro-detector (for example a photomultiplier) located far from the object (Pohl 1991, 1992). Today, many experimental configurations based on this concept of nano-detection provide us with an increasing amount of optical information about the nano-world. A recent comprehensive review of the experimental aspects of near-field microscopy appeared in this journal (Courjon and Bainier 1994).

1.1. Optics in the nanoscopic and mesoscopic regimes

The adjective *mesoscopic* is used to define the situations where the sizes are of the order of the incident wavelength λ. For visible light, it corresponds roughly to the length range between 0.1 and 1 μm. By *nanoscopic*, one usually means low-dimensional structures smaller than 100 nm. If these structures can be identified with single molecules, the nanoscopic regime also means the molecular range. However, structures smaller than 1 nm are commonly viewed as belonging to the atomic range.

When λ is much smaller than the size of the scatterers, one speaks of the macroscopic regime. Geometrical optics is a first approximation which describes the scattering of light by macroscopic objects. On a more refined level, Kirchhoff's diffraction theory uses a scalar field to account for phenomena where light displays a wave character on a macroscopic scale. Kirchhoff's theory attributes ideal properties to the scatterers such as a perfect conductivity or a real refractive index.

Microscopic systems are objects which are so small when compared to the incident wavelength that the non-retarded approximation becomes applicable. This approximation considers the scatterers as dipoles or a set of dipoles whose susceptibilities may include dissipative effects. For visible wavelengths, this regime corresponds to the atomic range.

Near-field optics deals with phenomena involving evanescent electromagnetic waves which becomes significant when the sizes of the objects are of the order of λ or smaller. By object, we also mean void structures carved in a surrounding material such as the vacuum gap basic to ATR (attenuated total reflection) experiments or the air gap separating a local probe and a sample surface. In view of the above classification, it is clear that near-field optics is thus concerned with the scattering of electromagnetic waves by meso- and nanoscopic systems. Even in the situations where atomic size structure is involved, near-field optical detection is affected by the nano- and mesoscopic system embedding the atomic size structure. Evanescent waves are important in near-field optics because the typical size of the objects is comparable to λ and the decay of evanescent waves occurs within a range given by λ.

1.2. The theoretical challenge

Today's theory of electromagnetic waves describes satisfactorily their interaction with objects which are macroscopic or microscopic relative to the incident wavelength λ. However, the theoretical knowledge about the scattering of electromagnetic waves by mesoscopic systems remains limited. Since many situations involve nanoscopic and mesoscopic systems simultaneously, incomplete information about the mesoscopic range impedes our understanding of nanoscopic systems. Most approximations are not appropriate for studying mesoscopic systems. Unlike macroscopic systems (successfully described by Kirchhoff theory) and microscopic systems (for which retardation is negligible), mesoscopic systems require the detailed solution of the full set of Maxwell equations.

A numerical method for solving Maxwell equations is needed because both geometries and dielectric responses of typical mesoscopic systems display a high degree of complexity. However, numerical methods traditionally used in electrodynamics are not well suited to cope with mesoscopic structures. Cumbersome procedures appear to be uncertain and produce unreliable outputs. The main origin of their problems can be traced to the crucial role played by the evanescent components of the field in the near-field zone close to mesoscopic scatterers. In complete analogy with the tunnel effect for electrons, these evanescent components can lead to optical tunnel effects. In the mesoscopic range, the accurate treatment of evanescent waves requires one to deal carefully with the electromagnetic boundary conditions at each interface and to include realistic dielectric responses.

1.3. Early theoretical works

The physics of evanescent electromagnetic waves, which is the central concept used in near-field optics, was a poorly developed research area before the mid 1960s. The analysis of the skin depth effect at metallic surfaces by Zenneck (1907) and Sommerfeld (1909) was probably the first recognition of the existence of evanescent electromagnetic waves. The famous papers of Mie (1908) and Debye (1909) about the scattering of electromagnetic waves by a sphere contained, at least in principle, the solution of the vector wave equation in the near-field zone not only for the microscopic regime but also for the mesoscopic one. However, these works were little exploited in the study of near-field optics. Indeed, for coordinates located in the near-field zone, the formulae of Mie and Debye display convergence problems which are even more difficult to solve in the mesoscopic regime. Much later, Fano (1941) realized that the anomalies observed in the diffraction of light by metal gratings were related to the excitation of evanescent electromagnetic modes bound to the surface.

The classical problem of diffraction by an aperture in a perfectly conducting screen was treated by numerous different approaches (see the review published in this journal by Bouwkamp (1954)). In 1944, Bethe proposed a curious model involving magnetic charges in order to extract a solution to the problem of the diffraction of light by a subwavelength aperture. The assessment of the near-field zone at the exit of the aperture was feasible in principle but was ignored because it was not experimentally relevant at that time. Much later, Leviatan (1986) and Roberts (1987, 1989, 1991) applied Bethe's theory to apertures sizes which were significant for near-field microscopes. Such computations confirmed unambiguously the existence of an important exponentially decaying field near the exit of the aperture. In the approximation of a perfectly conducting screen, the generation of this near-field may be understood qualitatively with the help of the Heisenberg uncertainty

principle (Dürig *et al* 1986, Vigoureux and Courjon 1992). Let us consider the simpler case of the diffraction of an incident plane wave $e^{ik_z z}$ propagating in the z direction by an aperture of diameter a carved in an infinitely thin screen covering the x–y plane. The Heisenberg uncertainty

$$\Delta s \, \Delta k_s \geqslant 1 \qquad (s = x, y) \tag{1}$$

states that passage through a slit implies that the transmitted field acquires a non-zero angular spectrum

$$\Delta k_s \geqslant a^{-1}. \tag{2}$$

In the case of a subwavelength aperture, $k \ll a^{-1}$ and we find

$$k \ll \Delta k_s. \tag{3}$$

The dispersion relation

$$k^2 = k_x^2 + k_y^2 + k_z^2 = \frac{\omega^2}{c^2} \epsilon \tag{4}$$

where ϵ is the dielectric function of the background reference medium, shows that k_z^2 may become either positive or negative. The Fourier components of the diffracted field may be classified according to the sign of k_z^2. For positive values, the behaviour $e^{ik_z z}$ corresponds to radiative (or propagative) waves along z. These waves reach the far-field. For negative values, the Fourier component takes the evanescent form $e^{-|k_z|z}$. The set of imaginary values of k_z defines the non-radiative waves existing in the near-field zone. Their exponental decay prevents them from reaching the far-field (Vigoureux *et al* 1989, 1992).

An important breakthrough was achieved by Levine and Schwinger (1950) who established the non-trivial form of the free-space Green dyadic. This dyadic allowed them to formulate a variational method of resolution to the puzzling problem of diffraction by small apertures. As in the case of Bethe's theory, the near-field zone at the exit of the aperture was not assessed at that time. However, the work of Levine and Schwinger inspired the later development of electromagnetic scattering theory (see section 4 later).

As discussed above, the assumption of a perfectly conducting screen allows one to reproduce the generation of the near-field as a result of the Heisenberg's uncertainty and can account for polarization effects. However, this assumption is much too restrictive when considering near-field optical phenomena since it hides resonance phenomena. Indeed, according to their frequency dependent dielectric properties, localized eigenmodes characterized by evanescent wavefunctions may be sustained by small objects and even by surfaces. The understanding of the physical content of the dielectric function was triggered by the work of Huang (1951). Huang brought to the fore the fact of how the parameters driving the infra-red values of the dielectric constant of a polar crystal are closely related to the coupling of light with vibrational eigenmodes (phonons) of the crystal. Hopfield (1958) and Pekar (1960) introduced a similar idea for the range of visible wavelengths by invoking the coupling of light to the excitons of the crystal. In particular, Hopfield developed the concept of polaritons and was the first to observe their dispersion relations by using Raman spectroscopy (Henry and Hopfield 1965). Polaritons are the polarization waves of a crystal which are excited by incident light. They are the electromagnetic eigenmodes of condensed matter.

The clarification brought by the concept of polariton allowed one to identify the conditions of existence of evanescent electromagnetic eigenmodes bound to a surface (Ferrell 1958, Stern and Ferrell 1960) and to the interfaces of a thin film (Kliewer and Fuchs 1966, Fuchs *et al* 1966). In 1968, Otto invented and explained theoretically the attenuated total

reflection spectroscopy (ATR) which allowed him to measure the dispersion relations of interface polaritons (plasmon-polaritons in the original experiment) by a simple reflectivity measurement. The principle of ATR consists in approaching a sample surface in the decay range of the evanescent wave produced by total reflexion on a prism surface. By the optical tunnel effect, the evanescent wave can then excite the non-radiative interface modes of the sample. This results in the frustration of the total reflection. Today's near-field optical microscopes using internal illumination (STOM or PSTM devices) rely on the basic discovery of Otto (Reddick *et al* 1989, 1990, Courjon *et al* 1989, Vigoureux *et al* 1989).

The success of the polariton concept encouraged further development. Although the subject of light scattering by small particles had already been thoroughly investigated (Mie 1908, van de Hulst 1957), electromagnetic eigenmodes of small spherical particles were revisited from the point of view of their frequency-dependent dielectric function (Englman and Ruppin 1966, Fuchs and Kliewer 1968). It appeared that, due to the curvature of the particles, the distinction between radiative and non-radiative modes was not so clear as it was in the case of planar interfaces. Nevertheless, it was rapidly recognized that the coupling of such particles with a planar surface changed dramatically as a function of the distance between them. This coupling can be identified as a true near-field effect since it occurs for separation distances smaller than the wavelength so that the evanescent waves scattered by the particle have not yet decayed.

The related phenomenon of enhanced Raman scattering of molecules adsorbed on metallic surfaces proved to be of electromagnetic origin (Otto 1984). In the non-retarded approximation, the particle is modelled by a point dipole and the surface plasmon is reduced to the image dipole. The local electric field at the coordinate of the molecule is the sum of the external field and the field due to the image dipole. This approximation reproduces the red-shift observed in the absorption spectrum (Ruppin 1983).

The inclusion of retardation in the case of metal particles approaching a metal surface demonstrates that a hybrid plasmon can show up as the result of the coupling of the particle plasmon to the surface one (Takemori *et al* 1987). The absorption spectrum of the sphere is much more red-shifted than expected by the electrostatic approximation. The finite size of the sphere makes possible the excitations of several plasmons related to higher multipole modes of the sphere.

If a dielectric (i.e. non-absorbing) sphere is substituted for the metallic one, the absorption can only be due to the surface plasmon of the metal surface. The incident field excites the resonant but undamped modes of the dielectric sphere which are distributed as evanescent waves around the sphere. As the sphere approaches the surface, a larger number of surface plasmons with shorter wavelengths can be excited. Therefore, absorption occurs in the frequency range located below the highest surface plasmon frequency and above the lowest frequency which can sustain an eigenmode of the sphere. Near-field optical microscopes using or detecting the resonance effects of small particles operate by this principle.

Concurrent with the study of the optical properties of small particles deposited on surfaces, several works brought to the fore the fact of how the presence of a surface alters the light emission of dipole and multipole sources (Lukosz and Kunz 1977a, b, Lukosz 1979, Lukosz and Meier 1981). The physical effects may also be understood as a consequence of the coupling of the dipole or multipole source to the planar surface. In the presence of a perfect conductor, the emitted intensity varies according to the orientation of the dipole: a dipole placed vertically radiates more than twice the value emitted in a homogeneous environment; for a dipole placed horizontally, the radiated intensity vanishes as the distance to the surface is reduced to contact. As the dipole is approaching a real (absorbing) metal,

a larger fraction of the power emitted by the dipole is dissipated in the metal so that the far-field radiation becomes quenched. When the dipole is facing a dielectric, the power transmitted into the dielectric increases as the distance is reduced. This is due to the conversion of evanescent waves into radiative waves in the medium with a larger index of refraction. The above mechanisms apply when modelling the electromagnetic coupling of molecules to solid surfaces. They constitute the basic principles of the recent trend towards near-field fluorescence microscopy (Betzig and Chichester 1993, Pedarnig *et al* 1995).

1.4. Perturbative or self-consistent approach?

Understanding the optical tip–sample interaction presented surely one of the most serious challenges for the beginning of near-field optical microscopy research. Empirical steps contributed to the progress in designing tips which provide a good imaging quality. Different tip designs evolved according to the type of experimental set-up. Internal illumination devices (STOM or PSTM) exploit bare and sharply elongated optical fibres (Reddick *et al* 1989, Courjon *et al* 1989, 1990, Adam *et al* 1993, van Hulst *et al* 1992, 1993) while external illumination scanning near-field optical microscopes (SNOM) favour metallized tips with a subwavelength aperture at the apex (Betzig *et al* 1986, 1987, 1991, Lieberman *et al* 1990). A few wavelengths away from the apex, the general shape of such tips is usually smooth. However, recent configurations involving tips with a tetrahedral termination (entirely or partly coated with metal) were also successful (Fischer and Köglin 1995). Metal tips were also found appropriate for the scanning surface plasmon microscope (SSPM) (Specht *et al* 1992).

   All these developments and their numerous variations were supported exclusively by instrumental intuition since classical optical theories were ineffective in describing the basic features of the tip–sample interaction. Moreover, the early theoretical works on near-field optics did not approach this interaction self-consistently (Vigoureux *et al* 1989). The pioneering investigations provided insights into the field distribution behind subwavelength apertures but without any sample present. After these exploratory studies, the first methods applied to simulate near-field microscopes images computed the electromagnetic field diffracted above a non-planar sample surface but ignored the presence of any probe. Such procedures were later improved by the use of ideal probes which did not disturb the near-field above the sample. Indeed, these probes were introduced after the computation of the field above the bare surface basically to model devices which integrate the optical near-field over a finite volume. Such ideal probes are thus not coupled to the sample when solving Maxwell's equations. Neglecting this coupling was assumed to be justified for large values of the tip-to-sample distance and for non-resonant wavelengths. It was hoped that less favourable situations could be handled within the first Born approximation. However, all experiments conducted with various devices confirmed that improving the resolution requires approaching the tip very close to the sample. Adding to this fact the recent development towards near-field spectroscopy and the exploration of radiation pressure effects, it became clear that a realistic computation should at least take the tip–sample coupling properly into account. Various numerical techniques were then applied in order to include this coupling successfully.

2. The finite-difference time domain scheme

Originally developed in the context of radar scattering in aeronautics, this method is a Maxwell's equations solver derived from the finite-element method. This purely numerical

scheme was recently applied to near-field optical problems (Kann *et al* 1995a, b). The technique directly solves the time-dependent Maxwell equations. This feature imposes time averaging over a period in order to consider harmonically oscillating fields which accurately model the time-dependence of the laser used in near-field optics. Typically, such a procedure requires a supercomputer in order to assess even relatively simple problems. Up to now, it has been able to reproduce some results found previously by the methods that we detail in sections 3 and 4.

3. Theories based on matching boundary conditions

Up until now, practically all applications of near-field optics have been running under stationary laser illumination. This experimental mode allows one to restrict the theoretical description to electromagnetic fields which depend harmonically on time. With this $\exp(-i\omega t)$ time-dependence, the Maxwell equations in the absence of any external source read (SI units)

$$\nabla \cdot \epsilon(\boldsymbol{r}, \omega)\boldsymbol{E}(\boldsymbol{r}, \omega) = 0 \tag{5}$$

$$\nabla \cdot \boldsymbol{B}(\boldsymbol{r}, \omega) = 0 \tag{6}$$

$$\nabla \times \boldsymbol{E}(\boldsymbol{r}, \omega) = i\omega\mu_0\boldsymbol{H}(\boldsymbol{r}, \omega) \tag{7}$$

$$\nabla \times \boldsymbol{H}(\boldsymbol{r}, \omega) = -i\omega\epsilon_0\epsilon(\boldsymbol{r}, \omega)\boldsymbol{E}(\boldsymbol{r}, \omega). \tag{8}$$

This set of equations is the starting point of a *macroscopic* approach of near-field optical phenomena where the response of matter to exciting electromagnetic fields is described by the dielectric function $\epsilon(\boldsymbol{r}, \omega)$. Roughly speaking, the dielectric function, which is equal to the square of the complex index of refraction, allows one to model the response of a large number of atoms to an external electric field. It is physically meaningful to address problems where the scatterers' size is large enough to justify the use of such a global property. In visible light near-field optics, it allows one to model the response of mesoscopic and nanoscopic objects larger than about 10 nm.

A first class of numerical methods follows the traditional approach of matching electromagnetic boundary conditions at interfaces. Such methods are typically based on well-established techniques previously developed for other purposes. Solutions are written as linear expansions of a set of eigenfunctions where the coefficients are the unknowns to be found numerically (Van Labeke and Barchiesi 1992, 1993a, b, Barchiesi and Van Labeke 1993, 1994, 1995, Sentenac and Greffet 1992, Bernsten *et al* 1993).

3.1. Expansion in plane waves: grating and diffraction theory

Building on the results of the grating theory, the expansions in Fourier series were proposed by Van Labeke and Barchiesi (1992) in order to model near-field phenomena above gratings. It is well suited to the study of periodic dielectric surface profiles. Such samples were frequently used some years ago for testing the resolution limit of near-field optical microscopes. In order to account for non-periodic and well localized scatterers, Van Labeke and Barchiesi (1993) later used the expansion in a continuum of plane waves typical of diffraction theory which we summarize below.

