



**HAL**  
open science

# Propagation, diffusion et emission thermique de rayonnement electromagnetique aux courtes echelles

Remi Carminati

► **To cite this version:**

Remi Carminati. Propagation, diffusion et emission thermique de rayonnement electromagnetique aux courtes echelles. Physique [physics]. Université Paris Sud - Paris XI, 2002. tel-00002748

**HAL Id: tel-00002748**

**<https://theses.hal.science/tel-00002748>**

Submitted on 17 Apr 2003

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## **HABILITATION A DIRIGER DES RECHERCHES**

**Spécialité : Physique**

**présentée par Rémi CARMINATI**

**Maître de Conférences à l'Ecole Centrale Paris**

**Sujet : Propagation, diffusion et émission thermique de rayonnement électromagnétique aux courtes échelles**

**Composition du Jury :**

<b>MM. Alain ASPECT</b>	
<b>Claude BOCCARA</b>	<b>rapporteur</b>
<b>Jean-Jacques GREFFET</b>	
<b>Juan José SAENZ</b>	
<b>Vahid SANDOGHDAR</b>	<b>rapporteur</b>
<b>Daniel VAN LABEKE</b>	<b>rapporteur</b>

**Date de soutenance : 20 novembre 2002**

# Table des matières

<b>I</b>	<b>Curriculum vitae</b>	<b>3</b>
<b>II</b>	<b>Liste des publications</b>	<b>7</b>
<b>III</b>	<b>Résumé des activités de recherche et d'encadrement</b>	<b>16</b>
<b>1</b>	<b>Optique de champ proche</b>	<b>18</b>
1.1	Un modèle général et adaptable pour le SNOM . . . . .	19
1.2	Des aspects fondamentaux éclaircis . . . . .	20
1.3	Recherches actuelles et perspectives . . . . .	22
<b>2</b>	<b>Propagation et imagerie en milieu diffusant</b>	<b>24</b>
2.1	Diffraction d'une onde diffusive par un objet . . . . .	25
2.2	Transfert instationnaire à travers un milieu diffusant. Transition entre régimes . . . . .	26
2.3	Perspectives . . . . .	29
<b>3</b>	<b>Rayonnement thermique aux courtes échelles</b>	<b>31</b>
3.1	Cohérence du champ thermique émis en champ proche . . . . .	32
3.2	Emission directionnelle et monochromatique par des réseaux . . . . .	34
3.3	Transferts radiatifs en champ proche . . . . .	34
3.4	Perspectives . . . . .	35
<b>IV</b>	<b>Annexes</b>	<b>41</b>

Première partie

**Curriculum vitae**

# Curriculum vitae

## CARMINATI Rémi

Né le 17 février 1971 à Thonon (Haute-Savoie)

Nationalité française

Marié

*Adresse personnelle :*

24, rue Marceau

92130 Issy les Moulineaux

Téléphone : 01 46 38 21 91

*Adresse professionnelle :*

Laboratoire EM2C

Ecole Centrale Paris, 92295 Châtenay-Malabry Cedex

Téléphone : 01 41 13 10 66

E-mail : remi@em2c.ecp.fr

---

## Situation actuelle

Maître de Conférences à l'Ecole Centrale Paris (depuis le 01/09/1997).

**Disciplines d'enseignement :** Physique et Energétique.

**Activités de recherche :** chercheur dans l'équipe "Optique électromagnétique et rayonnement thermique" du Laboratoire d'Energétique Moléculaire et Macroscopique; Combustion (EM2C), UPR 288 du CNRS.

---

## Formation et titres universitaires français

- Diplôme de l'Ecole Centrale Paris obtenu en 1993.
  - DEA Physique des Transferts et Combustion (Ecole Centrale Paris) obtenu en 1993. Mention bien.
  - Doctorat de l'Ecole Centrale Paris, thèse soutenue le 9 septembre 1996. Sujet : "*Analyse de la formation des images en optique de champ proche*". Mention très honorable avec félicitations. Directeur de thèse: Jean-Jacques Greffet, Professeur à l'ECP.
  - Séjour post-doctoral au Département de Théorie de la Matière Condensée (groupe du Prof. M. Nieto-Vesperinas) de l'Institut de Sciences des Matériaux de Madrid (CSIC) d'octobre 1996 à septembre 1997. Financement dans le cadre d'un réseau Européen.
- 

## Activités d'enseignement (effectuées avant juin 2002)

### *Enseignement à l'Ecole Centrale Paris*

- Cours (en 2001-2002) et Travaux Dirigés (depuis 1994-1995) d'Electromagnétisme, 2ème année.
- Cours de Rayonnement et Propriétés des Surfaces (depuis 2001-2002), 3ème année, option "Mécanique-Aéronautique-Energie" (MAE).

- Cours optionnel (Module Thématique) de Pratique de la Modélisation en Physique (depuis 2000-2001), 2ème année.
- Cours optionnel (Module Thématique) d'Optique de Fourier (de 1997-1998 à 1999-2000), 1ère année.
- Travaux Dirigés de Transferts Thermiques (depuis 1997-1998), 2ème année.
- Travaux Dirigés de Physique Statistique (de 1997-1998 à 2000-2001), 1ère année.
- Séances de Méthodologie-Dimensionnement en Transferts Thermiques (depuis 1997-1998), 3ème année, option MAE.
- Responsabilité de la branche "Energie-Propulsion" de l'option MAE (depuis 2001-2002, 19 élèves en 2001-2002).

#### *Enseignement à l'Ecole Supérieure d'Optique*

- Travaux Dirigés d'Electromagnétisme (de 1994-1995 à 1997-1998), 1ère année.

#### *Enseignements de DEA*

- Cours de Microscopies de Champ Proche (depuis 1999-2000), DEA d'Optique et Photonique et Module de Thèse de l'Ecole Doctorale Ondes et Matière d'Orsay.
- Cours de Rayonnement et Propriétés des Surfaces (depuis 2001-2002), DEA Physique des Transferts et Combustion, Ecole Centrale Paris (cours commun avec l'option MAE de 3ème année).

### **Encadrement de travaux de recherche**

#### *Stages de DEA et de fin d'études*

- Mars-juillet 1998: co-encadrement (80%) du stage de fin d'Etudes (et équivalent DEA) de Rachid ELALOUI. *Modélisation de la microscopie optique de champ proche en mode illumination.*
- Janvier-juillet 2001: co-encadrement (80%) du stage de DEA de Marjorie THOMAS. *Emission dipolaire au voisinage d'un nano-objet métallique.*

#### *Thèses*

- Octobre 1996 - janvier 1999: participation à l'encadrement (15%) de la thèse de Jean-Baptiste THIBAUT. *Propagation de la lumière en milieu aléatoire. Fondements et limites de la description radiométrique. Application à l'imagerie.*
- Depuis octobre 1999: co-encadrement (70%) de la thèse de Rachid ELALOUI. *Propagation et transfert diffusif du rayonnement en milieu fortement diffusant. Application à l'imagerie.*
- Depuis janvier 2000: participation à l'encadrement (15%) de la thèse de Jean-Philippe MULET. *Emission thermique et transferts radiatifs et conductifs aux courtes échelles de longueur.*

- Depuis octobre 2001 : co-encadrement (70%) de la thèse de Marjorie THOMAS. *Emission dipolaire en géométrie confinée et diffusion résonante par un système de particules.*
- Depuis octobre 2001 : participation à l'encadrement (15%) de la thèse de François MARQUIER. *Etude expérimentale et théorique de l'émission thermique cohérente par des surfaces microstructurées.*

---

### Langues étrangères

Anglais (courant), Espagnol (lu, parlé), Allemand (lu, parlé).

---

### Séminaires invités

R. Carminati, "Theory and modelling of apertureless near-field optical microscopy", Département de Physique, Groupe de Nano-Optique et Optique Quantique, Université de Constance, Allemagne, février 2000.

R. Carminati, K. Joulain, J.-P. Mulet, J.-J. Greffet, "Emission thermique en champ proche : vers le STM photonique?", Laboratoire de Physique, Université de Bourgogne, Dijon, 19/01/2001.

R. Carminati, "Principes et modélisation de la microscopie optique de champ proche", Laboratoire de Physico-Chimie des Matériaux Luminescents, Université Claude Bernard Lyon I, 12/06/2001.

R. Carminati, K. Joulain, J.-P. Mulet, J.-J. Greffet, "Transfert radiatif à échelle mésoscopique : de l'optique de champ proche au rayonnement thermique", Laboratoire d'Optique Physique, ESPCI, Paris, 4/10/2001.

R. Carminati, "Théorème de réciprocité et ondes évanescentes", Laboratoire Ondes et Acoustique, ESPCI, Paris, 7/03/2002.

R. Carminati, "Transferts radiatifs aux courtes échelles spatiales : limite des concepts classiques", Journée de la Société Française de Thermique, Paris, 19/06/2002.

---

### Divers

Rapporteur de revues de physique (*Physical Review*, *Physical Review Letters*, *Journal of Applied Physics*, *Applied Physics Letters*, *European Physical Journal*) et d'optique (*Journal of the Optical Society of America*, *Optics Letters*, *Optics Communications*, *Journal of Optics A*).

Séjour invité de deux mois en mai-juin 1999 à l'Université Autonome et à l'Institut de Sciences des Matériaux de Madrid (Prof. M. Nieto-Vesperinas et Prof. J.J. Sáenz).

Séjour invité d'une semaine en février 2000 à l'Université de Constance, Allemagne (Groupe de Nano-optique, Dr V. Sandoghdar).

## Deuxième partie

# Liste des publications



## Liste de publications (postérieures à la thèse de Doctorat)

### Revues internationales avec comité de lecture

- 1/ R. Carminati, “Phase properties of the optical near field”, *Phys. Rev. E* **55**, R4901-R4904 (1997).
- 2/ R. Carminati, A. Madrazo, M. Nieto-Vesperinas and J.-J. Greffet, “Optical content and resolution of near-field optical images: influence of the operating mode”, *J. Appl. Phys* **82**, 501-509 (1997).
- 3/ E.R. Méndez, J.-J. Greffet and R. Carminati, “On the equivalence between the illumination and collection modes of the scanning near-field optical microscope”, *Opt. Comm.* **142**, 7-13 (1997).
- 4/ J.-J. Greffet and R. Carminati, “Image formation in near-field optics”, *Prog. Surf. Science* **56**, 133–237 (1997).
- 5/ A. Madrazo, R. Carminati, M. Nieto-Vesperinas and J.-J. Greffet, “Polarization effects in the optical interaction between a nanoparticle and a corrugated surface : implications for apertureless near-field microscopy”, *J. Opt. Soc. Am. A* **15**, 109-119 (1998).
- 6/ P.J. Valle, R. Carminati and J.-J. Greffet, “Contrast mechanisms in illumination-mode SNOM”, *Ultramicroscopy* **71**, 39–48 (1998).
- 7/ R. Carminati, M. Nieto-Vesperinas and J.-J. Greffet, “Reciprocity of evanescent electromagnetic waves”, *J. Opt. Soc. Am. A* **15**, 706–712 (1998).
- 8/ R. Carminati and J.-J. Greffet, “Near-field effects in spatial coherence of thermal sources”, *Phys. Rev. Lett.* **82**, 1660–1663 (1999).
- 9/ J. Ripoll, M. Nieto-Vesperinas and R. Carminati, “Spatial resolution of diffuse photon-density waves”, *J. Opt. Soc. Am. A* **16**, 1466–1476 (1999).
- 10/ P.J. Valle, J.-J. Greffet and R. Carminati, “Optical contrast, topographic contrast and artifacts in illumination-mode scanning near-field optical microscopy”, *J. Appl. Phys.* **86**, 648–656 (1999).
- 11/ L. Aigouy, F.X. Andréani, A.C. Boccara, J.C. Rivoal, J.A. Porto, R. Carminati, J.-J. Greffet, V. Mathet and P. Beauvillain, “Near-field optical spectroscopy using an incoherent light source”, *Appl. Phys. Lett.* **76**, 397–399 (2000).
- 12/ S. Gómez-Moñivas, J.J. Sáenz, R. Carminati and J.-J. Greffet, “Theory of electrostatic probe microscopy : a simple perturbative approach”, *Appl. Phys. Lett.* **76**, 2955–2957 (2000).
- 13/ R. Carminati and J.J. Sáenz, “Scattering theory of Bardeen’s formalism for tunneling : new approach to near-field microscopy”, *Phys. Rev. Lett.* **84**, 5156–5159 (2000).
- 14/ J.B. Thibaud, R. Carminati and J.-J. Greffet, “Scattering of a diffusive wave by a subsurface object”, *J. Appl. Phys.* **87**, 7638–7646 (2000).
- 15/ R. Carminati, J.J. Sáenz, J.-J. Greffet and M. Nieto-Vesperinas, “Reciprocity, unitarity and time-reversal symmetry of the  $S$ -matrix of fields containing evanescent components”, *Phys. Rev. A* **62**, 12712 (2000).
- 16/ J.-J. Greffet and R. Carminati, “Radiative transfer at a nanometric scale: are the usual

- concepts still valid? ”, *Int. J. Heat and Technol.* **18**, Supplement N. 1, 81–85 (2000).
- 17/ A.V. Shchegrov, K. Joulain, R. Carminati and J.-J. Greffet, “Near-field spectral effects due to electromagnetic surface excitations”, *Phys. Rev. Lett.* **85**, 1548–1551 (2000).
- 18/ J.A. Porto, R. Carminati and J.-J. Greffet, “Theory of electromagnetic field imaging and spectroscopy in scanning near-field optical microscopy”, *J. Appl. Phys.* **88**, 4845–4850 (2000).
- 19/ C. Henkel, K. Joulain, R. Carminati and J.-J. Greffet, “Spatial coherence of thermal near fields”, *Opt. Comm.* **186** 57–67 (2000).
- 20/ J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, Comment on “Radiative transfer over small distances from a heated metal”, *Opt. Lett.* **26**, 480–481 (2001).
- 21/ J.N. Walford, J.A. Porto, R. Carminati, J.-J. Greffet, P.M. Adam, S. Hudlet, J.-L. Bijeon, A. Stashkevich and P. Royer, “Influence of tip modulation on image formation in scanning near-field optical microscopy”, *J. Appl. Phys.* **89**, 5159–5169 (2001).
- 22/ J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, “Nanoscale radiative heat transfer between a small particle and a plane surface”, *Appl. Phys. Lett.* **78**, 2931–2933 (2001).
- 23/ J. Ripoll, V. Ntziachristos, R. Carminati and M. Nieto-Vesperinas, “Kirchhoff approximation for diffusive waves”, *Phys. Rev. E* **64**, 51917 (2001).
- 24/ S. Gómez-Moñivas, L.S. Froufe, R. Carminati, J.-J. Greffet and J.J. Sáenz, “Tip-shape effects on electrostatic force microscopy resolution”, *Nanotechnology* **12**, 496–499 (2001).
- 25/ J.-J. Greffet, R. Carminati, K. Joulain, J.P. Mulet, S. Mainguy and Y. Chen, “Coherent emission of light by thermal sources”, *Nature* **416**, 61–64 (2002).
- 26/ J.N. Walford, J.A. Porto, R. Carminati and J.-J. Greffet, “Theory of near-field magneto-optical imaging”, *J. Opt. Soc. Am. A* **19** 572–583 (2002).
- 27/ R. Elaloufi, R. Carminati and J.-J. Greffet, “Pulse propagation through scattering media : from radiative transfer to diffusion”, *J. Optics A: Pure and Applied Optics* **4**, S103–S108 (2002).
- 28/ R. Elaloufi, R. Carminati and J.-J. Greffet, “Definition of the diffusion coefficient in scattering and absorbing media”, soumis à *J. Opt. Soc. Am. A* (2002).

### Contribution à des ouvrages

“Modélisation de la microscopie optique de champ proche”, chapitre 2 du livre *Le Champ Proche Optique : Théorie et Applications*, édité par D. Courjon and C. Bainier (Springer Verlag, Paris, 2001).

### Actes de congrès internationaux

R. Carminati, J.-J. Greffet and A. Sentenac, “A model for the radiative properties of opaque rough surface”, in *Heat Transfer 1998*, Proceedings of 11th IHTC, J.S. Lee ed. (Korean Soc. Mech. Eng., Kyongju, 1998), vol. 7, p. 427.

J.-J. Greffet and R. Carminati, “Radiative transfer at a nanometric scale : are the usual concepts still valid? ”, in *Microscale Heat Transfer*, Proceedings of Eurotherm Seminar N.57 (Poitiers, France, 1998), pp. 241–248.

K. Joulain, J.-P. Mulet, R. Carminati, J.-J. Greffet and A.V. Shchegrov, “Modification of the

thermal emission spectrum at short distances”, in *Heat Transfer and Transport Phenomena in Microsystems*, edited by G.P. Celata (Begell House, New York, 2000), p. 347–351.

J.-P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, “Enhanced radiative heat transfer at nanometric distances”, in *Heat Transfer and Transport Phenomena in Microsystems*, edited by G.P. Celata (Begell House, New York, 2000), p. 352–357.

J.-J. Greffet, R. Carminati, K. Joulain, J.-P. Mulet and C. Henkel, “Coherence properties of thermal near fields: implications for nanometer-scale radiative transfer”, Proceedings of the International School of Physics Enrico Fermi, Course CXLIV, edited by M. Allegrini, N. García and O. Marti (IOS Press, Amsterdam, 2001), pp. 375–392.

### **Communications invitées dans des conférences internationales**

Pour les communications orales, le nom de l’auteur ayant présenté le travail est souligné.

Les autres communications correspondent à des posters.

R. Carminati, K. Joulain, J.-P. Mulet and J.-J. Greffet, “The basic concepts of radiative heat transfer revisited at nanometric scale”, *Gordon Research Conference on Photoacoustic and Photothermal Phenomena*, Oxford (UK), august 2001.

R. Carminati, K. Joulain, J.-P. Mulet and J.-J. Greffet, “Thermal emission of light and radiative transfer at mesoscopic scale”, *Trends in Nanotechnology 2001*, Segovia (Spain), september 2001.

R. Carminati and J.-J. Greffet, “Radiative transfer at mesoscopic scale: the basic concepts revisited”, Keynote Lecture, *Micro/Nanoscale Energy Conversion and Transport*, ICHMT Conference, Antalya (Turquie), April 2002.

### **Communications présentées dans des congrès sans actes**

A. Madrazo, M. Nieto-Vesperinas, R. Carminati and J.-J. Greffet, “Numerical simulation of apertureless near-field optical microscopy”, *Fourth International Conference on Near Field Optics (NFO4)*, Jerusalem (Israel), 1997.

R. Carminati, J.-J. Greffet, A. Madrazo and M. Nieto-Vesperinas, “Comparison between three operating modes of the SNOM”, *Fourth International Conference on Near Field Optics (NFO4)*, Jerusalem (Israel), 1997.

R. Carminati, P.J. Valle and J.-J. Greffet, “Dielectric and topographic contrast in illumination-mode SNOM”, *Fourth International Conference on Near Field Optics (NFO4)*, Jerusalem (Israel), 1997.

J.-J. Greffet, P.J. Valle and R. Carminati, “Analysis of the near field scattered by small particles on a metallic film”, PIERS 97, Cambridge, Massachusetts (USA), July 1997.

A. Madrazo, R. Carminati and M. Nieto-Vesperinas, “Electromagnetic scattering from a particle in front of a conducting interface: influence of the cylinder-interface plasmon resonances”, PIERS 97, Cambridge, Massachusetts (USA), July 1997.

R. Carminati, “On the concept of phase in near-field optics”, PIERS 98, Nantes (France), July 1998.

J.-J. Greffet and R. Carminati, “Can a thermal source be spatially coherent?”, PIERS 98, Nantes (France), July 1998.

E.R. Méndez, J.-J. Greffet and R. Carminati, “Are the illumination and collection modes of the scanning near-field optical microscope fundamentally different?”, PIERS 98, Nantes (France), July 1998.

J.-J. Greffet and R. Carminati, “What is the signal measured by a scanning near-field optical microscope?”, PIERS 98, Nantes (France), July 1998.

J.-J. Greffet and R. Carminati, “Spatially-coherent thermal emission : effect of surface waves”, EOS Topical Meeting on Electromagnetic Optics, Hyeres (France), September 1998.

R. Carminati, M. Nieto-Vesperinas and J.-J. Greffet, “Reciprocity of evanescent electromagnetic waves”, EOS Topical Meeting on Electromagnetic Optics, Hyeres (France), September 1998.

R. Carminati, and J.-J. Greffet, “Near-field effects in thermal emission of light : theory for spatial coherence and suggestion of experiment”, *Fifth International Conference on Near Field Optics (NFO5)*, Shirahama (Japan), december 1998.

R. Carminati, “Phase imaging in near-field optics : polarization, confinement and filtering effects”, *Fifth International Conference on Near Field Optics (NFO5)*, Shirahama (Japan), december 1998.

P.J. Valle, R. Elaloufi, R. Carminati and J.-J. Greffet, “Optical contrast, topographic contrast and artifacts in illumination-mode SNOM”, *Fifth International Conference on Near Field Optics (NFO5)*, Shirahama (Japan), december 1998.

J.-J. Greffet and R. Carminati, “What is the signal measured by a scanning near-field optical microscope?”, *Fifth International Conference on Near Field Optics (NFO5)*, Shirahama (Japan), december 1998.

J.-J. Greffet and R. Carminati, “Reciprocity of evanescent electromagnetic waves”, *Fifth International Conference on Near Field Optics (NFO5)*, Shirahama (Japan), december 1998.

J.-J. Greffet and R. Carminati, “Near-field effects in spatial coherence of thermal sources of light : short-range and long range correlations”, 18th Congress of the International Commission for Optics (ICO), San Francisco (USA), june 1999.

R. Carminati, K. Joulain and J.-J. Greffet, “Near-field correlations of thermal light emitted by planar sources”, *Propagation and imaging in complex media*, Ecole de Physique de Cargèse (France), august 1999.

J.B. Thibaud, R. Carminati and J.-J. Greffet, “Scattering of a thermal wave by a subsurface object”, *Propagation and imaging in complex media*, Ecole de Physique de Cargèse (France), august 1999.

J.A. Porto, R. Carminati and J.-J. Greffet, “Modelling of the image formation of electromagnetic fields in scanning near-field optical microscopy”, Congrès Horizons de l’Optique/Coloq’6, Bordeaux (France), september 1999.

K. Joulain, R. Carminati and J.-J. Greffet, “Near-field correlations of thermal light emitted by planar sources”, Congrès Horizons de l’Optique/Coloq’6, Bordeaux (France), september 1999.

R. Carminati, K. Joulain et J.-J. Greffet, “Influence du champ proche sur la cohérence spatiale des sources thermiques”, Journées Thématiques du Cercle Français des Microscopies de Champ Proche, Troyes (France) september 1999.

J.A. Porto, R. Carminati et J.-J. Greffet, “Modélisation de la formation des images de champs

électromagnétiques en microscopie de champ proche”, Journées Thématiques du Cercle Français des Microscopies de Champ Proche, Troyes (France) september 1999.

K. Joulain, R. Carminati, J.-J. Greffet and A.V. Shchegrov, “Near-field spectroscopy of surface excitations”, CLEO/QELS 2000, Optical Society of America, San Francisco (USA), may 2000.

J.-J. Greffet, R. Carminati, S. Mainguy, Y. Chen and P.J. Valle, “Anisotropic coherent thermal emission by a SiC grating supporting a surface-phonon polariton : experimental and theoretical study”, PIERS 2000, Cambridge, Massachusetts, USA, July 2000.

L. Aigouy, A.C. Boccara, J.C. Rivoal, J.A. Porto, R. Carminati and J.-J. Greffet, “Near-field optical spectroscopy using a broadband light source”, *Sixth International Conference on Near Field Optics (NFO6)*, Twente (the Netherlands), august 2000.

J.N. Walford, J.A. Porto, R. Carminati and J.-J. Greffet, “Theory of near-field magneto-optical image formation” *Sixth International Conference on Near Field Optics (NFO6)*, Twente (the Netherlands), august 2000.

J.A. Porto, R. Carminati and J.-J. Greffet, “Theory of apertureless near-field optical microscopy with application to spectroscopy” *Sixth International Conference on Near Field Optics (NFO6)*, Twente (the Netherlands), august 2000.

J.-J. Greffet, K. Joulain and R. Carminati, “Near-field spectroscopy of surface excitations” *Sixth International Conference on Near Field Optics (NFO6)*, Twente (the Netherlands), august 2000.

R. Carminati and J.-J. Sáenz, “Scattering theory of Bardeen’s formalism for tunneling : a new approach to near-field microscopy”, *Fuerzas y Túnel 2000*, Santiago de Compostella (Spain), september 2000.

S. Gómez-Monivas, J.J. Sáenz, R. Carminati and J.-J. Greffet, “Theory of electrostatic force microscopy : a simple perturbative approach”, *Fuerzas y Túnel 2000*, Santiago de Compostella (Spain), september 2000.

S. Gómez-Monivas, J.J. Sáenz, R. Carminati and J.-J. Greffet, “A simple perturbative approach to Scanning Probe Microscopy”, *Trends on Nanotechnology 2000*, Toledo (Spain), october 2000.

J.-J. Greffet, R. Carminati, K. Joulain, J.-P. Mulet, Y. Chen and S. Mainguy, “Demonstration of a coherent thermal source”, *Coherence and Quantum Optics 8*, Rochester (USA), june 2001.

J.-P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, Y. Chen and S. Mainguy, “Radiative heat transfer between a small particle and a surface at nanometric distance”, Third International Symposium on Radiative Transfer, Antalya (Turkey), june 2001.

J.-P. Mulet, K. Joulain, R. Carminati, J.-J. Greffet, “An infrared thermal source with high directivity and peculiar spectral properties”, Third International Symposium on Radiative Transfer, Antalya (Turkey), june 2001.

R. Elaloufi, R. Carminati and J.-J. Greffet, “Pulse transmission and diffusive radiative transfer through scattering media”, Third International Symposium on Radiative Transfer, Antalya (Turkey), june 2001.

C. Henkel, J.-P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, “Thermal radiation : spatial coherence and mechanical effects in the near field”, 2nd EOS Topical Meeting on Electromagnetic Optics, Paris (France), August 2001.

J.-J. Greffet, R. Carminati, K. Joulain, J.-P. Mulet, S. Mainguy and Y. Chen, “Can a thermal

source be coherent?”, 2nd EOS Topical Meeting on Electromagnetic Optics, Paris (France), August 2001.

K. Joulain, R. Carminati, J.-P. Mulet and J.-J. Greffet, “Local thermal-emission spectroscopy”, 2nd EOS Topical Meeting on Electromagnetic Optics, Paris (France), August 2001.

J.N. Walford, J.A. Porto, R. Carminati and J.-J. Greffet, “Theory of near-field magneto-optical imaging”, 2nd EOS Topical Meeting on Electromagnetic Optics, Paris (France), August 2001.

J.-P. Mulet, K. Joulain, R. Carminati, J.-J. Greffet, “Modelling of nanoscale radiative heat transfer using an electromagnetic approach”, 2nd EOS Topical Meeting on Electromagnetic Optics, Paris (France), August 2001.

R. Elaloufi, R. Carminati and J.-J. Greffet, “Pulse transmission and diffusive radiative transfer through scattering media”, 2nd EOS Topical Meeting on Electromagnetic Optics, Paris (France), August 2001.

R. Elaloufi, R. Carminati and J.-J. Greffet, “Short-time radiative transfer through thin scattering slabs : a numerical and theoretical study”, *Micro/Nanoscale Energy Conversion and Transport*, ICHMT Conference, Antalya (Turquie), April 2002.

F. Marquier, J.-P. Mulet, K. Joulain, R. Carminati, J.-J. Greffet, S. Mainguy and Y. Chen, “Radiative coherent thermal emission by microstructured materials”, *Micro/Nanoscale Energy Conversion and Transport*, ICHMT Conference, Antalya (Turquie), April 2002.

J.-J. Greffet, R. Carminati, K. Joulain, J.-P. Mulet, F. Marquier, S. Mainguy and Y. Chen, “Emission of partially coherent light by thermal sources with periodic microstructures”, PIERS 2002, Cambridge, Massachusetts, USA, July 2002.

M. Thomas, R. Carminati, J.-J. Greffet, R. Arias and M. Nieto-Vesperinas, “Numerical study of the lifetime of an atom close to a lossy structure”, *Seventh International Conference on Near Field Optics (NFO7)*, Rochester (USA), august 2002.

R. Carminati, K. Joulain, J.P. Mulet and J.-J. Greffet, “Detection of the local density of states using near-field emission measurements. Application to local spectroscopy”, *Seventh International Conference on Near Field Optics (NFO7)*, Rochester (USA), august 2002.

R. Carminati, K. Joulain, J.P. Mulet and J.-J. Greffet, “Spontaneous coherent emission of light”, *Seventh International Conference on Near Field Optics (NFO7)*, Rochester (USA), august 2002.

J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, “Radiative heat transfer in the near field”, *Seventh International Conference on Near Field Optics (NFO7)*, Rochester (USA), august 2002.

S. Volz, R. Carminati and K. Joulain, “Thermal response of silicon crystal to picosecond heat pulse by molecular dynamics”, *12th International Heat Transfer Conference (IHTC12)*, Grenoble (France), august 2002.

J.-J. Greffet, R. Carminati, K. Joulain, F. Marquier, J.P. Mulet, S. Mainguy and Y. Chen, “Design of coherent thermal sources of radiation”, *12th International Heat Transfer Conference (IHTC12)*, Grenoble (France), august 2002.

## Rapports de contrats et rapports internes

R. Carminati, “Analyse de la formation des images en optique de champ proche”, Thèse de Doctorat, Ecole Centrale Paris, 1996.

R. Carminati et J.-J. Greffet, “Modèle de réflectivité bidirectionnelle de surfaces rugueuses opaques”, Rapport de contrat SNECMA/CRSA/Laboratoire EM2C, 2000.

## Publications relatives à la thèse de Doctorat

### Revues internationales avec comité de lecture

1/ R. Carminati, A. Madrazo and M. Nieto-Vesperinas, “Electromagnetic waves scattering from a cylinder in front of a conducting surface-relief grating”, *Opt. Comm.* **111**, 26–33 (1994).

2/ J.-J. Greffet, A. Sentenac and R. Carminati, “Surface profile reconstruction using near-field data”, *Opt. Comm.* **116**, 20–24 (1995).

3/ R. Carminati and J.-J. Greffet, “Two-dimensional numerical simulation of the Photon Scanning Tunneling Microscope. Concept of transfer function”, *Opt. Comm.* **116**, 316–321 (1995).

4/ R. Carminati and J.-J. Greffet, “Influence of dielectric contrast and topography on the near field scattered by an inhomogeneous surface”, *J. Opt. Soc. Am. A* **12**, 2716–2725 (1995).

5/ J.-J. Greffet and R. Carminati, “Relationship between the near-field speckle pattern and the statistical properties of a surface”, *Ultramicroscopy* **61**, 11–16 (1995).

6/ R. Carminati and J.-J. Greffet, “Reconstruction of the dielectric contrast profile from near-field data”, *Ultramicroscopy* **61**, 43–50 (1995).

7/ F. de Fornel, P.M. Adam, L. Salomon, J.-P. Goudonnet, A. Sentenac, R. Carminati and J.-J. Greffet, “Analysis of the image formation with a PSTM”, *J. Opt. Soc. Am. A* **13**, 35–45 (1996).

8/ R. Carminati, J.-J. Greffet, N. García and M. Nieto-Vesperinas, “Direct reconstruction of surfaces from near-field intensity under spatially incoherent illumination”, *Opt. Lett.* **21**, 501–503 (1996).

9/ R. Carminati and J.-J. Greffet, “Equivalence of constant-height and constant-intensity images in scanning near-field optical microscopy”, *Opt. Lett.* **21**, 1208–1210 (1996).

### Actes de congrès internationaux

J.-J. Greffet, R. Carminati and A. Sentenac, “Reconstruction of a surface profile using near field data”, dans *1994 International Symposium Digest, Antennas and Propagation* (IEEE Catalog Number 94CH3466-0, Piscataway, 1994) vol. 3, pp. 1640–1641.

J.-J. Greffet and R. Carminati, “Theory of imaging in near-field microscopy”, dans *Optics at the Nanometer Scale*, M. Nieto-Vesperinas and N. García eds., NATO ASI series (Kluwer, Dordrecht, 1996) pp. 1–26.

## Communications présentées dans des congrès sans actes

A. Sentenac, R. Carminati and J.-J. Greffet, “Numerical simulation of subwavelength structures”, *EOS International Conference on Optical Metrology and Nanotechnology*, Engelberg (Suisse), 1994.

J.-J. Greffet, R. Carminati and A. Sentenac, “Inversion des mesures obtenues en PSTM”, *Cercle Francais de Microscopie de Champ Proche*, Aix-en-Provence (France), 1994.

R. Carminati and J.-J. Greffet, “Numerical simulation of the STOM/PSTM”, *OSA Annual Meeting*, Dallas (USA), 1994.

R. Carminati and J.-J. Greffet, “Relationship between PSTM/STOM images contrast and sample dielectric contrast”, *OSA Annual Meeting*, Dallas (USA), 1994.

R. Carminati, A. Sentenac and J.-J. Greffet, “Reconstruction of a surface profile in near field optics”, *OSA Annual Meeting*, Dallas (USA), 1994.

R. Carminati and J.-J. Greffet, “A numerical simulation of the PSTM. Concept of transfer function”, *Near Field Optics and Related Techniques (NFO3)*, Brno (République Tchèque), 1995.

R. Carminati and J.-J. Greffet, “Recent results in the analysis of near field images”, *Workshop on Light Scattering and Related Phenomena*, Arcachon (France), 1995.

R. Carminati, J.-J. Greffet, N. García and M. Nieto-Vesperinas, “The inverse problem in near-field optics”, *Workshop Current state of near-field optical microscopy*, Dresden (Allemagne), 1996.

J.-J. Greffet and R. Carminati, “Recent progress in the analysis of image formation in near-field optical microscopy”, *Workshop on Rough Surface Scattering and Related Phenomena*, Yountville (USA), 1996.

J.-J. Greffet and R. Carminati, “Surface profile reconstruction using near-field detection”, *Progress In Electromagnetics Research Symposium (PIERS)*, Innsbruck (Autriche), 1996.

## Rapports de contrats et rapports internes

R. Carminati, “Etude de la diffusion des ondes électromagnétiques par des réseaux. Application à la conception d’une structure absorbante et d’un réseau de lumière”, *Rapport de stage de DEA*, Laboratoire EM2C, Ecole Centrale Paris, 1993.

R. Carminati et J.-J. Greffet, “Nouveaux mécanismes d’absorption par des réseaux”, *Rapport de contrat DGA/CRSA/Laboratoire EM2C*, 1995.



## Troisième partie

# Résumé des activités de recherche et d'encadrement

# Introduction

A la suite de la soutenance de ma thèse de Doctorat en septembre 1996, j'ai effectué un stage post-doctoral à l'Institut de Sciences des Matériaux de Madrid durant l'année universitaire 1996-1997. J'ai travaillé au Département de Théorie de la Matière Condensée, sous la responsabilité du Professeur M. Nieto-Vesperinas. Depuis septembre 1997, je suis Maître de Conférences à l'Ecole Centrale Paris et chercheur au Laboratoire EM2C (Energétique Moléculaire et Macroscopique, Combustion, UPR 288 du CNRS). Mon travail de recherche s'insère dans les activités de l'équipe "Optique Electromagnétique et Rayonnement Thermique" (responsable Jean-Jacques Greffet, Professeur à l'Ecole Centrale Paris).

Ce document présente une synthèse de mes activités de recherche et d'encadrement durant les six années de la période octobre 1996 - septembre 2002. Il constitue un texte autonome, dont l'objectif est de donner une vision d'ensemble du travail effectué. Des annexes regroupent les principales publications qui en résultent. Le lecteur pourra s'y reporter pour trouver un exposé plus détaillé de chacun des sujets abordés.

Mes activités de recherche se sont articulées autour de trois thèmes :

- Optique de champ proche
- Propagation et imagerie en milieu diffusant
- Rayonnement thermique aux courtes échelles.

Bien que différents par leurs domaines d'applications, ces trois thèmes font appel à un grand nombre de concepts et de méthodes communs. A titre d'exemple, la modélisation de l'émission thermique de rayonnement aux courtes échelles de longueur passe par le calcul du rayonnement électromagnétique émis à des distance sub-longueur d'onde des sources, ce qui est aussi à la base de la théorie de l'optique de champ proche. La propagation d'ondes en milieu aléatoire implique l'utilisation d'une approche statistique similaire à celle de la théorie de la cohérence optique, utilisée pour traiter du rayonnement thermique. La théorie de l'optique de champ proche était ma spécialité de thèse. Dès mon stage post-doctoral, j'ai pu m'ouvrir progressivement vers les deux autres thèmes, et mon activité de recherche actuelle s'équilibre entre les trois thèmes.

Dans cette synthèse, je m'attacherai à montrer ma contribution scientifique à chacun des thèmes, en soulignant les principaux résultats, et ma participation croissante à l'animation et à l'encadrement de la recherche.

# Chapitre 1

## Optique de champ proche

Articles se rapportant au thème et reproduits dans l'annexe 1<sup>1</sup> :

- R. Carminati, A. Madrazo, M. Nieto-Vesperinas and J.-J. Greffet, *Optical content and resolution of near-field optical images: influence of the operating mode*, J. Appl. Phys. **82**, 501–509 (1997).
- J.-J. Greffet and R. Carminati, *Image formation in near-field optics*, Prog. Surf. Science **56**, 133–237 (1997), section 7.
- J.A. Porto, R. Carminati and J.-J. Greffet, *Theory of electromagnetic field imaging and spectroscopy in scanning near-field optical microscopy*, J. Appl. Phys. **88**, 4845–4850 (2000).
- J.N. Walford, J.A. Porto, R. Carminati and J.-J. Greffet, *Theory of near-field magneto-optical imaging*, J. Opt. Soc. Am. A **19**, 572–583 (2002).
- R. Carminati, *Phase properties of the optical near field*, Phys. Rev. E **55**, R4901-R4904 (1997).
- R. Carminati, M. Nieto-Vesperinas and J.-J. Greffet, *Reciprocity of evanescent electromagnetic waves*, J. Opt. Soc. Am. A **15**, 706–712 (1998).
- R. Carminati, J.J. Sáenz, J.-J. Greffet and M. Nieto-Vesperinas, *Reciprocity, unitarity and time-reversal symmetry of the S matrix of fields containing evanescent components*, Phys. Rev. A **62**, 12712 (2000).
- R. Carminati and J.J. Sáenz, *Scattering theory of Bardeen's formalism for tunneling: new approach to near-field microscopy*, Phys. Rev. Lett. **84**, 5156–5159 (2000).
- S. Gómez-Moñivas, J.J. Sáenz, R. Carminati and J.-J. Greffet, *Theory of electrostatic probe microscopy: a simple perturbative approach*, Appl. Phys. Lett. **76**, 2955–2957 (2000).

---

1. Ces articles sont également indiqués en caractères gras dans la bibliographie à la fin de la troisième partie.

Datant d'une quinzaine d'années, les techniques d'optique de champ proche permettent d'obtenir des images de structures surfaciques avec une résolution sub-longueur d'onde (de l'ordre de la dizaine à la centaine de nanomètres actuellement), d'induire des modifications locales de propriétés d'échantillons (lithographie, stockage magnéto-optique) et de mesurer des champs électromagnétiques confinés (plasmons de surface, émission de molécules isolées, modes guidés dans des composants optoélectroniques). Pour une revue sur ces techniques, voir par exemple les réfs.[1]. Dans la suite, on utilisera parfois l'acronyme SNOM (Scanning Near-field Optical Microscopy) pour désigner l'ensemble de ces techniques.

L'interprétation des images obtenues dans ces techniques et l'optimisation des montages posent un problème de modélisation complexe, qui a été abordé au début des années 1990 [2, 3]. La difficulté vient du fait qu'il faut traiter l'interaction électromagnétique en champ proche entre des objets comportant des structures pouvant atteindre le nanomètre, et dont l'un au moins est macroscopique (la pointe utilisée pour détecter le champ confiné ou pour éclairer localement avec une source sub-longueur d'onde).

La modélisation de l'optique de champ proche était le sujet de ma thèse de Doctorat. Durant mon stage post-doctoral, j'ai dans un premier temps prolongé le travail de thèse par l'étude des couplages entre information optique et topographique (liés au balayage de la pointe en suivant le profil de l'échantillon). Ce problème était la sources d'artefacts qui posaient des problèmes d'interprétation des résultats expérimentaux [4]. Par une étude théorique et numérique, nous avons expliqué les conditions d'apparition de ces artefacts et illustré les différents cas par des simulations [5]. Ce travail a été complété lors du stage post-doctoral de Pedro Valle au Laboratoire EM2C en 1997 (encadré par Jean-Jacques Greffet) [6] et du stage de fin d'études (et équivalent DEA) de Rachid Elaloufi que j'ai encadré de mars à juillet 1998 [7].

En parallèle, mon travail sur l'optique de champ proche a pris des orientations nouvelles que je vais présenter plus précisément dans ce qui suit.

## 1.1 Un modèle général et adaptable pour le SNOM

Jean-Jacques Greffet et moi-même avons développé à partir de 1997 un modèle donnant une expression exacte du signal mesuré par un microscope optique de champ proche [8]. Ce modèle, fondé sur le théorème de réciprocité de l'électromagnétisme, a permis d'établir pour la première fois un lien explicite entre le signal mesuré et les propriétés optiques de l'échantillon (variations de constante diélectrique par exemple). Nous l'avons ensuite étendu au cas de la détection de champs électromagnétiques confinés [9], qui est un domaine d'application privilégié des techniques d'optique de champ proche (par exemple pour la spectroscopie locale). Ce modèle permet de prendre en compte le rôle de la pointe (géométrie, composition) et de toute l'optique de détection/illumination dans la formation du signal. Toute l'information est codée dans un champ appelé "champ réciproque", qui est le champ électrique créé à l'emplacement de l'échantillon, dans une situation fictive où l'échantillon est absent et le système est éclairé par une source ponctuelle positionnée en lieu et place du détecteur. Si ce champ peut-être évalué (par exemple à l'aide d'approximations justifiées où d'un calcul numérique), alors il est possible de calculer le signal au détecteur de manière exacte. Ce modèle prédit alors le rôle de la polarisation, de la fréquence, de la cohérence, *etc.*

Deux chercheurs post-doctorants (Juan-Antonio Porto et Julian Walford) ont travaillé avec nous, dans l'intervalle 1999-2001, dans le cadre du réseau européen TMR "Nano-SNOM", permettant

d'appliquer ce modèle à des problèmes concrets. Suite à des mesures effectuées par Lionel Aigouy dans le groupe d'optique de champ proche du Laboratoire d'Optique Physique de l'ESPCI, nous avons pu expliquer la réponse spectrale d'un microscope utilisant des pointes métalliques sans ouverture (technique dite "apertureless SNOM") [10]. Nous avons également étudié l'effet magnéto-optique en champ proche et comparé le potentiel de deux techniques de SNOM [11]. L'application visée par les études de l'effet magnéto-optique en champ proche est le stockage haute-densité. En effet, les techniques d'écriture/lecture optiques traditionnelles utilisent un spot lumineux dont la taille est limitée par la diffraction (de l'ordre du  $\mu\text{m}$ ). En utilisant une technique d'optique de champ proche, on peut réduire la taille du spot et augmenter ainsi la capacité de stockage du support (avec un spot de 10 nm, le gain serait de  $10^4$ ). Après les premières études expérimentales montrant la faisabilité du processus d'écriture/lecture [12], le potentiel (en termes de résolution) des techniques "apertureless" a semblé prometteur et des images de domaines magnéto-optiques (utilisant la rotation Kerr ou Faraday, ou le dichroïsme circulaire) ont été obtenues par les groupes de l'ESPCI et de l'Université de Troyes [13, 14]. Cependant, il est apparu que la résolution des techniques "apertureless", utilisant une pointe métallique, n'atteignait pas en imagerie magnéto-optique celle obtenue en imagerie classique avec les mêmes instruments. Les techniques en mode illumination utilisant des fibres optiques métallisées avec ouverture ne semblent pas souffrir de la même limitation (la résolution est la même dans les deux cas, mais de l'ordre de 100 nm) [12]. Notre étude a confirmé sur des bases théoriques rigoureuses ces observations expérimentales. En particulier, elle a permis d'expliquer l'origine de la perte de résolution des techniques "apertureless" dans le cas de l'imagerie magnéto-optique [11].

Enfin, en collaboration avec le Laboratoire de Nanotechnologies et d'Instrumentation Optique (LNIO) de l'Université de Technologies de Troyes, nous avons étudié l'influence de la vibration verticale de la pointe utilisée dans les techniques "apertureless SNOM" [15]. Cette vibration est nécessaire à l'obtention d'un rapport signal/bruit suffisant par détection synchrone. Cependant, son influence sur la formation de l'image n'avait pratiquement jamais été abordée.

Ce modèle a donc déjà montré son potentiel en termes pratiques. Notons qu'il permet d'introduire, sous certaines hypothèses, le concept de réponse impulsionnelle d'un microscope optique en champ proche [8, 9]. Ce concept est très utile car il permet une approche simple du mécanisme de formation des images, adaptée à un grand nombre de situations pratiques.

## 1.2 Des aspects fondamentaux éclaircis

En parallèle du développement du modèle général et de ses applications, différents aspects fondamentaux liés à l'optique de champ proche ont été étudiés et éclaircis.

### – Phase en champ proche :

L'utilisation de la phase en optique de champ proche a été mise en évidence dans des montages interférométriques, dans l'étude de la conjugaison de phase et dans la démonstration de la faisabilité de l'holographie en champ proche [16]. J'ai proposé en 1997 le premier modèle décrivant le comportement de la phase en optique de champ proche [17]. En particulier, ce travail a montré que la phase du champ pouvait présenter des confinements sub-longueur d'onde, similaires à ceux observés pour l'amplitude (ou l'intensité). Cette propriété de la phase a été utilisée récemment dans une expérience montrant qu'une mesure de phase locale portait une signature claire de la nature de l'échantillon à échelle sub-longueur d'onde [18].

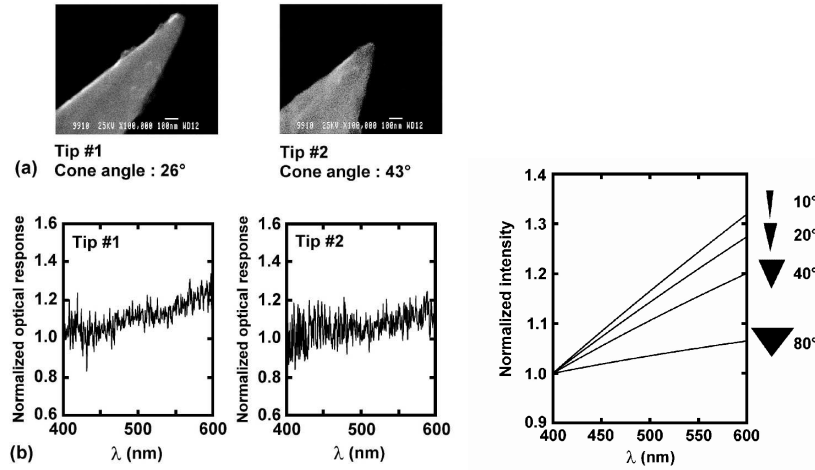


FIG. 1.1 – Comparaison entre une mesure de la réponse spectrale d’un microscope “apertureless” (figures de gauche) et la prédiction du modèle général (figure de droite). Le modèle prédit une dépendance en  $\lambda^{1-2\nu}$  où  $0 < \nu < 1$  est un paramètre dépendant de l’angle d’ouverture de la pointe. L’accord entre théorie et expérience est quantitatif. D’après [10].

– *Théorème de réciprocité :*

Le théorème de réciprocité nous a permis (en plus du développement du modèle général qui a déjà été mentionné) de démontrer l’égalité des performances potentielles des montages de SNOM utilisant la pointe en détection ou en illumination [19]. Une conséquence de ce travail est la preuve que les différences observées (notamment en termes de résolution) sont dues à des différences d’instrumentation (par exemple concernant les pointes utilisées), et non à des différences fondamentales entre les deux types de montages.

La validité du théorème de réciprocité en présence d’ondes évanescentes (ce qui est le cas en optique de champ proche) était cependant une question ouverte (conduisant même à une polémique). Nous avons démontré ce théorème pour les ondes évanescentes, scalaires et électromagnétiques [20, 21]. Ce travail a également permis de clarifier le lien entre réciprocité et renversement du temps, qui n’est pas toujours clair dans la littérature. Il a alors trouvé une application dans un domaine autre que l’optique de champ proche, qui est celui du retournement temporel d’ondes acoustiques [22]. Notre travail a confirmé la possibilité d’effectuer du retournement temporel au-delà de la limite de diffraction.

– *Formule de Bardeen généralisée :*

En 1961, Bardeen a établi une formule donnant la probabilité de transition électronique à travers une barrière tunnel [23]. Cette formule a été très utilisée depuis dans la modélisation du microscope à effet tunnel électronique (STM). En collaboration avec J.J. Sáenz, Professeur à l’Université Autonome de Madrid, nous avons montré que la formule de Bardeen peut être généralisée pour décrire la transmission d’ondes électromagnétiques à travers un système quelconque, en présence ou non d’effet tunnel [24]. Ce résultat permet d’unifier formellement le SNOM et le STM, avec comme conséquence la possibilité d’étendre au SNOM les modèles connus et validés en STM.

– *Vers la microscopie à force électrostatique :*

La collaboration avec J.J. Sáenz (soutenue en 1999-2000 par le programme franco-espagnol Picasso) a permis d’étendre en partie le champ d’application de notre travail sur le SNOM

à une microscopie voisine qui est la microscopie à force électrostatique. Nous avons montré qu'un modèle fondé sur une théorie perturbative et développé à l'origine pour le SNOM [8] pouvait s'appliquer également à cette technique et donner des résultats utiles [25, 26].

Etudier le rapprochement, d'un point de vue théorique, entre les différentes microscopies à sonde locale nous semble très intéressant et porteur d'idées nouvelles. Nous verrons notamment qu'il est possible d'ouvrir le sujet vers la microscopie thermique à sonde locale, et de rejoindre ainsi une partie des problématiques rencontrées dans l'étude des transferts thermiques aux courtes échelles. Nous y reviendrons dans le chapitre 3.

### 1.3 Recherches actuelles et perspectives

Une part croissante des recherches en optique de champ proche s'oriente vers l'étude de l'émission par des molécules uniques [27, 28]. L'intérêt est fondamental, pour la compréhension de l'interaction d'une source élémentaire avec un environnement nanostructuré [29], et pratique, pour la fabrication de sources ayant des propriétés d'émission contrôlées (spectrales, directionnelles, amplifiées ou atténuées). Les progrès des techniques expérimentales dans ce domaine ont été très impressionnants ces dernières années, et le développement d'une "nano-optique" est bien réel. Par exemple, la possibilité d'effectuer une image optique d'un échantillon en utilisant comme source une molécule fluorescente unique a été démontrée [30].

Avec le stage de DEA de Marjorie Thomas que j'ai encadré de janvier à juillet 2001, nous avons débuté une activité théorique et numérique visant à comprendre l'émission dipolaire d'une molécule (ou d'un atome) au voisinage d'un environnement comportant des nanostructures, en présence éventuellement de résonances (par exemple plasmons dans des particules métalliques). Un premier modèle simple a été étudié dans le cadre du stage, et des simulations numériques 2D, à partir d'une méthode intégrale de surface, ont été développées. Dans le cadre du stage, nous nous sommes restreints à l'étude de l'interaction avec un objet métallique bon conducteur (parfaitement conducteur ou modélisé en utilisant une condition d'impédance de surface).

Les méthodes intégrales de surface ont été utilisées en optique dès les années 1980 pour étudier la diffraction par des réseaux [31], puis pour étudier la diffusion par des surfaces rugueuses [32]. Associées à une méthode des moments pour la résolution numérique [33], les formulations intégrales de surface semblent bien adaptées à l'étude de la diffusion du rayonnement émis par un dipôle source ponctuel par une structure localisée.

La travail se prolonge actuellement dans le cadre de la thèse de Marjorie Thomas, que je co-encadre (à raison de 70%) avec Jean-Jacques Greffet, et qui a débuté en octobre 2001. L'objectif est double :

- (1) Proposer une modélisation simple de l'émission dipolaire au voisinage d'un nano-objet, permettant notamment de distinguer (dans l'interaction radiative) les effets d'absorption, de diffusion et l'influence des résonances. A terme, nous souhaitons prendre en compte l'influence des fluctuations thermiques du milieu (présentes dès qu'il y a de l'absorption) qui pourraient créer de nouveaux canaux de désexcitation. Notons que ce dernier point recoupe les travaux actuels sur la définition de la densité d'état du champ électromagnétique (et la quantification du champ) en milieu absorbant.
- (2) Développer un outil numérique 3D permettant de traiter divers problèmes (émission dipolaire au voisinage de nanostructures métalliques ou diélectriques, diffusion par un système de particules en interaction, recherche de configurations permettant de contrôler les propriétés d'émission).

Le travail déjà effectué concerne la simulation 2D, qui a été complétée afin de pouvoir traiter l'interaction de l'atome ou de la molécule (dipôle ponctuel) avec une structure de forme et de composition quelconque. Le calcul effectué permet de séparer la contribution à la modification de durée de vie due à l'absorption de celle due à la diffusion. Les résultats tendent à montrer que près d'un objet métallique, il existe une distance critique à partir de laquelle l'absorption domine toujours la chute de la durée de vie [34].

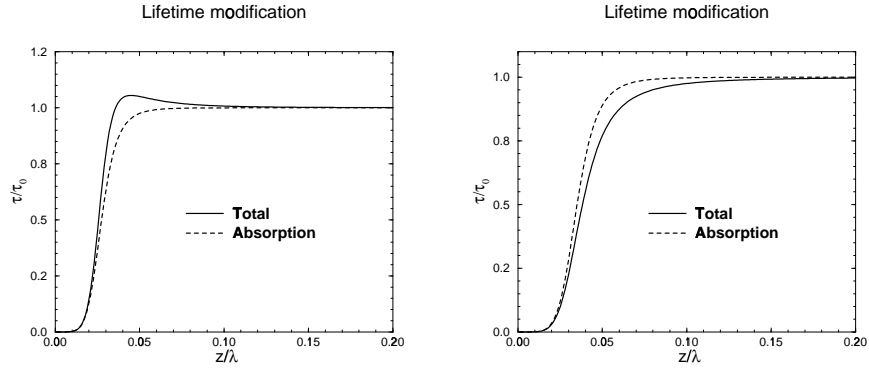


FIG. 1.2 – *Durée de vie normalisée d'un émetteur dipolaire (atome ou molécule) au voisinage d'une petite particule d'argent, en fonction de la distance à la particule. Emission à  $\lambda = 612\text{nm}$ . Figure de gauche : moment dipolaire orienté perpendiculairement à la direction dipôle-particule. Figure de droite : moment dipolaire pointant vers la particule. En trait pointillé, la modification de durée de vie due uniquement à l'absorption. A courte distance, l'absorption est le phénomène prépondérant.*



## Chapitre 2

# Propagation et imagerie en milieu diffusant

Articles se rapportant au thème et reproduits dans l'annexe 2 :

- J. Ripoll, M. Nieto-Vesperinas and R. Carminati, *Spatial resolution of diffuse photon density waves*, J. Opt. Soc. Am. A **16**, 1466–1476 (1999).
- J.B. Thibaud, R. Carminati and J.-J. Greffet, *Scattering of a diffusive wave by a subsurface object*, J. Appl. Phys. **87**, 7638–7646 (2000).
- J. Ripoll, V. Ntziachristos, R. Carminati and M. Nieto-Vesperinas, *Kirchhoff approximation for diffusive waves*, Phys. Rev. E **64**, 51917 (2001).
- R. Elaloufi, R. Carminati and J.-J. Greffet, *Time-dependent transport through scattering media: from radiative transfer to diffusion*, J. Optics A: Pure and Applied Optics **4**, S103–S108 (2002).
- R. Elaloufi, R. Carminati and J.-J. Greffet, *Definition of the diffusion coefficient in scattering and absorbing media*, soumis à J. Opt. Soc. Am. A (2002).

La propagation du rayonnement en milieu diffusant suscite un grand intérêt, principalement par son application à l'imagerie médicale [35]. En effet, les tissus vivants sont peu absorbants dans le rouge et le proche infrarouge sur des épaisseurs de quelques centimètres. Il est donc possible de développer des techniques d'imagerie médicale non intrusives et peu coûteuses, utilisant ces rayonnements. Cependant, les tissus vivants sont des milieux très diffusants, ce qui rend complexe la compréhension du transport radiatif et de la formation des images [36]. Notons que ce sujet est d'une portée très vaste, allant bien au-delà de l'imagerie optique. Citons, comme exemples d'autres domaines d'application, la prédiction des propriétés de réflexion et d'émission infrarouge de milieux chargés de particules (peintures), le transport d'ondes sismiques dans des milieux désordonnés comme la croûte terrestre, ou encore la propagation d'ondes thermiques ou acoustiques dans des solides comportant des imperfections [37].

Mon implication dans ce thème a débuté dès mon stage post-doctoral à Madrid, par une collaboration avec Jorge Ripoll, alors en thèse sur ce sujet avec M. Nieto-Vesperinas. A mon retour, je me suis progressivement impliqué dans des recherches sur ce thème que Jean-Jacques Greffet menait déjà. J'ai participé partiellement (15%) à l'encadrement de la thèse de Jean-Baptiste Thibaut durant l'année 1998-1999. Afin de me familiariser plus avec ce nouveau sujet, j'ai participé à deux Ecoles de Physique aux Houches en 1998 et à Cargèse en 1999. Nous avons ensuite proposé un sujet de thèse que je co-encadre (à raison de 70%) avec Jean-Jacques Greffet. Rachid Elaloufi a démarré le travail sur cette thèse en octobre 1999.

## 2.1 Diffraction d'une onde diffusive par un objet

Mon premier travail, mené avec Jorge Ripoll et M. Nieto-Vesperinas à Madrid a consisté en une étude théorique et numérique de la résolution spatiale en imagerie en milieu fortement diffusant. Dans de tels milieux, la densité d'énergie lumineuse  $U(\mathbf{r}, t)$  obéit à une équation de diffusion de la forme  $\partial U/\partial t - D\nabla^2 U = 0$  où  $D$  est le coefficient de diffusion [38]. Lorsqu'on utilise un éclairage modulé en intensité à une fréquence  $\omega$ , la densité d'énergie en régime établi  $U(\mathbf{r}, \omega)$  vérifie  $\nabla^2 U + k^2 U = 0$ , qui est l'équation de Helmholtz avec un vecteur d'onde complexe  $k^2 = i\omega/D$ . Les solutions constituent des ondes de diffusion, appelées "ondes diffusives" (*diffuse photon-density waves* en anglais). Ce phénomène est en tout point analogue à celui conduisant aux ondes thermiques de conduction obtenues en modulant l'échauffement dans un solide. En introduisant un développement en ondes planes pour les ondes diffusives, nous avons discuté la question de la résolution spatiale obtenue avec de telles ondes [39]. En particulier, ces ondes étant toujours évanescentes (le vecteur d'onde  $k$  étant toujours complexe), nous avons montré que les limitations à la résolution (notamment l'influence de la distance de détection) sont similaires à celles rencontrées en optique de champ proche. Les résultats ont été illustrés par des calculs numériques effectués par Jorge Ripoll, et montrant la répartition de densité d'énergie diffractée par deux objets enfouis dans un milieu diffusant, et séparés d'une distance  $d$  variable [39].

Après mon séjour à Madrid, je me suis impliqué dans le travail de thèse de Jean-Baptiste Thibaud. Nous avons développé une méthode numérique permettant de calculer la diffraction d'une onde diffusive (onde de diffusion de photons ou onde thermique) par un objet situé sous une interface [40]. Les applications visées étaient à la fois l'imagerie optique (médicale) en milieu diffusant et la détection par ondes thermiques de défauts dans des solides opaques, les formulations des deux problèmes étant identiques. La méthode utilisée repose sur une formulation intégrale de la diffraction par un objet, très utilisée pour la diffraction d'ondes électromagnétiques [33]. Nous avons utilisé deux formulations différentes. L'une d'elles fait intervenir uniquement des

intégrales de surface, et est propice au calcul numérique. L'autre fait intervenir une intégrale de surface et une intégrale de volume, et est plus adaptée à l'analyse physique et au développement de modèles simplifiés (à partir par exemple de l'approximation de Born). Pour limiter les temps de calcul, la méthode numérique a été programmée en géométrie 2D. Elle nous a servi notamment de référence pour une étude de la validité de l'approximation de Born pour les ondes diffusives. Cette approximation, utilisée par de nombreux auteurs, permet une formulation très simple du problème inverse, le réduisant à une déconvolution. Ce travail a constitué un chapitre de la thèse de Jean-Baptiste Thibaud, soutenue en mai 2000.

Finalement, je mentionnerai dans cette section un travail qui a été effectué au cours de l'année 2000 par Jorge Ripoll, suite à une idée que je lui avais suggérée à la fin de mon stage post-doctoral. Il s'agit de l'utilisation de l'approximation de Kirchhoff pour traiter de la diffraction d'une onde diffusive par un objet de grande taille. Nous avons écrit ensemble la formulation du problème. Jorge Ripoll l'a ensuite programmée, et comparée avec une résolution numérique exacte en géométrie 2D [41]. Les résultats montrent que dans une géométrie réaliste (échelles de longueur typiques rencontrées dans le cerveau humain), l'approximation fonctionne avec une précision de 10-15%, avec des temps de calculs beaucoup plus courts que ceux d'une résolution exacte. Il semble que ce type de calcul ait un intérêt pratique dans l'interprétation des images expérimentales, par exemple pour une première approche avant la résolution complète du problème inverse.

## **2.2 Transfert instationnaire à travers un milieu diffusant. Transition entre régimes**

Certaines techniques d'imagerie en milieu diffusant utilisent la transmission ou la réflexion d'impulsions lumineuses, ayant des durées allant de quelques dizaines de femtosecondes à quelques dizaines de picosecondes [42]. L'approximation de la diffusion étant très utilisée dans ce contexte pour interpréter les résultats, le problème de sa validité aux temps courts et/ou aux courtes échelles de longueur (systèmes optiquement fins) a été posé et étudié expérimentalement [43, 44, 45]. Des résultats parfois contradictoires ont été publiés, et plusieurs points ressortent, qui ont motivé notre travail : (1) Une étude précise des limites de l'approximation de diffusion aux temps courts et échelles spatiales courtes semble nécessaire. (2) L'utilisation de l'équation de transfert radiatif (ETR) comme outil permettant de traiter simplement la transmission et réflexion d'une impulsion par un milieu diffusant reste marginale dans ce domaine alors qu'elle semble appropriée lorsque l'on sort du domaine de l'approximation de diffusion. (3) L'ETR devrait permettre une étude précise des transitions entre régimes (du régime ballistique au régime diffusif), aussi bien en fonction des échelles spatiales que des échelles temporelles. Ces transitions ont été mises en évidence très précisément par des mesures [45, 46, 47]. Il manque cependant un modèle simple, fondé par exemple sur une théorie de transport de type ETR, permettant de les expliquer.

Un autre problème posé, et ouvert, était la définition du coefficient de diffusion en milieu diffusant et absorbant. Des expériences et des études théoriques publiés au cours des cinq dernières années ont créé une polémique sur ce sujet. En effet, certains résultats tendent à montrer que le coefficient de diffusion est indépendant de l'absorption [48], alors que d'autres affirment une dépendance, qui varie selon l'approche choisie pour établir l'approximation de diffusion [49, 50]. La référence [50] présente d'ailleurs un bon résumé du problème et une preuve convaincante de la dépendance du coefficient de diffusion avec l'absorption.

Ces observations ont motivé le travail effectué dans le cadre de la thèse de Rachid Elaloufi, devant être soutenue au dernier trimestre 2002. Le premier travail a été l'écriture d'un programme numérique permettant de calculer la réflexion et la transmission d'une impulsion lumineuse par une couche plane (*slab* en anglais) contenant un milieu diffusant, en résolvant l'ETR instationnaire. La méthode choisie est une méthode d'ordonnées discrètes standard, utilisée pour résoudre des problèmes statiques [51, 52]. Le principe de notre méthode est d'introduire la transformée de Fourier de l'ETR par rapport au temps. L'ETR instationnaire devient alors une ETR statique équivalente, avec une luminance et un coefficient d'extinction complexes [53]. Utilisant cette méthode, nous avons pu comparer la réflexion et la transmission à travers une couche obtenues en résolvant l'ETR (méthode de référence) et en résolvant l'équation de diffusion. Les résultats ont confirmé que l'approximation de la diffusion n'est valable qu'aux échelles spatiales grandes (devant les libres parcours moyens de diffusion  $l_{sc}$  et de transport  $l_{tr}$ ) et aux temps longs (devant  $l_{sc}/v$  et  $l_{tr}/v$ ,  $v$  étant la vitesse de l'énergie dans le milieu) [53].

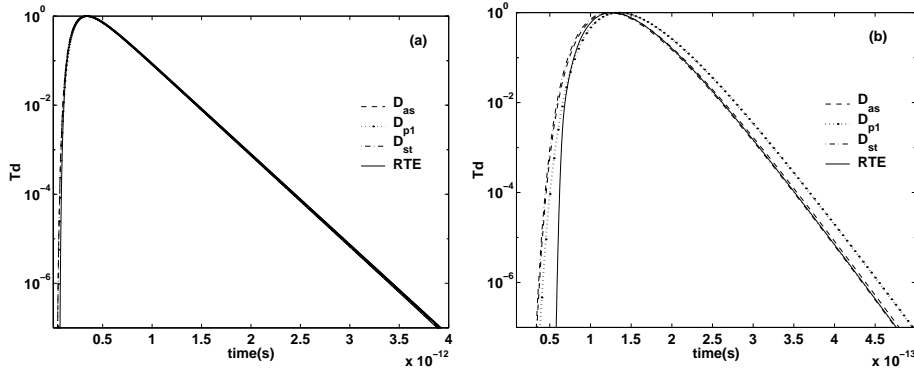


FIG. 2.1 – *Transmission diffuse d'une couche de milieu diffusant d'épaisseur  $L$  éclairée par une impulsion (de durée négligeable) en incidence normale. Comparaison entre la solution de l'ETR et la prédiction de l'approximation de la diffusion. Différentes expressions du coefficient de diffusion ont été utilisées :  $D_{P1} = 1/[3(\mu_s(1-g) + \mu_a)]$ ,  $D_{st} = vl_{tr}/3$ ,  $D_{as} = \nu_0\sqrt{\mu_a}$  (théorie asymptotique).  $L = 20\mu m$ ,  $l_{tr} = 0.95\mu m$ . (a): albedo  $a = 0.995$ , paramètre d'anisotropie  $g = 0.4$ . (b):  $a = 0.85$ ,  $g = 0$ . D'après [53].*

Un deuxième travail effectué dans le cadre de la thèse de Rachid Elaloufi a été de proposer une nouvelle définition du coefficient de diffusion photonique dans un milieu diffusant et absorbant. A partir de l'ETR statique en géométrie 1D (couche éclairée en incidence normale), il est possible d'introduire les modes propres du système de la forme  $\phi_m(z, \mu) = g_m^\pm(\mu) \exp(\pm k_m z)$ , où  $\phi_m(z, \mu)$  est la luminance à une distance  $z$  de l'entrée de la couche et  $\mu$  décrit la direction de propagation.  $g_m(\mu)$  et  $k_m$  apparaissent comme les vecteurs propres et valeurs propres d'un opérateur donné [52, 54]. Pour  $z$  grand, on a asymptotiquement dans le milieu  $L(z, \mu) \simeq g_0^+(\mu) \exp(-k_0 z) + g_0^-(\mu) \exp(+k_0 z)$  où  $k_0$  est la plus petite valeur propre. A partir de cette expression, nous avons montré que le flux d'énergie s'écrit sous la forme d'une loi de Fick, de laquelle le coefficient de diffusion peut être extrait [55]. Les résultats importants de notre travail sont : (1) le coefficient de diffusion dépend de l'absorption, d'une manière différente de celle prédite par l'approximation  $P_1$ , très souvent utilisée [38], (2) l'approximation de la diffusion peut être introduite même lorsque la luminance du mode fondamental (associé à la valeur propre  $k_0$ ) n'est pas quasi-isotrope, (3) le résultat obtenu pour l'expression du coefficient de diffusion est équivalent à celui obtenu par une autre approche dans les refs. [50, 54].

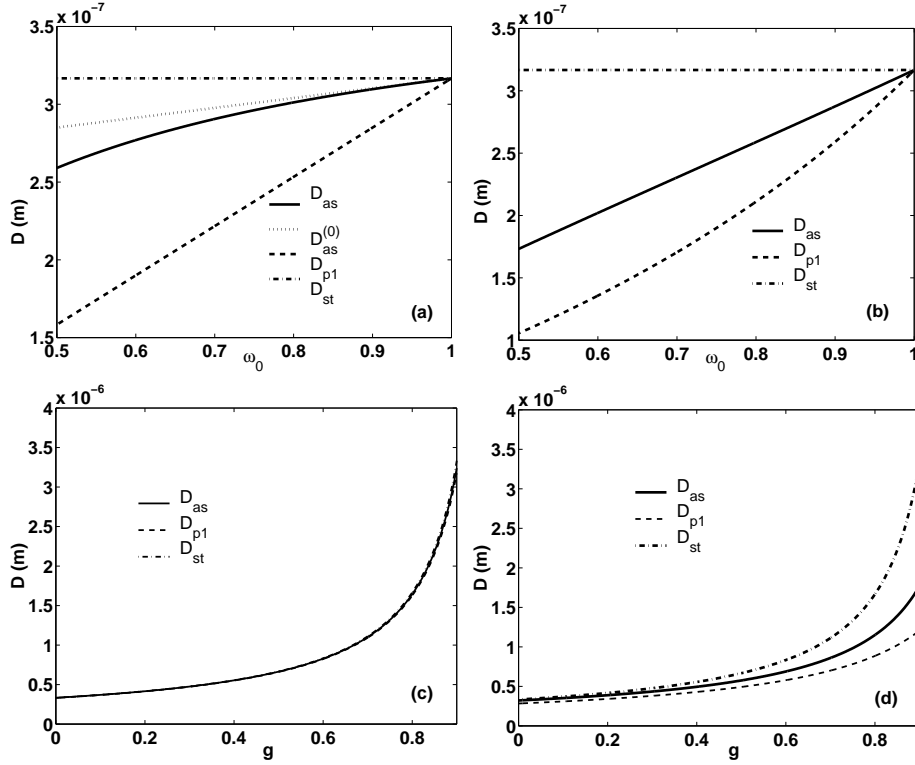


FIG. 2.2 – Dépendance en fonction de l'absorption [(a) and (b)] et de l'anisotropie de la diffusion [(c) and (d)] des différentes expressions du coefficient de diffusion :  $D_{P1} = 1/[3(\mu'_s + \mu_a)]$ ,  $D_{st} = 1/(3\mu'_s)$ ,  $D_{as}$  (résultat de notre étude). (a):  $l_{tr} = 1/\mu'_s = 0.95\mu m$ ,  $g = 0$ . (b):  $l_{tr} = 0.95\mu m$ ,  $g = 0.5$ . (c):  $l_{sca} = 1/\mu_s = 1\mu m$ ,  $a = 0.995$ . (d):  $l_{sca} = 1\mu m$ ,  $a = 0.85$ . D'après [55].

Finalement, nous avons étudié le problème de la transition entre régimes à partir de deux types de résultats expérimentaux publiés dans les cinq dernières années. Dans un premier temps, nous avons étudié le problème de la réduction du coefficient de diffusion effectif d'une couche en fonction de son épaisseur, sur la base des résultats de la réf. [45]. Ce coefficient est défini à partir de la décroissance exponentielle du flux transmis aux temps longs. Les expériences de la réf. [45] ont montré une réduction effective du coefficient de diffusion de moitié lorsque l'épaisseur  $L$  de la couche diminue, et entre dans le domaine  $L < l_{tr}$ . Nous avons montré que la résolution de l'ETR, prenant en compte correctement les conditions aux limites aux frontières de la couche, permet de rendre compte de cet effet. Celui-ci n'est donc pas la conséquence d'effets de diffusion dépendante, ou d'effets cohérents, comme le suggère la conclusion de la réf. [45].

Dans un second temps, nous nous sommes intéressés au calcul des fonctions de corrélations temporelles du champ  $g_1(\tau)$  et de l'intensité  $g_2(\tau)$  qui sont les grandeurs usuelles de la spectroscopie par ondes diffusives (DWS pour *Diffusive Wave Spectroscopy*) [56]. Des expériences récentes, utilisant des montages de tomographie par cohérence optique (OCT pour *Optical Coherence Tomography*) ont montré que la largeur  $\Omega$  des densités spectrales de puissances associées à  $g_1(\tau)$  et  $g_2(\tau)$  dépendait de la distance  $s$  parcourue par les photons dans le milieu diffusant [46, 47]. L'évolution de cette largeur en fonction de  $s$  montre différents régimes, allant du régime ballistique au régime diffusif. Des modèles permettent de prédire la forme de  $\Omega(s)$  dans ces deux cas extrêmes, mais pas la transition. Or, comprendre cette transition et savoir la modéliser

serait un pas important vers des techniques d'imagerie dans le régime de diffusion intermédiaire. Partant de la théorie de la DWS, nous l'avons étendue au cas où le transport des photons, bien qu'en diffusion multiple, n'est pas diffusif. Un outil central de la théorie DWS est la probabilité  $P(s)$  qu'un photon ait parcouru une longueur  $s$  donnée dans le milieu. A partir d'une résolution de l'ETR stationnaire en géométrie plane, nous sommes capables de calculer  $P(s)$  exactement, en prenant en compte la géométrie du système (conditions aux limites aux frontières) et les paramètres physiques du milieu (fonction de phase, coefficient d'absorption). Introduit dans un modèle similaire à celui utilisé en DWS, ce calcul de  $P(s)$  nous a permis de reproduire les résultats expérimentaux de la réf.[47], concernant des mesures de  $g_1(\tau)$ . On y reproduit la transition entre le régime ballistique et le régime diffusif, qui est retrouvé pour  $s$  grand. Le travail sur les résultats expérimentaux concernant  $g_2(\tau)$ , et publiés dans la réf.[46] sont en cours.

## 2.3 Perspectives

A court terme, nous souhaitons compléter le travail sur la propagation et l'imagerie optique en milieu diffusant sur deux aspects :

- étude de la propagation de la fonction de cohérence du champ à travers un milieu aléatoire à l'aide de l'ETR
- étude de la dépendance en fréquence (dispersion) du coefficient de diffusion en milieu diffusant et absorbant.

Sur le premier point, l'idée est d'utiliser l'ETR stationnaire comme équation de transport de la fonction de corrélation spatiale du champ à fréquence fixée. Plus précisément, on peut montrer que la transformée de Wigner du champ (i.e. la transformée de Fourier de la fonction de corrélation spatiale) est l'équivalent de la luminance du rayonnement, grandeur radiométrique et phénoménologique, et qu'elle vérifie l'ETR [57]. Cette approche permet de donner un fondement ondulatoire à la théorie des transferts radiatifs. Elle offre aussi un outil pratique pour étudier la modification du degré de cohérence spatiale du champ au cours de la propagation à travers un milieu diffusant aléatoire [58]. Le travail envisagé consisterait en l'étude, à partir d'un calcul numérique simple (résolution de l'ETR stationnaire en géométrie plane), de l'évolution du degré de cohérence spatiale au cours de la propagation. En particulier, un des objectifs serait de montrer comment des champs incidents avec différents degrés de cohérence sont affectés par la traversée d'un milieu aléatoire (comme par exemple une atmosphère turbulente ou un tissu biologique).

Concernant le second point, il s'agit de poursuivre l'étude sur la définition du coefficient de diffusion en milieu absorbant. La publication d'un article récent [59] intitulé "Diffusion coefficient depends on time, not on absorption" montre que le sujet n'est pas clos. Nous souhaitons étudier cette notion de dépendance en temps (ou dispersion) du coefficient de diffusion. Il semble en effet naturel que lors de l'étude de phénomènes aux temps courts, on aboutisse à une relation non locale dans le temps entre le flux et le gradient de densité d'énergie. Dans le domaine fréquentiel, ceci revient à obtenir une loi de Fick, mais avec un coefficient de diffusion dépendant de la fréquence. Une première étude d'ordres de grandeur montre qu'un tel régime existe pour l'imagerie biomédicale utilisant des impulsions lumineuses de l'ordre de 100 femtosecondes. Il semble donc utile d'essayer de clarifier ce point.

A moyen terme, notre travail sur la propagation optique en milieu diffusant nous servira dans l'étude de la conduction thermique dans des systèmes micro et nano-structurés. Une approche possible dans ce domaine est l'utilisation de l'équation de Boltzmann pour le transport de phonons. Actuellement, il n'existe pas de modèle général pour décrire la diffusion des phonons à l'aide d'une fonction de phase, et les approches habituelles utilisent un modèle de temps de relaxation (terme de collision de la forme  $(f - f^0)/\tau$  où  $f(\mathbf{r}, \mathbf{v}, \omega)$  est la fonction de distribution et  $f^0$  une fonction de distribution d'équilibre). L'étude des transitions entre régimes (du régime ballistique au régime diffusif correspondant à la loi de Fourier) est importante dans ce domaine, et de nombreuses questions sont analogues à celles rencontrées en optique (par exemple, la définition de la conductivité, ou de la conductance, d'un système mésoscopique). Ce sujet rejoint celui des transferts thermiques exposé dans la partie suivante, et sur lequel nous allons concentrer une partie de nos efforts dans les prochaines années.

## Chapitre 3

# Rayonnement thermique aux courtes échelles

Articles se rapportant au thème et reproduits dans l'annexe 3 :

- R. Carminati and J.-J. Greffet, *Near-field effects in spatial coherence of thermal sources*, Phys. Rev. Lett. **82**, 1660–1663 (1999).
- A.V. Shchegrov, K. Joulain, R. Carminati and J.-J. Greffet, *Near-field spectral effects due to electromagnetic surface excitations*, Phys. Rev. Lett. **85**, 1548–1551 (2000).
- J.-J. Greffet, R. Carminati, K. Joulain, J.P. Mulet, S. Mainguy and Y. Chen, *Coherent emission of light by thermal sources*, Nature **416**, 61–64 (2002).
- J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, *Nanoscale radiative heat transfer between a small particle and a plane surface*, Appl. Phys. Lett. **78**, 2931–2933 (2001).
- K. Joulain, R. Carminati, J.P. Mulet and J.-J. Greffet, *Near-field thermal-emission spectroscopy*, manuscrit en préparation (2002).



L'étude théorique et expérimentale des propriétés radiatives des surfaces est un thème présent depuis de nombreuses années au laboratoire EM2C. Avec l'émergence des nanotechnologies, le problème des transferts d'énergie (notamment par conduction et rayonnement) dans les microcomposants et les systèmes nanostructurés électroniques et optoélectroniques a été posé. Le contrôle des échauffements locaux est crucial dans la conception des composants. De plus, des techniques d'imagerie et de stockage haute-densité utilisant la mesure d'un flux thermique local ou le chauffage par une sonde locale ont été développées, posant là aussi le problème du transfert d'énergie entre une pointe et un objet pouvant présenter des structures nanométriques.

Profitant de l'expérience du laboratoire EM2C en transferts radiatifs et des compétences acquises en optique de champ proche (qui traite de problèmes de rayonnement électromagnétique à des échelles sub-longueur d'onde), il paraissait naturel de développer un thème nouveau traitant de l'émission thermique de rayonnement et des transferts radiatifs aux courtes échelles (spatiales et temporelles). Sous l'impulsion de Jean-Jacques Greffet, j'ai débuté mon travail sur ce thème dès mon retour de stage post-doctoral en septembre 1997.

### 3.1 Cohérence du champ thermique émis en champ proche

Notre premier travail a concerné l'étude théorique de l'émission thermique de rayonnement électromagnétique en champ proche par des sources planes (i.e. à des distances de la source très inférieures aux longueurs d'onde mises en jeu). Ce travail faisait suite à des résultats expérimentaux obtenus par deux équipes différentes et qui montraient que l'émission thermique de matériaux présentant des résonances de type phonon-polaritons (cristaux) ou plasmon-polaritons (métaux ou semi-conducteurs dopés) était, dans une certaine mesure, cohérente [60, 61]. L'émission pouvait présenter une certaine directivité (cohérence spatiale) et/ou une largeur spectrale étroite par rapport à la fonction de Planck (cohérence temporelle).

Nous nous sommes intéressés dans un premier temps à la cohérence spatiale. Nous avons utilisé une méthode de calcul du rayonnement thermique développée par Rytov dans les années 1950 [62]. Dans cette approche, le champ de rayonnement thermique est produit par des courants électriques fluctuants dans la matière, du fait de l'agitation thermique. A partir de ces courants, le calcul du champ rayonné se ramène à un calcul classique de rayonnement électromagnétique (comme celui du rayonnement d'une antenne). Cependant, les courants fluctuants ne sont pas connus précisément. Néanmoins, les grandeurs mesurables du rayonnement thermique étant en général des grandeurs quadratiques moyennées (flux ou fonctions de corrélation d'ordre deux), seule la fonction de corrélation spatiale (à fréquence fixée) de la densité de courant est nécessaire. Si on suppose la matière à l'équilibre thermodynamique local, celle-ci est donnée par le théorème fluctuation-dissipation [62].

Avec cette approche, nous avons calculé la fonction de corrélation spatiale, à fréquence fixée, du champ électrique émis par une source plane en champ proche [63]. Notre premier résultat a été de montrer que la longueur de cohérence spatiale du champ, à une longueur d'onde  $\lambda$  fixée, au voisinage d'un matériau présentant des ondes de surfaces résonantes (phonon ou plasmon polaritons) pouvait être très supérieure à  $\lambda/2$ , qui est la longueur de cohérence du rayonnement de corps noir [63]. Cette étude théorique a montré et expliqué pour la première fois l'apparition d'une cohérence spatiale conséquente sur des sources thermiques (qui sont souvent présentées comme des exemples types de sources incohérentes). De plus, nous avons montré que pour des matériaux ne présentant pas de résonances (par exemple le tungstène dans le visible), la longueur

de cohérence pouvait être, en champ proche, très inférieure à  $\lambda/2$ . En d'autres termes, nous avons montré qu'une source thermique en champ proche pouvait être "plus incohérente qu'un corps noir". Ces deux résultats, obtenus par des calculs numériques sans approximation, ont été retrouvés par une étude asymptotique, menée l'année suivante en collaboration avec C. Henkel de l'Université de Potsdam en Allemagne [64]. Cette seconde étude a permis de mettre en évidence les échelles caractéristiques influant sur la longueur de cohérence. En particulier, pour les matériaux présentant des ondes de surface, la longueur de cohérence spatiale est donnée par la longueur de propagation de l'onde de surface, qui peut être de plusieurs dizaines de longueurs d'onde. Pour les matériaux ne présentant pas d'ondes de surface à la longueur d'onde considérée, la longueur de cohérence est de l'ordre de la distance à la surface à la source (ce résultat n'étant valable qu'à une distance d'au moins quelques nanomètres pour que l'hypothèse de milieu continu et local, utilisée dans notre travail, reste valable).

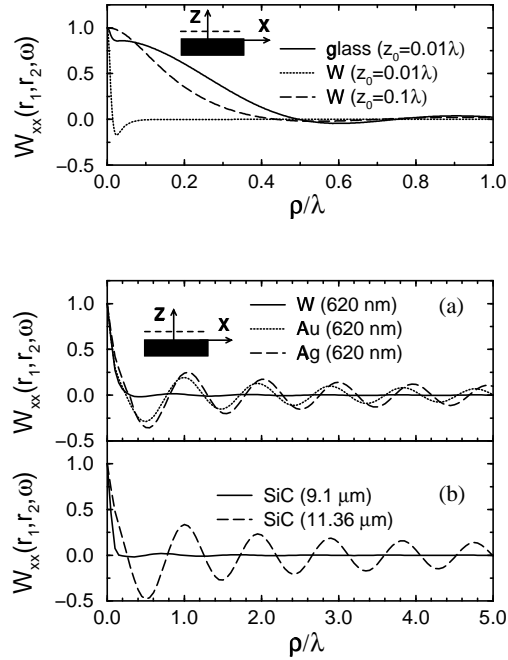


FIG. 3.1 – Fonction de corrélation spatiale  $W_{xx}(\mathbf{r}_1, \mathbf{r}_2, \omega)$  du champ électrique (composante  $x$ ) à fréquence (ou longueur d'onde) fixée, dans un plan  $z = z_0$  au-dessus d'une source plane, en fonction de  $\rho = |\mathbf{r}_1 - \mathbf{r}_2|$ .  $\lambda = 500\text{nm}$ . Figure du haut : cas d'un verre absorbant ( $z_0 = 0.01\lambda$ ) et du tungstène ( $z_0 = 0.01\lambda$  et  $z_0 = 0.1\lambda$ ). On constate que le verre se comporte à cette distance globalement comme un corps noir (longueur de cohérence  $\simeq \lambda/2$ ), et que le tungstène a une longueur de cohérence près de la surface très inférieure à  $\lambda/2$ . Figures du bas : cas de métaux (a) et d'une surface de SiC (b). Les métaux comme l'or ou l'argent présentant des plasmon-polaritons dans le visible ont un champ thermique émis présentant une longueur de cohérence très supérieure à la longueur d'onde. A  $\lambda = 11.36\mu\text{m}$ , l'excitation d'un phonon-polariton sur la surface de SiC engendre également une cohérence spatiale longue portée. D'après [63].

Suite à ce travail, en collaboration avec A. Shchegrov (à l'époque chercheur à l'Institute of Optics de Rochester, USA), nous avons étudié les propriétés spectrales du rayonnement thermique émis par le même type de sources en champ proche [65]. Ceci revient à étudier la cohérence temporelle de ces sources, le temps de cohérence temporelle étant l'inverse de la largeur spectrale. Là-aussi nous avons mis en évidence des effets très surprenants. Un résultat important est que le spectre d'émission thermique (dans l'infrarouge) au voisinage d'une source présentant une résonance de type phonon-polariton est quasi-monochromatique. Ce comportement, que nous avons mis en évidence d'abord par le calcul, a été expliqué par une divergence de la densité de modes électromagnétiques à la fréquence  $\omega_0$  correspondant à l'excitation résonante du phonon-polariton de surface (et donnée par  $\text{Re}[\epsilon(\omega_0)] = -1$  où  $\epsilon(\omega)$  est la constante diélectrique du matériau) [65]. Un autre résultat important est que la densité d'énergie est amplifiée (par rapport à sa valeur en champ lointain) de plusieurs ordres de grandeur. Ces deux comportements ont des conséquences importantes sur les transferts d'énergie par rayonnement en champ proche, comme nous le verrons un peu plus loin. De plus, ils laissent entrevoir la possibilité de concevoir des sources infrarouges en champ proche ayant des propriétés remarquables.

### 3.2 Emission directionnelle et monochromatique par des réseaux

Cette cohérence (spatiale et temporelle) amplifiée en champ proche a des conséquences mesurables sur les propriétés d'émission thermique en champ lointain de surfaces microstructurées. Par exemple, en gravant un réseau lamellaire sur une surface de SiC (qui présente une résonance phonon-polariton au voisinage de  $\lambda = 11\mu m$ , voir les réfs.[63, 65]), on peut créer une source émettant de manière directionnelle et monochromatique. Nous avons ainsi conçu un réseau de SiC par un calcul numérique, afin qu'il présente un pic d'émission très directionnel à  $\lambda = 11,36\mu m$ . Ce réseau a ensuite été réalisé au Laboratoire de Photonique et Nanostructures par Yong Chen, puis son émissivité a été mesurée au CEA-CESTA de Bordeaux. Les résultats sont spectaculaires, et montrent une source thermique émettant avec des lobes directionnels, comme une antenne infrarouge. Un excellent accord entre prédiction théorique et expérience a été obtenu [66]. Les premiers calculs numériques de réseaux ont été effectués par Jean-Philippe Mulet dans le cadre de sa thèse, débutée en janvier 2000. La suite du travail sur ce thème, visant à exploiter le phénomène pour concevoir et réaliser de nouvelles sources infrarouges, fait partie de la thèse de François Marquier débutée en octobre 2001. Ces deux thèses sont co-dirigées par Jean-Jacques Greffet et Karl Joulain, et je participe personnellement à l'encadrement de ces thèses à une hauteur d'environ 10-15%.

### 3.3 Transferts radiatifs en champ proche

Les résultats sur l'émission thermique en champ proche ont ouvert la voie à des travaux sur les transferts radiatifs entre objets à distance sub-longueur d'onde (nanométrique). Ces travaux se sont inscrits dans le cadre de la thèse de Jean-Philippe Mulet. Le premier système étudié consiste en deux milieux semi-infinis, séparés par un gap de largeur  $L$  (rempli d'un milieu transparent). Nous avons montré que pour des matériaux présentant des ondes de surface, le transfert radiatif est amplifié de plusieurs ordres de grandeur par rapport à sa valeur classique (obtenue pour des surfaces en champ lointain l'une de l'autre) [67, 68]. De plus, le transfert radiatif est quasi-monochromatique, ce qui n'est pas habituel en transfert thermiques entre deux corps chauffés.

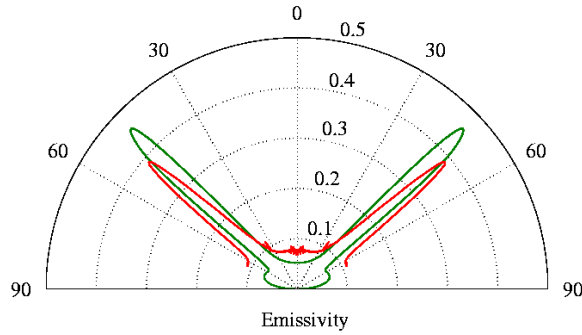


FIG. 3.2 – Diagramme polaire de l’émissivité d’un réseau lamellaire de  $\text{SiC}$  à  $\lambda = 11.36\mu\text{m}$ , et pour une émission polarisée  $p$ . Comparaison entre la courbe mesurée et la théorie. Le décalage vient du fait que la mesure est effectuée à  $T = 773\text{K}$  alors que le calcul utilise une valeur de constante diélectrique du matériau à  $T = 300\text{K}$ . On observe deux lobes d’émission similaires à ceux produits par une antenne. D’après [66].

Ces deux effets s’observent lorsque la distance  $L$  est très inférieure à la longueur d’onde du maximum de la fonction de Planck à la température considérée (typiquement, pour  $L < 100\text{nm}$  lorsque  $T \simeq 300\text{K}$ ). Ce résultat laisse entrevoir des applications intéressantes pour la production d’électricité par effet thermo-photovoltaïque (voir la section suivante “Perspectives”).

Un autre système étudié consiste en une pointe (dans notre modèle une petite sphère) échangeant de l’énergie par rayonnement avec un substrat plan. Une application visée par ce genre d’étude est le chauffage local d’un substrat par une pointe, en vue du stockage d’information haute-densité. Une autre application est la compréhension du mécanisme de formation du contraste en microscopie thermique à sonde locale (technique où la pointe mesure le flux thermique échangé avec l’échantillon par conduction et/ou rayonnement afin d’en déduire les propriétés locales). Dans les deux situations où la pointe chauffe un substrat froid, et où la pointe froide absorbe l’énergie radiative émise par le substrat chauffé, nous avons montré que pour les matériaux présentant des ondes de surface résonantes, le transfert est là-aussi fortement amplifié à courte distance et quasi-monochromatique [69]. L’amplification est importante lorsque la courbe de résonance (en fonction de la fréquence) du substrat et de la pointe (sphère) se recouvrent.

### 3.4 Perspectives

Les applications de ces travaux sont nombreuses, allant du contrôle des transferts radiatifs à la conception de nouvelles sources infrarouges. Il s’agit d’un thème très soutenu et porteur, dans lequel notre implication devrait croître ces prochaines années. Plus précisément, mon implication personnelle dans ce sujet à court et moyen terme concerne :

- L’étude d’une nouvelle structure pour la production d’électricité thermo-photovoltaïque.
- La modélisation de la microscopie thermique à sonde locale (et notamment son rapprochement

avec les techniques d'optique de champ proche utilisant du rayonnement infrarouge).

Ces deux thèmes entreront dans le cadre de la thèse de Marine Laroche (financement DGA), qui débutera en décembre 2002, et que je co-encadrerais avec Jean-Jacques Greffet à raison de 80%. La première application consiste à exploiter le fait qu'une source thermique en champ proche peut produire un rayonnement amplifié et quasi-monochromatique. L'éclairage d'une jonction à semi-conducteurs bien dimensionnée et accordée sur la fréquence d'émission pourrait permettre de concevoir de nouveaux types de cellules thermo-photovoltaïques ayant des rendements supérieurs aux cellules actuelles. Notons que l'idée d'utiliser du rayonnement en champ proche pour de tels systèmes existe déjà [70], mais que l'utilisation de matériaux présentant des résonances du type de celles que nous avons démontrées n'a jamais été proposée. Scientifiquement, ce projet est intéressant car il nécessite l'étude des couplages entre rayonnement et conduction thermiques dans des systèmes microstructurés (voire nanostructurés). En effet, la conduction interviendrait par exemple dans l'isolation de la source par rapport à la jonction qui doit rester à basse température pour ne pas faire chuter le rendement. Il nécessite également des compétences en physique des matériaux pour le photovoltaïque que nous ne possédons pas. Nous nous sommes donc associés au projet NOMACO-PV [71], coordonné par D. Mencaraglia (Supélec), et chargé de proposer des solutions pour les cellules photovoltaïques dites de 3ème génération. Une demande de soutien au programme Energie du CNRS a été faite, ainsi qu'une proposition d'intention pour un futur Réseau d'Excellence dans le cadre du 6ème PCRD de l'Union Européenne.

Concernant la seconde perspective, le travail a déjà partiellement démarré. Utilisant un modèle développé pour le SNOM [9, 24], nous avons montré qu'un microscope optique en champ proche mesurant le rayonnement thermique émis par un corps chauffé est en fait l'analogue optique du microscope électronique à effet tunnel [72]. De plus, nous avons montré qu'un microscope fonctionnant dans l'infrarouge (autour de  $\lambda = 10\mu m$ ) et utilisant une pointe métallique sans ouverture (technique dite "apertureless") devrait pouvoir mesurer un signal thermique émis par un substrat de SiC à  $T \simeq 300K$ . Une première tentative expérimentale est en cours, au Laboratoire d'Optique Physique de l'ESPCI (équipe de Claude Boccara). Nous avons également montré que ce type de mesure pourrait permettre de remonter à la constante diélectrique locale de l'échantillon mesuré, permettant ainsi d'effectuer une spectroscopie locale d'émission thermique [72]. Ce premier travail montre l'analogie qu'il peut y avoir entre le SNOM, le STM et un microscope thermique fonctionnant sous vide (et ne mesurant ainsi qu'un flux radiatif). Nous souhaitons pousser plus loin l'étude de la microscopie thermique traditionnelle, utilisant un transfert conductif et/ou radiatif, en utilisant notamment nos acquis sur la modélisation de la microscopie optique de champ proche. L'objectif est de mieux comprendre les mécanismes conduisant aux contrastes observés sur les images et de mieux percevoir les applications potentielles de ces techniques (et notamment leur complémentarité par rapport aux autres microscopies à sonde locale).

# Bibliographie

- [1] D.W. Pohl and D. Courjon (eds.), *Near-Field Optics*, NATO ASI Series (Kluwer, Dordrecht, 1993); M. Nieto-Vesperinas and N. García (eds.), *Optics at the Nanometer Scale*, NATO ASI Series (Kluwer, Dordrecht, 1996); M. Ohtsu (ed.), *Near-field Nano/Atom Optics and Technology* (Springer Verlag, Tokyo, 1998); S. Kawata, M. Ohtsu and M. Irie (eds.), *Nano-Optics* (Springer Verlag, Berlin, 2002).
- [2] J.M. Vigoureux, C. Girard and D. Courjon, *Opt. Lett.* **14**, 1039 (1989); C. Girard and D. Courjon, *Phys. Rev. B* **42**, 9340 (1990).
- [3] D. Van Labeke and D. Barchiesi, *J. Opt. Soc. Am. A* **9**, 732 (1992); *J. Opt. Soc. Am. A* **10**, 2193 (1993).
- [4] B. Hecht, H. Bielefeldt, Y. Inouye, D.W. Pohl and L. Novotny, *J. Appl. Phys* **81**, 2492–2498 (1997); V. Sandoghdar, S. Wegscheider, G. Krausch and J. Mlynek, *J. Appl. Phys* **81**, 2499–2503 (1997).
- [5] **R. Carminati, A. Madrazo, M. Nieto-Vesperinas and J.-J. Greffet**, *J. Appl. Phys* **82**, 501–509 (1997).
- [6] P.J. Valle, J.-J. Greffet and R. Carminati, *J. Appl. Phys.* **86**, 648–656 (1999).
- [7] P.J. Valle, R. Elaloufi, R. Carminati and J.-J. Greffet, “ Optical contrast, topographic contrast and artifacts in illumination-mode SNOM”, communicationn orale, *Fifth International Conference on Near Field Optics (NFO5)*, Shirahama (Japan), december 1998.
- [8] **J.-J. Greffet and R. Carminati**, *Prog. Surf. Science* **56**, 133–237 (1997), section 7.
- [9] **J.A. Porto, R. Carminati and J.-J. Greffet**, *J. Appl. Phys.* **88**, 4845–4850 (2000).
- [10] L. Aigouy, F.X. Andréani, A.C. Boccara, J.C. Rivoal, J.A. Porto, R. Carminati, J.-J. Greffet, V. Mathet and P. Beauvillain, *Appl. Phys. Lett.* **76**, 397–399 (2000).
- [11] **J.N. Walford, J.A. Porto, R. Carminati and J.-J. Greffet**, *J. Opt. Soc. Am. A* **19**, 572–583 (2002).
- [12] E. Betzig, J.K. Trautman, R. Wolfe, E.M. Gyorgy and P.L. Finn, *Appl. Phys. Lett.* **61**, 142–144 (1992).
- [13] L. Aigouy, S. Grésillon, A. Lahrech, A.C. Boccara, J.C. Rivoal, V. Mathet, C. Chappert, J.P. Jamet and J. Ferré, *J. Microsc.* **194**, 295–298 (1999); S. Grésillon, H. Cory, J.C. Rivoal and A.C. Boccara, *J. Opt. A* **1**, 178–184 (1999).