3.1.1. Diffracted field around surface corrugations. The method starts from the totally reflected electromagnetic wave $\boldsymbol{E}_0(\boldsymbol{r}, \omega)$ incident on a glass–air plane interface defined by

$z = 0$. The diffracted field $E_d(r, \omega)$ due to the surface corrugations is determined within the Rayleigh hypothesis by assuming the following plane wave expansion,

$$E_d(r, \omega) = \iint dk_x \, dk_y \, F_d(k, \omega) e^{ik \cdot r} \tag{9}$$

where $k = (k_x, k_y, k_z)$ represents the cartesian components of the different wave vectors associated with the field diffracted by the corrugated surface and $r = (x, y, z)$. The two-fold integral runs over k_x and k_y. Since the dispersion relation (4) must hold, the diffracted field generally contains both radiative and evanescent waves. Indeed, according to the Heisenberg uncertainties introduced in section 1.3, we know that the dispersion Δk_s (where $s = x, y$) is directly related to the lateral size of the surface corrugations with respect to the incident wavelength λ. When we deal with a sample that displays subwavelength details, the Fourier expansion (9) is mainly composed of evanescent components so that the resulting diffracted field $E_d(r, \omega)$ turns out to be confined around the surface corrugations.
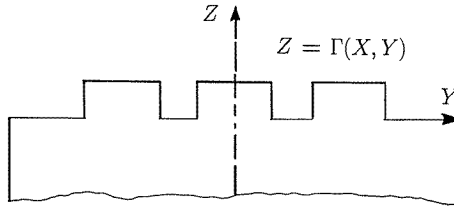

Figure 1. Schematic illustration of a sample limited by an arbitrary surface corrugation function $Z = \Gamma(x, y)$. In the simulation described in figures 2 and 3 this system is illuminated in total internal reflection (TIR).

3.1.2. Perturbative diffraction theory. The evaluation of the field amplitudes $F(k, \omega)$ is known as a rather difficult task. It requires the introduction of the Fourier transform of the function describing the surface profile $\Gamma(x, y)$ (cf figure 1):

$$\Gamma(x, y) = \iint dk_x \, dk_y \gamma(k_x, k_y) e^{i(k_x x + k_y y)}. \tag{10}$$

The application of the standard boundary conditions at the surface $z = \Gamma(x, y)$ leads to a complex relation between the incident and the diffracted field. This difficulty is reduced by working within the perturbative approximation introduced by Agarwal (1977), Toigo *et al* (1977) and Elson (1975) to study the far-field diffraction and the scattering properties of metallic corrugated surfaces. While applying the boundary conditions, this approximation consists of expanding the exponential function contained in (9) as a power series of $k_z \Gamma(x, y)$:

$$e^{i(k_x x + k_y y + k_z \Gamma(x,y))} = e^{i(k_x x + k_y y)}[1 + ik_z \Gamma(x, y) + \cdots]. \tag{11}$$

For surface corrugations with a weak amplitude, this expansion may be limited to first order. In this case, the diffracted amplitudes are proportional to the Fourier transform of the surface profile $\gamma(k_x, k_y)$ and depend linearly on the zeroth-order field $E_0(r, \omega)$:

$$F_d(k_x, k_y, \omega) \approx i(\epsilon' - \epsilon)\gamma(k_x - q_x, k_y - q_y)\mathbf{A}(k_x, k_y) \cdot E_0(r, \omega) \tag{12}$$

where q_x and q_y represent the (x, y) components of the incident wave vector. In this linear relation $\mathbf{A}(k_x, k_y)$ is the 3×3 transfer matrix defined by

$$\mathbf{A}(k_x, k_y) = \frac{\omega^2/c^2}{k_z + k_z'}\mathbf{1} - \frac{1}{k_z' + \epsilon' k_z}\begin{pmatrix} k_x^2 & k_x k_y & k_x k_z \\ k_y k_x & k_y^2 & k_y k_z \\ k_z' k_x & k_z' k_y & k_z' k_z \end{pmatrix} \tag{13}$$

where **1** represents the identity matrix and k_z' is a z component of the wave vectors diffracted inside the sample characterized by the dielectric function ϵ':

$$k_z'^2 = \frac{\omega^2}{c^2}\epsilon' - k_x^2 - k_y^2. \tag{14}$$

The total optical field generated near the surface protrusions is thus given by a correction to the result associated with a perfectly flat sample:

$$E(r, \omega) \approx E_0(r, \omega) + \mathrm{i}(\epsilon' - \epsilon) \iint \mathrm{d}k_x \, \mathrm{d}k_y \, \exp(\mathrm{i}k \cdot r)$$
$$\times \gamma(k_x - q_x, k_y - q_y)\mathbf{A}(k_x, k_y) \cdot E_0(r, \omega). \tag{15}$$

At this stage, the field may be determined numerically by using standard fast Fourier transform (FFT) routines. As within any scheme working in reciprocal space, the structural information about the object is thus contained in the Fourier transform $\gamma(k_x, k_y)$ of the surface profile $\Gamma(x, y)$. Consequently, for a given observation distance Z_0, the accuracy of the results will depend on the number of spatial harmonics introduced in the FFT.

   The method has a relatively low cost in term of computer time so that it is certainly an interesting tool to help the interpretation of massive amounts of near-field microscope images. Nevertheless, we have to emphasize that such an approximation has a range of validity restricted to surface corrugations of weak amplitude when compared with the incident wavelength. Figures 2 and 3 present two different numerical calculations based on this method. The system is a two-dimensional glass grating illuminated in total internal configuration. We note a strong variation of the optical energy distribution as a function of the polarization mode. For the small grating height investigated here, these results are in fairly good agreement with those issued from the accurate diffraction gratings theory developed by Nevière *et al* (see, for example, Goudonnet *et al* 1995). More recently several other experimental configurations were investigated with similar methods (Barchiesi and Van Labeke 1995, Van Labeke and Barchiesi 1995, Bernsten *et al* 1993). Also, some important issues concerning the problem of image reconstruction (inverse scattering) has been addressed in this context (Garcia and Nieto-Vesperinas 1993, 1994, 1995).

3.2. Expansion in multipoles

Inspired by a technique originally developed for antenna design at longer wavelengths (Hafner and Bomhodt 1993), Novotny *et al* (1994, 1995) and Novotny and Pohl (1995) applied the expansion of the solutions on multipolar eigenfunctions to study near-field optical phenomena.

   These multipolar eigenfunctions $F_n(r, \omega)$ do satisfy the vector wave equation for the eigenvalue q_n:

$$-\nabla \times \nabla \times F_n(r, \omega) + q_n^2 F_n(r, \omega) = 0. \tag{16}$$

According to early studies on the representation of electromagnetic fields in terms of scalar fields (Hansen 1937, Stratton 1941, Green and Wolf 1953, Morse and Feshbach 1953, Bouwkamp and Casimir 1954), they can be constructed from the eigenfunctions $\psi_n(r, \omega)$ of the scalar Helmholtz equation:

$$\nabla^2 \psi_n(r, \omega) + q_n^2 \psi_n(r, \omega) = 0. \tag{17}$$

Since

$$\int \mathrm{d}r \, \psi_n^*(r, \omega)\psi_{n'}(r, \omega) = \delta_{n,n'} \tag{18}$$
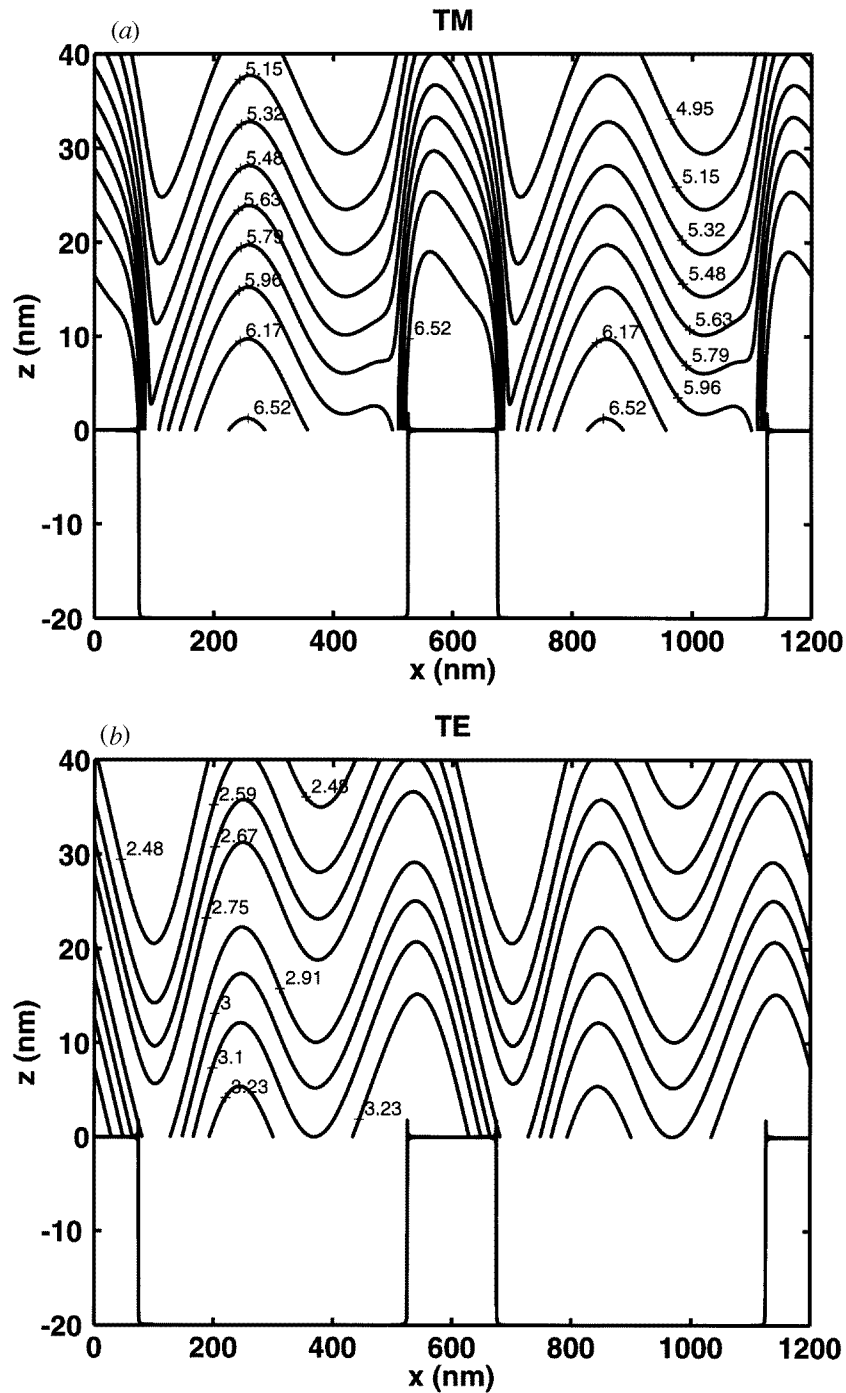
Figure 2. Iso-intensity lines calculated above a lamellar grating. The geometry of the grating is represented by rectangular-shaped surface protrusions. The system is lighted in the TIR configuration with an incident angle $\theta = 45°$, a wavelength $\lambda = 632$ nm and an optical index $n = 1.5$. The calculation has been performed with the perturbative approach as described in section 3.1.2: (*a*) p-polarized illumination mode; (*b*) s-polarized illumination mode (by courtesy of D Barchiesi and D Van Labeke).

these eigenfunctions form an orthonormal basis set in the Hilbert space. The simplest form is obtained in cartesian coordinates:

$$\psi_n(r, \omega) \equiv \psi_k^*(r, \omega) = \frac{1}{\sqrt{8\pi^3}} \exp(i k \cdot r). \tag{19}$$

Three-dimensional multipolar wavefunctions are formulated in spherical coordinates:

$$\psi_n(r, \omega) \equiv \psi_{\sigma, l, m, q_n}(r, \omega) \equiv P_l^m(\cos\theta) z_l(q_n r) \begin{Bmatrix} \cos m\phi \\ \sin m\phi \end{Bmatrix} \tag{20}$$

where $P_l^m(\cos\theta)$ stands for the associated Legendre polynomials and $z_l(q_n r)$ for the spherical Bessel functions. The index σ distinguishes between even (e) and odd (o) functions. Cylindrical symmetry characterizes the multipole functions used in two-dimensional geometries:

$$\psi_n(r, \omega) \equiv \psi_{\sigma, m, q_n}(r, \omega) \equiv B_m(\kappa\rho) \exp(i q_{n,z} z) \begin{Bmatrix} \cos m\phi \\ \sin m\phi \end{Bmatrix} \tag{21}$$

where the B_m are Bessel functions, $\kappa = \sqrt{q_{n,x}^2 + q_{n,y}^2}$ and $\rho = \sqrt{x^2 + y^2}$.

A first family of vector eigenfunctions is found by applying the gradient operator to the scalar functions (C is a normalization factor which depends on the coordinate system):

$$L_n(r, \omega) = C \nabla \psi_n(r, \omega). \tag{22}$$

A second set of eigenfunctions is built as follows,

$$M_n(r, \omega) = C \nabla \times \psi_n(r, \omega) a \tag{23}$$

where a is a constant vector of unit length, sometimes called the 'piloting vector'. The last group of eigenfunctions is given by

$$N_n(r, \omega) = \frac{C}{k} \nabla \times \nabla \times \psi_n(r, \omega) a. \tag{24}$$

Thanks to the property of the piloting vector and to the orthonormalization of the scalar eigenfunctions, the sets $L_n(r, \omega)$, $M_n(r, \omega)$, $N_n(r, \omega)$ are mutually orthogonal in the Hilbert sense. One can easily prove that the three above sets of vector eigenfunctions are sufficient to build the following completeness relationship valid for an infinite homogeneous system:

$$\sum_n [L_n(r, \omega) L_n^*(r', \omega) + M_n(r, \omega) M_n^*(r', \omega) + N_n(r, \omega) N_n^*(r', \omega)] = \mathbf{1}\delta(r - r'). \tag{25}$$

The first family of eigenfunctions, $L_n(r, \omega)$, are longitudinal eigenfunctions which correspond to physical solutions of the wave equation only if $k = 0$. This is only possible if $\epsilon = 0$, which occurs at longitudinal optical frequencies in polar materials or at the plasma frequency in metals. For other frequencies, these eigenfunctions have no physical meaning even if they are mathematically required to build the completeness relationship. Therefore, after dividing the space into subdomains where the index of refraction is constant, the multiple multipole method performs the expansion of the electric field in each subdomain α only on the sets of the transverse eigenfunctions $M_n(r, \omega)$ and $N_n(r, \omega)$:

$$E^\alpha(r, \omega) = \sum_n [a_n^\alpha M_n(r, \omega) + b_n^\alpha N_n(r, \omega)]. \tag{26}$$
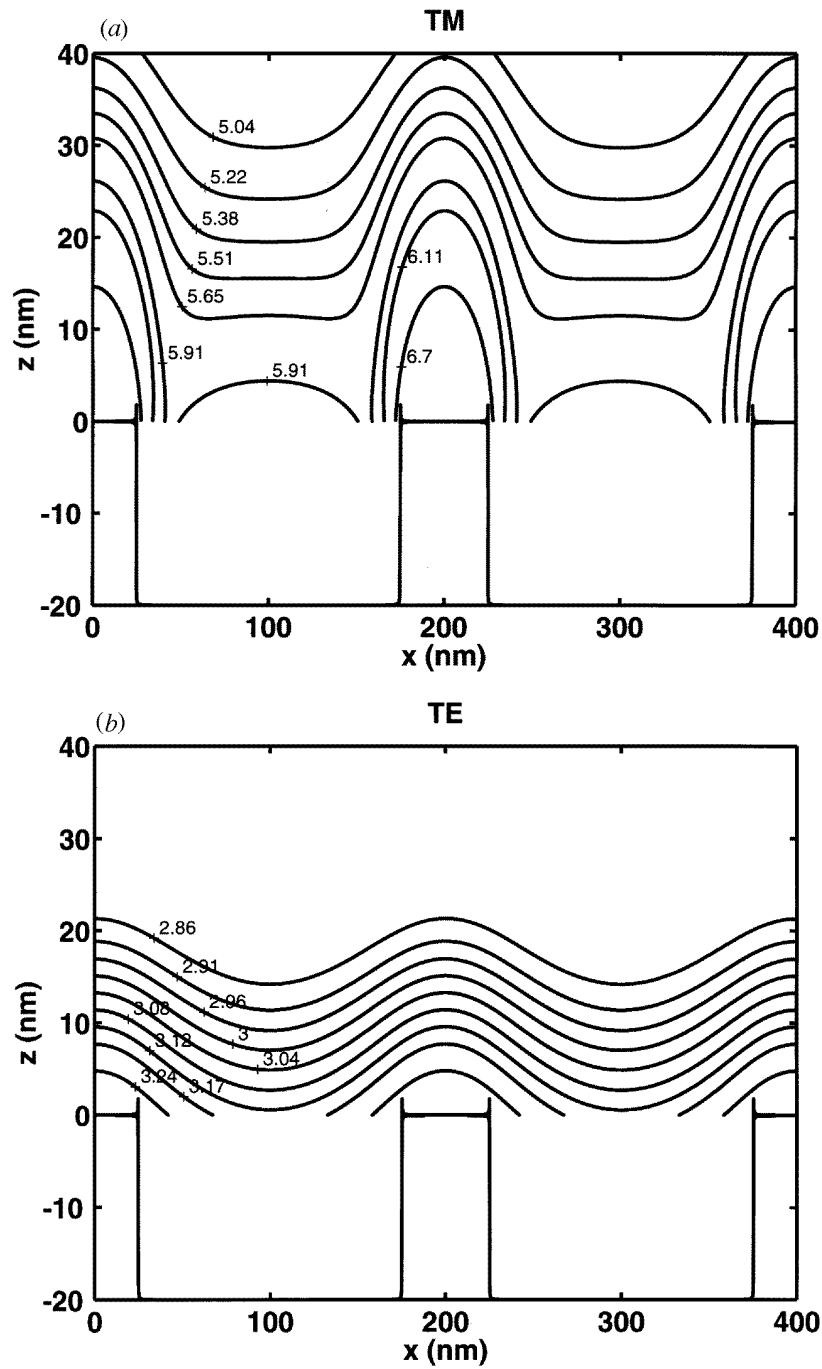
Figure 3. The same as figure 2, but with a smaller lateral extension of the surface structures (by courtesy of D Barchiesi and D Van Labeke).

In order to remedy the poor convergence of the above expansion for geometries far from the spherical (respectively cylindrical) geometry, the principle of the multiple multipole

method consists of using several different origins r_j in the multipole expansion:

$$\boldsymbol{E}^\alpha(\boldsymbol{r}, \omega) = \sum_j \sum_n [a^\alpha_{n,j} \boldsymbol{M}_n(\boldsymbol{r} - \boldsymbol{r}_j) + b^\alpha_{n,j} \boldsymbol{N}_n(\boldsymbol{r} - \boldsymbol{r}_j)]. \tag{27}$$

The unknowns coefficients $a^\alpha_{n,j}$ and $b^\alpha_{n,j}$ are then found from the electromagnetic boundary conditions on the interfaces between adjacent subdomains by a least-squares minimization. This optimization requires one to discretize the curves describing the interfaces. This discretization has some impact on the highest possible degree of a multipole centre.

Of course, due to the splitting of the geometry in subdomains and to the existence of several origins r_j, completeness and orthonormalization relationships over the entire composite system are not achieved anymore. Therefore, to avoid mutual dependences, the method relies on semi-empirical rules to fix the separation between the origins r_j.

The multipole functions used in the basis set of the multiple multipole method are rather short range so that they affect their close neighbourhood. The method is thus better suited to account for localized geometries than the expansion in plane waves. It is also well designed to describe complicated structures such as cylindrical waveguides coated with a realistic metal whose dielectric function is complex (Novotny and Hafner 1994) or the scattering inside sharpened cylindrical tips used in near-field microscopy. From a mathematical point of view, the idea of spreading a set of multipole functions and adjusting the coefficients is similar to quantum mechanical techniques used for computing electronic structures such as the linear combination of atomic orbitals (LCAO). The multipoles in electrodynamics thus play a role similar to the atomic orbitals in quantum mechanics. However, the situation is somewhat clearer in electron physics where the atomic orbitals are centred on each nucleus, whereas a physical meaning is lacking in the mathematical procedure (Regli 1992) which distributes the coordinates of the multipoles centres in electrodynamics.

Figures 4 and 5 display examples of numerical simulations performed using the MMP methods (Novotny and Pohl 1995).

4. Scattering theory

From a mathematical point of view, scattering theory (also known as the field susceptibility or Green dyadic technique) is developed upon the Green function theory applied to the wave equation where a source term is introduced. It simply casts the most general analytical solution of the inhomogeneous wave equation as an integral equation where the kernel is a Green function (Newton 1966, Keller 1986, 1988a, b). Scattering theory has been used for a long time in quantum mechanical problems. In electrodynamics, it has also been exploited extensively to solve engineering problems involving external sources of currents where the solution in the source region is not required. The application to the study of phenomena where the solution in the source region (such as in near-field optics) is of primary importance was hindered by the apparently divergent nature of the Green dyadic in the source region. As will be shown in section 4.3, this divergence is related to depolarization effects. These are well understood today so that an unambiguous renormalization procedure is available.

Several variants of electromagnetic scattering theory where applied successfully to the modelization of near-field optical phenomena. Although the Green function may be expanded in Fourier or multipoles series, most variants preferred a discretization in the direct space since near-field optical phenomena occur on a subwavelength scale.
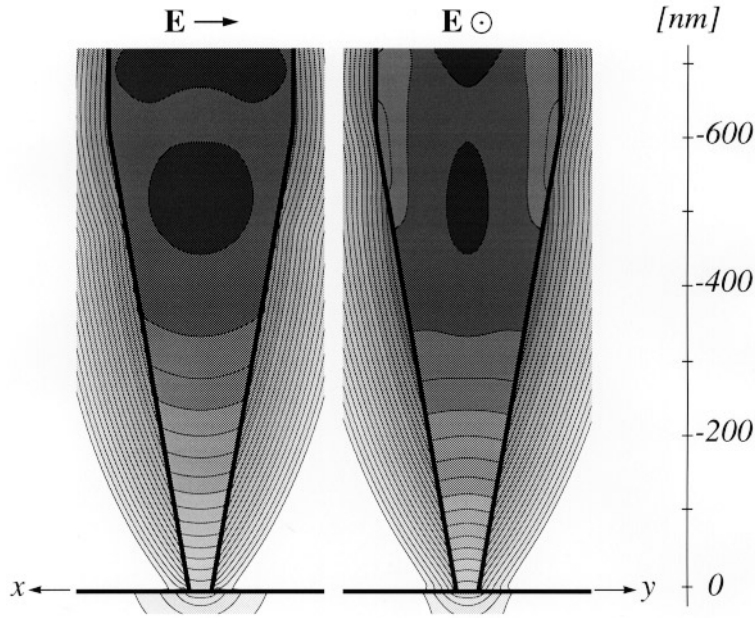
Figure 4. Examples of 3D near-field calculations performed with the MMP method (by courtesy of L Novotny (Novotny and Pohl 1995)). The physical system is the tip of an aperture SNOM consisting of a cylindrical part and a tapered part. The probe is excited by a waveguide mode of wavelength $\lambda = 488$ nm and there is a factor three between two successive lines.

4.1. Propagator or Green dyadic

This section outlines how the Green dyadic (or field susceptibility) method associated with the localized perturbation theory deals efficiently with the resolution of the self-consistent optical tip–sample interaction. A detailed account of this theory and its numerical implementation may be found in (Girard *et al* 1993, Dereux and Pohl 1993, Martin *et al* 1995a). A discussion of the convergence and stability of the algorithm is reported in Martin *et al* (1994). Therefore, this section will focus more on the physical content of the formulae introduced in the above-mentioned references.

With the usual $\exp(-i\omega t)$ time-dependence, the vector wave equation issued from Maxwell's equations (SI units),

$$- \nabla \times \nabla \times \boldsymbol{E}(\boldsymbol{r}, \omega) + \frac{\omega^2}{c^2} \epsilon(\boldsymbol{r}, \omega) \boldsymbol{E}(\boldsymbol{r}, \omega) = 0 \tag{28}$$

may be cast as

$$- \nabla \times \nabla \times \boldsymbol{E}(\boldsymbol{r}, \omega) + q^2 \boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{V}(\boldsymbol{r}, \omega) \boldsymbol{E}(\boldsymbol{r}, \omega) \tag{29}$$

with

$$q^2 = \frac{\omega^2}{c^2} \epsilon_{\text{ref}}. \tag{30}$$

Any complicated behaviour due to the anisotropy or to the low symmetry of the geometrical shape of the original dielectric tensor profile $\epsilon(\boldsymbol{r}, \omega)$ is described as a difference relative to the reference system ϵ_{ref}:

$$\boldsymbol{V}(\boldsymbol{r}, \omega) = \frac{\omega^2}{c^2} (\mathbf{1}\epsilon_{\text{ref}} - \epsilon(\boldsymbol{r}, \omega)). \tag{31}$$
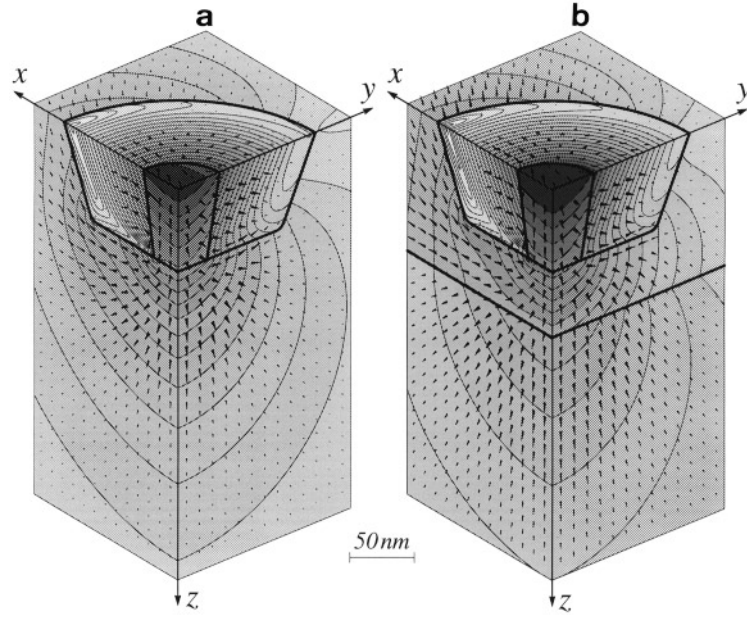
Figure 5. Contours of constant $|E|^2$ on three perpendicular planes near the aperture of the SNOM probe described in figure 4 (by courtesy of L Novotny (Novotny and Pohl 1995)). The arrows indicate the time-averaged Poynting vector. The polarization is in the plane $y = 0$. The transmission through the probe is increased when a dielectric substrate is approached (*b*).

As introduced in Dereux and Pohl (1993) and Girard *et al* (1993), the solution of (29) is obtained from the implicit Lippmann–Schwinger equation

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}_0(\boldsymbol{r}, \omega) + \int_D \mathrm{d}\boldsymbol{r}' \, \boldsymbol{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) \boldsymbol{V}(\boldsymbol{r}', \omega) \boldsymbol{E}(\boldsymbol{r}', \omega). \tag{32}$$

In scattering theory, the first term $\boldsymbol{E}_0(\boldsymbol{r}, \omega)$ is referred to as the incident field while the second term is called the scattered field obtained from the integration over the domain D where $\boldsymbol{V}(\boldsymbol{r}', \omega)$ is non-zero. D defines the volume of the scatterer relative to the reference system. Electromagnetic theory traditionally qualifies D as the source region.

To solve the Lippmann–Schwinger equation, we need to know the analytical solution $\boldsymbol{E}_0(\boldsymbol{r}, \omega)$ satisfying

$$-\nabla \times \nabla \times \boldsymbol{E}_0(\boldsymbol{r}, \omega) + q^2 \boldsymbol{E}_0(\boldsymbol{r}, \omega) = 0 \tag{33}$$

and the associated Green dyadic defined by

$$-\nabla \times \nabla \times \boldsymbol{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) + q^2 \boldsymbol{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) = \mathbf{1}\delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{34}$$

The reference structure ϵ_{ref} is usually a homogeneous background material or a semi-infinite surface system. For homogeneous media, the analytical form of $\boldsymbol{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega)$ is known from early studies (Morse and Feshbach 1953, Levine and Schwinger 1950). Its general form may be deduced as follows. We first seek a solution (temporarily written as $\boldsymbol{G}_0'$ for a reason which will appear later) of (34) as an expansion over the vector eigenfunctions already defined in the previous section when devising the expansion in multipoles:

$$\boldsymbol{G}_0'(\boldsymbol{r}, \boldsymbol{r}', \omega) = \sum_n \boldsymbol{L}_n(\boldsymbol{r}, \omega) \boldsymbol{X}_n^*(\boldsymbol{r}', \omega) + \boldsymbol{M}_n(\boldsymbol{r}, \omega) \boldsymbol{Y}_n^*(\boldsymbol{r}', \omega) + \boldsymbol{N}_n(\boldsymbol{r}, \omega) \boldsymbol{Z}_n^*(\boldsymbol{r}', \omega). \tag{35}$$

According to the completeness relationship (25), we can find out the unknowns $X_n^*(r', \omega)$, $Y_n^*(r', \omega)$ and $Z_n^*(r', \omega)$ by backsubstition of (35) in (34), so that

$$G_0'(r, r', \omega) = \sum_n \frac{L_n(r, \omega)L_n^*(r', \omega) + M_n(r, \omega)M_n^*(r', \omega) + N_n(r, \omega)N_n^*(r', \omega)}{q^2 - q_n^2}. \quad (36)$$

The above expansion is known as the spectral expansion of the Green dyadic. However, as mentioned in the previous section, the longitudinal wavefunctions are not physical in a homogeneous medium. One might think that discarding the longitudinal wavefunctions out of the spectral expansion would provide the appropriate cure in order to find $G_0(r, r', \omega)$. This would be a mistake since it does not take care of the singular behaviour of $G_0(r, r', \omega)$ as $r \to r'$. Indeed, when introduced in (34), the resulting Green dyadic expansion G_0^t containing only transverse eigenfunctions cannot build up the longitudinal part of $1\delta(r - r')$ which exists if $r \to r'$. Indeed, if

$$G_0^t(r, r', \omega) = \sum_n \frac{M_n(r, \omega)M_n^*(r', \omega) + N_n(r, \omega)N_n^*(r', \omega)}{q^2 - q_n^2} \quad (37)$$

we have

$$-\nabla \times \nabla \times G_0^t(r, r', \omega) + q^2 G_0^t(r, r', \omega) = D_t(r - r') \quad (38)$$

where

$$D_t(r - r') = \sum_n [M_n(r, \omega)M_n^*(r', \omega) + N_n(r, \omega)N_n^*(r', \omega)]. \quad (39)$$

A longitudinal part must then be added to G_0^t in order to build G_0:

$$G_0(r, r', \omega) = G_0^t(r, r', \omega) + G_0^l(r, r', \omega). \quad (40)$$

G_0^l must satisfy

$$-\nabla \times \nabla \times G_0^l(r, r', \omega) + q^2 G_0^l(r, r', \omega) = D_l(r - r') \quad (41)$$

where

$$D_l(r - r') = \sum_n [L_n(r, \omega)L_n^*(r', \omega)]. \quad (42)$$

The longitudinal character of G_0^l allows one to conclude that

$$G_0^l(r, r', \omega) = \frac{1}{q^2} D_l(r - r') \quad (43)$$

so that

$$G_0(r, r', \omega) = G_0^t(r, r', \omega) + G_0^l(r, r', \omega). \quad (44)$$

In Cartesian coordinates, the normalization factor in the vector wavefunctions is given by $C = k^{-1}$. Some algebra leads to

$$G_0(r, r', \omega) = \int dk \left[1 - \frac{1}{q^2}kk\right] \frac{e^{ik\cdot(r-r')}}{8\pi^3(q^2 - k^2)}. \quad (45)$$

Integration by the method of residues yields

$$G_0(r, r', \omega) = \left[1 - \frac{1}{q^2}\nabla\nabla\right] g(r, r', \omega) \quad (46)$$

where $g(r, r', \omega)$ is the Green function associated with the scalar Helmholtz equation. $g(r, r', \omega)$ is given by a spherical wave emitted at r' (Economou 1983):

$$g(r, r', \omega) = \frac{\exp(iq|r - r'|)}{4\pi|r - r'|}. \tag{47}$$

The last two equations are interpreted as follows. The dyadic Green function is not the direct extension of the scalar Green function. Indeed, at large distances (mathematically at infinite distances), one must pay attention to satisfying the Sommerfeld radiation condition which states that the electromagnetic fields are purely transverse. The negative terms inside the brackets of the last two equations take care of discarding any longitudinal components of the field at infinity.

For a surface system, the expression of the propagator is somewhat more elaborated (Agarwal 1975, Metiu 1984, Girard and Bouju 1992).

4.2. The Huygens–Fresnel principle revisited

The implicit character of this Lippmann–Schwinger equation ensures the self-consistency of the solution. This may be understood by revisiting the Huygens–Fresnel principle. The various Kirchhoff theories are themselves a mathematical translation of the Huygens–Fresnel principle. After an abstract discretization of the scatterer into elementary sources, this fundamental principle states that the scattering of light may be explained by the coherent sum of the fields emitted by each elementary source. It is easily checked that this basic feature is embedded into (32) if we assume that we know the value of the field attributed to each elementary source inside D. Moreover, one should notice that Huygens described the field emitted by each elementary source as spherical waves similar to (47). Relatively to the Lippmann–Schwinger equation (32), the different look of Kirchhoff integrals originates simply from the use of the Green theorem to transform the volume integrals into surface integrals.

One of the basic approximations in classical diffraction theories substitutes the incident field $E_0(r', \omega)$ into the values of the field in the source region $E(r', \omega)$. This procedure to solve (32) is known as the first Born approximation in scattering theory. The improvement brought by the self-consistent solution is thus related to the accurate determination of the field value in the source region. We discussed in Girard and Bouju (1991) and Dereux *et al* (1991) how the discretization of the Lippmann–Schwinger equation (32) allowed one to obtain the total field in the source region through the solution of

$$\sum_j [\mathbf{1}\delta(r_i - r_j) - G_0(r_i, r_j, \omega)V(r_j, \omega)w_j]E(r_j, \omega) = E_0(r_i, \omega). \tag{48}$$

where w_j is the volume (surface or length according to the dimensionality of the problem) of the mesh. In this linear system of equations, all r_i and r_j belong to the source region so that computation of the diagonal elements $G_0(r_i, r_i, \omega)$ is required. At first sight, this may cause some trouble since $G_0(r_i, r_j, \omega)$ looks singular in equation (46) when $r_i \to r_j$. A proper renormalization procedure to be explained in the next section is needed to resolve this singularity.