- [14] H. Wioland, O. Bergossi, S. Hudlet, K. Mackay and P. Royer, *EPJ Appl. Phys.* **5**, 289–295 (1999); O. Bergossi, H. Wioland, S. Hudlet, R. Deturche and P. Royer, *Jpn. J. Appl. Phys. Part 2* **38**, L655–L658 (1999).
- [15] J.N. Walford, J.A. Porto, R. Carminati, J.-J. Greffet, P.M. Adam, S. Hudlet, J.-L. Bijeon, A. Stashkevich and P. Royer, *J. Appl. Phys.* **89**, 5159–5169 (2001).
- [16] S.I. Bozhevolnyi, O. Keller and I.I. Smolyaninov, *Opt. Lett.* **19**, 1601–1603 (1994); S.I. Bozhevolnyi and B. Vohnsen, *Phys. Rev. Lett.* **77**, 3351–3354 (1996); C. Bainier, D. Courjon, F. Baida and C. Girard, *J. Opt. Soc. Am. A* **13**, 267–275 (1996).
- [17] **R. Carminati**, *Phys. Rev. E* **55**, **R4901–R4904** (1997).
- [18] R. Hillenbrand and F. Keilmann, *Phys. Rev. Lett.* **85**, 3029–3032 (2000).
- [19] E.R. Méndez, J.-J. Greffet and R. Carminati, *Opt. Comm.* **142**, 7–13 (1997).
- [20] **R. Carminati**, **M. Nieto-Vesperinas** and **J.-J. Greffet**, *J. Opt. Soc. Am. A* **15**, **706–712** (1998).
- [21] **R. Carminati**, **J.J. Sáenz**, **J.-J. Greffet** and **M. Nieto-Vesperinas**, *Phys. Rev. A* **62**, **12712** (2000).
- [22] M. Fink, *Physics Today* **50**, 34–40 (1997).
- [23] J. Bardeen, *Phys. Rev. Lett.* **6**, 57 (1961).
- [24] **R. Carminati** and **J.J. Sáenz**, *Phys. Rev. Lett.* **84**, **5156–5159** (2000).
- [25] **S. Gómez-Moñivas**, **J.J. Sáenz**, **R. Carminati** and **J.-J. Greffet**, *Appl. Phys. Lett.* **76**, **2955–2957** (2000).
- [26] S. Gomez-Monivas, L.S. Froufe, R. Carminati, J.-J. Greffet and J.J. Saenz, *Nanotechnology* **12**, 496–499 (2001).
- [27] R.X. Bian, R.C. Dunn and X.S. Xie, *Phys. Rev. Lett.* **75**, 4772 (1995); E.J. Sanchez, L. Novotny and X.S. Xie, *Phys. Rev. Lett.* **82**, 4014 (1999).
- [28] J. Azoulay, A. Débarre, A. Richard and P. Tchénio, *Europhys. Lett.* **51**, 374 (2000).
- [29] W.L. Barnes, *J. Mod. Opt.* **45**, 661 (1998); V.V. Klimov, M. Ducloy and V.S. Letokhov, *Quantum Electron.* **31**, 569 (2001).
- [30] J. Michaelis, C. Hettich, J. Mlynek and V. Sandoghdar, *Nature* **405**, 325 (2000).
- [31] D. Maystre, “Integral Methods”, in *Electromagnetic Theory of Gratings*, R. Petit (ed.) (Springer-Verlag, Berlin, 1980), p. 63.
- [32] M. Saillard and D. Maystre, *J. Opt. Soc. Am. A* **7**, 982 (1990); J.A. Sánchez-Gil and M. Nieto-Vesperinas, *J. Opt. Soc. Am. A* **8**, 1270 (1991); A.A. Maradudin, T. Michel, A.R. McGurn and E.R. Méndez, *Ann. Phys. (New York)* **203**, 255 (1990).
- [33] J.J.H. Wang, *Generalized Moment Methods in Electromagnetics* (Wiley, New York, 1991).

- [34] M. Thomas, R. Carminati, J.-J. Greffet, R. Arias and M. Nieto-Vesperinas, “Numerical study of the lifetime of an atom close to a lossy structure”, communication par poster, *Seventh International Conference on Near Field Optics (NFO7)*, Rochester (USA), august 2002.
- [35] A. Yodh and B. Chance, *Phys. Today* **48**, 34 (1995); S.K. Gayen and R.R. Alfano, *Opt. Photonics News* **7**, 17 (1996).
- [36] P. Sebbah (ed.), *Waves and Imaging through Complex Media* (Kluwer Academic, Dordrecht, 2001).
- [37] A. Mandelis, *Phys. Today* **53**, 29 (2000).
- [38] A. Ishimaru, *Wave Propagation and Scattering in Random Media* (IEEE Press, Piscataway, 1997).
- [39] **J. Ripoll, M. Nieto-Vesperinas and R. Carminati**, *J. Opt. Soc. Am. A* **16**, 1466–1476 (1999).
- [40] **J.B. Thibaud, R. Carminati and J.-J. Greffet**, *J. Appl. Phys.* **87**, 7638–7646 (2000).
- [41] **J. Ripoll, V. Ntziachristos, R. Carminati and M. Nieto-Vesperinas**, *Phys. Rev. E* **64**, 51917 (2001).
- [42] B.B. Das, K.M. Yoo and R.R. Alfano, *Opt. Lett.* **18**, 1092 (1993); F. Liu, K.M. Yoo and R.R. Alfano, *Opt. Lett.* **19**, 740 (1994).
- [43] K.M. Yoo, F. Liu and R.R. Alfano, *Phys. Rev. Lett.* **64**, 2647 (1990).
- [44] N.C. Bruce, F.E.W Schmidt, J.C. Dainty, N.P. Barry, S.C.W. Hyde and P.M.W. French, *Appl. Opt.* **34**, 5823 (1995).
- [45] R.H.J. Kop, P. de Vries, R. Sprik and A. Lagendijk, *Phys. Rev. Lett.* **79**, 4369 (1997).
- [46] K.K. Bizheva, A.M. Siegel and D.A. Boas, *Phys. Rev. E* **58**, 7664 (1998).
- [47] A. Wax, C. Yang, R.R. dasari and M.S. Feld, *Appl. Opt* **40**, 4224 (2001).
- [48] K. Furutsu and Y. Yamada, *Phys. Rev. E* **50**, 3634 (1994); M. Bassani, F. Martelli, G. Zaccanti and D. Contini, *Opt. Lett.* **22**, 853 (1997); T. Durduran, A.G. Yodh, B. Chance and D.A. Boas, *J. Opt. Soc. Am. A* **14**, 3358 (1997).
- [49] D.J. Durian, *Opt. Lett.* **23**, 1502 (1998); R. Graaff and J.J. Ten Bosch, *Opt. Lett.* **25**, 43 (2000).
- [50] R. Aronson and N. Corngold, *J. Opt. Soc. Am. A* **16**, 1066 (1999).
- [51] S. Chandrasekhar, *Radiative Transfer* (Dover, New York, 1960).
- [52] G.E. Thomas and K. Stamnes, *Radiative Transfer in the Atmosphere and Ocean* (Cambridge University Press, Cambridge, 1999), Chap. 8.
- [53] **R. Elaloufi, R. Carminati and J.-J. Greffet**, *J. Optics A: Pure and Applied Optics* **4**, S103–S108 (2002).



- [54] K.M. Case and P.F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Massachusetts, 1967).
- [55] **R. Elaloufi, R. Carminati and J.-J. Greffet, “Definition of the diffusion coefficient in scattering and absorbing media”, soumis à *J. Opt. Soc. Am. A* (2002).**
- [56] G. Maret and P.E. Wolf, *Z. Phys. B* **65**, 409 (1987); D.J. Pine, D.A. Weitz, P.M. Chaikin and E. Herbolzheimer, *Phys. Rev. Lett.* **60**, 1134 (1988); D.J. Pine, D.A. Weitz, J.X. Zhu and E. Herbolzheimer, *J. Phys. France* **51**, 2101 (1990).
- [57] L.A. Apresyan and Yu A. Kravtsov, *Radiation Transfer: Statistical and Wave Aspects* (Gordon and Breach, Amsterdam, 1996).
- [58] F.C. MacKintosh and S. John, *Phys. Rev. B* **40**, 2383 (1989).
- [59] W. Cai, M. Xu, M. Lax and R.R. Alfano, *Opt. Lett.* **27**, 731 (2002).
- [60] P.J. Hesketh, J.N. Zemel and B. Gebhart, *Nature* **325**, 549 (1986), *Phys. Rev. B* **37**, 10803 (1988).
- [61] J. Le Gall, M. Olivier and J.-J. Greffet, *Phys. Rev. B* **55**, 10105 (1997).
- [62] S.M. Rytov, Yu. A. Kravtsov and V.I. Tatarskii, *Principles of Statistical Radiophysics* (Springer Verlag, Berlin, 1989), vol. 3, chap. 3.
- [63] **R. Carminati and J.-J. Greffet, *Phys. Rev. Lett.* **82**, 1660–1663 (1999).**
- [64] C. Henkel, K. Joulain, R. Carminati and J.-J. Greffet, “Spatial coherence of thermal near fields”, *Opt. Comm.* **186** 57–67 (2000).
- [65] **A.V. Shchegrov, K. Joulain, R. Carminati and J.-J. Greffet, *Phys. Rev. Lett.* **85**, 1548–1551 (2000).**
- [66] **J.-J. Greffet, R. Carminati, K. Joulain, J.P. Mulet, S. Mainguy and Y. Chen, *Nature* **416**, 61–64 (2002).**
- [67] J.-P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, “Enhanced radiative heat transfer at nanometric distances”, in *Heat Transfer and Transport Phenomena in Microsystems*, edited by G.P. Celata (Begell House, New York, 2000), p. 352–357.
- [68] J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, Comment on “Radiative transfer over small distances from a heated metal”, *Opt. Lett.* **26**, 480–481 (2001).
- [69] **J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, *Appl. Phys. Lett.* **78**, 2931–2933 (2001).**
- [70] M.D. Whale, “Optimization of the spacing effect of thermal radiation for microscale thermophotovoltaic devices”, in *Heat Transfer and Transport Phenomena in Microsystems*, edited by G.P. Celata (Begell House, New York, 2000), p. 339.
- [71] D. Mancaraglia, “Nouveaux Matériaux et Concepts pour le Photovoltaïque”, Séminaire sur la Recherche et de le Développement Technique, CNRS-ADEME (Sophia Antipolis, novembre 2001).
- [72] **K. Joulain, R. Carminati, J.P. Mulet and J.-J. Greffet, “Near-field thermal-emission spectroscopy”, manuscrit en préparation (2002).**

## Quatrième partie

# Annexes

# Annexe 1

## Optique de champ proche

### Liste des articles reproduits :

- R. Carminati, A. Madrazo, M. Nieto-Vesperinas and J.-J. Greffet, *Optical content and resolution of near-field optical images: influence of the operating mode*, J. Appl. Phys. **82**, 501–509 (1997).
- J.-J. Greffet and R. Carminati, *Image formation in near-field optics*, Prog. Surf. Science **56**, 133–237 (1997), section 7.
- J.A. Porto, R. Carminati and J.-J. Greffet, *Theory of electromagnetic field imaging and spectroscopy in scanning near-field optical microscopy*, J. Appl. Phys. **88**, 4845–4850 (2000).
- J.N. Walford, J.A. Porto, R. Carminati and J.-J. Greffet, *Theory of near-field magneto-optical imaging*, J. Opt. Soc. Am. A **19**, 572–583 (2002).
- R. Carminati, *Phase properties of the optical near field*, Phys. Rev. E **55**, R4901-R4904 (1997).
- R. Carminati, M. Nieto-Vesperinas and J.-J. Greffet, *Reciprocity of evanescent electromagnetic waves*, J. Opt. Soc. Am. A **15**, 706–712 (1998).
- R. Carminati, J.J. Sáenz, J.-J. Greffet and M. Nieto-Vesperinas, *Reciprocity, unitarity and time-reversal symmetry of the S matrix of fields containing evanescent components*, Phys. Rev. A **62**, 12712 (2000).
- R. Carminati and J.J. Sáenz, *Scattering theory of Bardeen's formalism for tunneling: new approach to near-field microscopy*, Phys. Rev. Lett. **84**, 5156–5159 (2000).
- S. Gómez-Moñivas, J.J. Sáenz, R. Carminati and J.-J. Greffet, *Theory of electrostatic probe microscopy: a simple perturbative approach*, Appl. Phys. Lett. **76**, 2955–2957 (2000).

# Optical content and resolution of near-field optical images: Influence of the operating mode

R. Carminati,<sup>a)</sup> A. Madrazo, and M. Nieto-Vesperinas

*Instituto de Ciencia de Materiales, Consejo Superior de Investigaciones Científicas, Cantoblanco, 28049 Madrid, Spain*

J.-J. Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

(Received 2 January 1997; accepted for publication 1 April 1997)

Recent experimental work has shown that the contrast of near-field optical images depends on the path followed by the tip during the scan. This artifact may misguide the interpretation of the images and the estimation of the optical resolution. We provide a rigorous theoretical study of this effect based on three-dimensional perturbation theory and two-dimensional exact numerical calculations. We quantitatively study the dependence of the artifact on the illumination/detection conditions and on the scattering potential of the sample. This study should provide guidelines for future experimental work. © 1997 American Institute of Physics. [S0021-8979(97)06813-8]

## I. INTRODUCTION

Optical resolution beyond the Rayleigh limit has been demonstrated in the past ten years in scanning near-field optical microscopy (SNOM).<sup>1,2</sup> Among the various techniques that have been proposed, two categories can be distinguished: illumination-mode and collection-mode SNOM. In illumination-mode SNOM, a tip (nanosource) locally illuminates the sample and one collects the field scattered into the far zone.<sup>3,4</sup> Examples of nanosources are a tapered metal-coated optical fiber with an aperture at the tip apex,<sup>4</sup> or the tetrahedral tip introduced recently.<sup>5</sup> In collection-mode SNOM, the sample is illuminated by an extended field, as in classical microscopy, and the scattered near field is collected by a local probe. This probe can be the tip of an optical fiber,<sup>6</sup> or a scattering tip as that used in apertureless SNOM.<sup>7,8</sup> A particular case of collection-mode SNOM is the photon scanning tunneling microscope (PSTM or STOM) in which the sample is illuminated by an evanescent wave produced by total internal reflection.<sup>9</sup>

Let  $\mathcal{S}(x,y,z)$  be the optical signal that is detected when the tip is located at the point  $(x,y,z)$ . The  $z$  direction is chosen to be normal to the mean plane of the sample surface. In order to get sub-wavelength resolution, part of the signal must come from the conversion of evanescent waves into propagating waves. This holds whatever the technique. Thus, the tip (either illuminating or detecting) has to be kept at subwavelength distance from the sample during the scan. Three different operating modes have been used so far to regulate the tip-sample distance. (1) In the *constant-height mode*, the tip is moved in a plane  $z=z_0$ , and one records  $\mathcal{S}(x,y,z_0)$ . (2) In the *constant-intensity mode*, the optical signal  $\mathcal{S}$  is kept constant with a feedback system, forcing the tip to follow a surface  $z=h(x,y)$ . This surface does not in general reproduce the topography of the sample. Recording the motion of the tip [i.e., the surface  $z=h(x,y)$ ] produces the image. This mode has been used extensively in

PSTM/STOM.<sup>9</sup> (3) In the *constant-distance mode*, the tip is forced to follow a surface  $z=f(x,y)$  by an auxiliary non-optical distance-control mechanism. The optical signal that is recorded is  $\mathcal{S}[x,y,f(x,y)]$ . The distance-control mechanism can use a scanning tunneling microscope (STM),<sup>3,5</sup> an atomic force microscope (AFM),<sup>7</sup> or the lateral friction forces between the tip and the sample (shear forces).<sup>10,11</sup> In all these cases, the tip follows more or less the topography of the sample, and  $f(x,y)$  is the convolution of the sample profile by a function which describes the probe geometry.

A comparative study of modes (1) and (2) was presented in Ref. 12. It was shown that the two modes are equivalent, in the sense that a constant-intensity image  $z=h(x,y)$  and the constant-height image  $\mathcal{S}(x,y,z_0)$ , with  $z_0=\langle h(x,y) \rangle$ , are proportional. The brackets denote the background value of a function of  $(x,y)$ .

Concerning mode (3), the  $z$  motion of the probe, induced by a non-optical distance regulation mechanism, can couple to the purely optical information of the image.<sup>4,13</sup> This creates an artifact that may lead to a wrong interpretation of the images. It was demonstrated experimentally in Ref. 4 that the SNOM image may contain two different contributions: a purely optical one and one reproducing the motion of the tip. Recently, a systematic experimental study of this artifact was presented,<sup>13</sup> and showed that *many experimental images previously reported might be dominated by non-optical contrast mechanisms*. The origin of the artifact was discussed with a simple approach which we summarize here.<sup>13,14</sup>

In the constant-distance mode, we shall write the path followed by the tip as  $z=z_0+\delta f(x,y)$  where  $z_0=\langle f(x,y) \rangle$ . An expansion of the optical signal  $\mathcal{S}$  to first order in  $\delta f=\sup|\delta f(x,y)|$  leads to:

$$\begin{aligned} \mathcal{S}[x,y,z=f(x,y)] &= \langle \mathcal{S}(z_0) \rangle + \delta \mathcal{S}(x,y,z_0) \\ &+ \frac{\partial \mathcal{S}}{\partial z}(x,y,z_0) \delta f(x,y). \end{aligned} \quad (1)$$

The first term in Eq. (1) is the background value of the image, the second one gives a purely optical contrast and the

<sup>a)</sup>Electronic mail: rcarmina@icmm.csic.es

third one reflects the coupling between the optical information and the  $z$  motion of the tip. Due to this last term, the optical image will depend on the path followed by the tip  $\delta f(x,y)$ , leading to the presence of the artifact discussed in Ref. 13. Note that this problem does not occur in constant-height mode because  $\delta f(x,y)=0$  and the last term in Eq. (1) always vanishes.

As an illustration, let us consider a worst case scenario. Suppose, for example, that the third term dominates the right-hand side in Eq. (1) and that  $\partial \mathcal{A} / \partial z$  is a slowly varying function of  $(x,y)$ . Then, the detected signal  $\mathcal{S}$  given by the left-hand side of Eq. (1) is proportional to the motion of the tip  $\delta f(x,y)$ . The resulting image is mainly *an optical read-out of the motion of the tip*, as that obtained with a conventional AFM. This image does not contain any information on the optical properties of the sample. Moreover, the resolution of such an image does not result from optical mechanisms, but only from the interaction used to control the tip-sample distance.

The presence or not of this artifact, and its relative weight compared to the purely optical information, depend on both the experimental parameters (as the illumination conditions) and on the sample under study. As pointed out in Ref. 13, this makes the interpretation of constant-distance images a very difficult task. A precise study is necessary and constitutes the scope of the present work. Our purpose is to analyze rigorously the origin of the artifact and to identify the cases in which it may dominate the image contrast. The paper is organized as follows: In Sec. II, we consider the case of weak scattering samples, as that often used in SNOM. We use the three-dimensional perturbation theory to study analytically the origin of the artifact and to discuss the influence of both the experimental parameters and the sample. In Sec. III, we illustrate the discussion of Sec. II with exact two-dimensional numerical simulations based on a resolution of a volume integral (Lippmann–Schwinger) equation for the electric field. The samples studied in Sec. III consist of localized particles (dielectric or metallic) deposited on a flat dielectric substrate. In Sec. IV we study what happens when the sample is a very rough extended surface. In this case the scattered field cannot be described with perturbation theory. This discussion gives a complete picture of the scattering mechanism responsible for the presence of the artifact. Sec. V summarizes our conclusions.

## II. PERTURBATIVE MODEL FOR WEAKLY SCATTERING SAMPLES

In this section we analyze in detail the origin of the artifact presented in the introduction. We define this artifact as the presence in the detected signal of a cross term between the light scattered by the sample and the  $z$  motion of the tip. We will describe analytically the properties of the SNOM images in three dimensions, in the following context:

- (1) We focus the discussion on a collection-mode configuration in which the sample is illuminated by an extended field either in reflection or in transmission. After interaction with the sample, the near field is detected by a tip (see Fig. 1). This choice is in no way a limitation of the

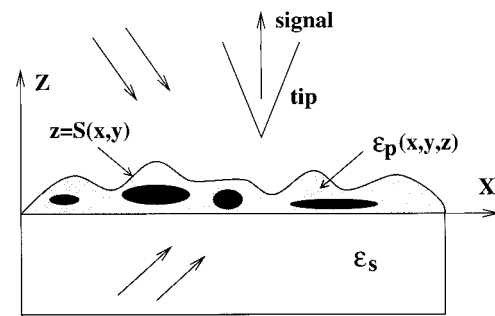


FIG. 1. Section of the three-dimensional geometry used in the perturbation theory.

generality of our study. Reciprocity can be used to extend all the results to the illumination-mode configurations. It has been shown that there exists an equivalent collection-mode setup for any illumination-mode setup.<sup>15</sup>

- (2) We assume that the probe is a passive point-like detector. This means that the signal  $\mathcal{S}(x,y,z)$  is proportional to the local near-field intensity, defined as the squared modulus of the electric field  $|\mathbf{E}(x,y,z)|^2$ . The passive probe assumption has been studied recently on a rigorous basis.<sup>16–18</sup> It was also demonstrated that the probe may be passive even if its presence modifies the near-field distribution around the sample.<sup>19,20</sup> On this basis, we do not take into account the presence of the tip. In what follows, we shall consider the structure of the near-field evaluated without the presence of the detecting (or illuminating) tip.

We point out that the coupling efficiency of the probe may depend on  $z$ . This occurs, for example, under  $p$ -polarized illumination, when the tip is very close to the sample (a precise study will be reported elsewhere). This  $z$  dependence may induce another kind of artifact. This artifact is not studied here because our model does not describe the coupling with the probe.

### A. Perturbative expression for the intensity

We consider a three-dimensional sample with variations in both topography and dielectric constant (Fig. 1). This sample is a layer of profile  $z=S(x,y)$  and inhomogeneous isotropic dielectric constant  $\epsilon_p(x,y,z)$ , deposited upon a semi-infinite homogeneous isotropic substrate of dielectric constant  $\epsilon_s$  (half space  $z<0$ ). The upper medium [ $z>S(x,y)$ ] is assumed to be a vacuum or air. The system is illuminated in transmission or reflection by a monochromatic field of wavelength  $\lambda$ . This field is either a plane wave (coherent illumination) or a set of uncorrelated plane waves (spatially incoherent illumination). In both cases, the intensity of the incident field depends only on  $z$ .

Let us write the total near field  $\mathbf{E}=\mathbf{E}^{(0)}+\mathbf{E}^{(1)}$  where  $\mathbf{E}^{(0)}$  is the field reflected or transmitted by the flat interface  $z=0$  and  $\mathbf{E}^{(1)}$  is the field scattered by the inhomogeneous layer. In many cases of practical interest in near-field optics (NFO), the object is weakly scattering so that the condition  $|\mathbf{E}^{(1)}|=\eta|\mathbf{E}^{(0)}|$ , with  $\eta\ll 1$ , is fulfilled. Instances in which

this condition is not satisfied will be examined in Sec. IV. When the tip is moved along a surface of equation  $z=f(x,y)$ , the detected intensity to first-order in  $\eta$  is:

$$I[x,y,z=f(x,y)] = I^{(0)}[f(x,y)] + I^{(1)}[x,y,f(x,y)], \quad (2)$$

where  $I^{(0)} = |\mathbf{E}^{(0)}|^2$  and  $I^{(1)} = 2 \operatorname{Re}[\mathbf{E}^{(0)*} \cdot \mathbf{E}^{(1)}]$ ,  $\operatorname{Re}$  denoting the real part and  $*$  the complex conjugate. A first-order expansion of Eq. (2) around  $z_0 = \langle f(x,y) \rangle$  leads to:

$$I[x,y,z=f(x,y)] = I^{(0)}(z_0) + \frac{dI^{(0)}}{dz}(z_0) \delta f(x,y) + I^{(1)}(x,y,z_0). \quad (3)$$

In order to determine the domain of validity of the preceding equation, we proceed as in Ref. 12. We introduce the length scales  $L_0$  and  $L_1$  of  $I^{(0)}$  and  $I^{(1)}$ , respectively, and  $\delta f = \sup|\delta f(x,y)|$ . Equation (3) is valid if  $\delta f \ll L_1$  and  $\delta f \sim \eta L_0$ . Note that these conditions involve only the structure of the near field, whatever the physical system which produces this near field. Thus, Eq. (3) applies to a large variety of problems. In the case of a sample with a linear inhomogeneous dielectric permittivity and an arbitrary surface profile, the conditions of validity of Eq. (3) are equivalent to those of first-order perturbation theory in the near field.<sup>12,21,22</sup>

Equation (3) reveals the content of the near-field optical image. The first term is independent on  $(x,y)$  and contributes to the background of the image. The two other terms are responsible for the contrast of the image, and two origins for this contrast are clearly identified. The second term is proportional to the path followed by the detecting tip  $\delta f(x,y)$ . The constant of proportionality depends only on the illumination conditions. Thus this term does not contain any information on the sample, and is only *an optical readout of the z motion of the tip*. In the best case, i.e., when the tip follows the topography of the sample, this term produces a signal proportional to the topography of the sample, and does not produce any additional information to the shear force, STM, or AFM signal. In contrast, the third term  $I^{(1)}(x,y,z_0)$  carries purely optical information on the sample properties (dielectric constant and topography). Its relationship to the topography and the dielectric constant variations of the sample is in general not simple. It can be described with the concepts of impulse response and equivalent surface profile.<sup>19</sup> This point will be useful in the following discussion.

## B. Optical content of the image

The right-hand side in Eq. (3) shows that the most general image is a superposition of a purely optical signal (third term) and a term proportional to the path followed by the tip (second term). Hence, Eq. (3) gives a rigorous theoretical basis to the experimental observations put forward in Refs. 4 and 13. The relative weight of these two terms determines the optical content of the near-field optical image. If the second term dominates, the use of NFO does not add any information to the AFM, STM, or shear force images. NFO is of interest only in the situations in which the contrast is dominated by the third term in Eq. (3).

Proceeding as in Ref. 13, we introduce the following parameter:

$$\Gamma = \frac{1}{I^{(1)}(x,y,z_0)} \frac{dI^{(0)}}{dz}(z_0) \quad (4)$$

which measures the optical content of the image.  $\Gamma=0$  corresponds to a purely optical image. A large value of  $|\Gamma|$  corresponds to a low optical content of the image, and a domination of the contrast by the artifact. In any relevant NFO image,  $|\Gamma|$  should be minimized.

The value of  $\Gamma$  depends on both the operating conditions and the properties of the sample under study. This makes the prediction of the presence of this artifact very difficult. We shall address separately the cases corresponding to an illumination in transmission with propagating waves, in transmission with evanescent waves, and in reflection.

### 1. Illumination in transmission with propagating waves

Let us first consider the situation in which the sample is illuminated in transmission from the lower medium (Fig. 1) with a field composed of one monochromatic plane wave (coherent illumination) or a set of uncorrelated plane waves (incoherent illumination), at an angle of incidence smaller than the critical one  $\theta_c = \arcsin[(\sqrt{\epsilon_s})^{-1}]$ . The illuminating field  $\mathbf{E}^{(0)}$  contains only homogeneous waves, and the illuminating intensity  $I^{(0)}$  does not depend on  $z$  ( $dI^{(0)}/dz=0$ ). Thus,  $\Gamma=0$  and no artifact is encountered.<sup>23</sup> Equation (3) shows that the image in this case is identical to a constant-height image, taken at the height  $z=z_0$ . This is rather surprising because one can have  $z_0 < \sup|S(x,y)|$ . In true constant-height mode, one always has  $z_0 > \sup|S(x,y)|$ . Thus, we expect that *the constant-distance image will look like a constant-height image, but with a better resolution*. This will be confirmed by the numerical simulation of Sec. III.

The previous remarks apply to collection-mode SNOM with an illumination in transmission with only homogeneous waves. By reciprocity<sup>15</sup> it also holds for illumination-mode SNOM in which the light is detected in transmission at angles smaller than the critical one  $\theta_c$  (“allowed light” in Ref. 24).

### 2. Illumination in transmission with evanescent waves

The situation in which some or all the plane waves of the incident field have an angle of incidence greater than the critical one is different. The corresponding zero-order transmitted waves are evanescent. Thus,  $\mathbf{E}^{(0)}$  contains inhomogeneous waves creating a  $z$  dependence in  $I^{(0)}$ . In the case of a single plane wave, the transmitted zero-order field is of the form  $\mathbf{E}_0 \exp(i\mathbf{k}_{\parallel}^{\text{inc}} \cdot \mathbf{r}_{\parallel} + i\gamma^{\text{inc}} z)$ , with  $\gamma^{\text{inc}} = (k_0^2 - \mathbf{k}_{\parallel}^{\text{inc}2})^{1/2}$ , and  $k_0 = \omega/c$ . We have used the notation  $\mathbf{r}_{\parallel} = (x,y)$ . An incidence in total internal reflection corresponds to  $|\mathbf{k}_{\parallel}^{\text{inc}}| > k_0$ . Thus  $\gamma^{\text{inc}}$  is imaginary, with the determination  $\operatorname{Im}(\gamma^{\text{inc}}) > 0$ ,  $\operatorname{Im}$  denoting the imaginary part. It follows that:

$$\Gamma = \frac{-2 \operatorname{Im}(\gamma^{\text{inc}})}{I^{(1)}(x,y,z_0)} |\mathbf{E}_0|^2 \exp[-2 \operatorname{Im}(\gamma^{\text{inc}}) z_0]. \quad (5)$$

It is worth noting that the numerator of  $\Gamma$  (i.e.,  $dI^{(0)}/dz$ ) is negative in this case. Hence the artifact appears as a superposition to the purely optical image of a signal proportional to the path followed by the tip in *inverse contrast*. This point will be illustrated by numerical simulation in Sec. III.

The present discussion applies to collection-mode techniques in which the sample is illuminated in transmission with part of (or all) the light being totally internally reflected, as in PSTM/STOM. By reciprocity,<sup>15</sup> it also applies to the illumination-mode configurations in which the light is detected in transmission at angles greater than the critical one (“forbidden light” in Ref. 24). The presence of the artifact in “forbidden light” images in illumination mode, and the fact that the non-optical signal represents the path followed by the tip in contrast reversal was demonstrated experimentally in Ref. 13. Our model explains this observation.

### 3. Illumination in reflection

Let us consider now an illumination in reflection from the upper medium (Fig. 1). The illuminating intensity  $I^{(0)}$  is an interference pattern between the incident field and the field reflected by the flat interface. Therefore  $I^{(0)}$  exhibits a  $z$  modulation which may lead to an artifact. This modulation will depend on the value of the reflection factor at the interface  $z=0$ . Keeping the same notations, in the case of a single incident plane wave, one obtains:

$$\Gamma = \frac{-4\gamma^{\text{inc}}}{I^{(1)}(x,y,z_0)} \text{Im}[\mathbf{E}_0^* \cdot \vec{\mathbf{r}} \mathbf{E}_0 \exp(2i\gamma^{\text{inc}}z_0)], \quad (6)$$

where  $\vec{\mathbf{r}}$  is a matrix of Fresnel reflection factors at the interface  $z=0$ . Here  $\gamma^{\text{inc}}$  is real. Equation (6) shows that the artifact may become important for large values of the reflection factors. Moreover, the sign of the numerator of  $\Gamma$  depends on the value of those factors, and may change from one sample to another. Thus the artifact may appear as a signal proportional to the path followed by the tip either in real or in inverse contrast.

The present discussion applies to collection-mode techniques in reflection. By reciprocity it also applies to the illumination-mode techniques in reflection.<sup>15</sup> In both cases, the artifact will be important if the substrate has a high reflectivity. We will illustrate this point in Sec. III.

### 4. Influence of the sample

The parameter  $\Gamma$  defined in Eq. (4) depends on the value of the purely optical contribution  $I^{(1)}$ . The aim of this section is to analyze the dependence of  $I^{(1)}$  on the sample properties.

$I^{(1)}$  is a complicated function of both the experimental conditions and the sample properties. The analysis through perturbation theory leads to the following expression:<sup>19</sup>

$$I^{(1)}(\mathbf{r}_{\parallel}, z_0) = \int H(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, \mathbf{k}_{\parallel}^{\text{inc}}, \epsilon_s, z_0) S_{\text{eq}}(\mathbf{r}'_{\parallel}) d\mathbf{r}'_{\parallel}. \quad (7)$$

$S_{\text{eq}}$  is an equivalent surface profile connecting the topography and the dielectric constant variations of the sample:

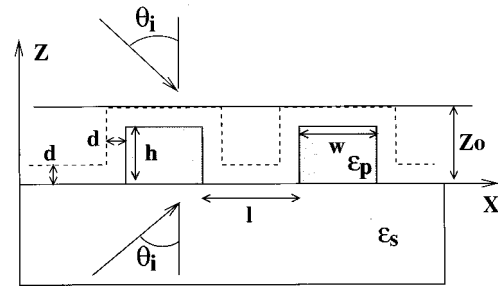


FIG. 2. Geometry of the two-dimensional sample used in the exact numerical calculations. Solid line: path followed in the constant-height calculations. Dashed line: path followed in the constant-distance calculations.

$$S_{\text{eq}}(\mathbf{r}_{\parallel}) = \frac{1}{\epsilon_s - 1} \int_0^{S(\mathbf{r}_{\parallel})} [\epsilon_p(\mathbf{r}_{\parallel}, z) - 1] dz. \quad (8)$$

Note that in the case of a homogeneous sample ( $\epsilon_p = \epsilon_s$ ),  $S_{\text{eq}}(\mathbf{r}_{\parallel})$  reduces to the true topographic profile  $S(\mathbf{r}_{\parallel})$ .  $H$  is an impulse response and is independent on the sample. Its existence is not postulated but comes out from the analysis through first-order perturbation theory.  $H$  is known analytically in Fourier space, its expression being given in Ref. 19.  $H$  depends on the illumination conditions (polarization, direction of incidence, coherence), on the dielectric constant of the substrate  $\epsilon_s$  and on the detection distance  $z_0$ . Therefore,  $H$  contains the dependence of  $I^{(1)}$  on the experimental conditions, while  $S_{\text{eq}}$  contains the properties of the sample. These concepts of impulse response and equivalent surface profile describe all the scattering process by weakly scattering samples and are very useful in the description of NFO imaging.<sup>19</sup>

Equation (7) shows that the value of  $I^{(1)}$  depends on the relative variations of the functions  $H(\mathbf{r}_{\parallel})$  and  $S_{\text{eq}}(\mathbf{r}_{\parallel})$ . Any situation may be encountered. At fixed experimental conditions (both  $H$  and  $dI^{(0)}/dz$  fixed), a sample may create a strong scattered intensity  $I^{(1)}$ , thus a small  $|\Gamma|$ , and another sample a low scattered intensity  $I^{(1)}$  and a large  $|\Gamma|$ . The same problem arises for a given sample by varying the illumination conditions. The conclusion is that it is not possible to give a universal rule governing the presence of the artifact in the image. Nevertheless, it is possible to study a typical sample in order to gain insight. This will be done in the next section.

## III. NUMERICAL RESULTS

In this section we study the near-field scattered by one or two particles deposited on a flat semi-infinite substrate. The scattering geometry is depicted in Fig. 2. This system is illuminated in transmission or in reflection by a monochromatic plane wave of wavelength  $\lambda = 633$  nm, with an angle of incidence  $\theta_i$ . We provide exact numerical calculations of the total near-field intensity either at a constant height or at a constant distance from the sample.

## A. Numerical method

The numerical scheme is based on a volume integral formulation of the electric field derived from Maxwell equations. At any point  $\mathbf{r}=(x,y,z)$ , the electric field is given by:<sup>25</sup>

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{(0)}(\mathbf{r}) + k_0^2 \int_{\Omega} [\epsilon_p(\mathbf{r}') - 1] \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}') \mathbf{E}(\mathbf{r}') d\mathbf{r}', \quad (9)$$

where  $\vec{\mathbf{G}}$  is the Green dyadic of the reference system consisting of a flat interface at  $z=0$  separating the substrate (half space  $z<0$ , dielectric constant  $\epsilon_s$ ) from vacuum (half space  $z>0$ ).  $\mathbf{E}^{(0)}$  is the field in the reference system and the integral gives an exact expression of the scattered field. It is restricted to the domain  $\Omega$  occupied by the particles, having a position-dependent dielectric constant  $\epsilon_p(\mathbf{r}')$ . In general Eq. (9) can only be solved numerically,<sup>26</sup> and several schemes have been proposed, in two-dimensional<sup>27</sup> and three-dimensional geometries.<sup>28</sup> Here we will consider a two-dimensional geometry, for both  $s$ - and  $p$ -polarized light, and solve Eq. (9) by a moment method.<sup>25</sup> In this method the volume  $\Omega$  is discretized in a mesh of rectangular cells of dimensions  $dx$  and  $dy$ . The field and the dielectric constant are assumed to be constant in each cell. Equation (9) is transformed into a linear system involving the integral of the Green dyadic over each cell. Note that this integration regularized the Green dyadic, which possesses a non-integrable singularity at the origin in  $p$  polarization.<sup>27</sup> In all the calculations presented here, the size of the cells is  $dx=dy=0.005\lambda$ .

## B. Images of one localized particle with different scattering potentials

We first consider the sample in Fig. 2 with one single particle. Its width is  $w=0.1\lambda$ , its height  $h=0.015\lambda$  and its dielectric constant  $\epsilon_p$ , assumed homogeneous, is a variable parameter. We compare the near-field intensity calculated along a line at a constant height (solid line in Fig. 2) and the intensity calculated at a constant distance from the surface profile (dashed line in Fig. 2). Our goal is to check the presence of the artifact due to the  $z$  motion of the tip in the constant-distance mode, in light of the discussion of Sec. II.

### 1. Illumination in transmission

The results in the case of an illumination in transmission at  $\theta_i=0^\circ$  are displayed in Fig. 3 ( $s$  polarization) and Fig. 4 ( $p$  polarization). We have used two different particles having the same dimensions but a different dielectric constant: (a)  $\epsilon_p=2.25$  (glass), (b)  $\epsilon_p=-9+i$  (gold). Varying  $\epsilon_p$  is a way to vary the scattering potential of the sample, and thus the level of the scattered intensity  $I^{(1)}$  [see Eqs. (7)–(8)]. The substrate is glass ( $\epsilon_s=2.25$ ). The solid curves correspond to constant-height calculations along the solid line in Fig. 2 with  $z_0=0.0225\lambda$  (14 nm). The dashed curves correspond to constant-distance calculations along the dashed line in Fig. 2 with  $d=0.0075\lambda$  (5 nm). The location of the particle is indicated at the bottom of the figures.

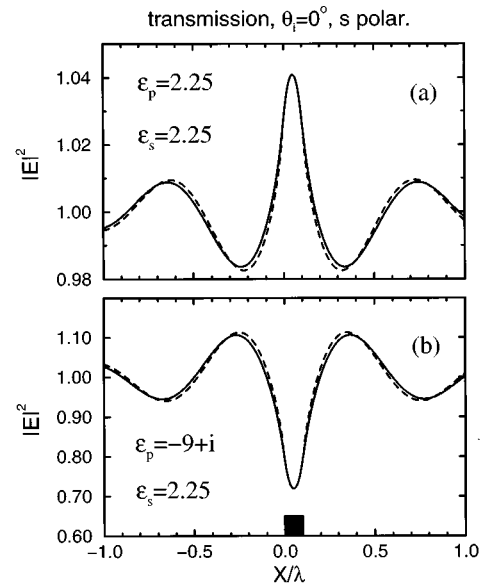


FIG. 3. Near-field intensity above the sample in Fig. 2 with only one surface particle.  $s$  polarization; illumination in transmission,  $\theta_i=0^\circ$ ; particle size:  $w=0.1\lambda$ ,  $h=0.015\lambda$ . Solid line: constant-height calculations with  $z_0=0.0225\lambda$ . Dashed line: constant-distance calculations with  $d=0.0075\lambda$ . Two values of the particle dielectric constant  $\epsilon_p$  are used. The dielectric constant of the substrate  $\epsilon_s=2.25$  is fixed.

In both polarizations and for the two particles, the constant-height and constant-distance curves are similar. This confirms the discussion of Sec. II. At normal incidence in transmission,  $dI^{(0)}/dz$  vanishes. Thus,  $\Gamma=0$  and the constant-distance image is purely optical. No  $z$ -motion artifact is to be expected and the constant-height and constant-distance images are similar. Yet the constant-distance curves exhibit a slightly better resolution than the constant-height curves. This is seen in Figs. 3 and 4 where the dashed curves exhibit faster variations than the solid curves. As discussed

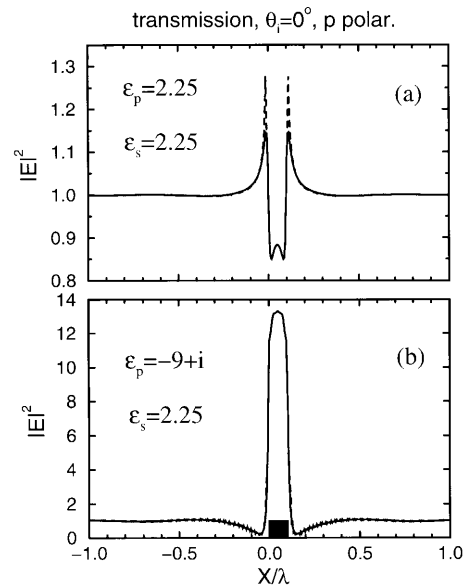


FIG. 4. Same as Fig. 3 for  $p$  polarization.



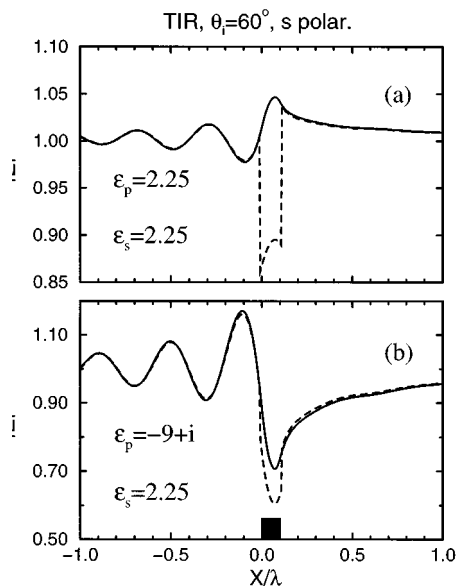


FIG. 5. Same as Fig. 3 with an illumination in total internal reflection ( $\theta_i = 60^\circ$ ).

in Sec. II, the constant-distant curve is a constant-height curve taken at the height  $\langle f(x, y) \rangle$  which is smaller than the height  $z_0$  of the constant-height curve. This explains the better resolution.

It is also worth noting the very high level of signal in the case of a metallic particle in  $p$  polarization [Fig. 4 (b)]. This is due to an enhancement of the field inside the particle, which creates a very strong scattered field. Even in this case there is no appreciable difference between the two kinds of images.

## 2. Illumination in total internal reflection

We show in Figs. 5 and 6 the same calculations as in Figs. 3 and 4, but for an incidence in total internal reflection ( $\theta_i = 60^\circ$ ). As seen in Sec. II,  $\Gamma$  does not vanish in this case, its value being given in Eq. (5). Moreover,  $dI^{(0)}/dz$  is negative, and one expects a contribution in the intensity of a term proportional to the path followed by the tip in inverse contrast [second term in Eq. (3)]. This is clearly seen in Figs. 5(a) and 6(a). Here the constant-height (solid line) and constant-distance (dashed line) curves look different. A superposition of a purely optical signal (looking like a constant-height signal) and a signal proportional to the path followed by the tip in contrast reversal (artifact) can be easily identified in the constant-distance curves. The scattering potential of the particle being small ( $\epsilon_p = 2.25$ ), the purely optical contribution  $I^{(1)}$  does not dominate the contrast of these curves. Thus they are strongly dominated by the  $z$ -motion artifact.

When the scattering potential increases [Figs. 5(b) and 6(b)], the contribution of the optical term  $I^{(1)}$  increases. For a gold particle [Figs. 5(b) and 6(b)], the purely optical contribution is so important that the presence of the artifact is strongly attenuated. This means that the contrast is now dominated by the third term in Eq. (3). The constant-height and constant-distance curves are practically identical. This

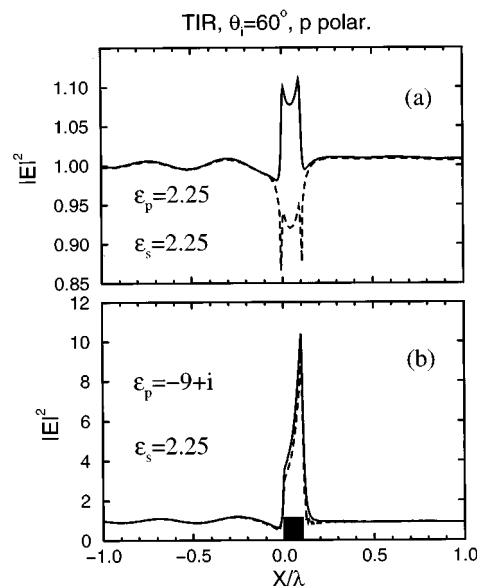


FIG. 6. Same as Fig. 5 for  $p$  polarization.

effect is particularly striking in  $p$  polarization [Fig. 6(b)] where the metallic particle creates a very strong scattered field. In this case the presence of the artifact is no longer visible.

## 3. Illumination in reflection

We now consider an illumination in reflection from the upper medium (see Fig. 2). We have shown in Sec. II that  $\Gamma$  does not vanish in this case. Its expression, given in Eq. (6), shows that its strength should increase with the reflectivity of the substrate. We thus present the calculations for only one kind of particle ( $\epsilon_p = 2.25$ ), but for two values of  $\epsilon_s$ : (a)  $\epsilon_s = 2.25$  and (b)  $\epsilon_s = 16$ . The result is displayed in Fig. 7 ( $s$  polarization). With the values of  $\epsilon_s$  used here, the Fresnel reflection factors appearing in the matrix  $\vec{r}$  in Eq. (6) are negative. Thus the numerator of  $\Gamma$  ( $dI^{(0)}/dz$ ) is positive and the  $z$ -motion artifact should appear as a superposition to the purely optical image of a signal proportional to the path followed by the tip (no contrast reversal). In Fig. 7(a) (low reflective substrate), the constant-height (solid line), and constant-distance (dashed line) curves are only slightly different. The contribution of the artifact in the constant-distance curve is lower than that of the purely optical term. Conversely, in Fig. 7(b) (high reflective substrate), the artifact appears clearly through the addition in the constant-distance intensity of a signal proportional to the path followed by the tip. This calculation confirms that in reflection the artifact is more important when the substrate has a high reflectivity.

Although we do not display the results for the sake of brevity, we have observed the same effect in  $p$  polarization. Moreover, increasing the scattering potential of the particle increases the contribution of the optical term in the detected signal. As in the case of TIR examined previously, this reduces the weight of the artifact in the image.

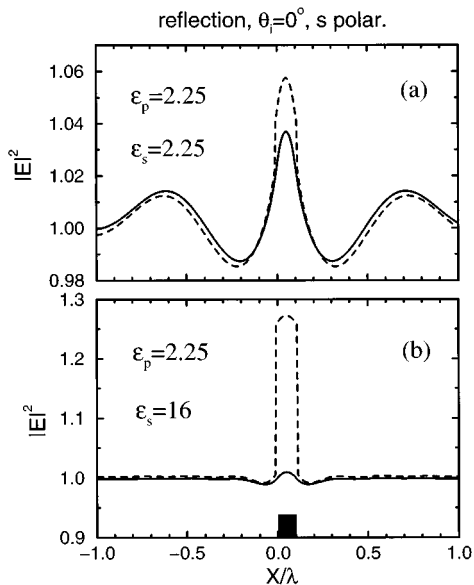


FIG. 7. Near-field intensity above the sample in Fig. 2 with only one surface particle;  $s$  polarization. Illumination in reflection from the upper medium,  $\theta_i=0^\circ$ . The particle dielectric constant  $\epsilon_p=2.25$  is fixed. Two values of the substrate dielectric constant  $\epsilon_s$  are used. Other parameters as in Fig. 3.

### C. Artifact and resolution

In the previous Section we have shown how the purely optical information of the image could be hidden by the artifact induced by the  $z$  motion of the probe. We shall now discuss the resolution issue.

It was pointed out in Ref. 13 that the artifact may lead to a wrong interpretation of the purely optical resolution. We illustrate this important point with exact numerical calculations of the field scattered by two particles ( $\epsilon_p=2.25$ ) deposited on a flat glass substrate ( $\epsilon_s=2.25$ ) and separated by a distance  $l=0.075\lambda$  (47 nm). The geometry is displayed in Fig. 2. The illumination is done in transmission.

Figures 8(a) and 8(b) show the constant-height (solid line) and constant-distance (dashed line) curves for  $\theta_i=0^\circ$  and  $\theta_i=60^\circ$ , respectively. The wave is  $s$  polarized. In Fig. 8(a) the presence of the artifact is not visible. As shown previously (see Fig. 3), at normal incidence the constant-height and constant-distance curves are almost identical. The constant-distance curve exhibits a slightly better resolution, as seen in Fig. 8(a). In contrast, in total internal reflection [Fig. 8(b)], the two images are clearly different. In the constant-height curve, the intensity distribution does not reproduce the surface profile. Even the presence of the two particles is not clear in this (purely optical) image. In contrast, the constant-distance curve exhibits strong variations at the precise location of the particles, with an extremely high resolution. But, these strong variations of the signal have their origin in the second term in Eq. (3), which is responsible for the artifact. In fact, a signal proportional to the path followed by the tip in inverse contrast can be easily recognized in the intensity at constant distance. This is the signature of the  $z$ -motion artifact with an illumination in total internal reflection in collection-mode SNOM (or by reciproc-

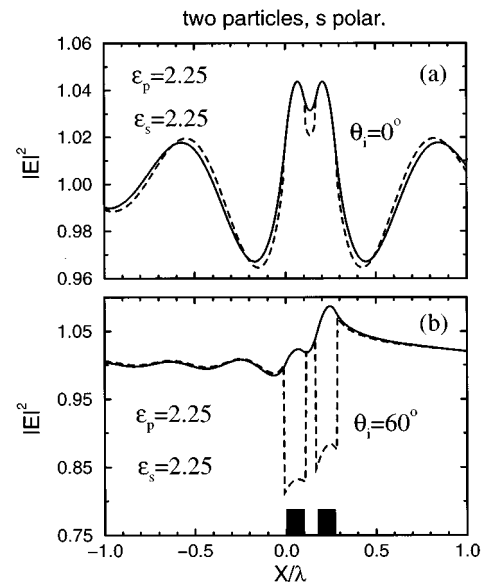


FIG. 8. Near-field intensity above the sample in Fig. 2 with two particles separated by a distance  $l=0.075\lambda$ . The dielectric constants  $\epsilon_s$  and  $\epsilon_p$  are fixed. (a):  $\theta_i=0^\circ$ ; (b):  $\theta_i=60^\circ$ . Other parameters as in Fig. 3.

ity with a ‘‘forbidden light’’ detection in illumination-mode SNOM). In conclusion, the resolution in the constant-distance image does not have its origin in an optical interaction with the sample, but in an optical readout of the  $z$  motion of the tip.

The same behavior is seen in Fig. 9 which is identical to Fig. 8 but for  $p$  polarization. Note that the separation between the two particles is clearly resolved in Fig. 9, even in the constant-height images. This is consistent with previous studies that showed that on dielectric substrates the light localization around the object was better in  $p$  than in  $s$

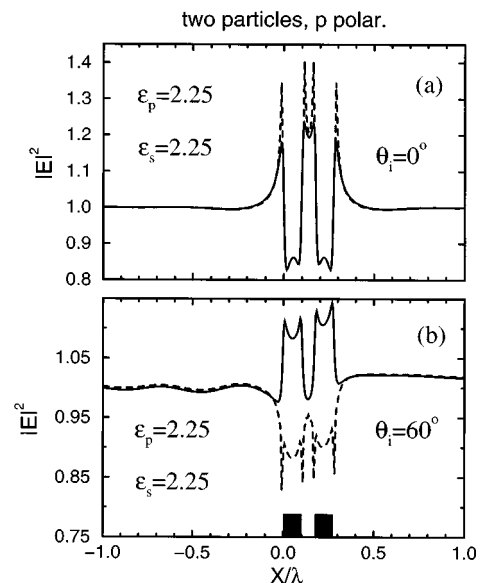


FIG. 9. Same as Fig. 8 for  $p$  polarization.

polarization.<sup>21,28</sup> Note that this polarization effect may be different with metallic substrates supporting polaritons.<sup>29</sup> In Fig. 9(b), due to the presence of the artifact, the constant-distance image appears in contrast reversal. But once again, this contrast reversal is a pure effect of the artifact, and does not reveal any particular optical property of the sample.

#### IV. CASE OF A STRONGLY SCATTERING EXTENDED SURFACE

The analysis in Sec. II, based on perturbation theory, shows that the origin of the  $z$ -motion artifact is the  $z$  variation of the illuminating intensity  $dI^{(0)}/dz$ . This term competes with the first-order scattered intensity  $I^{(1)}$  to produce the contrast of the image. In the perturbative analysis of Sec. II, the  $z$  variation of  $I^{(1)}$  was neglected because it was a second-order contribution. We have seen that this model predicts the behavior of the images of small localized particles, dielectric or metallic, as that studied numerically in Sec. III. This study is relevant for NFO applications.

Nevertheless, in order to get a complete picture of the scattering process responsible for the  $z$ -motion artifact, in this section we shall analyze the behavior of a strongly scattering extended non-flat surface. Note that, with constant-distance regulation, it is possible to study such highly corrugated samples with NFO microscopes.<sup>30</sup> In this case the scattered intensity is no longer weak compared to the illuminating intensity, and it may even dominate the total intensity. Hence, presently, the perturbative development used in Sec. II is not meaningful. This is a well-known fact in scattering from rough surfaces: when the roughness increases, the amount of energy in the specular direction (i.e.,  $I^{(0)}$ ) decreases, the energy being transferred to the scattered part of the field. In this case we may expect that the  $z$  variation of the scattered intensity may induce a new  $z$ -motion artifact.

In order to check this hypothesis, we have calculated the near-field scattered by a one-dimensional grating of profile  $z = S(x)$  with  $S(x) = h \cos(2\pi x/p)$ . The profile is represented on the top in Fig. 10. The upper medium is a vacuum, while the lower medium is assumed to be glass ( $\epsilon_s = 2.25$ ). The period is  $p = 0.5\lambda$ . The sample is illuminated in transmission from the lower medium with a monochromatic plane wave ( $\lambda = 633$  nm) at normal incidence.

We show in Fig. 10(a) the calculations when  $h = 0.01\lambda$ , in  $s$  polarization. The solid curve corresponds to a constant-height calculation, the dashed line to a constant-distance calculation. As expected, with this smooth grating, the conclusion of Secs. II and III remains valid. At normal incidence in transmission, the two curves are almost identical and no  $z$ -motion artifact may be detected in the constant-distance curve. It is so because the small grating height makes it a smooth scattering sample, for which the first-order perturbation theory is valid. Thus, Eq. (3) correctly describes the behavior of the near-field intensity and the  $z$  variation of the scattered intensity is a negligible second-order correction.

We show in Fig. 10(b) the same calculation with  $h = 0.15\lambda$ . The result is completely different. Even at normal incidence in transmission, the two curves are not identical. A  $z$ -motion artifact appears, but its origin is not the  $z$  variation

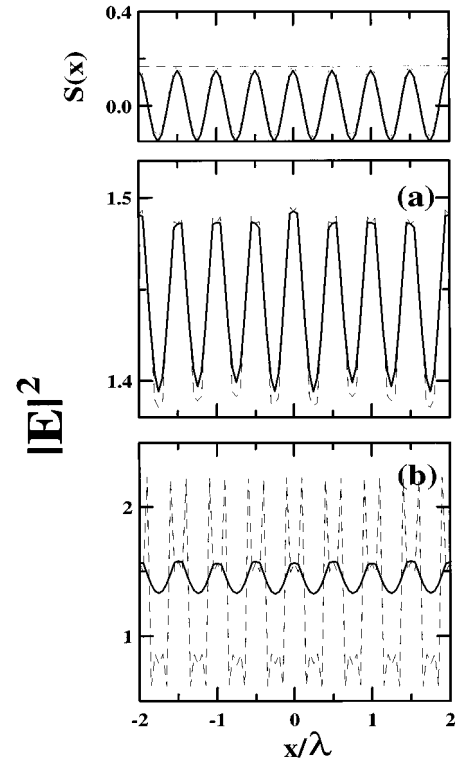


FIG. 10. Near-field intensity above a homogeneous surface of profile  $S(x) = h \cos(2\pi x/p)$ , with  $p = 0.5\lambda$ . Dielectric constant of the surface  $\epsilon_s = 2.25$ . Solid line: constant-height calculation. Dashed line: constant-distance calculation;  $s$  polarization:  $\theta_i = 0^\circ$ . (a):  $h = 0.01\lambda$ ; (b):  $h = 0.15\lambda$ .

of  $I^{(0)}$ . In fact, the mechanism is the following. The total electric field can be written  $\mathbf{E} = \mathbf{E}^{(0)} + \mathbf{E}_s$ . In this decomposition  $\mathbf{E}_s$  denotes the scattered field. The total intensity is now:

$$I(\mathbf{r}) = I^{(0)}(\mathbf{r}) + 2 \operatorname{Re}[\mathbf{E}^{(0)*}(\mathbf{r}) \cdot \mathbf{E}_s(\mathbf{r})] + I_s(\mathbf{r}), \quad (10)$$

where  $I_s = |\mathbf{E}_s|^2$ . The high-roughness grating creates a strong scattered field  $\mathbf{E}_s$ , and the last two terms in Eq. (10) contribute to the scattered intensity. The last one is not negligible, and may even dominate. Moreover, because  $\mathbf{E}_s$  contains both propagating and evanescent waves, the scattered intensity depends strongly on  $z$ , and induces the  $z$ -motion artifact proportional to the  $z$  derivative of the scattered intensity. Calculations of the total intensity versus  $z$  close to the surface (not displayed here) show that the derivative of the intensity is four times greater in the case in Fig. 10(b) than that in Fig. 10(a). This explains the visibility of a  $z$ -motion artifact in Fig. 10(b).

Also, not shown for the sake of brevity, the same results have been obtained in  $p$  polarization.

#### V. CONCLUSION

For weakly scattering samples, the scattering process does not induce any artifact in collection-mode SNOM when the sample is illuminated in transmission with propagating waves. This also holds for illumination-mode SNOM with a detection in transmission at angles smaller than the critical one. In collection-mode SNOM with some of the waves be-

ing totally internally reflected (like in PSTM/STOM), the artifact may become important. This is also true in illumination-mode SNOM with detection in transmission at angles greater than the critical one. In reflection SNOM, the artifact may also be encountered. It is important when the substrate has a high reflectivity. We have also shown that the purely optical contribution in the image increases with the scattering potential of the sample.

For strongly scattering extended surfaces, the scattering process induces an artifact even in transmission with illuminating (or detected) waves propagating with an angle of incidence smaller than the critical one.

The artifact we have studied is induced by the scattering process by the sample. As pointed out in the introduction, the probe coupling efficiency may depend on  $z$  and induces another kind of artifact, which was not taken into account here. Situations in which we did not find any artifact (of the first kind) may be affected in experiments by the probe-coupling artifact.

A safe way to avoid the artifact studied here would be to perform constant-height (or constant-intensity) images. Nevertheless, the constant-distance mode allows to perform a scan at a mean distance  $z_0$  smaller than the maximum height of the surface profile. Thus this mode provides the best potential optical resolution. We would like to address the problem of extracting the purely optical information from a constant-distance image exhibiting a  $z$ -motion artifact. In an experiment, one measures both the "optical signal" described by the left-hand side in Eq. (3) and the  $z$  motion  $\delta f(x,y)$  of the tip. The illuminating intensity  $I^{(0)}$  and its derivative  $dI^{(0)}/dz$  are known quantities that can be evaluated from the experimental parameters. Therefore solving for Eq. (3) gives a general procedure to extract the purely optical information  $I^{(1)}$  from the experimental signal. Such a procedure has been used experimentally in Ref. 31 on a collection-mode image obtained with an illumination in total internal reflection. A signal proportional to the path followed by the tip in inverse contrast was extracted from the image, leading to a purely optical signal. The main problem in this procedure is that the level of the purely optical signal in the input image has to be above the level of noise. Our study allows to find conditions in which the purely optical information is enhanced.

## ACKNOWLEDGMENTS

The authors thank D.W. Pohl for helpful discussions. This research has been supported by the Comision Interministerial de Ciencia y Tecnologia under Grant PB-95-0061A and by the EC. R.C. acknowledges financial support from the ESPRIT Program of the European Union. A.M. acknowledges a scholarship from Comunidad de Madrid.

- <sup>1</sup>*Near-Field Optics*, edited by D. W. Pohl and D. Courjon, NATO ASI Series E (Kluwer, Dordrecht, 1993), Vol. 242.
- <sup>2</sup>*Optics at the Nanometer Scale*, edited by M. Nieto-Vesperinas and N. García, NATO ASI Series E (Kluwer, Dordrecht, 1996), Vol. 319.
- <sup>3</sup>U. Dürig, D. W. Pohl, and F. Rohner, *J. Appl. Phys.* **59**, 3318 (1986).
- <sup>4</sup>E. Betzig, in *Near-Field Optics*, edited by D. W. Pohl and D. Courjon, NATO ASI Series E (Kluwer, Dordrecht, 1996), Vol. 242.
- <sup>5</sup>J. Koglin, U.Ch. Fischer, and H. Fuchs, in *Optics at the Nanometer Scale*, edited by M. Nieto-Vesperinas and N. García, NATO ASI Series E (Kluwer, Dordrecht, 1996), Vol. 319, p. 247.
- <sup>6</sup>E. Betzig, M. Isaacson, and A. Lewis, *Appl. Phys. Lett.* **51**, 2088 (1987).
- <sup>7</sup>N. F. van Hulst, M. H. P. Moers, and B. Bölger, *J. Microsc.* **171**, 95 (1993).
- <sup>8</sup>F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **65**, 1623 (1994); S. Kawata and Y. Inouye, *Ultramicroscopy* **57**, 313 (1995); R. Bachelot, P. Gleyzes, and A. C. Boccarda, *Opt. Lett.* **20**, 1924 (1995).
- <sup>9</sup>R. C. Reddick, R. J. Warmack, and T. L. Ferrel, *Phys. Rev. B* **39**, 767 (1989); D. Courjon, K. Sarayeddine, and M. Spajer, *Opt. Commun.* **71**, 23 (1989); F. de Fornel, J. P. Goudonnet, L. Salomon, and E. Lesniewska, *Proc. SPIE* **1139**, 77 (1989).
- <sup>10</sup>E. Betzig, P. L. Finn, and S. J. Weiner, *Appl. Phys. Lett.* **60**, 2484 (1992).
- <sup>11</sup>R. Toledo-Crow, P. C. Yang, and M. Vaez-Iravani, *Appl. Phys. Lett.* **60**, 2957 (1992).
- <sup>12</sup>R. Carminati and J.-J. Greffet, *Opt. Lett.* **21**, 1208 (1996).
- <sup>13</sup>B. Hecht, H. Bielefeldt, Y. Inouye, D. W. Pohl, and L. Novotny, *J. Appl. Phys.* **81**, 2492 (1997).
- <sup>14</sup>R. Carminati, Ph.D thesis, Ecole Centrale Paris, France, 1996.
- <sup>15</sup>E. R. Méndez, J.-J. Greffet, and R. Carminati, *Opt. Commun.* (in press).
- <sup>16</sup>R. Carminati and J.-J. Greffet, *Opt. Commun.* **116**, 316 (1995).
- <sup>17</sup>A. Madrazo and M. Nieto-Vesperinas, *J. Opt. Soc. Am. A* **13**, 785 (1996); *ibid.* **14**, 618 (1997).
- <sup>18</sup>J. C. Weeber, F. de Fornel, and J. P. Goudonnet, *Opt. Commun.* **126**, 285 (1996).
- <sup>19</sup>J.-J. Greffet and R. Carminati, in *Optics at the Nanometer Scale*, edited by M. Nieto-Vesperinas and N. García, NATO ASI Series E (Kluwer, Dordrecht, 1996), Vol. 319, p. 1.
- <sup>20</sup>A. Madrazo and M. Nieto-Vesperinas, *J. Opt. Soc. Am. A* (in press).
- <sup>21</sup>R. Carminati and J.-J. Greffet, *J. Opt. Soc. Am. A* **12**, 2716 (1995).
- <sup>22</sup>D. Barchiesi, C. Girard, O. J. F. Martin, D. Van Labeke, and D. Courjon, *Phys. Rev. E* **54**, 4285 (1996).
- <sup>23</sup>The  $z$  variation of  $I^{(1)}$  could induce an artifact, but this variation is of second order in  $\eta$  and is not taken into account in Eq. (3). It induces a visible effect only in the case of a highly scattering extended surface (see Sec. IV).
- <sup>24</sup>B. Hecht, H. Heinzelmann, and D. W. Pohl, *Ultramicroscopy* **57**, 228 (1995).
- <sup>25</sup>R. Harrington, *Field Computation by Moment Methods* (IEEE, Piscataway, NJ, 1993).
- <sup>26</sup>Approximate analytical expressions of the field can be derived from Eq. (9) by using the Born approximation. This leads to the first-order perturbation theory used in Sec. II. See Ref. 21 for more details.
- <sup>27</sup>F. Pincemin, A. Sentenac, and J.-J. Greffet, in *Near-Field Optics*, edited by D. W. Pohl and D. Courjon, NATO ASI Series E (Kluwer, Dordrecht, 1996), Vol. 242, p. 209; F. Pincemin, A. Sentenac, and J.-J. Greffet, *J. Opt. Soc. Am. A* **11**, 1117 (1994).
- <sup>28</sup>O. J. F. Martin, C. Girard, and A. Dereux, *Phys. Rev. Lett.* **74**, 526 (1995).
- <sup>29</sup>A. Madrazo and M. Nieto-Vesperinas, *Appl. Phys. Lett.* **70**, 31 (1997).
- <sup>30</sup>R. Toledo-Crow, B. Smith, J. Rogers, and M. Vaez-Iravani, *Proc. SPIE* **2196**, 62 (1994).
- <sup>31</sup>Y. Martin, F. Zenhausern, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **68**, 2475 (1996).



## IMAGE FORMATION IN NEAR-FIELD OPTICS

JEAN-JACQUES GREFFET\* and RÉMI CARMINATI†

Laboratoire d'Energétique Macroscopique et Moléculaire, Combustion.  
Ecole Centrale Paris, Centre National de la Recherche Scientifique,  
92295 Châtenay-Malabry Cedex, France

### Abstract

An overview is presented of the image formation theory in near-field optical microscopy. The emphasis is placed on the basic concepts and the understanding of the images. We briefly recall the general principles used in near-field optics to break the resolution limit. Since some of the concepts widely used in optics become meaningless in near field, a brief critical review of basic concepts is given. A theory of scattering of electromagnetic waves by inhomogeneous surfaces is then presented. For objects much smaller than the wavelength, a closed-form expression of the scattered field is derived, which provides a link between the near field and the structure of the sample. The different set-ups and their imaging capabilities are analysed. A general relationship between the signal and the induced currents in the sample is derived by means of the reciprocity theorem. The set-ups are compared and an equivalence between illumination and collection mode is proven. It is shown that, when multiple scattering between the sample and the rest of the system can be neglected, an impulse response can be defined for the three different types of set-ups : illumination mode, collection mode and apertureless. The importance of coherence in the near field is studied. Finally, the influence of the different control modes (constant height, constant intensity, constant tip-sample distance) is analysed and the existence of artifacts is discussed.

---

\*E-mail : jjg@em2c.ecp.fr

†E-mail : remi@em2c.ecp.fr

we have seen that the reciprocal set-up must use an incoherent source. In addition, the symmetry of the reciprocal source must be the same as the symmetry of the detector. In contrast, the typical PSTM/STOM uses a highly directional and coherent illumination, which explains why the images in illumination-mode SNOM resemble more closely the actual structure of the sample than those obtained in PSTM/STOM. It has been shown both experimentally [144] and theoretically [145] that, under partially coherent and symmetric illumination, the PSTM/STOM also produces images that closely resemble the sample structure [144, 145].

## E. Conclusion

Using the reciprocity theorem of electromagnetism, we have demonstrated that the illumination and collection modes of the SNOM are equivalent, which means that they have the same potential imaging capabilities and resolution. We have given a general rule to construct the equivalent collection-mode set-up to any illumination-mode set-up.

This result demonstrates that the differences that have been observed between the two modes are only due to differences in instrumentation and operating conditions. For example, the images in one case may present a better resolution, because of the use of coated fibers instead of uncoated ones, because the operating modes allows a smaller tip-sample distance, or because the degree of coherence of the illumination/detection scheme is not the same. But these differences are not fundamental. The same operating mode can be used in illumination and in detection mode. The same probes can also be used in both modes. Moreover, the coherence of the PSTM/STOM, for example, can be reduced [144, 146, 112, 147].

## 7. Image Formation

We now address the problem of the image formation with near-field optical microscopes. How the image is related to the optical properties and the topography of the sample has remained (to a large extent) an open question for the last decade. In this section, we derive a closed-form expression of the link between the detected signal and the structure of the sample.

Starting with the reciprocity theorem of electromagnetism, we obtain an *exact* expression for the signal, where all multiple interactions between the tip and the sample are taken into

account. We show that the signal is related to the induced current density in the sample, or to the induced polarization for non-magnetic materials. We give an explicit expression of the signal for the three main families of near-field optical set-ups, namely, collection-mode, illumination-mode and apertureless SNOM.

Finally, we address the problem of the existence of an impulse response (in direct space) or a transfer function (in Fourier space) in the image formation process in near-field optics. A number of authors have introduced different impulse responses or transfer functions [148, 136, 98, 149, 150, 151, 152, 153]. One of the major interest of the concept of impulse response is that it enables the inverse problem to be solved [98, 99, 145, 152, 153]. The existence of transfer functions in near-field optics is not trivial. One may define a transfer function between the signal and the near field (either a component or the square modulus of the electric field, for instance). It may also be possible to define a transfer function between the signal and the sample profile. The problem is more involved, if there is both topography and dielectric contrast. In that case, the concept of equivalent surface profile may be useful [99]. The necessary requirements for the existence of such a transfer function are linearity and translational invariance. The first requirement is satisfied in linear optics. The second implies that the effect of the tip is the same at any point of the sample. In other words, multiple scattering between the tip and the sample should be negligible. We will show that the concept of transfer function arises naturally, when using the reciprocity theorem and a single-scattering approximation (first Born approximation).

## A. Collection mode

In collection-mode SNOM, the sample is usually illuminated by a laser focussed onto a part of the sample. An optical fiber (waveguide) with a sharp tip is scanned along the sample, at subwavelength distance, and the signal delivered by an optical detector at the end of the fiber is recorded. The scan can be performed at a constant distance from the mean surface (constant-height mode), or following the topography (constant-distance mode). An alternative procedure is to keep the optical signal constant by varying the distance between the tip and the sample, and to record the tip position (constant-intensity mode). It will be shown in §9 that the constant-height and constant-intensity modes are equivalent, provided that the maximum height of the sample surface is much smaller than the wavelength [154]. Moreover, we will see

that the constant-distance mode may induce topographical artifacts in the signal, due to the  $z$ -motion of the probe [155, 156, 157]. For these reasons, we will assume here that the scan is performed in the *constant-height mode*.

The basic question addressed here is : “what is measured by a near-field microscope in the collection mode ?” This involves two problems, namely, how the electromagnetic field is coupled to the guide through the tip and how the near-field is related to the structure of the sample. It has been often assumed that the signal is related to the square modulus of the electric field above the sample (see e.g. ref. [50]). Yet, there is no reason to believe that the different components of the field will have the same coupling efficiency. Van Labeke and Barchiesi [135] have proposed a heuristic model that predicts different coupling factors for different polarizations. The present work confirms the need for accounting for polarization and explicitly shows how to do it. Moreover, it gives an expression of the signal, showing its dependence on the sample structure.

### (i) Basic expression of the signal

We establish the relationship between the signal delivered by a collection-mode SNOM and the current distribution inside the sample. To this aim, we use the reciprocity theorem of electromagnetism, and some standard relationships on guided waves, following the notation of ref.[80]. All the fields considered here are monochromatic, with frequency  $\omega$ . A temporal dependence  $\exp(-i\omega t)$  is assumed and omitted everywhere.

Inside an infinitely long waveguide with axis  $Oz$ , the field can be decomposed into modes  $\mathbf{E}_\lambda^{(+)}$  propagating towards  $z > 0$ , and modes  $\mathbf{E}_\lambda^{(-)}$  propagating towards  $z < 0$ . Thus, in a general situation, the field is :

$$\mathbf{E} = \sum_\lambda A_\lambda^{(+)} \mathbf{E}_\lambda^{(+)} + \sum_\lambda A_\lambda^{(-)} \mathbf{E}_\lambda^{(-)}, \quad (7.1)$$

where  $A_\lambda^{(\pm)}$  is the amplitude of the mode  $\mathbf{E}_\lambda^{(\pm)}$ . If a current distribution, described by the current density  $\mathbf{j}(\mathbf{r})$ , is present in the guide and excites several modes, the mode  $\mathbf{E}_\lambda^{(+)}$  has an amplitude  $A_\lambda^{(+)}$  given by [80]

$$A_\lambda^{(+)} = \frac{-2\pi Z_\lambda}{c} \int \mathbf{j}(\mathbf{r}) \cdot \mathbf{E}_\lambda^{(-)}(\mathbf{r}) d\mathbf{r}, \quad (7.2)$$

where the integral is performed over the volume of the current distribution. Equation (7.2) is



valid for a waveguide of arbitrary section, with perfectly conducting walls, and filled with a homogeneous material with dielectric constant  $\epsilon_p$  and relative permeability  $\mu_p$ . The impedance  $Z_\lambda$  is defined by  $Z_\lambda = ck_\lambda/\epsilon_p\omega$  for TM polarization and  $Z_\lambda = \mu_p\omega/ck_\lambda$  for TE polarization,  $k_\lambda$  being the wave vector of the mode  $\lambda$  along the  $z$ -axis. In what follows, we will consider a SNOM probe, i.e., a finite waveguide with a sharp tip at one extremity. For simplicity, we will assume that *far from the tip*, the guiding part of the probe is a waveguide with perfectly conducting walls. Thus, far from the tip, the field expansion (7.1), together with (7.2), is valid to describe the field reaching the detector.

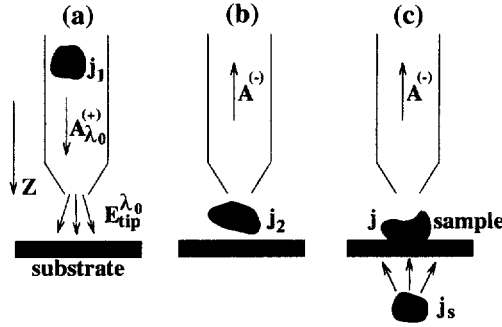


Fig. 13: Sketch of the system. (a) The current  $\mathbf{j}_1$  emits a field  $\mathbf{E}_{tip}^{\lambda_0}$  in the presence of the flat substrate, (b) A current distribution  $\mathbf{j}_2$  excites a field propagating into the waveguide, (c) A general situation where  $\mathbf{j}_2$  is splitted into a source with an externally driven current density  $\mathbf{j}_s$  and a sample with an induced current density  $\mathbf{j}$ .

Let us assume that the probe is placed in front of a flat substrate, with the tip apex, or the probe aperture, located at a point  $\mathbf{r}_{tip}$  (Fig. 13). This system (probe + flat substrate) is our reference system. In order to derive a general relationship between the amplitudes of the modes and the induced sources inside the sample, we apply the reciprocity theorem. To this aim, we consider two different configurations 1 and 2, involving the reference system with different sources [Figs. 13 (a) and (b)]. First, we assume that a *suitable current distribution*  $\mathbf{j}_1$  placed inside the guide and *far from the tip* excites a *single mode* propagating towards  $z > 0$ , with label  $\lambda_0$ , and amplitude  $A_{\lambda_0}^{(+)} = 1$  [Fig. 13(a)]. Using (7.2), this implies that :

$$\int_{V_1} \mathbf{j}_1(\mathbf{r}) \cdot \mathbf{E}_\lambda^{(+)}(\mathbf{r}) d\mathbf{r} = 0$$

$$\int_{V_1} \mathbf{j}_1(\mathbf{r}) \cdot \mathbf{E}_\lambda^{(-)}(\mathbf{r}) d\mathbf{r} = \frac{-c}{2\pi Z_\lambda} \delta_{\lambda\lambda_0}, \quad (7.3)$$

where  $V_1$  is the volume of the current distribution  $\mathbf{j}_1$ , and  $\delta_{\lambda\lambda_0} = 1$  if  $\lambda = \lambda_0$ ,  $\delta_{\lambda\lambda_0} = 0$  if  $\lambda \neq \lambda_0$ . Part of the mode is transmitted at the tip and produces a field denoted by  $\mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip})$ .

In a second configuration, we consider a current distribution  $\mathbf{j}_2(\mathbf{r}, \mathbf{r}_{tip})$ , of finite volume  $V_2$ , placed outside of the probe in the reference system, e.g., between the tip and the substrate [Fig. 13(b)]. This current distribution radiates a field  $\mathbf{E}_2(\mathbf{r}, \mathbf{r}_{tip})$ . Both  $\mathbf{j}_2$  and  $\mathbf{E}_2$  depend on  $\mathbf{r}_{tip}$  because the tip is a scatterer that modifies the field and reacts on the source (multiple scattering). The reciprocity theorem states that (see Appendix D) :

$$\int_{V_1} \mathbf{j}_1(\mathbf{r}) \cdot \mathbf{E}_2(\mathbf{r}, \mathbf{r}_{tip}) d\mathbf{r} = \int_{V_2} \mathbf{j}_2(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r}. \quad (7.4)$$

The field  $\mathbf{E}_2$  in the left-hand side in (7.4) is evaluated inside the volume  $V_1$ , i.e., inside the guide and far from the tip, so that expansion (7.1) can be used. Moreover, the dot product of  $\mathbf{j}_1$  with each mode of the expansion can be evaluated using (7.3). This leads to .

$$A_{\lambda_0}^{(-)}(\mathbf{r}_{tip}) = \frac{-2\pi Z_{\lambda_0}}{c} \int_{V_2} \mathbf{j}_2(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r}. \quad (7.5)$$

Equation (7.5) is a rigorous expression of the amplitude of the mode  $\mathbf{E}_{\lambda_0}^{(-)}$  excited by an arbitrary current distribution  $\mathbf{j}_2$ . The only restriction to the validity of this result is that the media should satisfy reciprocity [141] (see also Appendix D). Let us recall that  $\mathbf{E}_{tip}^{\lambda_0}$  is the field *emitted* by the tip used in illumination mode, when only the mode  $\mathbf{E}_{\lambda_0}^{(+)}$  is excited. Note that, in general,  $\mathbf{j}_2$  excites several modes and that (7.5) gives the amplitude of each mode. In (7.5),  $\mathbf{j}_2$  is the current density *in the presence of the tip and the substrate*. At this step, we have not neglected multiple interactions between the tip, the substrate, and the current distribution  $\mathbf{j}_2$ . Let us remark that the general expression of the amplitude (7.5) does not involve the near field at the tip position. It establishes a direct relationship between the amplitude of the mode and the *current distribution*  $\mathbf{j}_2$ .

## (ii) Case of a point dipole

Let us examine the coupling of a radiating point dipole  $\mathbf{p}_2$  with the probe/substrate system. This corresponds to a current density  $\mathbf{j}_2(\mathbf{r}) = -i\omega\mathbf{p}_2\delta(\mathbf{r} - \mathbf{r}_2)$ , where  $\mathbf{r}_2$  is the position of the

dipole. Equation (7.5) reduces to :

$$A_{\lambda_0}^{(-)} = 2\pi Z_{\lambda_0} \frac{i\omega}{c} \mathbf{p}_2 \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r}_2 - \mathbf{r}_{tip}) . \quad (7.6)$$

Equation (7.6) is an exact expression of the amplitude of the mode excited by a point dipole, in near-field coupling with the probe. Note that all multiple interactions are taken into account in this expression, namely,  $\mathbf{p}_2$  is the dipole moment in the presence of the probe and the substrate. This result could be useful, for instance, in the analysis of the fluorescence signal emitted by a single molecule, where the radiating dipole model is a good approximation. In particular, (7.6) involves a dot product, thus showing the sensitivity of the signal to the orientation of the molecule, as observed experimentally [38]. The dependence of the signal on the dipole orientation is also crucial when one uses the molecule lifetime as a source of contrast, this lifetime being strongly affected by the dipole orientation with respect to the surrounding objects [158].

### (iii) Case of a mesoscopic sample

The current density  $\mathbf{j}_2$  in (7.5) can be splitted into the current density  $\mathbf{j}_s$  of the sources illuminating the system and the current density  $\mathbf{j}$  of the sample under investigation, assumed to be deposited on the flat substrate [Fig. 13(c)]. In an experiment,  $\mathbf{j}_s$  is the *externally driven* current density of the sources, placed in the far field of the sample. In particular,  $\mathbf{j}_s$  is assumed to be unaffected by the presence of the sample. For the most general sample, the current density  $\mathbf{j}$ , which depends on the tip position, is :

$$\mathbf{j}(\mathbf{r}, \mathbf{r}_{tip}) = -i\omega \mathbf{P}(\mathbf{r}, \mathbf{r}_{tip}) + \nabla \times \mathbf{M}(\mathbf{r}, \mathbf{r}_{tip}) , \quad (7.7)$$

where  $\mathbf{P}$  and  $\mathbf{M}$  are the (position- and frequency-dependent) polarization and magnetization densities in the sample, respectively. For a local and linear material, they can be related to the electric and magnetic fields,  $\mathbf{E}$  and  $\mathbf{H}$ , by the local constitutive relations  $\mathbf{P} = \epsilon_0(\overset{\leftrightarrow}{\epsilon} - 1)\mathbf{E}$  and  $\mathbf{M} = (\overset{\leftrightarrow}{\mu} - 1)\mathbf{H}$ , where  $\overset{\leftrightarrow}{\epsilon}$  and  $\overset{\leftrightarrow}{\mu}$  are the dielectric and magnetic tensors of the sample, respectively. As usual in optics,  $\overset{\leftrightarrow}{\epsilon}$  is a generalized dielectric tensor, that accounts for both the free and bound charges of the material. Thus, (7.5) can describe the dielectric, metallic and magnetic response of the sample, which may be anisotropic.

For the sake of simplicity, we limit the following discussion to a non-magnetic, isotropic sample. In this case, the polarization density is  $\mathbf{P} = \epsilon_0(\epsilon - 1)\mathbf{E}$ , where  $\epsilon$  is the position-

and frequency-dependent dielectric constant of the sample. Dropping the superscript  $(-)$  for brevity, (7.5) can be cast in the form :

$$A_{\lambda_0}(\mathbf{r}_{tip}) = A_{\lambda_0}^{bg} + A_{\lambda_0}^{sa}(\mathbf{r}_{tip}), \quad (7.8)$$

where :

$$A_{\lambda_0}^{bg} = \frac{-2\pi Z_{\lambda_0}}{c} \int_{V_s} \mathbf{j}_s(\mathbf{r}) \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} \quad (7.9)$$

is the background component of the signal ( $V_s$  denotes the volume of the external sources) and

$$A_{\lambda_0}^{sa}(\mathbf{r}_{tip}) = \frac{2\pi Z_{\lambda_0}}{c} i\omega\epsilon_0 \int_V [\epsilon(\mathbf{r}) - 1] \mathbf{E}(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r}. \quad (7.10)$$

is the contribution of the sample ( $V$  is the volume of the sample).

The first term (7.9) does not depend on the sample properties. When the tip scans the sample *at a constant height*, this term contributes to the background of the mode amplitude only. Nevertheless, if the tip is forced to move along the  $z$  direction, for instance by the tip-sample distance regulation mechanism, this term may vary and induce a variation in the mode amplitude, this variation being *independent on the sample*. This  $z$ -motion artifact will be studied in details in §9 (see also refs.[155, 156, 157]).

The second term (7.10) describes the purely optical contribution. Assuming that the scan is performed at a constant height, this term is the only dynamic part of the mode amplitude. It relates the mode amplitude to the dielectric contrast in the sample. Let us stress that the dielectric contrast, as written in (7.10), shows the dependence on both the optical properties (dielectric constant) and the topography of the sample. Indeed, the function  $[\epsilon(\mathbf{r}) - 1]$  vanishes out of the sample domain.

#### (iv) What is measured ?

In optics, one detects the time-averaged intensity rather than the amplitude of the field. Thus, the detected signal is the energy flux in the guide. For simplicity, we consider a monomode guide, so that the signal is proportional to  $|A_{\lambda_0}|^2$ . Accounting for the difference of magnitude between the two terms  $A_{\lambda_0}^{(back)}$  and  $A_{\lambda_0}^{(dyn)}$ , the signal can be written to first order :

$$|A_{\lambda_0}|^2 = |A_{\lambda_0}^{bg}|^2 + 2\text{Re}\{A_{\lambda_0}^{bg*} A_{\lambda_0}^{sa}\}, \quad (7.11)$$

where  $\text{Re}$  denotes the real part of a complex number. Only the second term  $S = 2\text{Re}\{A_{\lambda_0}^{bg*} A_{\lambda_0}^{sa}\}$  varies when the tip is scanned at a constant height. Inserting (7.10) into (7.11), we obtain :

$$S = -4\pi\epsilon_0 \frac{\omega}{c} \text{Im} \left\{ Z_{\lambda} A_{\lambda_0}^{bg*} \int_V [\epsilon(\mathbf{r}) - 1] \mathbf{E}(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} \right\} . \quad (7.12)$$

Note that the signal  $S$  has the structure of an interference term between the field  $\mathbf{E}_{tip}^{\lambda_0}$  and the field  $\mathbf{E}$  in the sample. The major consequence of this remark is that both the *phase* and the *amplitude* of the near field are recorded, as was pointed out earlier [97, 100] (see also §5), and it was recognized that the output is essentially a hologram of the sample. Yet, note that this conclusion was derived from the heuristic assumption that the signal is proportional to the *near-field intensity*, an assumption that turns out to coincide with the exact result (7.12) only when the vectorial nature of the field disappears. In fact, (7.12) accounts for the polarization effect in the detection process. The signal produced by the mode  $\lambda_0^{(-)}$  propagating towards the detector is related to the projection of the near field on the direction of the field  $\mathbf{E}_{tip}^{\lambda_0}$ , emitted through the tip by the mode  $\lambda_0^{(+)}$ , in the absence of the sample. It is clear that  $\mathbf{E}_{tip}^{\lambda_0}$  is the fundamental quantity to measure in order to interpret the image. Note that several methods have been proposed to characterize the field emitted by a SNOM probe [159, 38].

Finally, (7.12) shows explicitly the dependence of the signal on the sample structure (dielectric constant and topography). The result could be easily extended to a magnetic material, as mentioned previously. This expression is exact in the sense that all multiple interactions are taken into account. Indeed, the field  $\mathbf{E}(\mathbf{r}, \mathbf{r}_{tip})$  in (7.12) is the self-consistent field inside the sample. In particular, it depends on the tip position.

### (v) Existence of an impulse response

Equation (7.12) provides a linear relationship between the signal and the dielectric contrast in the sample. Nevertheless, the dependence of the field  $\mathbf{E}$  on the tip position breaks the translational invariance, so that it is *a priori* not possible to introduce an impulse response (or a transfer function). Nevertheless, we *demonstrate* that a single scattering approximation, which is accurate in NFO for a broad class of samples [112, 115, 153, 160, 151], allows to introduce naturally an impulse response describing the imaging system. This result gives a rigorous theoretical basis to the concepts of impulse response and transfer functions introduced

previously in NFO, either heuristically or numerically [148, 136, 98, 149, 150, 152, 153, 95].

Assuming that only single scattering takes place inside the sample, the first Born approximation can be used in (7.12). It amounts to approximating the field  $\mathbf{E}(\mathbf{r}, \mathbf{r}_{tip})$  by the field  $\mathbf{E}^{(0)}$  illuminating the sample, i.e., the field radiated by the sources  $\mathbf{j}_s$  in the reference system (tip + substrate), in the absence of the sample. Due to the extended illumination used in collection-mode SNOM (the sources are in the far field, and the incident field is uniform across the scanning area), the illuminating field has the form  $\mathbf{E}^{(0)}(\mathbf{r} - \mathbf{r}_{tip})$ . Note that  $\mathbf{E}^{(0)}$  is in general non uniform, because of the interaction of the tip with the *substrate*. For example, a local enhancement may be observed. In other words, the tip may interact strongly with the *substrate*, but we assume a weak interaction with the *sample*. An example of such a situation would be a dielectric aggregate deposited on a metallic substrate. The remarkable consequence of this approximation is that an impulse response  $H(\mathbf{r} - \mathbf{r}_{tip})$  exists between the dielectric contrast and the signal :

$$S = \int_V H(\mathbf{r} - \mathbf{r}_{tip}) [\epsilon(\mathbf{r}) - 1] d\mathbf{r} , \quad (7.13)$$

where the impulse response is given by :

$$H(\mathbf{r} - \mathbf{r}_{tip}) = -4\pi\epsilon_0 \frac{\omega}{c} \text{Im} \left\{ Z_\lambda A_{\lambda_0}^{bgr} \mathbf{E}^{(0)}(\mathbf{r} - \mathbf{r}_{tip}) \cdot \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) \right\} . \quad (7.14)$$

Let us use this result to draw some conclusions on the imaging capabilities of the technique :

1. Obviously, the contribution to the signal comes from regions where the field  $\mathbf{E}_{tip}^{\lambda_0}$  takes significant values. In this respect, the field  $\mathbf{E}_{tip}^{\lambda_0}$  is a weighting factor. Yet, this condition is not sufficient, since the illuminating field  $\mathbf{E}^{(0)}$  may be perpendicular to  $\mathbf{E}_{tip}^{\lambda_0}$ , so that the contribution would be zero.
2. The nice feature that appears is that one may select, as a weighting factor, any cartesian component of  $\mathbf{E}_{tip}^{\lambda_0}$  by illuminating with a linearly polarized field  $\mathbf{E}^{(0)}$ . This can be used to enhance the resolution.
3. An important consequence of the existence of this impulse response is that the inverse problem can be solved, in principle, by a *deconvolution* [98, 99, 145].
4. Finally, (7.14) shows that *multiple scattering (i.e., strong interactions) between the tip and the sample is not necessary to obtain good resolution*. The achievable resolution is given

by the width of the impulse response. A sufficient condition, to achieve high resolution, is to have a high confinement of at least one component of  $\mathbf{E}_{tip}^{\lambda_0}$ . Moreover, the illuminating field  $\mathbf{E}^{(0)}$  may also be confined, due to the interaction of the tip and the flat substrate.

Several experiments have been done by both illuminating and detecting with the same tip, which is known as the *reflection mode* [161, 162, 163]. In that case, the fields  $\mathbf{E}_{tip}^{\lambda_0}$  and  $\mathbf{E}^{(0)}$  coincide and the overlapping integral is maximal. An interesting property is that this configuration yields an improvement of the resolution very similar to what happens in confocal microscopy. Let us suppose, for the sake of illustration, that the field  $\mathbf{E}_{tip}^{\lambda_0}$  is Gaussian with a width  $a$ . Then the product  $\mathbf{E}_{tip}^{\lambda_0} \cdot \mathbf{E}^{(0)}$  is a Gaussian with width  $a$  for a plane-wave illumination and a Gaussian with width  $a/\sqrt{2}$  for an illumination by  $\mathbf{E}_{tip}^{\lambda_0}$ . Note that, in this particular case, the transfer function does not depend on the polarization. A slightly different set-up has been used by Bozhevolnyi et al.[163]. They detected the cross-polarized light, which amounts to using, as an impulse response, the product  $\mathbf{E}_{tip}^{\lambda_0} \cdot \mathbf{E}'_{tip}$ , where  $\mathbf{E}'_{tip}$  is the field emitted by the tip, when the polarization has changed by  $90^\circ$ . It is clear that the depolarizing properties of the tip are involved in this particular case.

Equation (7.14) is the explicit expression of the impulse response relating the signal to the dielectric contrast in the sample. Nevertheless, some attempts to define an impulse response (or a transfer function) between the signal and the *near-field intensity* illuminating the tip were reported in the literature [136, 151, 152, 95]. In the present framework, it is possible to derive an explicit expression of this transfer function. The derivation is reported in Appendix F. In particular, it shows that 1) such a transfer function exists under the same conditions as the impulse response defined in (7.14), and 2) the transfer function does not, in general, relates the signal to the square modulus of the electric field, due to polarization effects.

## B. Illumination mode

The illumination-mode SNOM closely follows the original idea of Synge [51]. The nanosource, achieved for instance by illuminating through a metal-coated optical fiber tip [33], is scanned along the surface and the scattered field is recorded by a detector in the far field, either in transmission or reflection. The link between the signal and the properties of the sample can be established along similar lines : we use the reciprocity theorem to derive an exact relationship

between the induced current in the sample and the field at the detector position. Then, we show that a single scattering approximation leads to the introduction of an impulse response to describe the imaging properties of the microscope.

### (i) Basic equation

To proceed, we consider two different situations, as depicted in Figs. 14 (a) and (b). The reference geometry involves a flat substrate and the probe at subwavelength distance. We use the same notations as in the previous paragraph.

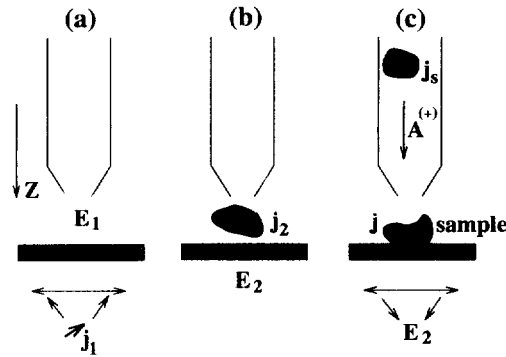


Fig. 14: Sketch of the system. (a) The current  $\mathbf{j}_1$  (assumed punctual) produces a field  $\mathbf{E}_1$  in the reference system, (b) A current distribution  $\mathbf{j}_2$  produces a field  $\mathbf{E}_2$  in the same system, (c) A general situation where  $\mathbf{j}_2$  is splitted into a source with an externally driven current density  $\mathbf{j}_s$  and a sample with an induced current density  $\mathbf{j}$ .

In the first configuration [Fig. 14(a)], a current density  $\mathbf{j}_1$ , placed in the far field (i.e. at the location of the detector), produces a field  $\mathbf{E}_1$ . This field is a function of  $\mathbf{r} - \mathbf{r}_{tip}$  because the source current density is placed in the far field, so that the *incident* field on the system is assumed to be uniform. Thus, the system is translationally invariant. In practice, this condition is fulfilled on the scanning area. For simplicity, we assume that the detector in the experiment is a *point detector* placed in the far field (i.e., a small detector in the focal plane of a lens), and we can use a current density  $\mathbf{j}_1(\mathbf{r}) = \mathbf{j}_1 \delta(\mathbf{r} - \mathbf{r}_1)$ . Note that we could also introduce the dipole moment  $\mathbf{p}_1$ , such that  $\mathbf{j}_1 = -i\omega \mathbf{p}_1$ . But in order to emphasize the similarity with the



study of the collection-mode set-up, it is preferable to work with the current  $\mathbf{j}_1$ . In the second configuration [Fig. 14(b)], a current distribution  $\mathbf{j}_2(\mathbf{r}, \mathbf{r}_{tip})$ , placed for instance between the tip and the substrate, produces a field  $\mathbf{E}_2(\mathbf{r}, \mathbf{r}_{tip})$ . The reciprocity theorem yields (see Appendix D) :

$$\mathbf{j}_1 \cdot \mathbf{E}_2(\mathbf{r}_1, \mathbf{r}_{tip}) = \int_{V_2} \mathbf{j}_2(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} . \quad (7.15)$$

We use the notation  $\mathbf{E}_1^{\mathbf{j}_1}$  for the field  $\mathbf{E}_1$  to keep in mind that this field depends on the choice of the current  $\mathbf{j}_1$  (especially on its orientation which controls the polarizaxtion of  $\mathbf{E}_1^{\mathbf{j}_1}$ ). This notation is similar to that used for the collection-mode set-up,  $\mathbf{E}_{tip}^{\lambda_0}$ , where the mode  $\lambda_0$  was chosen. Splitting the current density  $\mathbf{j}_2$  into the contribution of the sources  $\mathbf{j}_s$  located inside the guide and far from the tip (illumination mode) and the contribution of the sample  $\mathbf{j} = -i\omega\epsilon_0[\epsilon(\mathbf{r}) - 1] \mathbf{E}(\mathbf{r}, \mathbf{r}_{tip})$  [see Fig. 14(c)], one obtains :

$$A_{\mathbf{j}_1}(\mathbf{r}_{tip}) = A_{\mathbf{j}_1}^{bg} + A_{\mathbf{j}_1}^{sa}(\mathbf{r}_{tip}) , \quad (7.16)$$

where  $A_{\mathbf{j}_1} = \mathbf{p}_1 \cdot \mathbf{E}_2(\mathbf{r}_1, \mathbf{r}_1)$  is proportional to the component of the *field at the detector position* along the direction of  $\mathbf{j}_1$ . It contains a *background* contribution  $A_{\mathbf{j}_1}^{bg}$ , independent on the sample, and which does not vary with  $\mathbf{r}_{tip}$  when the scan is performed at a constant height, and a *dynamic* contribution  $A_{\mathbf{j}_1}^{sa}$ , which contains the information on the sample :

$$A_{\mathbf{j}_1}^{bg} = \int_{V_s} \mathbf{j}_s(\mathbf{r}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} \quad (7.17)$$

$$A_{\mathbf{j}_1}^{sa}(\mathbf{r}_{tip}) = -i\omega\epsilon_0 \int_V [\epsilon(\mathbf{r}) - 1] \mathbf{E}(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} . \quad (7.18)$$

Equations (7.17) and (7.18) give the field at the detector position, produced in the illumination-mode configuration in Fig. 14(c). The second term (7.18) gives the optical information on the sample, and depends on both its dielectric constant and topography. Note that (7.17) and (7.18) do not give the expression of the field  $\mathbf{E}_2$ , but of its *projection on the direction of  $\mathbf{p}_1$* . This describes the *polarization effect* in the imaging process in illumination-mode SNOM. For example, to determine the  $x$ -component of  $\mathbf{E}_2$ , one needs to know the field  $\mathbf{E}_1^{\mathbf{j}_1}$  produced when the reference system (probe + substrate) is illuminated by a point current  $\mathbf{j}_1$  of unit amplitude, polarized along  $Ox$ , and located at the detector position.

**(ii) What is measured ?**

The dynamic part  $S$  of the signal produced by the component of  $\mathbf{E}_2$  along the direction of  $\mathbf{j}_1$  is  $S = 2\text{Re}\{A_{\mathbf{j}_1}^{bg*} A_{\mathbf{j}_1}^{sa}\}$ , assuming that  $|A_{\mathbf{j}_1}^{sa}| \ll |A_{\mathbf{j}_1}^{bg}|$ . Using (7.17) and (7.18), we obtain :

$$S = 2\omega\epsilon_0 \text{Im} \left\{ A_{\mathbf{j}_1}^{bg*} \int_V [\epsilon(\mathbf{r}) - 1] \mathbf{E}(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} \right\} . \quad (7.19)$$

Equation (7.19) gives an exact expression of the signal in illumination-mode SNOM, when a point detector is used, and when only one component of the field is measured (polarized detection). Let us stress that all multiple interactions between the sample and the reference system are taken into account in this expression. In particular,  $\mathbf{E}(\mathbf{r}, \mathbf{r}_{tip})$  is the field inside the sample, which depends on the tip position  $\mathbf{r}_{tip}$ .

**(iii) Impulse response**

As for the collection-mode configuration, an impulse response appears naturally when the first Born approximation is used to describe the field in the sample. In this particular case, this approximation amounts to replacing  $\mathbf{E}(\mathbf{r}, \mathbf{r}_{tip})$  in (7.19) by the field produced by the tip in the reference system. Assuming for simplicity that only one mode, denoted by  $\lambda_0$ , is excited, this field is  $\mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip})$  [see Fig. 13(a)]. One obtains :

$$S = \int_V H(\mathbf{r} - \mathbf{r}_{tip}) [\epsilon(\mathbf{r}) - 1] d\mathbf{r} , \quad (7.20)$$

where

$$H(\mathbf{r} - \mathbf{r}_{tip}) = 2\omega\epsilon_0 \text{Im} \left\{ A_{\mathbf{j}_1}^{bg*} \mathbf{E}_{tip}^{\lambda_0}(\mathbf{r} - \mathbf{r}_{tip}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} \right\} . \quad (7.21)$$

**(iv) Similarity with the impulse response in collection-mode SNOM**

Except for a constant factor, (7.21) is identical to (7.14), obtained for collection-mode SNOM. In fact,  $A_{\mathbf{j}_1}^{bg}$ , and  $\mathbf{E}_1^{\mathbf{j}_1}$  in (7.21) are exactly  $A_{\lambda_0}^{bg}$  and  $\mathbf{E}^{(0)}$  in (7.14). This result is not surprising because in §6, we have demonstrated that the collection-mode and illumination-mode instruments are equivalent. Since the two modes are equivalent, their impulse response (when they exist) are identical, as expected. Note that some precise equivalence rules (involving the coherence and polarization properties of the illumination/detection scheme) are given in §6, in order to obtain the same signals in both modes. As a consequence of this equivalence, all the conclusions

we obtained from the structure of the impulse response in collection-mode SNOM (localization of the field emitted by the probe and resolution, polarization effect, inverse problem and role of multiple scattering) are valid in illumination-mode SNOM.