4.3. Renormalization procedure and the depolarizing dyadic

In the derivation of equation (32) from the vector wave equation, the commutation of the integration over the perturbation domain $\int_D dr'$ and the application of $-\nabla \times \nabla \times$ was assumed. Although this permutation is regular in most cases, it is not correct when the

integrand is singular like a spherical wave (Yaghjian 1980). In the source region, when $r \to r'$, the interchange of the linear operators $\int_D dr'$ and $-\nabla \times \nabla \times$ in front of a cylindrical or a spherical wavefunction supplies a point source term as proved in classical textbooks (Born and Wolf 1964). Moreover, this extra source term depends critically on the shape of the mesh used in the discretization procedure. This somewhat confusing feature has hindered the use of the Green dyadic technique in numerical applications. Yaghjian clarified this matter and tabulated the corrective terms for various discretization meshes (Yaghjian 1980). We now introduce this correction by discussing further the physical meaning of the self-consistent step which occurs in the numerical scheme. In order to be more intuitive, let us imagine that the perturbation is made of a single discretization mesh centred around r_i. Equation (48) allows one to find the electric field inside this mesh $E(r_i, \omega)$ as a function of the applied field $E_0(r_i, \omega)$:

$$[1 - G_0(r_i, r_i, \omega)V(r_i, \omega)w_i]E(r_i, \omega) = E_0(r_i, \omega). \tag{49}$$

Due to the finite size of the mesh, this evaluation must include the correction arising from the polarization of the mesh. In other words, the self-consistent step (49) must be in agreement with

$$E(r_i, \omega) = E_0(r_i, \omega) - \mathbf{L}_\alpha \frac{P(r_i, \omega)}{\epsilon_0} \tag{50}$$

where $\alpha = 1, 2, 3$ according to the dimensionality of the problem and $P(r_i, \omega)$ is the polarization vector which is related to the dielectric property of the mesh through

$$\frac{P(r_i, \omega)}{\epsilon_0} = [\epsilon(r_i, \omega) - 1\epsilon_{\text{ref}}]E(r_i, \omega). \tag{51}$$

The resulting field depends critically on the depolarizing dyadic $\mathbf{L}_\alpha$ associated with the shape of the mesh (Yaghjian 1980, Kittel 1976). Here, we quote only the results which arise in Cartesian coordinates. When addressing the scattering of electromagnetic waves by multilayers, the shape of the discretization mesh is a very thin plate so that

$$\mathbf{L}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{52}$$

In two-dimensional problems, considering infinitely long square rods along the x direction implies the following structure of the depolarizing dyadic:

$$\mathbf{L}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix}. \tag{53}$$

In three-dimensional scattering, the depolarizing dyadic associated with a cubic mesh reads

$$\mathbf{L}_3 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}. \tag{54}$$

Equation (50) evolves then as

$$\left[1 - \frac{c^2}{\omega^2}\mathbf{L}_\alpha V(r_i, \omega)\right] E(r_i, \omega) = E_0(r_i, \omega) \tag{55}$$

which is clearly not included in (49). Indeed, although equation (55) is related to the shape of the mesh, it does not depend on its size w_i whereas the matrix to be inverted in (49) depends on w_i. Thus, for a vanishing value of the mesh size w_i, equation (49) does not

take any depolarization effect into account. In order to include the depolarization properly, we need to modify the definition of the Green dyadic as follows:

$$\mathbf{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) = \left[\mathbf{1} - \frac{1}{q^2}\nabla\nabla\right] g(\boldsymbol{r}, \boldsymbol{r}', \omega) + \mathbf{L}_\alpha \delta(\boldsymbol{r} - \boldsymbol{r}')\frac{c^2}{\omega^2}. \tag{56}$$

The renormalization procedure outlined above is applied to discretize a large piece of continuous matter. It is based on the bulk dielectric properties of the discretized material which are known to be accurate so as to take particles larger than about 10 nm into account (Bohren and Huffman 1983). The introduction of smaller particles (clusters, molecules or even atoms) in the computational scheme is possible provided that its polarizability be known from some other source like an experimental measurement or a quantum mechanical calculation. In this case, the corrective delta function term of (56) must not be included (i.e. $L_\alpha = 0$) since the polarizabilities obtained by these means account for the depolarization effect inside the particles.

Let us point out that the computational procedure introduced in Dereux and Pohl (1993) and Girard *et al* (1993) deliberately avoids any step involving matching boundary conditions. The ability of the method in dealing with scatterers of arbitrary shapes and dielectric responses is of course a fundamental advantage when studying the optical interaction between elongated probe tips and non-planar samples. Not only does this feature enable the handling of arbitrary shaped objects made of continuous matter but, as explained in the preceding paragraph, it also allows one to treat discrete particles in the same framework. This will be a fundamental advantage when addressing the optical interaction of tips with particles whose shapes are intrinsically not sharp. This approach is thus ready to include fuzzy quantum systems (electron gas inside a metallic cluster, molecules or atoms). This is better understood by discussing briefly the relationship between the Green dyadic technique and the concept of field susceptibility (Agarwal 1975).

4.4. Alternative derivation of the Lippmann–Schwinger equation: the field-susceptibility method

Another meaningful equivalent approach may be applied to derive Lippmann–Schwinger equation for optical fields. This alternative derivation is based on the concept of field susceptibility that describes the response of the optical field itself to an infinitesimal fluctuating volume of polarized matter. As illustrated in the fundamental paper of Agarwal (1975), starting from the *microscopic* Maxwell equations expressed in terms of both charge and current densities is a direct route to derive the classical expression of the free-space field susceptibility. Let us consider a physical system characterized by its time-dependent charge density $\rho(\boldsymbol{r}, t)$ and its current density $\boldsymbol{j}(\boldsymbol{r}, t)$. In the ω-space, the Maxwell equations read

$$\nabla \cdot \boldsymbol{E}(\boldsymbol{r}, \omega) = \frac{\rho(\boldsymbol{r}, \omega)}{\epsilon_0} \tag{57}$$

$$\nabla \cdot \boldsymbol{B}(\boldsymbol{r}, \omega) = 0 \tag{58}$$

$$\nabla \times \boldsymbol{E}(\boldsymbol{r}, \omega) = \mathrm{i}\omega\mu_0\boldsymbol{H}(\boldsymbol{r}, \omega) \tag{59}$$

$$\nabla \times \boldsymbol{H}(\boldsymbol{r}, \omega) = -\mathrm{i}\omega\epsilon_0\boldsymbol{E}(\boldsymbol{r}, \omega) + \boldsymbol{j}(\boldsymbol{r}, \omega). \tag{60}$$

The vectorial wave equation for the electric field is readily obtained by taking the curl of equation (59). After some straightforward algebraic manipulations, one gets the well known result

$$\Delta\boldsymbol{E}(\boldsymbol{r}, \omega) + q_0^2\boldsymbol{E}(\boldsymbol{r}, \omega) = \nabla\frac{\rho(\boldsymbol{r}, \omega)}{\epsilon_0} - \mathrm{i}\omega\mu_0\boldsymbol{j}(\boldsymbol{r}, \omega) \tag{61}$$

where $q_0 = \omega/c$.

We now express both charge and current densities in terms of the local polarization $\boldsymbol{P}(\boldsymbol{r}, \omega)$ of the material system:

$$\rho(\boldsymbol{r}, \omega) = -\nabla \cdot \boldsymbol{P}(\boldsymbol{r}, \omega) \tag{62}$$

and

$$\boldsymbol{j}(\boldsymbol{r}, \omega) = -\mathrm{i}\omega \boldsymbol{P}(\boldsymbol{r}, \omega). \tag{63}$$

We now rewrite the non-homogeneous equation (61) as follows:

$$\Delta \boldsymbol{E}(\boldsymbol{r}, \omega) + q_0^2 \boldsymbol{E}(\boldsymbol{r}, \omega) = q_0^2 \frac{\boldsymbol{P}(\boldsymbol{r}, \omega)}{\epsilon_0} + \nabla \left[\nabla \cdot \frac{\boldsymbol{P}(\boldsymbol{r}, \omega)}{\epsilon_0} \right]. \tag{64}$$

Let $\boldsymbol{E}_0(\boldsymbol{r}, \omega)$ be the solution of the following homogeneous equation,

$$\Delta \boldsymbol{E}_0(\boldsymbol{r}, \omega) + q_0^2 \boldsymbol{E}_0(\boldsymbol{r}, \omega) = 0 \tag{65}$$

the general solution of (64) is the sum of the homogeneous field $\boldsymbol{E}_0(\boldsymbol{r}, \omega)$ plus a particular solution $\boldsymbol{E}_\mathrm{m}(\boldsymbol{r}, \omega)$. This particular solution can be derived from the knowledge of the free-space scalar Green function (47)

$$\boldsymbol{E}_\mathrm{m}(\boldsymbol{r}, \omega) = \int \boldsymbol{S}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) \cdot \boldsymbol{P}(\boldsymbol{r}', \omega) \, \mathrm{d}\boldsymbol{r}' \tag{66}$$

where $\boldsymbol{S}_0(\boldsymbol{r}, \boldsymbol{r}', \omega)$ defines the free-space dyadic field susceptibility

$$\boldsymbol{S}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) = (q_0^2 + \nabla\nabla)g(\boldsymbol{r}, \boldsymbol{r}', \omega). \tag{67}$$

At this stage it is interesting to note that the free space Lippmann–Schwinger equation can be deduced from these last relations. First, we write the complete solution of equation (64) as the sum of both homogeneous and inhomogeneous solutions:

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}_0(\boldsymbol{r}, \omega) + \boldsymbol{E}_\mathrm{m}(\boldsymbol{r}, \omega) \tag{68}$$

Second, we introduce the usual constitutive equation for a local medium:

$$\frac{\boldsymbol{P}(\boldsymbol{r}, \omega)}{\epsilon_0} = \chi(\boldsymbol{r}, \omega) \cdot \boldsymbol{E}(\boldsymbol{r}, \omega). \tag{69}$$

Finally, by substituting this expression into (66), and the resulting formula into (68), we find the implicit equation

$$\boldsymbol{E}(\boldsymbol{r}, \omega) = \boldsymbol{E}_0(\boldsymbol{r}, \omega) + \int \boldsymbol{S}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) \cdot \chi(\boldsymbol{r}', \omega) \cdot \boldsymbol{E}(\boldsymbol{r}', \omega) \, \mathrm{d}\boldsymbol{r}' \tag{70}$$

which recovers the result of section 4.1 for an arbitrary particle in free space. Indeed, the kernels appearing in equations (32) and (70) are identical,

$$\boldsymbol{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega)V(\boldsymbol{r}', \omega) = \boldsymbol{S}_0(\boldsymbol{r}, \boldsymbol{r}', \omega)\chi(\boldsymbol{r}', \omega). \tag{71}$$

since, in the case of *continuous matter* where the dielectric susceptibility $\chi(\boldsymbol{r}', \omega)$ is related to the dielectric function according to

$$\chi(\boldsymbol{r}', \omega) = \epsilon(\omega) - 1 \tag{72}$$

which is consistent with the fact that a simple constant factor is the only difference between the Green dyadic and the field susceptibility:

$$\boldsymbol{S}_0(\boldsymbol{r}, \boldsymbol{r}', \omega) = -q_0^2 \boldsymbol{G}_0(\boldsymbol{r}, \boldsymbol{r}', \omega). \tag{73}$$

Now, if we introduce a somewhat more complicated surrounding such as, for example, the presence of an extended medium (the surface of a semi-infinite material, a macroscopic

sized particle and so on), we just have to replace the free-space dyadic $S_0(r, r', \omega)$ by the following (Girard and Bouju 1991, Keller *et al* 1992, 1993),

$$S(r, r', \omega) = S_0(r, r', \omega) + S_s(r, r', \omega) \tag{74}$$

where the new contribution $S_s(r, r', \omega)$ accounts for the dynamical response of such extended system.

The factorization on the right-hand side of (71) was originally introduced to deal with atoms or molecules adsorbed on a surface. Indeed, if one considers, for example, a system formed of p individual molecules, $\chi(r, \omega)$ can be expressed as

$$\chi(r', \omega) = \sum_{i=1}^{p} \alpha^{(i)}(\omega)\delta(r' - r_i) + \cdots \tag{75}$$

where r_i represents the position vectors of the molecules and $\alpha^{(i)}(\omega)$ defines their optical dipolar polarizabilities. For more details about this factorization and its application to molecular near-field optics, we refer the reader to section 6.

However, the most important feature of all variants of scattering theory is the correct description of the self-consistent coupling between all scatterers which raises the numerical implementation to the rank of an accurate predictive procedure. In order to illustrate this versatility, we review the results of numerical applications to the study of optical tip–sample interactions arising in different contexts: optical near-field distributions, near-field spectroscopy and radiation pressure effects.

4.5. Near-field distributions

Near-field distributions obtained by scattering theoretical calculations where the discretization procedure is applied on pertubation volumes have been obtained by several groups (Girard *et al* 1993, Greffet *et al* 1995, Carminati and Greffet 1995a, b, Martin *et al* 1995a, b).

The traditional way of exploiting scattering theory by transforming volume integrals into surface integrals has been applied to near-field optical problems by Carminati *et al* (1994) and Nieto-Vesperinas and Madrazo (1995). These authors established an exact numerical method that can account for the interaction of the near-field scattered by a rough surface placed in an interaction with a local probe.

Our first numerical applications of the Green function technique will be presented in the framework of the total internal reflection configuration. In this mode, the transparent glass surface bearing the object (surface defect) is illuminated by a monochromatic optical field of frequency ω_0 so that total reflection may occur at the surface of the material. From the experimental point of view, this illuminated mode corresponds to the STOM/PSTM configuration (Reddick *et al* 1989, Courjon *et al* 1989, 1990, Adam *et al* 1993, van Hulst *et al* 1992, 1993, Jiang *et al* 1991). Let us consider the three-dimensional (3D) localized surface protrusion schematized in figure 6. Depending on its own dimensions with respect to the incident wavelength, such an object will behave as a more or less efficient obstacle to the propagation of the surface wave. The behaviour of the normalized field intensity defined by

$$I = \frac{|E|^2}{|E_0|^2} \tag{76}$$

is described in figure 7. In this ratio, $|E|^2$ represents the intensity of the self-consistent field and $|E_0|^2$ the intensity of the incident field (i.e. in the absence of the surface defect).

The calculation of I is performed in an observation plane parallel to the reference system, placed a distance $Z_0 = 70$ nm above the square-shaped surface defect. This observation plane is introduced merely for the convenience of data visualization, since the field can be computed for any arbitrary position outside and inside the system by using the Lippmann–Schwinger equation (70). In the internal reflection set-up described in figure 6, the zeroth-order solution $E_0(r, \omega)$ is the evanescent field created by total reflection at the surface $Z = 0$. Such an illumination configuration eliminates propagating waves along the z direction. It may be seen in figure 7 that the structure of the large scale 2000×2000 nm^2 calculated image is very complex. It displays a complicated standing field pattern currently observed in the STOM configuration, namely strong confined field effects observed just above the scatterer, scattering along its lateral sides and usual interference patterns due to the interaction between the travelling surface wave and the reflected wave by the surface defect. In particular, this 3D perspective view indicates that at 10 nm from the top of the object the enhancement factor of the field intensity reaches 2.1. As described in the experimental works of van Hulst and collaborators, for larger objects generally the interference phenomenon dominates and makes the detection of subwavelength features difficult (van Hulst *et al* 1992, 1993).
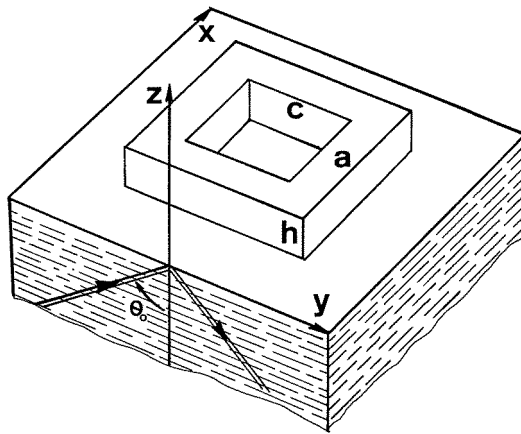

Figure 6. Perspective representation of the square-shaped surface protrusion used in the simulation of figure 7. The system, of optical index $n = 1.5$, is illuminated in total reflection, θ_0 represents the incident angle and the incident wavelength in vacuum is equal to 620 nm. Three geometrical parameters a, c and h have been introduced to define the spatial extension of the protrusion: a represents the external length of the pattern, c the length of the internal side and h its height.

4.6. Transition from the mesoscopic to the nanoscopic regime: confinement of light near surface defects

We have proved with the simulation of figure 7 that, first, the electromagnetic field is more and less confined around surface structures and, second, that the relation between the object profile and the resulting spatial field distribution may be very complex. In fact, the study of the gradual transition between mesoscopic and nanoscopic regimes is of interest for experimentalists working in SNOM since it might allow one to precisely find the fundamental difference between pure topographic signals and artefacts originating from interference and scattering phenomena. In order to get more insight about this important
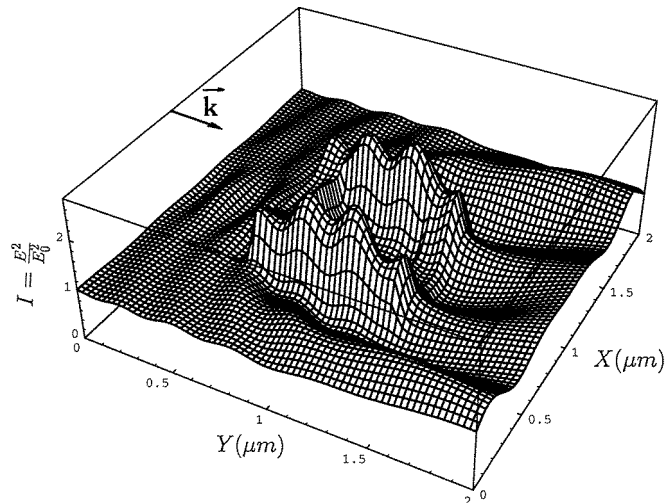
Figure 7. Perspective view of the normalized electric field intensity $I(X, Y) = |E|^2/|E_0|^2$ calculated above the scattering system schematized in figure 6. The sizes of the object are mesoscopic: $a = 750$ nm, $c = 450$ nm and $h = 60$ nm. The calculation is performed in the p-polarized mode. For the convenience of the data visualization, the numerical data $I(X, Y)$ have been calculated in the observation plane located at distance $Z_0 = 70$ nm from the flat surface.
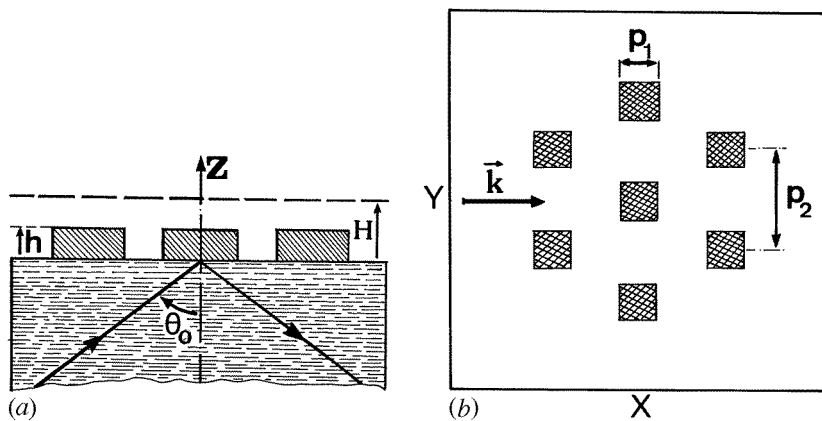

Figure 8. Geometry of a 3D object composed of several identical square-shaped glass protrusions. The dielectric parameters are the same than those used in the previous application (cf figure 6). The centre of each pad is located at the nodes of an hexagonal pattern of side P_2. P_1 represents the dimension of each individual protrusion. The system is illuminated in internal reflection configuration and k represents the surface wave vector. The height of the pads is h (from Girard *et al* 1995a); (*a*) side view, (*b*) top view.

question, we consider a second application with a more complex system composed of seven identical square-shaped pads (cf figure 8). The dielectric parameters are the same as those used in the previous application.