### C. Apertureless set-up

A number of groups have built near-field optical microscopes by using the apertureless technique, where detection is performed in the far field. In order to send light to the detector, a tip is brought into the vicinity of the sample. It is mechanically excited and vibrates in the  $z$ -direction at a well-controlled frequency. The amplitude of the scattered light is related to the amplitude of the near-field that locally illuminates the tip. Differences in the set-ups are due to the nature of the tip (AFM or STM tips [76, 110, 74, 77, 90], levitating particles[73, 164]), operating mode (constant-height or constant-distance mode), illumination, coherence of the illumination, polarization control, etc.

Since the first papers were published, the understanding of the mechanism has considerably changed. Basically, the tip may play two roles. On one hand, it acts as a nanodetector, as depicted in Fig. 15(a). On the other hand, it acts as a nanosource, as shown in Fig. 15(b). Further interactions (multiple scattering) between the tip and the substrate may appear, if there is a strong coupling (e.g., a metallic particle above a metallic surface supporting surface plasmons polaritons). It has been argued very often that such multiple scattering is necessary to obtain good resolution [90]. Indeed, excellent resolutions have been demonstrated for metallic surfaces[71, 72]. Yet, as for the collection and illumination modes, we will see that strong interaction is not necessary. The present goal is to analyse the imaging capabilities of this type of set-up. We have shown that, for collection and illumination-mode SNOM, the signal can be related to the dielectric contrast of the sample. In this section, we outline a similar analysis, which is also based on the reciprocity theorem. We first establish an exact expression for the signal, and relate it to the sample properties. Then, we show that an impulse response can be introduced, when multiple scattering between the sample and the reference system (tip + substrate) can be neglected.

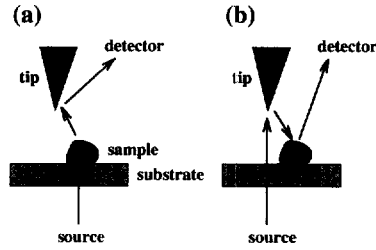


Fig. 15: Sketch of the apertureless set-up. (a) Nanodetector mechanism. (b) Nanosource mechanism.

### (i) Exact expression for the signal

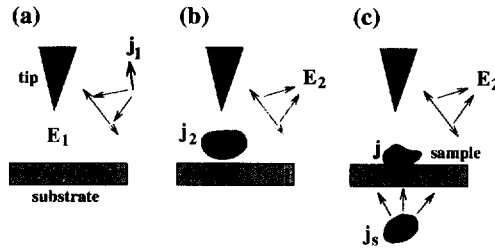


Fig. 16: Sketch of the system. (a) A current  $\mathbf{j}_1$  at the location of the detector produces a field  $\mathbf{E}_1$  in the presence of the tip and the flat substrate. (b) A current density  $\mathbf{j}_2$  produces a field  $\mathbf{E}_2$ . (c) A general experimental situation with a source current density  $\mathbf{j}_s$  and a sample.

Consider a tip placed at subwavelength distance from a flat substrate, at a position  $\mathbf{r}_{tip}$ . The tip may be either an AFM tip [76] or an STM tip [72, 77, 74, 110]. In a first situation [Fig. 16 (a)], a point dipole  $\mathbf{p}_1$  is placed in the far field, at the location of the detector (assumed to be point-like for simplicity). As for the illumination-mode SNOM, we work with the point current density  $\mathbf{j}_1 = -i\omega\mathbf{p}_1$ , rather than with the dipole moment  $\mathbf{p}_1$  itself. We denote by  $\mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip})$  the field produced in this situation. The field  $\mathbf{E}_1^{\mathbf{j}_1}$  can be highly confined, because of the tip effect. This type of tip effect can be described in the electrostatic approximation, and some analytical expressions of the field behavior close to the tip can be found, e.g., in ref. [165]. In a second situation [Fig. 16 (b)], a current density  $\mathbf{j}_2$ , placed in the near field, produces a field  $\mathbf{E}_2(\mathbf{r}, \mathbf{r}_{tip})$ . Using the reciprocity theorem, and splitting the current distribution  $\mathbf{j}_2$  into a source

distribution  $\mathbf{j}_s$  and the current density in the sample, we obtain an expression for the signal which is formally the same as (7.19), derived for the illumination-mode SNOM :

$$S = 2\omega\epsilon_0 \operatorname{Im} \left\{ A_{\mathbf{j}_1}^{bg*} \int_V [\epsilon(\mathbf{r}) - 1] \mathbf{E}(\mathbf{r}, \mathbf{r}_{tip}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) d\mathbf{r} \right\} . \quad (7.22)$$

The only differences between (7.22) and (7.19) are (1) in apertureless SNOM, the source distribution does not excite a mode in an optical fiber, used to produce a localized illumination of the sample, but the source directly illuminates the sample from the far field, as in conventional microscopy, and (2) the fields  $\mathbf{E}_1^{\mathbf{j}_1}$  are very different in both situations. Let us recall that in (7.22),  $S$  is the dynamic part of the signal produced at the detector position by the component of the electric field polarized along the direction of  $\mathbf{j}_1$ .

## (ii) Impulse response

As for the illumination-mode configuration, an impulse response appears naturally when the first Born approximation is used to describe the field in the sample. Within this approximation, we replace the field  $\mathbf{E}(\mathbf{r}, \mathbf{r}_{tip})$  in (7.22) by the field  $\mathbf{E}_{tip}^{(0)}(\mathbf{r} - \mathbf{r}_{tip})$  illuminating the sample. This field is the field radiated by the sources  $\mathbf{j}_s$  in the reference system (tip + flat substrate). Note that, due to the near field interaction between the subwavelength tip and the *substrate* and the tip effect, this field may exhibit subwavelength lateral variations, as well as a local enhancement [166]. We finally obtain :

$$S = \int_V H(\mathbf{r}, \mathbf{r}_{tip}) [\epsilon(\mathbf{r}) - 1] d\mathbf{r} , \quad (7.23)$$

where

$$H(\mathbf{r}, \mathbf{r}_{tip}) = 2\omega\epsilon_0 \operatorname{Im} \left\{ A_{\mathbf{j}_1}^{bg*} \mathbf{E}_{tip}^{(0)}(\mathbf{r} - \mathbf{r}_{tip}) \cdot \mathbf{E}_1^{\mathbf{j}_1}(\mathbf{r} - \mathbf{r}_{tip}) \right\} . \quad (7.24)$$

Let us emphasize that (7.24) accounts for the two roles of the tip, namely, nanosource and nanodetector (see Fig. 15). Note that in Fig. 15, only single scattering takes place inside the sample, thus the existence of an impulse response is not in contradiction with the presence of the two mechanisms. We see that the resolution depends on both the confinement of the illuminating field  $\mathbf{E}_{tip}^{(0)}$  and of the field  $\mathbf{E}_1^{\mathbf{j}_1}$ . As for the other modes, *multiple scattering in the sample is not necessary to achieve high resolution*. Note that a confocal type set-up can be used to run such an experiment, where the illuminating field  $\mathbf{E}^{(0)}$  coincides with the field  $\mathbf{E}_1$ , so that

the impulse response is even more localized, as discussed for the collection-mode set-up. In this particular arrangement, *the enhancement, due to the singularity close to the tip, is squared*. Such an experiment has been reported in the literature [110, 74].

Finally, we note that (7.24) accounts for the polarization properties of the imaging process in apertureless SNOM. The signal  $S$  is produced by a measurement of the field at the detector position, *projected on a direction*  $\mathbf{u}_1$ . This signal is proportional to the projection of the illuminating field  $\mathbf{E}^{(0)}$  on the field  $\mathbf{E}_1$ , produced in the reference system by a dipole  $\mathbf{p}_1$  oriented along  $\mathbf{u}_1$ , and placed at a detector position. The importance of polarization effects in apertureless SNOM was pointed out by several authors [84, 83]. The role of the polarization of the incident field and of the detector position was particularly emphasized. We hope that this model will be helpful in understanding the basic mechanisms leading to these observations.

### (iii) Nanosource or nanodetector ?

Recently, numerical calculations have been performed to investigate the importance of multiple scattering between the tip and the sample, and also to compare the relative importance of the two mechanisms : nanosource and nanodetector (see Fig. 15) [83]. The system studied was a small gold cylinder levitating above a one-dimensional dielectric rough surface, illuminated by a  $p$ -polarized plane wave (see Fig. 17).

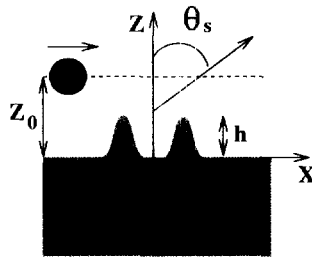


Fig. 17: Sketch of the system, where a small cylinder is placed above the surface.

The detector detects the flux of the Poynting vector in the far field for different observation angles. In a numerical simulation, it is possible to evaluate separately the field radiated by the tip (nanodetector mechanism) and the field radiated by the currents induced on the sample

(nanosource mechanism). If the field radiated by the sample shows high-resolution structures, it means that the cylinder has locally illuminated the sample, acting as a nanosource. The results displayed in Fig. 18(a) show that for a tip sample distance of  $0.017\lambda$ , the intensity scattered by the surface does not display any resolution. Thus, the small cylinder is not a nanosource. By contrast, for the same tip-sample distance, the signal scattered by the tip reproduces the surface profile (see Fig. 18), as shown in ref.[83]. Thus, in this case, the tip mainly acts as a nanodetector.

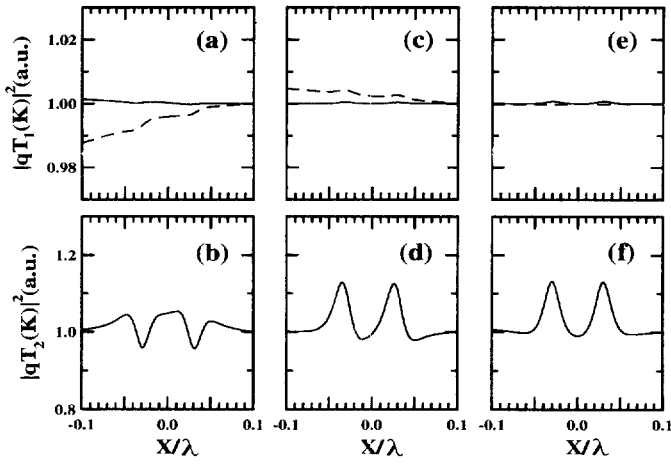


Fig. 18: Intensity scattered in the far field by the sample (nanosource mechanism) (a), (c), (e) and by the tip (nanodetector mechanism) (b), (d), (f) for different scattering angles when the tip is scanned at constant height  $z_0 = 0.017\lambda$  along the surface. The index of the substrate is 1.5. The dielectric constant of the tip is  $-9.89 + i1.05$ , its radius is  $\lambda/100$ . The topographical bumps are Gaussian, with height  $\lambda/200$  and width  $\lambda/100$ . The incident field is  $p$ -polarized and  $\theta_0 = 50^\circ$ . Left column,  $\theta_s = 0^\circ$ , central column  $\theta_s = 50^\circ$ , right column  $\theta_s = 90^\circ$ . From A. Madrazo et al., ref.[83]

The results in ref.[83] shows that, even with a metallic tip, the nanosource mechanism appears only for very small tip-sample distance, with a highly reflective substrate. Nevertheless, the calculations were carried out with a cylinder (i.e., the 2D equivalent of a sphere). With a sharp tip, a stronger local enhancement of the field could be obtained, due to the singularity of the field at the tip [165]. In this case, the nanosource mechanism may become more important.

## D. Conclusion

We have shown that the image formation in the different families of near-field optical microscopes can be described in the same framework. In each case, the signal can be related to the polarization density in the sample. The explicit form of the relationship has the structure of an overlapping integral of the dot product of the polarization density with a reference field  $\mathbf{E}_1$ . *This reference field is always the field produced in the reference system (tip + substrate, without sample), when the detector of the experiment is replaced by a source.* We emphasize that, in the derivation of this result, the only assumption is that the reciprocity theorem holds (see Appendix D). We also emphasize that the reference field  $\mathbf{E}_1$  depends on the polarization of the detection. By changing the polarization, the reference field is changed. The field  $\mathbf{E}_1$  plays a central role in the analysis. It is therefore important to be able to measure it experimentally. A possible solution is to have a point dipole acting as a local sensor of the field. The experiments described by Betzig *et al.* [38] are essentially this type of measurement. Another method was proposed recently to measure the near field emitted by a probe [159]. Note that producing an interference with a linearly polarized plane wave might allow to record separately the different components of the field  $\mathbf{E}_1$ .

At the cost of an approximation, the concept of impulse response can be introduced naturally. The only assumption is that multiple scattering between the sample and the rest of the system is negligible. For instance, consider a molecule placed between a coated optical fiber and a metallic substrate. There is obviously a strong interaction between the tip and the substrate, but multiple scattering between the *molecule* and the rest of the system can often be neglected, because the scattering cross-section is small. This condition is generally satisfied if the samples are small compared to the wavelength and if their dielectric contrast is low. The general form of the impulse response is then the product of the illuminating field  $\mathbf{E}^0$  (i.e., the



field produced in the reference system by the real sources of the experiment set-up) with the reference field  $\mathbf{E}_1$  :

$$H(\mathbf{r} - \mathbf{r}_{tip}) = K \operatorname{Im}\{K' \mathbf{E}_1 \cdot \mathbf{E}^0(\mathbf{r} - \mathbf{r}_{tip})\}, \quad (7.25)$$

where  $K$  and  $K'$  are constants that depend on the set-up. This formalism should be useful in designing and evaluating the performances of near-field experiments. A number of features can be studied with this approach, such as the role of polarization, illumination or detection scheme, etc. Moreover, the concept of impulse response provides a quantitative tool to compare the potential capabilities of different SNOM set-ups.

## 8. Influence of Coherence

Coherence theory allows to discuss the statistical properties of random fields [167]. For our purpose, we distinguish two main classes of random fields. On one hand, thermally produced fields are random in nature. Indeed, they are produced by the uncorrelated radiative desexcitation of a large number of excited states. On the other hand, a fully coherent beam can acquire a random nature after being scattered by a random system (e.g., a particle suspension or a rough surface). The coherence of a field is measured by the degree of correlation of the field with itself at different points and/or different times. We can distinguish between spatial coherence and temporal coherence. Spatial coherence is fully characterized by the correlation function  $\langle E_\alpha(\mathbf{r}, t) E_\beta^*(\mathbf{r}', t) \rangle$  whereas temporal coherence is described by the correlation function at the same point and different times  $\langle E_\alpha(\mathbf{r}, t) E_\beta^*(\mathbf{r}, t') \rangle$ , where the brackets denote an ensemble average. In what follows, we are mostly concerned by the spatial coherence of the light. In the first subsection, we study the structure of the near field above a slightly rough surface, illuminated by a plane wave. We show that the near-field intensity acquires a *wavy structure* that has been observed experimentally [168, 169, 146, 112], and which is in fact closely related to the statistical properties of the surface itself. In the second subsection, we analyse the role of spatial coherence in the image formation process, when using a microscope operating in the collection mode. We show that the best illumination conditions are obtained when using a partially *spatially incoherent symmetric beam*, as has been first observed experimentally by Chabrier *et al.* [144]. This discussion is intimately connected to the discussion in §6.D.

# Theory of electromagnetic field imaging and spectroscopy in scanning near-field optical microscopy

J. A. Porto, R. Carminati,<sup>a)</sup> and J.-J. Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

(Received 28 February 2000; accepted for publication 22 July 2000)

We derive a general expression for the signal in scanning near-field optical microscopy. This expression, based on the reciprocity theorem of electromagnetism, is an overlapping integral between the local electric field and a function that characterizes the tip. In particular, it includes the dependence on wavelength, illumination conditions, and polarization. To illustrate the potentiality of this theory, we discuss the polarization behavior and the spectral response of the apertureless setup. © 2000 American Institute of Physics. [S0021-8979(00)02821-8]

## I. INTRODUCTION

Scanning near-field optical microscopy (SNOM) has attracted considerable attention in the past fifteen years as a technique to obtain optical images of surfaces with subwavelength resolution.<sup>1,2</sup> In addition to surface-structure imaging, SNOM has proven its ability to generate and image confined electromagnetic fields such as surface plasmons,<sup>3</sup> to detect single molecule fluorescence,<sup>4</sup> to perform near-field spectroscopy,<sup>5,6</sup> or to observe light localization on disordered surfaces.<sup>7</sup> In addition, SNOM is a good tool to control light propagation in guiding microstructures<sup>8</sup> and optoelectronic components.<sup>9</sup> In these applications, where the main interest is in the detection of the electromagnetic field itself rather than in imaging a surface structure, SNOM appears as a privileged technique compared to other scanning probe microscopies.

Several theoretical studies about SNOM have been presented in the last ten years, based on numerical simulations<sup>10–12</sup> or analytical models.<sup>13,14</sup> Concerning imaging of confined electromagnetic fields, a first description is obtained by assuming that the signal is proportional to the square modulus of the electric field at the tip location.<sup>12,14</sup> Another point of view is to describe the tip by a pointlike scatterer that scatters the near field towards a far-field detector.<sup>10,13</sup> Although these approaches are well suited for some particular cases, they do not tackle important aspects, such as the tip shape effects and the nonlocality of the detection process. Hence, how the signal depends on the local electromagnetic field and on the tip in a real situation remains an open issue. Another important point, which remains unstudied, is the influence of the near-field detection and the tip properties on the spectral response of the SNOM setup. This issue is essential in order to understand the spectroscopic experiments. There is absolutely no reason to assume that a SNOM setup has a flat spectral response. Thus, normalizing a near-field spectrum by a far-field spectrum does not suppress all the instrument spectral properties.

In this article we introduce an exact and general expression for the signal as a function of both the local electromagnetic field and the tip properties. This expression (i) reveals which physical quantity is detected in a SNOM experiment and (ii) provides a useful tool to analyze experimental results and to calculate the SNOM signal. To illustrate the potentiality of this approach, we concentrate on the apertureless setup.<sup>15</sup> We describe the polarization behavior and the spectral response, in agreement with recent experimental results.<sup>6,16</sup> In view of these results, the approach looks particularly suitable to the description of near-field optical spectroscopy.

## II. THEORY

### A. General expressions of the SNOM signal

In order to obtain an expression for the signal, we use the reciprocity theorem of electromagnetism.<sup>17</sup> This theorem was the basis of a SNOM model derived previously for imaging of surface structures,<sup>14</sup> and of a model for light emission in scanning tunneling microscopy.<sup>18</sup> The reciprocity theorem involves two different situations.

The first one, called experimental situation, is a generic experimental SNOM setup, as illustrated in Fig. 1(a). In this experimental situation, a physical system described by a monochromatic current density  $\mathbf{j}_{\text{exp}}(\omega)$  radiates the field to be observed. This physical system is either a primary source (e.g., a molecule or an emitting optoelectronic device) or a secondary source excited by a primary point source with current density  $\mathbf{j}_{\text{sou}}(\omega)$  placed in the far field (e.g., a plasmon resonance on a metal surface or a guided mode in a microstructure excited by an incident laser beam). The field radiated by the physical system is probed by a local tip and the signal is recorded in the far field by a point detector. The region between the tip and the physical system is assumed to be homogeneous and free of sources (gap region).

In the second situation, called reciprocal situation and represented in Fig. 1(b), the collecting system (tip+detection optics) is illuminated by a hypothetical monochromatic point source  $\mathbf{j}_{\text{rec}}(\omega)$  placed at the detector position, in absence of the physical system and the primary source  $\mathbf{j}_{\text{sou}}(\omega)$ . The di-

<sup>a)</sup>Author to whom correspondence should be addressed; electronic mail: remi@em2c.ecp.fr

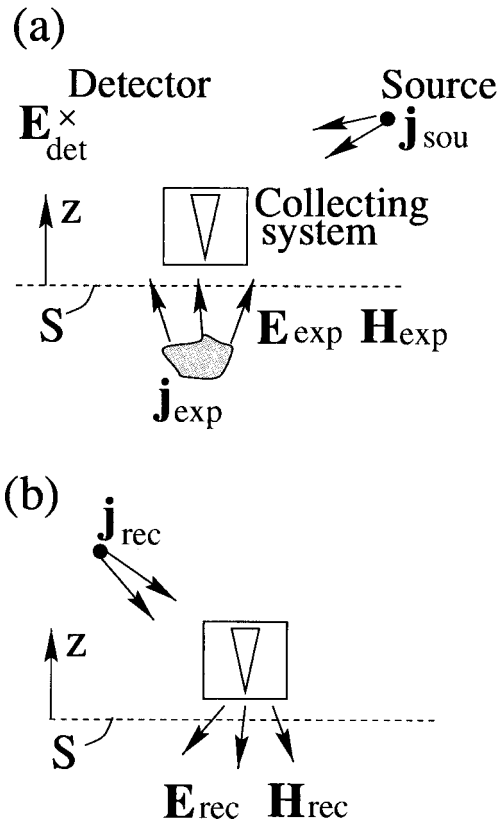


FIG. 1. Schematic views of the two situations considered for the application of the reciprocity theorem. (a): sketch of the experimental situation, where a generic experimental SNOM setup is considered. (b): reciprocal situation, where the collecting system is illuminated by a point source located at the detector position.

rection of  $\mathbf{j}_{\text{rec}}$  corresponds to that of an analyzer placed before the detector in the experimental situation. The application of the reciprocity theorem<sup>17</sup> leads to the following expression for the electric field  $\mathbf{E}_{\text{det}}$  at the detector position:

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} + \int_S (\mathbf{E}_{\text{exp}} \times \mathbf{H}_{\text{rec}} - \mathbf{E}_{\text{rec}} \times \mathbf{H}_{\text{exp}}) \cdot d\mathbf{S}, \quad (1)$$

where  $\mathbf{E}_{\text{exp(rec)}}$  and  $\mathbf{H}_{\text{exp(rec)}}$  are, respectively, the electric and magnetic fields in the experimental (reciprocal) situation,  $d\mathbf{S} = dS \hat{\mathbf{z}}$ , where  $\hat{\mathbf{z}}$  is the unit vector of the  $z$  axis, and  $\mathbf{r}_{\text{sou}}$  is the position of the source. In the gap region, the electromagnetic fields in both situations can be expressed in terms of an angular spectrum of plane waves.<sup>19</sup> For instance, at a point  $\mathbf{r} = (\mathbf{R}, z)$  in the gap region,  $\mathbf{E}_{\text{exp}}$  can be written in the form

$$\mathbf{E}_{\text{exp}}(\mathbf{R}, z) = \int \mathbf{e}_{\text{exp}}^+(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d\mathbf{K} + \int \mathbf{e}_{\text{exp}}^-(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d\mathbf{K}, \quad (2)$$

where  $\gamma(\mathbf{K}) = (k^2 - \mathbf{K}^2)^{1/2}$ , with  $k = \omega/c$  ( $\omega$  being the frequency and  $c$  being the speed of light in vacuum) and the determination  $\text{Re } \gamma > 0$  and  $\text{Im } \gamma > 0$ . The integrals in Eq. (2) are extended to  $0 < |\mathbf{K}| < +\infty$ , so that they contain both propagating ( $|\mathbf{K}| \leq k$ ) and evanescent ( $|\mathbf{K}| > k$ ) waves. Waves which propagate or decay towards  $z > 0$  have an am-

plitude denoted by  $\mathbf{e}_{\text{exp}}^+(\mathbf{K})$ , while  $\mathbf{e}_{\text{exp}}^-(\mathbf{K})$  corresponds to waves which propagate or decay towards  $z < 0$ . A similar angular spectrum can be written for the magnetic field

$$\mathbf{H}_{\text{exp}}(\mathbf{R}, z) = \int \mathbf{h}_{\text{exp}}^+(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d\mathbf{K} + \int \mathbf{h}_{\text{exp}}^-(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d\mathbf{K}, \quad (3)$$

with, as a consequence of Maxwell's equations

$$\mathbf{k}^\pm \times \mathbf{e}_{\text{exp}}^\pm(\mathbf{K}) = \omega \mu_0 \mathbf{h}_{\text{exp}}^\pm(\mathbf{K}), \quad (4)$$

where  $\mathbf{k}^\pm = (\mathbf{K}, \pm \gamma)$  and  $\mu_0$  is the magnetic permeability of vacuum.

Similar expressions can be used for the fields  $\mathbf{E}_{\text{rec}}$  and  $\mathbf{H}_{\text{rec}}$  in the reciprocal situation (see Appendix A). However, the angular spectra of these fields only contain waves which propagate or decay towards  $z < 0$ , because there are no sources below the probe in the reciprocal situation [see Fig. 1(b)]. Inserting the angular spectra of all fields into the integral in Eq. (1) and collecting terms, we obtain the following expression for the field at the detector position (see Appendix A for details):

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} - \frac{8\pi^2}{\omega \mu_0} \int \gamma(\mathbf{K}) \mathbf{e}_{\text{rec}}(-\mathbf{K}) \cdot \mathbf{e}_{\text{exp}}^+(\mathbf{K}) d\mathbf{K}, \quad (5)$$

where the integral is extended over  $0 < |\mathbf{K}| < +\infty$  and describes the detection of both propagating and evanescent components of the experimental field  $\mathbf{e}_{\text{exp}}^+(\mathbf{K})$ . An equivalent expression is obtained by transforming Eq. (5) into real space (see Appendix A)

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} - \frac{2i}{\omega \mu_0} \int_S \frac{\partial \mathbf{E}_{\text{rec}}}{\partial z}(\mathbf{R}, z) \cdot \mathbf{E}_{\text{exp}}^+(\mathbf{R}, z) d\mathbf{R}, \quad (6)$$

where the integral in the second term is now evaluated in a plane at a constant height  $z$  in the gap region (surface  $S$  in Fig. 1).  $\mathbf{E}_{\text{exp}}^+$  denotes the experimental field containing waves that propagate or decay towards  $z > 0$ , whose angular spectrum is given by the first integral in Eq. (2).

## B. Discussion

Both Eqs. (5) and (6) are exact expressions of the component of the electric field along the direction  $\mathbf{j}_{\text{rec}}$  of an analyzer at the detector position in the experimental situation. They rely on the validity of the reciprocity theorem, namely, the collecting system (tip + detection optics) has to be made of linear materials with symmetric constitutive tensors<sup>17</sup>. This restriction applies to the bodies entering the reciprocal situation in Fig. 1(b). In particular, it does not apply to the physical system, which may contain, for example, magnetic materials. We also put forward that both expressions take into account multiple scattering between the physical system and the tip. Indeed,  $\mathbf{E}_{\text{exp}}^+$  is the field illuminating the tip in the presence of the tip.

The first term  $\mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}}$  in the right-hand side of Eqs. (5) and (6) describes direct radiation of the illuminating source (e.g., a laser) towards the detector. It does not carry any information on the physical system. Moreover, it vanishes in experimental situations in which the physical system is itself the primary source (e.g., a fluorescent molecule or an emitting optoelectronic device). Both expressions can be useful to describe a given experiment:

- Expression (5) (Fourier space) describes how each spatial frequency  $\mathbf{K}$  of the experimental field is detected. The coupling factor is proportional to  $\mathbf{e}_{\text{rec}}(-\mathbf{K})$ . Hence, a given frequency  $\mathbf{K}$  is efficiently detected if it is present in the spectrum of  $\mathbf{E}_{\text{rec}}$ , the field produced by the collecting system (especially the tip) when it is illuminated from the detector. In other words, a tip is able to detect high spatial frequencies if it itself creates high spatial frequencies when illuminated from the far field.

- Expression (6) (real space) shows that the detected field is given by an overlapping integral between the experimental field and a term proportional to the derivative of the reciprocal field  $(\partial/\partial z)\mathbf{E}_{\text{rec}}$ . The latter is a response function of the instrument, describing the spatial localization of the detection process, the polarization effects, and the spectral response.

Finally, we note that an expression of the SNOM signal exhibiting the same structure as Eq. (6) was derived recently<sup>20</sup> as a generalization of Bardeen’s formula,<sup>21,22</sup> originally developed for electron tunneling between two weakly coupled electrodes. This result has shown that SNOM and scanning tunneling microscope (STM) can be handled using the same formalism. There are two differences between Eq. (6) and the result in Ref. 20: (i) The latter was derived under the assumption of weak tip–sample coupling whereas Eq. (6) is exact. (ii) The result in Ref. 20 was formally put in the symmetrized form of a current operator as in Bardeen’s original article.

### C. Electric or magnetic field?

Before closing this section, we address the question of whether a SNOM detects preferentially the electric field or the magnetic field. This question was raised recently in light of experiments using a photon scanning tunneling microscope (PSTM) with either dielectric or metal-coated probes.<sup>23</sup> Dielectric probes seem to detect a signal proportional to  $|\mathbf{E}_{\text{exp}}^+|^2$  whereas gold coated fiber tips seem to detect a signal proportional to  $|\mathbf{H}_{\text{exp}}^+|^2$ . We shall see that the preceding analysis allows to discuss this issue.

Starting from Eq. (1) and using Eqs. (2)–(4), it is possible to derive an expression of the electric field at the detector in terms of magnetic fields only. One obtains in this case

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} + \frac{2i}{\omega \epsilon_0} \times \int_S \frac{\partial \mathbf{H}_{\text{rec}}}{\partial z}(\mathbf{R}, z) \cdot \mathbf{H}_{\text{exp}}^+(\mathbf{R}, z) d\mathbf{R}. \quad (7)$$

Equations (6) and (7) are two equivalent expressions of the same quantity. The former connects the signal at the detector to the electric field illuminating the tip  $\mathbf{E}_{\text{exp}}^+$ , whereas the latter relates the signal to the magnetic field illuminating the tip  $\mathbf{H}_{\text{exp}}^+$ . In this case, the response function which characterizes the detection process is proportional to  $(\partial/\partial z)\mathbf{H}_{\text{rec}}$ . The fact that the signal may be expressed as a function of either the electric or the magnetic field is not surprising because both are linked by Maxwell’s equations. Nevertheless, depending on the shape of the response function  $(\partial/\partial z)\mathbf{E}_{\text{rec}}$  and  $(\partial/\partial z)\mathbf{H}_{\text{rec}}$ , the signal may resemble  $\mathbf{E}_{\text{exp}}^+$  or  $\mathbf{H}_{\text{exp}}^+$ . Note that the shape of the response function depends only on the tip and the collecting optics. In the case of the experiment described in Ref. 23, the dielectric tip may produce a response function  $(\partial/\partial z)\mathbf{E}_{\text{rec}}$  which is highly localized and symmetric, so that the signal  $I = |\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}}|^2$  closely follows the distribution of the electric field. Conversely, the gold coated fiber tip may produce a highly localized and symmetric response function  $(\partial/\partial z)\mathbf{H}_{\text{rec}}$ , so that the signal closely follows the distribution of the magnetic field.

## III. APPLICATION: APERTURELESS SNOM

Equation (6), or equivalently Eq. (5), can be used to analyze experimental results in SNOM. The key quantity, which characterizes the tip and the collecting system, is the reciprocal field  $\mathbf{E}_{\text{rec}}$ . Different models are available, that allow an approximate and practical description of this field. For instance, for an aperture SNOM, the Bethe–Bouwkamp model<sup>24</sup> could be used.<sup>13,25</sup> For an apertureless SNOM<sup>15</sup> using a conical metallic tip,  $\mathbf{E}_{\text{rec}}$  can be modeled by the field near the tip apex of a perfectly conducting cone.<sup>26,27</sup> With such models, the approach presented in this article provides a versatile and useful tool to analyze experimental results and to identify the key parameters.

### A. Tip model

In order to illustrate the potentiality of the theory, from now on we focus on the apertureless SNOM. As mentioned above, the field  $\mathbf{E}_{\text{rec}}$  can be modeled in this case by the field near the tip apex of a perfectly conducting cone illuminated by an electric dipole placed at the position of the detector (i.e., in the far field). At short distance  $r$  from the tip apex ( $kr \ll 1$ ), one has<sup>26,27</sup>

$$\mathbf{E} = k(kr)^{\nu-1} \sin \beta \left[ \mathbf{u}_r + \frac{\mathbf{u}_\theta}{\nu} \frac{\partial}{\partial \theta} \right] a(\theta_o, \theta, \alpha), \quad (8)$$

where  $a$  is a function of the angle of incidence  $\theta_o$ , the angle of observation  $\theta$ , and the semiangle of aperture of the cone  $\alpha$ . The other parameters are the wave vector  $k = \omega/c$  and the angle of polarization  $\beta$  of the incident wave ( $\beta = 0$  for an illumination in  $s$  or TE polarization,  $\beta = \pi/2$  for an illumination in  $p$  or TM polarization).  $\mathbf{u}_r$  and  $\mathbf{u}_\theta$  are the unit vectors in spherical coordinates.  $\nu$  is a positive number smaller than 1 which depends on the cone angle.<sup>26</sup> Note that although Eq. (8) is an asymptotic expression, it is not an electrostatic approximation. Therefore, it includes retardation effects. Remarkably, the field given by Eq. (8) is highly



enhanced near the tip apex, and its spatial distribution does not depend on the illumination conditions. This model for the tip was introduced in near-field optics by Cory *et al.*<sup>27</sup> We will show that, together with Eq. (6), it allows to explain the polarization dependence and the spectral response measured on an apertureless setup using a metallic tip.

## B. Polarization dependence

The polarization effect in apertureless SNOM has been recently studied experimentally.<sup>16</sup> In that work, the sample was a flat silicon surface, probed by a tungsten tip, illuminated at a wavelength  $\lambda = 647$  nm. The dependence of the signal on the polarization of the incident wave was measured. The polarization state of the incident wave is described by the angle  $\beta$ . The signal versus  $\beta$ , measured in reflection in the direction normal to the surface, is shown in Fig. 2 in Ref. 16. The result can be explained using Eq. (6), together with the tip model Eq. (8). The signal  $I$  is proportional to  $|\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}}|^2$ . This quantity is described by the integral term in Eq. (6), which involves the enhanced field at the tip apex [the first term in the right-hand side in Eq. (6) gives a negligible contribution in this setup]. When the tip is at a few nanometers from the surface, the field  $\mathbf{E}_{\text{exp}}^+$  illuminating the tip is mainly the enhanced field reflected by the surface. Therefore,  $\mathbf{E}_{\text{exp}}^+$  is proportional to the field given by Eq. (8), and thus to  $\sin \beta$ . Finally, the signal is  $I \propto \sin^2 \beta$ . This prediction is in agreement with the experimental result (see Fig. 2 in Ref. 16). Thus, Eq. (6), together with Eq. (8) correctly describes the polarization behavior of an apertureless SNOM using a metallic tip.

## C. Spectral response

We now turn to the study of the spectral response. It has been found experimentally very recently that the spectral response of an apertureless SNOM using a metallic tip is not flat, and that it depends on the tip shape.<sup>6</sup> This unexpected behavior is of great importance in near-field optical spectroscopy, where the recorded spectra have, in principle, to be corrected by the response function of the instrument. In Ref. 6, a confocal geometry was used. The sample was an aluminum mirror (flat spectral response in the visible, with a reflectivity  $R = 0.9$ ). The signal was measured versus the incident wavelength, and normalized by the far-field spectrum recorded under the same conditions. The results for two tips with different angles of aperture are shown in Fig. 3 in Ref. 6. We shall show that Eq. (6), together with the tip model Eq. (8), quantitatively describes this behavior. First, we note that although an electrostatic model can correctly describe the spatial structure of the field near the tip apex,<sup>28</sup> it cannot account for a spectral dependence due to geometrical effects. Therefore, the use of a tip model including retardation effects is of great importance for spectroscopic applications.

The field near the apex of a perfectly conducting cone illuminated by a point source placed in the far field has a frequency dependence  $E \propto \omega^\nu$ , where  $\nu$  only depends on the angle of aperture of the tip [see Eq. (8)]. This model correctly describes the reciprocal field  $\mathbf{E}_{\text{rec}}$ , the illuminating

dipole being, in this case, placed at the position of the detector. The experimental field  $\mathbf{E}_{\text{exp}}^+$  contains several contributions, e.g., the field reflected by the flat mirror (without interaction with the tip) and the enhanced field at the tip apex, reflected by the surface. The latter is given by the product of Eq. (8) and a reflection factor which does not depend on the frequency. In the case of a confocal detection, as in Ref. 6, the signal  $I$  is given by the interference between these two contributions. Therefore, it is proportional to the integral term in Eq. (6), in which both  $\mathbf{E}_{\text{rec}}$  and  $\mathbf{E}_{\text{exp}}^+$  are described by Eq. (8). Finally, we end up with  $I \propto \omega^{2\nu-1}$ . For different tip angles, the signal predicted by this model versus the incident wavelength is shown in Fig. 2 in Ref. 6. An excellent agreement with the experiment is found.

Finally, we note that if the tip apex were modeled by a small dipole sphere<sup>29</sup>, the spectral dependence expected for the signal would be  $\omega^4$ , in disagreement with the experimental results. Therefore, modeling a conical metallic tip by a small dipole in apertureless SNOM leads to wrong predictions, at least for spectroscopic applications. In view of this result, we believe that the approach in the present article should find broad applications in near-field spectroscopy.

## IV. CONCLUSIONS

In conclusion, by means of the reciprocity theorem of electromagnetism, we have derived an exact and general expression for the signal in SNOM. This expression connects the field at the detector position to the local field illuminating the tip. It is valid in the presence of multiple scattering and can be applied to any type of SNOM probe. We have illustrated the potentiality of this approach by analyzing the apertureless setup. We have described the polarization effect and found a result in agreement with experiments.<sup>16</sup> We have also studied the spectral response which was measured experimentally very recently<sup>6</sup> and we have shown that its dependence on the tip shape was fully described by our approach. Besides, it turns out that a dipole model for the tip does not account for the spectral response observed experimentally. Therefore, we believe that the general expression for the SNOM signal introduced in this article should be helpful for quantitative analysis of future experimental results.

## ACKNOWLEDGMENTS

The authors thank L. Aigouy, J. C. Rivoal, and A. C. Boccara for many helpful discussions. We acknowledge financial support from the TMR program Near-Field Optics for Nanotechnology under Contract No. ERBFMRXCT98-0242 from the European Union.

## APPENDIX A: DERIVATION OF EXPRESSIONS (5) AND (6)

Using the vector identity

$$\mathbf{a} \cdot (\mathbf{b} \times \mathbf{c}) = \mathbf{b} \cdot (\mathbf{c} \times \mathbf{a}) = \mathbf{c} \cdot (\mathbf{a} \times \mathbf{b}), \quad (\text{A1})$$

one can cast Eq. (1) in the form

$$\begin{aligned} \mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} &= \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} \\ &+ \int_S [\mathbf{E}_{\text{exp}} \cdot (\mathbf{H}_{\text{rec}} \times \hat{\mathbf{z}}) - \mathbf{H}_{\text{exp}} \cdot (\hat{\mathbf{z}} \times \mathbf{E}_{\text{rec}})] d\mathbf{R}. \end{aligned} \quad (\text{A2})$$

The experimental fields have angular spectra given by Eqs. (2) and (3), related by the Maxwell Eq. (4). The angular spectrum representations of the reciprocal fields read

$$\mathbf{E}_{\text{rec}}(\mathbf{R}, z) = \int \mathbf{e}_{\text{rec}}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d\mathbf{K}, \quad (\text{A3})$$

$$\mathbf{H}_{\text{rec}}(\mathbf{R}, z) = \int \mathbf{h}_{\text{rec}}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d\mathbf{K}. \quad (\text{A4})$$

They are related by the Maxwell equation

$$\mathbf{k}^- \times \mathbf{e}_{\text{rec}}(\mathbf{K}) = \omega \mu_0 \mathbf{h}_{\text{rec}}(\mathbf{K}). \quad (\text{A5})$$

Inserting Eqs. (2), (3), (A3) and (A4) into Eq. (A2) leads to Eq. (5) after little algebra without any approximation. During this manipulation, some simplifications occur due to the transversality of the fields in the gap region, which yields, in terms of angular spectra

$$\begin{aligned} \mathbf{k}^\pm \cdot \mathbf{e}_{\text{exp}}^\pm(\mathbf{K}) &= 0, \\ \mathbf{k}^\pm \cdot \mathbf{h}_{\text{exp}}^\pm(\mathbf{K}) &= 0, \\ \mathbf{k}^- \cdot \mathbf{e}_{\text{rec}}(\mathbf{K}) &= 0, \\ \mathbf{k}^- \cdot \mathbf{h}_{\text{rec}}(\mathbf{K}) &= 0. \end{aligned} \quad (\text{A6})$$

In order to obtain the real-space expression (6), one has to invert the angular-spectrum representation of both fields  $\mathbf{E}_{\text{exp}}^+$  and  $\mathbf{E}_{\text{rec}}$ . From Eq. (2), one obtains

$$\mathbf{e}_{\text{exp}}^+(\mathbf{K}) \exp(i\gamma z) = \frac{1}{4\pi^2} \int \mathbf{E}_{\text{exp}}^+(\mathbf{R}, z) \exp(-i\mathbf{K} \cdot \mathbf{R}) d\mathbf{R}, \quad (\text{A7})$$

and from Eq. (A3), one obtains

$$\begin{aligned} -i\gamma \mathbf{e}_{\text{rec}}(-\mathbf{K}) \exp(-i\gamma z) \\ = \frac{1}{4\pi^2} \int \frac{\partial \mathbf{E}_{\text{rec}}}{\partial z}(\mathbf{R}, z) \exp(i\mathbf{K} \cdot \mathbf{R}) d\mathbf{R}. \end{aligned} \quad (\text{A8})$$

Inserting Eqs. (A7) and (A8) into Eq. (5) leads to Eq. (6).

### APPENDIX B: INCLUSION OF A FLAT SUBSTRATE IN THE GEOMETRY

In many experimental situations, the real sample is deposited on (or included in) a substrate. It may be useful in the description of such experiments to clearly separate the influence of the two. The purpose of this appendix is to show how the presence of the substrate can be accounted for in the model.

The experimental geometry we consider is depicted in Fig. 2(a). It is identical to that in Fig. 1(a), except that the substrate is now separated from the real sample described by the current density  $\mathbf{j}_{\text{exp}}$ . The reciprocal situation, represented in Fig. 2(b), includes the substrate as well. In these conditions, the reciprocal fields  $\mathbf{E}_{\text{rec}}$  and  $\mathbf{H}_{\text{rec}}$  have angular spectra

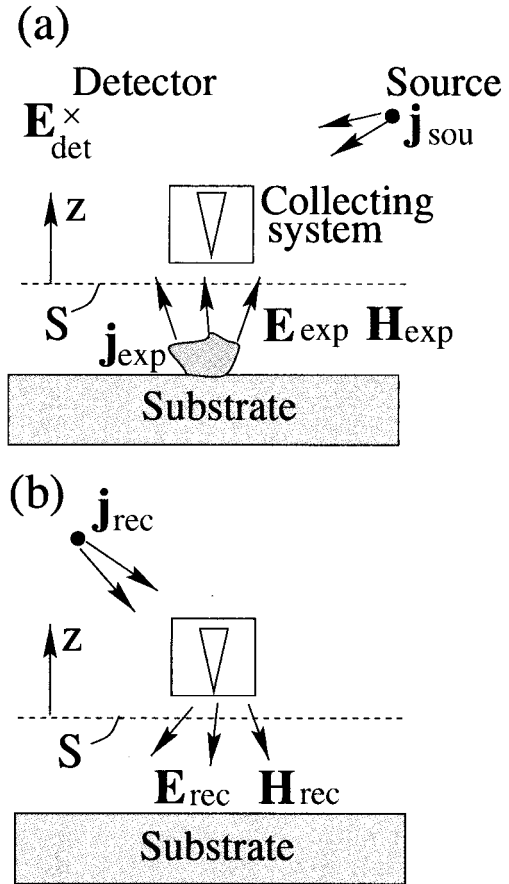


FIG. 2. Schematic views of the two situations considered for the application of the reciprocity theorem when the presence of the substrate is explicitly accounted for. (a): sketch of the experimental situation. (b): reciprocal situation, where the collecting system is illuminated by a point source located at the detector position in the presence of the substrate.

containing waves that propagate or decay towards  $z < 0$  and  $z > 0$ . Equations (A3) and (A4) have to be changed into

$$\begin{aligned} \mathbf{E}_{\text{rec}}(\mathbf{R}, z) &= \int \mathbf{e}_{\text{rec}}^+(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d\mathbf{K} \\ &+ \int \mathbf{e}_{\text{rec}}^-(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d\mathbf{K}, \end{aligned} \quad (\text{B1})$$

$$\begin{aligned} \mathbf{H}_{\text{rec}}(\mathbf{R}, z) &= \int \mathbf{h}_{\text{rec}}^+(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d\mathbf{K} \\ &+ \int \mathbf{h}_{\text{rec}}^-(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d\mathbf{K}. \end{aligned} \quad (\text{B2})$$

The angular spectra of both fields are related by

$$\mathbf{k}^\pm \times \mathbf{e}_{\text{rec}}^\pm(\mathbf{K}) = \omega \mu_0 \mathbf{h}_{\text{rec}}^\pm(\mathbf{K}). \quad (\text{B3})$$

Moreover, both fields are transverse, so that

$$\begin{aligned} \mathbf{k}^\pm \cdot \mathbf{e}_{\text{rec}}^\pm(\mathbf{K}) &= 0, \\ \mathbf{k}^\pm \cdot \mathbf{h}_{\text{rec}}^\pm(\mathbf{K}) &= 0. \end{aligned} \quad (\text{B4})$$

Following the same procedure as in Appendix A, one obtains the following expression for the electric field at the detector position:

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} - \frac{8\pi^2}{\omega\mu_o} \times \int \gamma(\mathbf{K}) [\mathbf{e}_{\text{rec}}^-(\mathbf{K}) \cdot \mathbf{e}_{\text{exp}}^+(\mathbf{K}) - \mathbf{e}_{\text{exp}}^-(\mathbf{K}) \cdot \mathbf{e}_{\text{rec}}^+(\mathbf{K})] d\mathbf{K}. \quad (\text{B5})$$

This expression is an extension of Eq. (5) to the case where the substrate is included in the reciprocal geometry. This means that its presence is completely described by the reciprocal field  $\mathbf{E}_{\text{rec}}$ . In particular, this field may account for a strong interaction between the tip and the substrate

(e.g., plasmon coupling). The main difference is that in this representation, both  $\mathbf{E}_{\text{exp}}^+$  and  $\mathbf{E}_{\text{exp}}^-$  enter the relationship.

The expression in real space is obtained, as in Appendix A, by inverting Eqs. (2) and (B1)

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} - \frac{2i}{\omega\mu_o} \times \int_S \left[ \frac{\partial \mathbf{E}_{\text{rec}}^-}{\partial z}(\mathbf{R}, z) \cdot \mathbf{E}_{\text{exp}}^+(\mathbf{R}, z) + \frac{\partial \mathbf{E}_{\text{rec}}^+}{\partial z}(\mathbf{R}, z) \cdot \mathbf{E}_{\text{exp}}^-(\mathbf{R}, z) \right] d\mathbf{R}. \quad (\text{B6})$$

As we discussed previously, it is also possible to derive an expression of the signal in terms of magnetic fields only. We obtain in this case

$$\mathbf{E}_{\text{det}} \cdot \mathbf{j}_{\text{rec}} = \mathbf{E}_{\text{rec}}(\mathbf{r}_{\text{sou}}) \cdot \mathbf{j}_{\text{sou}} + \frac{2i}{\omega\epsilon_o} \times \int_S \left[ \frac{\partial \mathbf{H}_{\text{rec}}^-}{\partial z}(\mathbf{R}, z) \cdot \mathbf{H}_{\text{exp}}^+(\mathbf{R}, z) + \frac{\partial \mathbf{H}_{\text{rec}}^+}{\partial z}(\mathbf{R}, z) \cdot \mathbf{H}_{\text{exp}}^-(\mathbf{R}, z) \right] d\mathbf{R}. \quad (\text{B7})$$

- <sup>1</sup>D. W. Pohl, W. Denk, and M. Lanz, *Appl. Phys. Lett.* **44**, 651 (1984); E. Betzig, A. Lewis, A. Harootunian, M. Isaacson, and E. Kratschmer, *Biophys. J.* **49**, 269 (1986).
- <sup>2</sup>*Near-Field Optics*, edited by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993); *Near-Field Nano/Atom Optics and Technology*, edited by M. Ohtsu (Springer, Tokyo, 1998).
- <sup>3</sup>O. Marti, H. Bielefeldt, B. Hecht, S. Herminghaus, P. Leiderer, and J. Mlynek, *Opt. Commun.* **96**, 225 (1993); B. Hecht, H. Bielefeldt, L. Novotny, Y. Inouye, and D. W. Pohl, *Phys. Rev. Lett.* **77**, 1889 (1996); S. I. Bozhevolnyi and F. A. Pudonin, *Phys. Rev. Lett.* **78**, 2823 (1997); J. R. Krenn, A. Dereux, J. C. Weeber, E. Bourillot, Y. Lacroute, J. P. Goudonnet, G. Schider, W. Gotschy, A. Leitner, F. R. Aussenegg, and C. Girard, *Phys. Rev. Lett.* **82**, 2590 (1999).
- <sup>4</sup>E. Betzig and R. Chichester, *Science* **262**, 1422 (1993); J. Azoulay, A. Debarre, A. Richard, and P. Tchénio, *J. Microsc.* **194**, 486 (1999).
- <sup>5</sup>E. Betzig and J. K. Trautman, *Science* **257**, 189 (1992); Y. Martin, F. Zenhausern, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **68**, 2475 (1996); C. L. Jahncke, H. D. Hallen, and M. A. Paesler, *J. Raman Spectrosc.* **27**, 579 (1996); J. Grausem, B. Humbert, A. Burneau, and J. Oswald, *Appl. Phys. Lett.* **70**, 1671 (1997).
- <sup>6</sup>L. Aigouy, F. X. Andréani, A. C. Boccara, J. C. Rivoal, J. A. Porto, R. Carminati, J.-J. Greffet, and R. Mégyry, *Appl. Phys. Lett.* **76**, 397 (2000).
- <sup>7</sup>S. I. Bozhevolnyi, *Phys. Rev. B* **54**, 8177 (1996); S. Grésillon, L. Aigouy, A. C. Boccara, J. C. Rivoal, X. Quelin, C. Desmarest, P. Gadenne, V. A. Shubin, A. K. Sarychev, and V. M. Shalaev, *Phys. Rev. Lett.* **82**, 4520 (1999).
- <sup>8</sup>S. Bourzeix, J. M. Moison, F. Mignard, F. Barthe, A. C. Boccara, C. Licoppe, B. Mersali, M. Allovon, and A. Bruno, *Appl. Phys. Lett.* **73**, 1035 (1998); M. L. M. Balistreri, A. Driessen, J. P. Korterik, L. Kuipers, and N. F. van Hulst, *Opt. Lett.* **25**, 637 (2000).
- <sup>9</sup>Ch. Lieneau, A. Richter, and T. Elsaesser, *Appl. Phys. Lett.* **69**, 325 (1996); R. Bachelot, G. Wurtz, and P. Royer, *Appl. Phys. Lett.* **73**, 3333 (1998).
- <sup>10</sup>C. Girard and D. Courjon, *Phys. Rev. B* **42**, 9340 (1990).
- <sup>11</sup>L. Novotny, D. W. Pohl, and P. Regli, *J. Opt. Soc. Am. A* **11**, 1768 (1994).
- <sup>12</sup>C. Girard and A. Dereux, *Rep. Prog. Phys.* **59**, 657 (1996).
- <sup>13</sup>D. van Labeke and D. Barchiesi, *J. Opt. Soc. Am. A* **9**, 732 (1992); **10**, 2193 (1993).
- <sup>14</sup>J. J. Greffet and R. Carminati, *Prog. Surf. Sci.* **56**, 133 (1997).
- <sup>15</sup>F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **65**, 1623 (1994); Y. Inouye and S. Kawata, *Opt. Lett.* **19**, 159 (1994); P. Gleyzes, A. C. Boccara, and R. Bachelot, *Ultramicroscopy* **57**, 318 (1995).
- <sup>16</sup>L. Aigouy, A. Lahrech, S. Grésillon, H. Cory, A. C. Boccara, and J. C. Rivoal, *Opt. Lett.* **24**, 187 (1999).
- <sup>17</sup>L. Landau, E. Lifchitz, and L. Pitaevskii, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984); see also R. Carminati, M. Nieto-Vesperinas and J.-J. Greffet, *J. Opt. Soc. Am. A* **15**, 706 (1998).
- <sup>18</sup>P. Johansson, R. Monreal, and P. Apell, *Phys. Rev. B* **42**, 9210 (1990).
- <sup>19</sup>M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991), Chap. 3, p. 83.
- <sup>20</sup>R. Carminati and J. J. Sáenz, *Phys. Rev. Lett.* **84**, 5156 (2000).
- <sup>21</sup>J. Bardeen, *Phys. Rev. Lett.* **6**, 57 (1961).
- <sup>22</sup>J. Tersoff and D. R. Hamann, *Phys. Rev. B* **31**, 805 (1985).
- <sup>23</sup>J. C. Weeber, Ph.D. thesis, Burgundy University, Dijon, France, 1996; A. Dereux, Ch. Girard, and J.-C. Weeber, *J. Chem. Phys.* **112**, 7775 (2000).
- <sup>24</sup>H. A. Bethe, *Phys. Rev.* **66**, 163 (1944); C. J. Bouwkamp, *Philips Res. Rep.* **5**, 321 (1950); **5**, 401 (1950).
- <sup>25</sup>D. van Labeke, F. Baida, D. Barchiesi, and D. Courjon, *Opt. Commun.* **114**, 470 (1995); D. van Labeke, D. Barchiesi, and F. Baida, *J. Opt. Soc. Am. A* **12**, 695 (1995).
- <sup>26</sup>*Electromagnetic and Acoustic Scattering by Simple Shapes*, edited by J.J. Bowman, T.B.A. Senior, and P.L.E. Uslenghi (North-Holland, Amsterdam, 1969), p. 662.
- <sup>27</sup>H. Cory, A. C. Boccara, J. C. Rivoal, and A. Lahrech, *Microwave Opt. Technol. Lett.* **18**, 120 (1998).
- <sup>28</sup>W. Denk and D. W. Pohl, *J. Vac. Sci. Technol. B* **9**, 510 (1991).
- <sup>29</sup>F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, *Science* **269**, 1083 (1995); C. J. Hill, P. M. Bridger, G. S. Picus, and T. C. McGill, *Appl. Phys. Lett.* **75**, 4022 (1999).

# Theory of near-field magneto-optical imaging

Julian N. Walford, Juan-Antonio Porto, Rémi Carminati, and Jean-Jacques Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion; Ecole Centrale Paris, Centre Nationale de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

Received March 14, 2001; accepted July 16, 2001

Scanning near-field optical microscopy has been recently applied to the imaging of magnetic samples. It was shown experimentally that an apertureless microscope suffers a substantial loss of resolution when used for magneto-optical imaging compared with that for conventional imaging. No such change is observed for aperture microscopes. We explain this observation by developing a model for the imaging process that incorporates the response of the probe. We calculate real observable properties such as the rotation of polarization at the detector or the circular dichroism signal and thus simulate magneto-optical images of a domain structure in cobalt for both aperture and apertureless microscopes. © 2002 Optical Society of America  
*OCIS codes:* 180.5810, 350.5730, 260.1960, 260.2110, 210.3820.

## 1. INTRODUCTION

Scanning near-field optical microscopy (SNOM) is a technique that has enabled the diffraction resolution limit in optical microscopy to be beaten through the use of subwavelength-sized probes scanned in close proximity to a sample.<sup>1–3</sup> The optical nature of the technique has led to applications in a wide range of areas, including fluorescence microscopy,<sup>4</sup> local spectroscopy,<sup>5</sup> plasmons,<sup>6,7</sup> and magneto-optical imaging.<sup>8,9</sup>

SNOM seems to be an ideally suited tool for magneto-optical imaging for two reasons. Unlike magnetic force microscopy, magneto-optical SNOM (MO-SNOM) imaging allows passive measurement of the sample field without introduction of an external magnetic field. MO-SNOM should also be able to provide a resolution superior to that of far-field optical techniques. Nevertheless, the imaging process is not completely understood.

Magneto-optical contrast is due to the rotation of polarization of the illuminating field caused by the magnetization in a sample. The magneto-optical signal can be distinguished from the conventional optical signal by measurement of the Faraday or Kerr rotation through polarization analysis at source and detector<sup>8–15</sup> or by measurement of circular dichroism induced by the sample magnetization.<sup>16–19</sup> In the latter, the illumination is modulated between left and right circular polarizations, and lock-in detection is used to measure a difference in absorption between the two polarization states. In the most commonly used geometry, the sample is locally illuminated by an aperture probe, and a signal is detected in the far field, through an analyzer oriented differently to the illumination polarization.<sup>8,10–12</sup>

Complete control of the polarization is difficult. No matter how well polarized is the light coupled into the fiber, the light emerging from the small aperture at its tip typically has an extinction ratio of the order of 1:20.<sup>10,11</sup> This is a limiting factor in the accuracy to which the angle of rotation can be measured.<sup>12</sup> Substantially better polarization control is achievable in an apertureless experi-

ment, in which the probe and the sample are illuminated by an external focused laser beam.<sup>19</sup>

While spatial resolutions as good as 10 nm have been obtained in conventional optical experiments with apertureless SNOM, apertureless MO-SNOM experiments have not demonstrated resolutions better than a few hundred nanometers.<sup>16,17,19</sup> This gross disparity in the achievable resolution is even observed when the same apparatus is used for both conventional optical and magneto-optical imaging.<sup>17</sup> On the other hand, aperture microscopes seem to obtain a similar resolution in optical and magneto-optical experiments, this being as good as 30–50 nm.<sup>8</sup> Understanding the response of the probe is clearly important if this problem is to be explained, since different results are obtained in aperture and apertureless experiments. A number of theoretical studies of near-field magneto-optical imaging have been performed previously.<sup>20–24</sup> Usually, the electric field distribution in the near field of a sample has been calculated, and a magneto-optical signal is determined based the angle of rotation of polarization or the absorption of different circular polarization states. However, none of these models studies the response of the probe, and therefore none of them can explain the observed loss of resolution.

The objective of this paper is to develop a model for the magneto-optical imaging process that takes into account the probe response, making it possible to answer some of the open topics regarding MO-SNOM, particularly that of understanding the loss of resolution of MO-SNOM for apertureless probes. The formalism will be applied to both aperture and apertureless experiments, and a response function to the sample magnetization will be developed for both cases.

## 2. DEVELOPMENT OF A GENERAL EXPRESSION FOR THE SIGNAL

In previous papers, Greffet and Carminati<sup>25</sup> and Porto *et al.*<sup>26</sup> have used the electromagnetic theorem of reci-



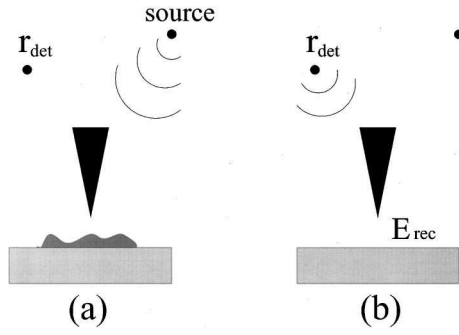


Fig. 1. (a) Scheme of a general SNOM setup and (b) fictitious reciprocal situation.

propensity to develop an expression for the response of a near-field microscope. A key feature of this approach is that it yields an exact expression for the signal that accounts for the properties of the tip. The theorem of reciprocity relates the electric and magnetic fields created by two different current distributions in the presence of a scattering object with linear and symmetric constitutive tensors.<sup>25,27,28</sup> In this paper, the *magnetically* induced currents in the sample are treated as an external source term, and the probe and the substrate are treated as the scattering object. The probe and the substrate, being nonmagnetic, have symmetric constitutive tensors, and thus the requirements of the reciprocity theorem are satisfied.

Let us consider a general SNOM setup, as depicted in Fig. 1(a). An inhomogeneous sample is deposited on (or embedded in) a flat homogeneous substrate. It is illuminated either through the tip (illumination-mode SNOM) or with an external beam (collection-mode and apertureless SNOM). The signal is recorded by a point detector placed in the far field at a position  $\mathbf{r}_{\text{det}}$ . We assume that an analyzer is placed in front of the detector, with a polarization direction defined by the unit vector  $\hat{\mathbf{p}}$ . Through application of the reciprocity theorem, the component  $A$  of the electric field at the detector along the direction of the analyzer has been shown to be<sup>25</sup>

$$A = \mathbf{E}_{\text{exp}}(\mathbf{r}_{\text{det}}) \cdot \hat{\mathbf{p}} = \frac{1}{i\omega} \int_V \mathbf{E}_{\text{rec}} \cdot \mathbf{J}_{\text{exp}} \, d\mathbf{r}. \quad (1)$$

In this expression,  $\mathbf{J}_{\text{exp}}$  and  $\mathbf{E}_{\text{exp}}(\mathbf{r}_{\text{det}})$  are the current density in the sample and the electric field at the detector position, respectively, in the experimental situation corresponding to Fig. 1(a).  $\mathbf{E}_{\text{rec}}$  is the electric field that would be produced by a dipole source of amplitude  $\hat{\mathbf{p}}$  placed at the detector position  $\mathbf{r}_{\text{det}}$  in the absence of the sample. This fictitious reciprocal situation is represented in Fig. 1(b). Note that the reciprocal situation contains the tip and all of the illumination–detection system (only the sample is removed). Therefore the reciprocal field  $\mathbf{E}_{\text{rec}}$  is the key quantity that contains all the information about the response of the setup to the excitation of a current  $\mathbf{J}_{\text{exp}}$  in a given sample.

### 3. RESPONSE FUNCTIONS FOR A MAGNETIC SAMPLE

The current density induced in a magnetic sample is

$$\mathbf{J}(\mathbf{r}, \omega) = -i\omega\epsilon_0[\vec{\epsilon}(\mathbf{r}, \omega) - \vec{\mathbf{I}}] \cdot \mathbf{E}_{\text{exp}}(\mathbf{r}, \omega), \quad (2)$$

where  $\vec{\epsilon}(\mathbf{r}, \omega)$  is the frequency-dependent dielectric tensor. It can be written as a sum of nonmagnetic and magnetically induced terms:

$$\mathbf{J} = \mathbf{J}_{\epsilon} + \mathbf{J}_M = -i\omega\epsilon_0[(\epsilon_1 - 1)\mathbf{E}_{\text{exp}} + ifM\hat{\mathbf{e}} \times \mathbf{E}_{\text{exp}}], \quad (3)$$

where  $\mathbf{M} = M\hat{\mathbf{e}}$  is the magnetization in the sample,  $\hat{\mathbf{e}}$  is a unit vector, and  $f$  is a constant of proportionality. The dependence on the magnetization is entirely within the second term, which is antisymmetric. If the magnetization is directed along the  $z$  axis, Eq. (3) corresponds to a dielectric tensor

$$\vec{\epsilon} = \begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_1 & 0 \\ 0 & 0 & \epsilon_1 \end{bmatrix} + \begin{bmatrix} 0 & -ifM & 0 \\ ifM & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \quad (4)$$

We will consider the magnetization to be in an arbitrary direction  $\hat{\mathbf{e}}$  in the following.

Isolating the component of the field at the detector that has a dependence on the magnetization, we obtain the following from Eqs. (1) and (3):

$$A_{\text{mag}} = -if\epsilon_0 \int_V M(\hat{\mathbf{e}} \times \mathbf{E}_{\text{exp}}) \cdot \mathbf{E}_{\text{rec}} \, d\mathbf{r}, \quad (5)$$

which can be rearranged to give

$$A_{\text{mag}} = -if\epsilon_0 \int_V M\hat{\mathbf{e}} \cdot (\mathbf{E}_{\text{exp}} \times \mathbf{E}_{\text{rec}}) \, d\mathbf{r}. \quad (6)$$

Note that both  $\mathbf{E}_{\text{rec}}$  and  $\mathbf{E}_{\text{exp}}$  depend implicitly on the position of the tip,  $\mathbf{r}_{\text{tip}}$ . If we define a constant-height amplitude response for the magnetization in the sample plane  $z$  (probe at height  $z_{\text{tip}}$ ),  $H_{\text{mag}}(x - x_{\text{tip}}, y - y_{\text{tip}}, z, z_{\text{tip}})$ , by

$$A_{\text{mag}}(\mathbf{r}_{\text{tip}}) = \int_V H_{\text{mag}}(\mathbf{R} - \mathbf{R}_{\text{tip}}, z, z_{\text{tip}}) M(\mathbf{r}) \, d\mathbf{r}, \quad (7)$$

with  $\mathbf{R} = (x, y)$ , then this response function is

$$H_{\text{mag}} \propto \hat{\mathbf{e}} \cdot (\mathbf{E}_{\text{exp}} \times \mathbf{E}_{\text{rec}}), \quad (8)$$

to within a constant factor. Similarly, a response function  $H_{\epsilon}$  for the variation of  $\epsilon_1$  in a nonmagnetic sample, defined by

$$A_{(M=0)}(\mathbf{r}_{\text{tip}}) = \int_V H_{\epsilon}(\mathbf{R} - \mathbf{R}_{\text{tip}}, z, z_{\text{tip}}) \epsilon_1(\mathbf{r}) \, d\mathbf{r}, \quad (9)$$

can be shown to be proportional to

$$H_{\epsilon} \propto (\mathbf{E}_{\text{exp}} \cdot \mathbf{E}_{\text{rec}}), \quad (10)$$

from Eqs. (1) and (3).

In the following sections, the response functions  $H_{\text{mag}}$  and  $H_{\epsilon}$  will be the key concepts. They will be evaluated for both apertureless and aperture microscopes, making it possible to discuss the magneto-optical imaging properties of these two experimental setups.

#### 4. OBSERVABLE MAGNETO-OPTICAL SIGNALS: ROTATION OF POLARIZATION AND CIRCULAR DICHOISM

The quantity  $A$  that we have associated with a signal up to here is the amplitude of the field at the position of the detector, projected along the axis of an analyzer. Of course, this is not what is actually measured in the course of a SNOM experiment. In conventional SNOM, it is the intensity of the field, either alone or with a coherent background. In a magneto-optical experiment, often the measurable quantity is the angle of rotation of polarization by the magnetic sample, or the dichroic signal as the incident polarization is modulated between left and right.

In this section, we will show how this theory makes it possible to completely determine the complex vectorial electric field at the detector, from which all measurable quantities can be determined. We demonstrate the existence of a response for such measurements as field polarization direction and circular dichroism signals. An expression for the ellipticity is given in Appendix A.

Equation (1) gives the component of the electric field at the detector directed along a unit vector  $\hat{\mathbf{p}}$ . We can thus determine the components of the field along two orthogonal axes ( $\hat{\mathbf{u}}$  and  $\hat{\mathbf{v}}$ ) in a transverse plane at the detector. These two field components are labeled  $E_{\text{det},\mathbf{u}}$  and  $E_{\text{det},\mathbf{v}}$ . The full electric field at the detector is given by

$$\mathbf{E}_{\text{det}} = E_{\text{det},\mathbf{u}}\hat{\mathbf{u}} + E_{\text{det},\mathbf{v}}\hat{\mathbf{v}}. \quad (11)$$

Given the amplitudes of the two vector components and their relative phase  $\delta$ , one can calculate the direction of polarization of the field, an angle  $\theta$  relative to the  $\hat{\mathbf{u}}$  axis, by using<sup>29</sup>

$$\tan 2\theta = \frac{2|E_{\text{det},\mathbf{u}}||E_{\text{det},\mathbf{v}}|}{|E_{\text{det},\mathbf{u}}|^2 - |E_{\text{det},\mathbf{v}}|^2} \cos \delta. \quad (12)$$

The circular dichroism signal can be approximated as the difference between the intensities measured when the experiment is illuminated with right and left circular polarizations,  $|\mathbf{E}_{\text{det}}^{(R)}|^2$  and  $|\mathbf{E}_{\text{det}}^{(L)}|^2$ . The dichroic signal can also be expressed in terms of the fields at the detector with  $s$ - and  $p$ -polarized illumination:

$$I_{\text{dichroic}} = |\mathbf{E}_{\text{det}}^{(R)}|^2 - |\mathbf{E}_{\text{det}}^{(L)}|^2 = 2 \operatorname{Re}[i\mathbf{E}_{\text{det}}^{(s)} \cdot \mathbf{E}_{\text{det}}^{(p)*}]. \quad (13)$$

A fuller development of these expressions is given in Appendix A.

#### 5. APPLICATION TO APERTURELESS SCANNING NEAR-FIELD OPTICAL MICROSCOPY

Magneto-optical apertureless SNOM experiments have been performed in both reflection and transmission modes.<sup>15,17</sup> We will discuss the reflection-mode experiment in this paper, but the same arguments are applicable to a transmission-mode experiment. A simplified illustration of a reflection-mode experiment is given in Fig. 1(a). In apertureless SNOM, both the illumination and the detection are external, and the tip acts as a local scatterer (no coupling with guided modes in a fiber). We use the Born approximation for the experimental field,

which is justified by the weak levels of magnetically induced fields (2 orders of magnitude smaller than the conventional optical fields induced in cobalt, for example, with  $\epsilon_1 = -12.3 + i18.4$  and  $ifM = -0.4 - i0.1$  at 633 nm). In this approximation, the experimental field is simply the field scattered by the probe and the substrate when illuminated by the experimental source in the absence of the magnetic sample, labeled  $\mathbf{E}_{\text{probe}}^{(\text{sou})}$ .

The reciprocal situation is depicted in Fig. 1(b). To determine the reciprocal field, we placed a dipole source at the detector position and removed the sample (i.e.,  $M$  is put to zero). The reciprocal field  $\mathbf{E}_{\text{rec}}$  is the field diffracted by the probe and the substrate with illumination from the detector position, labeled  $\mathbf{E}_{\text{probe}}^{(\text{det})}$ . Thus the field response function to magnetization for an apertureless experiment, from expression (8), with  $\mathbf{r}$  and  $\mathbf{r}_{\text{tip}}$  dependencies suppressed for clarity, is

$$H_{\text{mag}} \propto \hat{\mathbf{e}} \cdot [\mathbf{E}_{\text{probe}}^{(\text{sou})} \times \mathbf{E}_{\text{probe}}^{(\text{det})}]. \quad (14)$$

The field response function for the linear component of the dielectric tensor from expression (10) is

$$H_e \propto [\mathbf{E}_{\text{probe}}^{(\text{sou})} \cdot \mathbf{E}_{\text{probe}}^{(\text{det})}]. \quad (15)$$

To explore the consequences of this result, we will use a specific model for the probe, that of a perfectly conducting cone.<sup>30,31</sup> This has been experimentally validated<sup>32</sup> and is a good model for apertureless SNOM performed by using metallic tips.<sup>33–35</sup> One of the main features of this model is the existence of a singularity of the electric field at the cone tip. The field enhancement and confinement that this produces are responsible for the good signal and resolution normally obtained with this type of probe. The full field under the tip consists of a number of modes, of which only one contains the singularity. The other modes are much lower in amplitude, are less well confined near the probe tip, and do not provide a significant contribution to the imaging properties of the probe in conventional imaging.

Before we continue, it is worth briefly reviewing the origin of the magneto-optical signal. The theorem of reciprocity shows that the components of the field at the detector are given by the expression [Eq. (1)]

$$A = \frac{1}{i\omega} \int_V \mathbf{E}_{\text{probe}}^{(\text{det})} \cdot \mathbf{J}_{\text{exp}} \, d\mathbf{r}. \quad (16)$$

The reciprocal field represents the response of the probe to sample currents. In the Born approximation, the induced current density in the sample,  $\mathbf{J}_{\text{exp}}$ , is given by [Eq. (3)]

$$\mathbf{J}_{\text{exp}} = (-i\omega)\epsilon_0 \left[ \overbrace{(\epsilon_1 - 1)\mathbf{E}_{\text{probe}}^{(\text{sou})}}^{\mathbf{J}_e} + \overbrace{ifM\hat{\mathbf{e}} \times \mathbf{E}_{\text{probe}}^{(\text{sou})}}^{\mathbf{J}_M} \right]. \quad (17)$$

The magnetically induced current density  $\mathbf{J}_M$  is always *orthogonal* to the field that induces it,  $\mathbf{E}_{\text{probe}}^{(\text{sou})}$ , because of the cross product. In the following sections, we shall demonstrate that the reciprocal and experimental fields  $[\mathbf{E}_{\text{probe}}^{(\text{det})}$  and  $\mathbf{E}_{\text{probe}}^{(\text{sou})}]$  associated with the singularity are always *parallel* to each other, no matter what direction and polarization of detection or illumination are used. The

magnetically induced current density produced by the probe singularity ( $\mathbf{J}_M$ ) is thus orthogonal to the reciprocal field [ $\mathbf{E}_{\text{probe}}^{(\text{det})}$ ] everywhere. As a result, the field at the detector due to current density  $\mathbf{J}_M$  [Eqs. (16) and (17)] is identically zero.

The immediate conclusion that is to be drawn from this is that the probe singularity alone does *not* contribute to the magneto-optical signal. This is not to say that it is impossible to record a magneto-optical image by using a metallic apertureless probe. The experimental evidence<sup>15,17</sup> clearly contradicts this false conclusion. The magneto-optical signal that is recorded is due to non-singular components of the probe fields; these being less well confined, the attainable resolution is poorer.

This result is true no matter what detection technique is used: a measurement either of the polarization of the outgoing beam or of the dichroism in the sample. The field at the detector due to the singularity alone (which normally provides the good resolution) is completely insensitive to variations in magnetization in the sample.

To demonstrate this conclusion, we first discuss the mathematical origin of the singularity and look in some detail at the form of the electric field scattered from the probe. The consequences for imaging resolution are then illustrated in Subsection 5.C.

### A. Cone Model for the Probe

We give a brief outline of some of the relevant mathematical features of the electric field scattered by an infinite perfectly conducting cone,<sup>31</sup> a model that has given results in quantitative agreement with experiment.<sup>32</sup> In particular, the presence of a field singularity and the form of the field associated with it will be developed. We consider a cone illuminated by a plane-wave source, incident from an angle  $\theta_0$  to the positive vertical axis, forming an angle  $\phi_0$  with the  $x$ - $z$  plane, and polarized at an angle  $\beta$  with the normal to the plane of incidence ( $\beta = \pi/2$  corresponds to  $p$  polarization,  $\beta = 0$  corresponds to  $s$  polarization). This geometry is depicted in Fig. 2. For this situation, the total field is calculated from Debye potentials  $u$  and  $v$ , in polar coordinates, by using

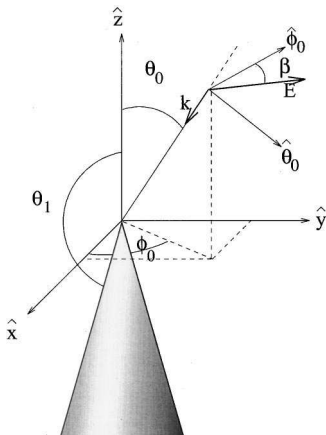


Fig. 2. Cone and illumination geometry: definitions of variables for calculation of Debye potentials for a cone.

$$\begin{aligned} E_r &= \left( \frac{\partial^2}{\partial r^2} + k^2 \right) (ru), \\ E_\theta &= \frac{1}{r} \frac{\partial^2}{\partial r \partial \theta} (ru) + \frac{ik}{\sin \theta} \left( \frac{\mu_0}{\epsilon_0} \right)^{1/2} \frac{\partial v}{\partial \phi}, \\ E_\phi &= \frac{1}{r \sin \theta} \frac{\partial^2}{\partial r \partial \phi} (ru) - ik \left( \frac{\mu_0}{\epsilon_0} \right)^{1/2} \frac{\partial v}{\partial \theta}. \end{aligned} \quad (18)$$

The Debye potentials for a cone illuminated by a plane wave can be written as

$$\begin{aligned} u(r, \theta, \phi) &= \sum_{m,p} f(r, \theta, \theta_1, m, p) \\ &\quad \times \left[ (m \sin m \phi \cos \beta) \frac{P_p^m(\cos \theta_0)}{\sin \theta_0} \right. \\ &\quad \left. + (\cos m \phi \sin \beta) \frac{\partial}{\partial \theta_0} P_p^m(\cos \theta_0) \right], \\ v(r, \theta, \phi) &= \sum_{m,p} g(r, \theta, \theta_1, m, p) \\ &\quad \times \left[ (\cos m \phi \cos \beta) \frac{\partial}{\partial \theta_0} P_p^m(\cos \theta_0) \right. \\ &\quad \left. - (m \sin m \phi \sin \beta) \frac{P_p^m(\cos \theta_0)}{\sin \theta_0} \right]. \end{aligned} \quad (19)$$

The field created by a transverse (no  $\hat{\mathbf{r}}$  component) unit dipole source at distance  $r_0$  ( $kr_0 \gg 1$ ) is the same but is multiplied by a factor  $k^2 \exp(ikr_0)/(4\pi\epsilon_0 r_0)$ . Further details are given in Appendix B and in Ref. 31.

The potentials, and consequently the fields, are a sum over a number of modes. Several of these are shown in Fig. 3 for increasing values of a mode index  $m$ . Two clear characteristics can be seen. First, the  $m$  index governs the azimuthal dependence of the field, with higher modes having higher orders of rotational symmetry. The field has a mixed  $\cos m\phi$  and  $\sin m\phi$  dependence on the azimuthal angle  $\phi$ . Second, the higher the mode number, the less well confined the field. For small  $r$  (i.e., close to the probe tip), the field depends on  $r$  like  $(kr)^{p-1}$ , where  $p$  is a second index that is always greater than  $m$ .

In fact, the first mode ( $m = 0$ ) is divergent at the probe tip. For a cone of interior half-angle  $30^\circ$ , the first value of  $p$  is approximately 0.346, giving a leading-order field dependence of  $(kr)^{-0.654}$ . The two dominant components of the electric field ( $E_r$  and  $E_\theta$ ) consequently diverge at the probe tip. This is the case for any cone. It can be seen in Fig. 3(a) that while the  $m = 0$  field is very large immediately beneath the probe tip, it falls to zero very rapidly. The presence of a singularity in the response function leads to strong signal levels and good resolution in the image.

The dielectric response function  $H_\epsilon$ , as defined above, is shown in Fig. 4(a). The component of this response due to the nondivergent modes is shown in Fig. 4(b). The dominance of the singular component can be clearly seen. This term has also been shown to be responsible for the

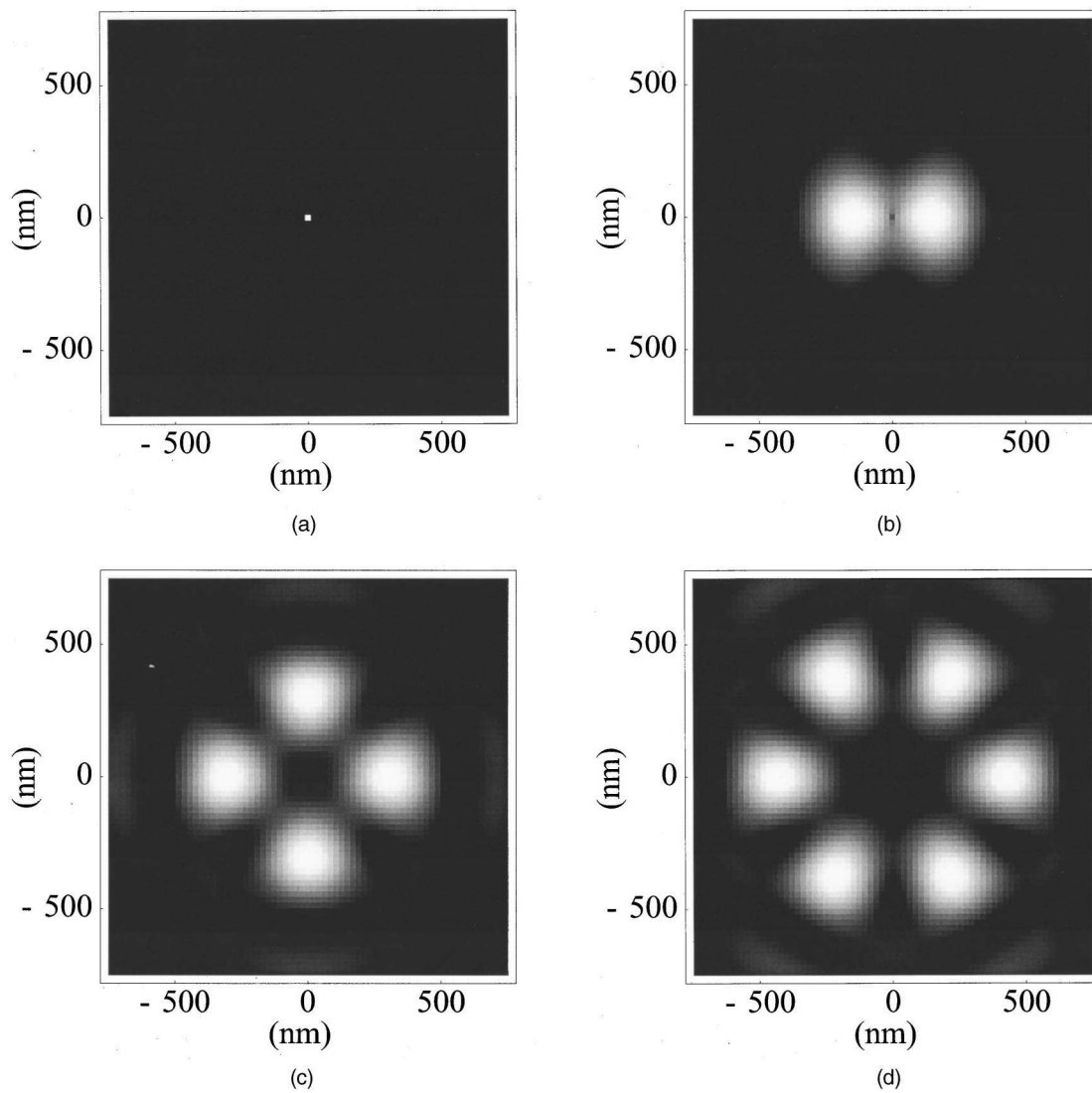


Fig. 3. Azimuthal field behavior for different values of  $m$ : intensity of the field components shown in a horizontal plane 1 nm below the probe. The  $m$  index determines the azimuthal symmetry of the field, and for increasing  $m$  the field is less well confined.

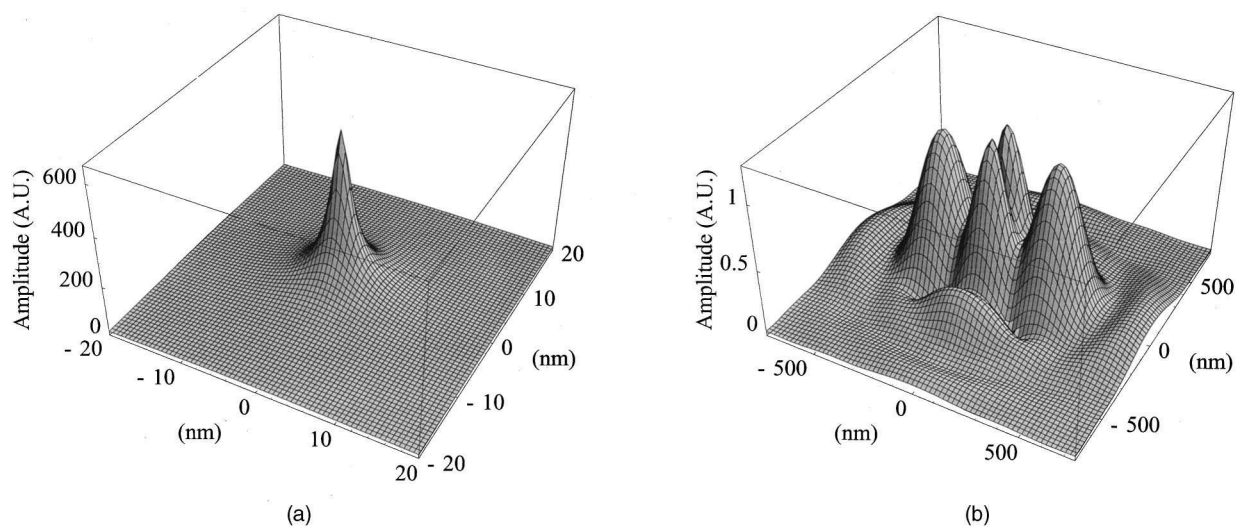


Fig. 4. Dielectric response function: field response to variation in the permittivity of a nonmagnetic material, evaluated for a plane 1 nm below the probe. (a) Component due to the singularity alone and (b) nonsingular component.



spectroscopic response of a metallic apertureless probe.<sup>32</sup> In conventional optical imaging, all the physics comes from the singularity at the probe tip.

### B. Null Magnetic Response Due to the Field Singularity

The magnetic response function, as we have seen [relation (14)], is proportional to

$$H_{\text{mag}} \propto \hat{\mathbf{e}} \cdot [\mathbf{E}_{\text{cone}}^{(\text{sou})} \times \mathbf{E}_{\text{cone}}^{(\text{det})}], \quad (20)$$

where  $\mathbf{E}_{\text{cone}}^{(\text{sou})}$  is the field below the cone when illuminated by the experimental source in the absence of sample and  $\mathbf{E}_{\text{cone}}^{(\text{det})}$  is the field below the cone when illuminated by the reciprocal source placed at the detector.

The singular electric field terms from Eqs. (18) and (19) are

$$\begin{aligned} E_r &= \left[ \left( \frac{\partial^2}{\partial r^2} + k^2 \right) r f(r, \theta, \theta_1, 0, p_1) \right] \\ &\quad \times (\sin \beta) \frac{\partial}{\partial \theta_0} P_p^0(\cos \theta_0), \\ E_\theta &= \left[ \frac{1}{r} \frac{\partial^2}{\partial r \partial \theta} r f(r, \theta, \theta_1, 0, p_1) \right] \\ &\quad \times (\sin \beta) \frac{\partial}{\partial \theta_0} P_p^0(\cos \theta_0), \end{aligned} \quad (21)$$

with the angle  $\theta_0$  equal to the angle of incidence from the experimental source or detector and the polarization  $\beta$  equal to the source polarization or direction of polarization at the detector for  $\mathbf{E}_{\text{exp}}$  and  $\mathbf{E}_{\text{rec}}$ , respectively. A first-order expression for these fields is given in Appendix B.

The spatial distribution of the field is determined entirely by the function  $f(r, \theta, \theta_1, 0, p_1)$ , which is independent of the illumination direction ( $\theta_0, \phi_0$ ) and polarization ( $\beta$ ). Changing the illumination conditions changes only the amplitude of the field.  $\mathbf{E}_{\text{cone}}^{(\text{sou})}$  and  $\mathbf{E}_{\text{cone}}^{(\text{det})}$  are identical except for an amplitude factor. The magneto-optical response in relation (20) due to the singularity alone is thus zero! This is contrary to the case of conventional optical imaging [relation (15)], where it is almost exclusively the field singularity that produces the image.

Any detectable magneto-optical signal is due to the full spectrum of nondivergent field modes below the probe. The higher-order modes being less well confined, it will be seen that the best attainable resolution (determined by the width of the response function) is much poorer for the magnetic signal than for the conventional optical signal. Subsection 5.D shows calculations of this response function for a few experimental situations.

### C. Magnetic Response Functions

We calculate the magnetic response functions for imaging of a magnetic sample with a magnetization aligned vertically, out of the sample plane. The response functions are evaluated for a horizontal plane 1 nm below the probe. The first two geometries considered here are shown in Fig. 5. Both are with  $p$ -polarized illumination from the right-hand side. The first response function is calculated for detection from the opposite side of the

probe from the illumination, with crossed polarization ( $s$ ). Figure 6 shows the response function for this situation. The sharp peak due to the overlap of the singularity with the higher-order modes can be seen in the response function, but it is not significantly stronger than the broad field around it. The width of the function is of the order of a few hundred nanometers for the probe-sample separation.

The second geometry uses the same illumination source, but with detection in a perpendicular direction, where the field component is polarized vertically, as illustrated in Fig. 5. Although not shown for the sake of brevity, the response function is of a similar width to that obtained in the first case.

These response functions for two different geometries show the same qualitative features: a broad function with a width of several hundred nanometers and no strong central peak. To illustrate their use, we have simulated magneto-optical images of an artificial cobalt sample, with  $\varepsilon_1 = -12.3 + i18.4$  and  $ifM = -0.4 - i0.1$ . For simplicity, we take a sample with no lower surface, i.e., a semi-infinite slab. The sample geometry is depicted in Fig. 7. It has been magnetically modified to contain three stripe domains with vertical magnetizations, of widths 180, 140, and 180 nm, respectively. Elsewhere, the magnetization is taken to be zero. The do-

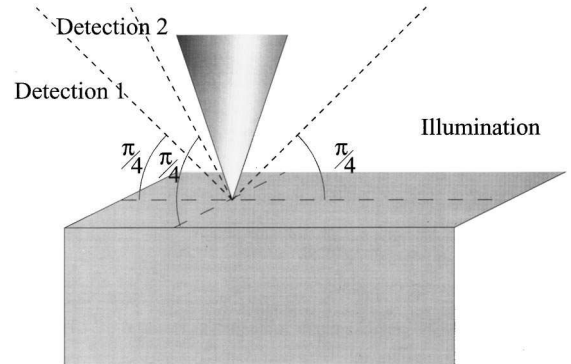


Fig. 5. Illumination and detection geometries for which magneto-optical impulse response functions have been calculated.

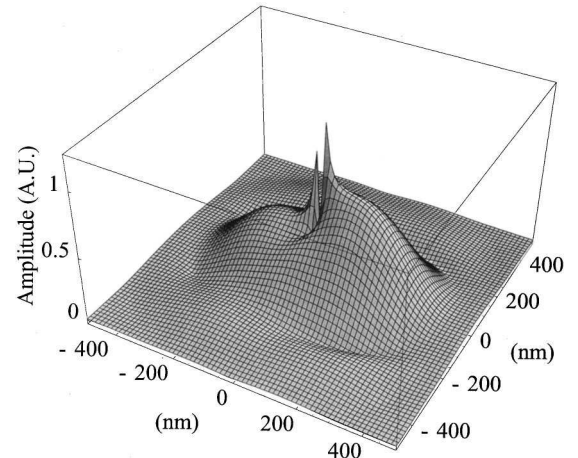


Fig. 6. Response function  $H_{\text{mag}}$  for detection and illumination on opposite sides of the probe, with crossed polarizations.

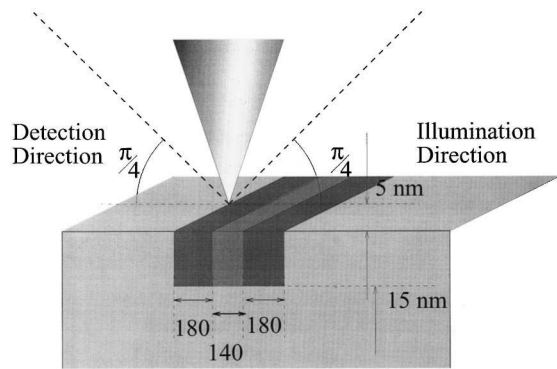


Fig. 7. Magnetic domain structure imaged in Figs. 8 and 9.

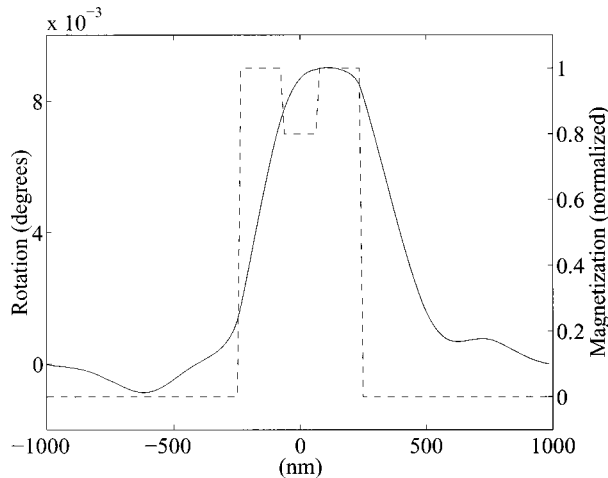


Fig. 8. Calculated image of the magnetic domain structure shown in Fig. 7, as measured through the rotation angle of the field at the detector relative to its direction in the absence of magnetization.

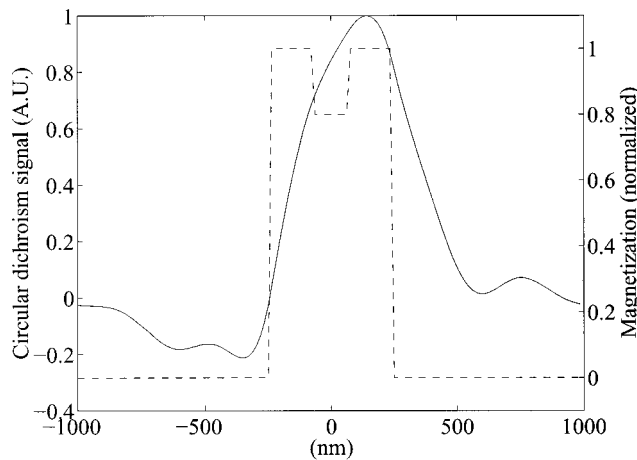


Fig. 9. Calculated image of the magnetic domain structure shown in Fig. 7, as measured by using circular dichroism as the imaging mechanism. The difference between intensities at the detector when using right and left circularly polarized illumination is given. These intensities have been calculated in the absence of a background at the detector.

mains lie at the surface and extend to a depth of 10 nm. This is a simplified representation of a thin magnetic film. We calculate rotation of the electric field at the detector as a function of probe position for a distance of 5 nm between

the probe and the upper surface of the sample. The resulting image is shown in Fig. 8.

This image has been calculated for *p*-polarized illumination from within the plane of the scan, at an angle of  $\pi/4$  to the vertical probe axis, and for detection from the symmetrically opposite position. The plot shows the angle of rotation of the electric field at the detector as a function of probe position during a constant-height scan across the domains. The rotation angle has been calculated from the complex field amplitude at the detector by using Eq. (12). Two comments can be made: First, that the shape of the structure seen in the rotation of the field bears little resemblance to the actual domain structure in the sample, and second, that the resolution in the image is very poor, of the order of a few hundred nanometers. The central domain, with a weaker magnetization, is not seen.

An image has also been calculated for the same sample by using circular dichroism as the imaging mechanism. Here the intensity at the detector has been calculated for both left and right circularly polarized illumination, and the difference in intensities is given as the signal, as shown in Eq. (A14). The result is shown in Fig. 9. It is of interest to note that the form of the measured profile is qualitatively similar to that obtained by measuring the field rotation at the detector but that there are nonetheless clear differences between the two signals. This underlines the fact that it is important to take not only the probe, but also the mode of detection, into account when calculating a SNOM image.

These results show that even if a sample does contain a nanometric domain wall or domain structure, it will be unresolvable with an apertureless near-field optical microscope and a metallic probe. The smallest resolvable structure in the image will be of the order of several hundred nanometers in width. This is a problem intrinsic to the response of the probe and will be the case for any magnetic sample.

Let us now look at the signal that will be recorded in an aperture experiment.

## 6. APPLICATION TO APERTURE SCANNING NEAR-FIELD OPTICAL MICROSCOPY

In this section, the formalism of Section 2 will be applied to aperture SNOM magneto-optical experiments. The example of an illumination-mode experiment will be given, as this is probably the more commonly used geometry, but the results are easily generalized to collection-mode or illumination-collection-mode experiments. An illustration of the experiment is given in Fig. 10(a). A source (depicted as being within the probe fiber) produces a field that is emitted from the probe aperture. This field excites currents in the sample, which in their turn produce an electric field, and the whole radiates toward a detector in the far field.

As in the apertureless case, we will use the first Born approximation to determine the field in the sample. This will be the field that would be present in the absence of the sample, the field from the source diffracted by the aperture,  $\mathbf{E}_{\text{probe}}^{(\text{sou})}$ ; i.e.,  $\mathbf{E}_{\text{exp}} = \mathbf{E}_{\text{probe}}^{(\text{sou})}$ .

The reciprocal situation is shown in Fig. 10(b). The sample is removed, and the sample and the probe are illuminated by a source placed at the actual position of the detector. The reciprocal field is the field produced by this dipole source; if the detector is in the far field (usually the case), then the reciprocal field can be approximated by a plane wave; i.e.,  $\mathbf{E}_{\text{rec}} = [k^2 \exp(ikr)/r]\hat{\mathbf{p}}$ .

If we take the reciprocal field to be a  $y$ -polarized plane wave (this corresponds to an analyzer oriented in the  $y$  direction), then the magnetic response function is simply reduced to the  $x$  component of the probe field:

$$H_{\text{mag}} \propto \hat{\mathbf{z}} \cdot [\mathbf{E}_{\text{probe}}^{(\text{sou})} \times \mathbf{E}_{\text{rec}}] = \mathbf{E}_{\text{probe}}^{(\text{sou})} \cdot \hat{\mathbf{x}}. \quad (22)$$

The nonmagnetic response function  $H_\varepsilon$  for this system is simply a function proportional to the  $y$  component of the probe field:

$$H_\varepsilon \propto \mathbf{E}_{\text{probe}}^{(\text{sou})} \cdot \mathbf{E}_{\text{rec}} = \mathbf{E}_{\text{probe}}^{(\text{sou})} \cdot \hat{\mathbf{y}}. \quad (23)$$

No matter what model we use to represent the probe, the response to the magnetization  $\mathbf{M}$  will be the same as the response to the permittivity  $\varepsilon_1$  that would be obtained by detection through an analyzer *aligned* with the illumination polarization. The magneto-optical response of an aperture probe will be the same as its response in a nonmagnetic experiment. This is in sharp contrast to the apertureless case, where the probe properties were drastically different for conventional and magneto-optical SNOM imaging.

The expression in Eq. (22) with Eq. (7) makes it possible to determine an image directly from the distribution of magnetization in the sample, with knowledge only of a component of the electric field distribution emitted by the probe.

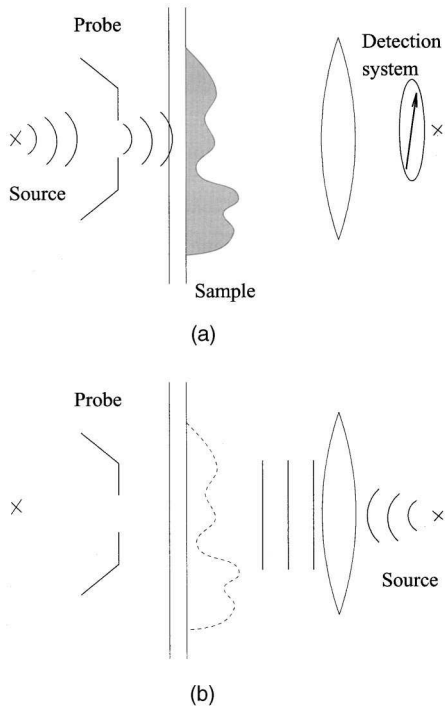


Fig. 10. (a) Experimental geometry of an illumination-mode MO-SNOM and (b) geometry of the reciprocal illumination-mode experiment.

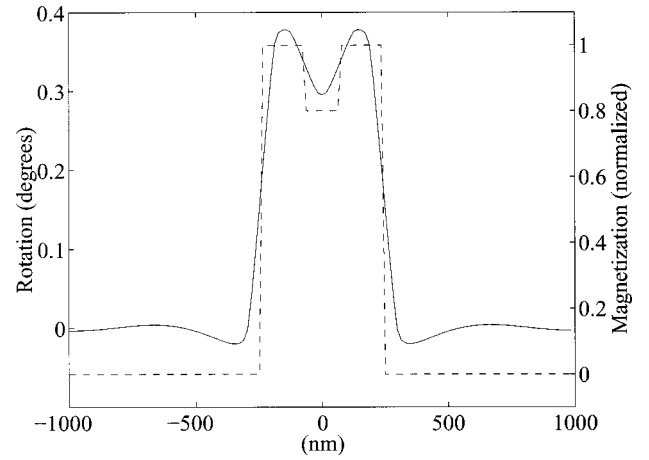


Fig. 11. Calculated image of the magnetic domain structure shown in Fig. 7, recorded with an aperture SNOM of aperture diameter 100 nm.

### A. Model for the Probe

To give an example of the application, we will use the Bethe–Bouwkamp model to simulate the field emitted by the probe, although it is clear that Eqs. (22) and (23) are easily applicable to any probe, provided that it is possible to calculate the emitted field. The Bethe–Bouwkamp model gives the electric field distribution produced by a small circular hole in an infinite conducting screen in the  $z = 0$  plane when illuminated by a polarized plane wave from above.<sup>36</sup> In the case of  $x$ -polarized, normally incident illumination, the field within an aperture of radius  $a$  is

$$E_x = \frac{2a^2 - x^2 - 2y^2}{(a^2 - x^2 - y^2)^{1/2}}, \quad E_y = \frac{xy}{(a^2 - x^2 - y^2)^{1/2}}. \quad (24)$$

The  $z$  component of the field is zero in the aperture, and the  $x$  and  $y$  components are zero outside the aperture.

### B. Response Functions

The response functions  $H_{\text{mag}}$  and  $H_\varepsilon$  for  $x$ -polarized illumination and detection along the  $y$  axis have been calculated for an aperture radius of 50 nm and at a distance of 5 nm from the aperture plane. Although not shown here, both functions have approximately the same width as the probe: in this case, 100 nm.

In this paper, we present for comparison a simulated image of the magnetic sample discussed in Subsection 5.C, using the response function calculated above. The rotation of the field at the detector as a function of probe position is shown in Fig. 11. Contrary to the image obtained with the apertureless microscope, all the domains are now clearly visible in the recorded image. The domain walls are also clearly localized and appear with much greater resolution in the image.

### C. Other Probe Models

The Bethe–Bouwkamp model for aperture near-field probes is a simplified one, which makes it possible to obtain a number of relatively straightforward results analytically. However, in reality, the field emitted by near-field aperture probes may vary from this model. For

example, small defects in the coating of a probe or in the shape of its aperture may lead to significant changes in the distribution of the emitted field.

Interactions between the probe and the substrate can also lead to depolarization of the emitted field,<sup>11,37</sup> which is a serious problem for magneto-optical imaging. These problems have not been dealt with in this paper. We have restricted ourselves to the fundamental demonstration that the significant resolution loss seen in apertureless imaging is not predicted for aperture MO-SNOM.

However, the procedure that has been presented is perfectly well suited to determining the imaging response in any of these more complicated situations. As the sample magnetization is handled as an external current source, the only requirement is to be able to calculate the field that would be present in the absence of magnetization.

The problem of simulating the field emitted by an aperture probe in three dimensions, has been rarely tackled in the past; most work has concentrated on two-dimensional simulations. Novotny *et al.* have calculated the field emitted by probes in both two and three dimensions by using the multiple multipole method.<sup>38–40</sup> These calculations account for the presence of a substrate below the probe and show that the Bethe–Bouwkamp model is no longer a good approximation for this situation. The finite-difference time-domain method has been applied to the study of the emission of an aperture probe above a surface<sup>41–43</sup> and is another technique for determining the field distribution below a probe in the presence of a substrate, even metallic.

With use of the results from models such as these, it is straightforward to calculate the magneto-optical signal as given by Eq. (22). With little additional calculation, the response function for a number of experimental geometries can be easily determined.

## 7. CONCLUSION

This paper has used the electromagnetic theorem of reciprocity to develop field response functions for both the dielectric constant and the magnetization in a sample. The magnetic sample is treated as an external current source rather than a scattering object, and thus the asymmetry of its permittivity tensor does not contradict the fundamental requirements of the theorem of reciprocity. This manner of treating the problem makes it possible to determine a linear response to the magnetization, even when dealing with metallic samples.

The field response functions are related to a reciprocal field, the field that would be present in the absence of magnetization with illumination from the detector. The response of the probe is thus directly taken into account, as is the experimental geometry. In the example of an apertureless magneto-optical experiment, the properties of this field, determined by scattering from the probe, are such that magneto-optical images differ greatly from their conventional optical counterpart.

Because of the existence of a response function for the complete electric field at the detector, it is possible to simulate images that would be obtained with a number of detection techniques. It is possible, for example, to calculate the rotation of the polarization of the field at the

detector for an arbitrary geometry of illumination, sample, and probe. A response function for the field obtained with circularly polarized illumination has also been illustrated. Images of a magnetic sample have been shown by using rotation of polarization at the detector for aperture and apertureless probes and using circular dichroism for an apertureless microscope.

The theory predicts that the best resolution attainable with an apertureless microscope with a metallic probe is 2 orders of magnitude worse in a magneto-optical experiment than in a conventional optical experiment. This is due both to the probe properties and to the asymmetric nature of the permittivity tensor. No such difference is predicted for aperture probe experiments. These predictions are in accordance with experimental observations. If this theory were combined with a numerical technique to evaluate the field below an aperture probe above a substrate, it would be possible to realistically simulate magneto-optical imaging, taking into account multiple scattering between probe and substrate and thus the depolarization effects that occur.

Finally, let us stress that these results indicate that it is essential to consider the properties of the probe when calculating the signal in a SNOM experiment and that a simple calculation of the electric field above the sample is inadequate for determining the signal that will be measured. The framework developed in this paper is easily applicable to any experimental geometry and makes possible a real characterization of the imaging properties of the system.

## APPENDIX A: EXPRESSIONS FOR THE FULL VECTORIAL FIELD AT THE DETECTOR AND OBSERVABLE SIGNALS

Equations (1), (7), and (9) give the projection of the field at the detector along an arbitrary direction  $\hat{\mathbf{p}}$ . Let us consider this direction to be in a transverse plane at the detector; this corresponds to detection of a field polarized within this plane.

We determine the polarization state of the field propagating toward the detector,  $\mathbf{E}_{\text{det}} = \mathbf{E}_{\text{exp}}(\mathbf{r}_{\text{det}})$ , for a fixed incident polarization. For uniformity of notation, we define two mutually orthogonal axes that are also orthogonal to the direction of propagation to the detector:  $\hat{\mathbf{u}}$  and  $\hat{\mathbf{v}}$ . These directions could correspond to (s) and (p) polarizations with respect to the plane of detection, or the  $\hat{\mathbf{x}}$  and  $\hat{\mathbf{y}}$  directions. We can calculate the projection of the field at the detector along either of these directions by using Eqs. (1), (7), and (9).

The complex field  $\mathbf{E}_{\text{det}}$ , projected along each of these vector directions, is found by using the reciprocal fields  $\mathbf{E}_{\text{rec}}^{(u)}$  and  $\mathbf{E}_{\text{rec}}^{(v)}$  created by a unit dipole oriented, respectively, along  $\hat{\mathbf{u}}$  and  $\hat{\mathbf{v}}$ :

$$E_{\text{det},\mathbf{u}} = \frac{1}{i\omega} \int \mathbf{E}_{\text{rec}}^{(u)} \cdot \mathbf{J}_{\text{exp}} dV, \quad (\text{A1})$$

$$E_{\text{det},\mathbf{v}} = \frac{1}{i\omega} \int \mathbf{E}_{\text{rec}}^{(v)} \cdot \mathbf{J}_{\text{exp}} dV, \quad (\text{A2})$$



where  $E_{\text{det},\mathbf{u}} = \mathbf{E}_{\text{det}} \cdot \hat{\mathbf{u}}$  and  $E_{\text{det},\mathbf{v}} = \mathbf{E}_{\text{det}} \cdot \hat{\mathbf{v}}$ . We can now specify the total electric field  $\mathbf{E}_{\text{det}}$ . It is

$$\mathbf{E}_{\text{det}} = \hat{\mathbf{u}}E_{\text{det},\mathbf{u}} + \hat{\mathbf{v}}E_{\text{det},\mathbf{v}}, \quad (\text{A3})$$

or

$$\mathbf{E}_{\text{det}} = \frac{1}{i\omega} \int [\hat{\mathbf{u}}\mathbf{E}_{\text{rec}}^{(u)} + \hat{\mathbf{v}}\mathbf{E}_{\text{rec}}^{(v)}] \cdot \mathbf{J}_{\text{exp}} dV. \quad (\text{A4})$$

The term in brackets is a tensor, not a scalar product. The entire expression could be written more concisely:

$$\mathbf{E}_{\text{det}} = \frac{1}{i\omega} \int \tilde{\mathbf{E}}_{\text{rec}}^{(\mathbf{uv})} \cdot \mathbf{J}_{\text{exp}} dV. \quad (\text{A5})$$

The tensor  $\tilde{\mathbf{E}}_{\text{rec}}^{(\mathbf{uv})}$  is the response function that relates the field  $\mathbf{E}_{\text{det}}$  to the current density  $\mathbf{J}_{\text{exp}}$  and is defined by

$$\tilde{\mathbf{E}}_{\text{rec}}^{(\mathbf{uv})} = (\hat{\mathbf{u}}\mathbf{E}_{\text{rec}}^{(u)} + \hat{\mathbf{v}}\mathbf{E}_{\text{rec}}^{(v)}). \quad (\text{A6})$$

With both field components,  $E_{\text{det},\mathbf{u}}$  and  $E_{\text{det},\mathbf{v}}$ , it is possible to completely characterize the state of polarization of the field at the detector.

### 1. Rotation of Polarization, Ellipticity

The complex field at the detector,

$$\mathbf{E}_{\text{det}} = a\hat{\mathbf{u}} + b \exp(i\delta)\hat{\mathbf{v}}, \quad (\text{A7})$$

with  $\exp(-i\omega t)$  time dependence, traces out an ellipse during each cycle of the wave. By knowing the amplitude of each component,  $a = |E_{\text{det},\mathbf{u}}|$  and  $b = |E_{\text{det},\mathbf{v}}|$ , and their relative phase  $\delta$ , we can determine the orientation of the major axis of the ellipse (the direction of polarization).<sup>29</sup> It is at an angle  $\theta$  with respect to the  $\hat{\mathbf{v}}$  axis, where  $\theta$  is defined by

$$\tan 2\theta = \frac{2ab}{b^2 - a^2} \cos \delta. \quad (\text{A8})$$

The ellipticity, defined as the ratio of minor axis to major axis of the ellipse,

$$\eta = \min|\mathbf{E}|/\max|\mathbf{E}|, \quad (\text{A9})$$

is given by

$$\eta = \frac{a^2 + b^2 - (a^2 - b^2)[1 + (4a^2b^2 \cos^2 \delta)/(a^2 - b^2)^2]^{1/2}}{a^2 + b^2 + (a^2 - b^2)[1 + (4a^2b^2 \cos^2 \delta)/(a^2 - b^2)^2]^{1/2}} \quad (\text{A10})$$

if  $a > b$ . The numerator and the denominator are interchanged for  $b < a$ .

### 2. Circular Dichroism Signal

When circular dichroism is used as a measurement technique, the incident polarization is modulated between left and right circular, while the variation in the signal is detected with a lock-in detector. The formalism presented here makes it possible to calculate the signal obtained with any state of incident polarization. We will write explicit statements for the signal with left and right circularly polarized illumination. As a first approximation, the difference between these signals gives the dichroic signal.

The field at the detector with circularly polarized illumination can be viewed as a superposition of the fields obtained with two orthogonal linearly polarized illumination states.

If we define the currents  $\mathbf{J}_{\text{exp}}^{(s)}$  and  $\mathbf{J}_{\text{exp}}^{(p)}$  as the currents induced in the sample with  $s$ - and  $p$ -polarized illumination, respectively, then the electric field at the detector due to each of these is

$$\begin{aligned} \mathbf{E}_{\text{det}}^{(s)} &= \frac{1}{i\omega} \int \tilde{\mathbf{E}}_{\text{rec}}^{(\mathbf{uv})} \cdot \mathbf{J}_{\text{exp}}^{(s)} dV, \\ \mathbf{E}_{\text{det}}^{(p)} &= \frac{1}{i\omega} \int \tilde{\mathbf{E}}_{\text{rec}}^{(\mathbf{uv})} \cdot \mathbf{J}_{\text{exp}}^{(p)} dV. \end{aligned} \quad (\text{A11})$$

For a geometry where the unit vectors ( $\hat{\mathbf{s}}$ ,  $\hat{\mathbf{p}}$ ,  $\hat{\mathbf{k}}_{\text{inc}}$ ) form a right-handed coordinate system, the circularly polarized basis is given by

$$\hat{\mathbf{R}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{s}} - i\hat{\mathbf{p}}), \quad \hat{\mathbf{L}} = \frac{1}{\sqrt{2}}(\hat{\mathbf{s}} + i\hat{\mathbf{p}}). \quad (\text{A12})$$

The currents induced by right and left circularly polarized illumination are, respectively,

$$\mathbf{J}_{\text{exp}}^{(R)} = \frac{1}{\sqrt{2}}[\mathbf{J}_{\text{exp}}^{(s)} - i\mathbf{J}_{\text{exp}}^{(p)}], \quad \mathbf{J}_{\text{exp}}^{(L)} = \frac{1}{\sqrt{2}}[\mathbf{J}_{\text{exp}}^{(s)} + i\mathbf{J}_{\text{exp}}^{(p)}]. \quad (\text{A13})$$

These currents produce fields  $\mathbf{E}_{\text{det}}^{(R)} = [\mathbf{E}_{\text{det}}^{(s)} - i\mathbf{E}_{\text{det}}^{(p)}]/\sqrt{2}$  and  $\mathbf{E}_{\text{det}}^{(L)} = [\mathbf{E}_{\text{det}}^{(s)} + i\mathbf{E}_{\text{det}}^{(p)}]/\sqrt{2}$  at the detector.

In the absence of a background, the measured intensities are  $|\mathbf{E}_{\text{det}}^{(R)}|^2$  and  $|\mathbf{E}_{\text{det}}^{(L)}|^2$ . The dichroic signal can be approximately represented as

$$I_{\text{dichroic}} = |\mathbf{E}_{\text{det}}^{(R)}|^2 - |\mathbf{E}_{\text{det}}^{(L)}|^2 = 2 \operatorname{Re}[i\mathbf{E}_{\text{det}}^{(s)}\mathbf{E}_{\text{det}}^{(p)*}]. \quad (\text{A14})$$

## APPENDIX B: FULL EXPRESSION OF THE FIELD AT THE CONE APEX

The expressions given here are to be found in Ref. 31. The cone and illumination geometry is illustrated in Fig. 2.

The dependence on cone geometry ( $\theta_1$ ) and coordinates ( $r$ ,  $\theta$ ) has been separated from the dependence on illumination conditions ( $\theta_0$ ,  $\phi_0$ ,  $\beta$ ) and azimuthal coordinate ( $\phi$ ) in Eqs. (19). For notational simplicity, the  $\phi_0$  term is also suppressed in this equation; this amounts to defining the coordinate system so that the source of illumination is above the positive  $x$  axis. The functions  $f$  and  $g$  are given by

$$f(r, \theta, \theta_1, m, p) = \frac{2i}{k \sin \theta_1} \varepsilon_m \frac{2p+1}{p(p+1)} \times \left[ \frac{\exp\left(-\frac{1}{2}ip\pi\right)}{\left(\frac{\partial}{\partial \theta_1}\right) P_p^m(\cos \theta_1) \left(\frac{\partial}{\partial p}\right) P_p^m(\cos \theta_1)} \right] \times j_p(kr) P_p^m(\cos \theta), \quad (\text{B1})$$

$$g(r, \theta, \theta_1, m, p) = \frac{-2i}{k \sin \theta_1} \left(\frac{\varepsilon_0}{\mu_0}\right)^{1/2} \varepsilon_m \frac{2q+1}{q(q+1)} \times \left[ \frac{\exp\left(-\frac{1}{2}iq\pi\right)}{\left[P_q^m(\cos \theta_1) \left(\frac{\partial^2}{\partial q \partial \theta_1}\right) P_q^m(\cos \theta_1)\right]} \right] \times j_q(kr) P_q^m(\cos \theta). \quad (\text{B2})$$

The singular electric field in Eqs. (21), to lowest order, is

$$E_r^{(\text{sing})} = \frac{i \exp\left(-\frac{1}{2}ip_1\pi\right)}{\sin \theta_1} \frac{\sqrt{\pi}(kr)^{p_1-1}}{2^{p_1-1} \Gamma\left(p_1 + \frac{1}{2}\right)} \times \frac{P_{p_1}(\cos \theta) P_{p_1}^1(\cos \theta_0) \sin \beta}{P_{p_1}^1(\cos \theta_1) \frac{\partial}{\partial p_1} P_{p_1}(\cos \theta_1)}, \quad (\text{B3})$$

$$E_\theta^{(\text{sing})} = \frac{i \exp\left(-\frac{1}{2}ip_1\pi\right)}{\sin \theta_1} \frac{\sqrt{\pi}(kr)^{p_1-1}}{2^{p_1-1} \Gamma\left(p_1 + \frac{1}{2}\right)} \times \frac{(1/p_1) \frac{\partial}{\partial \theta} P_{p_1}(\cos \theta) P_{p_1}^1(\cos \theta_0) \sin \beta}{P_{p_1}^1(\cos \theta_1) \frac{\partial}{\partial p_1} P_{p_1}(\cos \theta_1)}. \quad (\text{B4})$$

## ACKNOWLEDGMENTS

J. N. Walford and J. A. Porto gratefully acknowledge financial support from the European Union through Training Mobility and Research Program Near-Field Optics for Nanotechnology contract number ERBFMRXCT980242.

Address correspondence to R. Carminati at the location on the title page or by e-mail, remi@em2c.ecp.fr.

## REFERENCES

1. *Near Field Optics*, D. W. Pohl and D. Courjon, eds., NATO ASI Series (Kluwer Academic, Dordrecht, The Netherlands, 1993).
2. *Optics at the Nanometer Scale*, M. Nieto-Vesperinas and N. García, eds., NATO ASI Series (Kluwer Academic, Dordrecht, The Netherlands, 1996).
3. For a recent review, see *Ultramicroscopy* **202**, Part 1 (April 2001) and Part 2 (May 2001).
4. E. Betzig and R. Chichester, "Single molecules observed by near-field scanning optical microscopy," *Science* **262**, 1422–1425 (1994).
5. E. Betzig and J. K. Trautman, "Near-field optics: microscopy, spectroscopy, and surface modification beyond the diffraction limit," *Science* **257**, 189–195 (1992).
6. B. Hecht, H. Bielefeldt, L. Novotny, Y. Inouye, and D. W. Pohl, "Local excitation, scattering, and interference of surface plasmons," *Phys. Rev. Lett.* **77**, 1889–1892 (1996).
7. S. I. Bozhevolnyi and F. A. Pudonin, "Two-dimensional micro-optics of surface plasmons," *Phys. Rev. Lett.* **78**, 2823–2826 (1997).
8. E. Betzig, J. K. Trautman, R. Wolfe, E. M. Gyorgy, and P. L. Finn, "Near-field magneto-optics and high density data storage," *Appl. Phys. Lett.* **61**, 142–144 (1992).
9. T. J. Silva, S. Schultz, and D. Weller, "Scanning near-field optical microscope for the imaging of magnetic domains in optically opaque materials," *Appl. Phys. Lett.* **65**, 658–660 (1994).
10. F. Matthes, H. Brückl, and G. Reiss, "Near-field magneto-optical microscopy in collection and illumination mode," *Ultramicroscopy* **71**, 243–248 (1998).
11. C. Durkan, I. V. Shvets, and J. C. Lodder, "Observation of magnetic domains using a reflection-mode scanning near-field optical microscope," *Appl. Phys. Lett.* **70**, 1323–1325 (1997).
12. P. Fumagalli, A. Rosenberger, G. Eggers, A. Münnemann, N. Held, and G. Güntherodt, "Quantitative determination of the local Kerr rotation by scanning near-field magneto-optic microscopy," *Appl. Phys. Lett.* **72**, 2803–2805 (1998).
13. B. L. Petersen, A. Bauer, G. Meyer, T. Crecelius, and G. Kaindl, "Kerr-rotation imaging in scanning near-field optical microscopy using a modified Sagnac interferometer," *Appl. Phys. Lett.* **73**, 538–540 (1998).
14. V. I. Safarov, V. A. Kosobukin, C. Hermann, G. Lampel, C. Marlière, and J. Peretti, "Near-field magneto-optics with polarization sensitive STOM," *Ultramicroscopy* **57**, 270–276 (1995).
15. H. Wioland, O. Bergossi, S. Hudlet, K. Mackay, and P. Royer, "Magneto-optical Faraday imaging with an apertureless scanning near-field optical microscope," *EPJ Appl. Phys.* **5**, 289–295 (1999).
16. L. Aigouy, S. Grésillon, A. Lahrech, A. C. Boccara, J. C. Rivoal, V. Mathet, C. Chappert, J. P. Jamet, and J. Ferré, "Apertureless scanning near-field magneto-optical microscopy of magnetic multilayers," *J. Microsc.* **194**, 295–298 (1999).
17. S. Grésillon, H. Cory, J. C. Rivoal, and A. C. Boccara, "Transmission-mode apertureless near-field microscope: optical and magneto-optical studies," *J. Opt. A* **1**, 178–184 (1999).
18. V. Kottler, N. Essaidi, N. Ronarch, C. Chappert, and Y. Chen, "Dichroic imaging of magnetic domains with a scanning near-field optical microscope," *J. Magn. Magn. Mater.* **165**, 398–400 (1997).
19. O. Bergossi, H. Wioland, S. Hudlet, R. Deturche, and P. Royer, "Near field magneto-optical circular dichroism using an apertureless probe," *Jpn. J. Appl. Phys., Part 2* **38**, L655–L658 (1999).
20. Y. Chen, V. Kottler, C. Chappert, and N. Essaidi, "Magneto-optical contrast in near-field optics," *J. Microsc.* **194**, 495–499 (1999).
21. N. Richard, A. Dereux, E. Bourillot, T. David, J. P. Goudonnet, F. Scheurer, and E. Beaurepaire, "Near-field zone analysis of the Faraday rotation of magneto-optical thin films," *J. Appl. Phys.* **88**, 2541–2547 (2000).
22. N. Richard, "Optical properties of multilayer structures," *Eur. Phys. J. B* **17**, 11–14 (2000).
23. D. Van Labeke, A. Vial, and D. Barchiesi, "Near-field theoretical study of a magneto-optical grating," *Ultramicroscopy* **61**, 51–55 (1995).
24. F. Forati, A. Dereux, J.-P. Vigneron, C. Girard, and F. Scheurer, "Theory of Kerr effect in magnetic multilayered structures," *Ultramicroscopy* **61**, 57–62 (1995).

25. J.-J. Greffet and R. Carminati, "Image formation in near-field optics," *Prog. Surf. Sci.* **56**, 133–237 (1997).
26. J. A. Porto, R. Carminati, and J.-J. Greffet, "Theory of electromagnetic field imaging and spectroscopy in scanning near-field optical microscopy," *J. Appl. Phys.* **88**, 4845–4850 (2000).
27. L. Landau, E. Lifchitz, and L. Pitaevskii, *Electrodynamics of Continuous Media* (Pergamon, Oxford, UK, 1984).
28. R. Carminati, M. Nieto-Vesperinas, and J.-J. Greffet, "Reciprocity of evanescent electromagnetic waves," *J. Opt. Soc. Am. A* **15**, 706–712 (1998).
29. J. A. Stratton, *Electromagnetic Theory* (McGraw-Hill, New York, 1941).
30. H. Cory, A. C. Boccara, J. C. Rivoal, and A. Lahrech, "Electric field intensity variation in the vicinity of a perfectly conducting conical probe: application to near-field microscopy," *Microwave Opt. Technol. Lett.* **18**, 120–124 (1998).
31. J. J. Bowman, T. B. A. Senior, and P. L. E. Uslenghi, eds., *Electromagnetic and Acoustic Scattering by Simple Shapes*, rev. ed. (Hemisphere, Levittown, N.Y., 1987).
32. L. Aigouy, F. X. Andreani, A. C. Boccara, J. C. Rivoal, J. A. Porto, R. Carminati, J.-J. Greffet, and R. Megy, "Near-field optical spectroscopy using an incoherent light source," *Appl. Phys. Lett.* **76**, 397–399 (2000).
33. R. Bachelot, P. Gleyzes, and A. C. Boccara, "Reflection mode scanning near-field optical microscopy using an apertureless metallic tip," *Appl. Opt.* **36**, 2160–2170 (1997).
34. Y. Inouye and S. Kawata, "Near-field scanning optical microscopy with a metallic probe tip," *Opt. Lett.* **19**, 159–161 (1994).
35. G. Wurtz, R. Bachelot, and P. Royer, "A reflection mode apertureless scanning near-field optical microscope developed from a commercial scanning probe microscope," *Rev. Sci. Instrum.* **69**, 1735–1743 (1998).
36. C. J. Bouwkamp, "Diffraction theory," *Rep. Prog. Phys.* **17**, 35–100 (1954).
37. J. K. Trautman, E. Betzig, J. S. Weiner, D. J. DiGiovanni, T. D. Harris, F. Hellman, and E. M. Gyorgy, "Image contrast in near-field optics," *J. Appl. Phys.* **71**, 4659–4663 (1992).
38. L. Novotny, D. W. Pohl, and P. Regli, "Light propagation through nanometer-sized structures: the two-dimensional aperture scanning near-field optical microscope," *J. Opt. Soc. Am. A* **11**, 1768–1779 (1994).
39. L. Novotny and D. W. Pohl, "Light propagation in scanning near-field optical microscopy," in *Photons and Local Probes*, O. Marti and R. Möller, eds. (Kluwer Academic, Dordrecht, The Netherlands, 1995), Chap. 2, pp. 21–33.
40. L. Novotny, "Light propagation and light confinement in near-field optics," Ph.D. thesis (Swiss Federal Institute of Technology, Zurich, 1996).
41. D. A. Christensen, "Analysis of near field tip patterns including object interaction using finite-difference time-domain calculations," *Ultramicroscopy* **57**, 189–195 (1995).
42. J. L. Kann, T. D. Milster, F. Froehlich, R. W. Ziolkowski, and J. Judkins, "Numerical analysis of a two-dimensional near-field probe," *Ultramicroscopy* **57**, 251–256 (1995).
43. G. Parent, "Phénomènes transitoires en optique de champ proche: durée de vie de fluorescence et propagation d'impulsions dans des mésostructures," Ph.D. thesis (Université de France-Comté, Besançon, France, 2000).

## Phase properties of the optical near field

R. Carminati\*

*Instituto de Ciencia de Materiales, Consejo Superior de Investigaciones Científicas, Cantoblanco, 28049 Madrid, Spain*

(Received 30 January 1997)

This paper presents a theoretical and numerical study of the phase properties of the optical near field. A model based on the first Rytov approximation for three-dimensional electromagnetic vector fields describes the relationship between the phase variations and both the topographic and optical properties of the scatterer. It is shown that strong polarization effects can lead to subwavelength phase variations around nanometric structures. The conclusions of the model are illustrated by exact numerical calculations. This study should find broad experimental applications in near-field optical interferometric phase measurements. [S1063-651X(97)51705-3]

PACS number(s): 45.25.Fx, 07.79.Fc, 61.16.Ch, 03.80.+r

Optical resolution beyond the Rayleigh (or diffraction) limit can be achieved by detecting the electromagnetic field at subwavelength distance from the object [1]. This has opened new perspectives for light microscopy with the development of scanning near-field optical microscopy (SNOM) [2]. In SNOM, a tip of subwavelength dimension (either illuminating or detecting) is placed at subwavelength distance from the object. The scattering process transfers part of the light energy from the near zone to the far zone. Recording the far-field energy versus the relative tip-sample position provides the image. The key point in this technique is the conversion of evanescent waves into propagating waves, which allows one to overcome the diffraction limit. In order to understand the properties of the optical near field, a lot of work has been concentrated on the description of the light intensity (often assumed to be the square modulus of the electric field) in close proximity of scatterers of arbitrary shape and composition [3]. Light confinement and polarization effects around nanometric structures have been described by different theoretical approaches [4–6] and observed with a photon scanning tunneling microscope [7].

Recently, interferometric measurements have provided a way to record the phase of the near field, in the microwave regime [8], and with visible light [9,10]. These new kinds of near-field optical measurements are promising, since a nanometric resolution was obtained with the set up of Ref. [9]. The first theoretical study of the phase properties in SNOM was presented in Ref. [11]. A scalar model showed that such a resolution was strongly dependent on the sample properties (refractive index and topography). Moreover, this model put forward that the phase of the scattered near field should closely follow the surface profile (in the case of a homogeneous sample). It was stated that phase measurements could represent a breakthrough in SNOM [11].

In this paper, we will study the near-field phase properties with a model based on the Rytov approximation for three-dimensional vector fields [12]. It will be shown that, under certain conditions that are strongly dependent on the polarization of the incident field, the phase of the scattered field closely follows an “equivalent surface profile.” This equiva-

lent profile connects the topographic and dielectric properties of the scatterer. Subwavelength phase variations (“phase confinement”) and polarization effects will be demonstrated in order to put forward the power and the limitations of near-field phase imaging. The conclusions of our model will be illustrated by exact numerical calculations of the near field scattered by two-dimensional structures.

Let us consider a three-dimensional sample consisting of a flat interface separating a vacuum ( $z > 0$ ) from a homogeneous substrate of (frequency dependent) dielectric constant  $\epsilon_s$  ( $z < 0$ ). An inhomogeneous object described by its topographic profile  $z = S(x, y)$  and its (frequency-dependent) dielectric constant  $\epsilon(x, y, z)$  is deposited on the interface. An example of such a sample is shown in Fig. 1. When this system is illuminated by an incident monochromatic field of wavelength  $\lambda$ , the total field for  $z > S(x, y)$  obeys the Lippmann-Schwinger equation [13] (a temporal dependence  $\exp(-i\omega t)$  is assumed for all fields):

$$\mathbf{E}(\mathbf{r}) = \mathbf{E}^{(0)}(\mathbf{r}) + k_0^2 \int [\epsilon(\mathbf{r}') - 1] \vec{\mathbf{G}}(\mathbf{r}_{\parallel} - \mathbf{r}'_{\parallel}, z, z') \mathbf{E}(\mathbf{r}') d^3 \mathbf{r}'. \quad (1)$$

$\vec{\mathbf{G}}$  is the Green dyadic for the system with flat interface at  $z=0$ ,  $\mathbf{E}^{(0)}$  is the field that would exist in this system (i.e., without the object). The integral describes the scattered field

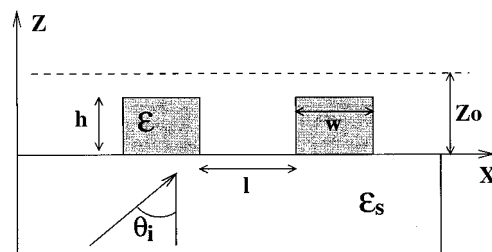


FIG. 1. Example of scattering system. The theoretical model applies to three-dimensional geometries. The system represented here is the one used in the two-dimensional numerical simulations, and is invariant in the  $y$  direction.

\*Electronic address: rcarmina@icmm.csic.es

and is extended to the volume of the object. The notations  $\mathbf{r}=(x,y,z)$ ,  $\mathbf{r}_{\parallel}=(x,y)$ , and  $k_0=\omega/c=2\pi/\lambda$  are used.

When the object has a smooth profile [ $h=\sup|S(x,y)|\ll\lambda$ ] and a low dielectric contrast  $\Delta\epsilon=\epsilon(\mathbf{r})-1$  (this is expected to be the case with most of the sample studied in SNOM), the first Rytov approximation can be used to describe the scattered field [14]. We write the  $\alpha$  component of the *total* field in the form

$$E_{\alpha}(\mathbf{r})=E_{\alpha}^{(0)}(\mathbf{r})\exp[\phi_{\alpha}(\mathbf{r})]\approx E_{\alpha}^{(0)}(\mathbf{r})[1+\phi_{\alpha}^{(1)}(\mathbf{r})]. \quad (2)$$

Note that Eq. (2) implies that each component of the electric field is scattered independently. This means that no energy is transferred from one component to another during the scattering process. This hypothesis is consistent with the weak scattering assumption, and will be confirmed below by the numerical simulations. Equations (1) and (2) lead to (in the first Rytov approximation) [14]:

$$\begin{aligned} \phi_{\alpha}^{(1)}(\mathbf{r}) &= \frac{k_0^2}{E_{\alpha}^{(0)}(\mathbf{r})} \int [\epsilon(\mathbf{r}')-1] G_{\alpha\alpha}(\mathbf{r}_{\parallel}-\mathbf{r}'_{\parallel}, z, z') \\ &\quad \times E_{\alpha}^{(0)}(\mathbf{r}') d^3\mathbf{r}'. \end{aligned} \quad (3)$$

The integral in Eq. (3) corresponds to the first Born approximation for the  $\alpha$  component in Eq. (1). To proceed further, we expand this integral to first order in  $h/\lambda$ . This leads to (see Ref. [15] for more details):

$$\begin{aligned} \phi_{\alpha}^{(1)}(\mathbf{r}) &= \frac{k_0^2}{E_{\alpha}^{(0)}(\mathbf{r})} (\epsilon_s-1) \int G_{\alpha\alpha}(\mathbf{r}_{\parallel}-\mathbf{r}'_{\parallel}, z, 0) \\ &\quad \times E_{\alpha}^{(0)}(\mathbf{r}'_{\parallel}, 0) S_{eq}(\mathbf{r}'_{\parallel}) d^2\mathbf{r}'_{\parallel}, \end{aligned} \quad (4)$$

where the integral is now a surface integral extended to the entire  $x$ - $y$  plane.  $S_{eq}$  is an equivalent surface profile connecting the dielectric constant variation and the topography of the object [15]:

$$S_{eq}(\mathbf{r}_{\parallel})=(\epsilon_s-1)^{-1} \int_0^{S(\mathbf{r}_{\parallel})} [\epsilon(\mathbf{r}_{\parallel}, z)-1] dz. \quad (5)$$

In the case of a homogeneous sample ( $\epsilon=\epsilon_s$ ),  $S_{eq}$  reduces to the true topographic profile.

Equation (4) is our starting point for a discussion of the phase properties in the near field.  $\text{Im}(\phi_{\alpha}^{(1)})$  ( $\text{Im}$  denoting the imaginary part) is the phase difference between the  $\alpha$  component of the *total* field and the  $\alpha$  component of the illuminating field  $\mathbf{E}^{(0)}$ . Equation (4) describes how this phase difference is connected to the properties of the object (the latter being described by  $S_{eq}$ ). The resemblance (or lack of it) between  $\text{Im}(\phi_{\alpha}^{(1)})$  and  $S_{eq}$  strongly depends on the illuminating field and the direction of the  $\alpha$  component of the field with respect to the (eventual) privileged directions of the equivalent surface profile. This leads to *phase* polarization and confinement effects, as those observed in the intensity (i.e.,  $|\mathbf{E}|^2$ ) [4–7]. Let us consider the simple case in which the illuminating field is a transmitted plane wave at normal incidence,  $\mathbf{E}^{(0)}(\mathbf{r})=\mathbf{E}_0\exp(ik_0z)$ . Equation (4) becomes

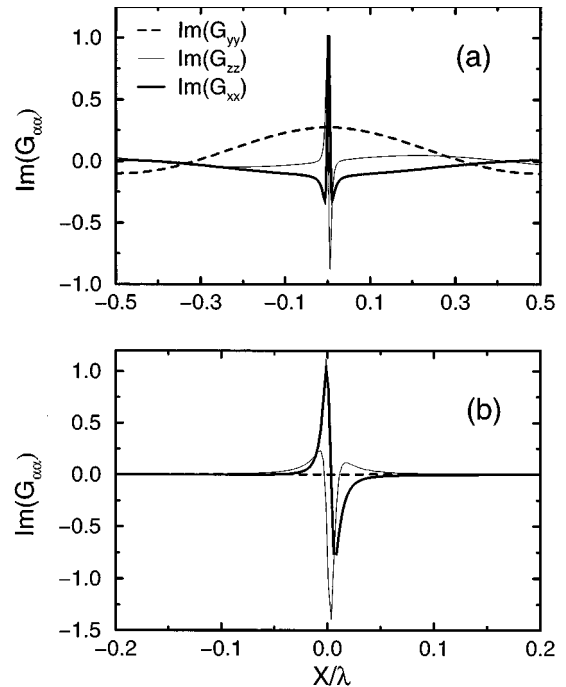


FIG. 2. Imaginary part of the components of the Green dyadic versus  $x$ ,  $y=0$  and  $z_0=6$  nm. (a)  $\theta_i=0^\circ$ . (b)  $\theta_i=50^\circ$  (total internal reflection).

$$\begin{aligned} \phi_{\alpha}^{(1)}(\mathbf{r}_{\parallel}, z_0) &= \frac{k_0^2}{\exp(ik_0z_0)} (\epsilon_s-1) \\ &\quad \times \int G_{\alpha\alpha}(\mathbf{r}_{\parallel}-\mathbf{r}'_{\parallel}, z_0, 0) S_{eq}(\mathbf{r}'_{\parallel}) d^2\mathbf{r}'_{\parallel}, \end{aligned} \quad (6)$$

where we have assumed that the phase was measured in a plane  $z=z_0$ . The relationship between the phase difference  $\text{Im}(\phi_{\alpha}^{(1)})$  and the object properties  $S_{eq}$  is governed by the imaginary part of the components  $G_{\alpha\alpha}$  of the Green dyadic. They are plotted in Fig. 2, versus  $x$ , for  $y=0$  and  $z_0=6$  nm. According to Eq. (6), the convolution of  $\text{Im}(G_{\alpha\alpha})$  by the equivalent surface profile  $S_{eq}$  gives the phase variation. At normal incidence [Fig. 2(a)],  $\text{Im}(G_{xx})$  is sharply peaked around  $x=0$  (and symmetric), so that the phase  $\text{Im}(\phi_x^{(1)})$  will closely follow the equivalent surface profile. Subwavelength phase variations (“phase confinement”) will be observed around the inhomogeneities of the object. On the contrary,  $\text{Im}(G_{yy})$  has a width of about one wavelength, eliminating the possibility of subwavelength resolution with phase imaging. The case of  $G_{zz}$  is not worth being discussed because at normal incidence, the  $z$  component of the scattered field is so weak that a measurement of  $\phi_z^{(1)}$  would not be appropriate. For an illumination in total internal reflection [Fig. 2(b)],  $\text{Im}(G_{zz})$  and  $\text{Im}(G_{xx})$  are peaked around  $x=0$ ,  $\text{Im}(G_{zz})$  being almost symmetric but in contrast reversal. Moreover, as for normal incidence,  $\text{Im}(G_{yy})$  only exhibits suprawavelength variations, with a very low contrast. In summary, Eq. (6) and Fig. 2 demonstrate a very strong polarization effect in near-field phase imaging. They also indicate the circumstances under which the phase variations will follow the *equivalent* surface profile of the object.

In order to check and illustrate the conclusions of the above model, we present exact numerical simulations of the

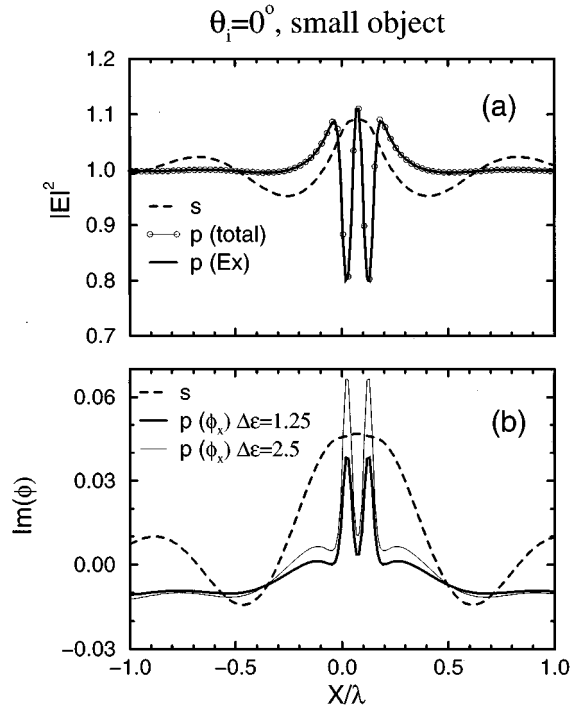


FIG. 3. Numerical calculation of the field along the line  $z_0=40$  nm above the sample in Fig. 1.  $w=l=h=30$  nm.  $\theta_i=0^\circ$ . (a)  $|E_y|^2$  in  $s$  polarization (dashed line),  $|E|^2$  (line with circles) and  $|E_x|^2$  (solid line) in  $p$  polarization. (b)  $\phi_y$  in  $s$  polarization (dashed line),  $\phi_x$  in  $p$  polarization with  $\Delta\epsilon=1.25$  (bold solid line) and  $\Delta\epsilon=2.5$  (thin solid line).

field scattered by the sample in Fig. 1. The numerical scheme consists of solving Eq. (1) using a moment method, without any approximation. This scheme is described in Ref. [16]. For the sake of computer time and memory space, the geometry is two-dimensional (i.e., invariant along  $y$ ). All quantities are calculated along a line at a constant height  $z=z_0$ .

We show in Fig. 3(a) the intensity ( $|E|^2$ ) calculated for  $z_0=40$  nm, in both  $s$  (TE) and  $p$  (TM) polarizations. The structure is homogeneous ( $\epsilon=\epsilon_s=2.25$ ), with  $w=h=l=30$  nm. It is illuminated at normal incidence with a monochromatic plane wave of wavelength  $\lambda=633$  nm. The light intensity is more confined around the structure in  $p$  (in contrast reversal) than in  $s$  polarization, in agreement with calculations previously reported [6]. Moreover, the intensity of the total vector field and that of the  $x$ -component alone are practically identical in  $p$  polarization. This confirms the hypothesis of weak cross-polarization scattering that was made in our model [see Eq. (2)]. At normal incidence, the incident field is polarized in the  $x$  direction, and the total field remains (in a very good approximation) polarized in the same direction.

Figure 3(b) represents the phase  $\text{Im}(\phi_y^{(1)})$  in  $s$  polarization (dashed curve) and  $\text{Im}(\phi_x^{(1)})$  in  $p$  polarization (solid bold curve). According to the model presented previously, the phase follows the object structure in  $p$  polarization [ $G_{xx}$  is implied; see Eq. (6) and Fig. 2(a)], and does not follow the structure in  $s$  polarization ( $G_{yy}$  is implied). In fact, the phase in  $s$  polarization does not exhibit any subwavelength variation. Moreover, in the case of an inhomogeneous

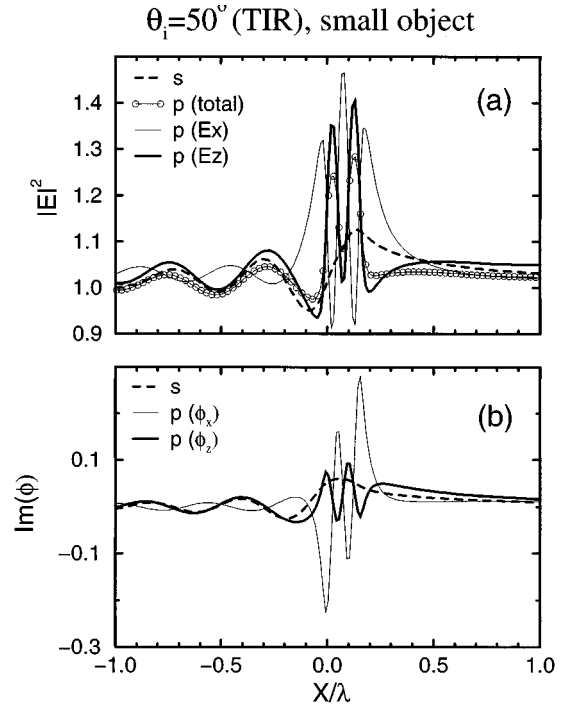


FIG. 4. same as Fig. 3 with  $\theta_i=50^\circ$  (total internal reflection).  $|E_z|^2$  and  $\text{Im}\phi_z^{(1)}$  are also displayed in  $p$  polarization.

sample with dielectric contrast  $\Delta\epsilon=\epsilon-1$  increased by a factor of 2 ( $\epsilon=3.5$ ,  $\epsilon_s=2.25$ ), the phase variation also increases by a factor of 2 (solid thin line). This is in agreement with Eq. (6), which shows that the phase variation is proportional to the *equivalent* surface profile.

We show in Fig. 4 the results for an illumination in total internal reflection. The illuminating field  $\mathbf{E}^{(0)}$  is in this case an evanescent wave, as in photon scanning tunneling microscopy [7,10]. The results for the intensity are plotted in Fig. 4(a). In  $s$  polarization, the situation is unchanged in comparison to the illumination in transmission. In  $p$  polarization, the incident field has two nonvanishing components  $E_x$  and  $E_z$ , and so has the total field. The square modulus of the electric vector field follows more or less the structure, without any contrast reversal. Moreover, the variations of  $|E_x|^2$  and  $|E_z|^2$  clearly demonstrate that this effect mainly stems from the  $z$  component. This was explained theoretically and demonstrated experimentally [5,7]. What is striking is that the same confinement occurs for the phase, as shown in Fig. 4(b). The phase of the  $z$  component of the field in  $p$  polarization follows the lateral variations of the structure (in contrast reversal, according to Fig. 2) with an excellent resolution. Note that the resemblance between the phase variations and the equivalent surface profile is perturbed by the phase of the illuminating field  $E_\alpha^{(0)}$ , which is nonzero at non-normal incidence [see Eq. (4)]. In addition, the phase in  $s$  polarization does not exhibit any subwavelength variation.

In conclusion, we have demonstrated that the near-field phase exhibits polarization and confinement effects, similar to those already known for the intensity. An important result is that, for an illumination at normal incidence, the phase of the parallel component of the total field follows the equivalent surface profile with an excellent resolution. For an inci-

dence in total internal reflection, the phase of the normal component of the total field follows the equivalent surface profile in inverse contrast. We have presented a model, based on the Rytov approximation for electromagnetic vector fields, which contains the essential physics of the phase behavior in the near-field zone. It describes the polarization effects. It also explains how the topographic and dielectric constant variations of the object influence the phase of the near field. This is a very important point in SNOM, where the purely optical properties of the sample are of great interest. The conclusions of our model have been illustrated by exact numerical simulations of the near field scattered by two-dimensional structures of nanometric dimensions.

Finally, we would like to show that the observed polarization and confinement effects are pure near-field effects that are encountered in the scattering by nanometric structures only. Figure 5 shows the phase  $\text{Im}\phi_y^{(1)}$  and  $\text{Im}\phi_x^{(1)}$  in  $s$  and  $p$  polarization, respectively [as in Fig. 3(b)], for the sample in Fig. 1 with  $h=63$  nm,  $w=633$  nm, and  $l=2.5$   $\mu\text{m}$ . It can be seen that the phase in  $s$  and  $p$  polarizations are very similar, both of them following the sample structure. This is precisely the result that is predicted by a scalar description of the field, as in physical optics [17]. Thus, with increasing the structure lateral size up to one wavelength or more, i.e., by

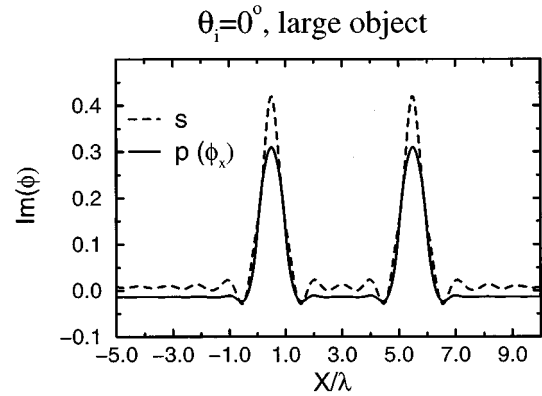


FIG. 5.  $\phi_y$  in  $s$  polarization and  $\phi_x$  in  $p$  polarization for the sample in Fig. 1 with  $w=633$  nm,  $h=63$  nm,  $l=2.5$   $\mu\text{m}$ , and  $z_0=175$  nm.  $\theta_i=0^\circ$ .

reaching the domain of physical optics, the polarization-dependent phase confinement effect disappears.

I would like to thank the EC for financial support. Helpful discussions with N. García, J.-J. Greffet, A. Sentenac, and M. Nieto-Vesperinas are also appreciated.

- 
- [1] E. A. Ash and G. Nicholls, *Nature (London)* **237**, 510 (1972).
  - [2] D. W. Pohl, W. Denk, and M. Lanz, *Appl. Phys. Lett.* **44**, 651 (1984); E. Betzig, M. Isaacson, and A. Lewis, *ibid.* **51**, 2088 (1987).
  - [3] C. Girard and A. Dereux, *Rep. Prog. Phys.* **59**, 657 (1996).
  - [4] D. Van Labeke and D. Barchiesi, *J. Opt. Soc. Am. A* **9**, 732 (1992); **10**, 2193 (1993).
  - [5] O. J. F. Martin, C. Girard, and A. Dereux, *Phys. Rev. Lett.* **74**, 526 (1995); C. Girard, A. Dereux, O. J. F. Martin, and M. Devel, *Phys. Rev. B* **52**, 2889 (1995).
  - [6] J.-J. Greffet and R. Carminati, in *Optics at the Nanometer Scale*, edited by M. Nieto-Vesperinas and N. García (Kluwer, Dordrecht, 1996), p. 1.
  - [7] J. C. Weeber, E. Bourillot, A. Dereux, J. P. Goudonnet, Y. Chen, and C. Girard, *Phys. Rev. Lett.* **77**, 5332 (1996).
  - [8] M. Totzeck and M. A. Krumbügel, *Opt. Commun.* **112**, 189 (1994); F. Keilmann *et al.*, *ibid.* **129**, 15 (1996).
  - [9] F. Zenhausern, Y. Martin, and H. K. Wickramasinghe, *Science* **269**, 1083 (1995).
  - [10] C. Bainier, D. Courjon, F. Baida, and C. Girard, *J. Opt. Soc. Am. A* **13**, 267 (1996).
  - [11] N. García and M. Nieto-Vesperinas, *Appl. Phys. Lett.* **66**, 3399 (1995).
  - [12] The Rytov approximation for scalar fields was used in Ref. [11].
  - [13] A. Poggio and E. Miller, in *Computer Techniques for Electromagnetics*, edited by R. Mittra (Pergamon Press, Oxford, 1973).
  - [14] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991), Chap. 3, p. 112.
  - [15] R. Carminati and J.-J. Greffet, *J. Opt. Soc. Am. A* **11**, 2716 (1995); *Ultramicroscopy* **61**, 43 (1995).
  - [16] F. Pincemin, A. Sentenac, and J.-J. Greffet, *J. Opt. Soc. Am. A* **11**, 1117 (1994); F. Pincemin, A. A. Maradudin, A. D. Boardman, and J.-J. Greffet, *Phys. Rev. B* **50**, 15261 (1994).
  - [17] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968), Chap. 5, p. 77.

# Reciprocity of evanescent electromagnetic waves

Rémi Carminati and Manuel Nieto-Vesperinas

*Instituto de Ciencia de Materiales de Madrid, Consejo Superior de Investigaciones Científicas, Cantoblanco, 28049 Madrid, Spain*

Jean-Jacques Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

Received May 27, 1997; revised manuscript received October 22, 1997; accepted October 27, 1997

We derive reciprocity relations for the generalized reflection and transmission coefficients of vector wave fields containing evanescent components. This is done by using Lorentz's reciprocity theorem with sources at finite distance from the scatterer. © 1998 Optical Society of America [S0740-3232(98)02103-6]

OCIS code: 260.2110.

## 1. INTRODUCTION

Reciprocity is a well-known property of wave propagation and scattering, whose first formulation is often attributed to Helmholtz.<sup>1</sup> In modern scattering theory, this property is expressed by the symmetry of the scattering (or  $S$ ) matrix. The  $S$  matrix was introduced to relate the far-field amplitudes of the incoming and outgoing states in quantum potential scattering<sup>2,3</sup> and in acoustic<sup>4</sup> and electromagnetic<sup>5</sup> scattering. In the  $S$ -matrix formalism, reciprocity appears as an *asymptotic* property of wave fields.<sup>2-6</sup> It is also well known that the Green function of the wave equation in the presence of a scatterer satisfies reciprocity.<sup>7</sup> This property holds whatever the distance between the scatterer and the observation points. Nevertheless, the role of the evanescent waves does not come out in a simple manner from this formalism. The main reason is that the Green function connects a field distribution to a source distribution, and not two field distributions as the  $S$  matrix does.

More recently, the angular spectrum representation of scalar wave fields led to the introduction of a partitioned  $S$  matrix, whose elements have the meaning of generalized transmission and reflection coefficients.<sup>8,9</sup> In this representation one chooses an arbitrary  $z$  direction and separates the entire space into two half-spaces  $\mathcal{R}^-$  ( $z < 0$ ) and  $\mathcal{R}^+$  ( $z > L$ ), the scatterer being included in the strip  $0 < z < L$  (see Fig. 1). In general, the fields in  $\mathcal{R}^-$  and in  $\mathcal{R}^+$  have an angular spectrum containing both homogeneous and evanescent waves. Nevertheless, reciprocity relations were obtained for the generalized transmission and reflection coefficients corresponding to *homogeneous waves only*.<sup>8-10</sup> The reason is the following: The derivations of reciprocity relations in electromagnetic scattering, either for the  $S$  matrix,<sup>5</sup> for plane-wave scattering,<sup>6</sup> or for the generalized transmission and reflection coefficients,<sup>8-10</sup> have in common the use of an integral theorem that is due to Lorentz.<sup>11</sup> In Refs. 9 and 10, Lorentz's theorem was applied *without sources* at finite distance from the scatterer. This implies dealing

with *source-free* fields,<sup>12</sup> namely, fields without evanescent components. Hence, whether or not the evanescent waves obey reciprocity is still an open issue.

In recent years the question of reciprocity of the evanescent components has acquired a renewed interest in optics. For example, reciprocity is implied in the surface-plasmon polariton mechanism, which leads to enhanced backscattering on weakly corrugated metallic rough surfaces,<sup>13</sup> and in the conversion of evanescent waves into propagating waves in the illumination and detection processes in near-field optical microscopy.<sup>14</sup>

To answer this question, we propose to follow the procedure of Ref. 10 but with a different starting point. We shall use Lorentz's reciprocity theorem *with sources* at finite distance from the scatterer. The derivation of this theorem can be found in Ref. 15. For reasons of comprehensiveness, and because the two forms of the theorem (with and without sources) are not clearly related in the literature, the derivation is reproduced in Appendix A of the present paper. With this procedure we shall demonstrate that reciprocity of the generalized transmission and reflection coefficients also holds for the evanescent components of the angular spectrum. Moreover, we shall consider vector fields, for which the generalized reflection and transmission coefficients are tensor operators. In view of this result, we shall also discuss the relationship between reciprocity and time reversibility, often addressed in the literature (see, e.g., Ref. 16).

## 2. GENERALIZED TRANSMISSION AND REFLECTION TENSOR COEFFICIENTS

In this section we define the generalized transmission and reflection tensor coefficients. These were introduced for scalar fields in Ref. 8.

Let us consider a scatterer of arbitrary shape and composition, made up of one or several bodies. Its properties are assumed to be linear and local. They are defined by the (frequency- and position-dependent) complex constitu-



tive tensors  $\vec{\epsilon}(\mathbf{r}, \omega)$  and  $\vec{\mu}(\mathbf{r}, \omega)$ .  $\vec{\epsilon}$  is a generalized dielectric tensor containing both the dielectric and the metallic response of the material.  $\vec{\mu}$  is the magnetic permeability tensor. The only restriction that we impose on  $\vec{\epsilon}$  and  $\vec{\mu}$  is their *symmetry*, which is a necessary condition for reciprocity (see Appendix A). To analyze the contribution of the evanescent waves in the reciprocity relations, we introduce the angular spectrum of the fields and thus choose an arbitrary  $z$  direction (see Fig. 1). We assume that the scatterer is situated within the strip  $0 < z < L$ , and we call  $\mathcal{R}^-$  and  $\mathcal{R}^+$  the half-spaces  $z < 0$  and  $z > L$ , respectively.

In a first situation, let  $\mathbf{E}_1^i$  be a monochromatic field of frequency  $\omega$  propagating toward  $z > 0$  in  $\mathcal{R}^-$ . We assume that this field is emitted by sources situated in the region  $z < z_1$  at *finite distance* from the scatterer (see situation 1 in Fig. 1). At any point  $\mathbf{r} = (\mathbf{R}, z)$  with  $z_1 < z < 0$ , one can write<sup>9</sup> [a temporal dependence  $\exp(-i\omega t)$  is assumed for all fields]

$$\mathbf{E}_1^i(\mathbf{r}) = \int \mathbf{e}_1^i(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} \quad \text{for } z_1 < z < 0, \quad (1)$$

where  $\mathbf{e}_1^i(\mathbf{K})$  is the angular spectrum of  $\mathbf{E}_1^i(\mathbf{r})$  and

$$\gamma(\mathbf{K}) = \sqrt{k^2 - |\mathbf{K}|^2} \quad \text{for } |\mathbf{K}| \leq k \text{ (homogeneous components),} \quad (2)$$

$$\gamma(\mathbf{K}) = i\sqrt{|\mathbf{K}|^2 - k^2} \quad \text{for } |\mathbf{K}| > k \text{ (evanescent components),} \quad (3)$$

with  $k = \omega/c$ ,  $c$  being the speed of light in vacuum. All integrals in this paper are extended to  $-\infty < K_x < +\infty$  and  $-\infty < K_y < +\infty$ , and we use the notation  $d^2\mathbf{K} = dK_x dK_y$ .  $\mathbf{e}_1^i(\mathbf{K})$  in Eq. (1) is assumed to decay as  $|\mathbf{K}|$  increases, in such a way that the integral converges. We shall see below that this is the case when the incident

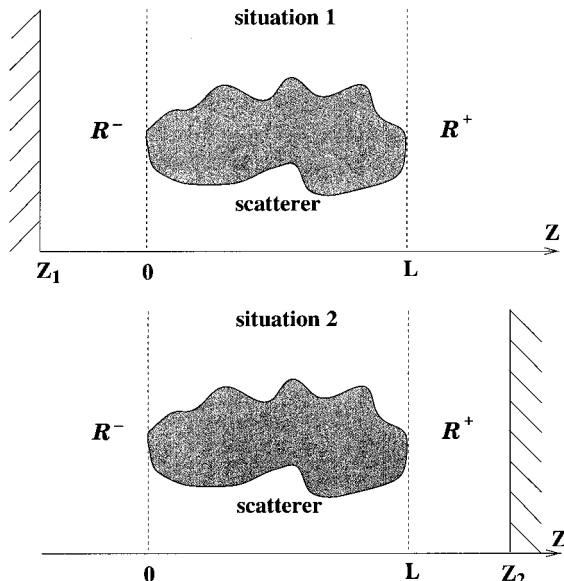


Fig. 1. Scattering geometry and definition of the half-spaces  $\mathcal{R}^-$  and  $\mathcal{R}^+$ .

field is created by a dipole source. The angular spectrum  $\mathbf{e}_1^i(\mathbf{K})$  is proportional to  $\exp(-i\gamma z_1)$ , ensuring the convergence of the integral.

Upon interaction with the scatterer, a transmitted field  $\mathbf{E}_1^t$  and a reflected field  $\mathbf{E}_1^r$  are created in  $\mathcal{R}^+$  and  $\mathcal{R}^-$ , respectively. They can be represented by their angular spectrum:

$$\mathbf{E}_1^t(\mathbf{r}) = \int \mathbf{e}_1^t(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} \quad \text{for } z > L, \quad (4)$$

$$\mathbf{E}_1^r(\mathbf{r}) = \int \mathbf{e}_1^r(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K} \quad \text{for } z_1 < z < 0. \quad (5)$$

Because the scatterer is linear, the angular spectra  $\mathbf{e}_1^t$  and  $\mathbf{e}_1^r$  are related to  $\mathbf{e}_1^i$  by<sup>8-10</sup>

$$\mathbf{e}_1^t(\mathbf{K}) = \int \vec{\mathbf{t}}(\mathbf{K}, \mathbf{K}') \cdot \mathbf{e}_1^i(\mathbf{K}') d^2\mathbf{K}', \quad (6)$$

$$\mathbf{e}_1^r(\mathbf{K}) = \int \vec{\mathbf{r}}(\mathbf{K}, \mathbf{K}') \cdot \mathbf{e}_1^i(\mathbf{K}') d^2\mathbf{K}', \quad (7)$$

where  $\vec{\mathbf{t}}$  and  $\vec{\mathbf{r}}$  are the generalized transmission and reflection tensor coefficients, respectively, for an incident field in  $\mathcal{R}^-$ .

In a second situation, let  $\mathbf{E}_2^i$  be a monochromatic field of frequency  $\omega$  propagating toward  $z < 0$  in  $\mathcal{R}^+$ . We assume that this field is emitted by sources situated in the region  $z < z_2$  at *finite distance* from the scatterer (see situation 2 in Fig. 1). At any point  $\mathbf{r} = (\mathbf{R}, z)$  with  $L < z < z_2$ , one can write

$$\mathbf{E}_2^i(\mathbf{r}) = \int \mathbf{e}_2^i(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K} \quad \text{for } L < z < z_2. \quad (8)$$

As in the previous situation,  $\mathbf{e}_2^i(\mathbf{K})$  in Eq. (8) is assumed to decay as  $|\mathbf{K}|$  increases, to ensure the convergence of the integral. We shall see below that when the incident field is created by a dipole source,  $\mathbf{e}_2^i(\mathbf{K})$  is proportional to  $\exp(i\gamma z_2)$ .

Let us call  $\mathbf{E}_2^t$  and  $\mathbf{E}_2^r$  the transmitted and reflected fields created in  $\mathcal{R}^+$  and  $\mathcal{R}^-$ , respectively. One can write

$$\mathbf{E}_2^t(\mathbf{r}) = \int \mathbf{e}_2^t(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K} \quad \text{for } z < 0, \quad (9)$$

$$\mathbf{E}_2^r(\mathbf{r}) = \int \mathbf{e}_2^r(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} \quad \text{for } L < z < z_2. \quad (10)$$

Similarly, one has

$$\mathbf{e}_2^t(\mathbf{K}) = \int \vec{\mathbf{\tau}}(\mathbf{K}, \mathbf{K}') \cdot \mathbf{e}_2^i(\mathbf{K}') d^2\mathbf{K}', \quad (11)$$

$$\mathbf{e}_2^r(\mathbf{K}) = \int \vec{\mathbf{\rho}}(\mathbf{K}, \mathbf{K}') \cdot \mathbf{e}_2^i(\mathbf{K}') d^2\mathbf{K}', \quad (12)$$

where  $\vec{\mathbf{\tau}}$  and  $\vec{\mathbf{\rho}}$  are the generalized transmission and reflection tensor coefficients, respectively, for an incident field in  $\mathcal{R}^+$ .

Note that the fields evaluated in vacuum satisfy  $\nabla \cdot \mathbf{E} = 0$ , so that their angular spectrum is transverse, i.e.,  $\mathbf{e}(\mathbf{K}) \cdot \mathbf{k} = 0$  with  $\mathbf{k} = (\mathbf{K}, \gamma)$ . Thus  $\mathbf{e}(\mathbf{K})$  has only two independent components, and the four generalized tensor coefficients have only four independent components.

We shall now derive reciprocity relations for the four generalized tensor coefficients, starting from Lorentz's reciprocity theorem with sources at finite distance from the scatterer. These relations are valid for both homogeneous and evanescent components of the field angular spectrum.

### 3. RECIPROCITY RELATIONS

#### A. Transmission

To illustrate our arguments, let us consider a dipole  $\mathbf{p}_1$  situated at a point  $\mathbf{r}_1$  in  $\mathcal{R}^-$  (situation 1 in Fig. 2). For  $z_1 < z < 0$  this point source radiates a field propagating toward  $z > 0$  (incident field),<sup>17</sup> given by Eq. (1), with the angular spectrum<sup>18</sup>

$$\mathbf{e}_1^i(\mathbf{K}) = \frac{i\mu_0\omega^2}{8\pi^2\gamma} \vec{\mathcal{T}}(\mathbf{K}) \cdot \mathbf{p}_1 \exp(-i\mathbf{K} \cdot \mathbf{R}_1 - i\gamma z_1), \quad (13)$$

where  $\gamma = \gamma(\mathbf{K})$ .

The operator  $\vec{\mathcal{T}}$  is the projection on the direction transverse to the  $\mathbf{k}$  vector:

$$\vec{\mathcal{T}}(\mathbf{K}) = \vec{\mathbf{I}} - \frac{\mathbf{k} \otimes \mathbf{k}}{k^2}, \quad (14)$$

where  $\otimes$  is a dyadic product, namely,  $(\mathbf{k} \otimes \mathbf{k})_{ij} = k_i k_j$ , and  $\vec{\mathbf{I}}$  denotes the unit tensor. Note that when  $\vec{\mathcal{T}}$  acts on

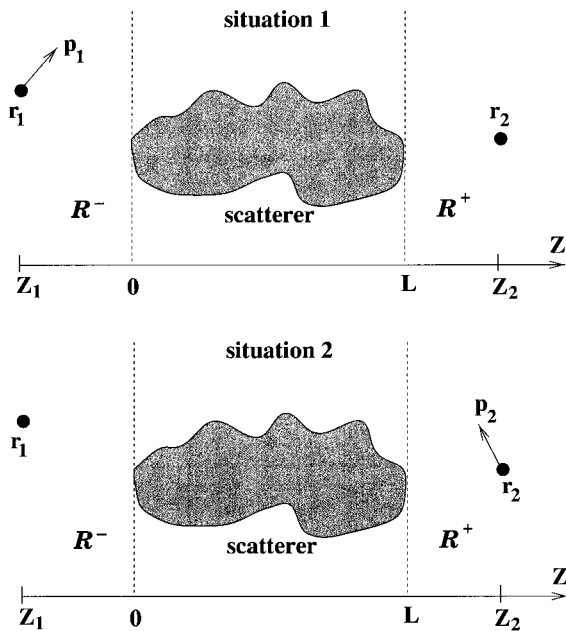


Fig. 2. Geometry considered in the demonstration of the reciprocity of the generalized transmission coefficient. Situation 1: The dipole source  $\mathbf{p}_1$  is in  $\mathcal{R}^-$ , and the scattered (transmitted) field is evaluated in  $\mathcal{R}^+$ . Situation 2: The dipole source  $\mathbf{p}_2$  is in  $\mathcal{R}^+$ , and the scattered (transmitted) field is evaluated in  $\mathcal{R}^-$ .

a transverse field (i.e., a field perpendicular to  $\mathbf{k}$ ),  $\vec{\mathcal{T}}$  becomes the unit tensor.<sup>9</sup> This property will be useful below.

Introducing Eq. (13) into Eq. (6), and the resulting equation into Eq. (4), leads to the expression of the transmitted field in  $\mathcal{R}^+$  in the presence of the scatterer:

$$\begin{aligned} \mathbf{E}_1^t(\mathbf{r}) &= \frac{i\mu_0\omega^2}{8\pi^2} \int d^2\mathbf{K} \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) \\ &\times \int d^2\mathbf{K}' \vec{\mathbf{t}}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{T}}(\mathbf{K}') \cdot \frac{\mathbf{p}_1}{\gamma'} \\ &\times \exp(-i\mathbf{K}' \cdot \mathbf{R}_1 - i\gamma' z_1), \end{aligned} \quad (15)$$

where  $\gamma' = \gamma(\mathbf{K}')$ .

Let us now consider a dipole  $\mathbf{p}_2$  situated at a point  $\mathbf{r}_2$  in  $\mathcal{R}^+$  (situation 2 in Fig. 2). For  $L < z < z_2$  this point source radiates a field propagating toward  $z < 0$  (incident field), given by Eq. (8), with the angular spectrum

$$\mathbf{e}_2^i(\mathbf{K}) = \frac{i\mu_0\omega^2}{8\pi^2\gamma} \vec{\mathcal{T}}(\mathbf{K}) \cdot \mathbf{p}_2 \exp(-i\mathbf{K} \cdot \mathbf{R}_2 + i\gamma z_2). \quad (16)$$

According to Eqs. (9), (11), and (16), the transmitted field in  $\mathcal{R}^-$  in the presence of the scatterer is given by

$$\begin{aligned} \mathbf{E}_2^t(\mathbf{r}) &= \frac{i\mu_0\omega^2}{8\pi^2} \int d^2\mathbf{K} \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) \\ &\times \int d^2\mathbf{K}' \vec{\tau}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{T}}(\mathbf{K}') \cdot \frac{\mathbf{p}_2}{\gamma'} \\ &\times \exp(-i\mathbf{K}' \cdot \mathbf{R}_2 + i\gamma' z_2). \end{aligned} \quad (17)$$

The application of Lorentz's reciprocity theorem with sources leads to [see Eq. (A11) of Appendix A]

$$\mathbf{p}_1 \cdot \mathbf{E}_2^t(\mathbf{r}_1) = \mathbf{p}_2 \cdot \mathbf{E}_1^t(\mathbf{r}_2). \quad (18)$$

Introducing Eqs. (15) and (17) into Eq. (18) yields, after a little algebra,

$$\gamma' [\vec{\tau}(-\mathbf{K}', -\mathbf{K}) \cdot \vec{\mathcal{T}}(-\mathbf{K})]^T = \gamma \vec{\mathbf{t}}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{T}}(\mathbf{K}'), \quad (19)$$

where the superscript T denotes the transposed tensor.

The generalized tensor coefficients  $\vec{\mathbf{t}}$  and  $\vec{\tau}$  are defined by their action on angular spectra  $\mathbf{e}$  that are transverse fields. For such fields the operator  $\vec{\mathcal{T}}$  is the unit tensor. Thus one always has  $\vec{\mathbf{t}}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{T}}(\mathbf{K}') \cdot \mathbf{e}(\mathbf{K}') = \vec{\mathbf{t}}(\mathbf{K}, \mathbf{K}') \cdot \mathbf{e}(\mathbf{K}')$ . This means that the tensors  $\vec{\mathbf{t}} \cdot \vec{\mathcal{T}}$  and  $\vec{\mathbf{t}}$  coincide. The same result holds for  $\vec{\tau}$ . Thus Eq. (19) may be rewritten in the form

$$\gamma' [\vec{\tau}(-\mathbf{K}', -\mathbf{K})]^T = \gamma \vec{\mathbf{t}}(\mathbf{K}, \mathbf{K}'), \quad (20)$$

which is the reciprocity relation for the generalized transmission tensor coefficient.

It should be remarked that the presence of the factors  $\gamma$  and  $\gamma'$  in Eq. (20) is a consequence of the definition of the angular spectra by integration over  $\mathbf{K}$ . For  $|\mathbf{K}| < k$  and  $|\mathbf{K}'| < k$ ,  $\gamma$  and  $\gamma'$  are only directional cosines. Note that with a definition of the angular spectra using angular variables, corresponding to spherical coordinates for the  $\mathbf{k}$  vector, these factors  $\gamma$  and  $\gamma'$  disappear from the

reciprocity relations.<sup>8,9</sup> One performs in this case an integration over the solid angle  $\Omega$ , with  $d\Omega = d^2\mathbf{K}/\gamma$ . Nevertheless, when the angular spectra contain evanescent waves, this representation involves complex angles. Hence a representation in  $\mathbf{K}$  looks more convenient.

### B. Reflection

Reciprocity relations for the generalized reflection tensor coefficients can be derived along similar lines. In the first situation, the dipole  $\mathbf{p}_1$  is situated at a point  $\mathbf{r}_1$  in  $\mathcal{R}^-$  (situation 1 in Fig. 3). The field created by this point source (incident field) is given by

$$\mathbf{E}_1^{i+}(\mathbf{r}) = \int \mathbf{e}_1^{i+}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} \quad \text{for } z_1 < z < 0, \quad (21)$$

with the angular spectrum

$$\mathbf{e}_1^{i+}(\mathbf{K}) = \frac{i\mu_0\omega^2}{8\pi^2\gamma} \vec{\mathcal{F}}(\mathbf{K}) \cdot \mathbf{p}_1 \exp(-i\mathbf{K} \cdot \mathbf{R}_1 - i\gamma z_1), \quad (22)$$

or by

$$\mathbf{E}_1^{i-}(\mathbf{r}) = \int \mathbf{e}_1^{i-}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K} \quad \text{for } z < z_1, \quad (23)$$

with the angular spectrum

$$\mathbf{e}_1^{i-}(\mathbf{K}) = \frac{i\mu_0\omega^2}{8\pi^2\gamma} \vec{\mathcal{F}}(\mathbf{K}) \cdot \mathbf{p}_1 \exp(-i\mathbf{K} \cdot \mathbf{R}_1 + i\gamma z_1). \quad (24)$$

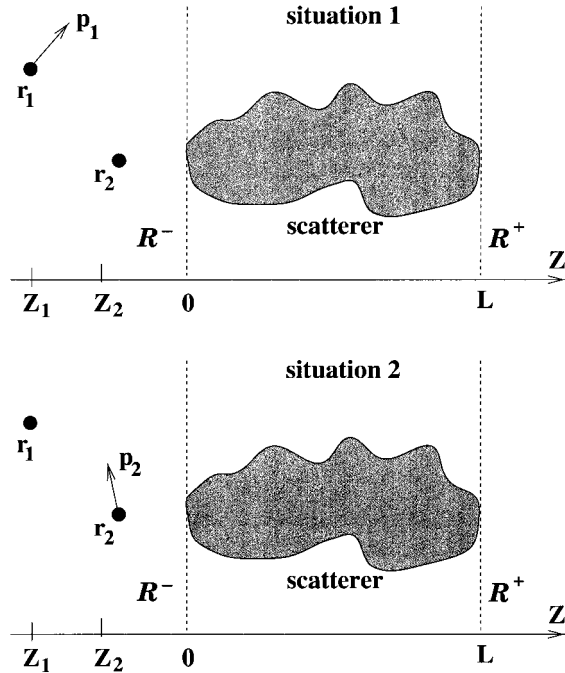


Fig. 3. Geometry considered in the demonstration of the reciprocity of the generalized reflection coefficient. Situation 1: The dipole source  $\mathbf{p}_1$  is in  $\mathcal{R}^-$ , and the scattered (reflected) field is evaluated in  $\mathcal{R}^-$ . Situation 2: The dipole source  $\mathbf{p}_2$  is in  $\mathcal{R}^-$ , and the scattered (reflected) field is evaluated in  $\mathcal{R}^-$ .

According to Eqs. (5), (7), and (22), in the presence of the scatterer, the reflected field in  $\mathcal{R}^-$ , for  $z \neq z_1$ , is given by

$$\begin{aligned} \mathbf{E}_1^r(\mathbf{r}) &= \frac{i\mu_0\omega^2}{8\pi^2} \int d^2\mathbf{K} \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) \\ &\times \int d^2\mathbf{K}' \vec{\mathbf{r}}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{F}}(\mathbf{K}') \cdot \frac{\mathbf{p}_1}{\gamma'} \\ &\times \exp(-i\mathbf{K}' \cdot \mathbf{R}_1 - i\gamma' z_1). \end{aligned} \quad (25)$$

In another situation a dipole  $\mathbf{p}_2$  is situated at a point  $\mathbf{r}_2$  in  $\mathcal{R}^-$  (situation 2 in Fig. 3). The field created by this point source (incident field) is given by

$$\mathbf{E}_2^{i+}(\mathbf{r}) = \int \mathbf{e}_2^{i+}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} \quad \text{for } z_2 < z < 0, \quad (26)$$

with the angular spectrum

$$\mathbf{e}_2^{i+}(\mathbf{K}) = \frac{i\mu_0\omega^2}{8\pi^2\gamma} \vec{\mathcal{F}}(\mathbf{K}) \cdot \mathbf{p}_2 \exp(-i\mathbf{K} \cdot \mathbf{R}_2 - i\gamma z_2), \quad (27)$$

or by

$$\mathbf{E}_2^{i-}(\mathbf{r}) = \int \mathbf{e}_2^{i-}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K} \quad \text{for } z < z_2, \quad (28)$$

with the angular spectrum

$$\mathbf{e}_2^{i-}(\mathbf{K}) = \frac{i\mu_0\omega^2}{8\pi^2\gamma} \vec{\mathcal{F}}(\mathbf{K}) \cdot \mathbf{p}_2 \exp(-i\mathbf{K} \cdot \mathbf{R}_2 + i\gamma z_2). \quad (29)$$

The reflected field in  $\mathcal{R}^-$ , for  $z \neq z_2$ , is given by

$$\begin{aligned} \mathbf{E}_2^r(\mathbf{r}) &= \frac{i\mu_0\omega^2}{8\pi^2} \int d^2\mathbf{K} \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) \\ &\times \int d^2\mathbf{K}' \vec{\mathbf{r}}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{F}}(\mathbf{K}') \cdot \frac{\mathbf{p}_2}{\gamma'} \\ &\times \exp(-i\mathbf{K}' \cdot \mathbf{R}_2 - i\gamma' z_2). \end{aligned} \quad (30)$$

Without any loss of generality, we suppose that  $z_1 < z_2$ . Lorentz's reciprocity theorem with sources reads in this case [see Eq. (A11) of Appendix A] as

$$\mathbf{p}_1 \cdot [\mathbf{E}_2^{i-}(\mathbf{r}_1) + \mathbf{E}_2^r(\mathbf{r}_1)] = \mathbf{p}_2 \cdot [\mathbf{E}_1^{i+}(\mathbf{r}_2) + \mathbf{E}_1^r(\mathbf{r}_2)]. \quad (31)$$

Note that the incident field also satisfies the reciprocity theorem

$$\mathbf{p}_1 \cdot \mathbf{E}_2^{i-}(\mathbf{r}_1) = \mathbf{p}_2 \cdot \mathbf{E}_1^{i+}(\mathbf{r}_2), \quad (32)$$

and thus, from Eq. (31), one finally has

$$\mathbf{p}_1 \cdot \mathbf{E}_2^r(\mathbf{r}_1) = \mathbf{p}_2 \cdot \mathbf{E}_1^r(\mathbf{r}_2). \quad (33)$$

Introducing Eqs. (25) and (30) into Eq. (33) leads, after a little algebra, to

$$\gamma' [\vec{\mathbf{r}}(-\mathbf{K}', -\mathbf{K}) \cdot \vec{\mathcal{F}}(-\mathbf{K})]^T = \gamma \vec{\mathbf{r}}(\mathbf{K}, \mathbf{K}') \cdot \vec{\mathcal{F}}(\mathbf{K}'). \quad (34)$$

Using the same argument as that in Subsection 3.A, one shows that the tensors  $\vec{\mathbf{r}} \cdot \vec{\mathcal{T}}$  and  $\vec{\mathbf{r}}$  coincide, so that Eq. (34) can be rewritten in the form

$$\gamma'[\vec{\mathbf{r}}(-\mathbf{K}', -\mathbf{K})]^T = \gamma\vec{\mathbf{r}}(\mathbf{K}, \mathbf{K}'). \quad (35)$$

Proceeding in a similar way with an incident field propagating toward  $z < 0$  in  $\mathcal{R}^+$ , one can derive

$$\gamma'[\vec{\rho}(-\mathbf{K}', -\mathbf{K})]^T = \gamma\vec{\rho}(\mathbf{K}, \mathbf{K}'). \quad (36)$$

Equations (35) and (36) are the reciprocity relations for the generalized reflection tensor coefficients.

### C. Discussion

Relations (20), (35), and (36) express the reciprocity relations that are valid for any value of  $|\mathbf{K}|$ . They include reciprocity between homogeneous waves ( $|\mathbf{K}| \leq k$  and  $|\mathbf{K}'| \leq k$ ), between evanescent waves ( $|\mathbf{K}| > k$  and  $|\mathbf{K}'| > k$ ), and between one homogeneous wave and one evanescent wave ( $|\mathbf{K}| \leq k$  and  $|\mathbf{K}'| > k$ ). We have thus extended to evanescent waves and to vector fields the reciprocity relations of the partitioned  $S$  matrix introduced in Ref. 8.

This work also brings together two usual formulations of the reciprocity theorem:

1. The symmetry of the  $S$  matrix,<sup>3-5</sup> or of the generalized transmission and reflection coefficients corresponding to homogeneous waves only.<sup>8-10</sup> Both of these formulations involved field amplitudes evaluated either at infinity or far from sources. Lorentz's reciprocity theorem without sources at finite distance from the scatterer was the basis of these formulations.

2. Relation (A11), which involves the global field (and not its angular spectrum), without any restriction on the distance between the scatterer and both the source and the observation point. This formulation is the one used, for instance, in antenna theory.<sup>19,20</sup> It is a consequence of Lorentz's reciprocity theorem applied with sources at finite distance from the scatterer.<sup>15</sup>

Our approach enlarges the generality of both formulations. It shows that each plane wave of the angular spectrum obeys reciprocity (and not only the global field), and this occurs whatever the distance from the source and the observation point to the scatterer. This is a more general statement than the one given by Eq. (A11) in formulation 2, which applies to the global field. This is also an extension of formulation 1 to vector fields containing evanescent components.

## 4. RECIPROCITY AND TIME REVERSIBILITY

When studying time reversibility of a scattering process in wave optics, one has to take care of the behavior of both homogeneous and evanescent waves. For example, in the domain of phase conjugation, the link between time reversal and phase conjugation was discussed a few years ago.<sup>21,22</sup> It was shown that homogeneous and evanescent waves are transformed differently under phase conjugation, and this was presented as a proof of the nonequivalence between phase conjugation and time reversal.<sup>22</sup>

Concerning the reciprocity theorem, a scrutiny of the literature shows that its relationship with time reversibility has been the subject of many discussions.

In quantum *elastic* scattering, the symmetry of the  $S$  matrix can be obtained as a consequence of time-reversal invariance of the process.<sup>3</sup> For instance, in a textbook,<sup>3</sup> it is stated: "In other words, time-reversal invariance implies that the  $S$  matrix is symmetric. The symmetry condition is also called reciprocity." In this case the probability waves undergo no absorption, and the  $S$  matrix is also unitary. When a complex potential is used to describe *inelastic* scattering, the unitarity condition is lost. Because the probability waves undergo absorption, time reversibility is also lost. Nevertheless, the  $S$  matrix remains symmetric, indicating that reciprocity still holds.<sup>23</sup>

Similarly, in electromagnetism, reciprocity and time reversibility are equivalent when only lossless media are involved.<sup>24</sup> In the optics of transparent media, one speaks of optical reversibility rather than reciprocity. Nevertheless, with Lorentz's theorem, reciprocity relations can be derived even for a scatterer with losses (i.e., with complex constitutive tensors).<sup>5,6,9,10,24,25</sup> Once again, reciprocity is conserved, even for a time-irreversible process.

The fact that reciprocity relations hold in irreversible processes is a consequence of general microscopic relations obtained from a statistical approach by Onsager.<sup>26</sup> A discussion of these relations and an application to the derivation of symmetry relations in heat conduction or electronic network theory can be found in papers by Casimir<sup>27</sup> and also in Ref. 24. As an example, reciprocity of the resistance tensor of four-poles was obtained.<sup>24,27</sup>

In conclusion, reciprocity and time reversibility in electromagnetic scattering are not equivalent. Nevertheless, as demonstrated in this paper, both homogeneous and evanescent waves satisfy the same reciprocity relations.

## APPENDIX A: LORENTZ'S RECIPROCITY THEOREM WITH SOURCES

We recall in this appendix the derivation of Lorentz's reciprocity theorem with sources at finite distance from the scatterer.<sup>15</sup>

In a first situation, let  $V_1$  be a source volume with a current density  $\mathbf{J}_1(\mathbf{r})$  radiating at a frequency  $\omega$ . Let us call  $\mathbf{E}_1(\mathbf{r})$  and  $\mathbf{H}_1(\mathbf{r})$  the fields created by this source in the presence of a scatterer described by its constitutive tensors  $\vec{\epsilon}(\mathbf{r}, \omega)$  and  $\vec{\mu}(\mathbf{r}, \omega)$ . In a second situation, let  $V_2$  be a source volume with a current density  $\mathbf{J}_2(\mathbf{r})$  radiating at the same frequency  $\omega$ . Let us call  $\mathbf{E}_2(\mathbf{r})$  and  $\mathbf{H}_2(\mathbf{r})$  the fields created by this source in the presence of the same scatterer (see Fig. 4).

The fields in each situation satisfy Maxwell's equations, that is, with  $k = 1, 2$ ,

$$\nabla \times \mathbf{E}_k = i\omega\mathbf{B}_k, \quad \nabla \times \mathbf{H}_k = \mathbf{J}_k - i\omega\mathbf{D}_k, \quad (A1)$$

together with the constitutive relations

$$\begin{aligned} \mathbf{D}_k(\mathbf{r}) &= \epsilon_0 \vec{\epsilon}(\mathbf{r}, \omega) \cdot \mathbf{E}_k(\mathbf{r}), \\ \mathbf{B}_k(\mathbf{r}) &= \mu_0 \vec{\mu}(\mathbf{r}, \omega) \cdot \mathbf{H}_k(\mathbf{r}). \end{aligned} \quad (A2)$$

From Eqs. (A1) one easily obtains for each point  $\mathbf{r}$  the following equality:

$$\begin{aligned} & (\mathbf{H}_2 \cdot \nabla \times \mathbf{E}_1 - \mathbf{E}_1 \cdot \nabla \times \mathbf{H}_2) \\ & + (\mathbf{E}_2 \cdot \nabla \times \mathbf{H}_1 - \mathbf{H}_1 \cdot \nabla \times \mathbf{E}_2) \\ & = i\omega(\mathbf{B}_1 \cdot \mathbf{H}_2 - \mathbf{H}_1 \cdot \mathbf{B}_2) \\ & - i\omega(\mathbf{D}_1 \cdot \mathbf{E}_2 - \mathbf{E}_1 \cdot \mathbf{D}_2) + \mathbf{J}_1 \cdot \mathbf{E}_2 - \mathbf{J}_2 \cdot \mathbf{E}_1. \end{aligned} \quad (\text{A3})$$

The left-hand side of Eq. (A3) can be rewritten in the form  $\nabla \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1)$ . With the use of Eqs. (A2), one shows that the first two terms on the right-hand side vanish, provided that  $\vec{\epsilon}$  and  $\vec{\mu}$  are *symmetric* tensors:

$$\vec{\epsilon}(\mathbf{r}) = [\vec{\epsilon}(\mathbf{r})]^T, \quad \vec{\mu}(\mathbf{r}) = [\vec{\mu}(\mathbf{r})]^T. \quad (\text{A4})$$

Finally, one obtains for each point  $\mathbf{r}$  the following:

$$\nabla \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) = \mathbf{J}_1 \cdot \mathbf{E}_2 - \mathbf{J}_2 \cdot \mathbf{E}_1. \quad (\text{A5})$$

Equation (A5) with  $\mathbf{J}_1 = \mathbf{J}_2 = 0$  gives the usual form of Lorentz's reciprocity theorem.<sup>11</sup> Thus Eq. (A5) is a generalization of Lorentz's theorem to the case where sources are present, and we shall refer to it as Lorentz's reciprocity theorem with sources.

By integrating Eq. (A5) over all space, we transform the left-hand side into a surface integral over a sphere whose radius tends to infinity. The asymptotic expressions of the fields for  $kr \rightarrow \infty$  in a direction defined by the vector  $\mathbf{k} = (\mathbf{K}, \gamma)$  are<sup>18</sup>

$$\mathbf{E}_k(kr \rightarrow \infty) = 2i\pi\gamma(\mathbf{K})\mathbf{e}_k(\mathbf{K}) \frac{\exp(ikr)}{r}, \quad (\text{A6})$$

$$\mathbf{H}_k(kr \rightarrow \infty) = 2i\pi\gamma(\mathbf{K})\mathbf{h}_k(\mathbf{K}) \frac{\exp(ikr)}{r}, \quad (\text{A7})$$

with, as a consequence of Maxwell's equations,

$$\omega\mu_0\mathbf{h}_k(\mathbf{K}) = \mathbf{k} \times \mathbf{e}_k(\mathbf{K}). \quad (\text{A8})$$

By making use of Eqs. (A6)–(A8), one shows that  $\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1$  vanishes identically in the far field, so

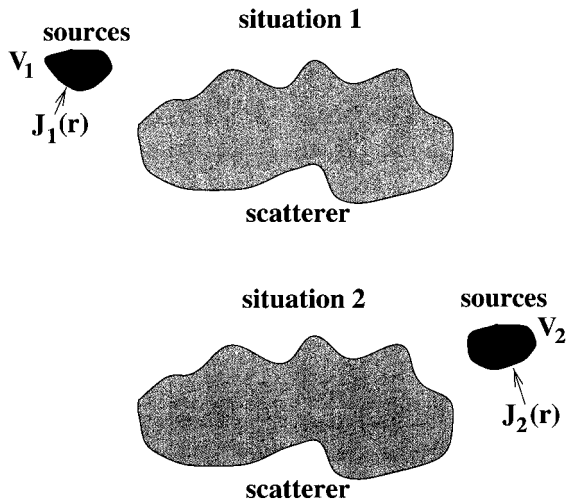


Fig. 4. Geometry considered in the derivation of Lorentz's reciprocity theorem with sources.

that its integral over a sphere of infinite radius disappears. Finally, one is left with

$$\int_{V_1} \mathbf{J}_1(\mathbf{r}) \cdot \mathbf{E}_2(\mathbf{r}) d^3\mathbf{r} = \int_{V_2} \mathbf{J}_2(\mathbf{r}) \cdot \mathbf{E}_1(\mathbf{r}) d^3\mathbf{r}, \quad (\text{A9})$$

which constitutes a reciprocity theorem with sources.<sup>15</sup>

In the case of dipole sources, the current densities are given by

$$\mathbf{J}_k(\mathbf{r}) = -i\omega\mathbf{p}_k\delta(\mathbf{r} - \mathbf{r}_k), \quad (\text{A10})$$

which gives, after introduction into Eq. (A9),

$$\mathbf{p}_1 \cdot \mathbf{E}_2(\mathbf{r}_1) = \mathbf{p}_2 \cdot \mathbf{E}_1(\mathbf{r}_2). \quad (\text{A11})$$

This last equation constitutes a customary statement of the reciprocity theorem, namely, the component of the electric field in the direction of polarization of the source is unchanged when the positions of source and detector are interchanged. This form of the reciprocity theorem is well known, for example, in antenna theory.<sup>19,20</sup> Equations (18) and (31) of the present paper are a direct application of Eq. (A11).

## ACKNOWLEDGMENTS

This paper is dedicated to Emil Wolf on the occasion of his 75th birthday.

R. Carminati acknowledges financial support from the European Union. M. Nieto-Vesperinas is supported by Dirección General de Investigación Científica y Técnica grant PB95-0061.

## REFERENCES AND NOTES

1. H. Helmholtz, *J. Reine Angew. Math.* **67**, 1 (1859); see also Lord Rayleigh, *The Theory of Sound* (Dover, New York, 1945), Sect. 294.
2. P. Roman, *Advanced Quantum Theory* (Addison-Wesley, Reading, Mass., 1965), pp. 282, 294.
3. R. G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), p. 46.
4. E. Gerjuoy and D. S. Saxon, "Variational principles for the acoustic field," *Phys. Rev.* **94**, 1445–1458 (1954).
5. D. S. Saxon, "Tensor scattering matrix for the electromagnetic field," *Phys. Rev.* **100**, 1771–1775 (1955).
6. A. T. de Hoop, "A reciprocity theorem for the electromagnetic field scattered by an obstacle," *Appl. Sci. Res. Sect. B* **8**, 135–140 (1959).
7. Ph. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953), Part I, Sect. 7.5.
8. M. Nieto-Vesperinas and E. Wolf, "Generalized Stokes reciprocity relations for scattering from dielectric objects of arbitrary shape," *J. Opt. Soc. Am. A* **3**, 2038–2046 (1986).
9. M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991), Chap. 5.
10. M. Nieto-Vesperinas, "Reciprocity of the impulse response for scattering from inhomogeneous media and arbitrary dielectric bodies," *J. Opt. Soc. Am. A* **5**, 360–365 (1988).
11. H. A. Lorentz, *Versl. Gewone Vergad. Afd. Natuurkd. K. Ned. Akad. Wet.* **4**, 176–188 (1896); H. A. Lorentz, *Collected Papers* (Nijhoff, Den Haag, The Netherlands, 1936), Vol. III; see also M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991), Chap. 1, p. 8.
12. G. C. Sherman, "Diffracted fields expressible by plane-wave expansions containing only homogeneous waves," *J. Opt. Soc. Am.* **59**, 697–711 (1969).
13. A. A. Maradudin, T. Michel, A. R. McGurn, and E. R.

- Méndez, "Enhanced backscattering of light from a random grating," *Ann. Phys. (N.Y.)* **203**, 255–307 (1990).
14. D. W. Pohl, "Scanning near-field optical microscopy," in *Advances in Optical and Electron Microscopy*, C. J. R. Sheppard and T. Mulvey, eds. (Academic, New York, 1991), pp. 243–311.
  15. L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Electrodynamics of Continuous Media* (Pergamon, Oxford, 1984), Sect. 89.
  16. M. Born, *Natural Philosophy of Cause and Chance* (Dover, New York, 1964), pp. 22–26.
  17. The incident field is here the field created by the dipole  $\mathbf{p}_1$ , where  $\mathbf{p}_1$  has to be understood as the value of the dipole *in the presence of the scatterer*. This corresponds to the usual definition of the incident field in scattering problems. Usually, the source is assumed to be located at infinity and also to be unaffected by the presence of the scatterer.
  18. M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991), Chap. 3, pp. 83–84.
  19. R. W. P. King, *Electromagnetic Engineering* (McGraw-Hill, New York, 1945), Vol. I, p. 311.
  20. D. S. Jones, *The Theory of Electromagnetism* (Pergamon, Oxford, 1964), pp. 59–65.
  21. E. Wolf, G. S. Agarwal, and A. T. Friberg, "Wavefront correction and scattering with phase-conjugated waves," in *Coherence and Quantum Optics V*, L. Mandel and E. Wolf, eds. (Plenum, New York, 1984), pp. 107–116.
  22. M. Nieto-Vesperinas and E. Wolf, "Phase conjugation and symmetries with wave fields in free space containing evanescent components," *J. Opt. Soc. Am. A* **2**, 1429–1434 (1985).
  23. R. J. Glauber, in *Lectures in Theoretical Physics*, W. E. Brittin and L. G. Dunham, eds. (Wiley-Interscience, New York, 1959), Vol. I, p. 315.
  24. P. Grivet, in *Proceedings of the Symposium on Modern Optics*, Vol. XVII of Microwave Research Institute Symposia Series (Wiley-Interscience, New York, 1967), pp. 467–479.
  25. R. Cadilhac, in *Electromagnetic Theory of Gratings*, R. Petit, ed. (Springer-Verlag, New York, 1980), pp. 58–59.
  26. L. Onsager, "Reciprocal relations in irreversible processes," *Phys. Rev.* **37**, 405–426 (1931).
  27. H. B. G. Casimir, "On Onsager's principle of microscopic reversibility," *Rev. Mod. Phys.* **17**, 343–350 (1945); "Reciprocity theorems and irreversible processes," *Proc. IEEE* **51**, 1570–1573 (1963).

## Reciprocity, unitarity, and time-reversal symmetry of the $S$ matrix of fields containing evanescent components

R. Carminati,<sup>1</sup> J. J. Sáenz,<sup>2</sup> J.-J. Greffet,<sup>1</sup> and M. Nieto-Vesperinas<sup>3</sup>

<sup>1</sup>*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

<sup>2</sup>*Departamento de Física de la Materia Condensada and Instituto de Ciencia de Materiales "Nicolás Cabrera," Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain*

<sup>3</sup>*Instituto de Ciencia de Materiales de Madrid, Consejo Superior de Investigaciones Científicas, Cantoblanco, 28049 Madrid, Spain*

(Received 7 February 2000; published 16 June 2000)

We derive general relationships of the  $S$  matrix of fields containing evanescent components. Our formalism covers time-independent quantum scattering as well as scattering of classical scalar waves. We show that reciprocity, energy (or probability) conservation, and time-reversal symmetry in the presence of evanescent waves lead to relationships that extend the well-known relations previously derived in asymptotic scattering. On this basis, we discuss the link between reciprocity and time-reversal symmetry. We also address the experimental feasibility of time reversal of a field containing evanescent components.

PACS number(s): 03.65.Nk, 03.50.-z, 42.25.Fx, 11.55.-m

### I. INTRODUCTION

The scattering matrix ( $S$  matrix) was introduced by Heisenberg to describe a scattering process without any assumption about the details of the interaction [1]. In this formalism, the process is thought of as a transformation of an incoming state  $\Psi_{in}$  into an outgoing state  $\Psi_{out}$ , which describe the system far away from the interaction potential. Hence, the  $S$  matrix describes the scattering process *asymptotically*. The mathematical transcription of this transformation is an operator relationship  $\Psi_{out} = \mathbf{S}\Psi_{in}$ , where  $\mathbf{S}$  is called the  $S$  matrix [2,3]. It is well known that the  $S$  matrix exhibits some properties that are independent of the specific problem under study. In particular, it is unitary and symmetric, these two properties reflecting probability (or energy) conservation in elastic scattering and reciprocity, respectively [3]. The general aim being to get maximum information about the  $S$  matrix with minimum knowledge about the interaction itself, other properties may be derived, based, for example, on dispersion relations and causality conditions [4]. The existence of such general properties of the  $S$  matrix is the reason why it has become a fundamental tool in most areas of theoretical physics, e.g., in quantum scattering [2,3], in particle physics [5], in field theory, and in statistical physics [6]. Its definition and its use have also been extended to scattering of classical (acoustic and electromagnetic) waves [7,8]. For example, the  $S$  matrix has become a fundamental tool (as well as a practical one) to compute scattered fields in physical optics [9]. This formalism has also found a wide range of applications with the development of random-matrix theory [10], which has recently acquired renewed interest through its use in quantum- and classical-wave transport in random media [11,12].

The  $S$  matrix was originally defined as an operator acting on *asymptotic* states. In scattering by a time-independent potential (we shall restrict our discussion to this case), this means that the  $S$  matrix relates the far-field amplitudes of the incoming and outgoing fields [2,7,8]. Nevertheless, in the last ten years, new effects have been observed and new tech-

niques developed that involve scattering and/or direct measurement of evanescent waves (near fields). For example, evanescent-wave scattering is involved in the emission process of an atom or a molecule close to a surface [13] or in the surface-plasmon-polariton (SPP) mechanism which leads to the enhanced backscattering of light on slightly rough metal surfaces [14]. It is also the basic principle of near-field scanning probe microscopies, using either electron [15] or photon [16] tunneling. Modeling the image formation in scanning near-field optical microscopy (SNOM) requires a precise description of a mechanism involving scattering of evanescent waves [17]. The advent of SNOM has also allowed a direct experimental study of SPP excitation and scattering [18] and Anderson localization of surface excitations [19] and stimulated theoretical works on SPP scattering by surface roughness or localized objects [20]. In all these fields, the description of the coupling between an incident evanescent wave and a scattered propagating or evanescent wave is of fundamental interest. An  $S$ -matrix formalism, with a definition including the near-field components, should be very useful in this context. The  $S$  matrix also provides a useful formalism to discuss time reversal of wave fields, and especially its link with reciprocity (symmetry of the  $S$  matrix) and probability or energy conservation (unitarity). In particular, the question of time reversal of fields containing evanescent components has recently received increasing attention, with the demonstration of phase conjugation of optical near fields [21] and of time reversal of acoustic waves [22]. In this last case, the  $S$ -matrix formalism was used to discuss the properties of a time-reversal acoustic cavity, without taking into account the role of evanescent waves [23]. Nevertheless, the question of subwavelength focusing of a field by time reversal was raised. This is an important issue, whose discussion requires the use of a formalism including the evanescent components of the field.