We present in figure 9 a first simulation by illuminating this system in TM polarization (see figure 8). Each dielectric pad is 100 nm high and has a section of 0.25×0.25 μm^2 and the calculation is performed in the plane $Z_0 = 120$ nm. In order to emphasize both
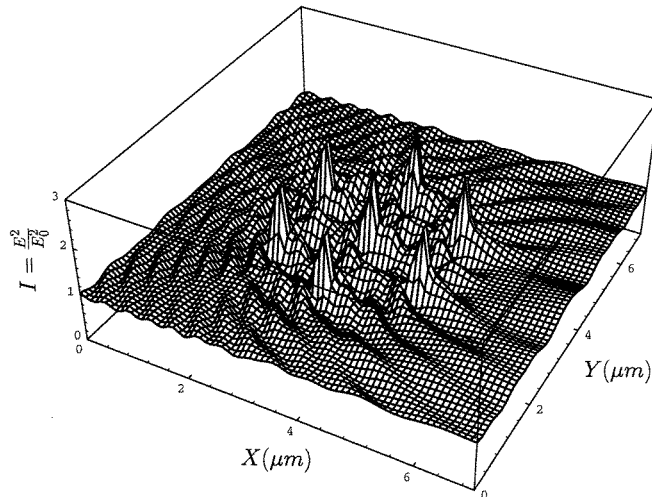
Figure 9. 3D perspective view of the normalized electric field intensity $I(X, Y) = |E|^2/|E_0|^2$ calculated above the scattering system schematized in figure 8. The observation plane is located at a distance $Z_0 = 120$ nm above the flat surface. This large scale calculation 7×7 μm^2 has been performed in the p-polarized mode. The parameters P_1 and P_2 are equal to 250 nm and 1750 nm, respectively, and the pads are 100 nm high (from Girard *et al* 1995a).

interference and scattering effects occurring around the obstacles, we have used a large computational window 7×7 μm^2. Due to the large spacing between each individual scatterer (1.75 μm), the resulting field pattern is a complex mixture of interference phenomena due to multiple reflections between the different pads. As expected, when the number of defects per unit area increases, the standing wave pattern arising from the multiple scattering effects gives rise to the well known 'speckle pattern' phenomenon.

Now, we study in figure 10 the evolution of the image upon reduction of the different geometrical parameters P_1, P_2 and H defined in figure 8. The two commonly used polarization modes TE and TM are simultaneously considered in figure 10. Three different typical sizes are successively investigated. In the first example (figure 10(*a*)) we start in the mesoscopic range ($P_1 = 250$ nm, $P_2 = 1750$ nm and $h = 100$ nm). In the two other examples (figures 10(*b*) and (*c*)), a reduction factor equal to two and four is applied and the position of the observation plane is reduced in the same proportion.

The evolution of the field pattern raises the following comments.

(i) First, when the object displays mesoscopic dimensions (see figures 9 and 10(*a*)), the field distribution is dominated by interference phenomena, so that the field lines do not follow the profile of the square-shaped protrusions.

(ii) Second, as the dimensions of the 3D objects enter the subwavelength range (see figures 10(*b*) and (*c*)), the interference pattern around the objects progressively collapses and the field intensity distribution tends to become perfectly symmetrical thereby reproducing the symmetry of the pads. Under such conditions, and in TE polarization, a highly localized field occurs just above the edges located in a perpendicular direction to the incident field E_0. In fact, when we deal with such subwavelength sized objects, the importance of retardation effects decreases dramatically, so that the symmetry of the field distribution is only governed by both the orientation of the incident field and the profile of the object itself. Actually, these features may help us to get more insight into the complex contrast phenomenon
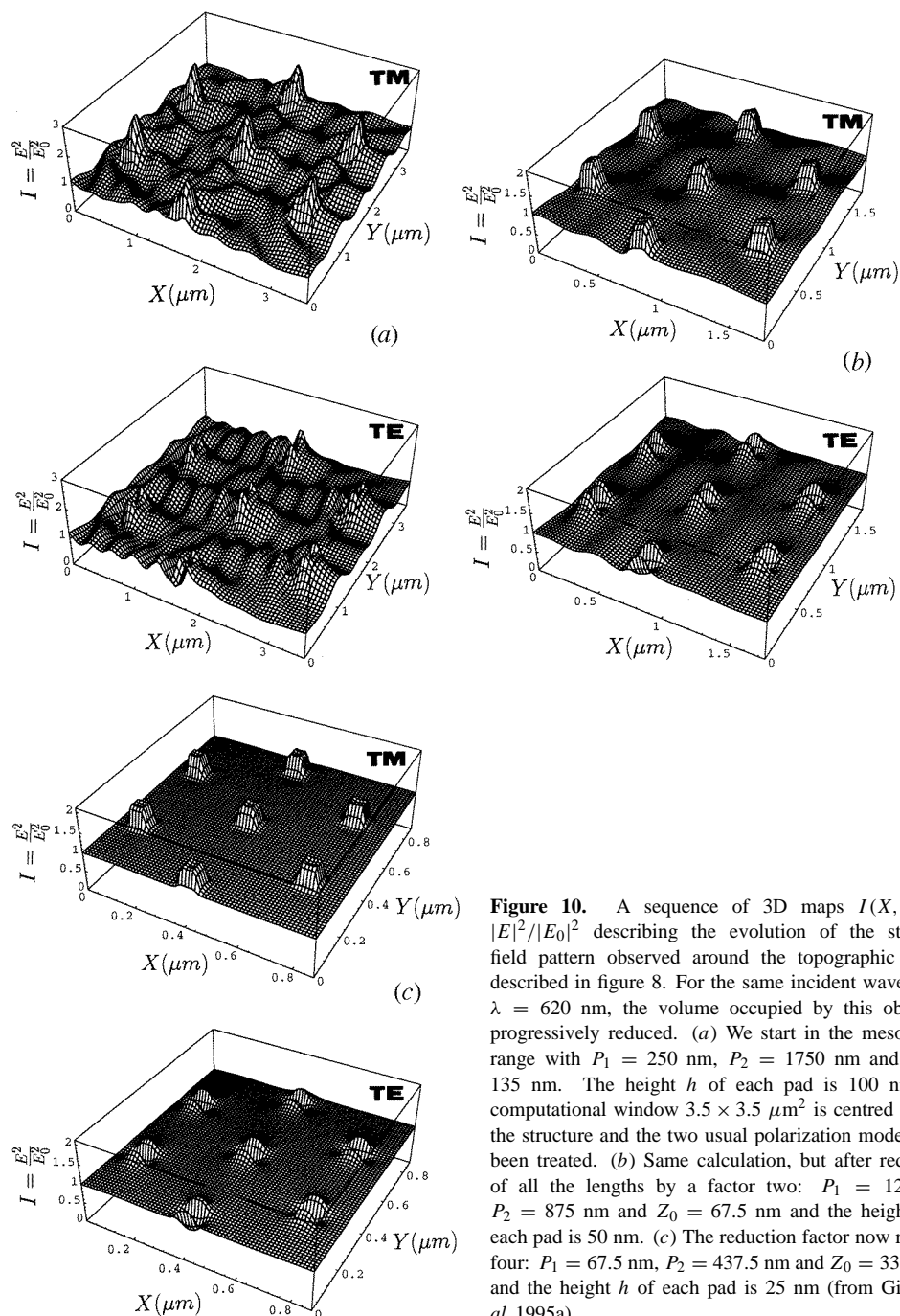
Figure 10. A sequence of 3D maps $I(X, Y) = |E|^2/|E_0|^2$ describing the evolution of the standing field pattern observed around the topographic object described in figure 8. For the same incident wavelength $\lambda = 620$ nm, the volume occupied by this object is progressively reduced. (*a*) We start in the mesoscopic range with $P_1 = 250$ nm, $P_2 = 1750$ nm and $Z_0 = 135$ nm. The height h of each pad is 100 nm, the computational window 3.5×3.5 μm^2 is centred around the structure and the two usual polarization modes have been treated. (*b*) Same calculation, but after reduction of all the lengths by a factor two: $P_1 = 125$ nm, $P_2 = 875$ nm and $Z_0 = 67.5$ nm and the height h of each pad is 50 nm. (*c*) The reduction factor now reaches four: $P_1 = 67.5$ nm, $P_2 = 437.5$ nm and $Z_0 = 33.75$ nm and the height h of each pad is 25 nm (from Girard *et al* 1995a).

observed in the TE mode. The field distribution is now governed by the depolarization effects which result from the conservation of the normal component of the displacement vector $\boldsymbol{D}(\omega) = \epsilon(\omega)\boldsymbol{E}(\omega)$ when crossing the surfaces of the dielectric protrusions. Due to the rapid variation of the dielectric constant between air and glass, this conservation imposes

a sharp variation of the field near the interfaces perpendicular to E_0.

A completely different behaviour is observed with the TM mode. In this polarization the surface wave is mainly dominated by the Z component of the incident field. The main resulting effect is, as expected, a better image–object relation in the subwavelength range. One can observe that, when the size of the square-shaped protrusions is gradually reduced, the field intensity distribution around the objects tends to reproduce their profiles.
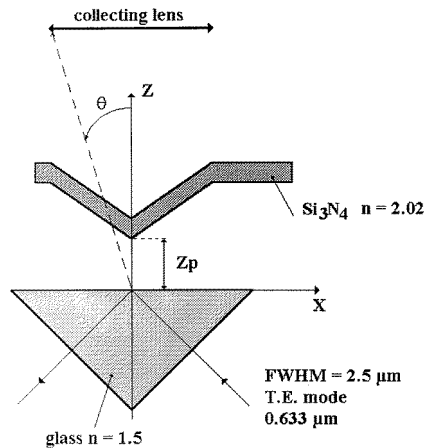

Figure 11. Schematic view of a two-dimensional model of SNOM: a silicon nitride AFM tip placed above a glass prism sample (by courtesy of Bouju *et al* (1996)). The incident focused optical wave has a FWHM of 2.5 μm, a wavelength of 0.633 μm and is in the s-polarized mode.

5. Applications of scattering theory

5.1. Tip–sample interaction and coupling to the far-field

In SNOM devices the use of a pointed detector allows the convertion of the non-radiative optical fields concentrated near the surface irregularities into radiative fields detectable in the far-field region. The amount of optical energy converted by such devices depends mainly on the size of the region of interaction with the confined optical fields described in the previous section. It is also very sensitive to the object parameters, the illumination conditions and to tip design. The tip–sample coupling can be included in the Green dyadic formalism without any formal difficulty. This can be done merely by adding a second perturbation in the self-consistent scheme described in section 4.1. In particular, the discretization procedure already used for taking into account the surface protrusions can be extended to the tip-apex of the detector. The conversion mechanism will then be analysed theoretically using the theory described previously. In fact, the knowledge of the effective field distribution inside the perturbation (tip-apex + surface defect) is sufficient to describe the far-field E_{far} crossing a given surface Σ located inside the wave zone of the detector. As described in Girard *et al* (1994b), outside the source region the Lippmann–Schwinger equation may be applied once more to derive the field radiated by the tip extremity. This numerical scheme has been used in Girard *et al* (1994b) to study the influence of the detector geometry and of the probing distance on image formation of subwavelength 3D topographic objects.

In this review, in order to illustrate the tip–sample coupling phenomenon simultaneously with the concommitant optical non-radiative transfer effect occurring around the contacting
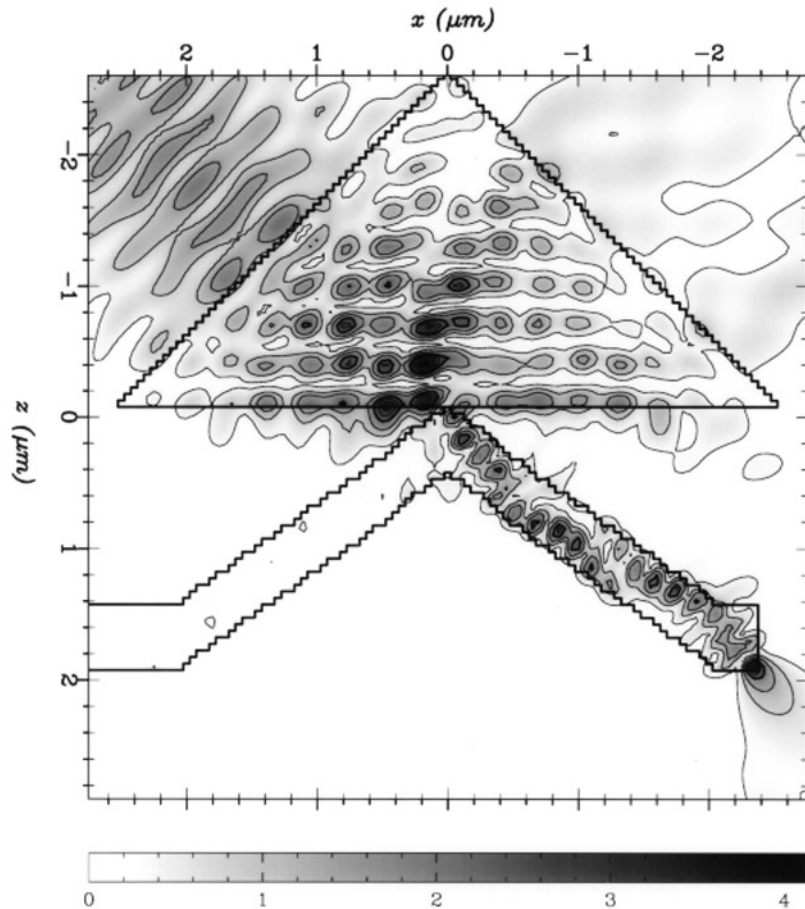
Figure 12. Illustration of the tip–sample optical interaction between a 2D silicon nitride tip and a glass prism. The parameters are described in figure 11 (by courtesy of Bouju *et al* (1996)). This simulation has been performed with a two-dimensionnal numerical code built with the Green function technique (see section 3.1). In this example the tip touches the surface of the prism.

zone, we report in figures 11–13 numerical computerized works on a two-dimensional model of a silicon nitride tip facing a glass sample (see figure 11). In fact, during the past few years, a lot of NFO probe designs have been proposed (Courjon and Bainier 1994): nano-apertures, sharpened optical fibres and, more recently, tetrahedral silicon nitride tips supported by cantilevers which are well known in AFM. The main advantage of this last probe is that one can rely on the force measurement to control accurately the tip–sample distance which, otherwise, is badly defined in usual SNOM configurations. A large lens or an optical fibre placed above the cantilever collects the far-field radiated by the SiN tip (see figure 11). Figure 12 represents the iso-intensity lines inside the entire SiN tip–sample system. It may be seen that when the tip touches the prism, a frustration of the surface wave occurs. In other words, the presence of the SiN probe allows the conversion of a non-radiative optical wave into a radiative one. A direct consequence of this coupling between the probe and the sample is the strong attenuation of the reflected wave. As expected, other
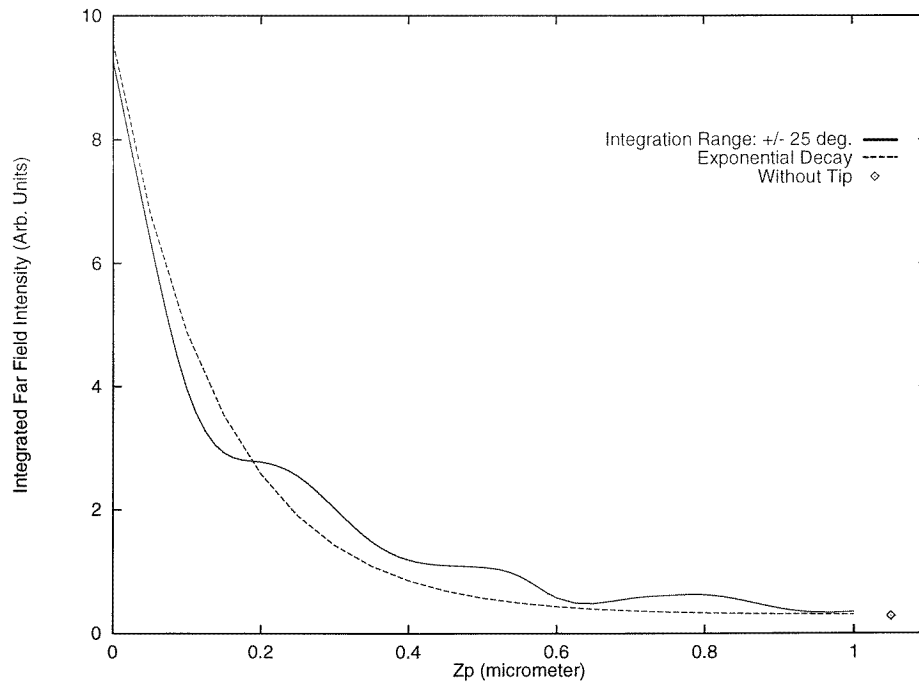
Figure 13. Another illustration of the effect of tip–sample coupling (by courtesy of Bouju *et al* (1996)). The full curve represents the integrated far-field intensity as a function of the tip–sample distance. The square dot located at the right-hand corner of the figure represents the integrated far-field value calculated without silicon nitride tip. The integration range in the far-field zone varies between $-25°$ and $25°$.

simulations indicate that when the tip is pushed higher the coupling between the tip and the microprism dramatically decreases (see figure 13). In order to simulate the experimental signal, the far-field intensity has been integrated around the angle θ (see figure 11). This calculation has been performed for various tip positions along the z-axis direction. The calculated approach curve represented in figure 13 is compared to the exponential decay of a pure evanescent wave whose amplitude is fitted to the self-consistent calculation. We note that such a self-consistent calculation restores the modulations generated by the coupled electromagnetic modes of the tip–sample system (Bouju *et al* 1996, Castiaux *et al* 1995).