Finding general properties of the  $S$  matrix, extended to evanescent waves, is of major importance in understanding and modeling all phenomena and devices involving near-field scattering. To our knowledge, this problem has received



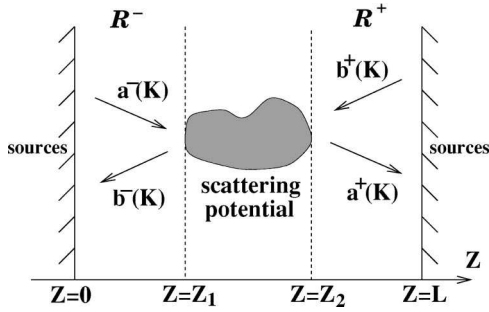


FIG. 1. Scattering geometry and notation.

little attention until now, except in electromagnetic waveguide theory where reciprocity and unitarity relations for the  $S$  matrix in the presence of evanescent modes have been briefly discussed [24]. In this work, we concentrate on reciprocity, unitarity, and time-reversal symmetry in the framework of scattering from a localized potential, in the presence of evanescent waves in both the incoming and outgoing fields. The formalism we use covers time-independent quantum scattering as well as scattering of classical scalar waves. In Sec. II, we define the  $S$  matrix based on the angular representation of the field, sometimes called the *partitioned*  $S$  matrix in the literature [9,25]. In Sec. III, we give a general derivation of the  $S$  matrix reciprocity for scalar (quantum and classical) fields, in the presence of evanescent waves. In Sec. IV, we show that energy (or probability) conservation in scattering of fields containing evanescent components leads to generalized unitarity relations of the  $S$  matrix. These relations extend those previously derived for source-free fields (i.e., fields without evanescent components) [9,25], and those obtained in electromagnetic waveguide theory [24]. In Sec. V, we address the problem of time-reversal symmetry of wave fields containing evanescent components. We show that the time-reversal invariance condition leads to a different relationship for the  $S$  matrix. On this basis, the link between reciprocity, unitarity, and time-reversal invariance is discussed. This problem is of fundamental and practical importance, due, for example, to its potential application to time reversal of acoustic waves [22]. Finally, we give a summary and a general conclusion in Sec. VI.

## II. DEFINITION OF THE $S$ MATRIX

Let us consider the scattering problem depicted in Fig. 1. The regions  $0 < z < z_1$  and  $z_2 < z < L$ , denoted by  $\mathcal{R}^-$  and  $\mathcal{R}^+$ , respectively, are assumed to be of constant potential, so that the wave field in these regions obeys the time-independent wave equation

$$\nabla^2 \Psi(\mathbf{r}) + k^2 \Psi(\mathbf{r}) = 0, \quad (1)$$

where  $\mathbf{r} = (x, y, z)$ . In Eq. (1),  $\Psi(\mathbf{r})$  is either the time-independent wave function of a state of energy  $E$  [ $k^2 = 2m/\hbar^2(E - V)$ , where  $V$  is the potential and  $m$  the mass of the particle], or a monochromatic classical wave of frequency  $\omega$  ( $k = \omega/c$ , where  $c$  is the phase velocity in the medium). The wave number  $k^2$  is real, but can be negative, e.g., in a tunneling barrier. The scattering potential is as-

sumed confined within the strip  $z_1 < z < z_2$ , and independent of the field (linear scattering). The regions  $z < 0$  and  $z > L$  contain sources, the presence of which, at finite distance from the scattering region, is necessary to ensure the existence of incident evanescent waves in both  $\mathcal{R}^-$  and  $\mathcal{R}^+$ . The evanescent waves explicitly appear when the angular-spectrum representation (or plane-wave expansion) of the field is used. In this representation, the fields  $\Psi^-$  in  $\mathcal{R}^-$  and  $\Psi^+$  in  $\mathcal{R}^+$  are written [9]

$$\begin{aligned} \Psi^\pm(\mathbf{r}) = & \int a^\pm(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} \\ & + \int b^\pm(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K}, \quad (2) \end{aligned}$$

where  $\gamma(\mathbf{K}) = \sqrt{k^2 - K^2}$  for  $K^2 \leq k^2$  (homogeneous or propagating components) and  $\gamma(\mathbf{K}) = i\sqrt{K^2 - k^2}$  for  $K^2 > k^2$  (inhomogeneous or evanescent components). We use the notations  $\mathbf{R} = (x, y)$  and  $K = |\mathbf{K}|$ . Except when the integration domain is specified, all integrals in this paper are extended to  $-\infty < (K_x, K_y) < +\infty$ . Note that Eq. (2) is a representation of the field valid in regions  $\mathcal{R}^-$  and  $\mathcal{R}^+$  where  $|z|$  remains finite, so that there is no divergence of evanescent waves when  $|z|$  increases.

In the angular-spectrum representation (2), the partitioned  $S$  matrix relates the outgoing vector  $\Psi^{out}(\mathbf{K}) = [b^-(\mathbf{K}) \ a^+(\mathbf{K})]$  to the incoming vector  $\Psi^{in}(\mathbf{K}) = [a^-(\mathbf{K}) \ b^+(\mathbf{K})]$  by the relation [9,25]

$$\Psi^{out}(\mathbf{K}) = \int \mathbf{S}(\mathbf{K}, \mathbf{K}') \Psi^{in}(\mathbf{K}') d^2\mathbf{K}', \quad (3)$$

where  $\mathbf{S}$  is a  $2 \times 2$  matrix, sometimes called the partitioned  $S$  matrix [9,25], which can be written in the form

$$\mathbf{S}(\mathbf{K}, \mathbf{K}') = \begin{bmatrix} r(\mathbf{K}, \mathbf{K}') & \tau(\mathbf{K}, \mathbf{K}') \\ t(\mathbf{K}, \mathbf{K}') & \rho(\mathbf{K}, \mathbf{K}') \end{bmatrix}. \quad (4)$$

The four elements  $r, t, \rho, \tau$  have the meaning of generalized reflection and transmission coefficients [9,25]. Their definition can easily be extended to vector fields, as in electromagnetic scattering. In this case, the four coefficients become tensor operators [26].

## III. RECIPROCITY RELATION

For incoming and outgoing fields without evanescent components (source-free fields), reciprocity and unitarity relations for the partitioned  $S$  matrix are well established [9,25]. They were derived as a consequence of the symmetry and unitarity of the asymptotic (far-field)  $S$  matrix. Extending these relations to general wave fields with evanescent components requires a different procedure. Note that reciprocity of evanescent waves was derived previously in specific cases, such as electromagnetic waveguide theory [24], electromagnetic (vector field) scattering [26], and elastic-wave scattering at a solid-solid interface [27]. In these examples, a suitable formulation of the reciprocity theorem was used for each particular case. In this paper, we give a proof of reciprocity of the  $S$  matrix for scattering of both homoge-



neous and evanescent scalar waves from a localized potential, starting from a general formulation of reciprocity valid for any kind of scalar wave.

Let  $\Psi_1$  and  $\Psi_2$  be two fields that are solutions of the scattering problem depicted in Fig. 1. With reference to Fig. 2, let us consider the volume  $V$  delimited by the closed surface  $\Sigma$  composed of two planes  $z=z^-$  and  $z=z^+$  and a sphere of radius  $R$  centered at the center of the potential region. The application of Green's second identity leads to

$$\int_V (\Psi_2 \Delta \Psi_1 - \Psi_1 \Delta \Psi_2) dV = \int_{\Sigma} \left( \Psi_2 \frac{\partial \Psi_1}{\partial n} - \Psi_1 \frac{\partial \Psi_2}{\partial n} \right) dS, \quad (5)$$

where  $\partial/\partial n = \mathbf{n} \cdot \nabla$  and  $\mathbf{n}$  is the outward normal on the surface  $\Sigma$ . Both  $\Psi_1$  and  $\Psi_2$  satisfy Eq. (1) in  $\mathcal{R}^-$  and  $\mathcal{R}^+$ , so that the integrand in the left-hand side in Eq. (5) vanishes. Moreover, in the far-field asymptotic limit  $|k|r \rightarrow \infty$ , one has  $\partial \Psi_j / \partial r = ikr^{-1} \Psi_j$ , so that  $\Psi_2 \partial \Psi_1 / \partial n - \Psi_1 \partial \Psi_2 / \partial n$  vanishes identically on the sphere surface when its radius  $R$  tends to infinity. Finally, Eq. (5) leads to the following equality:

$$\begin{aligned} & \int_{z=z^-} \left( \Psi_2 \frac{\partial \Psi_1}{\partial z} - \Psi_1 \frac{\partial \Psi_2}{\partial z} \right) d^2 \mathbf{R} \\ &= \int_{z=z^+} \left( \Psi_2 \frac{\partial \Psi_1}{\partial z} - \Psi_1 \frac{\partial \Psi_2}{\partial z} \right) d^2 \mathbf{R}. \end{aligned} \quad (6)$$

Equation (6) is a scalar version of Lorentz's reciprocity theorem, originally derived for the electromagnetic field [28]. In order to obtain a reciprocity theorem for the  $S$  matrix, we introduce the angular-spectrum representation Eq. (2) of the fields  $\Psi_1$  and  $\Psi_2$  into Eq. (6). After some algebra, one obtains the following expression:

$$\begin{aligned} & \int \int [a_1^-(\mathbf{K}) \quad b_1^+(\mathbf{K})] \{ \gamma(\mathbf{K}) \mathbf{S}(\mathbf{K}, \mathbf{K}') \\ & - \gamma(\mathbf{K}') \mathbf{S}^T(-\mathbf{K}', -\mathbf{K}) \} \begin{bmatrix} a_2^-(\mathbf{K}') \\ b_2^+(\mathbf{K}') \end{bmatrix} d^2 \mathbf{K} d^2 \mathbf{K}' = 0, \end{aligned} \quad (7)$$

where the superscript  $T$  denotes the transposed matrix. Because Eq. (7) must be satisfied for any incoming vectors  $[a_1^-(\mathbf{K}) \quad b_1^+(\mathbf{K})]$  in situation 1 and  $[a_2^-(\mathbf{K}) \quad b_2^+(\mathbf{K})]$  in situation 2, one finally obtains

$$\gamma(\mathbf{K}) \mathbf{S}(\mathbf{K}, \mathbf{K}') = \gamma(\mathbf{K}') \mathbf{S}^T(-\mathbf{K}', -\mathbf{K}). \quad (8)$$

Equation (8) is valid for  $0 < |\mathbf{K}| < +\infty$  and  $0 < |\mathbf{K}'| < +\infty$ , i.e., for propagating and evanescent waves. Note that the presence of the factors  $\gamma$  in Eq. (8) is a consequence of the definition of the angular spectrum of the field [Eq. (2)] by integration over the parallel wave vector  $\mathbf{K}$ . When using an integration over the solid angle  $\Omega$ , with  $k d\Omega = d^2 \mathbf{K} / \gamma$ , these factors disappear from the reciprocity relations. Nevertheless, the presence of evanescent waves would involve complex angles in the  $\Omega$  representation, so that the  $\mathbf{K}$  representation looks more appropriate.

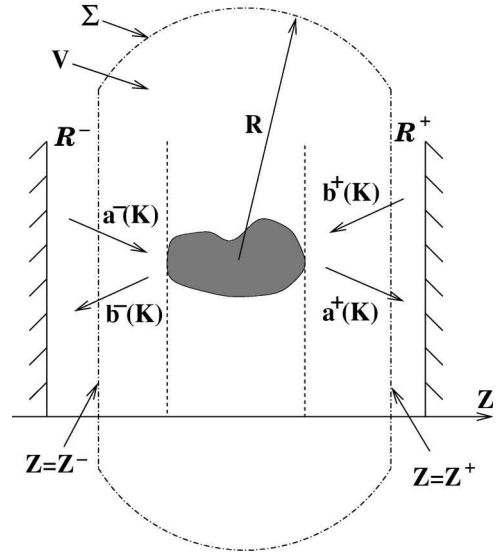


FIG. 2. Closed volume used for the application of Green's identity and the energy (or probability) balance.

#### IV. EXTENDED UNITARITY RELATIONS

The second basic property of the  $S$  matrix is unitarity. It is satisfied by the  $S$  matrix of a lossless system (elastic scattering). It is often assumed that the unitarity condition involves only propagating waves. This belief is, in fact, based on the use of asymptotic fields in the derivation. Indeed, unitarity has been studied extensively in far-field (asymptotic) scattering [2] or scattering of source-free fields in the angular-spectrum representation [9,25]. Conversely, the extension of unitarity relations to wave fields containing evanescent components has received little attention until now, except in the context of electromagnetic waveguide theory [24]. Nevertheless, as discussed in the Introduction of the present paper, extended unitarity relations could be helpful in various recent applications, such as time-reversed acoustics, near-field optics, or propagation through random media from the viewpoint of random-matrix theory. In this section, we show how such relations may be derived in the framework of free-space scattering from a localized potential.

In either  $\mathcal{R}^-$  or  $\mathcal{R}^+$ , the current density associated with the field is  $\mathbf{J}(\mathbf{r}) = A \text{Im} \{ \Psi^*(\mathbf{r}) \nabla \Psi(\mathbf{r}) \}$ , where  $A$  is a constant,  $\text{Im}$  denotes the imaginary part, and the asterisk is the complex conjugate. With reference to Fig. 2, energy (or probability) conservation states that the total flux  $\phi = \int \mathbf{J} \cdot \mathbf{n} dS$  flowing outside the volume  $V$  vanishes. When the radius  $R$  of the sphere tends to infinity, the contribution of the flux through the portions of the sphere surface between the two planes  $z=z^-$  and  $z=z^+$  vanishes. Finally, energy (or probability) conservation reads

$$\phi_{z^-} = \phi_{z^+}, \quad (9)$$

where

$$\phi_{z^\pm} = \int_{z=z^\pm} \mathbf{J}(\mathbf{R}, z^\pm) \cdot \mathbf{n} d^2 \mathbf{R}. \quad (10)$$

Using the angular-spectrum representation Eq. (2), the current  $\phi_{z^\pm}$  across a plane  $z=z^\pm$  in  $\mathcal{R}^\pm$  can be cast in the following form:

$$\begin{aligned} \phi_{z^\pm} = & A \int_{K^2 \leq k^2} \gamma[|a^\pm(\mathbf{K})|^2 - |b^\pm(\mathbf{K})|^2] d^2\mathbf{K} \\ & + A \int_{K^2 > k^2} \gamma[a^\pm(\mathbf{K})b^{\pm*}(\mathbf{K}) - a^{\pm*}(\mathbf{K})b^\pm(\mathbf{K})] d^2\mathbf{K}. \end{aligned} \quad (11)$$

This expression for the current deserves some comment. It explicitly shows two separated contributions, one stemming from propagating waves ( $K^2 \leq k^2$ ), and the other stemming from evanescent waves only ( $K^2 > k^2$ ). Note that the latter is a crossed term between counterdecaying evanescent waves. For a given  $\mathbf{K}$  such that  $K^2 > k^2$ , if either  $a(\mathbf{K})$  or  $b(\mathbf{K})$  vanishes, then the associated current also vanishes. For the aim pursued in this section, it is precisely the existence of this contribution to the current that leads to the extended unitarity relations of the  $S$  matrix. Introducing Eq. (11) into Eq. (9), and using the definition of the  $S$  matrix (3), one obtains the three following relations, involving scattering between two propagating waves, between two evanescent waves, and between one evanescent and one propagating wave:

$$\int_{K' \leq k} \frac{\gamma(\mathbf{K}'')}{\gamma(\mathbf{K}')} \mathbf{S}(\mathbf{K}, \mathbf{K}') \mathbf{S}^\dagger(\mathbf{K}'', \mathbf{K}') d^2\mathbf{K}' = \delta(\mathbf{K} - \mathbf{K}'') \mathbf{U} \quad \text{for } K \leq k, K'' \leq k, \quad (12)$$

$$= \mathbf{S}(\mathbf{K}, \mathbf{K}'') \quad \text{for } K \leq k, K'' > k, \quad (13)$$

$$= \mathbf{S}(\mathbf{K}, \mathbf{K}'') - \mathbf{S}^*(-\mathbf{K}, -\mathbf{K}'') \quad \text{for } K > k, K'' > k, \quad (14)$$

where  $\mathbf{U}$  is the  $2 \times 2$  unit matrix. The superscript  $\dagger$  denotes the conjugated and transposed matrix.

Equation (12) is the well-known unitarity condition of the  $S$  matrix restricted to the homogeneous components of the fields, which was obtained previously [2,3,25]. Using the partitioned form of the  $S$  matrix Eq. (4), this condition can be developed in terms of the generalized reflection and transmission coefficients. The resulting four expressions are the generalized Stokes relations of surface optics, which were derived in [25]. Equations (13) and (14) express, in the  $S$  matrix formalism, probability (or energy) conservation in a scattering process in which the incoming and/or the outgoing fields contain evanescent components. Hence, they can be considered as extended unitarity conditions for the  $S$  matrix of fields containing evanescent components. Because the  $S$ -matrix formalism is used in many fields of theoretical physics, and because there has been increasing interest in phenomena involving direct use and measurement of evanescent fields, we believe that this result may have important consequences and applications. For example, the development of Eqs. (13) and (14) in terms of generalized reflection and transmission coefficients leads to relationships that extend

the generalized Stokes relations to evanescent fields. Such relations may be useful, e.g., in the context of surfaces, thin films, and multilayers optics. The generalized Stokes relationships and their extension to fields containing evanescent components are given in the Appendix.

## V. TIME-REVERSAL INVARIANCE

In this section we discuss time-reversal invariance for classical scalar waves (we exclude from the discussion the question of time reversal in quantum mechanics which is difficult to separate from the measurement problem).

### A. Time reversal in angular-spectrum representation

Let  $\Psi(\mathbf{r}, t)$  be a classical scalar field *in the time domain*, and  $\Psi(\mathbf{r}, \omega)$  its frequency spectrum. Because  $\Psi(\mathbf{r}, t)$  is a real function, its frequency spectrum satisfies  $\Psi(\mathbf{r}, \omega) = \Psi^*(\mathbf{r}, -\omega)$ . From this condition, it is straightforward to show that the time-reversed field  $\Psi(\mathbf{r}, -t)$  has a frequency spectrum  $\Psi^*(\mathbf{r}, \omega)$ . Hence time reversal of  $\Psi(\mathbf{r}, t)$  is equivalent to complex conjugation of  $\Psi(\mathbf{r}, \omega)$  throughout all space. Note that time reversal is not equivalent to phase conjugation in only one plane [29]. We shall come back to this point later.

Let us see what time reversal means in terms of the angular spectrum of the monochromatic field  $\Psi(\mathbf{r})$  (the variable  $\omega$  is omitted in the following). From complex conjugation of Eq. (2) and the change of variable  $\mathbf{K} \rightarrow -\mathbf{K}$ , one obtains

$$\begin{aligned} \Psi^*(\mathbf{r}) = & \int a^*(-\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma^*z) d^2\mathbf{K} \\ & + \int b^*(-\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma^*z) d^2\mathbf{K}. \end{aligned} \quad (15)$$

The symbols  $\pm$  have been omitted in Eq. (15) because we do not need to specify at this stage whether the field propagates in  $\mathcal{R}^+$  or in  $\mathcal{R}^-$ . We see that, in terms of angular spectrum, time reversal is equivalent to the transformation  $a(\mathbf{K}) \exp(i\gamma z) \rightarrow a^*(-\mathbf{K}) \exp(-i\gamma^*z)$  and  $b(\mathbf{K}) \exp(-i\gamma z) \rightarrow b^*(-\mathbf{K}) \exp(i\gamma^*z)$  for all values of  $z$  in either  $\mathcal{R}^+$  or  $\mathcal{R}^-$ . Note that Eq. (15) is valid in regions of space for which  $|z|$  remains finite, so that there is no divergence of time-reversed evanescent waves when  $|z|$  increases.

### B. Time-reversal invariance: consequence for the $S$ matrix

In order to study the implication of time-reversal invariance for the  $S$  matrix, let us consider a monochromatic field  $\Psi_1(\mathbf{r})$  which is a solution of the scattering problem depicted in Fig. 1, the scattering potential being described by the  $S$  matrix  $\mathbf{S}$ . Let  $\Psi_2(\mathbf{r})$  be the time reversal of the field  $\Psi_1(\mathbf{r})$ . In terms of the angular spectrum, this means that

$$\begin{aligned} & a_1^+(\mathbf{K}) \exp(i\gamma z) + b_2^+(\mathbf{K}) \exp(-i\gamma z) \\ & = a_1^{+\dagger}(-\mathbf{K}) \exp(-i\gamma^*z) + b_1^{+\dagger}(-\mathbf{K}) \exp(i\gamma^*z) \end{aligned} \quad (16)$$

for all values of  $z$  in  $\mathcal{R}^+$  and

$$\begin{aligned} a_2^-(\mathbf{K})\exp(i\gamma z) + b_2^-(\mathbf{K})\exp(-i\gamma z) \\ = a_1^{-*}(-\mathbf{K})\exp(-i\gamma^*z) + b_1^{-*}(-\mathbf{K})\exp(i\gamma^*z) \end{aligned} \quad (17)$$

for all values of  $z$  in  $\mathcal{R}^-$ . The scattering problem is invariant under time reversal if, and only if,  $\Psi_2(\mathbf{r})$  is also a solution of the scattering problem, described by the same  $S$  matrix  $\mathbf{S}$ . This means that the outgoing vectors  $\Psi_1^{out}(\mathbf{K}) = [b_1^-(\mathbf{K}) \ a_1^+(\mathbf{K})]$  and  $\Psi_2^{out}(\mathbf{K}) = [b_2^-(\mathbf{K}) \ a_2^+(\mathbf{K})]$  are connected to the incoming vectors  $\Psi_1^{in}(\mathbf{K}) = [a_1^-(\mathbf{K}) \ b_1^+(\mathbf{K})]$  and  $\Psi_2^{in}(\mathbf{K}) = [a_2^-(\mathbf{K}) \ b_2^+(\mathbf{K})]$ , respectively, by relation (3). Introducing these conditions into Eqs. (16) and (17) leads to the following relations:

$$\begin{aligned} \int_{K' \leq k} \mathbf{S}(\mathbf{K}, \mathbf{K}') \mathbf{S}^*(-\mathbf{K}', \mathbf{K}'') d^2\mathbf{K}' \\ = \delta(\mathbf{K} + \mathbf{K}'') \mathbf{U} \quad \text{for } K \leq k, K'' \leq k, \end{aligned} \quad (18)$$

$$= -\mathbf{S}(\mathbf{K}, -\mathbf{K}'') \quad \text{for } K \leq k, K'' > k, \quad (19)$$

$$\begin{aligned} = \mathbf{S}^*(-\mathbf{K}, \mathbf{K}'') - \mathbf{S}(\mathbf{K}, -\mathbf{K}'') \\ \text{for } K > k, K'' > k, \end{aligned} \quad (20)$$

where the asterisk denotes the conjugated matrix. These relations express the condition of time-reversal invariance in terms of the  $S$  matrix of fields containing evanescent components.

The set of Eqs. (18)–(20) is very similar to the set of Eqs. (12)–(14), which describes energy conservation. In fact, it is easy to see that these two sets of equations are equivalent, provided that the reciprocity relation (8) is satisfied. Indeed, Eqs. (18)–(20) are transformed into Eqs. (12)–(14) by using Eq. (8) and changing  $\mathbf{K}'' \rightarrow -\mathbf{K}''$ . The result we have obtained can be summarized as follows: the condition of time-reversal invariance is equivalent to both energy conservation (extended unitarity condition) and reciprocity (symmetry of the  $S$  matrix). Although this result was already known for source-free fields [23], we have demonstrated that it holds for fields containing evanescent components.

### C. Time-reversal invariance and reciprocity

The results in this paper also provide a basis to discuss the link between time-reversal symmetry and reciprocity, which is sometimes confusing in the literature [3] (see also a discussion of this point in Ref. [26]). For a scattering system in which energy is conserved [Eqs. (12)–(14) are satisfied], the conditions of time-reversal invariance [Eqs. (18)–(20)] and reciprocity [Eq. (8)] are equivalent. This is probably the reason why time-reversal symmetry and reciprocity are often mistaken. In particular, reciprocity is often presented as a consequence of time-reversal invariance [3]. This is in general incorrect. For example, we have seen that imposing time-reversal invariance leads to Eqs. (18)–(20) and not to

the reciprocity condition Eq. (8). Moreover, a scattering system may be reciprocal, without being conservative [Eq. (8) is satisfied, but not Eqs. (12)–(14)]. In this case, the system is not invariant under time reversal [Eqs. (12)–(14) cannot be satisfied]. These conclusions hold for fields with or without evanescent components.

### D. Experimental feasibility of time reversal

An important problem is the experimental achievement of time reversal in a situation involving wave scattering. In optics, the development of phase conjugating mirrors provides a practical tool to produce fields that are conjugates of each other *in a given plane*. Nevertheless, it has been shown that this type of phase conjugation is not equivalent to time reversal when the fields involved contain evanescent components [29]. The subject of time reversal of fields containing evanescent components is of particular interest in the context of time-reversed acoustics [22]. In this technique, the acoustic field in a direct situation is recorded on a given surface  $\Sigma$ , after scattering by an arbitrary object. In the reversed situation, the time-reversed field is emitted from the surface  $\Sigma$ , in the presence of the same scattering object. In the frequency domain, the fields in the two situations are complex conjugates of each other on  $\Sigma$ . Thus, this experiment is equivalent to achieving acoustic phase conjugation on the surface  $\Sigma$ . In both optics [21] and acoustics [22], the possibility of achieving time reversal of both the homogeneous and evanescent components of the field by phase conjugation may be questioned.

The first part of the answer is given by showing that phase conjugation on the surface of a closed cavity (or equivalently along two planes  $z=z_1$  and  $z=z_2$ ) is equivalent to time reversal at all points *inside* the cavity (or in the strip  $z_1 < z < z_2$ ). This assertion is a consequence of the following result: two fields defined inside the strip  $z_1 < z < z_2$  that are complex conjugates in the two planes  $z=z_1$  and  $z=z_2$  are complex conjugates at all points within the strip  $z_1 < z < z_2$ . Therefore, they are time reversed from each other in the cavity. This result holds for fields containing evanescent components. It can be derived by extending the discussion in Ref. [29] to a situation involving phase conjugation along *two* planes  $z=z_1$  and  $z=z_2$ . Consequently, phase conjugation on a *closed* surface (or along two planes) may be a practical way to achieve complete time reversal of a field.

The second part of the answer must take into account the presence of sources inside the cavity in the direct experiment. In theory, reversing time leads automatically to the transformation of all primary sources into sinks. Therefore, to achieve time reversal experimentally, the field on the surface of the cavity has to be time reversed, *and* the sources have to be transformed into sinks. This is probably the greatest experimental challenge. This is also the necessary condition to obtain complete time reversal (i.e., with evanescent waves included) and achieve, for example, time-reversed focusing below the diffraction limit. The necessity of replacing sources by sinks in the time-reversed situation can be understood as follows. In the direct situation, a subwavelength source radiates a localized field whose angular spectrum con-

tains evanescent waves. In the reversed situation, the sink is equivalent to a source with opposite phase. This localized source also radiates evanescent waves which allow the time-reversed field to focus below the diffraction limit.

## VI. CONCLUSION

In summary, we have derived general properties of the  $S$  matrix of fields containing evanescent components. In particular, we have shown that energy (or probability) conservation leads to relationships that extend the well-known unitarity condition of the asymptotic  $S$  matrix. Using the partitioned  $S$  matrix, we have shown that these relationships lead to extended Stokes relations. We have also obtained different relationships as a consequence of time-reversal invariance. On this basis, we have discussed the link between unitarity, time-reversal symmetry, and reciprocity. With the increasing interest in techniques based on measurement and control of evanescent waves, we think that this work should find broad applications. In particular, we have briefly discussed its implications in time reversal of scattered fields by phase conjugation.

## ACKNOWLEDGMENTS

We wish to thank M. Fink, A. Garcia Martin, and J. L. Thomas for helpful discussions. This work was supported by the French-Spanish Integrated Program PICASSO. M.N.-V. acknowledges the support of Fundación Ramón Areces.

## APPENDIX: GENERALIZED STOKES RELATIONS FOR FIELDS CONTAINING EVANESCENT WAVES

In this Appendix, we summarize the relations that are obtained by inserting the coefficients of the partitioned  $S$  matrix Eq. (4) into relations (12)–(14). The first four relations are the generalized Stokes relationships obtained in Ref. [25]. The other relations are extensions of the Stokes relationships to fields containing evanescent components. We use the notations  $\gamma' = \gamma(\mathbf{K}')$  and  $\gamma'' = \gamma(\mathbf{K}'')$ .

### 1. Relations involving homogeneous waves only

Relations valid for  $K \leq k$  and  $K'' \leq k$ :

$$\int_{K' \leq k} d^2\mathbf{K}' [\rho(\mathbf{K}, \mathbf{K}') \tau^*(\mathbf{K}'', \mathbf{K}') + t(\mathbf{K}, \mathbf{K}') r^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = 0, \quad (\text{A1})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [\rho(\mathbf{K}, \mathbf{K}') \rho^*(\mathbf{K}'', \mathbf{K}') + t(\mathbf{K}, \mathbf{K}') t^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = \delta(\mathbf{K} - \mathbf{K}''), \quad (\text{A2})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [r(\mathbf{K}, \mathbf{K}') r^*(\mathbf{K}'', \mathbf{K}') + \tau(\mathbf{K}, \mathbf{K}') \tau^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = \delta(\mathbf{K} - \mathbf{K}''), \quad (\text{A3})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [r(\mathbf{K}, \mathbf{K}') t^*(\mathbf{K}'', \mathbf{K}') + \tau(\mathbf{K}, \mathbf{K}') \rho^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = 0. \quad (\text{A4})$$

### 2. Relations involving conversion of homogeneous to evanescent waves

Relations valid for  $K \leq k$  and  $K'' > k$ :

$$\int_{K' \leq k} d^2\mathbf{K}' [\rho(\mathbf{K}, \mathbf{K}') \tau^*(\mathbf{K}'', \mathbf{K}') + t(\mathbf{K}, \mathbf{K}') r^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = t(\mathbf{K}, \mathbf{K}''), \quad (\text{A5})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [\rho(\mathbf{K}, \mathbf{K}') \rho^*(\mathbf{K}'', \mathbf{K}') + t(\mathbf{K}, \mathbf{K}') t^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = \rho(\mathbf{K}, \mathbf{K}''), \quad (\text{A6})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [r(\mathbf{K}, \mathbf{K}') r^*(\mathbf{K}'', \mathbf{K}') + \tau(\mathbf{K}, \mathbf{K}') \tau^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = r(\mathbf{K}, \mathbf{K}''). \quad (\text{A7})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [r(\mathbf{K}, \mathbf{K}') t^*(\mathbf{K}'', \mathbf{K}') + \tau(\mathbf{K}, \mathbf{K}') \rho^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = \tau(\mathbf{K}, \mathbf{K}''). \quad (\text{A8})$$

### 3. Relations involving conversion of evanescent to evanescent waves

Relations valid for  $K > k$  and  $K'' > k$ :

$$\int_{K' \leq k} d^2\mathbf{K}' [\rho(\mathbf{K}, \mathbf{K}') \tau^*(\mathbf{K}'', \mathbf{K}') + t(\mathbf{K}, \mathbf{K}') r^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = t(\mathbf{K}, \mathbf{K}'') - t^*(-\mathbf{K}, -\mathbf{K}''), \quad (\text{A9})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [\rho(\mathbf{K}, \mathbf{K}') \rho^*(\mathbf{K}'', \mathbf{K}') + t(\mathbf{K}, \mathbf{K}') t^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = \rho(\mathbf{K}, \mathbf{K}'') - \rho^*(-\mathbf{K}, -\mathbf{K}''), \quad (\text{A10})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [r(\mathbf{K}, \mathbf{K}') r^*(\mathbf{K}'', \mathbf{K}') + \tau(\mathbf{K}, \mathbf{K}') \tau^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = r(\mathbf{K}, \mathbf{K}'') - r^*(-\mathbf{K}, -\mathbf{K}''), \quad (\text{A11})$$

$$\int_{K' \leq k} d^2\mathbf{K}' [r(\mathbf{K}, \mathbf{K}') t^*(\mathbf{K}'', \mathbf{K}') + \tau(\mathbf{K}, \mathbf{K}') \rho^*(\mathbf{K}'', \mathbf{K}')] \frac{\gamma''}{\gamma'} = \tau(\mathbf{K}, \mathbf{K}'') - \tau^*(-\mathbf{K}, -\mathbf{K}''). \quad (\text{A12})$$



- [1] W. Heisenberg, *Z. Phys.* **120**, 513 (1943); **120**, 673 (1943).
- [2] P. Roman, *Advanced Quantum Theory* (Addison-Wesley, Reading, MA, 1965), Chaps. 4 (S matrix formalism) and 5 (symmetry and invariance).
- [3] R.G. Newton, *Scattering Theory of Waves and Particles* (McGraw-Hill, New York, 1966), p. 44.
- [4] H.P. Stapp, *Rev. Mod. Phys.* **34**, 390 (1962).
- [5] G.F. Chew, *S-Matrix Theory of Strong Interaction* (Benjamin, New York, 1961).
- [6] A.A. Abrikosov, L.P. Gorkov, and I.E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963), Chaps. I–III.
- [7] E. Gerjuoy and D.S. Saxon, *Phys. Rev.* **94**, 1445 (1954).
- [8] D.S. Saxon, *Phys. Rev.* **100**, 1771 (1955).
- [9] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991), Chap. 5.
- [10] M.L. Mehta, *Random Matrices* (Academic, New York, 1991).
- [11] C.W.J. Beenakker, *Rev. Mod. Phys.* **69**, 731 (1997).
- [12] E. Bascones, M.J. Calderon, D. Castelo, T. López, and J.J. Sáenz, *Phys. Rev. B* **55**, R11 911 (1997); A. García-Martín, T. López-Ciudad, J.J. Sáenz, and M. Nieto-Vesperinas, *Phys. Rev. Lett.* **81**, 329 (1998).
- [13] R.R. Chance, A. Prock, and R. Silbey, in *Advances in Chemical Physics XXXVII*, edited by I. Prigogine and S.A. Rice (Wiley, New York, 1978); J.M. Wylie and J.E. Sipe, *Phys. Rev. A* **30**, 1185 (1984); C. Henkel and V. Sandoghdar, *Opt. Commun.* **158**, 250 (1998).
- [14] C.S. West and K.A. O'Donnell, *Opt. Lett.* **21**, 1 (1996); A.A. Maradudin, A.R. McGurn, and E.R. Méndez, *J. Opt. Soc. Am. A* **12**, 2500 (1995).
- [15] C.J. Chen, *Introduction to Scanning Tunneling Microscopy* (Oxford University Press, Oxford, 1993).
- [16] *Near-Field Optics*, edited by D.W. Pohl and D. Courjon (Kluwer Academic Publishers, Dordrecht, 1993); *Near-Field Nano/Atom Optics and Technology*, edited by M. Ohtsu (Springer-Verlag, Tokyo, 1998).
- [17] J.M. Vigoureux, C. Girard, and D. Courjon, *Opt. Lett.* **14**, 1039 (1989); D. van Labeke and D. Barchiesi, *J. Opt. Soc. Am. A* **10**, 2193 (1993); C. Girard and A. Dereux, *Rep. Prog. Phys.* **59**, 657 (1996); J.-J. Greffet and R. Carminati, *Prog. Surf. Sci.* **56**, 139 (1997).
- [18] O. Marti, H. Bielefeldt, B. Hecht, S. Herminghaus, P. Leiderer, and J. Mlynek, *Opt. Commun.* **96**, 225 (1993); P.M. Adam, L. Salomon, F. de Fornel, and J.P. Goudonnet, *Phys. Rev. B* **48**, 2680 (1993); S.I. Bozhevolnyi and F.A. Pudonin, *Phys. Rev. Lett.* **78**, 2823 (1997).
- [19] D.P. Tsai, J. Kovacs, Z. Wang, M. Moskovits, V.M. Shalaev, J.S. Suh, and R. Botet, *Phys. Rev. Lett.* **72**, 4149 (1994); S. Grésillon, L. Aigouy, A.C. Boccara, J.C. Rivoal, X. Quelin, C. Desmarest, P. Gadenne, V.A. Shubin, A.K. Saryshev, and V.M. Shalaev, *ibid.* **82**, 4520 (1999).
- [20] F. Pincemin, A.A. Maradudin, A.D. Boardman, and J.-J. Greffet, *Phys. Rev. B* **50**, 15 261 (1994); J.A. Sánchez-Gil, *ibid.* **53**, 10 317 (1996); A.V. Shchegrov, I.V. Novikov, and A.A. Maradudin, *Phys. Rev. Lett.* **78**, 4269 (1997); S.I. Bozhevolnyi and V. Coello, *Phys. Rev. B* **58**, 10 899 (1998).
- [21] S.I. Bozhevolnyi, O. Keller, and I.I. Smolyaninov, *Opt. Lett.* **19**, 1601 (1994).
- [22] M. Fink, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39** (5), 555 (1992); M. Fink, *Phys. Today* **50** (3), 34 (1997).
- [23] J.L. Thomas, Ph.D. thesis, University of Paris VI, 1994, Chap. 4; D. Cassereau and M. Fink, *IEEE Trans. Ultrason. Ferroelect. Freq. Control* **39** (5), 579 (1992).
- [24] C. Vassallo, *Théorie des Guides d'Ondes Electromagnétiques* (Eyrolles, Paris, 1985), Vol. 2, p. 169.
- [25] M. Nieto-Vesperinas and E. Wolf, *J. Opt. Soc. Am. A* **3**, 2038 (1986).
- [26] R. Carminati, M. Nieto-Vesperinas, and J.-J. Greffet, *J. Opt. Soc. Am. A* **15**, 706 (1998).
- [27] L. Kazandjian, *J. Acoust. Soc. Am.* **92**, 1679 (1992).
- [28] H.A. Lorentz, *Versl. Gewone Vergad. Afd. Natuurkd. K. Ned. Akad. Wet.* **4**, 176 (1896); Lorentz, *Collected Papers* (Nijhoff, Den Haag, 1936), Vol. III. See also M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Ref. [9]), Chap. 1, p. 8.
- [29] M. Nieto-Vesperinas and E. Wolf, *J. Opt. Soc. Am. A* **2**, 1429 (1985). See also M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Ref. [9]), Chap. 8.

## Scattering Theory of Bardeen's Formalism for Tunneling: New Approach to Near-Field Microscopy

R. Carminati<sup>1</sup> and J.J. Sáenz<sup>2</sup>

<sup>1</sup>*Laboratoire d'Énergétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

<sup>2</sup>*Departamento de Física de la Materia Condensada and Instituto de Ciencia de Materiales "Nicolás Cabrera," Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain*

(Received 19 October 1999)

We propose a new theoretical approach to near-field microscopy, which allows one to deal with scanning tunneling microscopy and scanning near-field optical microscopy with a unified formalism. Under the approximation of weak tip-sample coupling, we show that Bardeen's perturbation formula, originally derived for electron tunneling, can be derived from a scattering formalism which extends its validity to electromagnetic vector fields. This result should find broad applications in near-field imaging and spectroscopy.

PACS numbers: 61.16.Ch, 03.65.Nk, 07.79.Cz, 73.40.Gk

The development of scanning tunneling microscopy (STM) in the early eighties [1] opened the way to real-space surface study at the atomic scale. Since then, various techniques of scanning probe microscopy (SPM) have been proposed [2,3], based on local interaction between a sharp tip and the sample under study. Scanning near-field optical microscopy (SNOM) [4] is one of these techniques, which uses optical interaction in the visible or near-infrared range. SNOM has proven its ability to image optical fields and surface structure at a subwavelength scale [5]. In the field of microscopy, spectroscopy, and surface modification on the nanometer scale with visible or infrared light [6], SNOM looks complementary to other SPM techniques.

In the context of STM, some theories were developed shortly after the first experimental demonstrations, based on self-consistent methods and numerical calculations [7,8] or on analytical models [9–12]. Many of these theories [8–10] have in common the use of Bardeen's perturbation formula, originally derived for electron tunneling between two weakly coupled electrodes [13]. In particular, the approach of Tersoff and Hamann [9] remains an explicit and practical description of the STM. An important result in this approach was the direct interpretation of the STM signal as a measurement of the local electron density of state of the sample. Although this result is valid under weak tip-sample coupling, it was a breakthrough in understanding the STM images [2].

Similarly, in the context of SNOM, several theoretical methods and numerical simulations [14], as well as analytical models [15,16], have been developed, in order to improve the capability of the technique and to understand the measured signals. Although the underlying physics behind SNOM is understood to a certain extent, an overlook at the current state of SNOM leads to the two following remarks. (i) The analogy between STM and SNOM is often qualitatively put forward. In particular, some SNOM setups such as the photon scanning tunneling microscope

(PSTM) were introduced by analogy between optical and electron tunneling [17]. Nevertheless, there is no unified formalism and theoretical proof of a clear and general analogy. (ii) An explicit SNOM theory was developed some years ago [15], which gave an interpretation of the signal and clarified the role of spatial filtering and polarization effects. Nevertheless, a general formalism allowing to introduce in a natural way an appropriate tip model seems to be missing [18].

In this Letter, we propose a new approach to near-field microscopy which deals with both STM and SNOM with a unified formalism. We first derive an expression of the current in the gap [19] which is valid for STM and SNOM. This expression allows an original discussion of the tunneling contribution to the SNOM signal. Then, under the approximation of weak tip-sample coupling, we derive a general expression of the signal in SNOM, which generalizes Bardeen's formula to scattering of vector electromagnetic fields. This generalization allows one to deal with SNOM using the standard formalism of STM modeling.

Let us consider the general SNOM setup depicted in Fig. 1(a), and the general STM setup in Fig. 1(b). In the SNOM situation, the tip-sample system is illuminated by a light source of arbitrary size and state of coherence, and part of the scattered energy is collected by a detector of arbitrary shape. The gap region (between the sample and the tip) is assumed to be vacuum or air. At this stage of the discussion, we concentrate on the tunneling current in both STM and SNOM, and we do not take polarization effects into account. In the STM situation, we assume that the central part (with respect to the  $z$  direction) of the gap region is of constant potential  $V$ . The state of the electromagnetic field at a given frequency  $\omega$ , or a stationary state of the electron of energy  $E$ , are both represented by a scalar wave function  $\Psi(\mathbf{r})$ . We assume that the tip remains situated above the highest point of the surface topography (although the path followed during the scan may

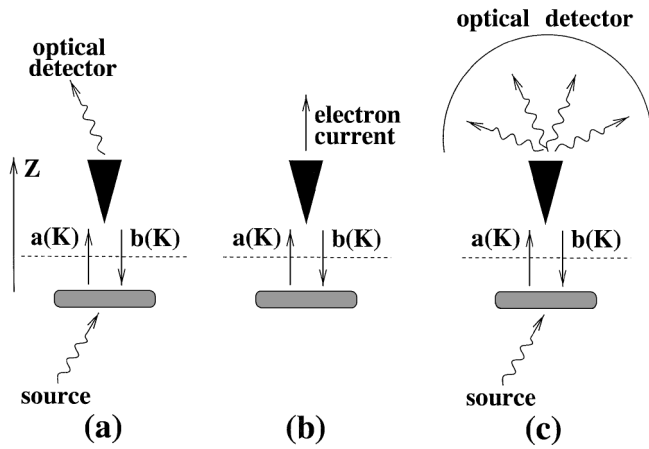


FIG. 1. (a) Scheme of a SNOM setup. Light coming from the source is scattered towards the detector through near-field coupling between the tip and the sample. (b) Scheme of a STM setup. The current is created by tunneling electrons between the tip and the sample. (c) SNOM setup with hemispherical detection.

be arbitrary). In the gap region, the wave field can be written in the form of an angular spectrum of plane waves [20]:

$$\Psi(\mathbf{r}) = \int a(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K} + \int b(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K}, \quad (1)$$

where  $\gamma(\mathbf{K}) = \sqrt{k^2 - K^2}$  for  $K^2 \leq k^2$  (homogeneous or propagating components) and  $\gamma(\mathbf{K}) = i\sqrt{K^2 - k^2}$  for  $K^2 > k^2$  (inhomogeneous or evanescent components). We use the notations  $\mathbf{r} = (x, y, z)$ ,  $\mathbf{R} = (x, y)$ , and  $K = |\mathbf{K}|$ . For the electromagnetic field,  $k = \omega/c$ ,  $c$  being the speed of light in vacuum. For the electron wave function,  $k^2 = 2m/\hbar^2(E - V)$ , where  $m$  is the electron mass and  $\hbar$  is Planck's constant. The integrals are extended to  $0 < K < +\infty$ . Note that, in the case of electron tunneling in STM ( $E < V$  and  $k^2 < 0$ ), the wave function in the gap region contains evanescent waves only.

The current density associated with the wave function  $\Psi$  is  $\mathbf{J}(\mathbf{r}) = A \text{Im}[\Psi^*(\mathbf{r})\nabla\Psi(\mathbf{r})]$ , where  $\text{Im}$  denotes the imaginary part and  $*$  the complex conjugate. This formula represents either the momentum density of the electromagnetic field in the scalar representation or the probability current in quantum mechanics [20]. The constant  $A$  may be determined by identifying the current flux at the detector with either the energy flux of the electromagnetic field (in the case of SNOM) or the electronic current (in the case of STM). Using Eq. (1), the total current  $\phi = \int J_z(\mathbf{r}) d^2\mathbf{R}$  across a plane at a constant  $z$  in the gap region (dashed line in Fig. 1) can be cast in the following form:

$$\phi = A \int_{K^2 \leq k^2} \gamma[|a(\mathbf{K})|^2 - |b(\mathbf{K})|^2] d^2\mathbf{K} + A \int_{K^2 > k^2} \gamma[a(\mathbf{K})b^*(\mathbf{K}) - a^*(\mathbf{K})b(\mathbf{K})] d^2\mathbf{K}. \quad (2)$$

Although Eq. (2) simply expresses the total current flowing through the gap region, it was never used before, to our knowledge, in the context of near-field microscopy. In STM, except for a constant factor,  $\phi$  is exactly the tunneling current which is measured in the experiment. In SNOM,  $\phi$  is proportional to the *total* energy flux, including the flux flowing through channels that do not end up at the detector. In many SNOM experiments, only part of this flux is actually collected and contributes to the signal.  $\phi$  would be an exact expression of the signal in situations in which a hemispherical detector is used to collect all the flux traveling in a half space, as shown in Fig. 1(c). An example of such a configuration is the tunneling near-field optical microscope [21] when complete hemispherical detection is performed, and its reciprocal setup, namely, a PSTM using hemispherical incoherent illumination [22,23].

Two separate contributions are clearly identified in Eq. (2). The first integral describes the contribution of waves that are homogeneous in the gap region. It simply expresses the balance between two currents flowing in opposite directions through propagating channels. The second integral describes a current flowing through evanescent (or tunneling) channels. In the case of STM, this is the only contribution to the current. This term simply reflects the net flux traveling through the tunneling channel  $\mathbf{K}$ , and vanishes if  $a(\mathbf{K}) = b(\mathbf{K})$ . Note that, if either  $a(\mathbf{K})$  or  $b(\mathbf{K})$  vanishes, then the contribution of this tunneling channel also vanishes. This reflects the fact that tunneling is essentially a consequence of the presence of two interfaces at close proximity (e.g., the sample and the tip). Equation (2) also demonstrates the existence of an optical tunneling contribution in any SNOM configuration. Moreover, it shows that the SNOM current travels through both propagating and tunneling channels in the gap, whereas in STM the current flows only through tunneling channels. This is a fundamental difference between SNOM and STM.

In practice, computing the SNOM or STM signal from Eq. (2) requires the knowledge of the angular spectra  $a(\mathbf{K})$  and  $b(\mathbf{K})$  of the wave function in the gap region. These are solutions of a difficult scattering problem in a confined geometry, which can, in general, be solved only numerically. Nevertheless, under the approximation of weak tip-sample coupling, it is known in STM modeling that Bardeen's formula can be used to describe the tunneling current [2,8–10]. We shall now give a new derivation of this formula, based on a scattering formalism. This approach generalizes Bardeen's original formula, by

showing that (i) it describes both the tunneling current and the current flowing through propagating channels, and that (ii) it also applies to vector electromagnetic fields.

Let us consider the general SNOM setup depicted in Fig. 2(a). The illumination is done by a plane wave with a wave vector  $\mathbf{K}_s$ , a unit amplitude, and a state of polarization described by the complex unit vector  $\mathbf{a}_s$ . The signal is recorded by a detector placed in the far field, in a direction defined by the wave vector  $\mathbf{K}_d$ . We assume that the detection is performed with a polarizer (analyzer) whose polarization direction is defined by the unit vector  $\mathbf{a}_d$ . Note that this represents the most general configuration, because an extended and/or unpolarized source or detector can be described by adding the contributions of a set of incoming or outgoing plane waves. Depending on the experimental setup, the summation should be done with a properly defined degree of coherence and/or polarization [23]. Without loss of generality, we have chosen the transmission geometry shown in Fig. 2(a), but the argument can be easily extended to any SNOM setup. Using a scattering formalism, we describe the sample, the tip, and the entire setup by their generalized transmission coefficients  $\vec{\tau}_s(\mathbf{K}, \mathbf{K}_s)$ ,  $\vec{\tau}_d(\mathbf{K}_d, \mathbf{K})$ , and  $\vec{T}(\mathbf{K}_d, \mathbf{K}_s)$ . These coefficients are elements of the scattering matrix of each system in a plane-wave basis [24]. The signal  $S$  is the flux of the Poynting vector (current density) at the detector position (i.e., in the far field). The far-field asymptotic expression of the electromagnetic field in the direction  $\mathbf{K}_d$  can be obtained by the stationary-phase technique [20]. In this condition, the expression of the signal is

$$S = 2\pi^2 \epsilon_0 c \gamma^2(\mathbf{K}_d) |\mathbf{a}_d \cdot \vec{T}(\mathbf{K}_d, \mathbf{K}_s) \cdot \mathbf{a}_s|^2. \quad (3)$$

This result shows that the basic quantity to compute is  $M_{ds} = \mathbf{a}_d \cdot \vec{T}(\mathbf{K}_d, \mathbf{K}_s) \cdot \mathbf{a}_s$ , which is analogous to the elastic tunneling matrix element in Bardeen's formalism [13]. We now assume that the coupling between the tip and the sample is weak. In the scattering picture, this means that the current in the gap results from fields that have been scattered once at the tip or at the sample. In this case, the transmission coefficient of the system is

$$\vec{T}(\mathbf{K}_d, \mathbf{K}_s) = \int \vec{\tau}_d(\mathbf{K}_d, \mathbf{K}) \cdot \vec{\tau}_s(\mathbf{K}, \mathbf{K}_s) d^2\mathbf{K}, \quad (4)$$

where the integral is extended to both propagating and tunneling channels. We see that, in the case of weak coupling,

$$M_{ds} = \frac{1}{8\pi^2 i \gamma(\mathbf{K}_d)} \int \left[ \Psi_d(\mathbf{r}) \cdot \frac{\partial \Psi_s}{\partial z}(\mathbf{r}) - \Psi_s(\mathbf{r}) \cdot \frac{\partial \Psi_d}{\partial z}(\mathbf{r}) \right] d^2\mathbf{R}, \quad (7)$$

where the integral is performed along a plane at a constant  $z$  in the gap region.

Equation (7) is the main result of this Letter. It is similar to Bardeen's formula for the elastic tunneling matrix element  $M_{\mu\nu}$  between a state  $\Psi_\mu$  of the probe and a state  $\Psi_\nu$  of the sample [see, e.g., Eq. (3) in Ref. [9]]. Note

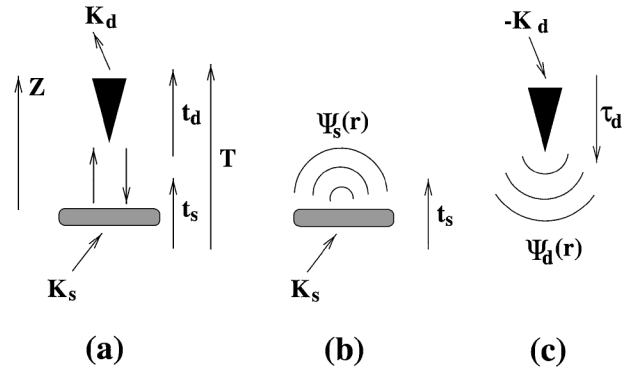


FIG. 2. (a) General SNOM setup with directional illumination and detection. (b) Illustration of the meaning of the sample wave function  $\Psi_s$ . (c) Illustration of the meaning of the tip wave function  $\Psi_d$ .

the signal can be calculated from the transmission coefficients of the sample and the tip, considered as independent systems. We will now transform Eq. (4) into an expression in direct space, involving two wave fields that are solutions of the two scattering problems in Figs. 2(b) and 2(c). This will lead to a generalization of Bardeen's formula to scattering of electromagnetic vector fields. Let  $\Psi_s(\mathbf{r})$  be the (vector) electric field, in the gap region, that results from scattering of the illuminating plane wave (wave vector  $\mathbf{K}_s$ , polarization state  $\mathbf{a}_s$ ) by the sample, in the absence of the tip. Let  $\Psi_d(\mathbf{r})$  be the (vector) electric field, in the gap region, that results from scattering by the tip of a plane wave of amplitude unity coming from the direction of the detector (wave vector  $-\mathbf{K}_d$ , polarization state  $\mathbf{a}_d$ ). The explicit expressions of these wave fields are

$$\Psi_s(\mathbf{r}) = \int \vec{\tau}_s(\mathbf{K}, \mathbf{K}_s) \cdot \mathbf{a}_s \exp(i\mathbf{K} \cdot \mathbf{R} + i\gamma z) d^2\mathbf{K}, \quad (5)$$

$$\Psi_d(\mathbf{r}) = \int \vec{\tau}_d(\mathbf{K}, -\mathbf{K}_d) \cdot \mathbf{a}_d \exp(i\mathbf{K} \cdot \mathbf{R} - i\gamma z) d^2\mathbf{K}, \quad (6)$$

where  $\vec{\tau}_d$  is related to  $\vec{\tau}_d$  by the reciprocity theorem  $\gamma(\mathbf{K}_d) \vec{\tau}_d(\mathbf{K}_d, \mathbf{K}) = \gamma(\mathbf{K}) \vec{\tau}_d^T(-\mathbf{K}, -\mathbf{K}_d)$ , the superscript  $T$  denoting the transposed tensor [24]. From Eqs. (4)–(6), one obtains the following expression for the matrix element  $M_{ds}$ :

that the complex conjugation of the tip wave function  $\Psi_d$  does not appear in Eq. (7). This point is not fundamental. Bardeen's formula is exactly retrieved when using a tip wave function  $\Psi'_d = \Psi_d^*$ , namely, the *time reversed* of the wave function  $\Psi_d$  introduced in Eq. (6). When the



tip is lossless, this difficulty can be overcome by choosing a real solution for  $\Psi_d$ . This is the usual choice that is made in STM modeling based on Bardeen's formula [9,12]. Our formalism clearly demonstrates that the tip wave function must be chosen as a solution of the scattering problem in Fig. 2(c), followed by time reversal, for the integrand in Eq. (7) to take the form of a current operator, as in Bardeen's original paper [13]. The general derivation of Eq. (7) given in this Letter shows that Bardeen's formula applies (i) to the general situation where the current in the gap flows through both propagating and tunneling channels, and (ii) to electromagnetic vector fields. From a fundamental point of view, this constitutes an important generalization of Bardeen's original formula. From a practical side, Eqs. (3) and (7) provide an expression of the SNOM signal using the same formalism as in standard STM modeling [9,10,12]. This unification could greatly improved the understanding of the SNOM signal.

Finally, let us comment on the choice of the tip wave field  $\Psi_d$  in SNOM modeling. The simplest model could consider the tip as a small sphere. In this case, it can be shown [25] that one retrieves the same result as that obtained in Ref. [15]. Except for some polarization effects, the signal in this model is proportional to  $|\Psi_s|^2$  at the location of the probe. This is also the result of the Tersoff and Hamann theory developed for STM [9]. Nevertheless, recent experimental studies of the apertureless setup have shown that this model is not sufficient to describe polarization and spectroscopic effects [18]. A theory including an appropriate tip model is needed to account for these effects [26]. The generalized Bardeen formula derived in this Letter provides a natural way to include realistic tip shapes in SNOM modeling.

In summary, we have presented a new formalism for near-field microscopy which unifies SNOM and STM. We have given an explicit expression of the total current in the gap region, which demonstrates the role of optical tunneling in SNOM. Under the approximation of weak tip-sample coupling, we have given a new and general derivation of Bardeen's perturbation formula, which applies to currents flowing through both propagating and tunneling channels, and to electromagnetic fields. The results in this Letter should find broad applications in near-field imaging and spectroscopy using visible or infrared light.

This work was supported by the French-Spanish Integrated Program PICASSO. We acknowledge helpful discussions with P. Chaumet, J.-J. Greffet, and M. Nieto-Vesperinas.

- 
- [1] G. Binnig *et al.*, Phys. Rev. Lett. **49**, 57 (1982); **50**, 120 (1983).  
 [2] C. J. Chen, *Introduction to Scanning Tunneling Microscopy* (Oxford University Press, Oxford, 1993).  
 [3] R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (Cambridge Univer-

sity Press, Cambridge, England, 1994); *Scanning Tunneling Microscopy and Scanning Force Microscopy of Biological Samples*, edited by O. Marti and M. Amrein (Academic Press, New York, 1993).

- [4] D. W. Pohl, W. Denk, and M. Lanz, Appl. Phys. Lett. **44**, 651 (1984); E. Betzig *et al.*, Biophys. J. **49**, 269 (1986).  
 [5] For a review, see Ultramicroscopy (special issues): **57**, 2–3 (1995); **61**, 1–4 (1995); **71**, 1–4 (1998).  
 [6] E. Betzig and J. K. Trautmann, Science **257**, 189 (1992); E. Betzig and R. Chichester, Science **262**, 1422 (1993); B. Hecht *et al.*, Phys. Rev. Lett. **77**, 1889 (1996); S. I. Bozhevolnyi and F. A. Pudonin, Phys. Rev. Lett. **78**, 2823 (1997); J. Massanell, N. García, and A. Zlatkin, Opt. Lett. **21**, 12 (1996).  
 [7] N. García, C. Ocal, and F. Flores, Phys. Rev. Lett. **50**, 2002 (1983); A. A. Lucas *et al.*, Phys. Rev. B **37**, 10708 (1988).  
 [8] N. D. Lang, Phys. Rev. Lett. **55**, 230 (1985).  
 [9] J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985).  
 [10] C. J. Chen, J. Vac. Sci. Technol. A **6**, 319 (1988).  
 [11] W. Sacks and C. Noguera, Phys. Rev. B **43**, 11 612 (1991).  
 [12] C. Bracher, M. Riza, and M. Kleber, Phys. Rev. B **56**, 7704 (1997).  
 [13] J. Bardeen, Phys. Rev. Lett. **6**, 57 (1961).  
 [14] C. Girard and A. Dereux, Rep. Prog. Phys. **59**, 657 (1996).  
 [15] D. van Labeke and D. Barchiesi, J. Opt. Soc. Am. A **10**, 2193 (1993).  
 [16] J.-J. Greffet and R. Carminati, Prog. Surf. Sci. **56**, 139 (1997).  
 [17] R. Reddick *et al.*, Phys. Rev. B **39**, 767 (1989); F. de Fornel *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **1139**, 77 (1989); D. Courjon *et al.*, Opt. Commun. **71**, 23 (1989).  
 [18] Recent experiments [L. Aigouy *et al.*, Opt. Lett. **24**, 187 (1998); Appl. Phys. Lett. **76**, 397 (2000)] demonstrated that understanding polarization and spectroscopic effects in apertureless SNOM requires a tip model going beyond the small-sphere approximation.  
 [19] The term “current” is used for both the STM signal (electron current) and the SNOM signal (electromagnetic energy flux).  
 [20] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley, New York, 1991).  
 [21] B. Hecht, H. Heinzelmann, and D. W. Pohl, Ultramicroscopy **57**, 228 (1995).  
 [22] G. Chabrier *et al.*, Opt. Commun. **107**, 347 (1994); N. García and M. Nieto-Vesperinas, Opt. Lett. **20**, 949 (1995); R. Carminati *et al.*, Opt. Lett. **21**, 501 (1996).  
 [23] E. R. Méndez, J.-J. Greffet, and R. Carminati, Opt. Commun. **142**, 7 (1997).  
 [24] R. Carminati, M. Nieto-Vesperinas, and J.-J. Greffet, J. Opt. Soc. Am. A **15**, 706 (1998).  
 [25] Assuming that the tip is a small sphere, behaving as a point dipole located at  $\mathbf{r}_t$ , the measured signal, proportional to  $|M_{ds}|^2$ , can be shown to be as follows [R. Carminati (unpublished)]:  $S \propto |\mathbf{a}_d \cdot \vec{h}(\mathbf{K}_d) \cdot \Psi_s(\mathbf{r}_t)|^2$ , where  $\vec{h}(\mathbf{K})$  is the projection operator on direction transverse to the wave vector  $\mathbf{k} = (\mathbf{K}, \gamma)$ , given in Ref. [24]. An integration over a solid angle leads to Eq. (19) of Ref. [15]. For scalar fields, one obtains  $S \propto |\Psi_s(\mathbf{r}_t)|^2$ , which is the result of Ref. [9].  
 [26] J. A. Porto, R. Carminati, and J.-J. Greffet (to be published).

# Theory of electrostatic probe microscopy: A simple perturbative approach

S. Gómez-Moñivas and J. J. Sáenz<sup>a)</sup>

*Departamento de Física de la Materia Condensada and Instituto de Ciencia de Materiales "Nicolás Cabrera," Universidad Autónoma de Madrid, Cantoblanco, 28049 Madrid, Spain*

R. Carminati and J. J. Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion; Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

(Received 26 January 2000; accepted for publication 24 March 2000)

A theoretical approach to electrostatic scanning probe microscopy is presented. We show that a simple perturbation formula, originally derived in the context of scattering theory of electromagnetic waves, can be used to obtain the capacitance and the electrostatic force between a metallic tip and an inhomogeneous dielectric sample. For inhomogeneous thin dielectric films, the scanning probe signal is shown to be proportional to the convolution between an effective surface profile and a response function of the microscope. This provides a rigorous framework to address the resolution issue and the inverse problem. © 2000 American Institute of Physics. [S0003-6951(00)04820-8]

Since the development of scanning tunneling microscopy<sup>1</sup> and atomic force microscopy<sup>2</sup> in the early eighties various techniques of scanning probe microscopy (SPM) have been proposed,<sup>3</sup> based on different local interactions between a sharp tip and the sample under study. The long range nature of electrostatic interactions makes them specially suitable to perform noncontact SPM imaging of both conducting and insulating materials. By applying a voltage between a force microscope tip and a sample, electrostatic force microscopy (EFM) has been used to study capacitance,<sup>4</sup> surface potential,<sup>5</sup> charge or dopant distribution,<sup>6</sup> topography and dielectric properties of metallic and insulating<sup>5,7</sup> surfaces and to deposit and image localized charges on insulators.<sup>8</sup> In analogy with the magnetic force microscope,<sup>9</sup> EFM has been used to image the domain structure of ferroelectric crystals.<sup>10</sup> Polarization forces have also been used to imaging weakly bound materials and liquids<sup>11</sup> and to perform electrostatic spectroscopy.<sup>12</sup>

As in other SPM techniques, the interpretation of the EFM images is not always evident. Since EFM is a nonlocal technique due to the long range nature of the electrostatic interaction, the detailed shape and dimensions of the tip must then be taken into account for a precise calculation of both force and capacitance.<sup>13</sup> Most of the theoretical work on EFM has been focused on a better understanding of tip shape effects on the electric field, force, and capacitance.<sup>13,14</sup> Although the influence of the tip shape is now more or less well understood for flat and homogeneous samples, there is no simple way to directly relate the electrostatic image with the dielectric and topographic properties of the sample. In this letter, we propose a theoretical approach to electrostatic probe microscopy that represents a first step to fill this gap. In analogy with previous theoretical work on scanning near-field optical microscopy (SNOM),<sup>15</sup> we will show that the EFM image is related to both the topography and dielectric inhomogeneities of the sample through a response function which describes all the instrument properties. Our perturba-

tive analytical approach is checked by comparison with exact numerical calculations. In the important case of imaging of thin dielectric films deposited on metallic substrates, we show that the force (or capacitance) signal closely follows an *equivalent surface profile*. This equivalent surface profile connects the film topography with the dielectric inhomogeneities, providing a simple physical picture of the contrast mechanism in EFM.

We consider a three-dimensional sample with both topographic and dielectric constant inhomogeneities (see Fig. 1). This sample is a finite layer of profile  $Z(x,y)=Z(\mathbf{r}_{\parallel})$  and dielectric constant  $\epsilon(\mathbf{r})$  on top of a reference sample. For simplicity, we will take a semi-infinite homogeneous ( $z > 0$ ) substrate of dielectric constant  $\epsilon_s$  as the reference sample. Our approach would equally apply to any reference sample surface with known dielectric response, however.

Under a constant tip-sample bias  $V$ , the electrostatic energy of the reference system (i.e., in the homogeneous case), is given by:

$$U_0 = \frac{1}{2} \int \epsilon_0 \mathbf{E}_0^2 d^3\mathbf{r} = \frac{1}{2} C_0 V^2, \quad (1)$$

where  $\mathbf{E}_0$  is the electric field and  $C_0$  is the capacitance. The electrostatic force (normal to the sample surface)  $F_{0z}$  can be written as the energy gradient:

$$F_{0z} = -\frac{\partial}{\partial z} U_0 = -\frac{1}{2} V^2 \frac{\partial}{\partial z} C_0. \quad (2)$$

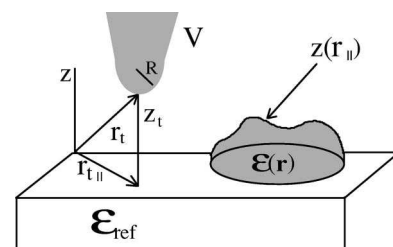


FIG. 1. Schematic configuration of an EFM.

<sup>a)</sup>Electronic mail: juanjo.saenz@uam.es

The presence of surface or volume inhomogeneities induces a change in the electrostatic energy (with respect to the reference sample),<sup>16</sup>

$$\Delta U = -\frac{1}{2} \int_V \mathbf{P} \cdot \mathbf{E}_0 d^3 \mathbf{r}, \quad (3)$$

where  $\mathbf{E}_0$  is the reference field and  $\mathbf{P} = \epsilon_0 [\epsilon(\mathbf{r}) - 1] \mathbf{E}$ , being  $\mathbf{E}$  the total field. In practice, computing the electrostatic energy (i.e., the force or/and the capacitance) from Eq. (3) requires the knowledge of the total self-consistent field in the gap region. These are solutions of a difficult Laplace problem in an open geometry, which can only be solved numerically. In order to handle this problem we will make use of a simple perturbative approach which was shown to be useful in scattering from rough surfaces.<sup>17</sup>

Following a simple Born-like approach one could replace the total field  $\mathbf{E}$  in Eq. (3) by the nonperturbed field  $\mathbf{E}_0$ . However, this simple approach is known to give wrong results in scattering from rough surfaces.<sup>17</sup> One way to improve this approximation is to take into account the discontinuity of the normal component of the field at the boundaries<sup>17</sup>

$$\Delta U = -\frac{1}{2} \epsilon_0 \int_{V_1} [\epsilon(\mathbf{r}) - 1] \left[ \frac{E_{0z}^2}{\epsilon(\mathbf{r})} + \mathbf{E}_{0\parallel}^2 \right] d^3 \mathbf{r}. \quad (4)$$

The force signal (or the capacitance) is directly obtained from  $\Delta U$  through  $\Delta F = \partial \Delta U / \partial z_t$  (or  $\Delta C = V^2 \Delta U / 2$ ). Equation (4) is the main result of our letter. This is an important result since the signal (image) is related to the topography and dielectric properties of the sample through a response function which depends only on the tip shape and the geometry of the reference sample. Although, in general, it is only a perturbative result, it is worth noticing that this equation gives the exact result for a parallel plate capacitor.

In order to check the validity of our perturbative approach we have compared our results with an extensive numerical calculation.<sup>18</sup> In this case, the reference system is a spherical tip of radius  $R = 50$  nm and a flat substrate with the geometry depicted in Fig. 2 (top). We have calculated the field  $\mathbf{E}_0$  in this reference system using the image-charge method. The force is then computed from Eq. (4) and a derivation with respect to  $z_t$ . The calculated EFM images at two different constant heights  $z_t$  are shown in Fig. 2. Forces are normalized to the force on the reference sample. Figure 3 shows scans along  $Y$  at  $X = 100$  nm (i.e., across two maxima). These results can be compared directly with numerical results obtained by using a self-consistent integral equation formalism (see Figs. 8 and 9 in Ref. 18). Taking into account the simplicity of our model, the agreement between the perturbative approach and the exact numerical results is remarkable.

In order to get a deeper understanding on the nature of the image contrast, let us consider a common experimental situation in which a dielectric soft sample is on a substrate with metallic character (i.e.,  $\epsilon_s \rightarrow \infty$ ). In this case, the electric field parallel to the substrate surface will be close to zero and the main contribution to the signal will come from the normal electric field. If the dielectric thickness is small com-

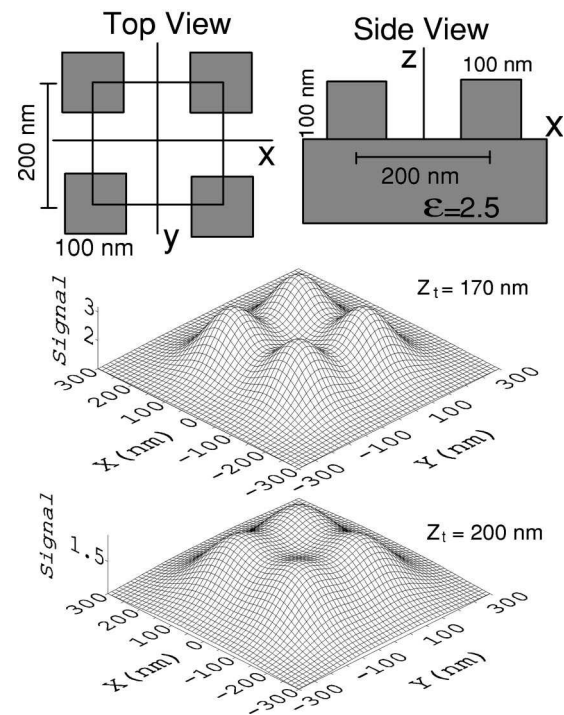


FIG. 2. (top) Geometry of a typical testing sample. Numerical “exact” results can be found in Ref. 18. (bottom) Calculated EFM image at two different constant heights. The force signal is normalized to the force of the reference surface.

pared with a typical field gradient length scale, i.e.,  $E_{0z}(\mathbf{r}_{\parallel} - \mathbf{r}_{\parallel}, z_t - z) \approx E_{0z}(\mathbf{r}_{\parallel} - \mathbf{r}_{\parallel}, z_t)$ , the energy will take the simple form of

$$\Delta U \approx \frac{1}{2} \epsilon_0 \int_{S_0} \{Z_{\text{eff}}(\mathbf{r}_{\parallel}) \cdot E_{0z}^2(\mathbf{r}_{\parallel} - \mathbf{r}_{\parallel}, z_t)\} d^2 \mathbf{r}_{\parallel}, \quad (5)$$

where

$$Z_{\text{eff}}(\mathbf{r}_{\parallel}) \equiv \int_0^{Z(\mathbf{r}_{\parallel})} \frac{\epsilon(\mathbf{r}) - 1}{\epsilon(\mathbf{r})} dz \quad (6)$$

is an *equivalent surface profile*<sup>15</sup> connecting the dielectric constant variation and the topography of the sample. The signal  $\Delta F = \partial \Delta U / \partial z_t$  (or  $\Delta C = V^2 \Delta U / 2$ ) will then be a simple two-dimensional convolution between the equivalent surface profile  $Z_{\text{eff}}$  and the *response function of the micro-*

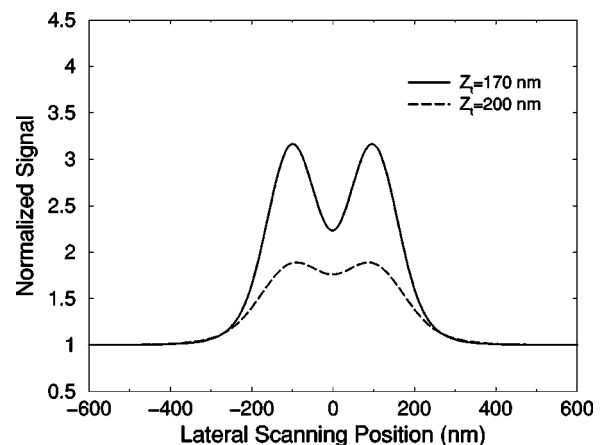


FIG. 3. Force scans corresponding to the images in Fig. 2, along  $Y$  at  $X = 100$  nm (i.e., across two maxima).

scope  $\mathcal{F}(\mathbf{r}_{\parallel}) = \partial |E_{0z}(\mathbf{r}_{\parallel}, z_t)|^2 / \partial z_t$ . Notice that the actual image would give information about  $Z_{\text{eff}}$ . For a homogeneous sample,  $Z_{\text{eff}}$  is directly proportional to the true topographic profile, while for a flat surface it reflects an average of the dielectric constant along the normal to the surface.<sup>19</sup>

In summary, we have presented a formalism for electrostatic force microscopy based on a modified first order perturbation theory. We have checked and illustrated our theory with exact numerical results for a model system. Our model describes how the topographic and dielectric constant variations of the sample influences the observed image in EFM. This is a very important point in EFM, where the purely dielectric properties of the sample are of great interest. In analogy with SNOM imaging,<sup>15</sup> we have introduced the concept of equivalent surface profile as the physical measured quantity in force microscopy. We believe that the results in this letter should find broad applications in the analysis of electrostatic imaging with scanning probe methods.

This work was supported by the French-Spanish Integrated Program PICASSO. The authors acknowledge helpful discussions with J. Colchero, A. García-Martín, and S. Lányi.

<sup>1</sup>G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* **49**, 57 (1982); G. Binnig, H. Rohrer, Ch. Gerber, and E. Weibel, **50**, 120 (1983).

<sup>2</sup>G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* **56**, 936 (1986).

<sup>3</sup>C. J. Chen, *Introduction to Scanning Tunneling Microscopy* (Oxford University Press, Oxford, 1993); R. Wiesendanger, *Scanning Probe Microscopy and Spectroscopy: Methods and Applications* (Cambridge University Press, Cambridge, 1994); *Scanning Tunneling Microscopy and Scanning Force Microscopy of Biological Samples*, edited by O. Marti and M. Amrein (Academic, New York, 1993).

<sup>4</sup>C. D. Bugg and P. J. King, *J. Phys. E* **21**, 147 (1988); C. C. Williams, W. P. Hough, and A. Rishton, *Appl. Phys. Lett.* **55**, 203 (1989).

<sup>5</sup>Y. Martin, D. W. Abraham, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **52**, 1103 (1988); M. Nonnenmacher, M. O'Boyle, and H. K. Wickramasinghe, *Ultramicroscopy* **42–44**, 268 (1992); T. Hochwitz, H. K. Henning, C. Leveg, C. Daghlian, and J. Slinkman, *J. Vac. Sci. Technol. B* **14**, 457 (1996).

<sup>6</sup>C. C. Williams, J. Slinkman, W. P. Hough, and H. K. Wickramasinghe, *Appl. Phys. Lett.* **55**, 1662 (1989); R. C. Barrett and C. F. Quate, *Ultramicroscopy* **42–44**, 262 (1992).

<sup>7</sup>R. Erlandsson, G. M. McClelland, C. M. Mate, and S. Chiang, *J. Vac. Sci. Technol. A* **6**, 266 (1988); S. Lányi, J. Török, and P. Rehurek, *J. Vac. Sci. Technol. B* **14**, 892 (1996).

<sup>8</sup>J. E. Stern, B. D. Terris, H. J. Mamin, and D. Rugar, *Appl. Phys. Lett.* **53**, 2717 (1998); B. D. Terris, J. E. Stern, D. Rugar, and H. J. Mamin, *Phys. Rev. Lett.* **63**, 2669 (1989); C. Schonenberger and S. Alvarado, *Phys. Rev. Lett.* **65**, 3162 (1990).

<sup>9</sup>Y. Martin and H. K. Wickramasinghe, *Appl. Phys. Lett.* **50**, 1455 (1987); J. J. Sáenz, N. García, P. Grütter, E. Meyer, H. Heinzelmann, R. Wiesendanger, L. Rosenthaler, H. R. Hidber, and H.-J. Güntherodt, *J. Appl. Phys.* **62**, 4293 (1987).

<sup>10</sup>F. Saurenbach and B. Terris, *Appl. Phys. Lett.* **56**, 1703 (1990); R. Lüthi, H. Haefke, K.-P. Meyer, E. Meyer, L. Howald, and H.-J. Güntherodt, *J. Appl. Phys.* **74**, 7461 (1992); T. Hidaka, T. Maruyama, M. Saitoh, N. Mikoshiba, M. Shimizu, T. Shiosaki, L. A. Wills, R. Hiskes, S. D. Dicarolis, and J. Amano, *Appl. Phys. Lett.* **68**, 2358 (1996); H. Bluhm, A. Wadas, R. Wiesendanger, K.-P. Meyer, and L. Szcześniak, *Phys. Rev. B* **55**, 4 (1997).

<sup>11</sup>J. Hu, X. D. Xiao, D. F. Ogletree, and M. Salmeron, *Science* **268**, 267 (1995); J. Hu, X. D. Xiao, and M. Salmeron, *Appl. Phys. Lett.* **67**, 476 (1995).

<sup>12</sup>D. Gekhtman, Z. B. Zhang, and G. Dresselhaus, *Phys. Rev. Lett.* **82**, 3887 (1999).

<sup>13</sup>H. W. Hao, A. M. Baró, and J. J. Sáenz, *J. Vac. Sci. Technol. B* **9**, 1323 (1991); S. Watanabe, K. Hane, T. Ohye, M. Ito, and T. Goto, *J. Vac. Sci. Technol. B* **11**, 1774 (1993); S. Belaidi, P. Girard, and G. Leveque, *J. Appl. Phys.* **81**, 1023 (1997); S. Hudlet, M. Saint Jean, C. Guthmann, and J. Berger, *Eur. Phys. J.* **2**, 5 (1998).

<sup>14</sup>G. Mesa, E. Dobado-Fuentes, and J. J. Sáenz, *J. Appl. Phys.* **79**, 39 (1996); G. Mesa and J. J. Sáenz, *Appl. Phys. Lett.* **69**, 1169 (1996).

<sup>15</sup>R. Carminati and J.-J. Greffet, *J. Opt. Soc. Am. A* **12**, 2716 (1995); *Ultramicroscopy* **61**, 11 (1995); J.-J. Greffet and R. Carminati, *Prog. Surf. Sci.* **56**, 133 (1997).

<sup>16</sup>J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), Chap. 4, p. 160.

<sup>17</sup>G. S. Agarwal, *Phys. Rev. B* **14**, 846 (1976).

<sup>18</sup>Z.-Y. Li, B.-Y. Gu, and G.-Z. Yang, *Phys. Rev. B* **57**, 9225 (1998).

<sup>19</sup>For the simple case of homogeneous flat films adsorbed on a semiinfinite substrate, this result can be obtained from the simpler electrostatic considerations of Hu, Xiao, and Salmeron following Eq. (1) in Ref. 11.

# Annexe 2

## Propagation et imagerie en milieu diffusant

Liste des articles reproduits :

- J. Ripoll, M. Nieto-Vesperinas and R. Carminati, *Spatial resolution of diffuse photon density waves*, J. Opt. Soc. Am. A **16**, 1466–1476 (1999).
- J.B. Thibaud, R. Carminati and J.-J. Greffet, *Scattering of a diffusive wave by a subsurface object*, J. Appl. Phys. **87**, 7638–7646 (2000).
- J. Ripoll, V. Ntziachristos, R. Carminati and M. Nieto-Vesperinas, *Kirchhoff approximation for diffusive waves*, Phys. Rev. E **64**, 51917 (2001).
- R. Elaloufi, R. Carminati and J.-J. Greffet, *Time-dependent transport through scattering media: from radiative transfer to diffusion*, J. Optics A: Pure and Applied Optics **4**, S103–S108 (2002).
- R. Elaloufi, R. Carminati and J.-J. Greffet, *Definition of the diffusion coefficient in scattering and absorbing media*, soumis à J. Opt. Soc. Am. A (2002).

# Spatial resolution of diffuse photon density waves

J. Ripoll and M. Nieto-Vesperinas

*Instituto de Ciencia de Materiales de Madrid, Consejo Superior de Investigaciones Científicas,  
Campus de Cantoblanco, 28049 Madrid, Spain*

Rémi Carminati

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris,  
Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

Received October 1, 1998; revised manuscript received February 5, 1999; accepted February 9, 1999

On comparison with the usual propagating scalar waves, the attenuation of diffuse photon density waves gives rise to important differences in structural information, such as higher spatial resolution in detection at short distances from objects and deviation from the Rayleigh limit at larger distances. This damping also establishes a minimum spatial resolution threshold for diffusive waves, which occurs by illumination in continuous mode, and demonstrates that in most cases spatial resolution is not improved by increasing the modulation frequency. Assessments of this formulation with numerical simulations of scattering and wave-front reconstruction in the presence of noise are given. © 1999 Optical Society of America [S0740-3232(99)01706-8]

OCIS codes: 170.5270, 290.1990, 170.7050, 100.6640.

## 1. INTRODUCTION

The study of light transport through strongly scattering media has recently received increasing attention because of its application to medical diagnosis.<sup>1,2</sup> In particular, much research is motivated by the ability of optical radiation to diagnose human breast cancer. In many practical situations the diffusion approximation is sufficiently accurate to describe visible or near-infrared light transport within turbid media such as human tissues.<sup>1-3</sup> Several imaging methods have been analyzed.<sup>4-23</sup> The frequency-domain methods use a light source modulated at a frequency  $\omega$ . In this case the average intensity  $U(\mathbf{r}, t) = U(\mathbf{r})\exp(-i\omega t)$  within the turbid medium obeys the Helmholtz equation with the wave number,

$$\kappa_0 = (-\mu_a/D + i\omega n/cD)^{1/2}, \quad (1)$$

where  $\mu_a$  is the absorption coefficient,  $c$  is the speed of light in vacuum,  $n$  is the index of refraction, and  $D = 1/3(\mu_a + \mu'_s)$  is the diffusion coefficient.  $\mu'_s$  is the reduced scattering coefficient, defined as  $(1 - g)\mu_s$ , where  $g$  is the average cosine of the scattering angle and  $\mu_s$  is the scattering coefficient. The solutions  $U(\mathbf{r})$  of the Helmholtz equation are called diffuse photon density waves (DPDW's). Their wave number  $\kappa_0 = \kappa_r + i\kappa_i$  is complex, with  $\kappa_r = 2\pi/\lambda_0$  and  $\kappa_i = 2\pi/l_a$ ,  $\lambda_0$  being their wavelength and  $l_a$  their decay length. In practice, with typical modulation frequencies and human tissues,  $\lambda_0$  is a few centimeters, whereas the attenuation takes place within distances shorter than a wavelength. Therefore DPDW's are strongly damped waves, and their detection is performed at subwavelength distances from the sources or the hidden objects, i.e., in the near field. This point is of great importance as far as the potential spatial resolution of the technique is concerned. In fact, it is well known that near-field imaging allows spatial resolution beyond the Rayleigh limit of  $\lambda_0/2$ . The problem of spatial

resolution with DPDW's has been addressed in both experimental and theoretical works.<sup>4-11,25</sup> Nevertheless, to our knowledge, no rigorous and general discussion of this topic can be found in the literature.

In this paper we present a rigorous theoretical analysis of the spatial resolution in imaging with DPDW's, and we illustrate the results by exact numerical simulation of the scattering of DPDW's by two objects hidden in a turbid medium. We address spatial resolution in its standard meaning (see Ref. 24), that is, as the ability to separate two object points, or fine details, on measurement at a certain distance from the scattering object. As will be shown, this spatial resolution depends only on the medium in which the wave propagates and on the detection-plane distance to the scattering object. The limits of recovery of the object's optical parameters by inverse scattering constitute a related subject that should not be confused with the term spatial resolution. Even so, a simple backpropagation scheme is presented to demonstrate the effects of this fundamental limit of spatial resolution. We find an analytical expression for the resolution limit of DPDW's, and we compare it with that known for the usual propagating scalar waves (PSW's). This concept can be applied for estimating the expected resolution in object reconstruction in any particular diffusive medium. As a consequence, we find that for DPDW's in a very few realistic cases, one obtains greater resolution by increasing the frequency of the incident wave, and in most practical cases illumination in the dc regime yields the same resolution as in the ac regime. On the other hand, if the main goal is to estimate the optical properties of the object, then ac or time-resolved measurements must be performed.<sup>26</sup> The issue of the effects of noise in both the resolution and the contrast is also studied, and exact numerical results are shown. In Section 2 we use the angular-spectrum representation of the wave field



$U(\mathbf{r})$  to derive a resolution criterion for DPDW's. In particular, we contrast the behavior of DPDW's with that of PSW's. Then we discuss the transition toward the electrostatic limit, defined as the region where the retarded effects are neglected, i.e., the dc regime for DPDW's. Following Rayleigh's resolution criterion, we find the analytical expressions for the full width at half-maximum (FWHM) of both the propagation transfer function and of its Fourier transform in real space, namely, the propagation impulse response. Hence these FWHM's give us directly the estimation of the spatial resolution limit either in frequency (FWHM of the transfer function) or in real space (FWHM of the impulse response).

The electrostatic limit can be found as one approaches the limit of infinite wavelength ( $\lambda_0 \rightarrow \infty$ ) in the expressions for the two FWHM's. Section 2 also shows a comparison of our results with experimental results previously presented by other authors. As an illustration of the discussion in Section 2, we present results of scattering numerical simulations in Section 3. The FWHM analytical expressions are verified with numerical examples. We also examine the effect of noise on the spatial resolution limit and illustrate it with numerical results with additive numerical noise, thus demonstrating how the nature of DPDW's allows a filtering of the image that substantially eliminates the noise contribution without much distortion of this image, and therefore reinforcing the use of the resolution limit put forward here. The effect of this filtering on the backpropagation of the scattered wave front is discussed in Subsection 3.C. Finally, in Section 4 we summarize the main conclusions.

## 2. THEORETICAL ANALYSIS

### A. Angular Spectrum for Diffuse Photon Density Waves

Let us consider a homogeneous multiple-scattering medium separated into two half-spaces  $z < 0$  and  $z > 0$  (see Fig. 1). It is assumed that the domain  $z < 0$  contains sources and scatterers (hidden objects), whereas the domain  $z > 0$  is source free. At any plane  $z = \text{constant}$ , with  $z > 0$ , we can express the scattered wave  $U(\mathbf{r})$  by its angular-spectrum representation of plane waves. That is, by a superposition of such waves of amplitude  $\mathcal{A}(\mathbf{K})$  and wave vector  $\mathbf{k} = (\mathbf{K}, q)$ ,  $|\mathbf{k}| = \kappa_0$  (Refs. 27–29):

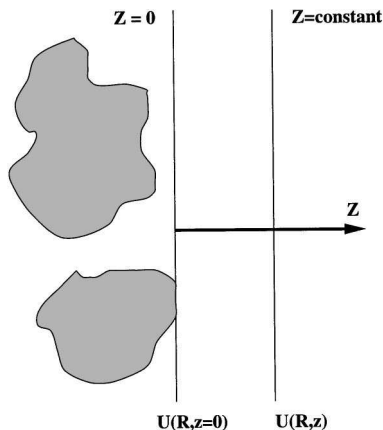


Fig. 1. Geometry used for the angular-spectrum representation.

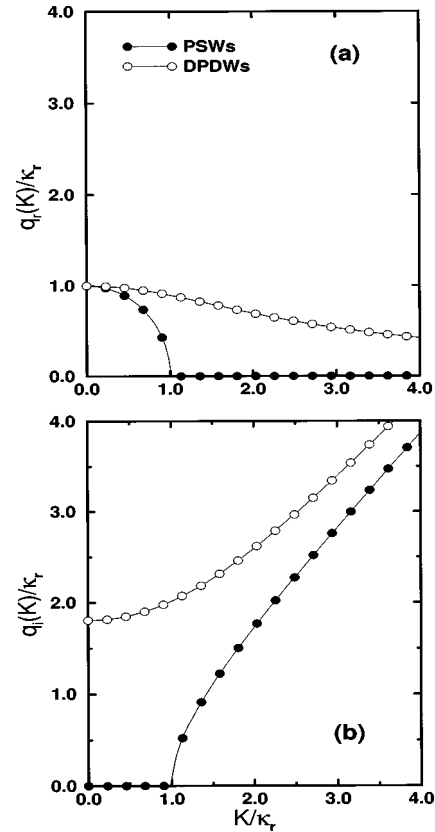


Fig. 2. Values of (a)  $q_r$  and (b)  $q_i$ , normalized to  $\kappa_r$ : solid circles, PSW's; open circles, DPDW's.

$$U(\mathbf{R}, z) = \int_{-\infty}^{+\infty} \mathcal{A}(\mathbf{K}) \exp[i\mathbf{K} \cdot \mathbf{R} + iq(\mathbf{K})z] d\mathbf{K}, \quad (2)$$

where  $\mathbf{R} = (x, y)$ ,  $|\mathbf{K}|^2 + q^2 = \kappa_0^2$ , i.e.,  $\mathbf{K} = (K_x, K_y)$  is a real vector and  $q(\mathbf{K}) = (\kappa_0^2 - |\mathbf{K}|^2)^{1/2}$ . For DPDW's, since  $\kappa_0$  is a complex number,  $q(\mathbf{K}) = q_r + iq_i$  is always complex; that is,  $q_r, q_i \neq 0$ . In Eq. (2) we choose  $q_r > 0$  and  $q_i > 0$  so that the field propagates toward  $z > 0$  and satisfies the radiation condition at infinity.

Note that for PSW's in the same geometry (the background turbid medium is replaced by a transparent dielectric),  $\kappa_0 = \kappa_r$  would be real. In this case,  $q(\mathbf{K}) = q_r = (\kappa_0^2 - |\mathbf{K}|^2)^{1/2}$  for  $|\mathbf{K}| \leq \kappa_0$  (homogeneous components), and  $q(\mathbf{K}) = iq_i = i(|\mathbf{K}|^2 - \kappa_0^2)^{1/2}$  for  $|\mathbf{K}| > \kappa_0$  (evanescent components). This difference is important for spatial resolution. For PSW's in a transparent medium, high spatial frequencies  $|\mathbf{K}|$  are exponentially attenuated, whereas low spatial frequencies always propagate. The cutoff frequency  $\kappa_0$  is well defined and underlines the diffraction limit of resolution in optical imaging. On the other hand, for DPDW's in a multiple-scattering medium, even without absorption there are always both propagation and attenuation at any spatial frequency. This difference is illustrated in Fig. 2, in which we plot  $q_r$  [Fig. 2(a)] and  $q_i$  [Fig. 2(b)] versus  $|\mathbf{K}|$  for both DPDW's and PSW's. The values taken for  $\kappa_r$  and  $\kappa_i$  correspond to breast tissue illuminated with light at a wavelength of 780 nm<sup>1</sup> and a modulation frequency  $\omega = 200$  MHz (where a  $2\pi$  factor is assumed), with parameters  $\lambda_0 = 7.53$  cm,  $l_a = 0.066$  cm,  $\mu_a = 0.035$  cm<sup>-1</sup>,  $\mu'_s$

= 15 cm<sup>-1</sup>, and  $n = 1.333$ . Figure 2 shows that for PSW's the regions of propagation ( $|\mathbf{K}| \leq \kappa_r$ ) and attenuation ( $|\mathbf{K}| > \kappa_r$ ) regimes are clearly separated; however, for DPDW's,  $q_r$  and  $q_i$  are of the same order of magnitude for  $|\mathbf{K}| \approx \kappa_r$ , so that there is no sharp transition between these two regimes. Nevertheless, as shown by these figures, DPDW's behave asymptotically like PSW's for  $|\mathbf{K}| \gg \kappa_r$ .

**B. Transfer Function and Impulse Response**

From Eq. (2) one obtains

$$A(\mathbf{K})\exp[iq(\mathbf{K})z] = \frac{1}{4\pi^2} \int_{-\infty}^{+\infty} U(\mathbf{R}, z)\exp(-i\mathbf{K} \cdot \mathbf{R})d\mathbf{R}. \tag{3}$$

Equation (3) shows that  $A(\mathbf{K})\exp[iq(\mathbf{K})z]$  is the two-dimensional Fourier transform of the wave field  $U(\mathbf{R}, z)$  in the plane  $z = \text{constant}$ . The spatial-frequency filter  $F(\mathbf{K}, z) = \exp[iq(\mathbf{K})z]$  constitutes the propagation transfer function.

The amplitude and the phase of  $F(\mathbf{K}, z)$  are represented in Fig. 3 for a two-dimensional geometry, namely,  $\mathbf{K} = (K, 0)$ , and for several propagation distances  $z$ . Both transfer functions for PSW's (left column) and DPDW's (right column) are shown. For DPDW's the values of  $\kappa_r$  and  $\kappa_i$  correspond to the breast tissues' parameters as in Fig. 2.

For PSW's the propagating and attenuation regions are clearly visible. For  $|\mathbf{K}| \leq \kappa_r$ , the transfer function is only a phase factor, whereas for  $|\mathbf{K}| > \kappa_r$ , it is a real low-pass filter. For large  $|\mathbf{K}|$ , one has  $q(\mathbf{K}) \approx |\mathbf{K}|$ , so that the transfer function is  $\exp(-|\mathbf{K}|z)$ , and high spatial frequencies are exponentially attenuated. Thus for PSW's a given spatial frequency  $\mathbf{K}$  has a decay length  $1/|\mathbf{K}|$ , and the cutoff frequency in the plane  $z = \text{constant}$  is  $1/z$ . These properties are well known in near-field optics.<sup>30</sup>

However, for DPDW's, the behavior of the transfer function is substantially different from that of PSW's. As we already noticed in Fig. 2, now there are no longer two separated propagation and attenuation regions. For a given observation distance  $z$ , the amplitude has its maximum at  $\mathbf{K} = 0$  and decreases for  $|\mathbf{K}| > 0$ . The peak value tends monotonically to zero as  $z$  increases, owing to the factor  $\exp[-q_i(\mathbf{K})z]$ , whereas its width broadens. The phase varies less abruptly than in the case of PSW's.

A description in direct space is also useful for discussing the spatial resolution conveyed in DPDW's. We introduce the impulse response  $H(\mathbf{R}, z)$ , namely, the Fourier transform with respect to  $\mathbf{K}$  of  $F(\mathbf{K}) = \exp[iq(\mathbf{K})z]$ :

$$H(\mathbf{R}, z) = \int_{-\infty}^{+\infty} F(\mathbf{K}, z)\exp(i\mathbf{K} \cdot \mathbf{R})d\mathbf{K}. \tag{4}$$

In terms of this impulse response, from Eq. (3) with  $z = 0$  and from Eq. (2), the wave function can be written as

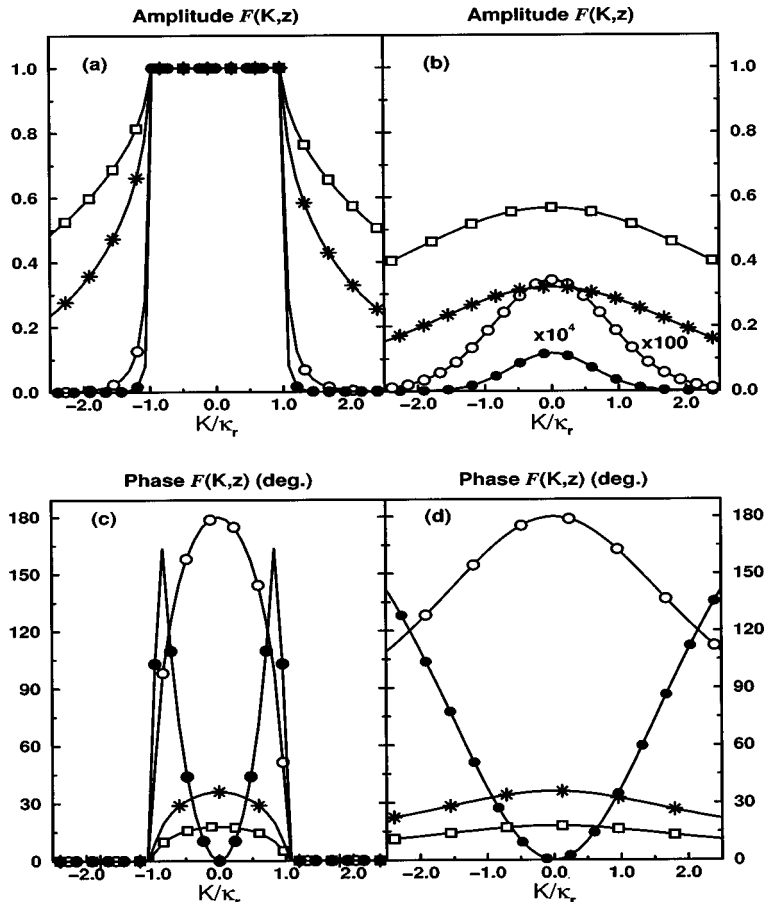


Fig. 3. Amplitude [(a) PSW's and (b) DPDW's] and phase [(c) PSW's and (d) DPDW's] of  $F(\mathbf{K}, z)$  for different values of  $z$ : solid circles,  $z = \lambda$ ; open circles,  $z = 0.5\lambda$ ; stars,  $z = 0.1\lambda$ ; squares,  $z = 0.05\lambda$ .



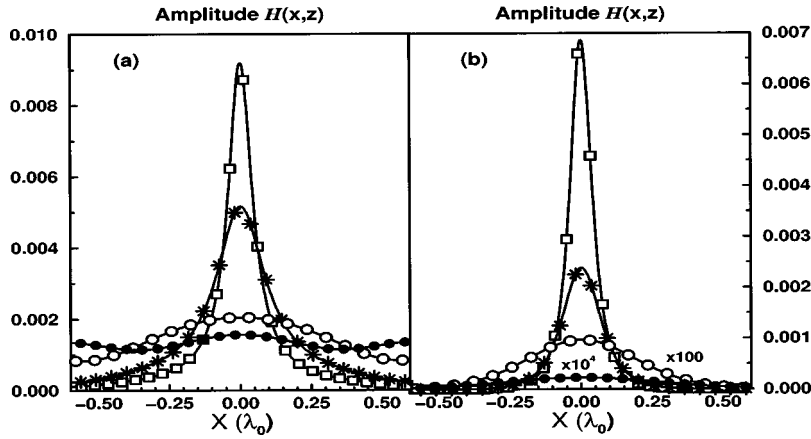


Fig. 4. Amplitude of  $H(\mathbf{R}, z)$  for (a) PSW's and (b) DPDW's for different values of  $z$ : solid circles,  $z = \lambda$ ; open circles,  $z = 0.5\lambda$ ; stars,  $z = 0.1\lambda$ ; squares,  $z = 0.05\lambda$ .

$$U(\mathbf{R}, z) = \int_{-\infty}^{+\infty} H(\mathbf{R} - \mathbf{R}', z) U(\mathbf{R}', z = 0) d\mathbf{R}'. \quad (5)$$

The FWHM of the amplitude of the impulse response yields the limit of spatial resolution in the wave field  $U(\mathbf{R}, z)$  on the plane  $z = \text{constant}$ . As an illustration, the amplitude of  $H(\mathbf{R}, z)$  for a two-dimensional configuration [ $\mathbf{R} = (x, 0)$ ] is shown in Fig. 4, at different values of  $z$ , for both DPDW's and PSW's. As  $z$  tends to zero,  $H$  tends to a delta function. It is also important that, at a given  $z$ , the width of this function is smaller for DPDW's than for PSW's. The consequences of this fact are discussed next.

The features of Figs. 3 and 4 can be further quantified in the following way. The FWHM of both the transfer function and the impulse response can be evaluated analytically. In the case of PSW's, the FWHM of the transfer function is obtained from the condition

$$|F(\mathbf{K}, z)| = \exp[-(|\mathbf{K}|^2 - \kappa_0^2 z)^{1/2}] = 1/2, \quad (6)$$

which, by taking logarithms on both sides, gives us

$$|\mathbf{K}|^2 - \kappa_0^2 = (\ln 2/z)^2. \quad (7)$$

It follows that the FWHM of  $F$  (denoted by  $\Delta|\mathbf{K}|$ ) and that of  $H$  (denoted by  $\Delta d$ ) are, respectively,

$$\Delta|\mathbf{K}| = 2[\kappa_0^2 + (\ln 2/z)^2]^{1/2}, \quad (8)$$

$$\Delta d/\lambda_0 = 1/2\{1 + [(\ln 2/2\pi z/\lambda_0)]^2\}^{-1/2}, \quad (9)$$

where we used the relationship between the FWHM of a function ( $H$  in our case) and that of its Fourier transform ( $F$ ):  $\Delta d = 2\pi/\Delta|\mathbf{K}|$ .

$\Delta d/\lambda_0$  is the spatial resolution limit in units of the wavelength. When  $z$  increases, we see in Eq. (9) that the spatial resolution limit tends to  $\lambda_0/2$ , that is, we retrieve the well-known Rayleigh limit, for  $z \gg \lambda_0$ , of optical imaging. In frequency space, Eq. (8) shows that as  $z$  increases,  $\Delta|\mathbf{K}|$  tends to  $2\kappa_0$ .

In the case of DPDW's, the value of  $\mathbf{K}$  that brings the normalized transfer function to its half-maximum is given by

$$\frac{|F(\mathbf{K}, z)|}{|F(\mathbf{K} = 0, z)|} = \exp\{-[q_i(\mathbf{K}) - \kappa_i]z\} = \frac{1}{2}, \quad (10)$$

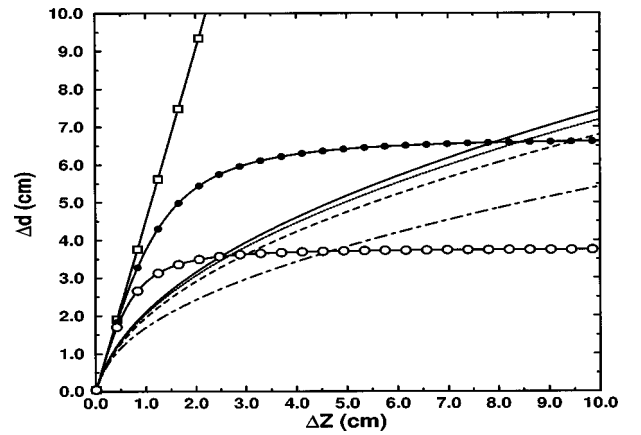


Fig. 5. Spatial resolution limit  $\Delta d$  in centimeters as  $z$  increases, for the following cases: DPDW's in breast tissue ( $\mu_a = 0.035 \text{ cm}^{-1}$ ,  $\mu_s' = 15 \text{ cm}^{-1}$ ): solid curve,  $\omega = 0$  [dc]; dotted curve,  $\omega = 100 \text{ MHz}$  ( $\lambda_0 = 13.38 \text{ cm}$ ); dashed curve,  $\omega = 200 \text{ MHz}$  ( $\lambda_0 = 7.53 \text{ cm}$ ); dotted-dashed curve,  $\omega = 300 \text{ MHz}$  ( $\lambda_0 = 5.60 \text{ cm}$ ). PSW's: squares, dc; solid circles,  $\lambda_0 = 13.38 \text{ cm}$ ; open circles,  $\lambda_0 = 7.53 \text{ cm}$ . In all cases  $n = 1.333$ .

so that, proceeding as with the PSW's, the FWHM of  $F$  and  $H$  are, respectively,

$$\Delta|\mathbf{K}| = 2 \left[ \left( \kappa_i + \frac{\ln 2}{z} \right)^2 + \kappa_r^2 - \kappa_i^2 - \frac{\kappa_i^2 \kappa_r^2}{(\kappa_i + \ln 2/z)^2} \right]^{1/2}, \quad (11)$$

$$\frac{\Delta d}{\lambda_0} = \frac{1}{2} \left[ \left( \frac{\lambda_0}{l_a} + \frac{\ln 2}{2\pi z/\lambda_0} \right)^2 - \left( \frac{\lambda_0}{l_a} \right)^2 + 1 - \left( 1 + \frac{\ln 2}{2\pi z/l_a} \right)^{-2} \right]^{-1/2}. \quad (12)$$

As  $z$  increases,  $z \gg \lambda_0$ , we see in Eq. (11) that  $\Delta|\mathbf{K}|$  tends to zero. Also, Eq. (12) shows that  $\Delta d/\lambda_0$  has no upper limit and tends to infinity as  $z$  increases, monotonically worsening the resolution. Hence, in contrast to the case of PSW's, all components of spatial frequencies  $\mathbf{K}$  always propagate into  $z > 0$ , even though attenuation of the signal exists in the whole  $\mathbf{K}$  range owing to diffusion.

The spatial resolution limit is seen in Fig. 5, which

shows the values of the spatial resolution limit both for DPDW's and PSW's versus the observation distance  $z$ . However, since in practical cases the values of  $z$  are small (of the order of a few centimeters), we find in Fig. 5 that in many cases the loss of resolution as  $z$  increases is less critical for DPDW's than for PSW's (compare in Fig. 5 the cases with  $\lambda = 7.53$  cm for PSW's and DPDW's in the interval  $[0, 3]$  cm, for example).

### C. Electrostatic Limit

Since DPDW's are damped waves, the detection of the wave field  $U(\mathbf{R}, z)$  is usually done in the near field, i.e., at subwavelength distance from the source object (considered either as a primary source or as a scattering object). In this range, if all distances involved are much smaller than the wavelength, retardation effects can be neglected. This property is well known, for example, in near-field optics.<sup>30</sup> When retardation effects are neglected, one is in the domain of the electrostatic limit.

At a given frequency  $\omega$ , this electrostatic limit is obtained when  $\lambda_0 \rightarrow \infty$ . Then, for PSW's, the limit of resolution within the electrostatic limit can be obtained from Eq. (9):

$$\Delta d = (\pi/\ln 2)z. \quad (13)$$

Whereas for DPDW's, we find from Eq. (12) that the resolution limit in the electrostatic limit is

$$\Delta d = \frac{1}{2} \left[ \left( \frac{1}{l_a} + \frac{\ln 2}{2\pi z} \right)^2 - \left( \frac{1}{l_a} \right)^2 \right]^{-1/2}. \quad (14)$$

It is interesting to note from Eqs. (13) and (14) that, in the electrostatic limit, resolution does not depend on the background medium for PSW's, whereas in the case of DPDW's, resolution still depends on the background medium, through the decay length  $l_a$ . In the limiting case, in which the absorption coefficient is negligible (i.e.,  $\mu_a \approx 0$ ), the expression  $\Delta d$  for DPDW's does not depend on the background medium, and we then obtain  $\Delta d \sim z$ , as

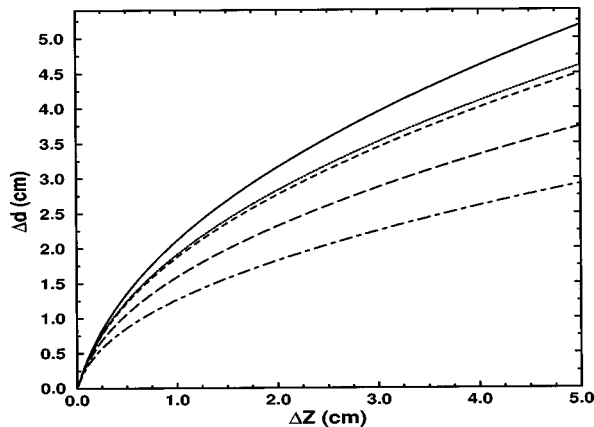


Fig. 6. DPDW's spatial resolution limit  $\Delta d$  in centimeters as we increment  $z$ , in dc regime ( $\omega = 0$ ), for the following cases: solid curve, breast parameters  $\mu_a = 0.035$  cm<sup>-1</sup>,  $\mu'_s = 15$  cm<sup>-1</sup>; dotted curve, abdomen parameters  $\mu_a = 0.09$  cm<sup>-1</sup>,  $\mu'_s = 9.5$  cm<sup>-1</sup>; short-dashed curve, back parameters  $\mu_a = 0.09$  cm<sup>-1</sup>,  $\mu'_s = 10.5$  cm<sup>-1</sup>; long-dashed curve, white matter  $\mu_a = 0.22$  cm<sup>-1</sup>,  $\mu'_s = 9.1$  cm<sup>-1</sup>; dotted-dashed curve, grey matter  $\mu_a = 0.27$  cm<sup>-1</sup>,  $\mu'_s = 20.6$  cm<sup>-1</sup>. In all cases  $n = 1.333$ .

previously found by Ref. 25 in the time domain. The transition to the electrostatic limit when  $z$  decreases can be clearly seen in Fig. 5, for both DPDW's and PSW's. As seen, at low  $z$ , one does no longer increase the resolution by lowering the incident wavelength, so that the behavior is similar to that observed with constant illumination, i.e., dc regime ( $\omega = 0$ ). However, it is important to note that, in the case of DPDW's, the electrostatic region extends to higher values of  $z$  and therefore is a good approximation even at modulation frequencies of the order of  $\omega = 100$  MHz. We also observe that, for modulation frequencies lower than 100 MHz, no increase in spatial resolution is obtained in ac, and therefore it is cheaper, and experimentally simpler, to perform measurements in dc if one is interested only in the location of the objects, that is, in obtaining an image. For modulation frequencies higher than 100 MHz we find that, even though the spatial resolution limit increases very quickly for such frequencies, the decay length  $l_a$  is considerably smaller and therefore the attenuation is much stronger, thus making detection at practical distances difficult.

The analysis presented here can be used to discuss experimental data, as those reported in Ref. 8. In Fig. 2 of Ref. 8, the authors characterize two diffusive objects with a relative diameter of  $\approx 0.1\lambda_0$ , 3.26 cm apart, embedded in a 0.75% Intralipid solution, illuminated by a modulated source of  $\omega = 100$  MHz (in this case  $\lambda_0 \approx 15.12$  cm and  $l_a \approx 7.7$  cm). If we take a look at our Fig. 5, approximating the diffusive parameters of the Intralipid solution to those of breast tissues, we see that two objects 3.26 cm apart can be resolved as long as we are measuring at distances  $z \leq 2$  cm. This resolution is what is observed in Fig. 2 of Ref. 8, in which the measurements are performed at a distance of 2 cm and are therefore within the limit of spatial resolution discussed above.

An important consequence of the existence of this electrostatic limit is that measurements in dc (i.e., at  $\omega = 0$ ), performed within the domain of validity of this limit, give the same spatial resolution as measurements in ac. To illustrate this point, we show in Fig. 6 the values of the spatial resolution limit  $\Delta d$  as  $z$  increases in several human tissues, corresponding to a dc illumination. These curves give the spatial resolution that can be reached at a given observation distance  $z$ , in each situation. If we once again refer to the situation depicted in Fig. 2 of Ref. 8, we see that at a distance of 2 cm, it is possible to resolve two objects 3.26 cm apart by measuring in dc (see the solid curve in Fig. 6). Thus we infer that in the case of Ref. 8, measurements in dc would have led to the same spatial resolution.

### 3. SCATTERING NUMERICAL RESULTS

To illustrate the discussion of Section 2 and to check the resolution criteria derived above, we now present rigorous numerical results on scattering of DPDW's. The geometry under consideration is two dimensional and depicted in Fig. 7. It consists in two diffusive infinite cylinders (the hidden objects), with axis along  $OY$ , both with radius  $R$  and separated by a distance  $d$ . The cylinders are embedded in an infinite, homogeneous, diffusive medium. Constant index of refraction  $n = 1.333$  is supposed

throughout all the media. A point source with modulation frequency  $\omega$  is located at  $\mathbf{r}_{\text{source}}$ , and the detection is performed in a plane  $z = z_{\text{detect}}$ . In this geometry we solve the diffusion equation without any approximation. The method uses a surface-integral formalism, which is an extension to diffusive media of the surface-integral method used in electromagnetic scattering from arbitrary interfaces.<sup>31–33</sup> The procedure is described in Ref. 34 in the case of diffusive scatterers in diffusive media and in Ref. 35 when index-mismatched interfaces are dealt with. This method consists basically in applying Green's theorem to the diffusion equation for the average intensity and to the corresponding equation for the Green function,

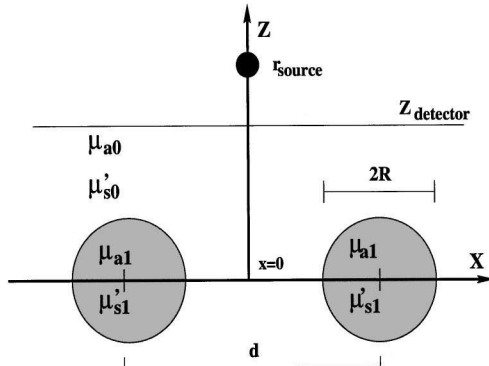


Fig. 7. Scattering geometry.

thus obtaining a closed set of coupled surface-integral equations. These integral equations are numerically solved without approximations, the numerical scheme being reduced to the solution of a linear system of equations. This method allows us to deal with multiple-scattering objects in interaction, and also with index-mismatch domains, although this case is not considered now (for a discussion on boundary conditions of DPDW's see Refs. 35 and 36, for example). In the following, we shall be interested in the amplitude of the scattered DPDW, defined by  $|U^{(\text{sc})}| = |U - U^{(\text{inc})}|$ , where  $U$  represents the total DPDW on interaction with the objects and  $U^{(\text{inc})}$  corresponds to the incident DPDW, namely, that created by the point source in the absence of the two objects.

### A. Diffuse-Wave Images of Two Hidden Objects

Following experimental procedures (see, for example, Refs. 37 and 38), we have considered a point source emitting light at a wavelength of 780 nm, with a modulation frequency  $\omega = 200$  MHz. The parameters chosen for the background medium correspond to breast tissue, with  $\mu_a = 0.035 \text{ cm}^{-1}$  and  $\mu_s' = 15 \text{ cm}^{-1}$ . For the cylinders, we have used the parameters of a breast tumor,  $\mu_a = 0.24 \text{ cm}^{-1}$  and  $\mu_s' = 10 \text{ cm}^{-1}$ . In all cases, the refractive index in the media is  $n = 1.333$ . To reach numerical convergence, owing to the small sizes of the cylinders un-

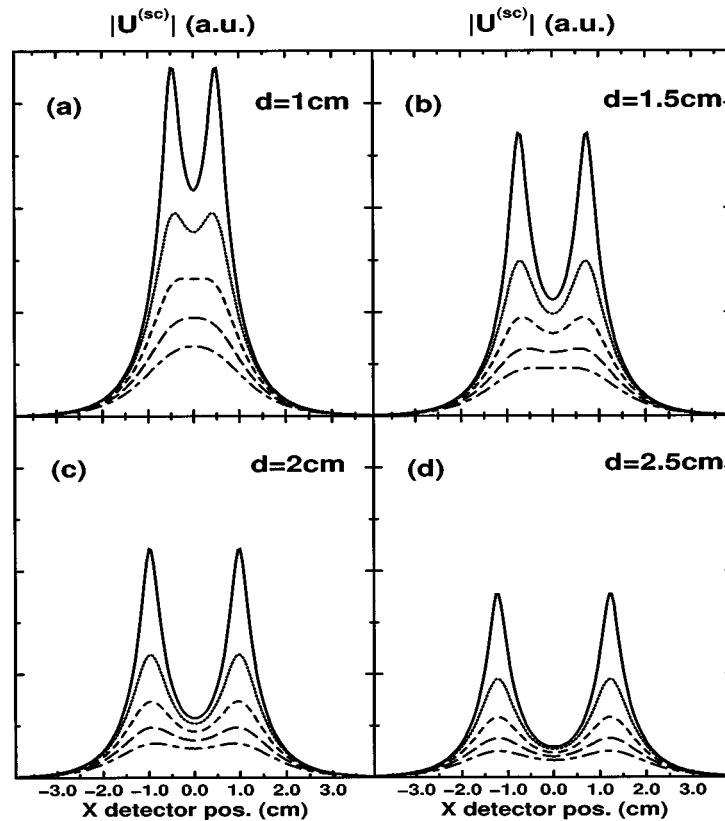


Fig. 8. Scattered amplitude corresponding to two diffuse cylinders of  $R = 0.1 \text{ cm}$  with breast tumor parameters  $\mu_a = 0.24 \text{ cm}^{-1}$ ,  $\mu_s' = 10 \text{ cm}^{-1}$ , embedded in breast tissue ( $\mu_a = 0.035 \text{ cm}^{-1}$ ,  $\mu_s' = 15 \text{ cm}^{-1}$ ), with the source located at  $\mathbf{r}_{\text{source}} = (0, 2.0 \text{ cm})$  with modulation frequency  $\omega = 200$  MHz, separated by distances (a)  $d = 1 \text{ cm}$ , (b)  $d = 1.5 \text{ cm}$ , (c)  $d = 2.0 \text{ cm}$ , (d)  $d = 2.5 \text{ cm}$  for the following  $Z$  detector distances: solid curve,  $z_{\text{detect}} = 0.2 \text{ cm}$ ; dotted curve,  $z_{\text{detect}} = 0.4 \text{ cm}$ ; short-dashed curve,  $z_{\text{detect}} = 0.6 \text{ cm}$ ; long-dashed curve,  $z_{\text{detect}} = 0.8 \text{ cm}$ ; dotted-dashed curve,  $z_{\text{detect}} = 1.0 \text{ cm}$ . In all cases  $n = 1.333$ .

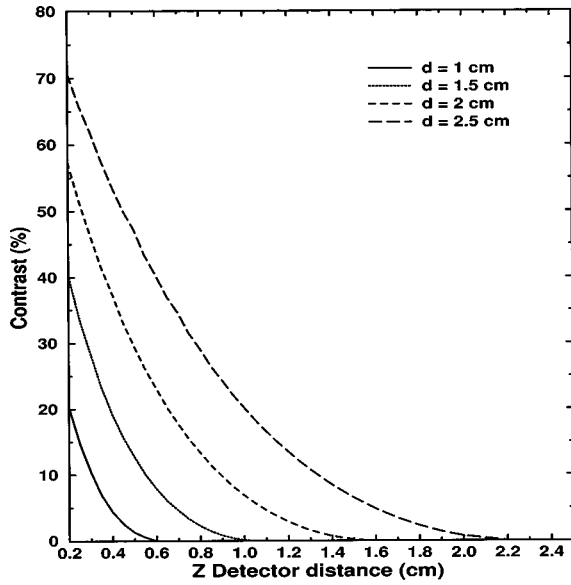


Fig. 9. Values of the noise-free contrast  $C_{\text{nf}}(\%)$  as we vary the detector-plane distance  $z_{\text{detect}}$  in the case of two cylinders of radius  $R = 0.1$  cm, with breast tumor parameters  $\mu_a = 0.24 \text{ cm}^{-1}$ ,  $\mu'_s = 10 \text{ cm}^{-1}$ , embedded in breast tissue ( $\mu_a = 0.035 \text{ cm}^{-1}$ ,  $\mu'_s = 15 \text{ cm}^{-1}$ ), with the source located at  $\mathbf{r}_{\text{source}} = (0, 2.0 \text{ cm})$  with modulation frequency  $\omega = 200$  MHz, separated by the following distances: solid curve,  $d = 1$  cm; dotted curve,  $d = 1.5$  cm; short-dashed curve,  $d = 2.0$  cm; long-dashed curve,  $d = 2.5$  cm. In all cases  $n = 1.333$ .

der study, we have used a discretization  $ds = 0.004$  cm for the surface of the cylinders.

Figure 8 shows the amplitude of the scattered DPDW,  $|U^{(\text{SC})}|$ , at different detection planes and for different cylinder distances, when the two diffusive cylinders have a radius  $R = 0.1$  cm. As expected, the farther apart the cylinders are from each other, the better they are resolved, and as we locate the detection plane farther away, this resolution power diminishes. To compare these results with the conclusions of the previous section, we refer again to Fig. 5, for breast tissue illuminated with a modulation frequency  $\omega = 200$  MHz (dashed curve). In the range of cylinder distances represented in Fig. 8, i.e., for  $d$  from 1 cm to 2.5 cm, Fig. 5 indicates that, to be within the spatial resolution limit, we must place the detection plane between 0.5 cm and 1.5 cm. This is precisely what is observed in Fig. 8. If we look at Fig. 8(a) for the case of  $z_{\text{detect}} = 1$  cm, we find that for a separation distance  $d = 1$  cm, the objects are not spatially resolved. In this ideal noiseless situation, the cylinders start being resolved at a detector-plane distance  $z_{\text{detect}} = 0.4$  cm [dotted curve in Fig. 8(a)].

Once data are above the spatial resolution limit, it is convenient to define another quantity that allows us to discriminate the image signal from a certain noise level present in the data, i.e., the contrast. It is also useful, in order to compare with previous definitions of contrast, to introduce first a noise-free contrast (nf), which we express in percentage (%) as

$$C_{\text{nf}}(\%) = \frac{|U_{\text{max}}^{(\text{SC})}| - |U_{\text{min}}^{(\text{SC})}|}{|U_{\text{max}}^{(\text{SC})}| + |U_{\text{min}}^{(\text{SC})}|} \times 100. \quad (15)$$

In Eq. (15)  $|U^{(\text{SC})}|$  is the noise-free scattered amplitude, and  $|U_{\text{min}}^{(\text{SC})}|$  is the minimum value of  $|U^{(\text{SC})}|$  between maxima  $|U_{\text{max}}^{(\text{SC})}|$ . In Fig. 9 we plot this noise-free contrast  $C_{\text{nf}}$  for two cylinders, in the same situation as in Fig. 8, versus the detection distance  $z_{\text{detect}}$ . The modulation frequency is 200 MHz. Results for different separation distances  $d$  of the two cylinders are shown. These curves give us the basis to deal with more realistic data, that is, with noise. Calculations for noise-free scattering data from cylinders with smaller radii  $R$  are not presented here because the resulting contrast curves are similar to those of Fig. 9. This similarity occurs because, to the extent that noise is not taken into consideration, the main effect of reducing the size of the scattering object is a decrease in the amplitude of the scattered wave, but the reduction has no effect in the resolution limit. An extensive study on the issue of noise can be found in Ref. 39. However, to derive useful consequences when Fig. 9 is applied to actual (noisy) experimental data, a contrast threshold must be introduced into Eq. (15) and, thus into the curves of this figure, below which no signal can be discriminated from the noise background. The effect of noise, and the threshold that it produces, is next addressed.

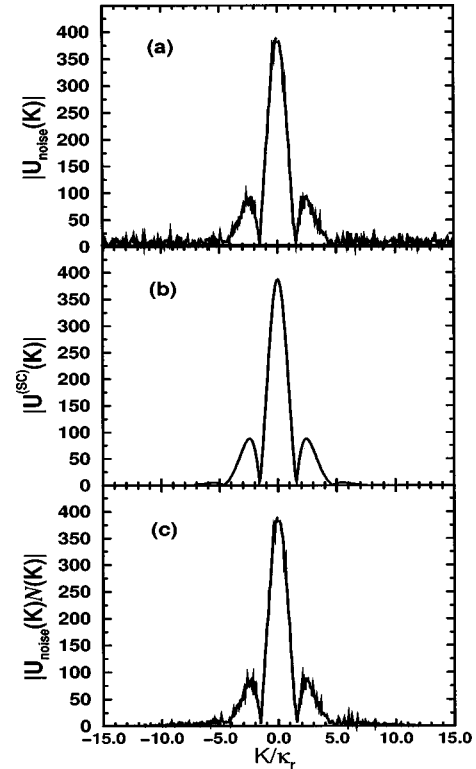


Fig. 10. Values of (a)  $\tilde{U}_{\text{noise}}(K, z = 1 \text{ cm})$ , (b)  $U^{(\text{SC})}(K, z = 1 \text{ cm})$ , (c)  $\tilde{U}_{\text{filt}}(K, z = 1 \text{ cm}) = U^{(\text{SC})}(K, z = 1 \text{ cm})N(K)$  for a detector-plane distance  $z_{\text{detect}} = 1$  cm in the case of two cylinders of radius  $R = 0.1$  cm, with breast tumor parameters  $\mu_a = 0.24 \text{ cm}^{-1}$ ,  $\mu'_s = 10 \text{ cm}^{-1}$ , embedded in breast tissue ( $\mu_a = 0.035 \text{ cm}^{-1}$ ,  $\mu'_s = 15 \text{ cm}^{-1}$ ), with the source located at  $\mathbf{r}_{\text{source}} = (0, 2.0 \text{ cm})$  with modulation frequency  $\omega = 200$  MHz, separated by a distance of  $d = 2.5$  cm.  $N(K)$  is a Hanning filter with  $K_{\text{cut}} = 15k_r$ . Noise parameters:  $\eta = 10\%$  and  $\sigma_\xi = 10^\circ$ . In all cases  $n = 1.333$ .

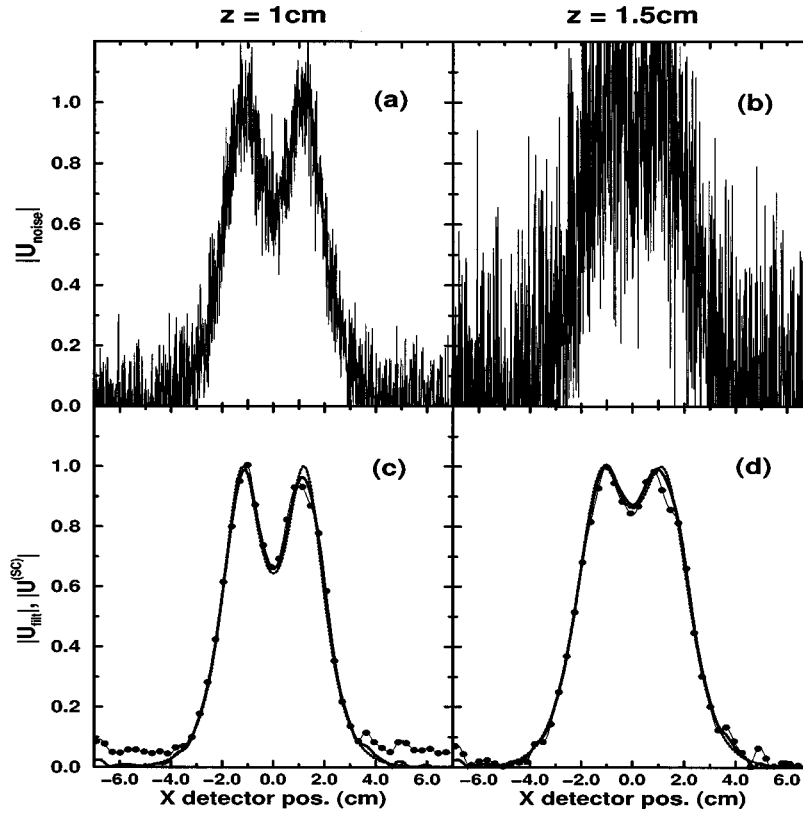


Fig. 11. Normalized scattered amplitude in the case of two cylinders of radius  $R = 0.1$  cm, with breast tumor parameters  $\mu_a = 0.24$  cm $^{-1}$ ,  $\mu'_s = 10$  cm $^{-1}$ , embedded in breast tissue ( $\mu_a = 0.035$  cm $^{-1}$ ,  $\mu'_s = 15$  cm $^{-1}$ ), with the source located at  $\mathbf{r}_{\text{source}} = (0, 2.0)$  cm with modulation frequency  $\omega = 200$  MHz, separated by a distance  $d = 2.5$  cm in the following cases. (a) Measured at a plane-detection distance of  $z_{\text{detect}} = 1$  cm with noise parameters  $\eta = 10\%$  and  $\sigma_\xi = 10^\circ$ . (b) Measured at a plane-detection distance of  $z_{\text{detect}} = 1.5$  cm with noise parameters  $\eta = 30\%$  and  $\sigma_\xi = 10^\circ$ . (c) Solid curve, after filtering by a Hanning filter with  $K_{\text{cut}} = 15\kappa_r$ , the image obtained in (a); dotted curve, direct measurement without noise at  $z_{\text{detect}} = 1$  cm; solid circle, after filtering by a Hanning filter with  $K_{\text{cut}} = 15\kappa_r$ , an image with noise parameters  $\eta = 30\%$  and  $\sigma_\xi = 10^\circ$ . (d) Solid curve, after filtering by a Hanning filter with  $K_{\text{cut}} = 15\kappa_r$ , an image with noise parameters  $\eta = 10\%$  and  $\sigma_\xi = 10^\circ$ ; dotted curve, direct measurement without noise at  $z_{\text{detect}} = 1.5$  cm; solid circle, after filtering by a Hanning filter with  $K_{\text{cut}} = 15\kappa_r$ , the image obtained in (b). In all cases  $n = 1.333$ .

## B. Effects of Noise on Resolution

Let  $U_{\text{noise}}$  be an image containing additive noise in amplitude  $N$  and in phase  $\xi$ . We express it in terms of the DPDW scattered by the object  $U^{(\text{SC})}$  as

$$U_{\text{noise}}(\mathbf{R}, z) = [|U^{(\text{SC})}(\mathbf{R}, z)| + N(\mathbf{R})] \times \exp\{i[\phi(\mathbf{R}, z) + \xi(\mathbf{R})]\}, \quad (16)$$

where  $\phi$  is the phase of the scattered DPDW. The random variables  $N(\mathbf{R})$  and  $\xi(\mathbf{R})$  are Gaussian distributed with correlation length  $T = 0$  (i.e., white noise) and root mean square  $\sigma_N$  and  $\sigma_\xi$ , respectively. The noise-to-signal ratio  $\eta(\%)$  of this image, in percentage, is introduced as

$$\eta(\%) = \sigma_N / |U_{\text{max}}^{(\text{SC})}| \times 100, \quad (17)$$

$|U_{\text{max}}^{(\text{SC})}|$  being the signal peak amplitude. Then we define the contrast  $\mathcal{C}(\%)$  in the presence of noise as

$$\mathcal{C}(\%) = \mathcal{C}_{\text{nf}}(\%) - \eta(\%). \quad (18)$$

With this definition of contrast,  $\eta$  is the aforementioned contrast threshold, as shown in Fig. 9. That is, resolving two objects requires placing the detector at such a distance  $z$  that the contrast remains above  $\eta$ . Once the largest detection distance  $z$  that yields a given resolution

limit  $\Delta d$  has been derived from Fig. 5, the maximum data uncertainty  $\eta$  that allows observation of signal contrast, and hence details with this resolution  $\Delta d$ , can be found from Fig. 9.

To estimate the scattered signal from the noisy data  $U_{\text{noise}}$ , the signal's high-frequency components are filtered out:

$$U_{\text{filt}}(\mathbf{R}, z) = \int_{-\mathbf{K}_{\text{cut}}}^{+\mathbf{K}_{\text{cut}}} \tilde{U}_{\text{noise}}(\mathbf{K}, z) \mathcal{N}(\mathbf{K}) \exp(i\mathbf{K} \cdot \mathbf{R}) d\mathbf{K}, \quad (19)$$

where  $U_{\text{filt}}$  is the filtered image and  $\tilde{U}_{\text{noise}}(\mathbf{K}, z)$  is the  $\mathbf{K}$ -Fourier transform of  $U_{\text{noise}}(\mathbf{R}, z)$ .  $\mathbf{K}_{\text{cut}}$  is the cut-off frequency.  $\mathcal{N}$  denotes a low-pass filter. In our computer simulations, this is a Hanning filter, which we define as

$$\mathcal{N}(\mathbf{K}) = \left[ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{K_x}{K_{\text{cut},x}} \pi\right) \right] \cdot \left[ \frac{1}{2} + \frac{1}{2} \cos\left(\frac{K_y}{K_{\text{cut},y}} \pi\right) \right], \quad (20)$$

bearing in mind that in the case of a two-dimensional configuration  $\mathbf{K} = (K, 0)$ ,  $\mathbf{K}_{\text{cut}} = (K_{\text{cut}}, 0)$ .

Owing to high damping and low reflectivity, DPDWs are not subjected to strong interference processes as

PSW's are. Therefore the scattering and diffraction patterns of DPDW's do not present high-frequency interference fringes, which means that the Fourier spectrum of DPDW's is mostly concentrated at frequencies  $\mathbf{K} \leq \mathbf{K}_{\text{cut}}$ ; thus the filtering neither substantially alters the image nor appreciably reduces resolution. Then, after filtering, one can estimate that  $C_{\text{filt}}(\%) \approx C_{\text{nf}}(\%)$ ,  $C_{\text{filt}}$  being the contrast of  $|U_{\text{filt}}^{(\text{SC})}|$ . This can be seen in Fig. 10 in which we plot the values of  $\tilde{U}_{\text{noise}}(\mathbf{K})$ ,  $\tilde{U}(\mathbf{K}, z)$ , and  $\tilde{U}_{\text{filt}}(\mathbf{K}, z)$ , where  $\tilde{U}_{\text{filt}}(\mathbf{K}, z) = \tilde{U}_{\text{noise}}(\mathbf{K}, z)\mathcal{N}(\mathbf{K})$  for  $z = 1$  cm and  $K_{\text{cut}} = 15\kappa_r$ . These quantities are computed for data from two diffusive cylinders with the same parameters as in Fig. 8, radius  $R = 0.1$  cm, separated a distance  $d = 2.5$  cm, with the detection plane at  $z_{\text{detect}} = 1.0$  cm. A numerical noise has been added to the scattered field, as described by Eq. (16). The noise amplitude  $N(\mathbf{R})$  has a ratio  $\eta = 10\%$ , and the phase noise has  $\sigma_\xi = 10^\circ$ . A comparison of Figs. 10(b) and 10(c) confirms that this filtering does not appreciably remove information in the spectrum, as we believe should be the case in most practical situations with DPDW's. The corresponding scattering amplitudes in real space, both before and after filtering, are shown in Fig. 11.

As regards the simulation of Fig. 11, notice that, according to Fig. 9, the contrast that we can expect in data taken at  $z = 1$  cm [Fig. 11(a)] and  $z = 1.5$  cm [Fig. 11(b)], for a cylinder separation distance  $d = 2.5$  cm, is approximately 20% and 10%, respectively. In Fig. 11(a) we have  $\eta = 10\%$ , and therefore we still have 10% of signal contrast above the threshold for  $z = 1$  cm. However, in Fig. 11(b) we have considered a rather extreme situation in which  $\eta = 30\%$ , which in the  $z = 1.5$  cm case places the contrast under the threshold. Even so, once the image is filtered, we find that in the  $z = 1$  cm case [Fig. 11(c)], which is above the threshold, the filtered image is very close to the noise-free image. Surprisingly, this outcome occurs also for the case  $z = 1.5$  cm [Fig. 11(d)], which is under the threshold for both values of  $\eta$ , i.e.,  $\eta = 10\%$  and  $\eta = 30\%$ . That is, even in very unfavorable signal detections, we find no effective threshold for the filtered image amplitude.

### C. Backpropagation

Let us now see the effect of the characteristics of the detected signal discussed so far, on the wave field close to the scattering objects. To reconstruct this wave front, we

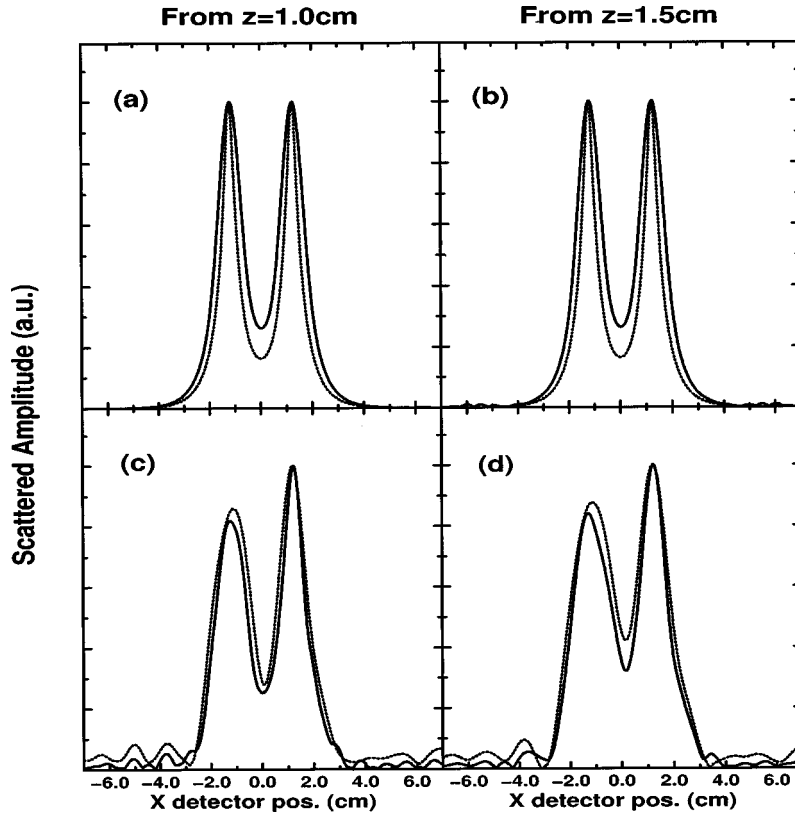


Fig. 12. Normalized scattered amplitude backpropagated onto  $z = 0.2$  cm in the case of two cylinders of radius  $R = 0.1$  cm, with breast tumor parameters ( $\mu_a = 0.24 \text{ cm}^{-1}$ ,  $\mu'_s = 10 \text{ cm}^{-1}$ ), embedded in breast tissue ( $\mu_a = 0.035 \text{ cm}^{-1}$ ,  $\mu'_s = 15 \text{ cm}^{-1}$ ), with the source located at  $\mathbf{r}_{\text{source}} = (0, 2.0 \text{ cm})$  with modulation frequency  $\omega = 200$  MHz, separated by a distance  $d = 2.5$  cm, for the following. (a) Solid line, noise-free image taken at  $z_{\text{detect}} = 1.0$  cm backpropagated with  $K_{\text{cut}} = 10\kappa_r$ ; dotted curve, direct measurement at  $z_{\text{detect}} = 0.2$  cm. (b) Solid curve, noise-free image taken at  $z_{\text{detect}} = 1.5$  cm backpropagated with  $K_{\text{cut}} = 10\kappa_r$ ; dotted curve, direct measurement at  $z_{\text{detect}} = 0.2$  cm. (c) Solid curve, image taken at  $z_{\text{detect}} = 1.0$  cm with noise parameters  $\sigma_\xi = 10^\circ$ ,  $\eta = 10\%$  backpropagated with  $K_{\text{cut}} = 10\kappa_r$ ; dotted curve, image taken at  $z_{\text{detect}} = 1.0$  cm with noise parameters  $\sigma_\xi = 10^\circ$ ,  $\eta = 30\%$  backpropagated with  $K_{\text{cut}} = 8\kappa_r$ . (d) Solid curve, image taken at  $z_{\text{detect}} = 1.5$  cm with noise parameters  $\sigma_\xi = 10^\circ$ ,  $\eta = 10\%$  backpropagated with  $K_{\text{cut}} = 7\kappa_r$ ; dotted curve, image taken at  $z_{\text{detect}} = 1.5$  cm with noise parameters  $\sigma_\xi = 10^\circ$ ,  $\eta = 30\%$  backpropagated with  $K_{\text{cut}} = 6\kappa_r$ . In all cases  $n = 1.333$ .

now backpropagate the image from the detection plane, which is done by means of Eqs. (2) and (3). A detailed description of this procedure can be found in Ref. 7. To carry out the backpropagation, we have filtered the backpropagated image in  $\mathbf{K}$  space once again, by means of a Hanning filter.

In Fig. 12 we show the backpropagated amplitudes from the detection planes  $z = 1$  cm and  $z = 1.5$  cm, onto the plane  $z = 0.2$  cm. This backpropagation is performed for all cases shown in Fig. 11, i.e., noise-free and filtered images. As stated above, the noise parameters for data shown in Figs. 12(c) and 12(d) are  $\sigma_\xi = 10^\circ$ ,  $\eta = 10\%$  and  $\sigma_\xi = 10^\circ$ ,  $\eta = 30\%$ , respectively. As shown in Fig. 12, and as already mentioned before, since the diffraction patterns from the scattered waves at  $z = 1$  cm and  $z = 1.5$  cm do not present appreciable interference fringes, the backpropagation in these cases basically constitutes an increase in contrast. Once again, we can see in Fig. 12 that the backpropagated filtered images corresponding to  $\eta = 30\%$  are approximately the same as those corresponding to  $\eta = 10\%$ . Also, since the image taken directly at  $z = 0.2$  cm does not have a high-frequency contribution, the backpropagated image is a very good approximation to this image [see Figs. 11(a) and 11(b)]. The reason for this is that the frequency cut for the filter does not have to be very high for retrieving information from the reconstructed wave. The asymmetry of the backpropagated noisy images with respect to  $x = 0$  is due to residual noise in the filtered images at the detection  $z$  plane [cf. Figs. 11(c) and 11(d)]. We wish to emphasize that we have not undertaken any additional processing of these data to filter out this effect. In practice, however, an averaging over several image recordings, together with any standard apodization procedure on these images before the backpropagation operation, can still considerably improve the results shown in Figs. 12(c) and 12(d). We do not pursue these aspects any further, since these are accessory to the main purpose of the present work.

#### 4. CONCLUSION

In this paper we have addressed the variation of the transfer function and the impulse response on propagation of DPDW's and have compared it with the known case of PSW's. We have put forward an analytical expression for the spatial limit of resolution of DPDW's, which is given by the FWHM of the impulse response. This spatial resolution has been studied versus the propagation distance from the scattering object, and its electrostatic limit has been discussed for DPDW's on the basis of near-field optics considerations. To illustrate the use of this limit of spatial resolution, we have presented an exact numerical computation, within the diffusion approximation, of the scattering of DPDW's from two cylinders with breast tumor parameters embedded in breast tissue. The effect of noise on the resolution contained in these data has been included.

The issue of whether it is more convenient to measure DPDW's in ac or in dc has been discussed as regards this spatial resolution limit. We demonstrate that in many practical cases, measuring in ac does not increase spatial

resolution, and therefore with spatial resolution (i.e., obtaining an image) one can take advantage of the cheaper and experimentally simpler measurements in dc. Once again we must state that if the main concern is to extract the diffusive properties of the objects, then ac or time-resolved measurements must be performed, as shown in Ref. 26. Detection of DPDW's is performed in the near field, and we show that in the majority of cases the contribution of the electrostatic limit dominates. On the basis of these results, we recommend the employment of this resolution-limit expression as a guide before undertaking experimental measurements. When measuring in dc, we obtain a lower value of the noise-to-signal ratio. This value increases as the modulation frequency grows. Therefore we conclude that, in most cases, measuring in dc not only does not reduce the spatial resolution but diminishes the noise-to-signal ratio, thus making detection of smaller objects more feasible.

The theory presented here constitutes a rigorous mathematical formulation that makes possible the understanding of the information content on propagation of DPDW's and underlines the interpretation of backpropagation results such as, for example, those of Ref. 7. Specifically, we have found that, owing to the high-damping property of DPDW's, their spectra is concentrated in the low-spatial-frequency range. Therefore considerably high noise levels can be filtered out with minimum loss of information, that is, of resolving power. We have shown that this is true even when the detected image contrast is under the threshold imposed by noise. That is, after filtering, we find no effective contrast threshold for the DPDW data.

#### ACKNOWLEDGMENTS

This research has been supported by Comisión Interministerial de Ciencia y Tecnología of Spain under grant PB95-0061 and by the Fundación Ramón Areces. J. Ripoll acknowledges a scholarship from Ministerio de Educación y Cultura of Spain. R. Carminati acknowledges the European Union for a postdoctoral grant in Madrid where part of this work was done.

The authors can be reached at the addresses on the title page or by e-mail, jripoll@everest.icmm.csic.es and remi@em2c.ecp.fr.

#### REFERENCES

1. A. Yodh and B. Chance, "Spectroscopy and imaging with diffusing light," *Phys. Today* **48**, 34 (1995), and references therein.
2. E. B. de Haller, "Time-resolved transillumination and optical tomography," *J. Biomed. Opt.* **1**, 7-17 (1996).
3. See related studies in *Advances in Optical Imaging and Photon Migration*, R. R. Alfano and J. G. Fujimoto, eds., Vol. 2 of OSA Trends in Optics and Photonics Series (Optical Society of America, Washington, D.C., 1996).
4. S. R. Arridge, P. van der Zee, M. Cope, and D. T. Delpy, "Reconstruction methods for near infra-red absorption imaging," in *Time-Resolved Spectroscopy and Imaging of Tissues*, B. Chance and A. Katzir, eds., *Proc. SPIE* **1431**, 204-215 (1991).
5. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Experimental images of heterogeneous turbid media by

- frequency-domain diffusing-photon tomography," *Opt. Lett.* **20**, 426–428 (1995).
6. C. P. Gonatas, M. Ishii, J. S. Leigh, and J. C. Schotland, "Optical diffusion imaging using a direct inversion method," *Phys. Rev. E* **52**, 4361–4365 (1995).
  7. C. L. Matson, N. Clark, L. McMackin, and J. S. Fender, "Three-dimensional tumor localization in thick tissue with the use of diffuse photon-density waves," *Appl. Opt.* **36**, 214–220 (1997).
  8. X. D. Li, T. Durduran, A. G. Yodh, B. Chance, and D. N. Pattanayak, "Diffraction tomography for biochemical imaging with diffuse-photon density waves," *Opt. Lett.* **22**, 573–575 (1997).
  9. H. Wabnitz and H. Rinneberg, "Imaging in turbid media by photon density waves: spatial resolution and scaling relations," *Appl. Opt.* **36**, 64–74 (1997).
  10. Y. Yao, Y. Wang, Y. Pei, W. Zhu, and R. L. Barbour, "Frequency-domain optical imaging of absorption and scattering distributions by a Born iterative method," *J. Opt. Soc. Am. A* **14**, 325–342 (1997).
  11. S. A. Walker, S. Fantini, and E. Gratton, "Image reconstruction by backprojection from frequency-domain optical measurements in highly scattering media," *Appl. Opt.* **36**, 170–179 (1997).
  12. H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Simultaneous reconstruction of optical absorption and scattering maps in turbid media from near-infrared frequency-domain data," *Opt. Lett.* **20**, 2128–2130 (1995).
  13. S. B. Colak, D. G. Papaioannou, G. W. 't Hooft, M. B. van der Mark, H. Schomberg, J. C. J. Paasschens, J. B. M. Mellissen, and N. A. A. J. van Asten, "Tomographic image reconstruction from optical projections in light-diffusing media," *Appl. Opt.* **36**, 180–213 (1997).
  14. P. N. den Outer, T. M. Nieuwenhuizen, and A. Lagendijk, "Location of objects in multiple-scattering media," *J. Opt. Soc. Am. A* **10**, 1209–1218 (1993).
  15. S. Feng, F. Zeng, and B. Chance, "Photon migration in the presence of a single defect: a perturbation analysis," *Appl. Opt.* **35**, 3826–3837 (1995).
  16. S. R. Arridge and J. C. Hebden, "Optical imaging in medicine. II. Modeling and reconstruction," *Phys. Med. Biol.* **42**, 841–853 (1997).
  17. J. C. Schotland, "Continuous-wave diffusion imaging," *J. Opt. Soc. Am. A* **14**, 275–279 (1997).
  18. S. R. Arridge and M. Schweiger, "A gradient-based optimization scheme for optical tomography," *Opt. Express* **2**, 213–226 (1998).
  19. S. J. Norton and T. Vo-Dinh, "Diffraction tomographic imaging with photon density waves: an explicit solution," *J. Opt. Soc. Am. A* **15**, 2670–2677 (1998).
  20. S. A. Walker, D. A. Boas, and E. Gratton, "Photon density waves scattered from cylindrical inhomogeneities: theory and experiments," *Appl. Opt.* **37**, 1935–1944 (1998).
  21. S. Fantini, S. A. Walker, M. A. Franceschini, M. Kaschke, P. M. Schlag, and K. T. Moesta, "Assessment of the size, position, and optical properties of breast tumor *in vivo* by non-invasive optical methods," *Appl. Opt.* **37**, 1982–1989 (1998).
  22. R. Cubeddu, A. Pifferi, P. Taroni, A. Torricelli, and G. Valentini, "Imaging with diffusing light: an experimental study of the effect of background optical properties," *Appl. Opt.* **37**, 3564–3573 (1998).
  23. B. DeBecker, A. Bulatov, and J. L. Birman, "Two-dimensional inverse problem of diffusion tomography: the approach applicable for small inclusions," *Appl. Opt.* **37**, 4294–4299 (1998).
  24. M. Born and E. Wolf, *Principles of Optics*, 6th ed. (Pergamon, New York, 1993).
  25. J. A. Moon, R. Mahon, M. D. Duncan, and J. Reintjes, "Resolution limits for imaging through turbid media with diffuse light," *Opt. Lett.* **18**, 1591–1593 (1993).
  26. S. R. Arridge and W. R. B. Lionheart, "Nonuniqueness in diffusion-bases optical tomography," *Opt. Lett.* **23**, 882–884 (1998).
  27. J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, New York, 1968), Chap. 3.
  28. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge U. Press, Cambridge, England, 1995), Chap. 3.
  29. M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Wiley—Interscience, New York, 1991), Chap. 2.
  30. For near-field optical methods see, e.g., D. W. Pohl and D. Courjon, eds., *Near Field Optics* (Kluwer Academic, Dordrecht, The Netherlands, 1993) or M. Nieto-Vesperinas and N. Garcia, eds., *Optics at the Nanometer Scale* (Kluwer Academic, Dordrecht, The Netherlands, 1996). For a recent review on the subject see J. J. Greffet and R. Carminati, *Image Formation in Near Field Optics*, *Prog. Surf. Sci.* **56**, 133 (1997).
  31. J. A. Sánchez-Gil and M. Nieto-Vesperinas, "Light scattering from random rough dielectric surfaces," *J. Opt. Soc. Am. A* **8**, 1270–1286 (1991).
  32. A. Madrazo and M. Nieto-Vesperinas, "Scattering of light and other electromagnetic waves from a body buried beneath a highly rough random surface," *J. Opt. Soc. Am. A* **14**, 1859–1866 (1997).
  33. J. Ripoll, A. Madrazo, and M. Nieto-Vesperinas, "Scattering of electromagnetic waves from a body over a random rough surface," *Opt. Commun.* **142**, 173–178 (1997).
  34. J. Ripoll and M. Nieto-Vesperinas, "Scattering integral equations for diffusive waves. Detection of objects buried in diffusive media in the presence of rough interfaces," *J. Opt. Soc. Am. A* (to be published).
  35. J. Ripoll and M. Nieto-Vesperinas, "Index mismatch for diffuse photon density waves at both flat and rough diffuse-diffuse interfaces," *J. Opt. Soc. Am. A* (to be published).
  36. See, for example, R. Aronson, "Boundary conditions for diffusion of light," *J. Opt. Soc. Am. A* **12**, 2532–2539 (1995).
  37. D. A. Boas, M. A. O'Leary, B. Chance, and A. G. Yodh, "Scattering of diffuse photon density waves by spherical inhomogeneities within turbid media: analytic solution and applications," *Proc. Natl. Acad. Sci. USA* **19**, 4887–4891 (1994).
  38. M. A. O'Leary, D. A. Boas, B. Chance, and A. G. Yodh, "Refraction of diffuse photon density waves," *Phys. Rev. Lett.* **69**, 2658–2661 (1992).
  39. D. A. Boas, M. A. O'Leary, B. Chance, and A. G. Yodh, "Detection and characterization of optical inhomogeneities with diffuse photon density waves: a signal-to-noise analysis," *Appl. Opt.* **36**, 75–92 (1997).



# Scattering of a diffusive wave by a subsurface object

J.-B. Thibaud, R. Carminati,<sup>a)</sup> and J.-J. Greffet

*Laboratoire d'Énergétique Moléculaire et Macroscopique, Combustion; Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

(Received 13 October 1999; accepted for publication 6 March 2000)

We present a theoretical and numerical study of the scattering of a diffusive wave by an object embedded in a semi-infinite substrate. We derive exact integral equations for the scattered wave, using Green's theorem and appropriate Green's functions. We show that two methods can be used, leading either to a purely surface-integral formalism or to a formalism involving a volume integral and a surface term. In the first case, we derive an extinction theorem for diffusive waves and present an efficient numerical procedure to solve exactly the scattering problem. In the second formalism, physically more explicit, we apply an improved Born approximation, and study its range of validity by comparison with rigorous numerical results. Our approach also suggests a simple way to determine the depth of the object. In this article, we focus on thermal waves. Yet the formalism can be applied to photon-density waves. © 2000 American Institute of Physics.

[S0021-8979(00)08811-3]

## I. INTRODUCTION

Thermal wave scattering has received a lot of attention in the past 20 years as a tool for noninvasive detection and imaging of defects buried in the bulk of an opaque material.<sup>1</sup> It has applications in numerous fields, such as microelectronics, material process control, or environmental monitoring.<sup>2</sup> Thermal waves are usually generated by irradiating the surface of an absorbing material with a light beam of harmonically modulated intensity. The absorption of light in a thin layer at the surface generates a local temperature fluctuation that diffuses into the bulk. Imaging techniques rely on measurements of the surface temperature field to detect subsurface heterogeneities. For the modulated part of the temperature field, the heat diffusion equation takes the form of the Helmholtz equation with a complex wave number.<sup>3</sup> From a theoretical point of view, thermal waves are very analogous to diffuse photon-density waves (DPDW) in turbid media that have received increasing attention in the past few years.<sup>4</sup> In the frequency domain and under the diffusion approximation, the photon-density transport in such media is also governed by the Helmholtz equation with a complex wave vector.<sup>5,6</sup> Although the present work refers to thermal waves, its formalism can easily be extended to DPDW.<sup>7</sup> Studies of thermal waves and DPDW have the same goal of detecting and imaging objects hidden in an opaque or turbid environment. Solving the direct problem is essential in this context, either for direct imaging or for inverse scattering, since any inversion method is based on a particular formulation of the direct problem.<sup>8-10</sup>

Apart from the case of one-dimensional (1D) structures where an analytical solution can be derived, the direct problem for thermal wave scattering has to be solved numerically and no exact solution for arbitrary scatterers has been pro-

posed yet. In the bustling field of DPDW, the direct problem has been addressed with various methods like finite elements<sup>8,11</sup> and Green's theorem.<sup>12</sup> The extinction theorem leads to a surface integral formalism worth noting, which is suitable for the exact numerical solution of the direct problem.<sup>7</sup> Green's techniques in the frequency domain have also been applied to thermal wave scattering.<sup>3,13,14</sup> More recently, integral equations have been used to derive perturbative solutions based on the Born approximation.<sup>15</sup> This approximation is valid when the scattered field is a small correction to the incident field and leads to a simple inversion scheme by deconvolution.<sup>16</sup> However, because of the discontinuity of the gradient of the temperature field on the boundary of the object, a difficulty arises in properly approximating the normal derivative of the temperature field.

The present article is devoted to the presentation of an efficient, exact numerical solution of the heat-diffusion equation in the frequency domain. Surface integrals are solved by a moment method inherited from electromagnetics<sup>17</sup> and provide an exact solution of thermal wave scattering by an arbitrary object. The article is organized as follows: Secs. II and III are devoted to the derivation of surface-integral equations for the temperature field and to the description of the numerical procedure. For numerical reasons, we have described and implemented this procedure in a two-dimensional (2D) geometry, but we provide equations to solve the three-dimensional (3D) problem as well. An alternative formulation involving both volume and surface integrals is outlined in Sec. IV. Although less convenient for exact numerical solution, this formulation allows us to differentiate the role of thermal conductivity variations from that of heat capacity variations in the scattering process and is well suited for the Born approximation. Special care is given to correctly evaluate the dominant term in the field normal derivative. Numerical simulations in Sec. V are used to validate the exact solution and to check the range of validity of the Born approximation. Simplistic analytical crite-

<sup>a)</sup>Author to whom correspondence should be addressed; electronic mail: remi@em2c.ecp.fr

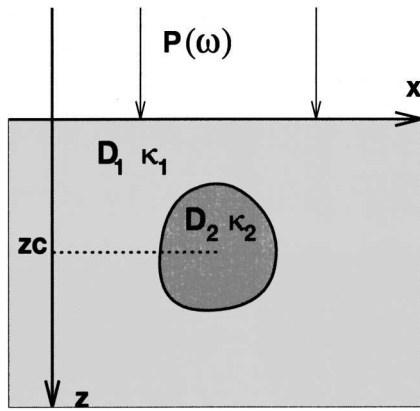


FIG. 1. Geometry of the system. The substrate (object) has a thermal conductivity  $\kappa_1$  ( $\kappa_2$ ) and a thermal diffusivity  $D_1$  ( $D_2$ ). The interface  $z=0$  is heated by a plane wave with intensity  $P$  modulated at a frequency  $\omega$ .

ria of validity of this approximation are also considered. We finally obtain a simple inversion scheme to estimate the depth of buried objects.

## II. STATEMENT OF THE PROBLEM

We consider a flat interface separating a vacuum (medium  $z < 0$ ) from a semi-infinite substrate (medium  $z > 0$ ) with uniform thermal conductivity  $\kappa_1$  and diffusivity  $D_1$  (see Fig. 1). An object of arbitrary shape, with uniform thermal conductivity  $\kappa_2$  and diffusivity  $D_2$ , is embedded in the substrate (subsurface object). This system is heated by a laser beam illuminating the interface  $z=0$ , the beam intensity  $P$  being modulated at a frequency  $\omega$ . The total temperature field at a given point  $\mathbf{r}=(x,y,z)$  with  $z > 0$  can be written

$$T_{\text{tot}}(\mathbf{r}, t) = T_0 + \text{Re}[T(\mathbf{r})\exp(-i\omega t)], \quad (1)$$

where  $T_0$  is a uniform background and  $T(\mathbf{r})$  is the complex amplitude of the temperature fluctuation induced by the modulated heating of the surface.

In the half-space  $z > 0$ , the temperature field  $T(\mathbf{r})$  obeys the partial differential equation

$$\nabla^2 T_j(\mathbf{r}) + k_j^2 T_j(\mathbf{r}) = -\frac{Q(\mathbf{r})}{\kappa_j}, \quad (2)$$

where  $k_j^2 = i\omega/D_j$ ,  $j=(1,2)$ , and  $T_j$  denotes the temperature field evaluated at a point inside the substrate ( $j=1$ ) or inside the object ( $j=2$ ).  $Q$  is the volume heat source produced by the absorption of the laser beam.

Solving Eq. (2) for  $j=(1,2)$  requires the specification of four boundary conditions. The temperature field and the heat flux are continuous at the surface  $\Sigma_2$  of the buried object. For  $\mathbf{r}$  on  $\Sigma_2$

$$T_1(\mathbf{r}) = T_2(\mathbf{r}), \quad (3)$$

$$\kappa_1 \frac{\partial T_1}{\partial n}(\mathbf{r}) = \kappa_2 \frac{\partial T_2}{\partial n}(\mathbf{r}), \quad (4)$$

where  $\partial/\partial n = \mathbf{n} \cdot \nabla$ ,  $\mathbf{n}$  being the outward or inward normal to the surface. Moreover, the temperature field is unperturbed for  $|\mathbf{r}| \rightarrow \infty$  inside the medium

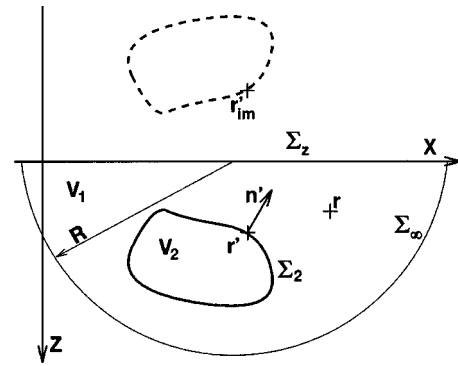


FIG. 2. Notations and surfaces used for the application of Green's theorem.

$$T_1(|\mathbf{r}| \rightarrow \infty) = 0. \quad (5)$$

The heat flux is also continuous at the interface  $z=0$ , and we assume that (1) the radiative and convective losses at  $z=0^-$  and (2) the conductive flux at  $z=0^+$  are negligible compared to the laser flux. This additional condition is written

$$\frac{\partial T_1}{\partial z}(x,y,z=0) = 0. \quad (6)$$

Using Green's theorem and appropriate Green's functions, we will now transform Eq. (2), together with the boundary conditions Eqs. (3)–(6) into a set of integral equations, involving either surface integrals (Sec. III), or a surface integral and a volume integral (Sec. IV). We will see that the first formulation is suitable for numerical calculations, whereas the second approach is helpful for a physical analysis of the scattering problem. We note that the problem defined by Eqs. (2)–(6) is very similar to the problem of photon-density waves. It suffices to replace the temperature  $T$  by the density of photons  $u$ ; the thermal conductivity  $\kappa$  by the inverse of the transport length  $l^*$ ; and the thermal diffusivity  $D$  by the photon-diffusivity  $cl^*/3$ , where  $c$  is the speed of light.

## III. SURFACE INTEGRAL FORMALISM: EXTINCTION THEOREM FOR THERMAL WAVES

### A. Green's theorem

Let  $u(\mathbf{r})$  and  $v(\mathbf{r})$  be arbitrary scalar fields defined in a volume  $V$  bounded by a closed surface  $\Sigma$ . Green's second identity states that

$$\int_V (u \nabla^2 v - v \nabla^2 u) d^3 \mathbf{r} = \int_\Sigma \left( u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) d^2 \mathbf{r}, \quad (7)$$

where the normal  $\mathbf{n}$  is chosen to be *outward* from the volume  $V$ .

With reference to Fig. 2, we denote by  $V_2$  and  $\Sigma_2$  the volume and the surface of the object, respectively. The complementary volume, enclosed by the plane  $z=0$  (denoted by  $\Sigma_z$ ), the surface  $\Sigma_2$  and an hemisphere  $\Sigma_\infty$  of infinite radius is denoted by  $V_1$ . We will apply Green's theorem in  $V_1$  and  $V_2$  successively, using two different Green's functions  $G_1$  and  $G_2$ . Let us first choose  $G_1$  obeying

$$\nabla^2 G_1(\mathbf{r}, \mathbf{r}') + k_1^2 G_1(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}') \quad (8)$$

and satisfying the boundary conditions Eqs. (5) and (6), namely,  $G_1(|\mathbf{r} - \mathbf{r}'| \rightarrow \infty) = 0$  and  $\partial G_1 / \partial z(x, y, z=0) = 0$ . In a 3D geometry, this Green's function is given by

$$G_1(\mathbf{r}, \mathbf{r}') = \frac{\exp(ik_1|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|} + \frac{\exp(ik_1|\mathbf{r} - \mathbf{r}'_{\text{im}}|)}{4\pi|\mathbf{r} - \mathbf{r}'_{\text{im}}|}, \quad (9)$$

where  $\mathbf{r}' = (x', y', z')$  and  $\mathbf{r}'_{\text{im}} = (x', y', -z')$ . In a 2D geometry (translational invariance along the  $y$  axis), it is given by

$$G_1(\mathbf{r}, \mathbf{r}') = \frac{i}{4} H_0^{(1)}(k_1|\mathbf{r} - \mathbf{r}'|) + \frac{i}{4} H_0^{(1)}(k_1|\mathbf{r} - \mathbf{r}'_{\text{im}}|), \quad (10)$$

where  $H_0^{(1)}$  is the Hankel function of first kind and order zero.

We now apply Eq. (7) in  $V_1$  with  $u = T_1$  and  $v = G_1$ . In the following, the normal  $\mathbf{n}'$  is outward with respect to the object (see Fig. 2). The left-hand side can be simplified using Eqs. (8) and (2) with  $j = 1$ , and we obtain

$$\begin{aligned} & \int_{V_1} G_1(\mathbf{r}, \mathbf{r}') \frac{Q(\mathbf{r}')}{\kappa_1} d^3 \mathbf{r}' - \int_{V_1} T_1(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') d^3 \mathbf{r}' \\ &= \int_{\partial V_1} \left[ T_1(\mathbf{r}') \frac{\partial G_1(\mathbf{r}, \mathbf{r}')}{\partial n'} - G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_1(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}', \end{aligned} \quad (11)$$

where  $\partial V_1 = \Sigma_z \cup \Sigma_2 \cup \Sigma_\infty$ . Due to the boundary conditions satisfied by  $T_1$  and  $G_1$ , the integrals on  $\Sigma_z$  and  $\Sigma_\infty$  vanish. Depending on whether the observation point  $\mathbf{r}$  is inside or outside  $V_1$ , we obtain a pair of integral equations.

For  $\mathbf{r} \in V_1$

$$\begin{aligned} T_1(\mathbf{r}) = T_{\text{ref}}(\mathbf{r}) + \int_{\Sigma_2} \left[ T_1(\mathbf{r}') \frac{\partial G_1(\mathbf{r}, \mathbf{r}')}{\partial n'} \right. \\ \left. - G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_1(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}' \end{aligned} \quad (12)$$

and for  $\mathbf{r} \in V_2$ ,

$$\begin{aligned} T_{\text{ref}}(\mathbf{r}) + \int_{\Sigma_2} \left[ T_1(\mathbf{r}') \frac{\partial G_1(\mathbf{r}, \mathbf{r}')}{\partial n'} \right. \\ \left. - G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_1(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}' = 0, \end{aligned} \quad (13)$$

where

$$T_{\text{ref}}(\mathbf{r}) = \int_{V_1} G_1(\mathbf{r}, \mathbf{r}') \frac{Q(\mathbf{r}')}{\kappa_1} d^3 \mathbf{r}' \quad (14)$$

is the field produced in the semi-infinite medium without the presence of the object (reference field). The expression of  $T_{\text{ref}}$  in the case of a plane-wave illumination of the surface is given in Appendix A. We now apply Green's theorem in  $V_2$  using the Green's function  $G_2$  obeying

$$\nabla^2 G_2(\mathbf{r}, \mathbf{r}') + k_2^2 G_2(\mathbf{r}, \mathbf{r}') = -\delta(\mathbf{r} - \mathbf{r}') \quad (15)$$

and satisfying  $G_2(|\mathbf{r} - \mathbf{r}'| \rightarrow \infty) = 0$ . In a 3D geometry, it is given by

$$G_2(\mathbf{r}, \mathbf{r}') = \frac{\exp(ik_2|\mathbf{r} - \mathbf{r}'|)}{4\pi|\mathbf{r} - \mathbf{r}'|}, \quad (16)$$

whereas in a 2D geometry, it is given by

$$G_2(\mathbf{r}, \mathbf{r}') = \frac{i}{4} H_0^{(1)}(k_2|\mathbf{r} - \mathbf{r}'|). \quad (17)$$

Applying Eq. (7) with  $u = T_2$  and  $v = G_2$ , and making use of Eqs. (15) and (2) with  $j = 2$ , we obtain

$$\begin{aligned} & \int_{V_2} G_2(\mathbf{r}, \mathbf{r}') \frac{Q(\mathbf{r}')}{\kappa_2} d^3 \mathbf{r}' - \int_{V_2} T_2(\mathbf{r}') \delta(\mathbf{r} - \mathbf{r}') d^3 \mathbf{r}' \\ &= \int_{\Sigma_2} \left[ T_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n'} - G_2(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}'. \end{aligned} \quad (18)$$

In the case of a plane-wave illumination, if we denote by  $P(\omega)$  the power density (per unit area) of the laser beam, the volume source term is

$$Q(\mathbf{r}) = \alpha_\nu P(\omega) \frac{\exp(-z/l_\nu)}{l_\nu}, \quad (19)$$

where  $\alpha_\nu$  is the absorptivity of the surface at the frequency  $\nu$  of the laser, and  $l_\nu$  is the absorption length (i.e., the inverse of the absorption coefficient). In a typical metal,  $l_\nu \sim 10 - 100$  nm. Equation (19) shows that  $Q$  vanishes inside the volume  $V_2$ , provided that the distance from the object to the boundary  $z = 0$  is much larger than the absorption length  $l_\nu$  in the substrate at the frequency  $\nu$  of the laser light. We assume that this condition is fulfilled in the following. Finally, we obtain a second pair of integral equations.

For  $\mathbf{r} \in V_2$

$$T_2(\mathbf{r}) = - \int_{\Sigma_2} \left[ T_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n'} - G_2(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}' \quad (20)$$

and for  $\mathbf{r} \in V_1$

$$\int_{\Sigma_2} \left[ T_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n'} - G_2(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}' = 0. \quad (21)$$

Equations (12) and (20) enable us to obtain the temperature field at any point  $\mathbf{r}$  with  $z > 0$ , provided that the source functions  $T_1$ ,  $T_2$ ,  $\partial T_1 / \partial n'$  and  $\partial T_2 / \partial n'$  are known on the object surface  $\Sigma_2$ . The integral in Eq. (12) represents the scattered field in the substrate. Equation (13) expresses how the reference field is cancelled inside the object by this scattered field. This constitutes the equivalent of the extinction theorem in electromagnetic scattering.<sup>18</sup>

## B. Equations for the source functions

Using Eq. (4), let us define two unknown source functions  $T(\mathbf{r})$  and  $\Phi(\mathbf{r})$  for  $\mathbf{r}$  on the surface  $\Sigma_2$  of the object

$$T(\mathbf{r}) = T_1(\mathbf{r}) = T_2(\mathbf{r}), \quad (22)$$

$$\Phi(\mathbf{r}) = \frac{\kappa_1}{\kappa_2} \frac{\partial T_1}{\partial n}(\mathbf{r}) = \frac{\partial T_2}{\partial n}(\mathbf{r}). \quad (23)$$

A pair of coupled integral equations for  $T$  and  $\Phi$  is obtained from Eqs. (12) and (13) and (20) and (21) by letting  $\mathbf{r}$  tend to a point on the surface  $\Sigma_2$ . In this operation, much care has to be taken for the singularities of both the Green's functions  $G_j$  and their normal derivatives  $\partial G_j/\partial n'$  at  $\mathbf{r}=\mathbf{r}'$ . In particular, the singularity of  $\partial G_j/\partial n'$  is nonintegrable, and its extraction is performed following Refs. 17 or 19. In the limit  $\mathbf{r}\rightarrow\Sigma_2$ , Eqs. (12) and (13) lead to the same integral equation involving  $T$  and  $\Phi$

$$\frac{T(\mathbf{r})}{2} = T_{\text{ref}}(\mathbf{r}) + \text{PV} \int_{\Sigma_2} T(\mathbf{r}') \frac{\partial G_1(\mathbf{r}, \mathbf{r}')}{\partial n'} d^2\mathbf{r}' - \frac{\kappa_2}{\kappa_1} \int_{\Sigma_2} G_1(\mathbf{r}, \mathbf{r}') \Phi(\mathbf{r}') d^2\mathbf{r}', \quad (24)$$

where PV stands for the principal value of the integral. Similarly, Eqs. (20) and (21) lead to a second integral equation involving  $T$  and  $\Phi$  for  $\mathbf{r} \in \Sigma_2$

$$\frac{T(\mathbf{r})}{2} = -\text{PV} \int_{\Sigma_2} T(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n'} d^2\mathbf{r}' + \int_{\Sigma_2} G_2(\mathbf{r}, \mathbf{r}') \Phi(\mathbf{r}') d^2\mathbf{r}'. \quad (25)$$

Once the source functions  $T$  and  $\Phi$  are calculated from the system of Eqs. (24) and (25), we can obtain the temperature field at any point in the substrate or in the object by introducing  $T$  and  $\Phi$  into Eqs. (12) and (20), respectively.

Except in very particular geometries, the system of Eqs. (24) and (25) has to be solved numerically. The details of the numerical procedure that we have used are given in the following section.

### C. Numerical procedure

To solve Eqs. (24) and (25), we convert them into a matrix equation, which is then solved numerically. We do this by introducing a set of  $N$  points on the surface  $\Sigma_2$ , each of them having coordinates  $\mathbf{r}_i=(x_i, z_i)$  and being the center of a cell with size  $\Delta\Sigma_i$ . The two unknown functions  $T$  and  $\Phi$  are assumed to be slowly varying functions at the scale of the cell size. When the object is regular enough, it is possible to choose equally spaced points, so that the cells have the same size  $\Delta\Sigma$ . We will assume this condition to be fulfilled in the following. Equations (24) and (25) are then rewritten as

$$\frac{T^i}{2} - \sum_{j \neq i} T^j \left[ \int_{\text{cell}j} \frac{\partial G_1(\mathbf{r}_i, \mathbf{r})}{\partial n_j} d^2\mathbf{r} \right] + \sum_j \Phi^j \left[ \frac{\kappa_2}{\kappa_1} \int_{\text{cell}j} G_1(\mathbf{r}_i, \mathbf{r}) d^2\mathbf{r} \right] = T_{\text{ref}}^i, \quad (26)$$

$$\frac{T^i}{2} + \sum_{j \neq i} T^j \left[ \int_{\text{cell}j} \frac{\partial G_2(\mathbf{r}_i, \mathbf{r})}{\partial n_j} d^2\mathbf{r} \right] - \sum_j \Phi^j \left[ \int_{\text{cell}j} G_2(\mathbf{r}_i, \mathbf{r}) d^2\mathbf{r} \right] = 0. \quad (27)$$

We have used the notations  $T^i=T(\mathbf{r}_i)$  and  $\Phi^i=\Phi(\mathbf{r}_i)$ . If we introduce the two following vectors of length  $2N$ :

$$V = [T^1, T^2, \dots, T^N, \Phi^1, \Phi^2, \dots, \Phi^N],$$

$$V_{\text{ref}} = [T_{\text{ref}}^1, T_{\text{ref}}^2, \dots, T_{\text{ref}}^N, 0, 0, \dots, 0], \quad (28)$$

we can transform Eqs. (26) and (27) into a linear system

$$V_{\text{ref}}^i = M^{ij} V^j. \quad (29)$$

The  $2N \times 2N$  matrix  $\mathbf{M}$  of the linear system has the form

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad (30)$$

where  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  and  $\mathbf{D}$  are  $N \times N$  matrices whose explicit expressions are

$$A_{ij} = \begin{cases} 1/2 & \text{if } i=j \\ -\frac{\partial G_1(\mathbf{r}_i, \mathbf{r}_j)}{\partial n_j} \Delta\Sigma & \text{if } i \neq j, \end{cases} \quad (31)$$

$$B_{ij} = \begin{cases} \frac{\kappa_2}{\kappa_1} \int_{\text{cell}i} G_1(\mathbf{r}_i, \mathbf{r}) d^2\mathbf{r} & \text{if } i=j \\ \frac{\kappa_2}{\kappa_1} G_1(\mathbf{r}_i, \mathbf{r}_j) \Delta\Sigma & \text{if } i \neq j, \end{cases} \quad (32)$$

$$C_{ij} = \begin{cases} 1/2 & \text{if } i=j \\ \frac{\partial G_2(\mathbf{r}_i, \mathbf{r}_j)}{\partial n_j} \Delta\Sigma & \text{if } i \neq j, \end{cases} \quad (33)$$

$$D_{ij} = \begin{cases} -\int_{\text{cell}i} G_2(\mathbf{r}_i, \mathbf{r}) d^2\mathbf{r} & \text{if } i=j \\ -G_2(\mathbf{r}_i, \mathbf{r}_j) \Delta\Sigma & \text{if } i \neq j. \end{cases} \quad (34)$$

The evaluation of the diagonal elements of the matrices  $\mathbf{B}$  and  $\mathbf{D}$  has to be performed with care, due to the singularity of the Green functions at the origin. The integral on the cell domain is performed analytically, in the limit where the cell size tends to zero. For the 2D case, the evaluation of the diagonal elements is performed in Appendix B of Ref. 20. Note that the procedure used to convert the integral equation into a linear system is known as the moment method. The reader is referred to Ref. 17 for more details on this method. Once the linear system (29) is solved by standard procedures, the source functions  $T$  and  $\Phi$  are known on each point of the object surface  $\Sigma_2$ . We can then calculate the temperature field at any point in the substrate or in the object by using Eqs. (12) and (20).

### IV. VOLUME INTEGRAL FORMALISM

Another formulation of the scattering problem may be obtained by using only the Green function  $G_1$ , and applying Green's theorem successively in  $V_1$  and  $V_2$ . This procedure leads to an expression of the temperature field involving a well-known volume integral and a previously unreported surface term. We will first derive the exact integral equation for the total field, and then show how the Born approximation can be used to obtain an explicit expression of the scattered field. We will also derive analytically some criteria of validity of this approximation, and compare the approximate results to exact numerical calculations in Sec. V.



## A. Alternative formulation

The use of Green's theorem in  $V_1$  with the Green function  $G_1$  leads to Eqs. (12) and (13) derived in Sec. III. Let us now apply Green's theorem in  $V_2$ , with  $u=T_2$  and  $v=G_1$ . Making use of Eqs. (8) and (2) with  $j=2$ , we obtain: for  $\mathbf{r} \in V_2$

$$T_2(\mathbf{r}) = (k_2^2 - k_1^2) \int_{V_2} G_1(\mathbf{r}, \mathbf{r}') T_2(\mathbf{r}') d^3 \mathbf{r}' - \int_{\Sigma_2} \left[ T_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n'} - G_2(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}' \quad (35)$$

and for  $\mathbf{r} \in V_1$

$$\int_{\Sigma_2} \left[ T_2(\mathbf{r}') \frac{\partial G_2(\mathbf{r}, \mathbf{r}')}{\partial n'} - G_2(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} \right] d^2 \mathbf{r}' = (k_2^2 - k_1^2) \int_{V_2} G_1(\mathbf{r}, \mathbf{r}') T_2(\mathbf{r}') d^3 \mathbf{r}'. \quad (36)$$

In Eqs. (35) and (36), the normal  $\mathbf{n}'$  is outward with respect to the object (see Fig. 2). As in the previous section, we have assumed that the source term  $Q$  is zero inside the object, so that the integrals over  $V_2$  involving  $Q$  in the integrand vanish. Subtracting Eq. (35) from Eq. (13) (for  $\mathbf{r} \in V_2$ ), and Eq. (36) from Eq. (12) (for  $\mathbf{r} \in V_1$ ), and using the boundary conditions Eq. (4), one obtains

$$T(\mathbf{r}) = T_{\text{ref}}(\mathbf{r}) + (k_2^2 - k_1^2) \int_{V_2} G_1(\mathbf{r}, \mathbf{r}') T_2(\mathbf{r}') d^3 \mathbf{r}' + \left( 1 - \frac{\kappa_2}{\kappa_1} \right) \int_{\Sigma_2} G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} d^2 \mathbf{r}'. \quad (37)$$

Equation (37) is an exact expression of the temperature field, valid at any point  $\mathbf{r}=(x,y,z)$  with  $z>0$ . It involves a volume integral, extended to the volume of the object, and a surface integral, extended to its boundary.

This formulation of the problem has the advantage of clearly presenting the diffusivity variation and the conductivity variation as the origin of scattering. Moreover, the two contributions are separated: the diffusivity variation yields a volume effect and the conductivity variation yields a surface effect. The existence of two such contributions to the scattered field has been discussed in the case of DPDW.<sup>7,9</sup>

## B. Born approximation

Although numerical procedures can be used to solve the scattering problem exactly (see Sec. V), explicit approximate solutions are also useful to get insight into the physics of a given system and/or to get a fast evaluation of the temperature field. Such a solution can be obtained by performing the first Born approximation in Eq. (37). This approximation also leads to a natural solution to the inverse problem, by deconvolution procedures.

The first Born approximation amounts to approximate the field inside the scattering object by the reference field  $T_{\text{ref}}$ . In the case of the volume integral in Eq. (37), one

simply makes the approximation  $T_2(\mathbf{r}') \approx T_{\text{ref}}(\mathbf{r}')$ . The surface integral has to be handled with care because  $\partial T_2 / \partial n'$  is discontinuous across the boundary  $\Sigma_2$ . Thus, replacing  $\partial T_2 / \partial n'$  by  $\partial T_{\text{ref}} / \partial n'$  would be incorrect. A careful analysis shows that  $\partial T_2(\mathbf{r}') / \partial n' \approx \alpha_B \partial T_{\text{ref}}(\mathbf{r}') / \partial n'$  with  $\alpha_B = 2 \eta \kappa_1 / [(2 \eta - 1) \kappa_1 + \kappa_2]$  where  $\eta=2$  for a 2D problem and  $\eta=3$  for a 3D problem (see Appendix B for details). In the Born approximation, Eq. (37) yields

$$T(\mathbf{r}) = T_{\text{ref}}(\mathbf{r}) + (k_2^2 - k_1^2) \int_{V_2} G_1(\mathbf{r}, \mathbf{r}') T_{\text{ref}}(\mathbf{r}') d^3 \mathbf{r}' + \alpha_B \left( 1 - \frac{\kappa_2}{\kappa_1} \right) \int_{\Sigma_2} G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_{\text{ref}}(\mathbf{r}')}{\partial n'} d^2 \mathbf{r}'. \quad (38)$$

Some criteria of validity of Eq. (38) may be obtained by requiring that the scattered field remains a first-order correction to the reference field on the integration domain, namely, that

$$\sup_{\mathbf{r} \in V_2} \left| \frac{T(\mathbf{r}) - T_{\text{ref}}(\mathbf{r})}{T_{\text{ref}}(\mathbf{r})} \right| \ll 1. \quad (39)$$

A rough estimation of the two integrals in Eq. (38) leads to sufficient conditions of validity (see Appendix C for details). For a 3D geometry, these conditions are

$$\mathcal{B}_v = |k_2^2 - k_1^2| \frac{V_2}{2 \pi d} \ll 1, \quad (40)$$

$$\mathcal{B}_s = \frac{6 |\kappa_1 - \kappa_2|}{5 \kappa_1 + \kappa_2} \frac{\Sigma_2 |k_1|}{2 \pi d} \ll 1, \quad (41)$$

whereas for a 2D geometry, they are

$$\mathcal{B}_v = |k_2^2 - k_1^2| \frac{V_2}{\pi} [|\ln(d/\delta_1)| + \exp(-d/\delta_1)] \ll 1, \quad (42)$$

$$\mathcal{B}_s = \frac{4 |\kappa_1 - \kappa_2|}{3 \kappa_1 + \kappa_2} \frac{|k_1| \Sigma_2}{\pi} [|\ln(d/\delta_1)| + \exp(-d/\delta_1)] \ll 1, \quad (43)$$

where  $\delta_1 = [\text{Im}(k_1)]^{-1}$  is the thermal diffusion length in medium 1,  $V_2$  is the volume of the object,  $\Sigma_2$  the surface of its boundary, and  $d$  its diameter. The preceding conditions give a rough estimate of the range of validity of the first Born approximation for thermal-wave scattering and in particular the influence of each physical parameter on this approximation.

## V. NUMERICAL SIMULATION

### A. Definition of the complex contrast

In a typical photothermal imaging experiment, the modulated temperature field at the surface of the sample is scanned at different times using radiometric, optical or other measurement methods.<sup>14,21</sup> Frequency filtering allows us to compute the phase and amplitude of the signal at each point. The absolute value of the temperature does not matter in detecting objects, but rather the departure of the temperature from the unperturbed reference field. Therefore, the quantity of interest at the interface  $z=0$  is the complex temperature contrast

$$C(\mathbf{r}) = \frac{T(\mathbf{r}) - T_{\text{ref}}(\mathbf{r})}{T_{\text{ref}}(\mathbf{r})}. \quad (44)$$

When the scattered field is a small correction to the reference field, we may use the approximation

$$C(\mathbf{r}) \approx \ln\left(\frac{T(\mathbf{r})}{T_{\text{ref}}(\mathbf{r})}\right). \quad (45)$$

Equation (45) shows that the real part of the contrast represents the logarithm of the amplitude of the temperature field, whereas the imaginary part represents its phase difference with the reference field.

### B. Numerical tests

We present in this section numerical results, based on the resolution of Eqs. (24) and (25) by a moment method, without approximation. We will also compare the results with those obtained within the first Born approximation, namely, by a direct calculation of the integrals in Eq. (38). Note that apart from the expression of the reference field—see Eq. (A3)—thermal conductivities appear only through the ratio  $\kappa_2/\kappa_1$  in all the equations of the problem. Since  $T_{\text{ref}}(z=0)$  is a scaling parameter for all the temperature fields, the conductivity ratio  $\kappa_2/\kappa_1$  is the relevant physical parameter to determine the contrast. In the following, all objects are centered at  $x=0$ .

In order to validate our numerical results, we compare in Fig. 3 the numerical calculation for a rectangular object such that its width  $L_x$  is much larger than  $\delta_1$  with the analytical solution for a layered medium (i.e., an object with  $L_x \rightarrow \infty$ , see Appendix D). The object in the numerical simulation is a rectangle centered at depth  $z_c = 1.5$  mm with dimensions  $L_x = 20$  mm,  $L_z = 0.5$  mm, and a diffusivity  $D_2 = 40$  mm<sup>2</sup> s<sup>-1</sup>. The modulation frequency is  $f = \omega/2\pi = 10$  Hz. The substrate has a diffusivity  $D_1 = 30$  mm<sup>2</sup> s<sup>-1</sup>. We see that the numerical solution tends towards the analytical result when the mesh size decreases, thus proving the good convergence of the method. Figure 3 also shows that the numerical result diverges near the boundaries of the object, where the normal derivative of the temperature has a discontinuity. This divergence appears numerically at a distance of the order of the mesh size.

We show in Fig. 4 the real and imaginary parts of the complex temperature contrast along the interface  $z=0$  for two different objects. The substrate and the objects have a diffusivity  $D_1 = D_2 = 30$  mm<sup>2</sup> s<sup>-1</sup> (conductivity variation only). The modulation frequency of the laser beam is  $f = 10$  Hz. The 2D object is a small disk. In Fig. 4(a), the diameter is  $L = 0.1$  mm and the conductivity ratio  $\kappa_2/\kappa_1 = 2$ . The calculation is plotted for different mesh sizes. We see that convergence is ensured when the mesh size is much smaller than the attenuation distance. Another interesting point is that the Born approximation coincides with the converged calculation, showing that with these parameters, the Born approximation gives accurate results. To illustrate the relevance of the factor  $\alpha_B$  introduced in Eq. (38), we also show the result of a Born approximation where  $\partial T_2/\partial n'$  is approximated by  $\partial T_{\text{ref}}/\partial n'$ . This second Born result is

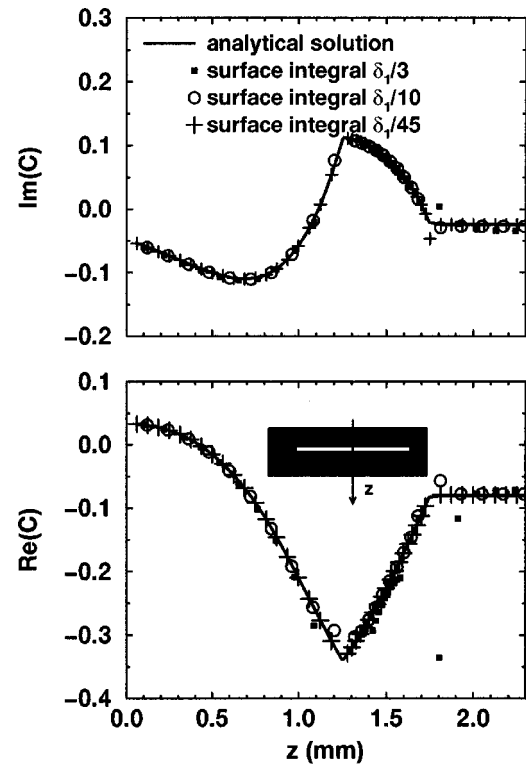


FIG. 3. Complex contrast along a line of constant  $x$  going through the center of an elongated rectangular object (see inset). We compare the surface integral calculation with the analytical solution for an infinitely wide object. The numerical solution is shown for several mesh-sizes  $\Delta\Sigma/\delta_1 = 1/3; 1/10; 1/45$ . Parameters of the calculation:  $L_x = 20$  mm,  $L_z = 0.5$  mm,  $z_c = 1.5$  mm,  $\delta_1 = 0.98$  mm,  $\delta_2 = 1.13$  mm,  $\kappa_2/\kappa_1 = 3$ .

clearly in a worse agreement with the converged result. We conclude that this factor significantly improves the accuracy of the Born approximation when this approximation is valid.

In Fig. 4(b), the diameter is doubled ( $L = 0.2$  mm) and the conductivity ratio is increased to  $\kappa_2/\kappa_1 = 50$ . In these

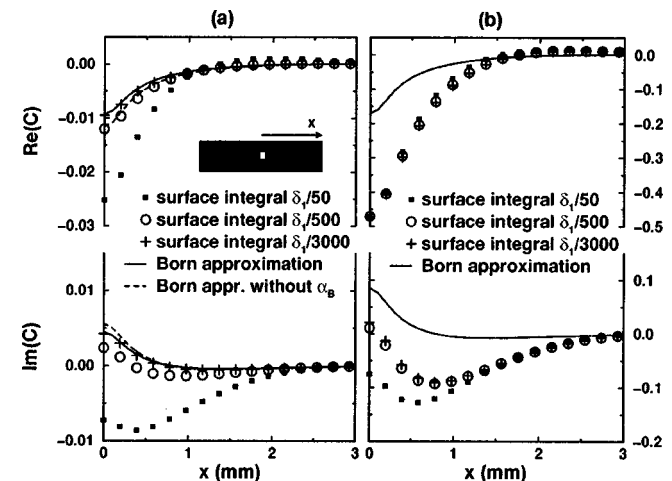


FIG. 4. Complex contrast on the surface in the case of a conductivity variation. The 2D object is a small disk with diameter  $L$ . The numerical solution is shown for several mesh-sizes  $\Delta\Sigma$ . (a)  $\kappa_2/\kappa_1 = 2$ ,  $L = 0.1$  mm. (b)  $\kappa_2/\kappa_1 = 50$ ,  $L = 0.2$  mm. Other parameters:  $z_c = 0.3$  mm,  $\delta_1 = \delta_2 = 0.98$  mm.

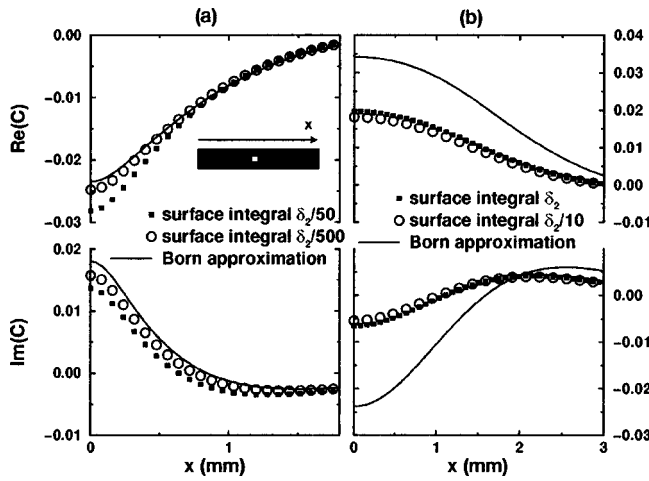


FIG. 5. Complex contrast on the surface in the case of a diffusivity variation:  $\delta_1=0.98$  mm,  $\delta_2=0.40$  mm. The 2D object is a square box with size  $L$ . The numerical solution is shown for several mesh-sizes  $\Delta\Sigma$ . (a)  $L=0.1$  mm,  $z_c=0.3$  mm. (b)  $L=1$  mm,  $z_c=2$  mm.

conditions, the Born approximation loses its accuracy, as predicted by the increase of the factor  $\mathcal{B}_s$  from 0.4 (a) to 2.6 (b).

The same calculation is presented in Fig. 5, for two objects with no conductivity variation ( $\kappa_2/\kappa_1=1$ ), diffusivity  $D_2=5$  mm<sup>2</sup>s<sup>-1</sup> in a substrate with diffusivity  $D_1=30$  mm<sup>2</sup>s<sup>-1</sup>. We see that convergence is ensured when the mesh size is much smaller than the attenuation distance in both media (with the above parameters,  $\delta_1=0.98$  mm and  $\delta_2=0.40$  mm). In Fig. 5(a), the object is a square of size  $L_x=L_z=0.1$  mm (see the inset in Fig. 5). The Born approximation yields a good result in this case. In Fig. 5(b), the object size is increased to 1 mm. The Born approximation is not accurate in that case, as predicted by the increase of the factor  $\mathcal{B}_v$  from 0.2 (a) to 4 (b).

We have proven the convergence and validity of our numerical method through comparison with an analytical result. We have used it as a reference method to check the validity of the Born approximation in several cases, thus validating simple coefficients that allow us to infer whether the Born approximation should be valid in a particular case. In the Born approximation, we have proven that the normal derivative of the field is better approximated when using a corrective factor that takes into account the discontinuity of the thermal conductivity at the boundary of the object.

**C. Determination of the object depth**

In the case of a layered medium, the complex contrast on the interface  $z=0$  has the form (see Appendix D)

$$C(f) = ik_2L_z \left( \frac{1}{r} - r \right) \exp(2ik_1z_c), \tag{46}$$

where  $r = (\kappa_1/\kappa_2)\sqrt{D_2/D_1}$ .  $L_z$  and  $z_c$  are the width and the depth of the layer, respectively. This expression is obtained in the limit of small width  $|k_2L_z| \ll [2r/(1+r)^2]$ .

Figure 6 shows the complex contrast as a function of the depth  $z_c$  of the object. The object is a layer of thickness  $L_z=0.1$  mm with diffusivity  $D_2=4$  mm<sup>2</sup>s<sup>-1</sup>, conductivity ratio

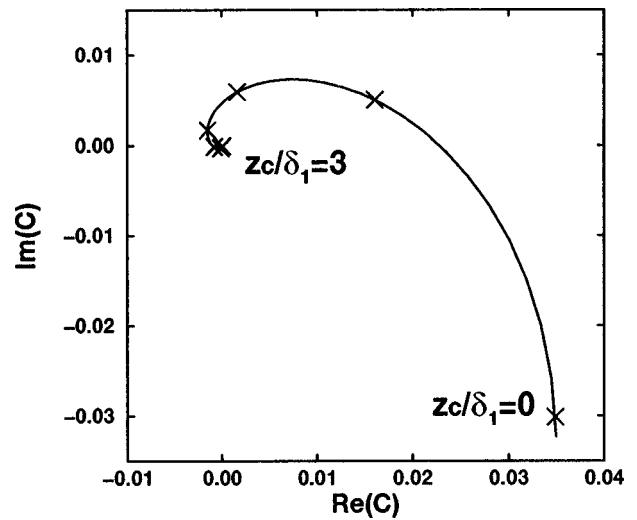


FIG. 6. Polar representation of the complex contrast as a function of the depth of the object in the monodimensional case.  $L_z/\delta_2=0.089$ ,  $\kappa_2/\kappa_1=0.3$ ,  $\delta_1=3.1$  mm,  $\delta_2=1.13$  mm,  $f=1$  Hz. Symbols (x) are separated by a 0.5 increment of  $z_c/\delta_1$ .

$\kappa_2/\kappa_1=0.3$  in a substrate with diffusivity  $D_1=30$  mm<sup>2</sup>s<sup>-1</sup>. The behavior of the contrast in this case can help to understand why the better way to image an object can be to measure either the amplitude or the phase of the temperature fluctuation at the surface. Indeed, depending on  $z_c$ , the dominant part of  $C$  is real (see, for example,  $z_c/\delta_1=0.4$ ) or imaginary ( $z_c/\delta_1=1$ ).

Considering  $C$  as a function of the frequency  $f$ , we see that  $z_c$  may be obtained from  $C$  by the slope of the curves

$$\ln|C(f)| - \frac{1}{2}\ln f = -2 \frac{z_c}{\delta_1} + Cte, \tag{47}$$

$$\text{Im}\{\ln[C(f)]\} = 2 \frac{z_c}{\delta_1} + \frac{3\pi}{4}. \tag{48}$$

For an object of arbitrary shape, using Eq. (47) or Eq. (48) in order to determinate its depth leads to reasonably accurate results that may be useful in practice. We show in Fig. 7 a numerical example of depth determination for a mineral inclusion in a steel plate. The modulation frequency varies from  $f=0.1$  to  $f=1.5$  Hz. We plot the logarithm of  $|C|$  and the phase of  $C$  vs  $1/\delta_1 = \text{Im}(k_1)$ , the latter being proportional to  $\sqrt{f}$ . From the slope of the curves, the estimated depth is  $z_e^a=2.52$  mm with the amplitude and  $z_e^p=2.93$  mm with the phase, whereas the exact value is  $z_c=3.0$  mm. Note that the size of the object does not strictly satisfy the condition of validity of Eq. (46) since  $|k_2L_z|$  varies between 0.5 and 2 in this range of frequency.

**VI. CONCLUSION**

We have presented a method to solve exactly the heat diffusion equation in the frequency domain within a homogeneous substrate containing an arbitrary object. We have derived an extinction theorem for thermal waves which is analogous to that used in electromagnetic scattering. This formulation leads to surface-integral equations which are

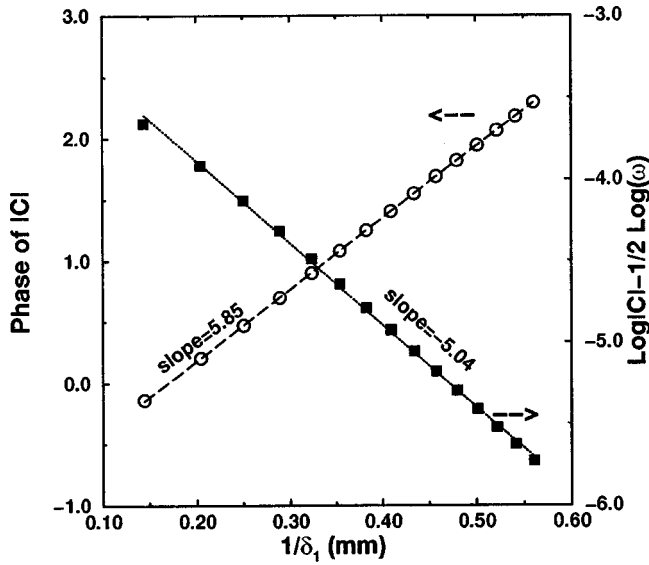


FIG. 7. Example of inversion of the complex contrast for the depth of the heterogeneity. The frequency  $f$  varies from 0.1 to 1.5 Hz. The estimated depth is 2.52 mm with the amplitude ( $\square$ ) and 2.93 mm with the phase ( $\circ$ ) of  $C$ . The exact value is  $z_c=3.0$  mm. Other parameters:  $D_1 = 15.0 \text{ mm}^2 \text{ s}^{-1}$ ,  $D_2=0.6 \text{ mm}^2 \text{ s}^{-1}$ ,  $\kappa_2/\kappa_1=2.0 \times 10^{-2}$ . The object is a rectangular box with dimensions:  $L_x=1.0$  mm,  $L_z=0.5$  mm.

well suited for an efficient numerical solution. Although we have presented only 2D numerical results, the implementation of our numerical procedure in 3D is easily feasible. Also, whereas we have only presented results for a single object, generalization to multiple domains is straightforward.

Second, we have derived a mixed surface and volume integral formulation and have applied the Born approximation. We have shown that a corrective factor arising from the discontinuity of the thermal conductivity at the boundary of the defect must be taken into account to approximate the normal derivative of the field. Exact solutions of the direct problem allow us to check and validate approximations that are often used to solve the inverse problem. Especially, the Born approximation is valid for small defects and small variations of the thermal conductivity and diffusivity. We have expressed simplistic analytical criteria of validity that correctly predict the influence of each physical parameter on the accuracy of the Born approximation.

Third, the analytical solution for the temperature contrast at the surface of a multilayered medium suggests a way to estimate the depth of an object with multiple frequency measurements. We have performed computations showing that such a method gives a precision that can reach a few percent.

### APPENDIX A

In this Appendix, we give the explicit expressions of the reference field  $T_{\text{ref}}$  appearing in Eqs. (2) and (14). We assume that the heating laser beam is a plane wave, incident on the surface  $z=0$  at normal incidence (see Fig. 1).

Introducing Eq. (19) into Eq. (14), we obtain

$$T_{\text{ref}}(\mathbf{r}) = \frac{\alpha_\nu P(\omega)}{l_\nu \kappa_1} \int_{V_1} G_1(\mathbf{r}, \mathbf{r}') \exp(-z'/l_\nu) d^3 \mathbf{r}'. \quad (\text{A1})$$

We now make use of the plane-wave representation of  $G_1$  (or Weyl expansion<sup>22</sup>)

$$G_1(\mathbf{r}, \mathbf{r}') = \frac{i}{8\pi^2} \int \exp[i\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}')] \times \frac{\exp(i\gamma|z - z'|) + \exp(i\gamma|z + z'|)}{\gamma} d^2 \mathbf{K}, \quad (\text{A2})$$

where  $\mathbf{R}=(x, y)$ ,  $\gamma=(k_1^2 - \mathbf{K}^2)^{1/2}$ , with  $\text{Re}(\gamma)>0$  and  $\text{Im}(\gamma)>0$ . Introducing Eq. (A2) into Eq. (A1) leads to

$$T_{\text{ref}}(\mathbf{r}) = \frac{i\alpha_\nu P(\omega)}{2l_\nu \kappa_1 k_1} \int_0^{+\infty} [\exp(i\gamma|z - z'|) + \exp(i\gamma|z + z'|)] \exp(-z'/l_\nu) dz'. \quad (\text{A3})$$

Equation (A3) is an exact expression of the reference temperature field when the interface  $z=0$  is illuminated by a plane wave.