5.2. Particular conditions for the validity of the Born approximations

As discussed earlier in this review, the Born approximation is generally not expected to work. However, it may be valid in some circumstances which have been cast mathematically by Carminati and Greffet (1995a, b). The Born approximation was found to be valid for low-relief surfaces such that the largest vertical steps h are much smaller than the incident wavelength λ *and* for small dielectric contrasts such that $\Delta \epsilon = \sup |\epsilon(r, \omega) - \epsilon_{\text{ref}}| \ll 1$.

For a localized defect, these conditions may be detailed as follows,

$$\frac{\pi \Delta \epsilon h \mathcal{S}}{\lambda^2 d} \ll 1 \tag{77}$$

where $\mathcal{S}$ is the area of the defect and d is the distance of observation of the scattered field.

The case of an extended surface profile was also analysed in terms of the Fourier transform of the surface profile. Different validity criteria evolve according to the content of the Fourier spectrum of the surface profile. For low-frequency profiles, the criterion is given by

$$\frac{2\pi \Delta \epsilon h}{\lambda} \ll 1. \tag{78}$$

The validity of the Born approximation for high-frequency profiles is strongly dependent on the polarization of the incident wave and on the distance to the plane of observation. In brief, the approximation holds if the observation distance is large and it is also more accurate for TE polarized incident waves.

5.3. Nanometre-scale optical spectroscopy

The simulations briefly presented in this section illustrate the ability of the Green dyadic technique to account for near-field spectroscopy of three-dimensional objects. A detailed discussion of this matter has been published recently (Girard and Dereux 1994).

The three-dimensional geometry considered in this application is depicted in figure 14. The incident beam is totally reflected at the prism surface. The perturbation of this surface system consists of small metallic spheres of nanometre size. The glass tip scans the surface at a constant height Z_p. This tip is a 70 nm high pyramidal volume whose apex is rounded in order to achieve a 20 nm curvature radius.
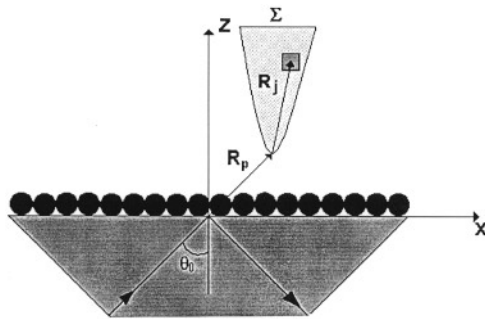

Figure 14. STOM set-up used for the spectroscopy of metal clusters.

Spectroscopic effects occur when the dielectric response of the observed particles show a strong dispersion as a function of the incident wavelength. For visible wavelengths, metals display this characteristic which gives rise to the phenomenon of plasmon resonances. In classical far-field optics, plasmon resonances of small metallic particles are well known phenomena. In near-field optics, they were used by Fischer and Pohl (1989) as the principal ingredient of a particular near-field optical microscope. The calculations in Girard and Dereux (1994) anticipated the near-field imaging of these resonances, a research area where very recent experimental developments have been reported (Krenn *et al* 1995).

The dielectric response of metallic clusters with a radius below 15 nm cannot be approximated by bulk dielectric function values. For the calculations displayed in this section, we have therefore used a somewhat more elaborate model for the polarizability of the metallic spheres obtained by extending the method of Newns (1970) to an electron gas inside a spherical well. A model of multipolar polarizabilities as a function of the size of the metal particle can be described from the knowledge of the interband electronic transitions

of the bulk metal. Such a model includes depolarization effects so that, not only $L_3 = 0$, but also the discretization of the metal spheres is no longer required.

Plasmon resonances are very sensitive to slight variations of the dielectric environment so that approaching a glass tip has unavoidable consequences on the detected spectrum. Near-field spectroscopy of metallic nanospheres is thus a typical example of strong tip–sample interactions (Girard *et al* 1993). Upon approaching the sample, the principal features of the tip–sample interaction are slight red-shifting and broadening of the resonance peak. These behaviours show up simultaneously as a dramatic increase of the intensity which reveals the presence of an enhanced electric near-field (figure 15).

A typical near-field spectroscopic effect shows up as illustrated in figure 16. The iso-intensity curves account for the detection above a small cluster of nine metallic particles. The sphere located at the lower left is made of silver while the others are made of gold. At an incident frequency which is far from the localized plasmon of a silver particle (figure 16), the image recorded above the nine spheres follows the square symmetry of the deposited structure. The intensity recorded around the aggregate displays a larger gradient in the direction orthogonal to the incident surface wave vector which is aligned along the *y* direction. When the incident frequency is tuned to the plasmon frequency of the silver sphere, the intensity is considerably enhanced close to this silver sphere. The resonating particles distort the field above the gold spheres so that the square symmetry of the aggregates is no longer recognizable. This optical resonance effect is interesting to discriminate between different kinds of metallic particles.

Pincemin *et al* (1994) investigated further aspects of the near-field detection of plasmons associated with subsurface particles.

5.4. Radiation pressure effects

The last aspect of the optical tip–sample interaction that we review in this paper deals with the force which results from light pressure effects. Indeed, it was recently proved that multiple scattering of light between two sufficiently close objects is able to induce an optical binding force between the two objects (Burns *et al* 1989, Dereux *et al* 1991). Obviously, this situation is typical of any probe tip approaching a sample so that questions about light pressure effects arose in the context of near-field microscopy. It was then roughly estimated that the magnitude of this light induced force should be experimentally accessible (Depasse and Courjon 1992). Accurate calculations accounting for a complete description of the phenomenon were published recently (Dereux *et al* 1994, Girard *et al* 1994).

Our discussion copes with electrically neutral objects which are kept sufficiently far from each other to avoid the overlap of electron wavefunctions. When the tip approaches the surface under laser beam illumination, the long-range interaction energy includes the van der Waals dispersion energy and the optical binding energy induced by the incident light beam (cf figure 17).

The van der Waals dispersion energy is due to quantum zero-point fluctuations of the coupled charge densities inside the probe tip and the substrate. Its computation may be achieved within a framework similar to the Lippmann–Schwinger formalism but where the evaluations are performed for imaginary frequencies.

The calculation of the optical binding energy induced by the incident beam requires that the spatial and temporal shape of the external field be specified. In the case of harmonically oscillating fields, the optical binding energy is given by the sum of the time-averaged inductive energies experienced by the sample and the probe tip. In Girard *et al* (1994a), we show that these inductive energies are obtained by integrating the electric field intensity
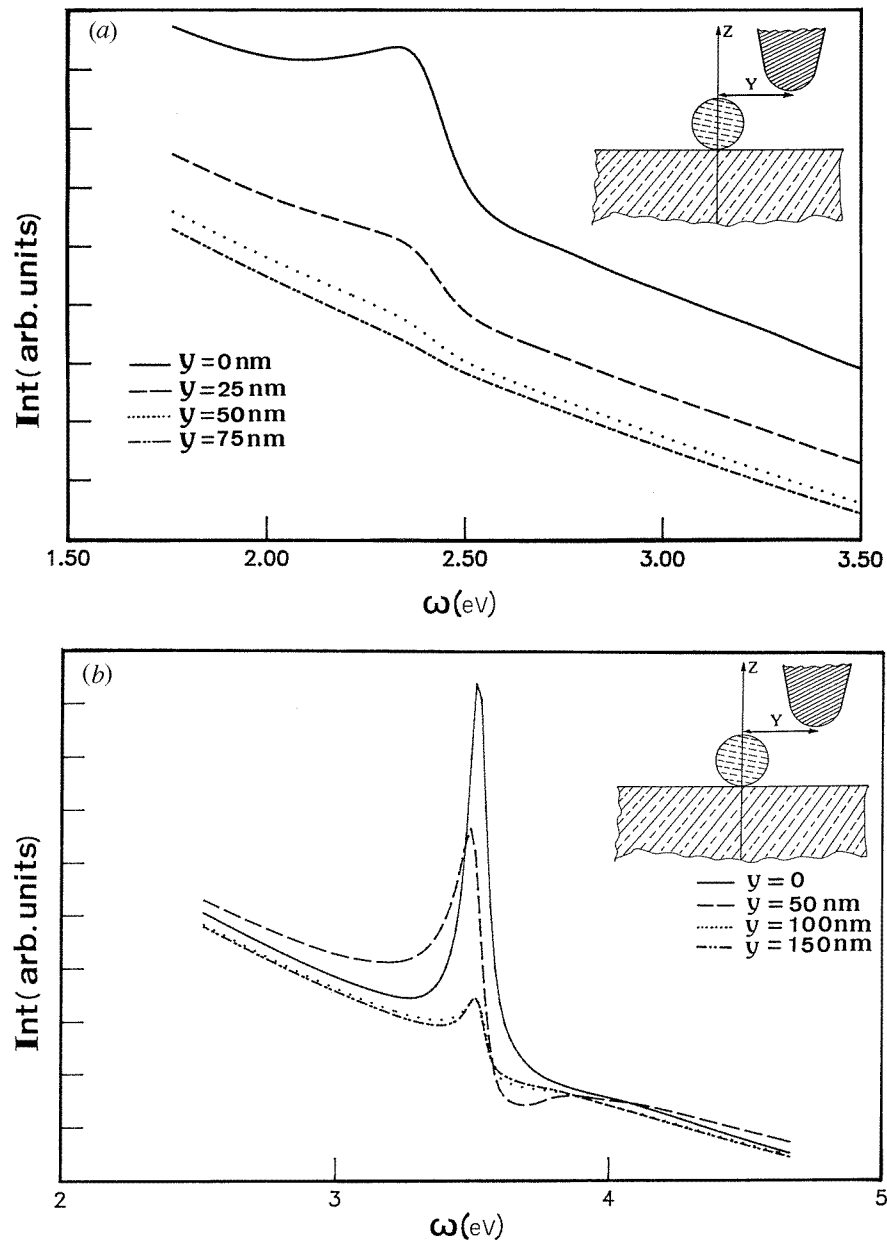
Figure 15. Study of the behaviour of the detected intensity as a function of the incident field frequency. The object is a single metallic particle of 15 nm radius located at the origin of the absolute frame. The approach distance is maintained constant ($Z_p = 32$ nm): (*a*) gold sphere; (*b*) silver sphere.

$|E(r)|^2$ multiplied by the linear susceptibility over the volumes of the substrate and the tip. The optical binding force is then deduced by a straightforward partial derivative with respect to the coordinate Z_p of the probe tip apex.

In the numerical applications of figure 18, a 400 nm high tetrahedral tip made of glass
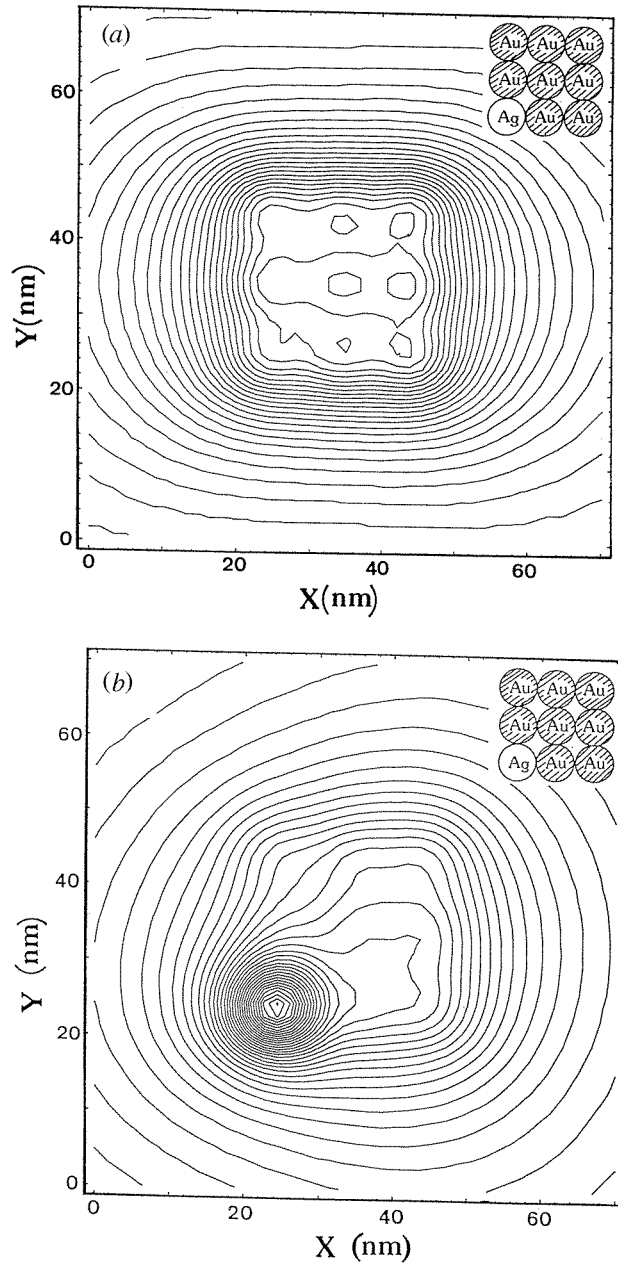
Figure 16. Constant height $Z_p = 11$ nm iso-intensity curves detected by a probe tip scanning an array of gold spheres deposited on a glass substrate. A silver sphere is introduced at the lower left-hand corner of the array. All spheres have a radius of 5 nm. (*a*) The frequency of the incident field is $\omega = 2$ eV. (*b*) Image simulated at the resonance frequency of the silver particle $\omega = 3.34$ eV.

terminated by a 40 nm curvature radius approaches a flat glass substrate. The incident p-polarized plane wave illuminates the tip–sample junction in external reflection (from the vacuum) with a mean power of 15 mW μm^{-2}. We computed that the force arising in

Figure 17. Geometry used in the numerical application of figures 18 and 19. The shape of the tip is tetrahedral like those currently used in scanning force microscopy. The tip aperture angle used in the simulation is equal to 90°. Z_p represents the tip–sample approach distance and the junction is lit in external configuration.

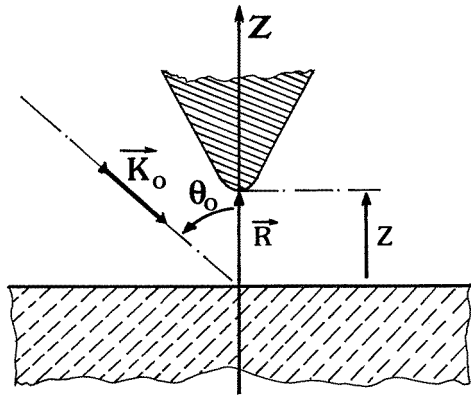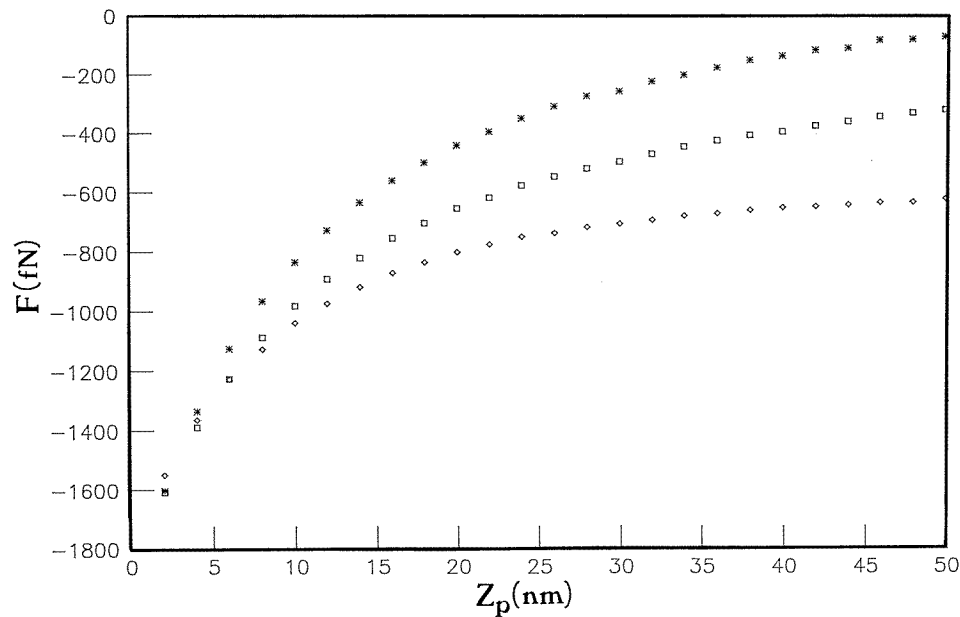**Figure 18.** Optically-induced force as a function of the approach distance Z_p and for various angles of incidence. The refraction index of both tip and substrate is equal to 1.5: ($\star$) $\theta_0 = 56.5°$ (the Brewster angle); ($\square$) $\theta_0 = 50°$; ($\lozenge$) $\theta_0 = 60°$.

this circumstance is of the order of a few piconewtons which is within the range of the experimental resolution of today's resonant scanning force microscopes. We also observed a critical dependence of the optical binding force on the angle of incidence. This variation is understood as follows. If we consider the glass surface alone, the electromagnetic field in the vicinity of the surface is reduced to the interference pattern arising from the superposition of the incident and the reflected fields. We call this pattern the zeroth-order solution in the description of the multiple scattering in the tip–sample junction. This zeroth-order

solution is perturbed by the optical near-field which shows up when approaching the tip. However, the optical field gradients imposed by the zeroth-order solution still dominate. The observation of the the optical binding force due only to the multiple scattering between the substrate and the tip is thus better observed in the absence of any reflected field. Such a situation occurs in the vicinity of the Brewster angle (56.5°). Finally, we note that the magnitude of the force increases with the index of refraction of the tip (see figure 19). This last phenomenon also demonstrates the near-field origin of the optical binding force since it was significant for approach distances below 50 nm.
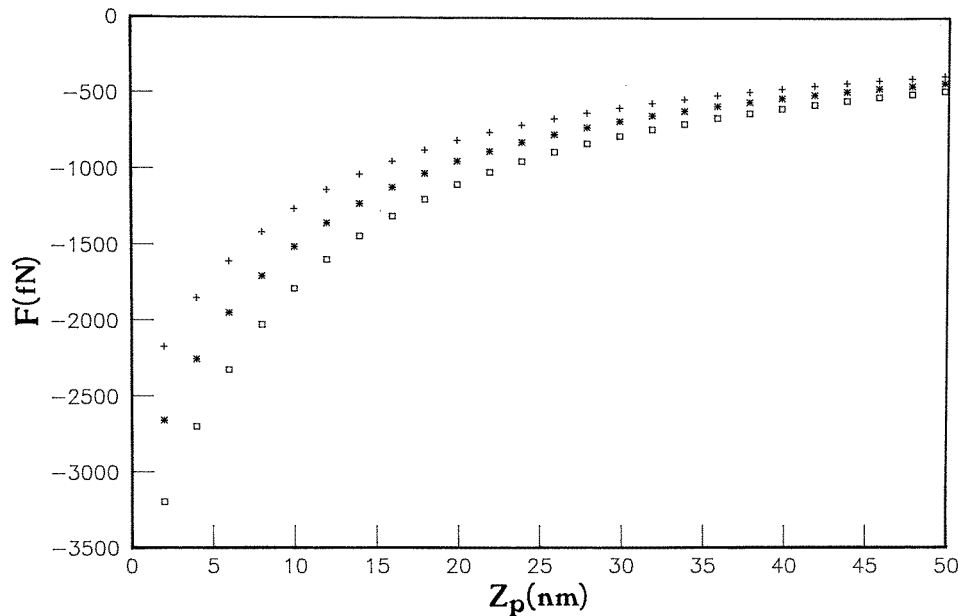
Figure 19. The same as figure 18, but calculated with different values of tip optical index: (+) $n_{pr} = 1.8$; ($\star$) $n_{pr} = 2$; ($\square$) $n_{pr} = 2.2$.