When the absorption in the substrate is high ( $l_\nu \rightarrow 0$ ), we can see that the laser beam is absorbed within an infinitely thin layer at  $z=0$ . In this case, one can derive a simplified expression for  $T_{\text{ref}}$ . The integral in Eq. (A3) can be split into two contributions: that of  $[0, z]$ , and that of  $[z, \infty]$ . In the limit  $l_\nu \rightarrow 0$ , one can easily show that the first integral equals  $2l_\nu \exp(ik_1 z)$ , whereas the second integral tends to zero. Finally, we end up with the expression for  $T_{\text{ref}}$  in the surface-absorption model

$$T_{\text{ref}}(\mathbf{r}) = \frac{i\alpha_\nu P(\omega)}{\kappa_1 k_1} \exp(ik_1 z). \quad (\text{A4})$$

This expression of  $T_{\text{ref}}$  is valid when  $z \gg l_\nu$ .

### APPENDIX B

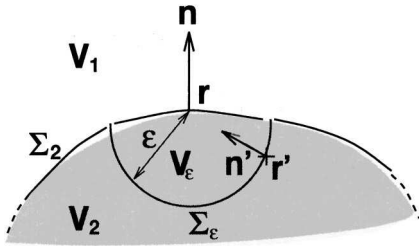
In this Appendix, we derive the correct form of the Born approximation in the surface-integral term appearing in Eq. (38). Taking the gradient of Eq. (37) leads to

$$\nabla T(\mathbf{r}) = \nabla T_{\text{ref}}(\mathbf{r}) + (k_2^2 - k_1^2) \int_{V_2} \nabla_{\mathbf{r}} G_1(\mathbf{r}, \mathbf{r}') T_2(\mathbf{r}') d^3 \mathbf{r}' + \left(1 - \frac{\kappa_2}{\kappa_1}\right) \int_{\Sigma_2} \nabla_{\mathbf{r}} G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_2(\mathbf{r}')}{\partial n'} d^2 \mathbf{r}'. \quad (\text{B1})$$

Because of the nonintegrable singularity of  $\nabla_{\mathbf{r}} G_1(\mathbf{r}, \mathbf{r}')$  at  $\mathbf{r} = \mathbf{r}'$  in the surface integral, we cannot simply state that the zeroth order approximation for  $\nabla T$  is  $\nabla T_{\text{ref}}$ . Instead, we need to extract this singularity in a similar fashion as in deriving Eqs. (24) and (25). Because this is a slightly different situation from that in Ref. 19, we detail this derivation hereafter. We consider a point  $\mathbf{r}$  on  $\Sigma_2$  and a half sphere  $V_\epsilon$  (half disk in 2D) inside  $V_2$  centered on  $\mathbf{r}$  with infinitely small radius  $\epsilon$ .  $\Sigma_\epsilon$  is the boundary of  $V_\epsilon$  (see Fig. 8). The surface integral in Eq. (B1) is split into the contribution of  $\Sigma_2/\Sigma_\epsilon$  and that of  $\Sigma_\epsilon$ . The singularity is evaluated by taking the limit at vanishing  $\epsilon$  of

$$I_\epsilon = \int_{\Sigma_\epsilon} \mathbf{n} \cdot \nabla_{\mathbf{r}} G_1(\mathbf{r}, \mathbf{r}') \mathbf{n}' \cdot \nabla T_2(\mathbf{r}') d^2 \mathbf{r}'. \quad (\text{B2})$$



FIG. 8. Geometry for the extraction of the singularity in  $\nabla_{\mathbf{r}}G_1(\mathbf{r},\mathbf{r}')$ .

Since  $T_2$  is analytical inside  $V_2$ , we may introduce the following Taylor expansion into Eq. (B2):

$$\nabla T_2(\mathbf{r}') = \nabla T_2(\mathbf{r}) + |\mathbf{r} - \mathbf{r}'| \mathbf{F}(\mathbf{r}, \mathbf{r}'), \quad (\text{B3})$$

where  $|\mathbf{F}|$  is finite. Using the asymptotic term

$$\nabla_{\mathbf{r}}G_1(\mathbf{r},\mathbf{r}') \sim \frac{1}{2(\eta-1)\pi} \frac{\mathbf{r}' - \mathbf{r}}{|\mathbf{r}' - \mathbf{r}|^\eta}, \quad (\text{B4})$$

where  $\eta=2$  in 2D and  $\eta=3$  in 3D, we obtain

$$\lim_{\epsilon \rightarrow 0} I_\epsilon = \nabla T_2(\mathbf{r}) \cdot \frac{1}{2\eta} \mathbf{n}. \quad (\text{B5})$$

Finally, the zeroth order solution of Eq. (B1) is

$$\frac{\partial T_2}{\partial n}(\mathbf{r}) \simeq \frac{2\eta\kappa_1}{(2\eta-1)\kappa_1 + \kappa_2} \frac{\partial T_{\text{ref}}}{\partial n}(\mathbf{r}). \quad (\text{B6})$$

### APPENDIX C

We give in this Appendix the detailed derivation of conditions (40)–(43) for the validity of the first Born approximation.

A sufficient condition for Eq. (39) to be satisfied is that for  $\mathbf{r}$  in  $V_2$

$$\left| (k_2^2 - k_1^2) \int_{V_2} G_1(\mathbf{r}, \mathbf{r}') T_{\text{ref}}(\mathbf{r}') d^3\mathbf{r}' \right| \ll T_{\text{ref}}(\mathbf{r}), \quad (\text{C1})$$

$$\left| \left( 1 - \frac{\kappa_2}{\kappa_1} \right) \int_{\Sigma_2} G_1(\mathbf{r}, \mathbf{r}') \frac{\partial T_{\text{ref}}(\mathbf{r}')}{\partial n'} d^2\mathbf{r}' \right| \ll T_{\text{ref}}(\mathbf{r}). \quad (\text{C2})$$

From Eqs. (9) and (10), we can estimate the modulus of the Green function. For a 3D geometry, we obtain:  $|G_1(\mathbf{r}, \mathbf{r}')| \ll (2\pi d)^{-1}$ , where  $d$  is a typical distance between two points of the object, namely its diameter. For a 2D geometry, we have:  $|G_1(\mathbf{r}, \mathbf{r}')| \simeq |H_0^{(1)}[(1+i)d/\delta_1]|/2$ , and  $|H_0^{(1)}[(1+i)x]| \ll 2[|\ln(x)| + \exp(-x)]/\pi$ . Inserting those approximations into Eqs. (C1) and (C2) leads to the four sufficient conditions (40)–(43).

### APPENDIX D

In this Appendix, we derive the exact expression of the contrast at the surface in the case of a one-dimensional problem. The object is a layer of thickness  $L_z$  centered at depth  $z_c$ . The modulated temperature field in each homogeneous domain labeled by  $j$  ( $j=1$  in the substrate above the object,  $j=2$  in the object,  $j=3$  in the substrate below the object), can be expressed as the sum of a wave going down and a wave going up

$$T_j(z) = A_j \exp(ik_j z) + B_j \exp(-ik_j z) \quad (\text{D1})$$

with  $k_3 = k_1$ . The boundary conditions of the problem are

$$\kappa_1 \frac{dT_1}{dz}(0) = \kappa_1 \frac{dT_{\text{ref}}}{dz}(0), \quad (\text{D2})$$

$$\lim_{z \rightarrow \infty} \left( \kappa_1 \frac{dT_3}{dz} \right) = 0 \quad (\text{D3})$$

and the continuity of the temperature field and of the normal flux at  $z = z_c \pm L_z/2$ , that yield four additional relations. By substituting the expression (D1) into the boundary conditions, six linear relations between the unknowns  $A_1, A_2, A_3, B_1, B_2, B_3$  are obtained. After some algebra, one gets

$$C(0) = \frac{2\mathcal{R}c^2(b^2 - 1)}{1 - b^2\mathcal{R}^2 - \mathcal{R}c^2(b^2 - 1)}, \quad (\text{D4})$$

where  $b = \exp(ik_2 L_z)$ ,  $c = \exp(ik_1[z_c - L_z/2])$ , and  $\mathcal{R} = (1 - r)/(1 + r)$  with  $r = (\kappa_1/\kappa_2)\sqrt{D_2/D_1}$ . Expression (46) is obtained by approximating the denominator of Eq. (D4) by  $(1 - \mathcal{R}^2)$  under the condition  $|b^2 - 1| \ll 1 - \mathcal{R}^2$ .

<sup>1</sup>A. Rosenzweig and A. Gersho, J. Appl. Phys. **47**, 64 (1976); G. Busse, *ibid.* **21**, 107 (1982); Z. Sodnik and H. Tiziani, Opt. Commun. **58**, 295 (1986).

<sup>2</sup>*Photoacoustic and Thermal Wave Phenomena in Semiconductors*, edited by A. Mandelis (North-Holland, Amsterdam, 1987); *Eighth International Meeting on Photoacoustic and Photothermal Phenomena*, edited by D. Fournier and J. Roger (Les Éditions de Physique, Les Ulis, 1994).

<sup>3</sup>P. Kuo and L. Favro, Appl. Phys. Lett. **40**, 1012 (1982); A. Mandelis, J. Opt. Soc. Am. A **6**, 298 (1989).

<sup>4</sup>A. Yodh and B. Chance, Phys. Today **48**, 34 (1995).

<sup>5</sup>A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978).

<sup>6</sup>J. Fishkin and E. Gratton, J. Opt. Soc. Am. A **10**, 127 (1993); J. Ripoll, M. Nieto-Vesperinas, and R. Carminati, *ibid.* **16**, 1466 (1999).

<sup>7</sup>J. Ripoll and M. Nieto-Vesperinas, J. Opt. Soc. Am. A **16**, 1453 (1999).

<sup>8</sup>Y. Yao, Y. Wang, Y. Pei, W. Zhu, and R. Barbour, J. Opt. Soc. Am. A **14**, 325 (1997).

<sup>9</sup>M. Ostermeyer and S. Jacques, J. Opt. Soc. Am. A **14**, 255 (1997).

<sup>10</sup>E. Miller, L. Nicolaides, and A. Mandelis, J. Opt. Soc. Am. A **15**, 1545 (1998).

<sup>11</sup>H. Jiang, K. Paulsen, U. Osterberg, B. Pogue, and M. Patterson, J. Opt. Soc. Am. A **13**, 253 (1996).

<sup>12</sup>S. Norton and T. Vo-Dinh, J. Opt. Soc. Am. A **15**, 2670 (1998).

<sup>13</sup>A. Mandelis, J. Phys. A **24**, 2485 (1991); A. Mandelis, J. Appl. Phys. **78**, 647 (1995).

<sup>14</sup>A. Sánchez-Lavega and A. Salazar, J. Appl. Phys. **74**, 536 (1993).

<sup>15</sup>D. Crowther, L. Favro, P. Kuo, and R. Thomas, J. Appl. Phys. **74**, 5828 (1993). O. Pade and A. Mandelis, in part 2 of Ref. 2, pp. 99–102. O. Pade and A. Mandelis, Inverse Probl. **10**, 185 (1994).

<sup>16</sup>U. Seidel, H. Walther, and J. Burt, in part 2 of Ref. 2, pp. 551–554.

<sup>17</sup>J. Wang, *Generalized Moment Method in Electromagnetics: Formulation and Computer Solution of Integral Equations* (Wiley, New York, 1991).

<sup>18</sup>M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1980).

<sup>19</sup>A. J. Poggio and E. K. Miller, in *Computer Techniques for Electromagnetics*, edited by R. Mittra (Pergamon, Oxford, 1973), Vol. 7 of *Monographs in Electrical Engineering*.

<sup>20</sup>A. Maradudin, T. Michel, A. McGurn, and E. Mendez, Ann. Phys. (N.Y.) **203**, 255 (1990).

<sup>21</sup>M. Munidasa and A. Mandelis, J. Opt. Soc. Am. A **8**, 1851 (1991). C. Bennett and R. Patty, Appl. Opt. **21**, 49 (1982).

<sup>22</sup>L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, UK, 1995).

**Kirchhoff approximation for diffusive waves**

Jorge Ripoll\*

*Institute for Electronic Structure and Laser, Foundation for Research and Technology-Hellas, P.O. Box 1527, 71110 Heraklion, Crete, Greece*

Vasilis Ntziachristos

*Center For Molecular Imaging Research, Massachusetts General Hospital & Harvard Medical School, Building 149, 13th Street 5406, Charlestown, Massachusetts 02129-2060*

Remi Carminati

*Laboratoire d'Énergetique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche, Scientifique, 92295 Chatenay-Malabry Cedex, France*

Manuel Nieto-Vesperinas

*Instituto de Ciencia de Materiales, Consejo Superior de Investigaciones Científicas, Campus de Cantoblanco, 28049 Madrid, Spain*

(Received 28 May 2001; revised manuscript received 9 August 2001; published 26 October 2001)

Quantitative measurements of diffuse media, in spectroscopic or imaging mode, rely on the generation of appropriate forward solutions, independently of the inversion scheme employed. For complex boundaries, the use of numerical methods is generally preferred due to implementation simplicity, but usually results in great computational needs, especially in three dimensions. Analytical expressions are available, but are limited to simple geometries such as a diffusive slab, a sphere or a cylinder. An analytical approximation, the Kirchhoff approximation, also called the tangent-plane method is presented for the case of diffuse light. Using this approximation, analytical solutions of the diffusion equation for arbitrary boundaries and volumes can be derived. Also, computation time is minimized since no matrix inversion is involved. The accuracy of this approximation is evaluated on comparison with results from a rigorous numerical technique calculated for an arbitrary geometry. Performance of the approximation as a function of the optical properties and the size of the medium is examined and it is demonstrated that the computation time of the direct scattering model is reduced at least by two orders of magnitude.

DOI: 10.1103/PhysRevE.64.051917

PACS number(s): 87.10.+e, 42.25.Fx, 42.30.Wb, 42.62.Be

**I. INTRODUCTION**

The study of light transport through highly scattering media, such as tissue, has been the focus of recent research geared towards medical diagnostics [1–9]. This has been motivated by the fact that light offers unique contrast mechanisms while probing structural and functional tissue characteristics. Furthermore, the associated technology employs nonionizing radiation and is generally low cost. Imaging through tissues using light in the near infrared (NIR) spectral region offers penetration capability of several centimeters due to the low absorption by tissue in the 700–850 nm spectral region. Lately, rigorous mathematical modeling of light propagation in tissue (see Ref. [10] for a review), combined with technological advancements in photon sources and detection techniques, has made possible the application of tomographic principles [11] for NIR three-dimensional imaging of the internal optical contrast of tissue, using a technique generally termed diffuse optical tomography (DOT) [5–22].

At the moment, powerful numerical methods are available for accurately solving the direct scattering problem [17,18,23] for arbitrary geometries, but these methods are

computationally costly. A fast method that can be applied to arbitrary geometries is needed for real time diagnostics. A good candidate is the Kirchhoff approximation (KA), also called the tangent-plane method [25,26]. This approximation is a linear method that does not involve matrix inversion while solving the forward problem. The KA can be used to generate the sensitivity functions (or weights) of the system, so that inversion schemes such as algebraic reconstruction techniques (ART) [11], amongst others, may be applied. Also, since it generates the complete Green function for any three-dimensional (3D) geometry, it is possible to apply it to improve the already existing reconstruction methods that use the Born or Rytov approximations [6–9,11–15].

The KA is a well-known approximation in physical optics that has been under study for over 30 years, and, in particular, extensively employed in studies of scattering from rough surfaces (see Ref. [26] and references therein). In these cases, the validity of the KA has been usually studied versus the angle of incidence. We here study the performance of the KA for a point source in an arbitrary diffusive medium in order to demonstrate the potential of the KA in diffuse optical tomography. We would like to state that a more rigorous study of the limits of validity of the KA would imply calculating the error for each frequency component of the incident wave, but this is out of the scope of the present paper.

In this work, we present the theory of the Kirchhoff ap-

\*Email address: jripoll@iesl.forth.gr

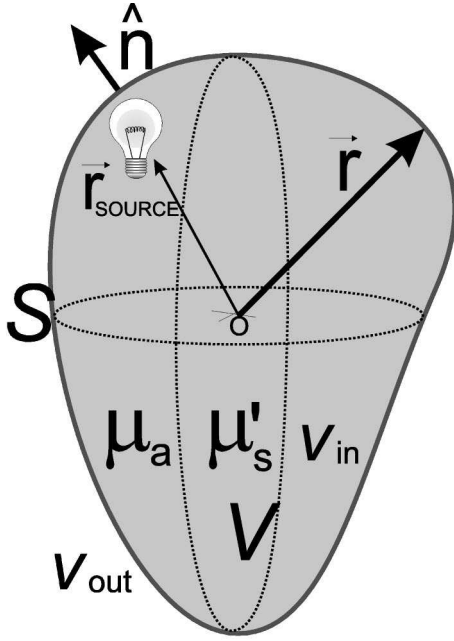


FIG. 1. Scattering geometry.

proximation in the diffusion equation context, and study its limits of validity. In Sec. II we present the exact expression for the Green function for arbitrary diffusive volumes. In Sec. III we introduce the KA specifics and derive the expression for the approximate Green function for an arbitrary geometry from the exact expression. The limits of validity of the KA are studied in Sec. IV as a function of the medium's size and optical properties. We demonstrate that these limits are independent of the geometry and depend mainly on the size of the system in diffusion length units. KA is applied to an arbitrary geometry, and compared with results obtained when employing an accurate numerical method and the infinite homogeneous Green function. We investigate the accuracy of the KA and compare the computational times of both methods, demonstrating that the KA is more than two orders of magnitude faster than accurate numerical methods and, therefore, could be a very useful tool for DOT. Finally, we conclude in Sec. V.

## II. THEORY: EXACT EXPRESSION FOR THE GREEN FUNCTION

Let us consider the geometry shown in Fig. 1, consisting of a diffusive volume  $V$  bounded by surface  $S$ , which separates it from an outer nondiffusive medium of refractive index  $\nu_{\text{out}}$ . This diffusive medium is characterized by its absorption coefficient  $\mu_a$ , the diffusion coefficient  $D = \frac{1}{3}(\mu'_s + \alpha\mu_a)$  (where  $\mu'_s$  is the reduced scattering coefficient), and the refractive index  $\nu_{\text{in}}$ . In the expression for  $D$ , the factor  $\alpha$  does not have a closed-form expression and has values that range from  $\alpha=0$  to  $\alpha=1$  depending on the approximation used (see Ref. [24] and references therein for more insight on the subject). Even so, since we have not found significant differences in the results presented here when introducing the dependence of  $D$  on absorption, we shall use  $\alpha=0$  for sim-

plicity. If the incident light impinging on the medium is modulated at a frequency  $\omega$ , the average intensity  $U(\mathbf{r}, t) = U(\mathbf{r})\exp[-i\omega t]$  detected at  $\mathbf{r}$  represents a diffuse photon density wave (DPDW) [1] and obeys the Helmholtz equation with a wave number  $\kappa = (-\mu_a/D + i\omega\nu/cD)^{1/2}$ , where  $c$  is the speed of light in vacuum, and  $\nu$  is the refractive index of the medium. In an infinite homogeneous medium, the Green function  $g$  satisfies

$$\nabla^2 g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) + \kappa^2 g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) = -4\pi\delta(\mathbf{r}_s - \mathbf{r}_d), \quad (1)$$

where  $\mathbf{r}_s$  and  $\mathbf{r}_d$  represent the source and detector points, respectively. In 3D it is well known to be

$$g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) = \frac{\exp[i\kappa|\mathbf{r}_s - \mathbf{r}_d|]}{|\mathbf{r}_s - \mathbf{r}_d|}. \quad (2)$$

In terms of the *complete* Green function  $G(\mathbf{r}_s, \mathbf{r}_d)$  that corresponds to the full geometry in Fig. 1 with boundaries, the expression of the average intensity at a point  $\mathbf{r}_d$  inside the medium is

$$U(\mathbf{r}_d) = \frac{1}{4\pi} \int_V \frac{S(\mathbf{r}')}{D} G(\mathbf{r}', \mathbf{r}_d) d\mathbf{r}', \quad \mathbf{r}_d \in V, \quad (3)$$

where  $S(\mathbf{r}')$  represents the source distribution and  $V$  is the volume occupied by the diffusive medium. Of course, for a source in infinite space  $G(\mathbf{r}_s, \mathbf{r}_d) = g(\kappa|\mathbf{r}_s - \mathbf{r}_d|)$ .

The complete Green function inside the diffusive medium can be expressed in terms of its surface integral by means of Green's Theorem (see Refs. [23, 27] for a detailed derivation) as

$$G(\mathbf{r}_s, \mathbf{r}_d) = g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) - \frac{1}{4\pi} \int_S \left[ G(\mathbf{r}_s, \mathbf{r}') \frac{\partial g(\kappa|\mathbf{r}' - \mathbf{r}_d|)}{\partial n'} - g(\kappa|\mathbf{r}' - \mathbf{r}_d|) \frac{\partial G(\mathbf{r}_s, \mathbf{r}')}{\partial \mathbf{n}'} \right] dS', \quad (4)$$

where  $\mathbf{n}'$  is the surface unit outward normal pointing into the nondiffusive medium (see Fig. 1), and  $\partial/\partial \mathbf{n}' = \mathbf{n}' \cdot \nabla_{\mathbf{r}'}$ . The boundary condition between the diffusive and nondiffusive medium in the diffusion approximation is obtained by assuming that all the flux traversing the interface is outwards into the nondiffusive medium (see Ref. [28] for a detailed derivation). This is always true as long as no sources are located outside the diffusive medium. In terms of the Green function this boundary condition is expressed as [28–30]

$$G(\mathbf{r}_s, \mathbf{r}')|_S = -C_{\text{nd}} D \frac{\partial G(\mathbf{r}_s, \mathbf{r}')}{\partial \mathbf{n}'} \Big|_S, \quad \mathbf{r}' \in S, \quad (5)$$

where the coefficient  $C_{\text{nd}}$  takes into account the refractive index mismatch between both media [28]. In the case of index matched media, i.e.,  $\nu_{\text{out}} = \nu_{\text{in}}$ ,  $C_{\text{nd}} = 2$ . Introducing Eq. (5) into Eq. (4), we obtain

$$G(\mathbf{r}_s, \mathbf{r}_d) = g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) + \frac{1}{4\pi} \int_S \left[ C_{nd} D \frac{\partial g(\kappa|\mathbf{r}' - \mathbf{r}_d)}{\partial \mathbf{n}'} \right. \\ \left. + g(\kappa|\mathbf{r}' - \mathbf{r}_d|) \right] \frac{\partial G(\mathbf{r}_s, \mathbf{r}')}{\partial \mathbf{n}'} dS'. \quad (6)$$

A rigorous solution to  $G(\mathbf{r}_s, \mathbf{r}_d)$  in Eq. (6) is found by determining the boundary value  $\partial G/\partial \mathbf{n}'$  by discretizing the surface  $S$  into a number of surface elements and inverting the resulting matrix (see Ref. [23] and references therein). Similarly to Ref. [23], Eq. (6) makes an indirect use of the extinction theorem in order to solve the system; hereon we will refer to this method as the *extinction theorem* (ET) method. The ET method gives a rigorous numerical solution to the forward problem, but is time consuming since it involves matrix inversion, and, therefore, is also limited to surfaces that can be segmented to a moderate number of discretisation points. For example, solving for more than 5000 surface points is generally excessive while considering the inverse problem, requiring about 1 h for one forward calculation on a Pentium III running at 650 MHz with 256 Mb memory. Even so, it must be understood that the computation times considered in this paper correspond to Eq. (6), which has only one unknown variable. In the case of a diffusive volume within a diffusive medium, the existence of two unknown variables (the average intensity and its derivative) increases the number of unknown variables by a factor of two. Hence, assuming that in order for the ET to give accurate results, the minimum distance between two discretization points must be in the order of the transport mean free path  $l_{tr} = 1/\mu'_s$ , the ET method would become inappropriate to solve the inverse problem for diffusive/nondiffusive surface areas in the order of 50 cm<sup>2</sup>, or 25 cm<sup>2</sup> in the case of diffusive/diffusive profiles. This fact limits the applicability of exact methods in large geometries, such as the adult head. Anyhow, the use of exact methods such as the ET is fundamental in order to validate approximate methods (see Sec. 5 of Ref. [23], where a brief discussion of the need for exact methods in optical tomography is presented). Conversely, the computing time required is practically independent of the number of detector points since Eq. (6) provides for a direct solution of detector readings along the boundary simultaneously. This is general for all surface-value dependent methods.

### III. THE KIRCHHOFF APPROXIMATION

When many forward solution calculations are required, such as in most tomographic schemes (except those reported in Ref. [16]), an approximation to Eq. (6) that can handle *arbitrary* 3D geometries in a linear fashion is needed, to reduce computing time and memory requirements. One such approximation is the KA, also known as the physical-optics or the tangent-plane method [25], which is well known and used in both optics and acoustics. This approximation assumes that the surface is replaced at each point by its tangent plane. This means that the value of the total intensity  $U$  at any point  $\mathbf{r}_p$  of the surface  $S$  is given by the sum of the incident field  $U^{(inc)}$  and the wave reflected from the *local plane* defined by the surface normal  $\mathbf{n}(\mathbf{r}_p)$  at that surface

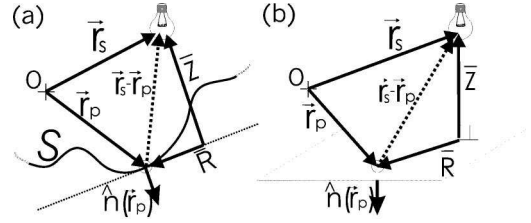


FIG. 2. (a) Detail of the local tangent plane used in the KA. (b) Representation in the coordinates of the tangent plane.

point. In terms of the Green function this is expressed as

$$G^{KA}(\mathbf{r}_s, \mathbf{r}_p) = g(\kappa|\mathbf{r}_s - \mathbf{r}_p|) * [1 + R_{ND}(\mathbf{r}_p)], \quad (7)$$

where  $*$  denotes convolution, and  $R_{ND}$  is the reflection coefficient for diffusive waves in real space, which in Fourier space has the expression for each plane wave component of  $g$  [31],

$$R_{ND}(\mathbf{K}) = \frac{i C_{nd} D \sqrt{\kappa^2 - \mathbf{K}^2} + 1}{i C_{nd} D \sqrt{\kappa^2 + \mathbf{K}^2} - 1}. \quad (8)$$

In a similar manner, the gradient of the Green function is

$$\frac{\partial G^{KA}(\mathbf{r}_s, \mathbf{r}_p)}{\partial \mathbf{n}_p} = \frac{\partial g(\kappa|\mathbf{r}_s - \mathbf{r}_p|)}{\partial \mathbf{n}_p} * [1 - R_{ND}(\mathbf{r}_p)], \quad (9)$$

the minus sign takes into consideration the different propagation directions of the incident and reflected wave with respect to the local plane. Equations (7) and (9) are directly expressed in Fourier space as

$$G^{KA}(\mathbf{r}_s, \mathbf{r}_p) = \int_{-\infty}^{+\infty} [1 + R_{ND}(\mathbf{K})] \bar{g}(\mathbf{K}, \bar{Z}) \exp[i\mathbf{K} \cdot \bar{\mathbf{R}}] d\mathbf{K}, \\ \frac{\partial G^{KA}(\mathbf{r}_s, \mathbf{r}_p)}{\partial \hat{\mathbf{n}}_p} = \int_{-\infty}^{+\infty} [1 - R_{ND}(\mathbf{K})] \frac{\partial \bar{g}(\mathbf{K}, \bar{Z})}{\partial \bar{Z}} \exp[i\mathbf{K} \cdot \bar{\mathbf{R}}] d\mathbf{K}. \quad (10)$$

In order to numerically perform the Fourier transforms in Eq. (10), a typical number of values for  $\mathbf{K}$  is 512 for each dimension, with  $|d\mathbf{K}| \sim 0.123 \text{ cm}^{-1}$ , which corresponds to a spatial discretization value of  $|d\mathbf{R}| = 0.1 \text{ cm}$ . The need for a low number of  $\mathbf{K}$  values is due to the fact that DPDW's are highly damped and do not contain high spatial frequencies. 256 values for  $\mathbf{K}$  were also tested, finding differences smaller than 1%. In all cases presented in this work 512  $\mathbf{K}$  values were employed. In Eq. (10)  $(\bar{\mathbf{R}}, \bar{Z})$  are the coordinates of  $|\mathbf{r}_s - \mathbf{r}_p|$  with respect to the plane defined by  $\mathbf{n}(\mathbf{r}_p)$  as shown in Fig. 2, namely,

$$\bar{Z} = (\mathbf{r}_s - \mathbf{r}_p) \cdot [-\hat{\mathbf{n}}(\mathbf{r}_p)], \\ \bar{\mathbf{R}} = \bar{Z} - (\mathbf{r}_s - \mathbf{r}_p). \quad (11)$$

In Eq. (10) the Fourier transform  $\bar{g}(\mathbf{K}, z)$  of the 3D homogeneous Green function  $g(\kappa|\mathbf{r}_s - \mathbf{r}_p|)$  is given by [20,21,31,32]



$$\bar{g}(\mathbf{K}, \bar{Z}) = \frac{i}{2\pi} \frac{\exp[i\sqrt{\kappa^2 - \mathbf{K}^2}|\bar{Z}|]}{\sqrt{\kappa^2 - \mathbf{K}^2}},$$

$$\frac{\partial \bar{g}(\mathbf{K}, \bar{Z})}{\partial \bar{Z}} = \frac{1}{2\pi} \exp[i\sqrt{\kappa^2 - \mathbf{K}^2}|\bar{Z}|]. \quad (12)$$

Therefore, if we discretise the surface  $S$  in Eq. (6) into  $N$  area elements  $\Delta S$ , we can write the complete Green function given by Eq. (6) inside the volume  $V$  in terms of the KA as

$$G^{\text{KA}}(\mathbf{r}_s, \mathbf{r}_d) = g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) + \frac{\Delta S}{4\pi} \sum_{p=1}^N \left[ C_{\text{nd}} D \frac{\partial g(\kappa|\mathbf{r}_p - \mathbf{r}_d|)}{\partial \mathbf{n}_p} + g(\kappa|\mathbf{r}_p - \mathbf{r}_d|) \right] \frac{\partial G^{\text{KA}}(\mathbf{r}_s, \mathbf{r}_p)}{\partial \mathbf{n}_p}. \quad (13)$$

Equation (13) is an explicit expression of the Green function where the computation time will increase only linearly with the system size. Also, one of the main advantages of Eq. (13) is that the values of  $\partial G^{\text{KA}}/\partial \mathbf{n}_p$  given by Eq. (10) need only to be calculated once for all possible source-plane distances  $\bar{Z}$  and  $\bar{\mathbf{R}}$  values present in the geometry, recalling or interpolating its value each time the source and plane positions  $\mathbf{r}_s$  and  $\mathbf{r}_p$  hold Eq. (11). This considerably increases the computation speed by reducing the number of Fourier transforms, especially in the cases in which many different source positions need to be generated, such as in DOT. We would like to state that an analogous expression to Eq. (13) can be easily found for diffusive/diffusive interfaces by means of the diffusive/diffusive reflection and transmission coefficients [33].

#### IV. NUMERICAL RESULTS

In order to study the limits of validity of the KA, we compare the performance of the exact solution, based on Eq. (6), with the approximate solution, based on Eq. (13), using the geometry shown in Fig. 3. The cylinder has a radius  $R$ , length  $h$ , and is illuminated by an infinite longitudinal light source running parallel to the cylinder at ( $R_s = R - l_{\text{tr}}$ ,  $\theta = 0$ ), where  $l_{\text{tr}} = 1/\mu'_s$  is the transport mean free path. The refractive indices inside and outside the diffusive volume are that of water, i.e.,  $\nu_{\text{in}} = 1.333$ , and of air,  $\nu_{\text{out}} = 1$ , respectively. An angular scan is performed at ( $R_d = R - l_{\text{tr}}$ ,  $z = 0$ ). In order to quantify the accuracy of the approximation, we shall define the error in percentage as

$$(\text{Error}) = 100 \times \int_{2\pi} |1 - U^{\text{KA}}(R_d, \theta)/U^{\text{ET}}(R_d, \theta)| d\theta, \quad (14)$$

where  $U^{\text{ET}}$  is the exact solution obtained from the ET [23] using 2D Green functions (corresponding to an infinitely long cylinder), and  $U^{\text{KA}}$  is the solution obtained from the KA using a 3D geometry (corresponding to a cylinder of length  $h$ ). In order to solve for the ET by means of Eq. (6) for a cylinder and a line source, we used the corresponding Green

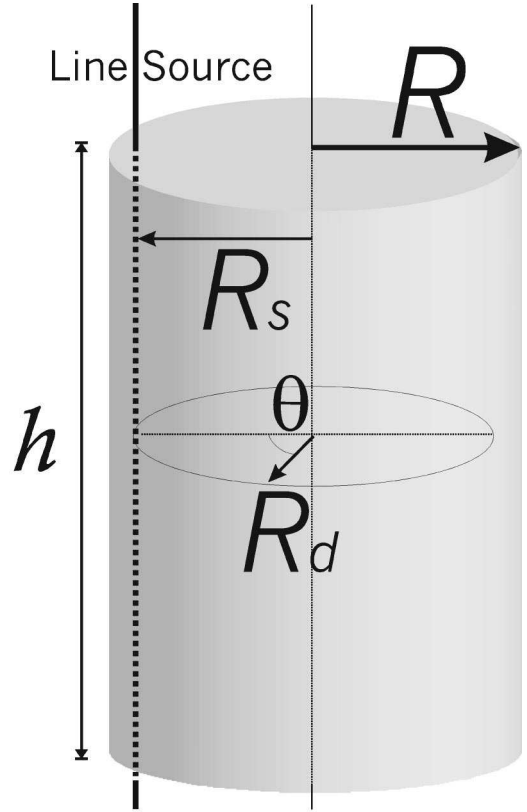


FIG. 3. Geometry used for the study of the limits of validity of the KA.

function in 2D,  $g(\kappa|\mathbf{r}_s - \mathbf{r}_d|) = \pi i H_0^{(1)}(\kappa|\mathbf{r}_s - \mathbf{r}_d|)$ , where  $H_0^{(1)}$  is the zero-order Hankel function of the first kind.

In all cases, the KA results will be generated for a cylinder of height  $h = 10$  cm and no lids, consisting of  $N = 9191$  surface discretisation points. The value of  $h$  was such that no variation in the results was found by increasing its value. The results generated with the ET in 2D consisted of  $N = 360$  points. In these cases, a lookup table consisting of 257 values for  $\bar{\mathbf{R}}$  and a maximum value for  $\bar{Z}$  of  $R$  was generated, with a distance of 0.1 cm between values. We performed the study in the continuous illumination mode (CW), since in this modality light suffers less attenuation. For higher attenuation values the multiple reflections between the surface boundary decays, and, therefore, the limits of validity here found will apply to all frequency modulation values. A similar study was performed for different modulation frequencies, finding the error in amplitude in the order or smaller than in its corresponding CW case, and a difference in phase in the order of 1 to 5 degrees.

In Fig. 4 we show the error committed by the KA for different values of  $R$ , absorption, and scattering coefficients as compared to the ET solution. The results shown here are representative of biologically relevant optical properties. On the whole, the approximation works well ( $< 5\%$  error) for  $R > 3L_d$ , where  $L_d = \sqrt{D/\mu_a}$  is the diffusion length in CW ( $\omega = 0$ ). That is, to maintain an error below 5% for  $R = 1.5$  cm,  $L_d$  should be larger than 0.5 cm for  $\mu'_s = 5 \text{ cm}^{-1}$ , which gives  $\mu'_s > 0.13 \text{ cm}^{-1}$  (see Fig. 4). When diffusive/

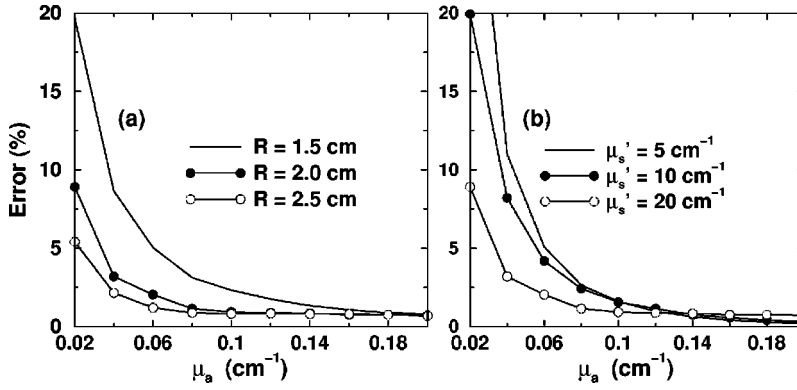


FIG. 4. (a) Error in percentage committed for different values of  $\mu_a$  and cylinder radii  $R$ . In all cases  $\mu_s' = 10 \text{ cm}^{-1}$ . (b) Same as (a) but for different values of  $\mu_s'$  for  $R = 1 \text{ cm}$ .

diffusive interfaces are considered the approximation works much better, and is valid for  $2R > 3L_d$  because lower reflectivity is attained in this case (similar results were obtained in Ref. [31]).

In order to establish the effect of the surface topography, we have studied the same configuration as in Fig. 4, but have added a sine profile of amplitude 0.5 cm and period  $\pi/4$  to the surface. In this case, a lookup table consisting of 257 values for  $\bar{\mathbf{R}}$  and 76 values for  $\bar{\mathbf{Z}}$  was generated, with a distance of 0.1 cm between values. The error between the ET and the KA, is now depicted in Fig. 5 for two cases of source position (Fig. 5 top row  $R_s = 2.3 \text{ cm}$ , bottom row  $R_s = 1.5 \text{ cm}$ ). In addition, we have also represented this error when, instead of the KA, we simply use the homogeneous Green function, Eq. (2), to calculate the source radiation using Eq. (3). Generally, the KA approximation calculates the average intensity with errors that are less than 5% [see Fig. 5(a) and Fig. 5(c)], except in the shadow regions of the corrugations. This *shadowing effect* appears when certain surface areas are blocked from the source by the geometry of the interface. Since these shadow areas are not taken into account in the KA, the KA predicts higher values of the intensity. A first approximation to this problem would be to

include a geometry-dependent constant that will assign a zero value to those surface points not visible from the source position. Such a geometry factor was included and the results presented in Fig. 5 repeated, finding no important improvement. Another way of improving the KA is to include second order reflections, but this would render the method time consuming and thus would lose its potential as a fast analytical tool.

In Fig. 5 we also see that the approximation yields errors in the order of 5–10% close to the boundaries, where the Green function has low values due to the boundary conditions that force the average intensity to zero at approximately one extrapolated distance ( $l_{tr}$ ) from the interface [28]. When the error obtained from the KA is compared with that obtained by using a mere homogeneous Green function [see Fig. 5(b) and Fig. 5(d)], we see that the KA is more accurate by one order of magnitude. Similar figures to those represented in Fig. 5 were generated for a modulation frequency of  $\omega = 200 \text{ MHz}$ . In this case we found that the error distribution in amplitude was very similar to the cases presented in Fig. 5, with slightly smaller values, and thus results are not shown. This is expected due to the lower reflectivity of the boundaries. The maximum phase difference found for the

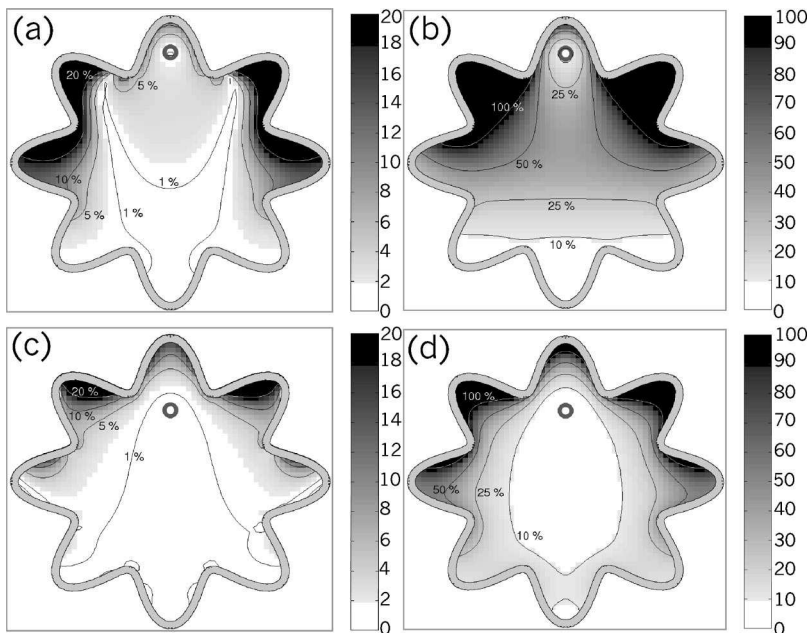


FIG. 5. Error committed in percentage when using the KA [(a) and (c)] and when using the homogeneous Green function [(b) and (d)] for a cylinder of  $R = 2.5 \text{ cm}$  with a sine profile on the boundary of amplitude 0.5 cm, and period  $\pi/4$ . The following source locations are considered: ( $R_s = 2.3 \text{ cm}$ ,  $\theta = 0$ ) [(a) and (b)]; ( $R_s = 1.5 \text{ cm}$ ,  $\theta = 0$ ) [(c) and (d)]. In all cases  $\mu_s' = 10 \text{ cm}^{-1}$ ,  $\mu_a = 0.1 \text{ cm}^{-1}$ .

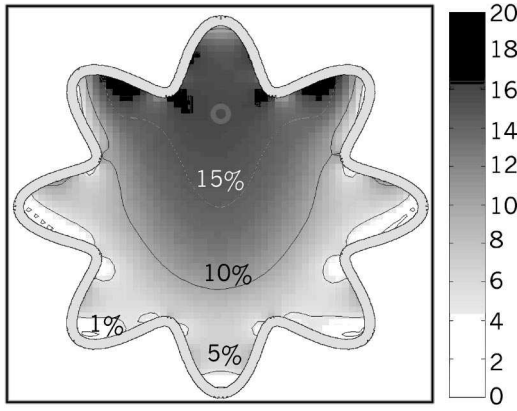


FIG. 6. Error committed in percentage for the scattered Green function [see Eq. (15)] when using the KA for a cylinder of  $R = 2.5$  cm with a sine profile in the boundary of amplitude 0.5 cm, and period  $\pi/4$ . Source locations ( $R_s = 1.5$  cm,  $\theta = 0$ ),  $\mu'_s = 10$  cm $^{-1}$ ,  $\mu_a = 0.1$  cm $^{-1}$ .

KA was 2 deg at the shadow regions and 1 deg near the boundaries. When compared to the infinite Green function we found 60 deg difference at the shadow regions and a minimum phase difference of 10 deg.

In general terms of the KA, it is expected that lower spatial frequencies  $\mathbf{K}$  (or angles of incidence close to the normal in the electromagnetic case [25,26]) will yield more accurate solutions than high spatial frequencies (or grazing angles of incidence in the electromagnetic case [25,26]). This may be translated to diffusive waves in the following manner. In the cases in which the point source is close to the interface, high spatial frequencies play an important role. It is in these cases in which the KA is expected to fail for diffusive waves, since then multiple reflections are predominant. On the other hand, when the source is far (more than one diffusion length) from the interface, due to high damping the high frequencies components of the incident wave do not contribute significantly to the incident wave at the interface, reducing the multiple reflections. This effect is shown in Fig. 6, where we represent the error when considering the perturbation caused by the interface, i.e., the scattered wave  $G_{SC}$  [see Eq. (13)]:

$$G_{SC}^{KA}(\mathbf{r}_s, \mathbf{r}_d) = G^{KA}(\mathbf{r}_s, \mathbf{r}_d) - g(\kappa|\mathbf{r}_s - \mathbf{r}_d|). \quad (15)$$

As seen in Fig. 6, the error obtained from the KA at long distances from the source where only low values of  $\mathbf{K}$  contribute to the incident field is very low ( $<5\%$ ). On the other hand, values of the scattered wave in the regions where the source is near the interface present larger errors (20%), due to the higher contribution of large values of  $\mathbf{K}$ .

We have tested other values of the period and the amplitude of the sine profile, reaching the same conclusion: outside the shadow regions, and for  $R > 3L_d$ , the error is consistently less than 5%. This also holds true for calculations performed for a rough surface plane, such as in Ref. [34].

As mentioned before, besides its ability to handle arbitrary geometries the KA is attractive due to its computation efficiency. In Fig. 7 we present a comparison of the computation times obtained by using the ET and the KA with a

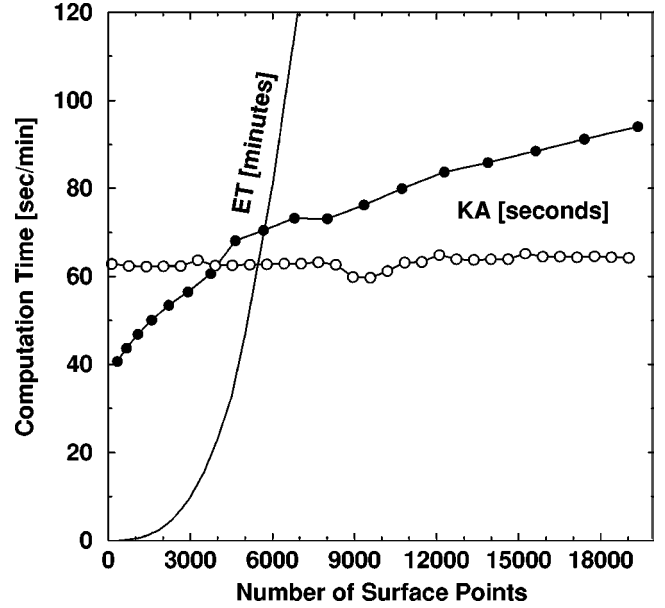


FIG. 7. Computation times for one source-detector pair versus the number of surface points  $N$  [see Eq. (13)] for the ET represented in minutes [solid line] and for the KA represented in seconds for the cases: cylinder of radius  $R = 2$  cm, with its height increased from 0–15 cm [open circles]; and a cylinder with its radius increased from  $R = 0.5$ –4 cm and its height as  $h = 2R$  [full circles]. Results obtained from a Pentium III running at 650 MHz with 256 Mb memory.

Pentium III running at 650 MHz, with 256 Mb memory. These computation times include the generation of the lookup table for  $(\bar{\mathbf{R}}, \bar{\mathbf{Z}})$  [see Eq. (11)] aforementioned in Sec. III. That is, the computation times presented make use of no *a priori* calculations. In all cases shown here, the lookup table for  $(\bar{\mathbf{R}}, \bar{\mathbf{Z}})$  is generated by finding the range of values  $[\min\{\bar{\mathbf{R}}\}, \max\{\bar{\mathbf{R}}\}]$  and  $[\min\{\bar{\mathbf{Z}}\}, \max\{\bar{\mathbf{Z}}\}]$  present in the geometry, and generating all the corresponding values of Eq. (10), with an increase of one transport mean free path ( $l_{tr}$ ) between  $(\bar{\mathbf{R}}, \bar{\mathbf{Z}})$  values. This discretisation value can be understood since the diffusion equation in itself has no meaning when considering distances smaller than  $l_{tr}$ . Once the lookup table for Eq. (10) is built, the different values present in Eq. (13) are found by interpolation. As mentioned in Sec. III, in order to numerically perform the Fourier transforms 512  $\mathbf{K}$  values were used for each dimension (i.e.,  $512 \times 512$ ), with  $|d\mathbf{K}| \sim 0.123$ . The computation times are represented in Fig. 7 versus the number of discretisation points  $N$  [see Eqs. (6) and (13)], which in the ET are independent on the dimension and shape of the geometry. In the case of the KA, since the computing time is dependent on the number  $(\bar{\mathbf{R}}, \bar{\mathbf{Z}})$  values, we present two cases: a case in which we have a cylinder of radius  $R = 2$  cm, and increase its height from 0–15 cm (open circles in Fig. 7); and a cylinder that is increased in radius  $R$  from 0.5–4 cm and its height as  $h = 2R$  (full circles in Fig. 7). In all the KA cases the discretized areas are kept to be  $dS = 0.1 \times 0.1$  cm $^2$ . Due to the fact that both the KA and the ET perform equivalently for any number of source-detector pairs once the surface values are found, the computation

times presented in Fig. 7 correspond to the forward problem for one source-detector pair. When considering the computational times for the ET in Fig. 7, we see that there is an approximately quadratic increase with respect to  $N$  (note the difference in scale between the KA and the ET). On the other hand, when considering the KA, we see that the increase is approximately linear for both cases, the difference in slope due to the dependence on the number of  $(\bar{\mathbf{R}}, \bar{Z})$  values generated. As a practical example, the number of discretization points for a sphere of radius 2 cm needed in order to maintain a  $1/l_{tr}$  distance between points is in the order of 5000. If we compare the speed of the KA and the ET in this case we obtain 70 s and 50 min, respectively, yielding the KA as approximately 40 times faster. A more realistic surface such as the adult head, would imply an equivalent radius of at least 4 cm, and thus  $N \sim 20\,000$ . In this case, the KA takes in the order of 90 s, whereas the ET takes in the order of 45 h for one only forward solution. In this more realistic case the KA is 1800 times faster.

## V. CONCLUSIONS

We have presented an approximate method that solves the 3D diffusion equation in geometries of arbitrary shape and size in a linear fashion. This approximation has been compared to the ET solution of the diffusion equation [23], a boundary-value dependent numerical method that has been extensively used in physical optics due to its high degree of accuracy [25]. We have found that when the average radius of the geometry considered is  $R > 3(D/\mu_a)^{1/2}$ , the method performs with an error less than 5%. Therefore, with the KA we can generate general Green functions that take into account complex geometries, the computation times of this ap-

proximation being very fast compared to the rigorous solution, and increasing linearly with the size of the system. The implications of this approximation are several: In the first place, these KA Green functions can be employed in more complex numerical schemes such as the ET [23], so as to reduce the number of discretization points needed to solve the forward problem. As an example, the problem of an object embedded in an arbitrary volume would be reduced to an object on its own by using the KA Green function. In a similar manner, it can be used to improve the reconstruction schemes based on Rytov or Born approximations, such as algebraic reconstruction technique (ART) and simultaneous iterative reconstructive technique (SIRT) [11–15]. Second, since the computation times and the memory requirements increase linearly with the size of the system, the KA may be used to describe light propagation in large volumes such the adult head, the calf, etc. It is in these large volumes where rigorous numerical methods have problems due to the great amount of memory required for matrix inversion, and the need of extremely large computational times to solve the inverse problem. We believe that this approximation will aid to the development of real time diagnostics with diffuse light in the presence of complex boundaries.

## ACKNOWLEDGMENTS

J.R. acknowledges a European TMR grant under Project No. FMRX-CT96-. V.N. acknowledges support in part from DRG-1638 of the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation and the U.S. Army (CDMRP BC995360). V.N. and M.N.-V. also acknowledge a TMR contract from the EU. This work has received partial support from the Spanish DGICYT.

- 
- [1] A. Yodh and B. Chance, *Phys. Today* **48**(3), 38 (1995).
  - [2] S. K. Gayen and R. R. Alfano, *Opt. Photonics News* **7**, 17 (1996).
  - [3] E. B. de Haller, *J. Biomed. Opt.* **1**, 7 (1996).
  - [4] M. A. Franceschini, K. T. Moesta, S. Fantini, G. Gaida, E. Gratton, H. Jess, W. W. Mantulin, M. Seeber, P. M. Schlag, and M. Kaschke, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 6468 (1997).
  - [5] D. Grosenick, H. Wabnitz, H. H. Rinneberg, T. Moesta, and P. M. Schlag, *Appl. Opt.* **38**, 2927 (1999).
  - [6] V. Ntziachristos, A. G. Yodh, M. Schnall and B. Chance, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 2767 (2000).
  - [7] V. Ntziachristos and B. Chance, *Breast Cancer Res. Treat.* **3**, 41 (2001).
  - [8] V. Ntziachristos, X. H. Ma, A. G. Yodh, and B. Chance, *Rev. Sci. Instrum.* **70**, 193 (1999).
  - [9] V. Ntziachristos, X. H. Ma, and B. Chance, *Rev. Sci. Instrum.* **69**, 4221 (1998).
  - [10] S. R. Arridge, *Inverse Probl.* **15**, R41 (1999).
  - [11] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging* (IEEE, New York, 1988).
  - [12] M. A. O’Leary, D. A. Boas, B. Chance, and A. G. Yodh, *Opt. Lett.* **20**, 426 (1995).
  - [13] D. A. Boas, L. E. Campbell, and A. G. Yodh, *Phys. Rev. Lett.* **75**, 1855 (1995).
  - [14] M. A. O’Leary, D. A. Boas, X. D. Li, B. Chance, and A. G. Yodh, *Opt. Lett.* **21**, 158 (1996).
  - [15] X. De Li, T. Durduran, A. G. Yodh, B. Chance, and D. N. Pattanayak, *Opt. Lett.* **22**, 573 (1997).
  - [16] C. P. Gonatas, M. Ishii, J. S. Leigh, and J. C. Schotland, *Phys. Rev. E* **52**, 4361 (1995).
  - [17] S. R. Arridge, *Appl. Opt.* **34**, 7395 (1995).
  - [18] H. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, *J. Opt. Soc. Am. A* **13**, 253 (1996).
  - [19] J. Chang, H. L. Graber, and R. L. Barbour, *IEEE Trans. Biomed. Eng.* **44**, 810 (1997).
  - [20] T. Durduran, J. P. Culver, M. J. Holboke, X. D. Li, L. Zubkov, B. Chance, D. N. Pattanayak, and A. G. Yodh, *Opt. Express* **4**, 247 (1999).
  - [21] X. Li, D. N. Pattanayak, T. Durduran, J. P. Culver, B. Chance, and A. G. Yodh, *Phys. Rev. E* **61**, 4295 (2000).
  - [22] D. L. Everitt, S. Wei, and X. D. Zhu, *Phys. Rev. E* **62**, 2924 (2000).
  - [23] J. Ripoll and M. Nieto-Vesperinas, *J. Opt. Soc. Am. A* **16**, 1453 (1999).



- [24] R. Aronson, J. Opt. Soc. Am A **116**, 61 (1999).
- [25] P. Beckmann, Prog. Opt. **61**, 55 (1961).
- [26] J. M. Soto-Crespo and M. Nieto-Vesperinas, J. Opt. Soc. Am. A **6**, 367 (1989).
- [27] M. Nieto-Vesperinas, *Scattering and Diffraction in Physical Optics* (Pergamon, New York, 1996).
- [28] R. Aronson, J. Opt. Soc. Am. A **12**, 2532 (1995).
- [29] N. G. Chen and J. Bai, Phys. Rev. Lett. **80**, 5321 (1998).
- [30] G. Popescu, C. Mujat, and A. Dogariu, Phys. Rev. E **61**, 4523 (2000).
- [31] J. Ripoll, V. Ntziachristos, J. P. Culver, D. N. Pattanayak, A. G. Yodh, and M. Nieto-Vesperinas, J. Opt. Soc. Am. A **18**, 821 (2001).
- [32] J. Ripoll, M. Nieto-Vesperinas, and R. Carminati, J. Opt. Soc. Am. A **16**, 1466 (1999).
- [33] J. Ripoll and M. Nieto-Vesperinas, Opt. Lett. **24**, 796 (1999).
- [34] J. Ripoll and M. Nieto-Vesperinas, J. Opt. Soc. Am. A **16**, 1947 (1999).

# Time-dependent transport through scattering media: from radiative transfer to diffusion

Rachid Elaloufi, Rémi Carminati and Jean-Jacques Greffet

Laboratoire d’Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France

Received 17 January 2002

Published 14 August 2002

Online at [stacks.iop.org/JOptA/4/S103](http://stacks.iop.org/JOptA/4/S103)

## Abstract

We study the propagation of light pulses through scattering media using the time-dependent radiative transfer equation. A standard discrete-ordinate method is used to solve this equation in the space-frequency domain. We present calculations of diffuse transmission through scattering slabs, in the presence of absorption and anisotropic scattering. We show that the diffusive regime is attained at long times only for thick slabs. Comparisons with diffusion theory show that the proper choice of the diffusion constant is an important issue for time-dependent transport.

**Keywords:** Light diffusion, radiative transfer equation, scattering media, biomedical imaging, diffusion coefficient

## 1. Introduction

Radiative transfer through scattering media has attracted considerable interest recently, particularly for imaging applications through turbid media [1]. Several techniques have been developed in order to determine the location of objects in strongly scattering biological tissues, using visible or near-infrared light [2, 3]. Pulse transmission measurements on short timescales and optical coherence tomography give promising results [4–7]. In several other areas, diffusion waves—such as thermal, acoustic or elastic waves—form the basis of imaging and measurement techniques [8]. With the rapid development of micro- and nano-technologies, understanding the propagation of such waves at short (time and length) scales has become a key issue. Heat conduction at short scales in solids is also handled on the basis of a Boltzman transport equation for phonons, which undergo scattering, emission and absorption [9, 10]. Transport theories in all these topics have in common the use of a Boltzman-like transport equation for the wave intensity which describes scattering, absorption and emission by the medium. This equation is usually referred to as the radiative transfer equation (RTE), whose formalism was first developed for astrophysics [11] and neutron transport [12].

Solving the RTE in complex geometries and in the presence of scattering remains a challenging issue. For time-dependent light transport, some methods have been developed

recently, see e.g. [13–15]. The diffusion approximation [16] offers a great simplification of the problem and a practical tool to describe the diffuse part of the radiation intensity. This approximation is widely used, for example, in optical imaging for biomedical applications [2, 3]. Yet, the use of the diffusion approximation to handle short time and length scales is questionable, and its domain of validity has to be studied carefully. Moreover, although the diffusion approximation may be derived starting from different approaches, the proper definition of the diffusion coefficient is still an open issue. In particular, its dependence on absorption was questioned recently [17].

In this work, we study radiative transfer of visible or near-infrared light through strongly scattering slabs, in the presence of absorption. We first describe a method to solve the time-dependent RTE, based on a discrete-ordinate method [11] in the space-frequency domain. We show that this approach is well suited to the study of pulse transmission and reflection through scattering slabs of arbitrary optical thicknesses. Then, we study the validity of the diffusion approximation for time-dependent transport. We show that the diffusive regime is recovered for the long-time behaviour of transmitted pulses *and* for thick slabs. We compare the results obtained for different expressions of the diffusion coefficient. We show that the use of the proper expression is a crucial issue for time-dependent transport, especially in the presence of absorption and anisotropic scattering.

## 2. The radiative transfer equation

The basic quantity of radiative transfer theory is the specific intensity  $I_\nu(\mathbf{r}, \mathbf{u}, t)$ , from which the power radiated at point  $\mathbf{r}$ , through a surface element  $dS$ , of unit normal  $\mathbf{n}$ , at time  $t$ , in the frequency interval  $[\nu, \nu + d\nu]$  and in a solid angle  $d\Omega$  around the direction  $\mathbf{u}$  writes

$$dP = I_\nu(\mathbf{r}, \mathbf{u}, t) \mathbf{u} \cdot \mathbf{n} dS d\Omega d\nu. \quad (1)$$

The specific intensity obeys a transport equation, the RTE, which describes its variations due to absorption, scattering and emission [11, 18]. In this work, we study the propagation of visible and near-infrared light in absorbing and scattering media, such as those encountered, for example, in optical imaging through biological tissues. At room temperature, thermal emission is negligible at these wavelengths, so that only absorption and scattering has to be considered. Note that the validity of the RTE to study light transport through scattering slabs was recently studied numerically by comparison with exact electromagnetic simulations [19]. It was shown that, except for coherent effects such as back-scattering enhancement, the RTE gives very accurate results even for a system whose geometric thickness is of the order of one wavelength.

### 2.1. Radiative transfer equation in a slab geometry

The geometry we consider is depicted in figure 1. A slab of width  $L$  containing a scattering and absorbing medium is illuminated at normal incidence by a plane-wave pulse. The absorption and diffusion coefficients of the medium are  $\mu_a$  and  $\mu_s$ , respectively. The associated absorption and scattering mean-free paths are  $l_a = \mu_a^{-1}$  and  $l_s = \mu_s^{-1}$ . The real part of its effective index, accounting both for a homogeneous background medium and scattering particles, is denoted by  $n_2$ . The half-spaces  $z < 0$  and  $z > L$  are filled with homogeneous and transparent materials of refractive indices  $n_1$  and  $n_3$ , respectively. For this geometry, the RTE gives

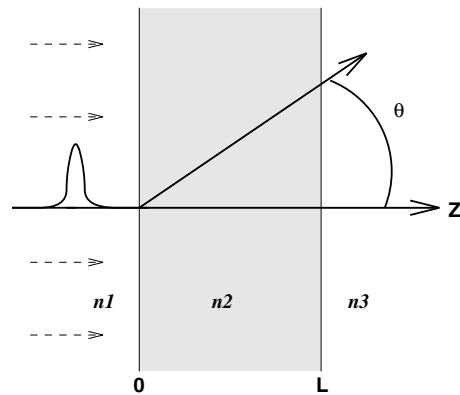
$$\frac{1}{\beta v} \frac{\partial I_\nu(\tau, \mu, t)}{\partial t} + \mu \frac{\partial I_\nu(\tau, \mu, t)}{\partial \tau} = -I_\nu(\tau, \mu, t) + \frac{a}{2} \int_{-1}^{+1} p^{(0)}(\mu, \mu') I_\nu(\tau, \mu', t) d\mu' \quad (2)$$

where  $v$  is the transport velocity,  $\mu = \cos \theta$ ,  $\beta = \mu_a + \mu_s$  is the extinction coefficient,  $a = \mu_s/\beta$  is the albedo and  $\tau = \beta z$  is the optical depth.  $p^{(0)}$  is the phase function integrated over the azimuthal angle:

$$p^{(0)}(\mu, \mu') = \frac{1}{2\pi} \int_0^{2\pi} p(\mathbf{u} \cdot \mathbf{u}') d\phi \quad (3)$$

where  $\mathbf{u}$  and  $\mathbf{u}'$  are unit vectors corresponding to directions  $(\theta, \phi)$  and  $(\theta', \phi')$ , respectively. A useful parameter is the averaged scattering angle (or anisotropy factor)  $g$  defined by  $g = (4\pi)^{-1} \int (\mathbf{u} \cdot \mathbf{u}') p(\mathbf{u} \cdot \mathbf{u}') d\Omega'$ . Strongly forward-scattering corresponds to  $g = 1$ , and isotropic scattering to  $g = 0$ . In order to solve the RTE in the space-frequency domain, we introduce the time-domain Fourier transform of the specific intensity:

$$I_\nu(\tau, \mu, t) = \int_{-\infty}^{+\infty} I_\nu(\tau, \mu, \omega) \exp(-i\omega t) d\omega. \quad (4)$$



**Figure 1.** The geometry of the system. The scattering slab is illuminated from the left by a plane-wave pulse, at normal incidence.

Note that the hypothesis of a quasi-monochromatic pulse is apparent in equation (4). The subscript  $\nu$  refers to the central frequency of the pulse, and the variable  $\omega$  describes the frequency analysis of the temporal pulse shape. We assume non-dispersive materials within the pulse bandwidth, so that the optical properties of the medium are evaluated at the central frequency  $\nu$ . In the following, we omit the subscript  $\nu$  in order to simplify the notations. Introducing equation (4) into (2) leads to

$$\mu \frac{\partial I(\tau, \mu, \omega)}{\partial \tau} = -\left(1 - i \frac{\omega}{\beta v}\right) I(\tau, \mu, \omega) + \frac{a}{2} \int_{-1}^{+1} p^{(0)}(\mu, \mu') I(\tau, \mu', \omega) d\mu'. \quad (5)$$

This equation has the same structure as the static RTE, with a complex specific intensity and a complex extinction coefficient  $\alpha(\omega) = 1 - i\omega/(\beta v)$ . It can be solved by standard methods developed for the static RTE for each frequency  $\omega$ . An inverse Fourier transform allows one to recover afterwards the time-dependent solution. A similar approach was used in [20] in the case of a strongly forward-scattering medium, and more recently in [21] in the case of polarized light transfer through a scattering slab.

Assuming an illumination by a plane wave (representing, for example, a collimated laser beam) it is useful to separate the collimated and the diffuse components of the specific intensity inside the medium. One writes

$$I(\tau, \mu, \omega) = I_{coll}^+(\tau, \omega) \delta(\mu - 1) + I_{coll}^-(\tau, \omega) \delta(\mu + 1) + I_d(\tau, \mu, \omega) \quad (6)$$

where  $\delta(x)$  is the Dirac distribution. For the sake of clarity, the two components of the collimated beam, propagating towards  $z > 0$  and  $z < 0$ , have been separated. Inserting equation (6) into (5) leads to

$$\frac{dI_{coll}^\pm(\tau, \omega)}{d\tau} = -\alpha(\omega) I_{coll}^\pm(\tau, \omega) \quad (7)$$

for the collimated components and to

$$\mu \frac{\partial I_d(\tau, \mu, \omega)}{\partial \tau} = -\left(1 - i \frac{\omega}{\beta v}\right) I_d(\tau, \mu, \omega) + \frac{a}{2} \int_{-1}^{+1} p^{(0)}(\mu, \mu') I_d(\tau, \mu', \omega) d\mu' + S(\tau, \mu, \omega) \quad (8)$$

for the diffuse component, where  $S(\tau, \mu, \omega)$  is a source term which describes the transfer of energy from the collimated to the diffuse component by scattering. Equations (7) and (8) are the RTE for the collimated and diffuse components of the specific intensity, respectively.

## 2.2. The source term and boundary conditions

Taking into account the internal reflections at the boundary of the slab, the expressions of the collimated components are

$$I_{coll}^+(\tau, \omega) = T_{12}(\mu = 1)I_0(\omega) \exp[-\alpha(\omega)\tau]\Gamma \quad (9)$$

$$I_{coll}^-(\tau, \omega) = T_{12}(\mu = 1)I_0(\omega) \times \exp[-\alpha(\omega)(2\beta L - \tau)]R_{23}(\mu = 1)\Gamma \quad (10)$$

where  $\Gamma = [1 - R_{12}(\mu = 1)R_{23}(\mu = 1) \exp(-2\alpha(\omega)\beta L)]^{-1}$ , and  $R_{ij}(\mu)$  and  $T_{ij}(\mu)$  are the Fresnel reflection and transmission factors in energy, at the interface between two media of refractive indices  $n_i$  and  $n_j$ . Their expression is given in the appendix.  $I_0(\omega)$  is the time-domain Fourier transform of the incident pulse, at the boundary  $z = 0$ .

The source term in equation (8) is given by

$$S(\tau, \mu, \omega) = \frac{a}{2}p^{(0)}(\mu, 1)I_{coll}^+(\tau, \omega) + \frac{a}{2}p^{(0)}(\mu, -1)I_{coll}^-(\tau, \omega). \quad (11)$$

For the diffuse components of the specific intensity, the boundary conditions at the slab boundaries are

$$I_d(\tau = 0, \mu, \omega) = R_{21}(\mu)I_d(\tau = 0, -\mu, \omega) \quad \text{for } \mu > 0 \quad (12)$$

$$I_d(\tau = \beta L, \mu, \omega) = R_{23}(|\mu|)I_d(\tau = \beta L, -\mu, \omega) \quad \text{for } \mu < 0. \quad (13)$$

## 2.3. Numerical calculation

In order to solve equation (8), we have used a discrete-ordinate method. The first step in this method is to replace the integral involving the phase function by a quadrature [11]. The integro-differential equation (8) is thus replaced by a system of linear differential equations, one equation for each direction  $\mu_i$  used in the quadrature. To solve this system of differential equations, we have used the matrix eigenproblem approach described in [24].

In our case, the entire procedure is as follows:

- (i) Calculate the Fourier transform of the incident pulse, which appears in the source term in equation (8);
- (ii) Solve equation (8) for all relevant frequencies, using the discrete-ordinate method;
- (iii) Perform an inverse Fourier transform to recover the time-domain evolution of the specific intensity at each boundary of the slab.

## 2.4. Observable quantities

In an experiment, the observable quantities are the transmitted and reflected fluxes, either directional or integrated over a half-space. The transmitted and reflected collimated fluxes are defined by

$$T_{coll}(t) = T_{23}(\mu = 1)I_{coll}^+(\tau = \beta L, t) \quad (14)$$

$$R_{coll}(t) = R_{12}(\mu = 1)I_0(t) + T_{21}(\mu = 1)I_{coll}^-(\tau = 0, t) \quad (15)$$

where  $I_0(t)$  is the incident pulse shape at the boundary  $z = 0$ . In the following, we will concentrate on the diffuse transmitted and reflected fluxes. They are defined by

$$T_d(t) = \frac{n_3^2}{n_2^2} \int_0^1 \mu T_{23}(\mu') I_d(\tau = \beta L, \mu', t) d\mu \quad (16)$$

$$R_d(t) = \frac{n_1^2}{n_2^2} \int_{-1}^0 \mu T_{21}(|\mu'|) I_d(\tau = 0, \mu', t) d\mu \quad (17)$$

where  $\mu' = [1 + (n_3^2/n_2^2)(\mu^2 - 1)]^{1/2}$  in equation (16) and  $\mu' = -[1 + (n_1^2/n_2^2)(\mu^2 - 1)]^{1/2}$  in equation (17).

## 3. Diffusion approximation

The diffusion approximation is usually obtained from transport theory in the limit of smooth space and time variations of the specific intensity, compared to the scattering mean-free path  $l_s$  and the microscopic timescale  $(l_s/v)$  [16]. In this approximation, the energy density  $U(\mathbf{r}, t) = v^{-1} \int I(\mathbf{r}, \mathbf{u}, t) d\Omega$  obeys the diffusion equation

$$\frac{\partial U(\mathbf{r}, t)}{\partial t} - D\nabla^2 U(\mathbf{r}, t) + v\mu_a U(\mathbf{r}, t) = q(\mathbf{r}, t) \quad (18)$$

where  $D$  is the diffusion coefficient and  $q(\mathbf{r}, t)$  a source term.

In the diffusion approximation, the diffuse transmission through a scattering slab can be evaluated analytically using the method of images and extrapolated boundary conditions [22]. For a source term of the form  $q(\mathbf{r}, t) = \delta(z)\delta(t)$ , one obtains

$$T(t) = \frac{H(t)D}{d} \exp(-\mu_a vt) \sum_{m=1}^{\infty} \frac{\pi m}{d} \sin\left(\frac{\pi mL}{d}\right) \times \exp\left(-\frac{\pi^2 m^2 Dt}{d^2}\right) \quad (19)$$

where  $H(t)$  is the Heavyside step function,  $d = L + 2z_0$ ,  $z_0 = 0.71l_{tr}$  being the extrapolation distance and  $l_{tr} = l_s/(1 - g)$  the transport mean-free path. The result in equation (19) is the transmission Green function (or impulse response) of the diffusion equation in a slab geometry.

The diffusion approximation is very robust, in the sense that it can be derived from any transport theory as the limit of smooth spatial and time variations [12, 16, 18, 23]. However, all derivations do not necessarily lead to the same expression for the diffusion coefficient  $D$ . In particular, the dependence on absorption may change from one expression to the other [17]. Using numerical results, we will show in the following that the correct definition of the diffusion coefficient is a crucial issue for time-dependent transport.

## 4. Numerical calculations of impulse responses

In this section, we present numerical calculations of diffuse transmission, and compare them to the results of diffusion theory. In all cases, the calculated quantities are impulse responses, namely, the incident pulse is a delta function in time. The validity of diffusion theory for time-dependent light transport through scattering media has been studied previously, either by comparison to RTE

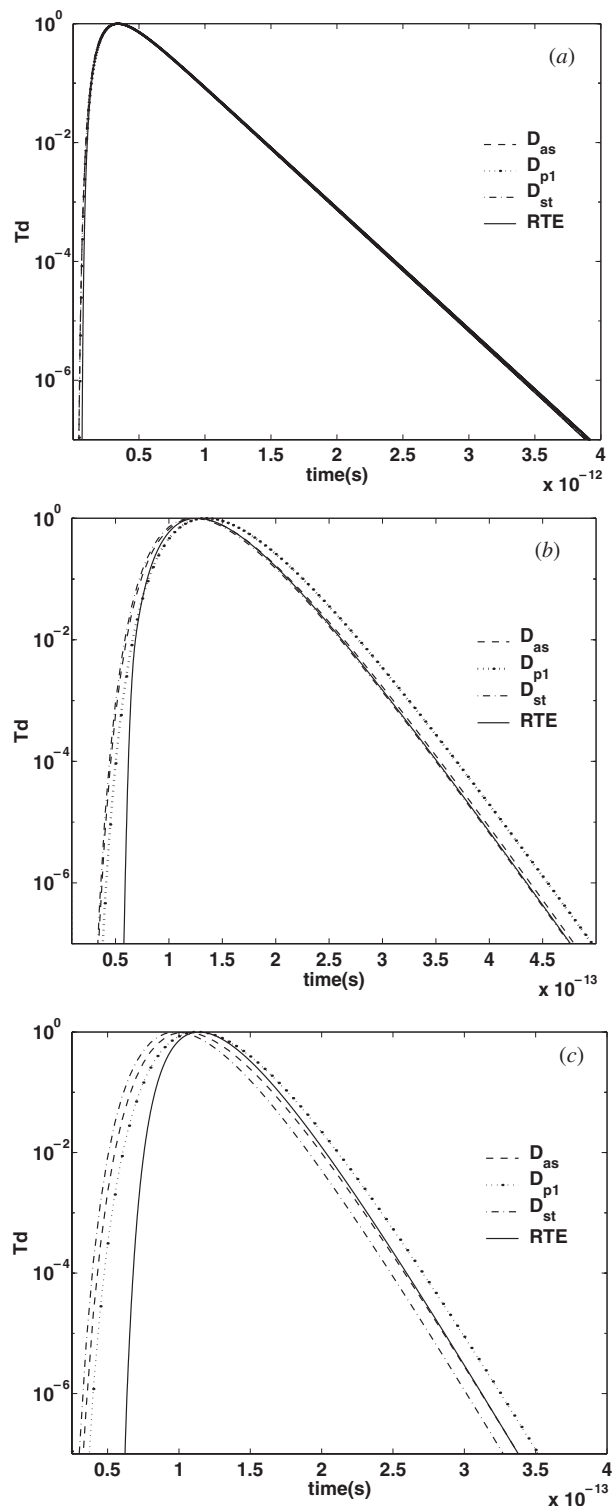
calculations [25] or to experiments [26]. Nevertheless, it seems that the interplay between space scales and timescales on the validity of diffusion results is still an open issue. On the one hand, experiments reported in [27] seem to show that diffusion theory gives an accurate prediction for the long-time decay of transmitted pulses for both thin (optical depth  $< 10$ ) and thick systems, whereas it fails for the short-time behaviour. On the other hand, experiments reported in [28] show that diffusion theory fails for thin systems (optical depth  $< 8$ ), even for the long-time behaviour. In this paper we address this issue by means of numerical examples. We study the short- and long-time behaviour of transmitted pulses, for both thin and thick slabs. In particular, we address the issue of the definition of the diffusion coefficient and its relevance for time-dependent transport.

#### 4.1. The influence of the diffusion coefficient

We show in figure 2 the normalized diffuse impulse response for a slab of thickness  $L = 20 \mu\text{m}$ , and for a transport mean-free path  $l_{tr} = 0.95 \mu\text{m}$ . For such thick slabs, the diffusion approximation is expected to give correct results, at least for the long-time behaviour of the transmitted flux. As briefly discussed in the previous section, several expressions for the diffusion coefficient have been proposed in the literature. Its correct dependence on absorption and on the scattering properties of the medium has been the subject of a controversy in the domain of optical imaging through turbid media [17], and the problem does not seem to be completely solved. In figure 2, we compare the results obtained with three commonly used expressions of the diffusion coefficient:  $D_{P1} = [3(\mu_s(1-g) + \mu_a)]^{-1}$  is the result of the standard  $P_1$  approximation to the RTE [16],  $D_{st} = vl_{tr}/3$  is the result obtained from a statistical approach (and usually in a non-absorbing medium) [23] and  $D_{as}$  is the result obtained from an asymptotic analysis of the RTE, based on an eigenmode expansion [12, 17]. In this last approach, each eigenmode has a spatial decay  $\exp(-\nu z)$ , where  $\nu$  is the associated eigenvalue. The diffusion coefficient is defined by  $D_{as} = \mu_a \nu / \nu_0^2$ , where  $\nu_0$  is the smallest eigenvalue.

We see in figure 2(a) that for a weakly absorbing medium ( $a = 0.995$ ) and for anisotropic scattering ( $g = 0.4$ ), the diffusion approximation correctly describes the transmitted pulse, no matter which diffusion coefficient is used. The same result was observed for isotropic scattering. When absorption increases (see figures 2(b) and (c), for  $a = 0.85$ ) some discrepancies appear between the different diffusion results and the RTE solution. At short time, none of the diffusion results are able to predict the correct arrival of the transmitted pulse. This was expected because the diffusion theory itself fails at short time. Note that diffusion theory always underestimates the arrival time of the first part of the pulse. This is a consequence of the instantaneous (and non-causal) response predicted by diffusion theory, which does not describe the propagation of energy at finite velocity.

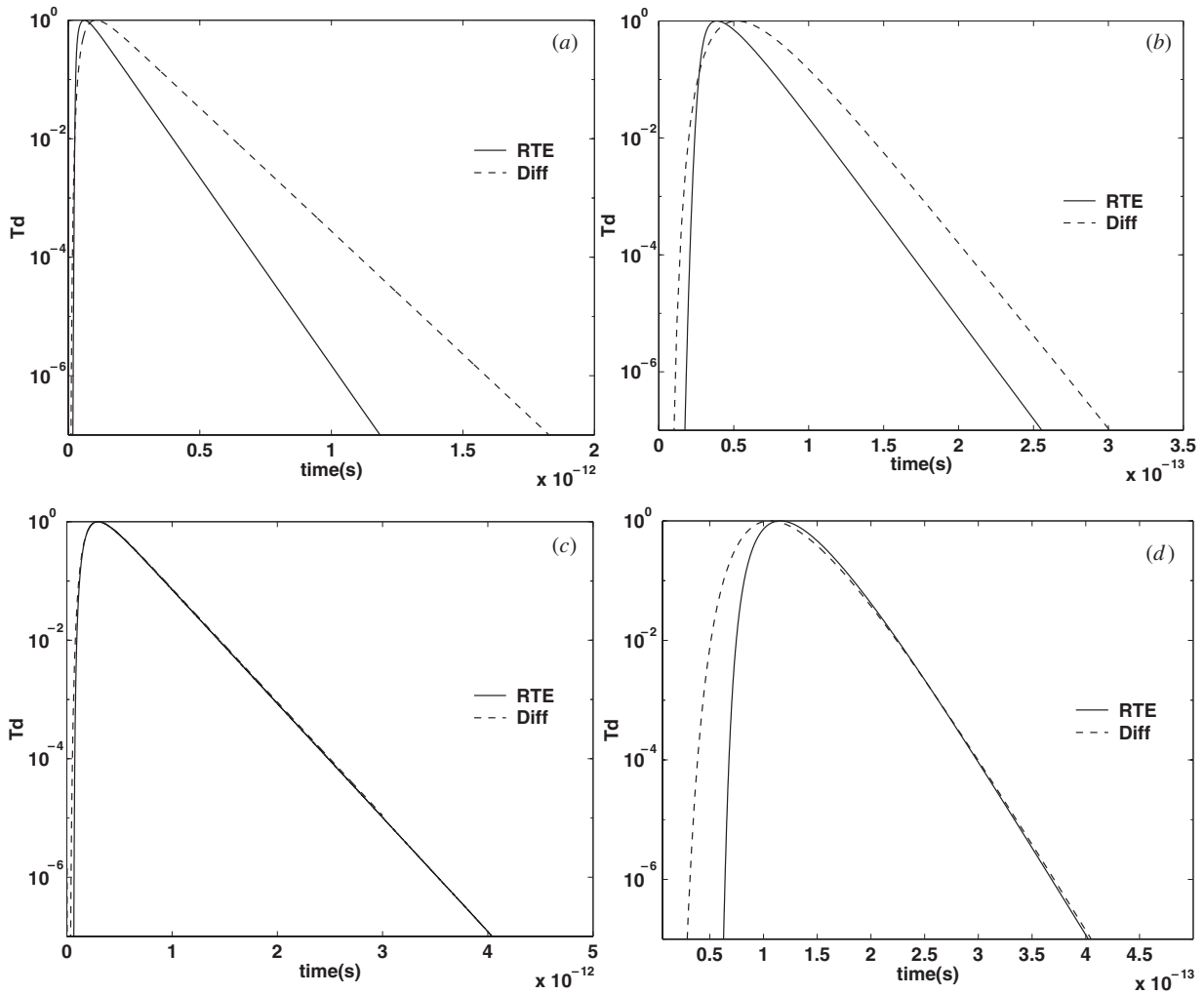
A more surprising result is found for the long-time behaviour of the pulse. Diffusion theory gives a correct result only when  $D_{as}$  is used. This is particularly true in the presence of both absorption and anisotropic scattering (figure 2(c)). We thus conclude that even in a domain where diffusion theory



**Figure 2.** Diffuse transmission normalized by its peak value. Comparison between RTE and diffusion theory. Different expressions of the diffusion coefficient are used:  $D_{P1} = 1/[3(\mu_s(1-g) + \mu_a)]$ ,  $D_{st} = vl_{tr}/3$ ,  $D_{as} = \mu_a \nu / \nu_0^2$  (asymptotic theory).  $L = 20 \mu\text{m}$ ,  $l_{tr} = 0.95 \mu\text{m}$ . (a)  $a = 0.995$ ,  $g = 0.4$ . (b)  $a = 0.85$ ,  $g = 0$ . (c)  $a = 0.85$ ,  $g = 0.4$ . Phase function: Henyey Greenstein. Effective index  $n_2 \simeq 1$ .

is expected to give accurate results (long-time transmission through a thick slab) the proper diffusion coefficient has to





**Figure 3.** Diffuse transmission normalized by its peak value. Comparison between RTE and diffusion theory, with the diffusion coefficient from asymptotic theory. (a)  $L = 6 \mu\text{m}$ ,  $a = 0.995$ ,  $g = 0.4$ . (b) The same as (a) with  $a = 0.85$ . (c)  $L = 15 \mu\text{m}$ ,  $a = 0.995$ ,  $g = 0.4$ . (d) The same as (c) with  $a = 0.85$ . The transport mean-free path is fixed  $l_{tr} = 0.95 \mu\text{m}$ . Phase function: Henyey Greenstein. Effective index  $n_2 \simeq 1$ .

be used. In view of our numerical results, it seems that  $D_{as}$  provides the more accurate results.

These results clearly put forward the dependence of the diffusion coefficient on absorption and show that this dependence is more complicated than that predicted by the  $P_1$  approximation. The asymptotic definition of the diffusion coefficient correctly describes the dependence on absorption, this dependence being implicit and hidden in the determination of the eigenvalue  $\nu_0$ . In the following, we shall use  $D_{as}$  in all calculations based on diffusion theory.

#### 4.2. Short-time and long-time behaviour

In this section we compare diffusion theory and RTE calculations, for both thin and thick slabs, and for both weakly and strongly absorbing media. We plot, in figure 3, the normalized diffuse impulse response, obtained from the RTE and from diffusion theory (equation (19)) with the diffusion coefficient  $D_{as}$ . Figures 3(a) and (b) correspond to a thin slab ( $L = 6 \mu\text{m}$ ,  $l_{tr} = 0.95 \mu\text{m}$ ) and figures 3(c) and (d) to a thick slab ( $L = 20 \mu\text{m}$ ,  $l_{tr} = 0.95 \mu\text{m}$ ). Cases 3(a)

and (c) correspond to a weakly absorbing medium ( $a = 0.995$ ), and cases 3(b) and (d) to strong absorption ( $a = 0.85$ ). The calculations are performed for anisotropic scattering ( $g = 0.4$ ) but all the conclusions remain valid for isotropic scattering.

For thin slabs (figures 3(a) and (b)), we clearly see that diffusion theory fails to describe both the short-time and the long-time parts of the transmitted pulse. This is to be expected from the derivation of the diffusion approximation itself (which clearly assumes large scales for both space and time). This result is in accordance with the experimental results of [28] and contradicts the assertion made in [27], which seemed to show that diffusion theory was able to predict the long-time behaviour of the transmitted pulses for thin slabs. When absorption increases (figure 3(b)) the difference between the RTE result and the diffusion approximation increases at short time. It is also clear that diffusion theory underestimates the arrival time of the early part of the pulse, and overestimates the arrival time of the maximum as well as the long-time decay.

For thick slabs (figures 3(c) and (d)) the situation is substantially different. Although the same result as for thin

slabs is observed at short time, the long-time exponential decay is well predicted by diffusion theory, even in the case of large absorption (figure 3(d)) and anisotropic scattering. This result is also in accordance with the experiments reported in [28]. In both cases, diffusion theory is found to be valid for the prediction of the long-time diffuse transmission for slabs with optical thicknesses  $L/l_{tr} > 8$ .

## 5. Conclusion

We have presented a simple and efficient procedure to solve the time-dependent RTE in a slab geometry, based on a time-domain Fourier transform and a standard discrete-ordinate method. This procedure is well suited to study the transmission and reflection of short (picosecond) light pulses through scattering and absorbing media. The method could be extended to dispersive media, in order to deal with ultrashort (wide-band) pulses.

Using the numerical solutions of the RTE as a reference, we have studied the transition towards the diffusion regime, for different slab thicknesses, and for different levels of absorption and anisotropy in the scattering properties. We have shown that the various expressions of the diffusion coefficient lead to substantial differences in the results obtained by diffusion theory in the presence of absorption. The diffusion coefficient defined from the eigenmode asymptotic approach of neutron diffusion theory seems to give the most accurate results.

We have also shown that diffusion theory always fails for short slabs (optical depth  $< 8$ ), even for the prediction of the long-time behaviour of transmitted pulses. For thick slabs, the long-time behaviour is correctly described by diffusion theory, even in the case of strong absorption and anisotropy, when the diffusion coefficient defined from the eigenmode asymptotic approach is used.

## Appendix

The expressions of the amplitude Fresnel reflection and transmission factors at a flat boundary between two media of refractive indices  $n_i$  and  $n_j$  are

$$r_{ij}^{\perp}(\cos \theta_i) = \frac{n_i \cos \theta_i - n_j \cos \theta_j}{n_i \cos \theta_i + n_j \cos \theta_j} \quad (\text{A.1})$$

$$r_{ij}^{\parallel}(\cos \theta_i) = \frac{n_j \cos \theta_i - n_i \cos \theta_j}{n_j \cos \theta_i + n_i \cos \theta_j} \quad (\text{A.2})$$

$$t_{ij}^{\perp}(\cos \theta_i) = \frac{2n_i \cos \theta_i}{n_i \cos \theta_i + n_j \cos \theta_j} \quad (\text{A.3})$$

$$t_{ij}^{\parallel}(\cos \theta_i) = \frac{2n_i \cos \theta_i}{n_j \cos \theta_i + n_i \cos \theta_j} \quad (\text{A.4})$$

where  $\perp$  and  $\parallel$  stand for the TE and TM polarizations, respectively.  $\theta_i$  is the incidence angle. Note that Snell's law gives  $\cos \theta_j = [1 - (1 - \cos^2 \theta_i)n_i^2/n_j^2]^{1/2}$ , so that the each factor is a function of  $\cos \theta_i$  only. For unpolarized light, the reflection and transmission factors in energy are

$$R_{ij}(\cos \theta_i) = \frac{|r_{ij}^{\perp}(\cos \theta_i)|^2 + |r_{ij}^{\parallel}(\cos \theta_i)|^2}{2} \quad (\text{A.5})$$

$$T_{ij}(\cos \theta_i) = \frac{n_j \cos \theta_j}{n_i \cos \theta_i} \frac{|t_{ij}^{\perp}(\cos \theta_i)|^2 + |t_{ij}^{\parallel}(\cos \theta_i)|^2}{2}. \quad (\text{A.6})$$

## References

- [1] Sebbah P (ed) 2001 *Waves and Imaging through Complex Media* (Dordrecht: Kluwer)
- [2] Yodh A and Chance B 1995 *Phys. Today* **48** 34
- [3] Gayen S K and Alfano R R 1996 *Opt. Photonics News* **7** 17
- [4] Patterson M S, Chance B and Wilson B C 1989 *Appl. Opt.* **28** 2331
- [5] Das B B, Yoo K M and Alfano R R 1993 *Opt. Lett.* **18** 1092
- [6] Hee M R et al 1993 *Opt. Lett.* **18** 950
- [7] Liu F, Yoo K M and Alfano R R 1994 *Opt. Lett.* **19** 740
- [8] Mandelis A *Phys. Today* **53** 29
- [9] Majumdar A 1993 *J. Heat Transfer* **115** 7
- [10] Chen G 1999 *J. Heat Transfer* **121** 945
- [11] Chandrasekhar S 1960 *Radiative Transfer* (New York: Dover)
- [12] Case K M and Zweifel P F 1967 *Linear Transport Theory* (Reading, MA: Addison-Wesley)
- [13] Liu P and Kruger R A 1994 *Opt. Eng.* **33** 2783
- [14] Mitra K and Kumar S 1999 *Appl. Opt.* **38** 188
- [15] Guo Z and Kumar S 2001 *Appl. Opt.* **40** 3156
- [16] Ishimaru A 1997 *Wave Propagation and Scattering in Random Media* (Piscataway, NJ: IEEE)
- [17] Aronson R and Corngold N 1999 *J. Opt. Soc. Am. A* **16** 1066
- [18] Apresyan L A and Kravtsov Yu A 1996 *Radiation Transfer: Statistical and Wave Aspects* (Amsterdam: Gordon and Breach)
- [19] Roux L et al 2001 *J. Opt. Soc. Am. A* **18** 374
- [20] Whitman G M et al 1996 *IEEE Trans. Antennas Propag.* **44** 118
- [21] Ishimaru A et al 2001 *Appl. Opt.* **40** 5495
- [22] Lax M et al 1987 *Proc. Symp. on Laser Optics of Condensed Matter* ed J L Birman, H Z Cummins and A A Kaplyanskii (New York: Plenum) p 229
- [23] Lagendijk A and van Tiggelen B A 1996 *Phys. Rep.* **270** 143
- [24] Thomas G E and Stamnes K 1999 *Radiative Transfer in the Atmosphere and Ocean* (Cambridge: Cambridge University Press) ch 8
- [25] Kim A D and Ishimaru A 1998 *Appl. Opt.* **37** 5313
- [26] Bruce N C et al 1995 *Appl. Opt.* **34** 5823
- [27] Yoo K M, Liu F and Alfano R R 1990 *Phys. Rev. Lett.* **64** 2647
- [28] Kop R H J et al 1997 *Phys. Rev. Lett.* **79** 4369

# Definition of the diffusion coefficient in scattering and absorbing media

Rachid Elaloufi, Rémi Carminati and Jean-Jacques Greffet

Laboratoire d'Energétique Moléculaire et Macroscopique,  
Combustion; Ecole Centrale Paris, Centre National de la  
Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France.

*remi@em2c.ecp.fr*

We revisit the definition of the diffusion coefficient for light transport in scattering and absorbing media. From an asymptotic analysis of the transport equation, we present a novel derivation of the diffusion coefficient, which is restricted neither to low absorption nor to a situation in which the specific intensity is quasi-isotropic. Our expression agrees with previous in the appropriate limit. Using numerical simulations, we discuss the implications of the proper choice of the diffusion coefficient for time-dependent transport. © 2002 Optical Society of America

**OCIS codes:** 290.1990, 290.7050, 290.4210, 170.5280, 170.3660

## 1. Introduction

Light transport through scattering media has attracted considerable interest for imaging applications, especially through biological tissues.<sup>1-4</sup> The standard transport theory for such a problem is based on a Boltzman-like transport equation for the specific intensity, usually referred to as the Radiative Transfer Equation (RTE), whose formalism was first developed for astrophysics<sup>5</sup> and neutron transport.<sup>6</sup> Because solving the RTE in complex geometries remains a challenging issue, the diffusion approximation<sup>7</sup> offers a great simplification of the problem and is widely used.<sup>1-3</sup> In the presence of both anisotropic scattering and absorption, which is the case for biological tissues, the definition of the diffusion coefficient is a non-trivial problem. Indeed, depending on the way the diffusion equation is derived from transport theory, one obtains different expressions for this coefficient. In particular, its dependence on absorption, has been the subject of a controversy.<sup>8</sup> After some claims that the diffusion coefficient should be independent on absorption,<sup>9-11</sup> recent theoretical works have given convincing arguments proving that  $D$  definitely depends on absorption.<sup>8,12,13</sup> Based on an asymptotic analysis of the time-independent RTE<sup>8,13</sup> or on the telegrapher's equation,<sup>12</sup> one arrives at expressions of the form<sup>8,12</sup>  $1/[3(\alpha\mu_a + \mu'_s)]$ , or of the form<sup>13</sup>  $1/[3(\mu_a + \beta\mu'_s)]$ , where  $\mu_a$  is the absorption coefficient and  $\mu'_s = (1 - g)\mu_s$  is the reduced scattering coefficient, with  $\mu_s$  the scattering coefficient and  $g$  the anisotropy factor (average cosine of the scattering angle). Although numerical values for  $\alpha$  and  $\beta$  can be given in some particular cases,<sup>8,12,13</sup> such expressions of  $D$  are only partially satisfactory because the correction factors  $\alpha$  and  $\beta$  depend strongly on both



$\mu_a$  and the phase function.

In this paper, we present a novel derivation of the correct expression of the diffusion coefficient in a medium with absorption and anisotropic scattering. Although we make use of the asymptotic analysis of the RTE developed in the context of neutron transport theory,<sup>6</sup> our approach is different from that in refs.<sup>8,13</sup> In these approaches, the diffusion coefficient is defined by comparison of the spatial decay of the intensity at large scales given by the RTE with that given by diffusion theory. In our approach, starting from the RTE and the asymptotic analysis, we directly derive the radiative flux in the form of a Fick law from which the diffusion coefficient is identified. The details of the derivation are given in section 2, and the resulting expression is discussed qualitatively in section 3. In section 4, we illustrate the dependence of the diffusion coefficient on absorption ( $\mu_a$ ) and scattering anisotropy ( $g$ ). In section 5, we show by numerical calculations that the correct definition of the diffusion coefficient is an important issue for time-dependent transport through scattering and absorbing media. Finally, we summarize our conclusions in section 6.

## 2. Derivation of the expression of the diffusion coefficient

For simplicity, we consider a slab geometry, with the  $z$ -axis taken normal to the slab boundaries. We assume the slab to be illuminated by a monochromatic plane wave propagating towards  $z > 0$ , at normal incidence. The specific intensity  $I(z, \mu)$  inside the medium obeys the stationary RTE:<sup>5</sup>

$$\mu \frac{\partial I(z, \mu)}{\partial z} = -(\mu_a + \mu_s) I(z, \mu) + \frac{\mu_s}{2} \int_{-1}^{+1} p^{(0)}(\mu, \mu') I(z, \mu') d\mu' \quad (1)$$

where  $\mu = \cos \theta$  with  $\theta$  the angle between the propagation direction and the  $z$ -axis,  $p^{(0)}$  is the phase function integrated over the azimuthal angle  $p^{(0)}(\mu, \mu') = \frac{1}{2\pi} \int_0^{2\pi} p(\mathbf{u} \cdot \mathbf{u}') d\phi$ ,  $\mathbf{u}$  and  $\mathbf{u}'$  being unit vectors corresponding to directions  $(\theta, \phi)$  and  $(\theta', \phi')$ , respectively. The RTE results from an energy balance in a given direction<sup>5</sup> (specified only by the directional cosine  $\mu$  in our slab geometry). Energy is taken from this direction due to absorption and scattering (first term in the right-hand side in Eq. (1)). Energy is also added to this direction due to scattering: Energy propagating in a direction  $\mu'$  is scattered into the direction  $\mu$  (second term in the right-hand side in Eq. (1)), the probability for this mechanism to occur being given by the phase function.

In order to find the diffusion coefficient, we compute the radiative flux across a plane normal to the  $z$ -axis. By definition, it is given by<sup>5</sup>  $q_z = 2\pi \int_{-1}^{+1} \mu I(z, \mu) d\mu$ . Starting from Eq. (1), one can establish another general (and rigorous) expression of the radiative flux (see Appendix A for details):

$$q_z(z) = \frac{-2\pi}{\mu_a + \mu'_s} \frac{\partial}{\partial z} \int_{-1}^{+1} \mu^2 I(z, \mu) d\mu \quad (2)$$

We now introduce an approximation. When the thickness  $L$  of the slab is much larger than the transport mean-free path  $l_{tr} = 1/\mu'_s$ , the specific intensity inside the medium takes the asymptotic form:<sup>6</sup>

$$I(z, \mu) = c_0^+ g_0^+(\mu) \exp(-k_0 z) + c_0^- g_0^-(\mu) \exp(k_0 z) \quad (3)$$

where  $k_0$  is the smallest eigenvalue and  $g_0^\pm(\mu)$  the associated eigenfunctions of the eigenvalue problem obtained when seeking solutions of the RTE of the form  $I(z, \mu) = g^\pm(\mu) \exp(\pm k z)$  (see Ref.<sup>14</sup> for a practical numerical solution of the eigenvalue problem for arbitrary absorption and phase function). For the eigenfunctions, the normalization condition  $\int_{-1}^{+1} g_0^\pm(\mu) d\mu = 1$  is used. The constants  $c_0^\pm$  depend on the boundary conditions at the slab boundaries, and do not influence the subsequent derivation. It is known that the (degenerate) mode associated with the smallest eigenvalue  $k_0$  determines the diffusion behavior.<sup>6</sup> This is the only mode that survives at large depths.

In the following, we shall refer to it as the *fundamental mode*. Inserting (3) into (2), one obtains:

$$q_z(z) = \frac{2\pi}{\mu_a + \mu'_s} k_0 \left[ c_0^+ \langle g_0 \rangle \exp(-k_0 z) - c_0^- \langle g_0 \rangle \exp(k_0 z) \right] \quad (4)$$

where  $\langle g_0 \rangle$  is given by

$$\langle g_0 \rangle = \int_{-1}^{+1} \mu^2 g_0^\pm(\mu) d\mu, \quad (5)$$

the integral taking the same value for  $g_0^+(\mu)$  and  $g_0^-(\mu)$ . The factor  $\langle g_0 \rangle$  is a measure of the anisotropy of the fundamental mode  $g_0^\pm$ . For an isotropic mode, one has  $g_0^\pm(\mu) = 1/2$  for all values of  $\mu$ , which leads to  $\langle g_0 \rangle = 1/3$ . A deviation from this value is a signature of the anisotropy of the fundamental mode, which may be due the anisotropy of the phase function (scattering) and/or to absorption. We shall come back to this point and show some examples in section 3.