6. Perspectives: from classical to quantum electrodynamics

Since the pioneering works performed in the ATT group by Betzig and Chichester (1993) concerning the observation of single molecules by near-field optical detection, an increasing interest has been devoted to this new single molecular detection (SMD) technique. These contributions raised several important questions concerning the mechanisms underlying such experiments: How does the macroscopic surrounding formed by the tip–sample junction modify the behaviour of the embedded quantum system? In particular, what kind of relation between image and object may be expected by measuring the lifetime changes of a molecular aggregate attached to the probing tip?

From a theoretical point of view, this new experimental methodology is challenging because of the limited number of predictive models able to describe the modification of the intrinsic spectroscopic properties of molecules in the presence of mesoscopic structures dressed with complex optical surroundings. In fact, the presence of any microscopic system (here the molecules) placed in interaction with a mesoscopic environment breaks down the symmetry. This makes the application of a standard boundary conditions based method

extremely difficult. This difficulty can be overcome by a combination of microscopic and macroscopic descriptions, where the response function of the dielectric surrounding medium is derived from an appropriate iterative numerical algorithm (Girard *et al* 1995b).

Let us consider a single molecule trapped inside a confined geometry. For example, it may be the junction formed by the tip of a scanning near field optical microscope (SNOM) facing the surface of a sample. When such a low symmetry system is submitted to an external excitation, different kinds of phenomena may be expected.

—First a highly-confined optical field may be observed in the gap. The magnitude and the shape of the field depends both on the sharpness of the detector and the chemical properties of the junction (metal, semiconductor or dielectric).

—Second, the intrinsic polarizability of the molecule is modified. This effect then introduces a significant modification of the lifetime of the excited molecule which is always accompanied by a small shift of the corresponding occupied state. These spectroscopic changes can be understood by saying that the molecule is 'dressed' by the surroundings, or in other words, that the molecule responds with an *effective polarizability* $\alpha^{\text{eff}}(\omega)$ depending on its location with respect to the material system.

6.1. Spectroscopic properties of a single molecule in confined geometry

The decay of the excited molecule will be considered from the point of view of electromagnetic theory. All effects originating from some specific chemical interactions between the molecule and the substrate will be neglected.

The lifetime of an excited molecule located near spherical and planar surfaces has been the subject of many experimental and theoretical researches for the past 20 years (Drexhage *et al* 1968, Metiu 1984).

In the case of a planar surface, the variation of the lifetime excited state depends on the distance between the molecule and the surface in a complex manner. First, at a very large distance, the variation displays an oscillatory regime due to interference phenomena. In the near zone, we observe a strong decrease of the lifetime due to the non-radiative energy transfer between the excited molecule and the surface. There are actually some similarities between such phenomena and the non-radiative optical energy transfer occurring in the optical tunnelling effect. Indeed, both of them are governed by the evanescent optical fields.

Now let us examine how the intrinsic response properties of the molecule are perturbed by the surroundings.

6.2. Response of the isolated molecule

In the framework of quantum theory, the dynamical polarizability $\alpha_0(\omega)$ of a physical system (atom, molecule, small metallic particle and so on) can be expressed as follows:

$$\alpha_0(\omega) = \frac{1}{\hbar} \sum_r \left(\frac{2\omega_{r0}\mu^{0r}\mu^{r0}}{\omega_{r0}^2 - \omega^2 - i\omega\Gamma_{r0}} \right) \tag{79}$$

where μ^{r0} represent the matrix elements of the polarization molecular operator between the fundamental state and the excited states. Moreover, Γ_{r0} represents the natural lifetime widths of the molecule.

In what will follow, we will restrict ourselves to the case of a two-level molecular system (Metiu 1984). In a first stage, this simplication avoids the huge computational difficulties involved in a realistic treatment of the dynamical properties of the molecule.

Moreover, it provides a comprehensive scheme to analyse the physical mechanisms at the origin of the lifetime changes induced by the presence of highly complex optical systems. This approximation leads to

$$\alpha_0(\omega) = \frac{1}{\hbar} \left(\frac{2\omega_0 \mu^{01} \mu^{10}}{\omega_0^2 - \omega^2 - i\omega\Gamma_0} \right) \tag{80}$$

where μ^{01} is now the matrix element of the polarization operator between the two levels.

6.3. Molecule–surrounding medium coupling: towards the concept of effective polarizability

In this subsection we introduce the electromagnetic coupling between our molecule and the 3D system described in the inset of figure 20. When this system is perturbed by an optical field $E_0(r, t)$ coming from an external laser source, the field $E(r, t)$ in the gap may be derived from the different numerical schemes described in previous sections. It is important to note that this field is an 'observable' since it has already been averaged on the quantum states of the whole primary system (dielectric surrounding medium).
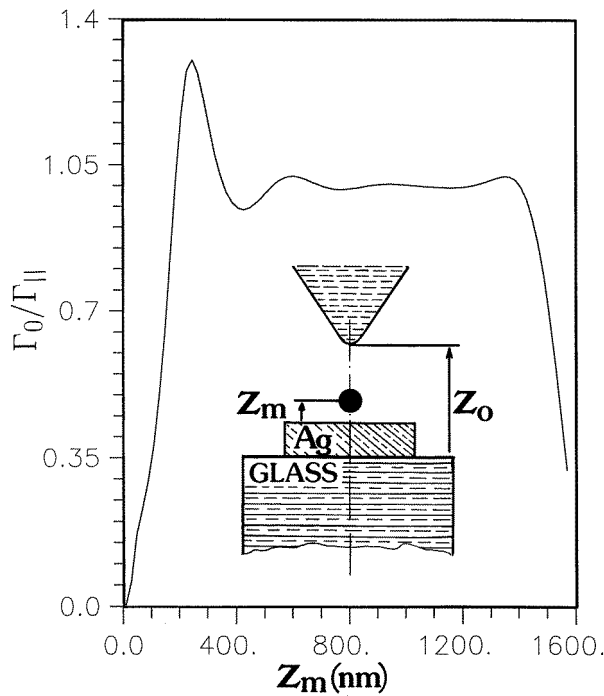
Figure 20. Distance-dependence of the normalized lifetime $\Gamma_0/\Gamma_\parallel$ change of a molecule trapped inside a dielectric–metal–dielectric junction, as a function of the position of the molecule (from Girard *et al* 1995b). The geometry used in the simulation is described in the inset and the tip–sample separation is maintained at $Z_0 = 1600$ nm. The glass sample is covered with a square-shaped silver layer 300 nm long and 20 nm high.

Let us write the interaction Hamiltonian coupling between the primary system and the molecule

$$H(t) = -(E(r_m, t) + \mathcal{E}(r_m, t)) \cdot \mu(t) \tag{81}$$

where $\mu(t)$ and $\mathcal{E}(\boldsymbol{r}_{\mathrm{m}}, t)$ are the polarization operator of the molecule and the electric field operator associated with the 3D dielectric surrounding medium, respectively. The vector $\boldsymbol{r}_{\mathrm{m}}$ denotes the position of the molecule. These operators are written here in the interaction representation

$$\mathcal{E}(\boldsymbol{r}_{\mathrm{m}}, t) = \exp[\mathrm{i}\hbar^{-1}H_0 t]\mathcal{E}_{\mathrm{m}}(\boldsymbol{r})\exp[-\mathrm{i}\hbar^{-1}H_0 t] \tag{82}$$

and

$$\mu(t) = \exp[\mathrm{i}\hbar^{-1}H_0 t]\mu\exp[-\mathrm{i}\hbar^{-1}H_0 t] \tag{83}$$

where H_0 represents the Hamiltonian for the 3D dielectric surrounding medium. At this stage, since one neglects all chemical interactions between the molecule and its support, one can assume that there is no significant modification of the whole wavefunction $|\psi\rangle$ of the system due to the short-range interaction between the molecule and the 3D dielectric surrounding medium. In this situation it is worthwhile applying the time-dependent Hartree (TDH) approximation in which one assumes that each part of the system moves under the combined effect of the external force and the average displacement of the other system (Maclachlan *et al* 1963). Within this approximation one can then consider that $|\psi\rangle$ is a tensorial product of the two wavefunctions $|\psi_{\mathrm{mol}}\rangle$ and $|\psi_{\mathrm{sur}}\rangle$ associated with the molecule and the surrounding medium, respectively.

A straightforward application of the perturbation theory shows that the linear response of the two variables $\mu(t)$ and $\mathcal{E}(\boldsymbol{r}, t)$ is given by

$$\begin{aligned}\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, t) &= \langle \boldsymbol{E}(\boldsymbol{r}_{\mathrm{m}}, t) + \mathcal{E}(\boldsymbol{r}_{\mathrm{m}}, t)\rangle \\ &= \boldsymbol{E}(\boldsymbol{r}_{\mathrm{m}}, t) + \int_{-\infty}^{t} \boldsymbol{S}(\boldsymbol{r}_{\mathrm{m}}, \boldsymbol{r}_{\mathrm{m}}, t - t') \cdot \langle \mu(t')\rangle \,\mathrm{d}t'\end{aligned} \tag{84}$$

and

$$\begin{aligned}\mu_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, t) &= \langle \mu(t)\rangle \\ &= \int_{-\infty}^{t} \alpha_0(t - t') \cdot \langle \boldsymbol{E}(\boldsymbol{r}_{\mathrm{m}}, t') + \mathcal{E}(\boldsymbol{r}_{\mathrm{m}}, t')\rangle \,\mathrm{d}t'\end{aligned} \tag{85}$$

where $\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, t)$ and $\mu_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, t)$ represent both the temporal variation of the effective field and of the dipole moment at the position of the molecule. In addition, the dyadic tensors $\boldsymbol{S}(\boldsymbol{r}_{\mathrm{m}}, \boldsymbol{r}_{\mathrm{m}}, t - t')$ and $\alpha_0(t - t')$ are nothing but the temporal representation of the field susceptibility of the 3D dielectric surrounding medium and the polarizability of the molecule (Agarwal 1975). These quantities can be expressed in terms of the quantum averages of the commutators of the operators $\mathcal{E}(\boldsymbol{r}, t)$ and $\mu(t)$:

$$\boldsymbol{S}(\boldsymbol{r}, \boldsymbol{r}', t - t') = \frac{\mathrm{i}}{\hbar}\langle\psi|[\mathcal{E}(\boldsymbol{r}, t), \mathcal{E}(\boldsymbol{r}', t')]|\psi\rangle \tag{86}$$

and

$$\alpha_0(t - t') = \frac{\mathrm{i}}{\hbar}\langle\psi|[\mu(t), \mu(t')]|\psi\rangle. \tag{87}$$

Finally, by replacing equation (85) in equation (84) one obtains the time-dependent self-consistent equation for the molecular electric field:

$$\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, t) = \boldsymbol{E}(\boldsymbol{r}_{\mathrm{m}}, t) + \int_{-\infty}^{t} \mathrm{d}t' \int_{-\infty}^{t'} \mathrm{d}t''\, \boldsymbol{S}(\boldsymbol{r}_{\mathrm{m}}, \boldsymbol{r}_{\mathrm{m}}, t - t')\alpha_0(t' - t'')\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, t''). \tag{88}$$

Solving this implicit integral equation requires one to pass into the ω-space

$$\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, \omega) = \mathbf{M}(\boldsymbol{r}_{\mathrm{m}}, \omega) \cdot \boldsymbol{E}(\boldsymbol{r}_{\mathrm{m}}, \omega) \tag{89}$$

where $\mathbf{M}(\boldsymbol{r}_{\mathrm{m}}, \omega)$ is a 3×3 matrix defined by

$$\mathbf{M}(\boldsymbol{r}_{\mathrm{m}}, \omega) = [\boldsymbol{I} - \boldsymbol{S}(\boldsymbol{r}_{\mathrm{m}}, \boldsymbol{r}_{\mathrm{m}}, \omega) \cdot \alpha_0(\omega)]^{-1}. \tag{90}$$

Note that from this equation one obtains in a first stage the molecular effective field $\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, \omega)$. The field $\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}, \omega)$ generated by the molecule far away from the emitting zone can be described by using once again the Lippmann–Schwinger equation

$$\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}, \omega) = \boldsymbol{E}(\boldsymbol{r}, \omega) + \boldsymbol{S}(\boldsymbol{r}, \boldsymbol{r}_{\mathrm{m}}, \omega) \cdot \alpha_0(\omega) \cdot \mathbf{M}(\boldsymbol{r}_{\mathrm{m}}, \omega) \cdot \boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, \omega) \tag{91}$$

that can be rewritten as

$$\boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}, \omega) = \boldsymbol{E}(\boldsymbol{r}, \omega) + \boldsymbol{S}(\boldsymbol{r}, \boldsymbol{r}_{\mathrm{m}}, \omega) \cdot \alpha^{\mathrm{eff}}(\omega) \cdot \boldsymbol{E}_{\mathrm{mol}}(\boldsymbol{r}_{\mathrm{m}}, \omega) \tag{92}$$

where $\alpha^{\mathrm{eff}}(\omega)$ defines the effective polarizability of the molecule in the presence of the dielectric surrounding medium:

$$\alpha^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) = \alpha_0(\omega) \cdot \mathbf{M}(\boldsymbol{r}_{\mathrm{m}}, \omega). \tag{93}$$

Many spectroscopic experiments performed near a solid body require the knowledge of the effective molecular polarizability. In fact, this new response function contains all the dynamical information about the coupling with the dielectric surrounding medium. In other words, the molecule radiates optical energy with a polarizability 'dressed' by the dielectric surrounding medium. In the past, several theoretical works have been devoted to its calculation when the molecule interacts with systems of simple symmetry (spheres, cylinders, planes and so on). The symmetry of the tensor $\alpha^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega)$ is governed mainly by the symmetry of the molecule–surrounding medium super system even if the molecular polarizability $\alpha_0(\omega)$ is initially isotropic. For example, in the particular case of a single molecule interacting with a perfectly planar surface, the dyadic tensor $\alpha^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega)$ belongs to the $C_{\infty v}$ symmetry group, and consequently, may be described with two independent components $\alpha_{xx}^{\mathrm{eff}}$ and $\alpha_{zz}^{\mathrm{eff}}$. In the case of a SNOM surrounding medium (corrugated surface + pointed detector), the effective polarizability tensor will be more complex and all components should be taken into account in a realistic calculation.

6.4. Fluorescence lifetime change in complex optical systems

Two important electrodynamic effects are included in the expression of the effective polarizability. The first is a small shift of the excited state, more precisely the frequency ω_0 is shifted towards a lower frequency. This effect can be characterized by the ratio

$$\Omega(\boldsymbol{r}_{\mathrm{m}}) = \omega^{\mathrm{eff}}/\omega_0. \tag{94}$$

Under usual conditions this coefficient is weak and, with respect to other effects, it may be chosen close to unity. Moreover, we have to keep in mind the fact that $\Omega(\boldsymbol{r}_{\mathrm{m}})$ depends on the polarizability component under consideration.

The second effect, much more sensitive to the location of the molecule in the junction, is the lifetime change defined by the ratio

$$\eta(\boldsymbol{r}_{\mathrm{m}}) = \Gamma^{-1}(\boldsymbol{r}_{\mathrm{m}})/\Gamma_0^{-1}. \tag{95}$$

This coefficient also depends on all effective polarizability components. Nevertheless, with a good approximation, we may assume that the dressed molecule belongs to the $C_{\infty v}$ group with its main axis in the z direction. This assumption leads to

$$\alpha^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) = \begin{pmatrix} \alpha_{\perp}^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) & 0 & 0 \\ 0 & \alpha_{\perp}^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) & 0 \\ 0 & 0 & \alpha_{\parallel}^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) \end{pmatrix} \tag{96}$$

where the two independent components $\alpha_\perp^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega)$ and $\alpha_\parallel^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega)$ may then be identified by a two-levels polarizability expression similar to that associated with the isolated molecule (cf equation (80)). This procedure yields

$$\alpha_\perp^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) = \left(\frac{2\omega_\perp A_\perp}{\omega_\perp^2 - \omega^2 - \mathrm{i}\omega\Gamma_\perp} \right) = \alpha_0(\omega) M_{xx}(\boldsymbol{r}_{\mathrm{m}}, \omega) \tag{97}$$

and

$$\alpha_\parallel^{\mathrm{eff}}(\boldsymbol{r}_{\mathrm{m}}, \omega) = \left(\frac{2\omega_\parallel A_\parallel}{\omega_\parallel^2 - \omega^2 - \mathrm{i}\omega\Gamma_\parallel} \right) = \alpha_0(\omega) M_{zz}(\boldsymbol{r}_{\mathrm{m}}, \omega). \tag{98}$$

For each equation, the three different parameters $\omega_{\perp/\parallel}$, $A_{\perp/\parallel}$ and $\Gamma_{\perp/\parallel}$ may be defined by identification. After some algebra one finds

$$\Gamma_{\perp/\parallel} = \mathrm{Re}\left(\frac{\Gamma_0 M_{xx/zz}(\boldsymbol{r}_{\mathrm{m}}, 0)}{M_{xx/zz}(\boldsymbol{r}_{\mathrm{m}}, \omega_0)} \right). \tag{99}$$

In conclusion, one obtains a very compact result, only depending on the dynamical matrix $\boldsymbol{M}(\boldsymbol{r}_m, \omega_0)$ at the resonance frequency of the molecule. Note that the spatial variation of $\Gamma_{\perp/\parallel}$ with respect to the dielectric surrounding medium is contained in the field susceptibility $\boldsymbol{S}(\boldsymbol{r}_{\mathrm{m}}, \boldsymbol{r}_{\mathrm{m}}, \omega)$. In the vicinity of a highly complex system this dyadic may be derived self-consistently by a recursive sequence of Dyson's equations (Girard *et al* 1995b). Moreover, within the first Born approximation one recovers the well known result (Metiu 1984)

$$\Gamma_{\perp/\parallel} \simeq \Gamma_0 + \frac{2|\mu^{01}|^2}{\hbar} \mathrm{Im}[\boldsymbol{S}_{xx/zz}(\boldsymbol{r}_m, \boldsymbol{r}_{\mathrm{m}}, \omega_0)]. \tag{100}$$

In figure 20 we have used relation (99) to investigate the distance-dependence of the normalized lifetime $\Gamma_0/\Gamma_\parallel$ for a fluorescing molecule located in a tip–sample junction. The molecular parameters used in this simulation are $\Gamma_0 = 2 \times 10^6 \mathrm{~s}^{-1}$ and $\alpha_0 = 10 \mathrm{~\AA}^3$, and the fluorescing wavelength is $\lambda_0 = 612$ nm. The geometry of the junction is sketched in the inset of figure 20; it consists of a glass support with a thin square silver protrusion, facing a tetrahedral dielectric tip with sharp edges and a 10 nm ending curvature radius. When the molecule approaches the metal pad one first observes the usual decay which is then followed by the fluorescence quenching. For intermediate distances, the lifetime variation in the gap region (300 nm $\leqslant Z_{\mathrm{m}} \leqslant$ 1300 nm) displays standard quasi-periodic oscillations with a period close to the half-fluorescing wavelength λ_0. It may be seen that towards the dielectric tip the decay is less abrupt, and a magnification of the evolution of the coefficient $\Gamma_0/\Gamma_\parallel$ as the molecule approaches the tip extremity indicates that the lifetime drops by about one order of magnitude when the molecule becomes adsorbed on the tip surface.

This simulation clearly indicates that the fluorescing molecule behaves as a highly sensitive nano-probe to the external environment. In particular, working in the near-field zone just before the quenching effect occurs should make it possible to increase the lateral SNOM resolution (Ambrose *et al* 1994, Xie and Dunn 1994).

7. Conclusion

In this review we have shown that near-field optics is not restricted to the improvement of the resolution of optical microscopy. We emphasize that the principles on which near-field optics is developed are related to the physics of evanescent electromagnetic waves which have been approached in several different contexts since the mid 1960s. We also insist on the fact that the theoretical description of near-field optical phenomena involves

the problem of light diffraction by subwavelength structures which is neglected in most classical textbooks. Finally, since near-field optics is still in rapid development, we give some perspective about trends, for example the study of lifetime effects and their relation to quantum electrodynamics.

Acknowledgments

References

Adam P M, Salomon L, de Fornel F and Goudonnet J P 1993 *Phys. Rev.* B **48** 2680–3
Agarwal G S 1975 *Phys. Rev.* A **11** 230–42
——1977 *Phys. Rev.* B **15** 2371–83
Ambrose W P, Goodwin P M, Martin J C and Keller R A 1994 *Science* **265** 364–7
Barchiesi D and Van Labeke D 1993 *J. Mod. Opt.* **40** 1239–54
——1994 *Microsc. Microanal. Microstruct.* **5** 1–10
——1995 *Ultramicroscopy* **57** 196–203
Bernsten S, Bozhevolnaya E and Bozhevolnyi S 1993 *J. Opt. Soc. Am.* A **10** 878–85
Bethe H A 1944 *Phys. Rev.* **66** 163–82
Betzig E and Chichester R J 1993 *Science* **262** 1422–5
Betzig E, Harootunian A, Lewis A and Isaacson M 1986 *Appl. Opt.* **25** 1890–900
Betzig E, Isaacson M and Lewis A 1987 *Appl. Phys. Lett.* **51** 2088–90
Betzig E, Trautman J K, Harris T D, Weiner J S and Kostelak R L 1991 *Science* **251** 1468–471
Bohren C and Huffman D 1983 *Absorption and Scattering of Light by Small Particles* (New York: Wiley)
Born M and Wolf E 1964 *Principles of Optics* appendix V (New York: Pergamon)
Bouju X, Dereux A, Vigneron J P and Girard C 1996 Scattering of electromagnetic waves by silicon nitride tips
     Stm 95 *J. Vac. Sci. Technol.* B in press
Bouwkamp C J 1954 *Rep. Prog. Phys.* **XVII** 35
Bouwkamp C J and Casimir H B G 1954 *Physica* **20** 539–54
Burns M, Fournier J and Golovchenko J 1989 *Phys. Rev. Lett.* **63** 1233
Carminati R and Greffet J J 1995a *J. Opt. Soc. Am.* A **12** 2716–25
——1995b *Opt. Commun.* **116** 316–21
Carminati R, Madrazo A and Nieto-Vesperinas M 1994 *Opt. Commun.* **111** 26–33
Castiaux A, Dereux A, Vigneron J P, Girard C and Martin O J F 1995 *Ultramicroscopy* **60** 1–9
Courjon D and Bainier C 1994 *Rep. Prog. Phys.* **57** 989–1028
Courjon D, Sarayedine K and Spajer M 1989 *Opt. Commun.* **71** 23–8
Courjon D, Vigoureux J-M, Spajer M, Sarayedine K and Leblanc S 1990 *Appl. Opt.* **29** 3734–40
Debye P 1909 *Ann. Phys. IV* **30** 57–136
Depasse F and Courjon D 1992 *Opt. Commun.* **87** 79
Dereux A, Girard C, Martin O J F and Devel M 1994 *Europhys. Lett.* **26** 37–42
Dereux A and Pohl D 1993 The 90 degree prism edge as a model SNOM probe: near-field, photon tunneling, and
     far-field properties *Near-field Optics (NATO ASI Series E242)* ed D Pohl and D Courjon (Dordrecht: Kluwer)
     pp 189–98
Dereux A, Vigneron J P, Lambin P and Lucas A A 1991 *Physica* B **175** 65–7
Drexhage K H, Kuhn H and Schäfer F P 1968 *Ber. Bunsenges. Phys. Chem.* **72** 329–34

Dürig U, Pohl D W and Rohner F 1986 *J. Appl. Phys.* **59** 3318–27

Economou E 1983 *Green's Functions in Quantum Physics (Springer Series in Solid-State Science 7)* 2nd edn (Berlin: Springer) chs 1–3

Elson J M 1975 *Phys. Rev.* B **12** 2541–2

Englman R and Ruppin R 1966 *Phys. Rev. Lett.* **16** 898

Fano U 1941 *J. Opt. Soc. Am.* **31** 213

Ferrell R A 1958 *Phys. Rev.* **111** 1214

Fischer U C and Pohl D 1989 *Phys. Rev. Lett.* **62** 458–61

Fischer U C and Köglin J 1995 Private communication

Fuchs R and Kliewer K L 1968 *J. Opt. Soc. Am.* **58** 319–30

Fuchs R, Kliewer K L and Pardee W J 1966 *Phys. Rev.* **150** 589–96

Garcia N and Nieto-Vesperinas M 1993 *Opt. Lett.* **18** 2090–2

——1994 *Photon and Local Probe* ed O Marti and R Moller (Dordrecht: Kluwer)

——1995 *Opt. Lett.* **20** 949–51

Girard C and Bouju X 1991 *J. Chem. Phys.* **95** 2056–64

Girard C and Bouju X 1992 *J. Opt. Soc. Am.* B **9** 298–305

Girard C, Bouju X and Dereux A 1993 Optical near-field detection and local spectroscopy of a surface: a self-consistent theoretical study *Near-Field Optics (NATO ASI Series E242)* ed D Pohl and D Courjon (Dordrecht: Kluwer) pp 199–208

Girard C and Dereux A 1994 *Phys. Rev.* B **49** 11 344–51

Girard C, Dereux A and Martin O J F 1994a *Phys. Rev.* B **49** 13 872–81

Girard C, Dereux A, Martin O J F and Devel M 1994b *Phys. Rev.* B **50** 14 467–73

——1995a *Phys. Rev.* B **52** 2889–98

Girard C, Martin, O J F and Dereux A 1995b *Phys. Rev. Lett.* **75** 3098–111

Goudonnet J P, Bourillot E, Adam P M, de Fornel, Salomon L, Vincent P, Nevière M and Ferrell T L 1995 *J. Opt. Soc. Am.* **12** 1749–65

Green H S and Wolf E 1953 *Proc. R. Soc.* A **56** 1129–37

Greffet J-J, Sentenac A and Carminati R 1995 *Opt. Commun.* **116** 20–4

Hafner C and Bomhodt L H 1993 *The 3d Electromagnetic Wave Simulator* (Chichester: Wiley)

Hansen W W 1937 *Phys. Rev.* **47** 139–43

Henry C and Hopfield J J 1965 *Phys. Rev. Lett.* **15** 964

Hopfield J J 1958 *Phys. Rev.* **112** 1555

Huang K 1951 *Proc. R. Soc.* A **208** 352

Jiang S, Tomita N, Ohsawa H and Ohtsu M 1991 *Japan. J. Appl. Phys.* **30** 2107–111

Kann J L, Milster T D, Froehlich F F, Ziolkowski R W and Judkins J B 1995a *J. Opt. Soc. Am.* A **12** 1677–82

——1995b *J. Opt. Soc. Am.* A **12** 501–12

Keller O 1986 *Phys. Rev.* B **34** 3883–99

——1988a *Phys. Rev.* B **38** 8041–60

——1988b *Phys. Rev.* B **37** 10 588–607

Keller O, Bozhevolnyi S and Xiao M 1993 On the resolution limit of near-field optical microscopy *Near-field Optics (NATO ASI Series E242)* ed D Pohl and D Courjon (Dordrecht: Kluwer) pp 229–37

Keller O, Xiao M and Bozhevolnyi S 1992 *Surf. Sci.* **280** 217–30

Kittel C 1976 *Introduction to Solid State Physics* 5th edn (New York: Wiley) ch 13

Kliewer K L and Fuchs R 1966a *Phys. Rev.* **144** 495–503

——1966b *Phys. Rev.* **150** 573–88

Krenn J R, Gotschy W, Somitsch D, Leitner A and Aussenegg F R 1995 *Appl. Phys.* A **61** 541–5

Leviatan Y 1986 *J. Appl. Phys.* **60** 1577–83

Levine H and Schwinger J 1950 *Comm. Pure App. Math.* **3** 355–91

Lieberman K, Harush S, Lewis A and Kopelman R 1990 *Science* **247** 59–61

Lukosz W 1979 *J. Opt. Soc. Am.* **69** 1495–503

Lukosz W and Kunz R E 1977a *J. Opt. Soc. Am.* **67** 1607–14

——1977b *J. Opt. Soc. Am.* **67** 1615–19

Lukosz W and Meier M 1981 *Opt. Lett.* **6** 251–3

Maclachlan A D, Gregory R D and Ball M A 1963 *Mol. Phys.* **7** 119–24

Martin O J F, Dereux A and Girard C 1994 *J. Opt. Soc. Am.* A **11** 1073–80

Martin O J F, Girard C and Dereux A 1995a *Phys. Rev. Lett.* **74** 526–9

——1995b *Helv. Phys. Acta* **10** 12–13

Metiu H 1984 *Prog. Surf. Sci.* **17** 153–320

Mie G 1908 *Ann. Phys. IV* **25** 377–445

Morse P and Feshbach H 1953 *Methods of Theoretical Physics* (New York: McGraw-Hill) ch 13

Newns D 1970 *Phys. Rev.* B **1** 3304

Newton R G 1966 *Scattering Theory of Waves and Particles* (New York: McGraw-Hill) ch 4

Nieto-Vesperinas M and Madrazo A 1995 A theoretical study of near-field interactions with local probes *Photon and Local Probe (NATO ASI Series E)* ed O Marti and R Möller (Dordrecht: Kluwer)

Novotny L and Hafner C 1994 *Phys. Rev.* E **50** 4094–106

Novotny L and Pohl D W 1995 Light propagation in scanning near-field optical microscopy *Photons and Local Probes (NATO ASI Series E)* ed O Marti and R Möller (Dordrecht: Kluwer)

Novotny L, Pohl D W and Regli P 1994 *J. Opt. Soc. Am.* A **11** 1768–79

——1995 *Ultramicroscopy* **57** 180–8

Otto A 1968 *Z. Phys.* **216** 398–410

——1984 *Light Scattering in Solids IV (Topics in Applied Physics 54)* (Berlin: Springer) ch 6, pp 289–418

Pedarnig J D, Specht M and Hänsch T W 1995 Fluorescence lifetime variations and local spectroscopy in scanning near-field optical microscopy *Photons and Local Probes (NATO ASI Series E)* ed O Marti and R Möller (Dordrecht: Kluwer)

Pekar S I 1960 *Sov. Phys.–JETP* **11** 1286

Pincemin F, Sentenac A and Greffet J J 1994 *J. Opt. Soc. Am.* A **11** 1117–27

Pohl D and Courjon D (ed) 1993 *Near-field Optics (NATO ASI Series E242)* (Dordrecht: Kluwer)

Pohl D W 1991 *Adv. Opt. El. Micr.* **12** 243–312

——1992 *Scanning Tunneling Microscopy* **2** 233–71

Reddick R C, Warmack R J, Chilcott D W, Sharp S L and Ferrell T L 1990 *Rev. Sci. Instrum.* **61** 3669–77

Reddick R C, Warmack R J and Ferrell T L 1989 *Phys. Rev.* B **39** 767–70

Regli P 1992 Automatische Wahl der sphaerischen Entwicklungsfunktionen für die 3D-MMP Methode *PhD Thesis* ETH Zürich, Switzerland

Roberts A 1987 *J. Opt. Soc. Am.* A **4** 1970–83

——1989 *J. Appl. Phys.* **65** 2896–9

——1991 *J. Appl. Phys.* **70** 4045–9

Ruppin R 1983 *Surf. Sci.* **127** 108

Sentenac A and Greffet J-J 1992 *J. Opt. Soc. Am.* A **9** 996–1006

Sommerfeld A 1909 *Ann. Phys. IV* **28** 665–737

Specht M, Pedarnig J D, Heckl W M and Hänsch T W 1992 *Phys. Rev. Lett.* **68** 476–9

Stern E A and Ferrell R A 1960 *Phys. Rev.* **120** 130

Stratton J A 1941 *Electromagnetic Theory* (New York: McGraw-Hill)

Takemori T, Inoue M and Ohtaka K 1987 *J. Phys. Soc. Japan* **56** 1587

Toigo F, Marvin A, Celli V and Hill N R 1977 *Phys. Rev.* B 5618–26

van de Hulst H 1957 *Light Scattering by Small Particles* (New York: Wiley)

van Hulst N F, Moers M H and Bölger B 1993 *J. Microsc.* **171** 95–105

van Hulst N F, Segerink F B, Achten F and Bolger B 1992 *Ultramicroscopy* **42** 416–21

Van Labeke D and Barchiesi D 1992 *J. Opt. Soc. Am.* A **9** 732–9

——1993a *J. Opt. Soc. Am.* A **10** 2193–201

——1993b Theoretical problems in scanning near-field optical microscopy *Near-field Optics (NATO ASI Series E242)* ed D Pohl and D Courjon (Dordrecht: Kluwer) pp 157–78

——1995 *Opt. Commun.* **114** 470–80

Vigoureux J, Depasse F and Girard C 1992 *Appl. Opt.* **31** 3036–45

Vigoureux J M and Courjon D 1992 *Appl. Opt.* **31** 3170–9

Vigoureux J M, Girard C and Courjon D 1989 *Opt. Lett.* **14** 1039–42

Xie X S and Dunn R C 1994 *Science* **265** 361–4

Yaghjian A 1980 *Proc. IEEE* **68** 248–63

Zenneck J 1907 *Ann. Phys. IV* **23** 846