In order to transform Eq. (4) into a Fick law from which a diffusion coefficient can be deduced, we introduce the energy density  $U(z)$  defined by  $U(z) = 2\pi/v \int_{-1}^{+1} I(z, \mu) d\mu$  where  $v$  is the energy velocity. Inserting the asymptotic expression of  $I(z, \mu)$ , Eq. (3), into this relation, one obtains the following expression for the energy density:

$$U(z) = \frac{2\pi}{c} \left[ c_0^+ \exp(-k_0 z) + c_0^- \exp(k_0 z) \right] \quad (6)$$

Calculating the first derivative of  $U(z)$  from Eq. (6) and comparing the result with the right-hand side of Eq. (4) leads to the following relation between the energy flux and the gradient of the energy density:

$$q_z(z) = -\frac{\langle g_0 \rangle}{\mu_a + \mu'_s} v \frac{\partial U(z)}{\partial z} \quad (7)$$

Equation (7) takes the form of a Fick law, from which the diffusion coefficient can be identified. At this stage, a remark has to be done concerning the units used to define the diffusion coefficient for light transport. By writing  $q_z = -D' \partial U / \partial z$ , we naturally end up with a diffusion coefficient having the unit  $m^2.s^{-1}$ , and given by

$$D' = \langle g_0 \rangle \frac{v}{\mu_a + \mu'_s}. \quad (8)$$

This diffusion coefficient is consistent with a diffusion equation written in the form  $\partial U / \partial t - D' \nabla^2 U + \mu_a v U = 0$ , as, for example, in ref.<sup>11</sup> Nevertheless, in optics, a diffusion coefficient  $D$  in unit of length ( $m$ ) is often used, as for example in refs.<sup>8,13</sup> and is consistent with a diffusion equation written in the form  $(1/v) \partial U / \partial t - D \nabla^2 U + \mu_a U = 0$ . In this case, the Fick law writes  $q_z = -D v \partial U / \partial z$ , and our approach leads to a diffusion coefficient given by

$$D = \frac{\langle g_0 \rangle}{\mu_a + \mu'_s}. \quad (9)$$

The advantage of this second definition of the diffusion coefficient is to remove any dependence of  $D$  on the energy velocity. This is particularly relevant in the case of a stationary situation, for which the energy velocity disappears from the diffusion equation. In this work, we shall study mainly the dependence of  $D$  on absorption and on the scattering properties of the medium. We will use the the expression  $D'$  only in section 5, when we study a time-dependent situation.

Equation (9) is the main result of this paper. Provided that the specific intensity can be represented using only the fundamental mode as in Eq. (3), the derivation of Eq. (9) follows immediately.

The validity of Eq. (3) is the crucial issue in our derivation, and may be considered carefully. It relies on the fact that the medium is thick enough so that the fundamental mode dominates. This occurs under two conditions: (i) the second mode with eigenvalue  $k_1$  ( $k_1 > k_0$ ) should be such that  $k_1 L \gg 1$ , with  $L$  the size of the medium, and (ii)  $k_1$  should not be too close from  $k_0$ . Whether the second condition is valid is not obvious. We have verified numerically that the difference between the values of  $k_0$  and  $k_1$  is always large enough for the asymptotic expression to be valid. Some numerical values of the first eigenvalues  $k_0$ ,  $k_1$  and  $k_2$  for realistic parameters in biomedical optics are given in Table 1 and illustrate this conclusion. It is also interesting to note that the value of  $k_0$  decreases when increasing the anisotropy factor  $g$ . For  $g > 0$ , forward scattering is stronger than for  $g = 0$ , so that the fundamental mode has a larger decay length  $1/k_0$ .

### 3. Physical analysis of the expression of $D$

Before elaborating on the significance of Eq. (9), we shall show that it is equivalent to that obtained by the usual procedure based on the asymptotic approach.<sup>6,8,13</sup> To do so, we first recall that energy conservation, for a stationary situation, writes:

$$\partial q_z(z)/\partial z = -\mu_a v U(z) \quad (10)$$

This relation can be established by integrating Eq. (1) over  $\mu$  and using the definition of both the radiative flux  $q_z$  and the energy density  $U(z)$  previously given in the text. Inserting Eq. (7) into (10) leads to:

$$\frac{-1}{\mu_a + \mu'_s} \langle g_0 \rangle \frac{\partial^2 U(z)}{\partial z^2} = -\mu_a U(z) \quad (11)$$

Using (6) to replace both the energy flux and its second derivative, one obtains  $\langle g_0 \rangle = \mu_a(\mu_a + \mu'_s)/k_0^2$ . The introduction of this result into Eq. (9) leads to  $D = \mu_a/k_0^2$ , which is the form of  $D$  obtained by Case *et al.*<sup>6,8,13</sup>

Equation (9) shows that the diffusion coefficient is determined by the second moment  $\langle g_0 \rangle$  of the angular distribution of the fundamental mode. The derivation presented here has several advantages. (1) It is a direct derivation in the sense that the diffusion behavior for the radiative flux is directly obtained from the asymptotic expression of the specific intensity. We do not have to use the solution of the diffusion equation to extract  $D$  by comparison of the spatial decay of the intensity as is usually done in the asymptotic approach.<sup>6,8,13</sup> (2) The expression of  $D$  takes a form similar to that obtained in the  $P1$ -approximation,<sup>7</sup> with  $\langle g_0 \rangle$  appearing as a correction factor. The  $P1$ -approximation assumes a quasi-isotropy of the specific intensity inside the medium. For an isotropic fundamental mode, one has  $g_0^\pm(\mu) = 1/2$  and thus  $\langle g_0 \rangle = 1/3$ . One recovers in this case the result of the  $P1$ -approximation. (3) The derivation is not restricted to low absorption media. (4) Last but not least, the derivation shows that the diffusion regime is valid even for a specific intensity which is far from the quasi-isotropy assumed in the  $P1$ -approximation.

To clarify the last point, let us discuss the interplay between losses and isotropy. A simple picture can be obtained using a random walk model. If the absorption length is shorter than the transport mean-free path, the light cannot reach the quasi-isotropic regime because photons are absorbed before undergoing a sufficient number of scattering events. In this case, the fundamental mode is not isotropic. Yet, the flux can still be cast in the form of a Fick law as shown above. To illustrate this qualitative discussion, the angular behavior of  $g_0^+(\mu)$  has been calculated by numerically solving the RTE, using the eigenvalue matrix formulation of ref.<sup>14</sup> It is represented in Fig. 1(a), for an anisotropically scattering medium (Henyey-Greenstein phase function with  $g = 0.5$ ) and for different levels of absorption. We see that for low absorption (dashed curve), the fundamental mode is quasi-isotropic. A weak anisotropy is necessary to provide a non-vanishing net

flux, which for the mode  $g_0^-$  is directed from left to right. For increasing absorption, the fundamental mode becomes more and more anisotropic (solid and dotted curves). The  $P1$ -approximation is no longer valid but we can still introduce a diffusion coefficient using Eq. (9). This is due to the fact that only the fundamental mode of the RTE contributes to the specific intensity.

Finally, let us discuss the limit of a non-absorbing medium, with arbitrary scattering properties. In this case, the result  $D = 1/(3\mu'_s)$  is known to be the exact one,<sup>8,12,13</sup> and the diffusion coefficient defined in Eq. (9) has to give this value. In other words, in the limit of a vanishing absorption, the factor  $\langle g_0 \rangle$  has to be  $1/3$ , whatever the scattering properties of the medium. We have verified numerically that this is the case. We show in Fig. 1(b) the variation of  $\langle g_0 \rangle$  versus absorption for two different media, with either isotropic scattering ( $g = 0$ ) or anisotropic scattering ( $g = 0.5$ ). We see that for a vanishing absorption ( $\omega_0 = 1$ ),  $\langle g_0 \rangle$  takes the value  $1/3$  in both cases. Although only two curves are represented for the sake of clarity, we have verified that this result holds for any value of  $g$  and different phase functions.

#### 4. Dependence of the diffusion coefficient on absorption and on the scattering properties

In order to illustrate the strong dependence of  $D$  on absorption and scattering anisotropy, we plot in Fig. 2 the behavior of  $D$  versus the albedo  $\omega_0 = \mu_s/(\mu_a + \mu_s)$  for a fixed value of  $g$  and the transport mean-free path  $l_{tr} = 1/\mu'_s$  [Figs. 2(a) and 2(b)], and versus  $g$  for a fixed value of the absorption and transport mean-free paths  $l_{abs} = 1/\mu_a$  and  $l_{tr}$  [Figs. 2(c) and 2(d)]. The result obtained with  $D$  given by Eq. (9) is denoted by  $D_{as}$  (solid line). It is obtained by calculating numerically the fundamental mode of the RTE, following the eigenvalue matrix formulation of ref.<sup>14</sup> A Henyey-Greenstein phase function is used in the calculation. For comparison, we show the dependence predicted by other expressions of the diffusion coefficient:  $D_{P1} = 1/[3(\mu_a + \mu'_s)]$  is the result of the (widely-used)  $P1$ -approximation<sup>7</sup> and  $D_{st} = 1/(3\mu'_s)$  is the expression obtained from a (microscopic) statistical approach in an non-absorbing medium.<sup>15</sup> Note that this last expression is the exact value for a non-absorbing medium. It was put forward by some authors as the exact one even in the presence of absorption,<sup>9-11</sup> a result which was shown to be incorrect by recent theoretical considerations.<sup>8,12,13</sup> In Fig. 2(a), we also represent the result given by  $D_{as}^{(0)} = 1/[3(0.2\mu_a + \mu'_s)]$ , a first-order approximation of expression (9) valid for isotropic scattering and low absorption.<sup>8</sup> For isotropic scattering [Fig. 2(a)], the dependence on absorption predicted by the  $P1$ -approximation and the asymptotic approach are clearly different. The  $P1$ -approximation strongly overestimates the dependence on absorption. Note that the approximate expression  $D_{as}^{(0)}$  is valid for an albedo  $\omega_0 > 0.8$ , and gives accurate results in this case. When the scattering anisotropy increases [Fig. 2(b)], the dependence on absorption becomes more important. For exemple, between  $\omega_0 = 1$  and  $\omega_0 = 0.5$ , the value of  $D_{as}$  decreases by a factor of two. The  $P1$ -approximation predicts a division by three on the same interval. These results prove the strong dependence of  $D$  on absorption, and the failure of  $D_{P1}$  and  $D_{st}$  to predict the correct dependence. The  $P1$ -approximation always overestimates the dependence on absorption, while (obviously), the coefficient  $D_{st}$  underestimates it. This behavior was reported previously for isotropic scattering using a different approach.<sup>13</sup> It is also consistent with the result in ref.<sup>8</sup> which shows that an approximate expression of the form  $1/[3(\alpha\mu_a + \mu'_s)]$ , with  $0.2 < \alpha < 0.8$ , could be used. This expression gives a dependence on absorption which is smaller than that given by the  $P1$ -approximation (which would correspond to  $\alpha = 1$ ) and larger than that predicted by  $D_{st}$  (which would correspond to  $\alpha = 0$ ).

The influence of the anisotropy factor  $g$ , for a fixed value of the absorption and transport mean-free paths  $l_{abs} = 1/\mu_a$  and  $l_{tr} = 1/\mu'_s$ , is shown in Fig. 2(c). Both  $D_{P1}$  and  $D_{st}$ , which only depend on  $\mu_a$  and  $\mu'_s$ , are constant in this case. The expression  $D_{as}$  predicts a slight dependence

on  $g$ , even for a fixed value of  $l_{tr} = 1/\mu'_s$ . This dependence is given by the factor  $\langle g_0 \rangle$  in Eq. (9). This is an interesting property of the new approach, although in practice this dependence might be negligible as seen by the numerical values in Fig. 2(c).

In Fig. 2(d), we show the influence of the anisotropy factor  $g$ , for a fixed value of the absorption and scattering lengths  $l_{abs} = 1/\mu_a$  and  $l_{sca} = 1/\mu_s$ . In this case, changing  $g$  amounts to change both  $l_{tr} = 1/\mu'_s$  and the factor  $\langle g_0 \rangle$ . With a high level of absorption ( $\omega_0 = 0.85$ ), the three expressions of  $D$  give different results. Compared to the dependence given by  $D_{as}$ , the  $P1$ -approximation overestimates the dependence on  $g$ , leading to an error on the order of 40% for  $g = 0.8$ . The diffusion coefficient  $D_{st}$ , independent on absorption, underestimates the dependence on  $g$  in an absorbing medium, the error being on the order of 25% for  $g = 0.8$ . These results demonstrate that the expressions  $D_{P1}$  and  $D_{st}$  are unable to predict the dependence of the diffusion coefficient on the scattering properties of the medium in the presence of absorption.

## 5. Implications for time-dependent transport

We shall now demonstrate the relevance of the definition of the diffusion coefficient for time-dependent transport. This is an important issue for time-resolved techniques in biomedical imaging using diffusing light, which seem to give promising results.<sup>1-4,16-19</sup> To do so, we compare calculations of the diffuse transmission through scattering slabs obtained by solving the RTE (reference solution) and by solving the diffusion equation. The time dependent RTE for a slab geometry writes:

$$\frac{1}{v} \frac{\partial I(z, \mu, t)}{\partial t} + \mu \frac{\partial I(z, \mu, t)}{\partial z} = -(\mu_a + \mu_s) I(z, \mu, t) + \frac{\mu_s}{2} \int_{-1}^{+1} p^{(0)}(\mu, \mu') I(z, \mu', t) d\mu' \quad (12)$$

Assuming an illumination of the slab from the left by a plane-wave pulse at normal incidence, we have solved this equation using a space-frequency domain discrete-ordinate method, which is described in ref.<sup>20</sup> This provides the reference solution. In order to compute a quantity which can be directly compared with diffusion theory, we compute the impulse response for diffuse transmission, by using an incident pulse of the form  $I_{inc}(z, t) = \delta(z)\delta(t)$  and calculating the total diffuse transmission across the slab. In the diffusion approximation, this impulse response can be evaluated analytically using the method of images and extrapolated boundary conditions.<sup>21</sup> One obtains:

$$T(t) = \frac{H(t)D'}{d} \exp(-\mu_a vt) \sum_{m=1}^{\infty} \frac{\pi m}{d} \sin\left(\frac{\pi m L}{d}\right) \exp\left(-\frac{\pi^2 m^2 D' t}{d^2}\right) \quad (13)$$

where  $H(t)$  is the Heaviside step function,  $d = L + 2z_0$ ,  $L$  being the width of the slab and  $z_0$  the extrapolation distance. This distance may be determined from the Milne problem with isotropic scattering and no absorption,<sup>6</sup> which leads to the value  $z_0 = 0.71 l_{tr}$ . In the presence of absorption, this distance may change, especially for slabs with internal reflection at the boundaries.<sup>22</sup> Note that the diffusion coefficient entering Eq. (13) is  $D' = Dv$ , and that the energy velocity also enters the exponential factor  $\exp(-\mu_a vt)$ . In the present work, we do not address the issue of the expression of  $v$ , which is by itself a whole subject of investigation.<sup>15,23-26</sup> In the subsequent calculations, we have assumed a constant velocity  $v$ , with a value given by the speed of light in vacuum (we have assumed that the effective index of refraction of the system was close to one).

The results of the comparison are shown in Fig. 3, for a medium with absorption and transport mean-free paths  $l_{abs} = 5.4\mu m$  and  $l_{tr} = 0.95\mu m$  and for both isotropic [ $g = 0$ , Fig. 3(a)] and anisotropic [ $g = 0.5$ , Fig. 3(b)] scattering. The optical width of the slab (on the order of 15) is chosen so that the diffusion approximation should be valid at long times.<sup>20,27</sup> The results of the diffusion approximation are calculated using Eq. (13), with three different expressions of the diffusion coefficient ( $D_{as}$ ,  $D_{P1}$  and  $D_{st}$ , with the same notations as in Fig. 2). The medium is

assumed to have an optical index  $n = 1$ , so that there are no internal reflections at the boundaries. We have used the value  $z_0 = 0.71 l_{tr}$  for the extrapolation distance, which is approximate in the presence of absorption. We have verified that modifying this value does not change significantly the long-time behavior of the transmitted pulse predicted by diffusion theory.

We see in Fig. 3(a) and 3(b) that the diffusion approximation does not accurately predict the short-time behavior of the transmitted pulse. This is an expected result, because the diffusion equation is not expected to be valid at short time.<sup>20,29,30</sup> In the present work, we shall concentrate on the prediction of the long-time behavior. In the case of isotropic scattering [Fig. 3(a)], the long-time (exponential) decay of the pulse is correctly predicted by the diffusion approximation, except when the expression of  $D_{P1}$  is used. This shows that in this case, the coefficient  $D_{st}$  gives better results (compared to the rigorously defined coefficient  $D_{as}$ ) than the widely-used coefficient  $D_{P1}$ , even in the presence of absorption. For anisotropic scattering [Fig. 3(b)], the long-time decay is correctly described by diffusion theory only when  $D_{as}$  is used. The result using  $D_{st}$  underestimates the slope, while the result with  $D_{P1}$  overestimates it. This result demonstrates the relevance of the definition of  $D$  for time-dependent transport through absorbing media with anisotropic scattering. None of the standard coefficients  $D_{P1}$  and  $D_{st}$  give correct results for the long-time transmission through thick slabs, for which the diffusion approximation is supposed to be valid.<sup>20,27</sup>

## 6. Conclusion

In summary, we have presented a novel derivation of the rigorous expression of the diffusion coefficient  $D$  for light transport in a scattering and absorbing medium. The result shows that the dependence of  $D$  on absorption and on phase function is described by the second moment of the angular behavior of the fundamental mode. This provides a new physical interpretation of the diffusion coefficient and confirms on a rigorous basis its strong dependence on absorption. It also yields a simple way to derive its correct form. Using a numerical solution of the unsteady RTE in a slab geometry as a reference, we have shown that the use of the correct expression of  $D$  in solutions based on the diffusion approximation is a crucial issue for time-dependent transport.

*Corresponding author:* Dr Rémi Carminati, remi@em2c.ecp.fr

## Acknowledgements

We acknowledge an anonymous referee for valuable comments and suggestions.

## Appendix A

In this appendix, we give the explicit derivation of the general expression of the radiative flux Eq. (2). Multiplying Eq. (1) by  $\mu$  and integrating over  $\mu$ , one obtains:

$$\int_{-1}^{+1} \mu^2 \frac{\partial I(z, \mu)}{\partial z} d\mu = -(\mu_a + \mu_s) \int_{-1}^{+1} \mu I(z, \mu) d\mu + \frac{\mu_s}{2} \int \int_{-1}^{+1} \mu p^{(0)}(\mu, \mu') I(z, \mu') d\mu' d\mu \quad (\text{A1})$$

We now make use of the relationship  $\int_{-1}^{+1} \mu p^{(0)}(\mu, \mu') d\mu = 2g\mu'$  which can be found, for exemple, in ref.,<sup>5</sup> and whose derivation is reproduced in Appendix B. Equation (A1) can be transformed into:

$$\frac{\partial}{\partial z} \int_{-1}^{+1} \mu^2 I(z, \mu) d\mu = -(\mu_a + \mu_s) \int_{-1}^{+1} \mu I(z, \mu) d\mu + \mu_s g \int_{-1}^{+1} \mu' I(z, \mu') d\mu' \quad (\text{A2})$$

In this last equation, the integrals in the right-hand can be replaced by the radiative flux  $q_z$  across the slab, thanks to the relationship  $q_z = 2\pi \int_{-1}^{+1} \mu I(z, \mu) d\mu$ . One finally obtains:

$$q_z(z) = \frac{-2\pi}{\mu_a + \mu_s(1-g)} \frac{\partial}{\partial z} \int_{-1}^{+1} \mu^2 I(z, \mu) d\mu \quad (\text{A3})$$

which is Eq. (2).

## Appendix B

In this appendix, we give the derivation of formula  $\int_{-1}^{+1} \mu p^{(0)}(\mu, \mu') d\mu = 2g\mu'$  which was used in Appendix A. The phase function  $p(\mathbf{u} \cdot \mathbf{u}')$  can be expanded on a basis of Legendre polynomials:<sup>5</sup>

$$p(\mathbf{u} \cdot \mathbf{u}') = \sum_{n=0}^{+\infty} a_n P_n(\mathbf{u} \cdot \mathbf{u}') \quad (\text{B1})$$

with  $a_0 = 1$  and  $a_1 = 3g$ ,  $g = (4\pi)^{-1} \int p(\mathbf{u} \cdot \mathbf{u}') d\Omega$  being the average cosine of the scattering angle. The addition theorem of spherical harmonics<sup>5</sup> allows to express the Legendre polynomial versus the directional cosines  $(\mu, \mu')$  and the azimuthal angles  $(\phi, \phi')$ :

$$P_n(\mathbf{u} \cdot \mathbf{u}') = P_n(\mu)P_n(\mu') + 2 \sum_{m=1}^n \frac{(n-m)!}{(n+m)!} P_n^m(\mu)P_n^m(\mu') \cos(\phi - \phi') \quad (\text{B2})$$

where  $P_n^m$  are associated Legendre functions. Using Eqs. (B1) and (B2), the phase function  $p^{(0)}$  integrated over  $\phi$  and  $\phi'$  can be written:

$$p^{(0)}(\mu, \mu') = \sum_{n=0}^{+\infty} a_n P_n(\mu)P_n(\mu') \quad (\text{B3})$$

From Eq. (B3), we can calculate the following integral:

$$\begin{aligned} \int_{-1}^{+1} \mu p^{(0)}(\mu, \mu') d\mu &= \sum_{n=0}^{+\infty} a_n P_n(\mu') \int_{-1}^{+1} \mu P_n(\mu) d\mu \\ &= \sum_{n=0}^{+\infty} a_n P_n(\mu') \int_{-1}^{+1} \mu P_n(\mu) P_0(\mu) d\mu \end{aligned} \quad (\text{B4})$$

where we have introduced  $P_0(\mu) = 1$  in the last line. The last integral vanishes, except when  $n = 1$ , where it takes the value  $2/3$  (see for example ref.<sup>28</sup>). We finally end-up with :

$$\begin{aligned} \int_{-1}^{+1} \mu p^{(0)}(\mu, \mu') d\mu &= \frac{2}{3} a_1 P_1(\mu') \\ &= 2g\mu' \end{aligned} \quad (\text{B5})$$

where we have used  $P_1(\mu') = \mu'$ . This is the expected relationship.

## References

1. A. Yodh and B. Chance, "Spectroscopy and imaging with diffusing light", *Phys. Today* **48**, 34-40 (1995).
2. S.K. Gayen and R.R. Alfano, "Biomedical imaging techniques", *Opt. Photonics News* **7**, 17-22 (1996).
3. A. Mandelis, "Diffusion waves and their uses", *Phys. Today* **53**, 29-34 (2000).
4. P. Sebbah, ed., *Waves and Imaging through Complex Media* (Kluwer Academic, Dordrecht, 2001).
5. S. Chandrasekhar, *Radiative Transfer* (Dover, New York, 1960).
6. K.M. Case and P.F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Massachusetts, 1967).
7. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (IEEE Press, Piscataway, 1997).
8. R. Aronson and N. Corngold, "Photon diffusion coefficient in an absorbing medium", *J. Opt. Soc. Am. A* **16**, 1066-1071 (1999).
9. K. Furutsu and Y. Yamada, "Diffusion approximation for a dissipative random medium and the applications", *Phys. Rev. E* **50**, 3634-3640 (1994).
10. M. Bassani, F. Martelli, G. Zaccanti and D. Contini, "Independence of the diffusion coefficient from absorption: experimental and numerical evidence", *Opt. Lett.* **22**, 853-855 (1997).
11. T. Durduran, A.G. Yodh, B. Chance and D.A. Boas, "Does the photon-diffusion coefficient depend on absorption?", *J. Opt. Soc. Am. A* **14**, 3358-3365 (1997).
12. D.J. Durian, "The diffusion coefficient depends on absorption", *Opt. Lett.* **23**, 1502-1504 (1998).
13. R. Graaff and J.J. Ten Bosch, "Diffusion coefficient in photon diffusion theory", *Opt. Lett.* **25**, 43-45 (2000).
14. G.E. Thomas and K. Stamnes, *Radiative Transfer in the Atmosphere and Ocean* (Cambridge University Press, Cambridge, 1999), Chap. 8.
15. A. Lagendijk and B.A. van Tiggelen, "Resonant multiple scattering of light", *Phys. Rep.* **270**, 143 (1996).



16. F. Liu, K.M. Yoo and R.R. Alfano, "Transmitted photon intensity through biological tissues within various time windows", *Opt. Lett.* **19**, 740-742 (1994).
17. D.J. Pine, D.A. Weitz, P.M. Chaikin and E. Herbolzheimer, "Diffusive-wave spectroscopy", *Phys. Rev. Lett.* **60**, 1134-1137 (1988).
18. D.A. Boas, L.E. Campbell and A.G. Yodh, "Scattering and imaging with diffusing temporal field correlations", *Phys. Rev. Lett.* **75**, 1855-1858 (1995).
19. K.K. Bizheva, A.M. Siegel and D.A. Boas, "Path-length resolved dynamic light scattering in highly scattering random media: the transition to diffusing wave spectroscopy", *Phys. Rev. E* **58**, 7664-7667 (1998).
20. R. Elaloufi, R. Carminati and J.-J. Greffet, "Time-dependent transport through scattering media: from radiative transfer to diffusion", *J. Opt. A: Pure Appl. Opt.* **4**, S103-S108 (2002).
21. M. Lax, V. Nayaranamurti and R.C. Fulton, "Classical diffusive photon transport in a slab", in *Proc. Symposium on Laser Optics of Condensed Matter*, J.L. Birman, H.Z. Cummins and A.A. Kaplyanskii, eds. (Plenum, New York, 1987), pp. 229-235.
22. R. Aronson, "Boundary conditions for diffusion of light", *J. Opt. Soc. Am. A* **12**, 2532-2539 (1995).
23. M.P. va, Albada, B.A. van Tiggelen, A. Lagendijk and A. Tip, "Speed of propagation of classical waves in strongly scattering media", *Phys. Rev. Lett.* **66**, 3132-3135 (1991).
24. B.A. van Tiggelen, A. Lagendijk M.P. va, Albada and A. Tip, "Speed of light in random media", *Phys. Rev. B* **45**, 12233-12243 (1992).
25. Y. Kuga, A. Ishimaru and D. Rice, "Velocity of coherent and incoherent electromagnetic waves in a dense strongly scattering medium", *Phys. Rev. B* **48**, 13155-13158 (1993).
26. H.P. Schriemer, M.L. Cowan, J.H. Page, P. Sheng, Z. Liu and D.A. Weitz, "Energy velocity of diffusing waves in strongly scattering media", *Phys. Rev. Lett.* **79**, 3166-3169 (1997).
27. R.H.J. Kop, P. de Vries, R. Sprik and A. Lagendijk, "Observation of anomalous transport of strongly multiple scattered light in thin disordered slabs", *Phys. Rev. Lett.* **79**, 4369-4372 (1997).
28. J.D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1975), p. 90.
29. D.J. Durian and J. Rudnick, "Photon migration at short times and distances and in case of strong absorption", *J. Opt. Soc. Am. A* **14**, 235-245 (1997).

30. A.D. Kim and A. Ishimaru, "Optical diffusion of continuous wave, pulsed, and density waves in scattering media and comparisons with radiative transfer", *Appl. Opt.* **37**, 5313-5319 (1998).

Fig. 1. (a): Angular behavior (polar diagram with  $\theta$  in degrees) of the fundamental mode  $g_0^+(\mu = \cos \theta)$ , for different values of the albedo  $\omega_0$ .  $l_{tr} = 0.95\mu m$  and  $g = 0.5$ . (b): variation of  $\langle g_0 \rangle$  versus absorption, for two values of the anisotropy factor  $g$ .  $l_{sca} = 1\mu m$ . Phase function: Henyey-Greenstein.

Fig. 2. Dependence on absorption [(a) and (b)] and scattering anisotropy [(c) and (d)] of the diffusion coefficients obtained from three different approaches:  $P1$ - approximation  $D_{P1}$ , random-walk in a non-absorbing medium  $D_{st}$  and asymptotic approach  $D_{as}$ .  $D_{as}^{(0)}$  in Fig. 2(a) is a first-order approximation of  $D_{as}$  valid for isotropic scattering and low absorption. (a):  $l_{tr} = 0.95\mu m$ ,  $g = 0$ . (b):  $l_{tr} = 0.95\mu m$ ,  $g = 0.5$ . (c):  $l_{tr} = 0.95\mu m$ ,  $l_{abs} = 0.1mm$ . (d):  $l_{sca} = 1\mu m$ ,  $\omega_0 = 0.85$ . Phase function: Henyey-Greenstein.

Fig. 3. Diffuse-transmission impulse response through a slab of width  $L = 15\mu m$ , with  $l_{tr} = 0.95\mu m$  and  $l_{abs} = 5.4\mu m$ . (a):  $g = 0$ . (b):  $g = 0.5$ . Phase function: Henyey-Greenstein. The energy velocity is assumed to be a constant:  $v = 3.10^8 m.s^{-1}$ . The long-time decay is correctly predicted by diffusion theory only when the diffusion coefficient given by Eq. (9) is used.

Table 1. Numerical values of the three first eigenvalues  $k_0$ ,  $k_1$  and  $k_2$  obtained when seeking solutions of Eq. (1) of the form  $I(z, \mu) = g^\pm(\mu) \exp(\pm kz)$ . Four different cases are shown. The scattering length is fixed:  $l_{sca} = 0.95\mu m$ . Phase function: Henyey-Greenstein.

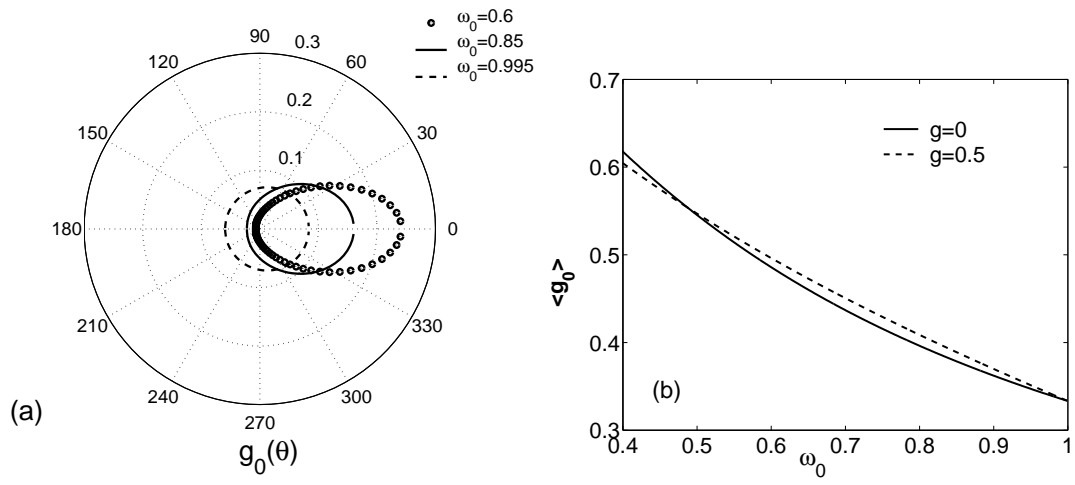


Figure 1, R. Elaloufi et al.

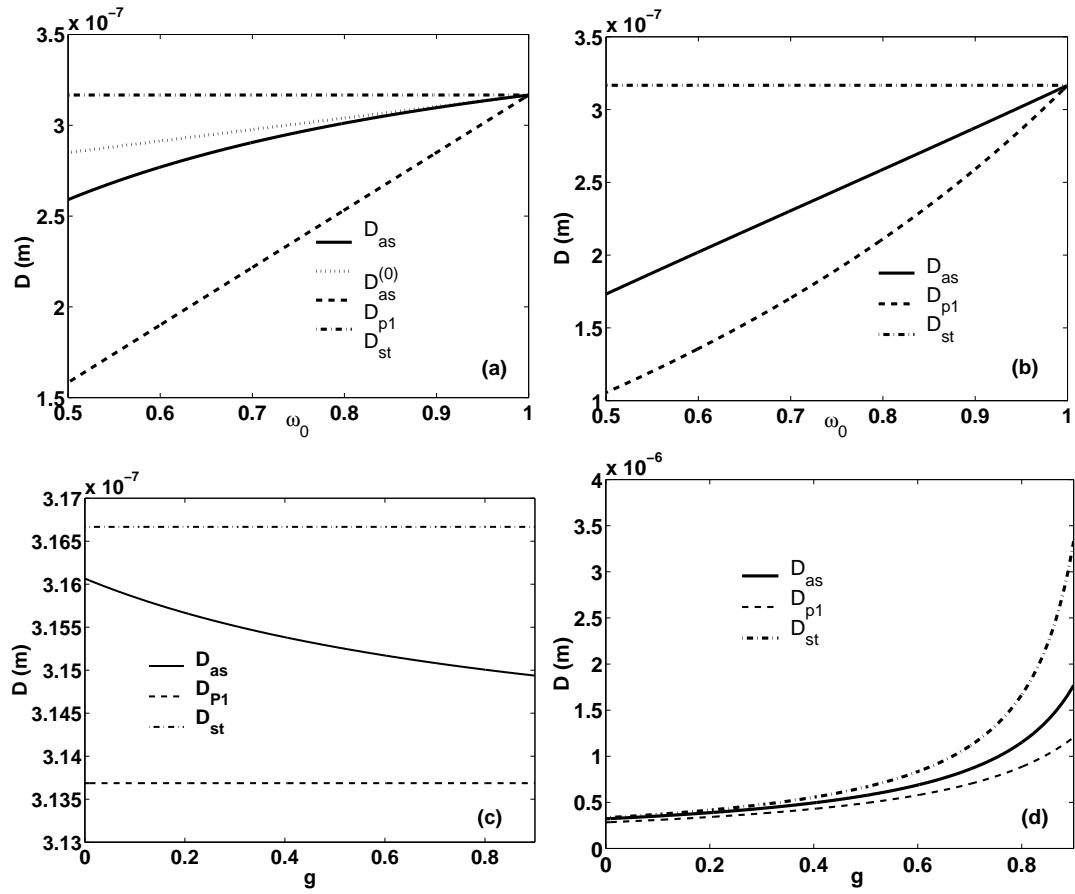


Figure 2, R. Elaloufi et al.

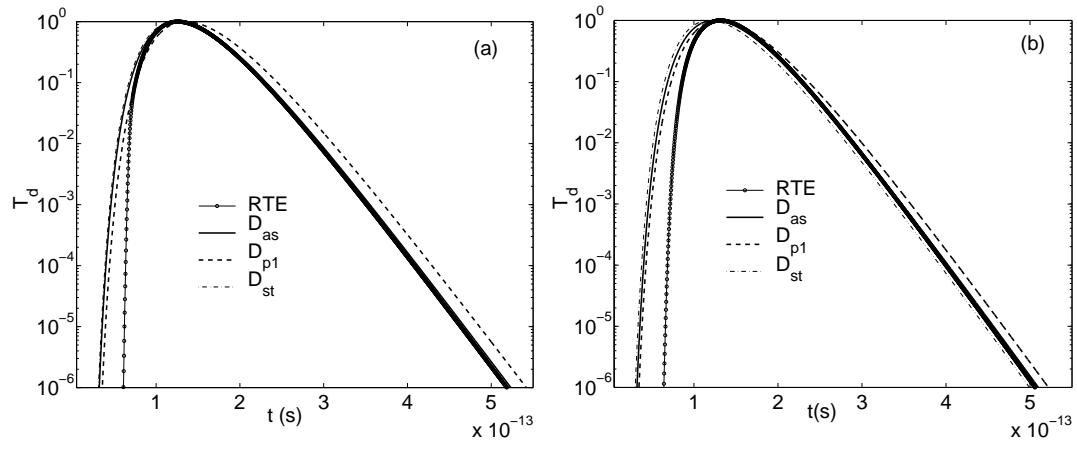


Figure 3, R. Elaloufi et al.

	$\omega_0 = 0.988$		$\omega_0 = 0.840$	
	$g = 0$	$g = 0.343$	$g = 0$	$g = 0.343$
$k_0$ ( $\mu m^{-1}$ )	0.1989	0.1616	0.8100	0.6800
$k_1$ ( $\mu m^{-1}$ )	1.0662	1.0661	1.2539	1.2539
$k_2$ ( $\mu m^{-1}$ )	1.0700	1.0697	1.2585	1.2582

Table 1, R. Elaloufi et al.

# Annexe 3

## Rayonnement thermique aux courtes échelles

Liste des articles reproduits :

- R. Carminati and J.-J. Greffet, *Near-field effects in spatial coherence of thermal sources*, Phys. Rev. Lett. **82**, 1660–1663 (1999).
- A.V. Shchegrov, K. Joulain, R. Carminati and J.-J. Greffet, *Near-field spectral effects due to electromagnetic surface excitations*, Phys. Rev. Lett. **85**, 1548–1551 (2000).
- J.-J. Greffet, R. Carminati, K. Joulain, J.P. Mulet, S. Mainguy and Y. Chen, *Coherent emission of light by thermal sources*, Nature **416**, 61–64 (2002).
- J.P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, *Nanoscale radiative heat transfer between a small particle and a plane surface*, Appl. Phys. Lett. **78**, 2931–2933 (2001).
- K. Joulain, R. Carminati, J.P. Mulet and J.-J. Greffet, *Near-field thermal-emission spectroscopy*, manuscrit en préparation (2002).



## Near-Field Effects in Spatial Coherence of Thermal Sources

Rémi Carminati and Jean-Jacques Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris,  
Centre National de la Recherche Scientifique, 92295 Châtenay-Malabry Cedex, France*

(Received 16 October 1998)

We present an exact calculation of the cross-spectral density tensor of the near field thermally emitted into free space by an opaque planar surface. The approach, based on fluctuational electrodynamics and the fluctuation-dissipation theorem, yields novel near-field correlation properties. We show that the spatial coherence length of the field close to the surface at a given wavelength  $\lambda$  may be much smaller than the well-known  $\lambda/2$  of blackbody radiation. We also show that a long-range correlation may exist, when resonant surface waves, such as surface-plasmon or surface-phonon polaritons, are excited. These results should have important consequences in the study of coherence in thermal emission and in the modeling of nanometer scale radiative transfer. [S0031-9007(99)08515-4]

PACS numbers: 42.25.Kb, 05.40.-a, 71.36.+c, 78.20.Ci

Thermal emission of radiation from opaque materials is often regarded as an incoherent phenomenon. Indeed, thermal light is chiefly generated by the uncorrelated process of spontaneous emission. Nevertheless, since the development of optical coherence theory, it has been known that the field emitted by thermal sources may exhibit a certain degree of temporal and spatial coherence. For example, it was shown that light across a planar quasi-homogeneous Lambertian source, at a given wavelength  $\lambda$ , is spatially correlated over a distance on the order of  $\lambda/2$  [1]. This result was found to be in agreement with known spatial correlation properties of free blackbody radiation [2].

In deriving this result, the nonradiating near-field part of the emitted light was ignored [1], because it plays no role in the far-field properties of emission from planar sources. Nevertheless, recent interest in microscale and nanoscale radiative transfers [3], together with the development of local-probe thermal microscopy [4] and the observation of coherent thermal emission from doped silicon and silicon carbide (SiC) gratings [5,6], has raised new challenges. In fact, all these topics have in common the substantial role of the nonradiating (evanescent) thermal fields.

In this Letter, we concentrate on the role of the nonradiating field in the coherence properties of thermal sources, with particular emphasis on the spatial coherence. We revisit this concept by studying the spatial correlation of the vector thermal field at close proximity of an emitting body. Our calculation is based on macroscopic fluctuational electrodynamics and the fluctuation-dissipation theorem [7,8]. Novel properties are obtained, showing that near-field effects may dramatically modify the spatial coherence of thermal sources.

Let us consider a homogeneous opaque material filling the half space  $z < 0$ , in local thermodynamic equilibrium at a uniform temperature  $T$ . The electrodynamic properties of the material, assumed to be isotropic and nonmagnetic, are macroscopically described by its complex frequency-dependent dielectric constant  $\epsilon(\omega)$ . The

thermal fluctuations of the current density at a point  $\mathbf{r} = (x, y, z)$  inside the body are described by a random process  $\mathbf{j}(\mathbf{r}, t)$ , which is stationary in time [7,8]. The field radiated into the half space  $z > 0$ , i.e., the thermally emitted field, is itself a fluctuating quantity described by a time-stationary random process  $\mathbf{E}(\mathbf{r}, t)$ . In this Letter, we focus on the spatial coherence of the emitted field at a given frequency  $\omega$ . The basic quantity of second-order coherence theory of vector fields in the space-frequency domain is the electric-field cross-spectral density tensor  $W_{jk}$  defined by [9]

$$\langle E_j(\mathbf{r}_1, \omega) E_k^*(\mathbf{r}_2, \omega') \rangle = W_{jk}(\mathbf{r}_1, \mathbf{r}_2, \omega) \delta(\omega - \omega'), \quad (1)$$

where  $\mathbf{E}(\mathbf{r}, \omega)$  is the time-domain Fourier transform of the electric field  $\mathbf{E}(\mathbf{r}, t)$  and the superscript  $*$  denotes the complex conjugate. The brackets denote a statistical ensemble average. The tensor  $W_{jk}$  is a measure of the spatial correlation of the electric field at a given frequency  $\omega$  and at two different points  $\mathbf{r}_1$  and  $\mathbf{r}_2$ . Note that the presence of the delta function in Eq. (1) is a consequence of the stationarity of the field in the time domain. Recently, calculation of the cross-spectral density tensor  $W_{jk}$  was reported for an infinite blackbody surface in the far field or for a finite planar surface in the radiometric limit [10]. In this Letter, we address near-field effects—in some cases under resonant conditions—so that we do not use any approximation in our calculation. The only assumptions are the local thermal equilibrium of the body and the use of macroscopic electrodynamics. To proceed, we introduce the time-domain Fourier transform  $\mathbf{j}(\mathbf{r}, \omega)$  of the random current density  $\mathbf{j}(\mathbf{r}, t)$ . The electric field at a given point  $\mathbf{r}$  in the half space  $z > 0$  is given by

$$\mathbf{E}(\mathbf{r}, \omega) = i\mu_o\omega \int_V \vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}', \omega) \cdot \mathbf{j}(\mathbf{r}', \omega) d^3\mathbf{r}', \quad (2)$$

where the integration is performed on the volume  $V$  of the body.  $\vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}', \omega)$  is the Green dyadic of the vector Helmholtz equation in the considered geometry, namely, a flat interface separating a semi-infinite homogeneous

medium with dielectric constant  $\epsilon(\omega)$  (medium  $z < 0$ ) from a vacuum (medium  $z > 0$ ) [11].

In order to calculate the cross-spectral density tensor  $W_{jk}$ , we need to know the spatial correlation function of the thermal current fluctuations in the frequency domain. It is given by the fluctuation-dissipation theorem [7,8]:

$$\langle j_m(\mathbf{r}, \omega) j_n^*(\mathbf{r}', \omega') \rangle = \frac{\omega}{\pi} \epsilon_o \epsilon''(\omega) \Theta(\omega, T) \delta(\mathbf{r} - \mathbf{r}') \delta_{mn} \delta(\omega - \omega'), \quad (3)$$

where  $\delta_{mn}$  is the Kronecker symbol,  $\Theta(\omega, T) = \hbar\omega/2 + \hbar\omega/[\exp(\hbar\omega/kT) - 1]$  is the mean energy of the quantum harmonic oscillator in thermal equilibrium at temperature  $T$ , and  $2\pi\hbar$  is Planck's constant.  $\epsilon''(\omega)$  is the imaginary part of the dielectric constant  $\epsilon(\omega)$ .

Inserting Eq. (2) into Eq. (1) and using Eq. (3), we obtain the following expression for the cross-spectral tensor of the electric field in the half space  $z > 0$ :

$$W_{jk}(\mathbf{r}_1, \mathbf{r}_2, \omega) = \frac{\omega^3}{\pi} \mu_o^2 \epsilon_o \epsilon'' \Theta(\omega, T) \int_V G_{jm}(\mathbf{r}_1, \mathbf{r}', \omega) G_{km}^*(\mathbf{r}_2, \mathbf{r}', \omega) d^3\mathbf{r}'. \quad (4)$$

The expression of the Green dyadic  $\vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}', \omega)$  connecting a source current inside the body (medium  $z < 0$ ) to the electric field in vacuum (medium  $z > 0$ ) is [11]

$$\vec{\mathbf{G}}(\mathbf{r}, \mathbf{r}', \omega) = \frac{i}{8\pi^2} \int \frac{1}{\gamma_2} (\hat{s}t_s\hat{s} + \hat{p}_1t_p\hat{p}_2) \exp[i\mathbf{K} \cdot (\mathbf{R} - \mathbf{R}')] \times \exp(i\gamma_1z - i\gamma_2z') d^2\mathbf{K}, \quad (5)$$

where  $\mathbf{r} = (\mathbf{R}, z)$ ,  $\hat{s} = \hat{\mathbf{K}} \times \hat{z}$ ,  $\hat{p}_j = (|\mathbf{K}|\hat{z} + \gamma_j\hat{\mathbf{K}})/k_j$ , the symbol  $\hat{\phantom{a}}$  denoting a unit vector,  $k_1 = k = \omega/c$ ,  $k_2 = \sqrt{\epsilon}k$ , and  $\gamma_j = (k_j^2 - \mathbf{K}^2)^{1/2}$ , with the determination  $\text{Re}(\gamma_j) > 0$  and  $\text{Im}(\gamma_j) > 0$ .  $t_s(\mathbf{K})$  and  $t_p(\mathbf{K})$  are the Fresnel transmission factors for  $s$  and  $p$  polarizations, respectively [11].

Equation (4), together with Eq. (5), gives an exact expression of the electric cross-spectral density tensor  $W_{jk}$ . This expression is valid for any position of the two observation points  $\mathbf{r}_1$  and  $\mathbf{r}_2$ , especially at close proximity from the surface (near-field zone). Inserting (5) into (4), tedious but straightforward algebra allows us to express the elements  $W_{jk}(\mathbf{r}_1, \mathbf{r}_2, \omega)$  of the cross-spectral tensor in terms of a single integral over the wave vector  $|\mathbf{K}|$  which is evaluated numerically.

Let us first compare the spatial correlation of the field emitted by lossy glass and tungsten, the latter being a metal which does not exhibit surface-polariton resonances in the visible part of the spectrum. We plot in Fig. 1 the diagonal element  $W_{xx}(\mathbf{r}_1, \mathbf{r}_2, \omega)$  of the cross-spectral density tensor, at a wavelength  $\lambda = 2\pi/k = 500$  nm and normalized by its value at  $\rho = 0$ . Note that this normalization amounts to plotting the component  $\mu_{xx}(\mathbf{r}_1, \mathbf{r}_2, \omega)$  of the (tensor) spectral degree of spatial coherence [9]. At this wavelength, the dielectric constant of a lossy glass is  $\epsilon =$

$2.25 + 0.001i$  and that of tungsten is  $\epsilon = 4.35 + 18.05i$ . The calculation is performed in a plane  $z = z_0$  above the surface of the emitting material. Both  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are along the  $x$  axis, and the result is plotted versus  $\rho = |\mathbf{r}_1 - \mathbf{r}_2|$ . In the very near field ( $z_0 = 0.01\lambda$ ), the curve corresponding to glass (solid line) drops to negligible values after  $\rho = \lambda/2$ , showing that the correlation length of the  $x$  component of the field is  $\lambda/2$ . In fact, a semi-infinite medium of lossy glass is a good approximation of a planar black-body radiator, and the solid curve in Fig. 1 strongly resembles the  $\sin(kr)/kr$  shape of the cross-spectral density in the source plane of a Lambertian source, previously obtained in the scalar approximation [1]. Note that, although not shown here for brevity, we have observed this behavior for the three diagonal elements  $W_{jj}$  of the cross-spectral density tensor. In comparison, the case of tungsten (dotted curve) is completely different. *The correlation length is much smaller than  $\lambda/2$ , on the order of  $0.06\lambda$ .* This distance is comparable to the skin depth  $\delta = [k \text{Im}(\sqrt{\epsilon})]^{-1}$  of tungsten at this wavelength. This is actually not surprising because  $\delta$  is, in addition to the wavelength, the relevant length scale for the propagation of electromagnetic waves inside the material. Therefore, it is physically sound that the induced currents, and thus the field close to the surface, be correlated over a distance which is the smallest of  $\delta$  and  $\lambda$ . This result, which, to our knowledge, was never pointed out before, has important consequences in the modeling of radiative transfer at small scales [3,4]. Moreover, it allows us to revisit the concept of emissivity at subwavelength scale, useful in the study of the radiative properties of rough surfaces [12]. Indeed, this macroscopic concept can be defined on a length scale larger than the field correlation length.

Finally, we show that this subwavelength correlation length is a pure near-field effect, due to nonradiative evanescent fields. At a distance  $z_0 = 0.1\lambda$ , we see that the correlation length for tungsten (dashed curve in Fig. 1) is much larger (on the order of  $0.4\lambda$ ) than that obtained with

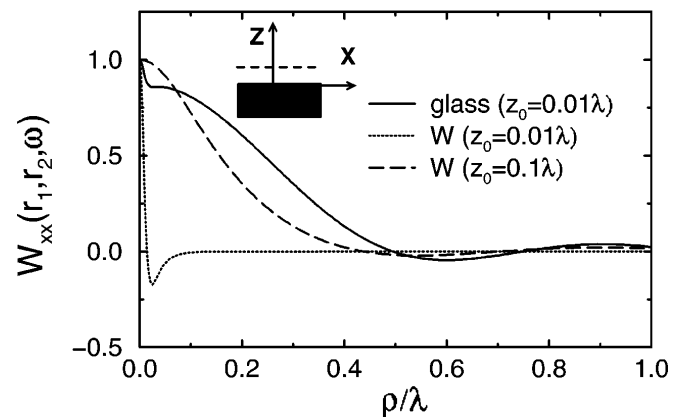


FIG. 1.  $W_{xx}(\mathbf{r}_1, \mathbf{r}_2, \omega)$  in the plane  $z = z_0$  versus  $\rho = |\mathbf{r}_1 - \mathbf{r}_2|$ .  $\mathbf{r}_1$  and  $\mathbf{r}_2$  are on the  $x$  axis.  $\lambda = 500$  nm. Two materials are considered: lossy glass ( $z_0 = 0.01\lambda$ ) and tungsten ( $z_0 = 0.01\lambda$  and  $z_0 = 0.1\lambda$ ). All curves are normalized by their maximum value at  $\rho = 0$ .

$z_0 = 0.01\lambda$  (dotted curve). Thus, during the transition from the near to the far zone, the correlation length increases because evanescent waves vanish. The value  $\lambda/2$  of blackbody radiation is reached when  $z \geq 0.1\lambda$ . Note that this increase in spatial coherence is *not described by the far-field form of the van Cittert-Zernike theorem* [9,13], which describes an increase of spatial coherence due to propagation.

We now turn to the study of spatial coherence in light emission from materials supporting resonant surface waves, such as surface-plasmon or surface-phonon polaritons [14]. In the formalism used here, the existence of polaritons is taken into account in the Green dyadic  $\bar{\mathbf{G}}$  which describes all the electrodynamic properties of the interface. In fact, it is well known that the  $p$ -polarized Fresnel transmission factor  $t_p(\mathbf{K})$  appearing in Eq. (5) possesses a pole when  $|\mathbf{K}|$  approaches the value  $K_{sp} = k[\epsilon/(\epsilon + 1)]^{1/2}$  of the wave vector of the polariton along the interface [14]. The thermal excitation of a surface polariton induces some spatial correlation in the field close to the surface, and we may expect a large increase of the correlation length.

We illustrate in Fig. 2 the effect of surface-plasmon [Fig. 2(a)] and surface-phonon [Fig. 2(b)] polaritons on the spatial coherence of the thermal near field. We plot in Fig. 2(a) the element  $W_{xx}$  of the cross-spectral density tensor at the wavelength  $\lambda = 620$  nm, and in the plane  $z_0 = 0.05\lambda$ , for three different metals. At this wavelength, their dielectric constants are  $\epsilon = 4.6 + i20.5$  for tungsten,  $\epsilon = -8.26 + i1.12$  for gold, and  $\epsilon = -15.04 + i1.02$  for silver. Both gold and silver exhibit surface-plasmon resonances at this wavelength. We clearly see that whereas the spatial correlation length for tungsten is a fraction of the wavelength (as in Fig. 1), the correlation length for gold and silver is much larger. In fact, although Fig. 2 is limited to  $\rho < 5\lambda$  for the sake of visibility, the correlation extends on a distance given by the attenuation length of the surface-plasmon polariton. For gold and silver, the attenuation lengths are  $16\lambda$  and  $65\lambda$ , respectively. The same effect is seen in Fig. 2(b) for a SiC crystal, which exhibits a surface-phonon polariton resonance at the wavelength  $\lambda = 11.36 \mu\text{m}$  ( $\epsilon = -7.56 + i0.41$ ) and no resonance at  $\lambda = 9.1 \mu\text{m}$  ( $\epsilon = 1.80 + i4.07$ ). The emission of SiC gratings and the effect of surface-phonon waves were studied experimentally and numerically in Ref. [6]. The difference of behavior of this material at the two different wavelengths is striking in Fig. 2(b). The correlation length is much higher in the presence of the resonant surface wave (dashed line) than in the case where no surface wave is excited (solid line). The propagation distance of the surface-phonon polariton in this case is  $36\lambda$ . In summary, we have shown how the delocalized electromagnetic surface mode, either coupled to a plasmon or a phonon in the material, correlates the near field on distances on the order of its propagation length, which easily reaches several tenths of wavelengths in the visible part of the spectrum. Physically, one could say that the collective resonance of the free electrons in the metals for

plasmons, or of the crystal for phonons (lattice vibrations), transmits its spatial coherence to the electromagnetic field.

In order to demonstrate more precisely the role of the surface wave on the long-range correlation, we show in Fig. 3 the three diagonal elements  $W_{kk}$  of the cross-spectral density tensor, for gold, in the same conditions as in Fig. 2(a). It is well known that a surface polariton propagating along the  $x$  axis is polarized in the  $x$ - $z$  plane [14]. Thus, when looking at the correlation between two points  $\mathbf{r}_1$  and  $\mathbf{r}_2$  that are along the  $x$  axis, we expect to see a surface-wave induced correlation for the  $x$  and  $z$  components of the field only. This is exactly what is observed in Fig. 3. The elements  $W_{xx}$  and  $W_{zz}$  exhibit long-range correlation due to the surface wave, whereas  $W_{yy}$  exhibits the same behavior as that of a metal without surface-wave resonance. This is a clear signature of the role of the surface polariton in increasing the spatial coherence. Although not shown here for brevity, the same result is obtained for SiC with surface-phonon excitation.

Another signature of the surface-wave excitation is the  $z$  dependence of the electric energy density  $u_e(z, \omega) = \sum W_{kk}(\mathbf{r}, \mathbf{r}, \omega)$ . Note that at a given frequency  $\omega$ ,  $u_e$  is a function of  $z$  only due to the translational invariance of the geometry in the  $x$ - $y$  plane. The energy density  $u_e$ , normalized by its far-field value, is plotted in Fig. 4 versus  $z$ . For materials without surface-wave resonances, such as glass and tungsten, the energy density increases at short distance ( $z < 0.1\lambda$ ). In fact, it is known that  $u_e$  behaves like  $1/z^3$  in the very near field, due to nonradiating fields [8]. This result is retrieved in our calculation.

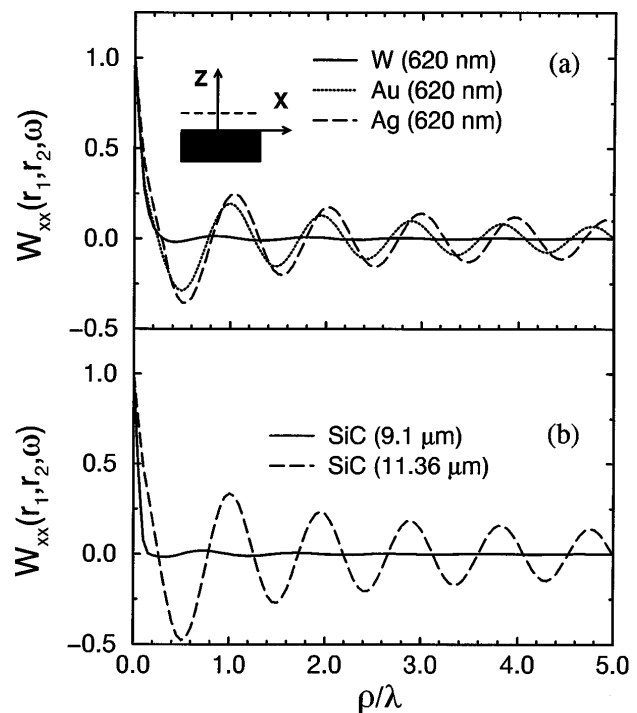


FIG. 2. Same as Fig. 1, with  $z_0 = 0.05\lambda$ . (a) The case of three metals (tungsten, gold, and silver),  $\lambda = 620$  nm. (b) The case of SiC with  $\lambda = 9.1 \mu\text{m}$  and  $\lambda = 11.36 \mu\text{m}$ .

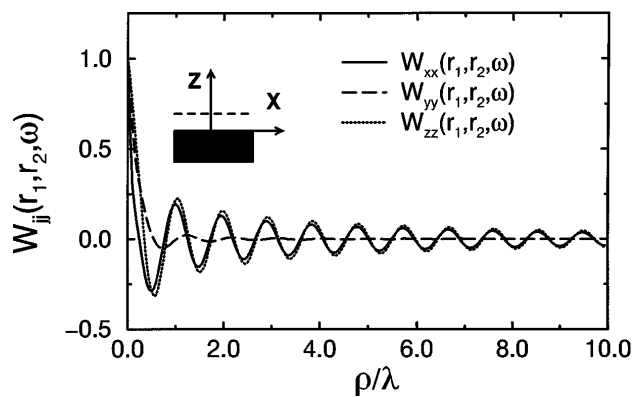


FIG. 3. Diagonal elements  $W_{jj}(\mathbf{r}_1, \mathbf{r}_2, \omega)$  in the plane  $z_0 = 0.05\lambda$  versus  $\rho = |\mathbf{r}_1 - \mathbf{r}_2|$ , with  $\mathbf{r}_1$  and  $\mathbf{r}_2$  on the  $x$  axis. Gold surface:  $\lambda = 620$  nm.

Also note that the existence of these nonradiating fields at short distances  $z$  is responsible for the small peak observed at  $\rho = 0$  for glass in Fig. 1 (solid line). This is a striking difference between free blackbody radiation and the radiation produced at short distance by a blackbody radiator, which contains evanescent waves. The novel property is that for a material supporting a surface-plasmon polariton, such as gold in Fig. 4, the energy density starts to increase as soon as  $z < \lambda$ . This is so because the decay length of the surface wave in the  $z$  direction is on the order of  $\lambda$ . The near-field enhancement is up to 100 times higher than that observed without surface-wave excitation. Although not shown here, we have observed the same effect with a surface-phonon polariton on a SiC surface. This behavior of the electric energy density shows that evanescent waves play a crucial role in radiative heat transfer at subwavelength scale when surface waves are excited.

We have shown that the thermal excitation of an electromagnetic surface mode dramatically changes the spatial correlation of the emitted field close to the surface. By using a grating as a coupler, the surface wave can be converted into propagating waves. This give rise to *highly directional thermal emission* at a given wavelength. This is a very unusual behavior for thermal sources. Since such a source emits light in particular directions and at particular frequencies, *the spectrum of the emitted light strongly depends on the direction of observation*. Thus, such a material is an example of a *natural* thermal source for which the field correlation in the source plane does not obey the scaling law established by Wolf [15]. Note that temporal and spatial coherence in thermal emission by silicon and SiC gratings has already been observed [5,6]. The analysis presented in this Letter explains the physical origin of this phenomenon.

To summarize, we have found novel correlation properties of fields produced by thermal sources in the near field: (i) Materials that do not support surface waves may display coherence lengths much shorter than  $\lambda/2$ , (ii) ma-

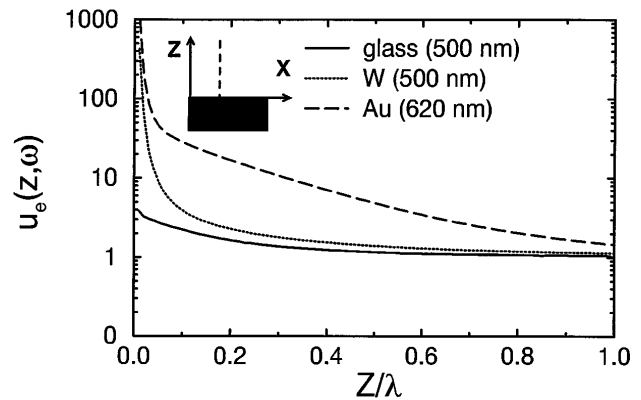


FIG. 4. Electric energy density  $u_e(\mathbf{r}, \omega)$  versus  $z$ . The case of glass ( $\lambda = 500$  nm), tungsten ( $\lambda = 500$  nm), and gold ( $\lambda = 620$  nm). All curves are normalized by their far-field value at  $z = 10\lambda$ .

terials that support resonant surface waves exhibit long-range spatial correlations, (iii) the spectral electric density of energy is strongly enhanced close to an interface when a surface wave is excited, and (iv) an interface with a shallow grating supporting a surface wave is a natural thermal source whose spectrum depends on the emission direction.

- [1] W.H. Carter and E. Wolf, J. Opt. Soc. Am. **65**, 1067 (1975); E. Wolf, J. Opt. Soc. Am. **68**, 6 (1978).
- [2] *Inverse Source Problems*, edited by H.P. Baltes (Springer-Verlag, Berlin, 1978), p. 123.
- [3] C.L. Tien and G. Chen, J. Heat Transfer **114**, 799 (1994).
- [4] C.C. Williams and H.K. Wickramasinghe, Appl. Phys. Lett. **49**, 1587 (1986).
- [5] P.J. Hesketh *et al.*, Nature (London) **324**, 549 (1986); Phys. Rev. B **37**, 10803 (1988).
- [6] J. Le Gall *et al.*, Phys. Rev. B **55**, 10105 (1997).
- [7] L.D. Landau, E.M. Lifshitz, and L.P. Pitaevskii, *Statistical Physics* (Pergamon Press, Oxford, 1980), 3rd ed., Part 1, Chap. XII and Part 2, Chap. VIII.
- [8] S.M. Rytov, Yu.A. Kravtsov, and V.I. Tatarskii, *Principles of Statistical Radiophysics* (Springer-Verlag, Berlin, 1989), Vol. 3, Chap. 3.
- [9] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, England, 1995).
- [10] D.C. Bertilone, J. Mod. Opt. **43**, 207 (1996); J. Opt. Soc. Am. A **14**, 693 (1997).
- [11] J.E. Sipe, J. Opt. Soc. Am. B **4**, 481 (1987).
- [12] R. Carminati *et al.*, in *Heat Transfer 1998*, edited by J.S. Lee, Proceedings of the 11th International Heat Transfer Conference (Korean Society of Mechanical Engineering, Kyongju, 1998), Vol. 7, p. 427.
- [13] J.W. Goodman, *Statistical Optics* (Wiley, New York, 1985), p. 206.
- [14] *Surface Polaritons*, edited by V.M. Agranovich and D.L. Mills (North-Holland, Amsterdam, 1982).
- [15] E. Wolf, Phys. Rev. Lett. **56**, 1370 (1986).

## Near-Field Spectral Effects due to Electromagnetic Surface Excitations

Andrei V. Shchegrov\*

*Rochester Theory Center for Optical Science and Engineering  
and Department of Physics and Astronomy, University of Rochester, Rochester, New York 14627-0171*

Karl Joulain, Rémi Carminati, and Jean-Jacques Greffet

*Laboratoire d'Energétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris, Centre National de la Recherche Scientifique, 92295, Châtenay-Malabry Cedex, France*

(Received 14 March 2000)

We demonstrate theoretically that the spectra of electromagnetic emission of surface systems can display remarkable differences in the near and the far zones. The spectral changes occur due to the loss of evanescent modes and are especially pronounced for systems which support surface waves.

PACS numbers: 78.20.-e, 05.40.-a, 44.40.+a, 87.64.Xx

Spectroscopy of electromagnetic radiation is perhaps the most powerful exploration tool employed in natural sciences: astronomy, atomic and molecular physics and chemistry, materials science, biology, etc. The central question considered in this paper—whether the spectral content of the radiation emitted by an object can change on propagation to the observer—usually does not arise, since it seems natural that nothing can happen to waves traveling through empty space. Surprisingly similar failure of common sense was put forward by the recent progress of near-field optical microscopy [1,2], which achieves sub-wavelength resolution exactly because evanescent modes carrying subwavelength spatial information do not propagate far away from the object. However, a great deal of work devoted to such an irreversible *change of spatial information* on propagation has been accompanied with significantly lesser interest in the possibility of the *change of spectral information*.

The subject of spectral changes on propagation has, however, been addressed. In the 1980's, Wolf [3] predicted that the spectrum of light can be changed on propagation from the source to the observer, even through empty space. This effect, whose origin lies in the fluctuating nature of the source, has been intensively studied in a variety of systems [4,5]. Typically, the Wolf effect is manifested in small spectral shifts and can be viewed as a redistribution of the weights of different spectral components.

In this Letter, we demonstrate spectral changes, whose physical origin is very different from that of Wolf spectral shifts and lies in the presence of the evanescent component in the emitted field. We show that the near-field and far-field spectra of emitted electric fields can display drastic differences. For a broadband emission, such as thermal emission which is considered in detail in this paper, the near-field spectrum dominated by evanescent modes can be entirely different from the far-field spectrum of propagating modes. These spectral changes occur not due to the statistical nature of the source (as in the Wolf effect) but due to the loss of evanescent components on propagation. We analyze how such spectral changes can be enhanced

by electromagnetic surface waves (SW) near the interface. These SW are known to play an important role in the enhancement of interaction between nanoparticles near the surface [6], in localization effects on random surfaces [7], in surface-enhanced Raman scattering [8], in extraordinary transmission of light through subwavelength holes [9], etc. In this Letter, we show that SW provide the leading contribution to the density of energy in the near-field zone of electromagnetic emission.

We now proceed with analyzing near-field effects in the spectra of thermal emission, which provides an easy way to excite both propagating and evanescent electromagnetic modes in a wide range of frequencies (at least, in the infrared [10–12]). Thermal emission is frequently associated with the textbook example of equilibrium blackbody radiation. The Planck spectrum  $I_{BB}(\omega)$  of such radiation is obtained by multiplying the thermal energy  $\theta(\omega, T) = \hbar\omega / [\exp(\hbar\omega/k_B T) - 1]$  of a quantum oscillator by the density of oscillations (modes) per unit volume  $N(\omega)d\omega = \omega^2 d\omega / (\pi^2 c^3)$  in the frequency interval  $(\omega, \omega + d\omega)$ , and dividing the result by  $d\omega$  [13],

$$I_{BB}(\omega) = \theta(\omega, T)N(\omega) = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\exp(\hbar\omega/k_B T) - 1}. \quad (1)$$

Here  $T$  is the body temperature,  $k_B$  is Boltzmann's constants,  $\hbar$  is Planck's constant divided by  $2\pi$ , and  $c$  is the speed of light in vacuum. A well-known representation of blackbody radiation is the equilibrium radiation in a closed cavity with lossy walls when only *propagating* modes of the field are taken into account.

To demonstrate the importance of near-field effects, we consider a somewhat more sophisticated example of thermal emission from a semi-infinite ( $z < 0$ ) slab of homogeneous, nonmagnetic material held in local thermodynamic equilibrium at a uniform temperature  $T$ , into the empty half-space  $z > 0$ . We will describe the macroscopic dielectric properties of the material by a frequency-dependent, complex dielectric function  $\varepsilon(\omega) = \varepsilon'(\omega) + i\varepsilon''(\omega)$ . The

Fourier component  $\mathbf{E}(\mathbf{r}, \omega)$  of the electric field  $\mathbf{E}(\mathbf{r}, t)$  at a point  $\mathbf{r} = (x, y, z)$  in the empty half space  $z > 0$  is generated by thermal currents with density  $\mathbf{j}(\mathbf{r}', \omega)$ , which is nonzero only for  $z' < 0$ . It can be computed following the procedure outlined in Refs. [12,14],

$$E_\alpha(\mathbf{r}, \omega) = i\mu_0\omega \sum_{\beta=x,y,z} \int_V d^3r' G_{\alpha\beta}(\mathbf{r}, \mathbf{r}', \omega) j_\beta(\mathbf{r}', \omega), \quad (2)$$

where  $V$  is the volume of the hot body which occupies the half-space  $z' < 0$ , and  $G_{\alpha\beta}(\mathbf{r}, \mathbf{r}', \omega)$  is the electromagnetic Green tensor for the system of two homogeneous materials separated by a planar interface  $z = 0$ . According to the fluctuation-dissipation theorem [14], the fluctuations of thermal currents are described by the correlation function

$$\langle j_\alpha(\mathbf{r}, \omega) j_\beta(\mathbf{r}', \omega') \rangle = \frac{\omega\theta(\omega, T)}{\pi} \varepsilon_0 \varepsilon''(\omega) \delta_{\alpha\beta} \delta(\mathbf{r} - \mathbf{r}') \times \delta(\omega - \omega'), \quad (3)$$

where the angle brackets denote the statistical ensemble average. The Kronecker symbol  $\delta_{\alpha\beta}$  and the spatial  $\delta$  function in this formula follow from the assumption that the dielectric function is isotropic, homogeneous, and local [14].

The energy density  $I(\mathbf{r}, \omega)$  of the emitted electric field at the point  $\mathbf{r}$  is defined by the formula

$$\sum_{\alpha=x,y,z} \frac{\varepsilon_0}{2} \langle E_\alpha^*(\mathbf{r}, \omega) E_\alpha(\mathbf{r}, \omega') \rangle \equiv I(\mathbf{r}, \omega) \delta(\omega - \omega'). \quad (4)$$

Using Eqs. (2) and (3) into (4), we obtain, for  $I(\mathbf{r}, \omega)$ ,

$$I(\mathbf{r}, \omega) = 8\pi^3 \frac{\omega^3}{c^4} \theta(\omega, T) \varepsilon''(\omega) \sum_{\alpha,\beta=x,y,z} \int_{-\infty}^0 dz' \times \int \frac{d^2k_\parallel}{(2\pi)^2} |g_{\alpha\beta}(\mathbf{k}_\parallel, \omega | z, z')|^2, \quad (5)$$

where  $g_{\alpha\beta}(\mathbf{k}_\parallel, \omega | z, z')$  is the analytically known [12] 2D spatial Fourier transform (in  $x$  and  $y$ ) of the Green's tensor  $G_{\alpha\beta}(\mathbf{r}, \mathbf{r}', \omega)$ . Note that  $I(\mathbf{r}, \omega)$  in Eq. (5) is independent of  $x$  and  $y$ , due to the translational invariance of the system in  $x$  and  $y$  directions.

We now assume that the interface  $z = 0$  between the material and a vacuum can support electromagnetic SW. The dispersion relation between the wave number  $k_\parallel = |\mathbf{k}_\parallel|$  and frequency  $\omega$  of SW is

$$[k_\parallel^{\text{SW}}(\omega)]^2 = (\omega^2/c^2)\varepsilon(\omega)/[\varepsilon(\omega) + 1]. \quad (6)$$

Such waves exist for materials having  $\varepsilon'(\omega) < -1$  in one or several frequency ranges [10]. We consider SiC, which supports SW known as surface phonon polaritons and which has been used in previous experimental [11] and theoretical [12] investigations of thermal emission. The dielectric function of this material is given by the expression  $\varepsilon(\omega) = \varepsilon_\infty(\omega_L^2 - \omega^2 - i\gamma\omega)/(\omega_T^2 - \omega^2 - i\gamma\omega)$  with

$\varepsilon_\infty = 6.7$ ,  $\omega_L = 182.7 \times 10^{12} \text{ s}^{-1}$ ,  $\omega_T = 149.5 \times 10^{12} \text{ s}^{-1}$ , and  $\gamma = 0.9 \times 10^{12} \text{ s}^{-1}$  [11]. By substituting  $\varepsilon(\omega)$  into Eq. (5) and performing a straightforward evaluation of integrals (only the integral over the magnitude of  $\mathbf{k}_\parallel$  has to be calculated numerically, the other two integrals can be evaluated analytically), we obtain the spectra of thermal emission for SiC at different heights  $z$  above the surface. We plot the results in Fig. 1 in the frequency range  $0 < \omega < 400 \times 10^{12} \text{ s}^{-1}$  for  $T = 300 \text{ K}$  at three different heights.

Although one could expect to find differences of the SiC spectra with the blackbody spectrum (1), it is striking that near-field and far-field spectra of the same SiC sample are so dramatically different, as seen in Fig. 1. An observer doing a traditional far-zone spectroscopic measurement (Fig. 1a) would detect the spectrum with a rather wide dip in the range  $150 \times 10^{12} \text{ s}^{-1} < \omega < 180 \times 10^{12} \text{ s}^{-1}$  due to the low emissivity of SiC in that range [11]. Note that the sample effectively acts as a *nonradiating* source in this frequency range. However, when the probe moves within a *subwavelength* distance from the material (typical thermal emission wavelengths at  $T = 300 \text{ K}$  are of the order of  $10 \mu\text{m}$ ), the spectrum starts to change rapidly. In Fig. 1b, showing the emission spectrum at  $2 \mu\text{m}$  above the surface, this change is seen as a peak emerging at  $\omega = 178.7 \times 10^{12} \text{ s}^{-1}$ . At very close distances (Fig. 1c), the peak becomes so strong that an observer would surprisingly see almost *monochromatic* emission with photon energies not represented in the far zone.

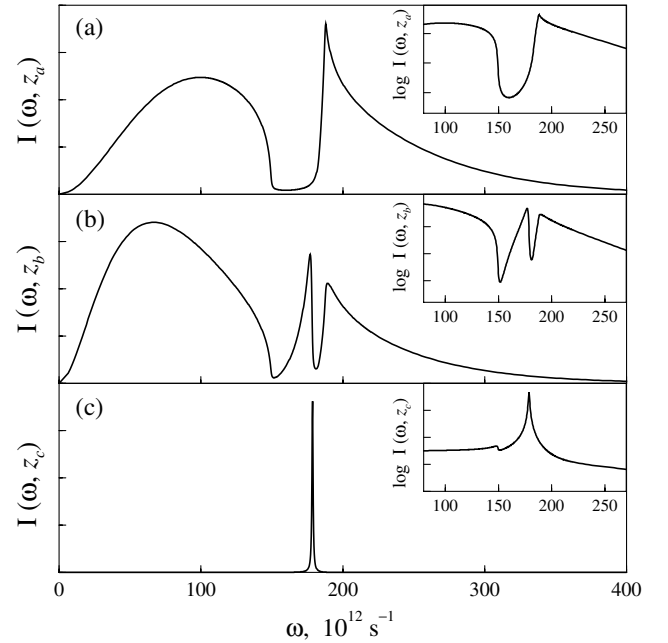


FIG. 1. Spectra of thermal emission of a semi-infinite sample of SiC at  $T = 300 \text{ K}$  and three different heights above the surface: (a)  $z_a = 1000 \mu\text{m}$ , (b)  $z_b = 2 \mu\text{m}$ , (c)  $z_c = 0.1 \mu\text{m}$ . The insets magnify the spectra plotted on a semilog scale in the range of strong contribution from evanescent surface modes.

Thus, we find that the spectrum changes *qualitatively* on propagation. The occurrence of such striking spectral changes is related to SW existing in the region  $150 \times 10^{12} \text{ s}^{-1} < \omega < 180 \times 10^{12} \text{ s}^{-1}$ . We shall now clarify the mechanism of the formation of spectra  $I(\mathbf{r}, \omega)$  at different distances from the surface.

We note that the spectrum of the electric field (5) has a similar structure as that of blackbody radiation (1),

$$I(z, \omega) = \theta(\omega, T)N(z, \omega). \quad (7)$$

In Eq. (1),  $N(\mathbf{r}, \omega)$  does not depend on  $\mathbf{r}$ , while in Eq. (5) it depends on  $z$ . Equation (7) is a pivotal point of our paper since it accounts for the evolution of the spectrum on propagation by relating it to the local density of electromagnetic modes  $N(\mathbf{r}, \omega)$ . Note that  $N(\mathbf{r}, \omega)$  includes only relevant modes excited in the material ( $z < 0$ ) and emitted into vacuum ( $z > 0$ ). The correct counting of modes is done automatically in the integral in Eq. (5), which has a typical structure that relates the density of modes to the Green's tensor of the system. The function  $|g_{\alpha\beta}(\mathbf{k}_{\parallel}, \omega | z, z')|^2$  in Eq. (5) represents a mode that is excited in the plane  $z'$  with a 2D wave vector  $\mathbf{k}_{\parallel}$  and polarization  $\beta$  and arrives at the plane  $z$  with the same 2D wave vector (due to translational invariance in  $x$  and  $y$ ) and polarization  $\alpha$ . The sum over  $\beta$  and the integrals over  $\mathbf{k}_{\parallel}$  and  $z' < 0$  take into account all possible modes that are initially excited.

The origin of a sharp peak seen in the near-field emission spectrum (Fig. 1c) becomes clear when we analyze the dispersion relation (6) for SW (Fig. 2). Near the frequency  $\omega_{\text{max}}$  defined by the condition  $\varepsilon'(\omega_{\text{max}}) = -1$ , there exists a large number of surface modes with different wave numbers but with frequencies that are very close to each other. Therefore, the density of surface modes will necessarily display a strong peak at  $\omega = \omega_{\text{max}}$ . However, since SW decay exponentially away from the surface, this peak is not seen in the far zone (Fig. 1a).

To achieve a detailed understanding of the  $z$  dependence of the emission spectrum, we calculate an approximate expression for the density of modes  $N(z, \omega)$  from Eqs. (5) and (7) for small distances  $z$  and in the limit of large  $k_{\parallel}$

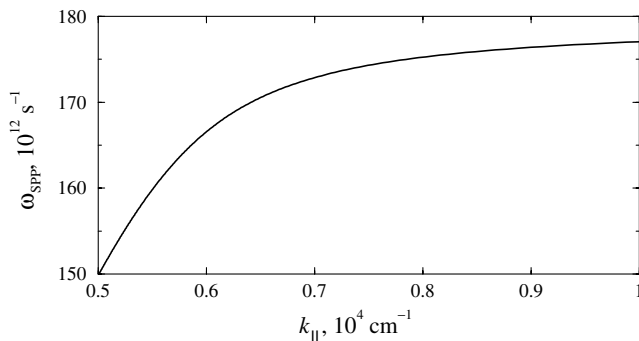


FIG. 2. Dispersion curve for SW on the vacuum-SiC planar interface.  $\text{Re}(\omega_{\text{SPP}})$  is calculated for a given real  $k_{\parallel}$ .

(according to Fig. 2, the modes with large  $k_{\parallel}$  define the behavior near the peak):

$$N(z, \omega) = \frac{\varepsilon''(\omega)}{|1 + \varepsilon(\omega)|^2} \frac{1}{16\pi^2 \omega z^3}. \quad (8)$$

This  $1/z^3$  contribution can be recognized as a well-known quasistatic behavior exhibited near the surfaces of all materials [12,14]. Note that if  $\varepsilon''$  is not very large at  $\omega = \omega_{\text{max}}$  then the density of modes exhibits a resonance at that frequency. This is the origin of the peak in the near-field spectrum of SiC at  $\omega = 178.7 \times 10^{12} \text{ s}^{-1}$ . The presence of a resonance in the density of modes  $N(z, \omega)$  is, however, not required for observing spectral changes caused by the loss of evanescent modes. Indeed, in the short distance regime, the spectrum is given by Eq. (8), whereas, in the far field, the spectrum is given by Eq. (1) multiplied by the emissivity of the surface. Thus, even in the absence of resonant SW, the near-field spectrum is different from the far-field spectrum, but the changes are less dramatic.

The result (8) is valid only in the limit of distances much smaller than the wavelength. We show in Fig. 3 the variation of the spectral density  $I(\omega, z)$  with the distance  $z$  from the surface. We consider two different frequencies none of which is very close to the resonance at  $\omega_{\text{max}}$ . In agreement with Eq. (8), we observe that the spectral density increases sharply for  $z < 1 \mu\text{m}$ , i.e., when the distance to the surface is much smaller than the wavelength. However, the decay behavior of the two curves for larger values of  $z$  exhibits an essential difference. The exponential decay seen in the solid curve for the values of  $z$  between 1 and  $5 \mu\text{m}$  is a signature of the presence of a SW whose energy decays exponentially with  $z$ . This SW is a surface phonon polariton that exists at  $\omega = 166 \times 10^{12} \text{ s}^{-1}$  (solid curve) but not at  $\omega = 210 \times 10^{12} \text{ s}^{-1}$  (dashed curve). The different  $z$  dependence of the  $I(\omega, z)$  for different values of  $\omega$  causes the spectrum to change on propagation of emitted radiation from the surface to the far zone.

This analysis allows us to conclude that the spectral changes in thermal emission should be observable in a wide variety of solid-state systems supporting evanescent

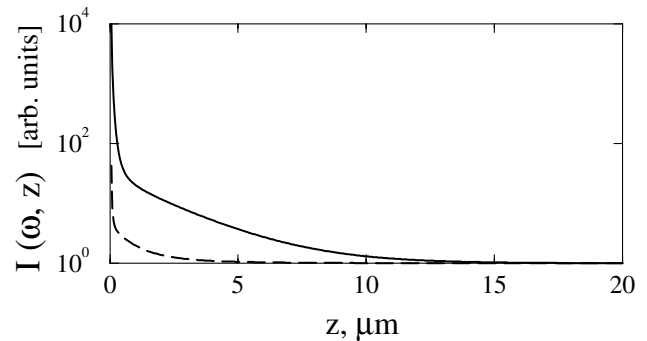


FIG. 3. The variation of the spectral density of the thermal emission for SiC at  $T = 300 \text{ K}$  with the distance  $z$  from observer to the surface. Solid line:  $\omega = 166 \times 10^{12} \text{ s}^{-1}$  ( $\lambda = 11.4 \mu\text{m}$ ), dashed line:  $\omega = 210 \times 10^{12} \text{ s}^{-1}$  ( $\lambda = 9.0 \mu\text{m}$ ).

surface waves or guided waves (in layered structures [5]). Resonant features in the near-field spectra (such as in Fig. 1c) correspond to resonances in the local density of surface modes  $N(z, \omega)$  and appear when the dispersion curve for SW has a flat portion (as in Fig. 2). In addition to our example of SiC, such resonances  $N(z, \omega)$  are displayed by metals (supporting surface plasmon polaritons), semiconductors (supporting surface exciton polaritons  $\Phi$ ), and several other materials. Yet, a peak in the emission intensity  $I(z, \omega)$  will be observable only if  $\theta(\omega, T)$  is not too small. For example, the near-field spectrum of thermal emission from amorphous glass near-field spectrum has a sharp peak for  $\omega = 9.24 \times 10^{13} \text{ s}^{-1}$  ( $\lambda = 20.4 \text{ }\mu\text{m}$ ) visible at room temperature. All of the III-V and II-VI semiconductors can support surface waves in the midinfrared. However, although the number of modes (8) has a resonance in the case of silver at about  $\omega = 5.57 \times 10^{15} \text{ s}^{-1}$  ( $\lambda = 0.339 \text{ }\mu\text{m}$ ), no sharp peak is seen if the temperature silver sample is lower than 4000 K.

Equations (7) and (8) also suggest a new application for near-field spectroscopy. As a near-field spectrum at a given distance to the interface gives access to  $\varepsilon''(\omega)/|1 + \varepsilon(\omega)|^2$ , one can hope to retrieve the material dielectric constant, similar to the method usually used to obtain  $\varepsilon$  from reflectivity measurements [15]. With the rapid development of near-field optical microscopy, such near-field spectra can be measured. This could open the way to a new technique of local solid-state spectroscopy. Finally, we anticipate that the effects reported in this paper should significantly improve our understanding of the radiative heat transfer at nanometric scale with particular applications in the field of near-field microscopies. This might have applications for high density storage where the local control of temperature is essential in the writing process. Also, note that the effect of near-field thermal fluctuations was measured recently using the induced brownian motion on an atomic force microscope tip [16].

To summarize, we have demonstrated that the spectrum of thermal emission can undergo significant, qualitative changes on propagation due to the loss of evanescent modes. Such novel spectral changes are caused by the change in the local density of emitted electromagnetic modes, and are especially pronounced in the systems supporting surface waves.

We thank Professor Emil Wolf and Dr. Carsten Henkel for useful discussions. The work of A. V. S. was supported

by the Engineering Research Program of the Office of Basic Energy Science at the Department of Energy under Grant No. DE-FG02-90ER 14119 and by the Air Force Office of Scientific Research under Grant No. F49620-00-1-0125.

---

\*Present address: KLA-Tencor Corporation, 160 Rio Robles, San Jose, CA 95134.

Email address: Andrei.Shchegrov@KLA-Tencor.com

- [1] *Near Field Optics*, edited by D. W. Pohl and D. Courjon (Kluwer, Dordrecht, 1993); J.-J. Greffet and R. Carminati, *Prog. Surf. Sci.* **56**, 133 (1997).
- [2] Q. Wu, R. D. Grober, D. Gammon, and D. S. Katzer, *Phys. Rev. Lett.* **83**, 2652 (1999).
- [3] E. Wolf, *Phys. Rev. Lett.* **56**, 1370 (1986); *Nature (London)* **326**, 363 (1987).
- [4] E. Wolf and D. F. V. James, *Rep. Prog. Phys.* **59**, 771 (1996).
- [5] T. A. Leskova, A. A. Maradudin, A. V. Shchegrov, and E. R. Méndez, *Phys. Rev. Lett.* **79**, 1010 (1997); A. V. Shchegrov and A. A. Maradudin, *Phys. Rev. B* **59**, 7732 (1999).
- [6] H. R. Stuart and D. G. Hall, *Phys. Rev. Lett.* **80**, 5663 (1998).
- [7] A. R. McGurn, A. A. Maradudin, and V. Celli, *Phys. Rev. B* **31**, 4866 (1985); S. I. Bozhevolnyi, I. I. Smolyaninov, and A. V. Zayats, *Phys. Rev. B* **51**, 17916 (1995).
- [8] *Surface-Enhanced Raman Scattering*, edited by R. K. Chang and T. E. Furtak (Plenum, New York, 1982).
- [9] T. W. Ebbesen, H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, *Nature (London)* **391**, 667 (1998).
- [10] E. A. Vinogradov, G. N. Zhizhin, and V. I. Yudson, in *Surface Polaritons*, edited by V. M. Agranovich and D. L. Mills (North-Holland, Amsterdam, 1982), p. 145.
- [11] P. J. Hesketh, J. N. Zemel, and B. Gebhart, *Nature (London)* **324**, 549 (1986); J. Le Gall, M. Olivier, and J.-J. Greffet, *Phys. Rev. B* **55**, 10105 (1997).
- [12] R. Carminati and J.-J. Greffet, *Phys. Rev. Lett.* **82**, 1660 (1999).
- [13] L. D. Landau, E. M. Lifshitz, and L. P. Pitaevskii, *Statistical Physics, Part I* (Pergamon, Oxford, 1980), p. 183.
- [14] S. M. Rytov, Yu. A. Kravtsov, and V. I. Tatarskii, *Principles of Statistical Radiophysics 3* (Springer-Verlag, Berlin, 1989), p. 118.
- [15] E. W. Palik, *Handbook of Optical Constants of Solids* (Academic Press, San Diego, 1985).
- [16] I. Dorofeyev, H. Fuchs, G. Wenning, and B. Gotsmann, *Phys. Rev. Lett.* **83**, 2402 (1999).



because the actual masses of massive protostars are poorly determined. Our approach is to predict the properties of some well studied massive protostars in terms of their bolometric luminosities. The bolometric luminosity  $L_{\text{bol}}$  has contributions from main-sequence nuclear burning  $L_{\text{ms}}$ , deuterium burning  $L_{\text{D}}$ , and accretion  $L_{\text{acc}}$ . The accretion luminosity  $L_{\text{acc}} = f_{\text{acc}} G m_* \dot{m}_* / r_*$ , where  $f_{\text{acc}}$  is a factor of order unity accounting for energy radiated by an accretion disk, advected into the star or converted into kinetic energy of outflows, and where the stellar radius  $r_*$  may depend sensitively on the accretion rate  $\dot{m}_*$ . Massive stars join the main sequence during their accretion phase at a mass that also depends on the accretion rate<sup>23</sup>. To treat accelerating accretion rates, we have developed a simple model for protostellar evolution based on that of refs 6 and 24. The model accounts for the total energy of the protostar as it accretes and dissociates matter and, if the central temperature  $T_c \geq 10^6$  K, burns deuterium. We have modified this model to include additional processes, such as deuterium shell burning, and we have calibrated these modifications against the more detailed calculations of refs 23 and 25.

Our model allows us to make predictions for the masses and accretion rates of embedded protostars that are thought to power hot molecular cores (C.F.M. and J.C.T., manuscript in preparation). Figure 2 compares our theoretical tracks with the observed bolometric luminosities of several sources. We find that uncertainties in the value of the pressure create only small uncertainties in  $m_*$  for  $L_{\text{bol}}$  in excess of a few times  $10^4$  solar luminosities.

The infrared and submillimetre spectra of accreting protostars and their surrounding envelopes have been modelled in ref. 5, modelling the same sources shown in Fig. 2. We note that uncertainties in the structure of the gas envelope and the possible contributions from additional surrounding gas cores or diffuse gas will affect the observed spectrum. Comparing results, our inferred stellar masses are similar, but our accretion rates are systematically smaller by factors of  $\sim 2-5$ . The modelled<sup>5</sup> high accretion rates of  $\sim 10^{-3} M_{\odot} \text{yr}^{-1}$  for stars with  $m_* \approx 10 M_{\odot}$  would be difficult to achieve unless the pressure was increased substantially; for example, if the stars are destined to reach  $m_{\text{sf}} \approx 30 M_{\odot}$ , pressure increases of a factor  $\sim 40$  are required.  $\square$

Received 17 September 2001; accepted 2 January 2002.

1. Bernasconi, P. A. & Maeder, A. About the absence of a proper zero age main sequence for massive stars. *Astron. Astrophys.* **307**, 829–839 (1996).
2. McLaughlin, D. E. & Pudritz, R. E. Gravitational collapse and star formation in logotropic and nonisothermal spheres. *Astrophys. J.* **476**, 750–765 (1997).
3. Stahler, S. W., Palla, F. & Ho, P. T. P. in *Protostars & Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S. S.) 327–351 (Univ. Arizona Press, Tucson, 2000).
4. Behrend, R. & Maeder, A. Formation of massive stars by growing accretion rate. *Astron. Astrophys.* **373**, 190–198 (2001).
5. Osorio, M., Lizano, S. & D'Alessio, P. Hot molecular cores and the formation of massive stars. *Astrophys. J.* **525**, 808–820 (1999).
6. Nakano, T., Hasegawa, T., Morino, J.-I. & Yamashita, T. Evolution of protostars accreting mass at very high rates: Is Orion IRC2 a huge protostar? *Astrophys. J.* **534**, 976–983 (2000).
7. Plume, R., Jaffe, D. T., Evans, N. J. II, Martin-Pintado, J. & Gomez-Gonzalez, J. Dense gas and star formation: Characteristics of cloud cores associated with water masers. *Astrophys. J.* **476**, 730–749 (1997).
8. Bertoldi, F. & McKee, C. F. Pressure-confined clumps in magnetized molecular clouds. *Astrophys. J.* **395**, 140–157 (1992).
9. Hillenbrand, L. A. & Hartmann, L. W. A preliminary study of the Orion nebula cluster structure and dynamics. *Astrophys. J.* **492**, 540–553 (1998).
10. van der Tak, F. F. S., van Dishoeck, E. F., Evans, N. J. II & Blake, G. A. Structure and evolution of the envelopes of deeply embedded massive young stars. *Astrophys. J.* **537**, 283–303 (2000).
11. Vázquez-Semadeni, E., Ostriker, E. C., Passot, T., Gammie, C. F. & Stone, J. M. in *Protostars & Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S. S.) 3–28 (Univ. Arizona Press, Tucson, 2000).
12. Richer, J. S., Shepherd, D. S., Cabrit, S., Bachiller, R. & Churchwell, E. in *Protostars & Planets IV* (eds Mannings, V., Boss, A. P. & Russell, S. S.) 867–894 (Univ. Arizona Press, Tucson, 2000).
13. Matzner, C. D. & McKee, C. F. Efficiencies of low-mass star and star cluster formation. *Astrophys. J.* **545**, 364–378 (2000).
14. McLaughlin, D. E. & Pudritz, R. E. A model for the internal structure of molecular cloud cores. *Astrophys. J.* **469**, 194–208 (1996).
15. Larson, R. B. Turbulence and star formation in molecular clouds. *Mon. Not. R. Astron. Soc.* **194**, 809–826 (1981).
16. Stahler, S. W., Shu, F. H. & Taam, R. E. The evolution of protostars. I—Global formulation and results. *Astrophys. J.* **241**, 637–654 (1980).
17. Crutcher, R. M. Magnetic fields in molecular clouds: Observations confront theory. *Astrophys. J.* **520**, 706–713 (1999).

18. Jijina, J. & Adams, F. C. Infall collapse solutions in the inner limit: Radiation pressure and its effects on star formation. *Astrophys. J.* **462**, 874–887 (1996).
19. Shu, F. H. Self-similar collapse of isothermal spheres and star formation. *Astrophys. J.* **214**, 488–497 (1977).
20. Palla, F. & Stahler, S. W. Star formation in the Orion nebula cluster. *Astrophys. J.* **525**, 772–783 (1999).
21. Wolfire, M. G. & Cassinelli, J. Conditions for the formation of massive stars. *Astrophys. J.* **319**, 850–867 (1987).
22. Bonnell, I. A., Bate, M. R., Clarke, C. J. & Pringle, J. E. Competitive accretion in embedded stellar clusters. *Mon. Not. R. Astron. Soc.* **323**, 785–794 (2001).
23. Palla, F. & Stahler, S. W. The evolution of intermediate-mass protostars. II—Influence of the accretion flow. *Astrophys. J.* **392**, 667–677 (1992).
24. Nakano, T., Hasegawa, T. & Norman, C. The mass of a star formed in a cloud core: Theory and its application to the Orion A cloud. *Astrophys. J.* **450**, 183–195 (1995).
25. Palla, F. & Stahler, S. W. The evolution of intermediate-mass protostars. I—Basic results. *Astrophys. J.* **375**, 288–299 (1991).
26. Schaller, G., Schaerer, D., Meynet, G. & Maeder, A. New grids of stellar models from 0.8 to 120 solar masses at  $Z = 0.020$  and  $Z = 0.001$ . *Astron. Astrophys. Suppl.* **96**, 269–331 (1992).
27. Hunter, T. R. et al. G34.24+0.13MM: A deeply embedded proto-B-star. *Astrophys. J.* **493**, L97–L100 (1998).
28. Molinari, S., Testi, L., Brand, J., Cesaroni, R. & Palla, F. IRAS 23385+6053: A prototype massive class 0 object. *Astrophys. J.* **505**, L39–L42 (1998).
29. Kaufman, M. J., Hollenbach, D. J. & Tielens, A. G. G. M. High-temperature molecular cores near massive stars and application to the Orion hot core. *Astrophys. J.* **497**, 276–287 (1998).
30. Wyrowski, F., Schilke, P., Walmsley, C. M. & Menten, K. M. Hot gas and dust in a protostellar cluster near W3(OH). *Astrophys. J.* **514**, L43–L46 (1999).

#### Acknowledgements

We thank S. Stahler, R. Pudritz, M. Walmsley and M. Krumholz for discussions. This work was supported by the NSF, by NASA (which supports the Center for Star Formation Studies) and (for J.C.T.) by a Spitzer-Cotsen fellowship from Princeton University.

Correspondence and requests for materials should be addressed to C.F.M. (e-mail: cmckee@mckee.berkeley.edu).

## Coherent emission of light by thermal sources

Jean-Jacques Greffet\*<sup>†</sup>, Rémi Carminati\*, Karl Joulain\*, Jean-Philippe Mulet\*, Stéphane Mainguy<sup>†</sup> & Yong Chen<sup>‡</sup>

\* Laboratoire EM2C, CNRS, Grande Voie des Vignes, Châtenay-Malabry 92295 Cedex, France

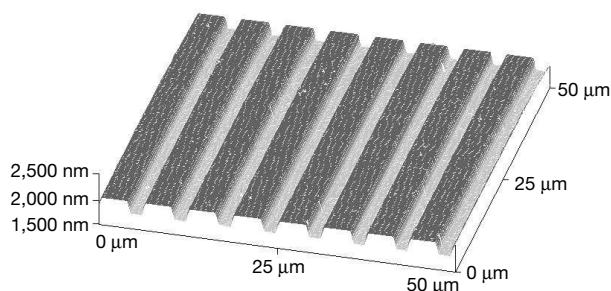
† CEA CESTA, Le Barp, 33114, France

‡ Laboratoire de Microstructures et de Microélectronique, CNRS, av. H. Ravera, 92220 Bagneux, France

✉ Present address: The Institute of Optics, University of Rochester, New York 14627, USA

A thermal light-emitting source, such as a black body or the incandescent filament of a light bulb, is often presented as a typical example of an incoherent source and is in marked contrast to a laser. Whereas a laser is highly monochromatic and very directional, a thermal source has a broad spectrum and is usually quasi-isotropic. However, as is the case with many systems, different behaviour can be expected on a microscopic scale. It has been shown recently<sup>1,2</sup> that the field emitted by a thermal source made of a polar material is enhanced by more than four orders of magnitude and is partially coherent at a distance of the order of 10 to 100 nm. Here we demonstrate that by introducing a periodic microstructure into such a polar material (SiC) a thermal infrared source can be fabricated that is coherent over large distances (many wavelengths) and radiates in well defined directions. Narrow angular emission lobes similar to antenna lobes are observed and the emission spectra of the source depends on the observation angle—the so-called Wolf effect<sup>3,4</sup>. The origin of the coherent emission lies in the diffraction of surface-phonon polaritons by the grating.

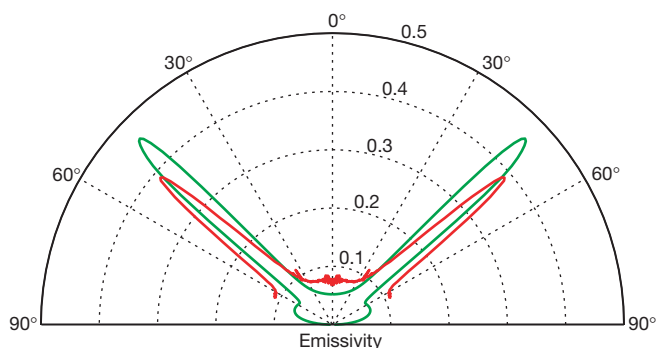
It is usually taken for granted that light spontaneously emitted by



**Figure 1** Image of the grating obtained by atomic force microscopy. Its period  $d = 0.55\lambda$  ( $\lambda = 11.36 \mu\text{m}$ ) was chosen so that a surface wave propagating along the interface could be coupled to a propagating wave in the range of frequencies of interest. The depth  $h = \lambda/40$  was optimized so that the peak emissivity is 1 at  $\lambda = 11.36 \mu\text{m}$ . It was fabricated on SiC by standard optical lithography and reactive ion-etching techniques.

different points of a thermal source cannot interfere. In contrast, different points of an antenna emit waves that interfere constructively in particular directions producing well defined angular lobes. The intensity emitted by a thermal source is the sum of the intensities emitted by different points so that it cannot be directional. However, it has been shown<sup>1,2</sup> recently that some planar sources may have a spectral coherence length in the plane much larger than a wavelength and can be quasi-monochromatic in the near-field. This paves the way for the construction of a thermal source that could radiate light within narrow angular lobes as an antenna instead of having the usual quasi-lambertian angular behaviour.

Here we report experimental measurements demonstrating that it is possible to build an infrared antenna by properly designing a periodic microstructure on a polar material. Such an antenna radiates infrared light in a narrow solid angle when it is heated. Another unusual property of this source is that its emission spectrum depends on the observation direction. This property was first predicted by Wolf as a consequence of spatial correlations

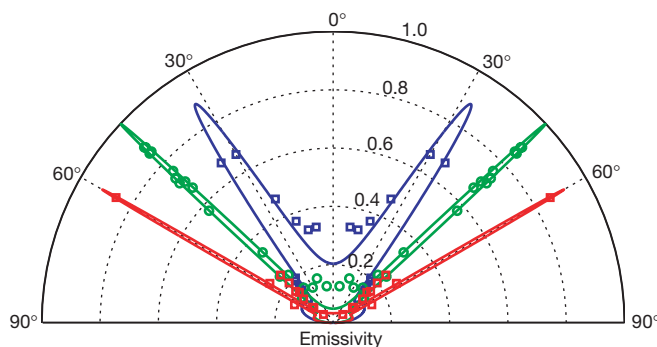


**Figure 2** Polar plot of the emissivity of the grating depicted in Fig. 1 at  $\lambda = 11.36 \mu\text{m}$  and for  $p$ -polarization. Red, experimental data; green, theoretical calculation. The measurements were taken by detecting the intensity emitted by the sample in the far field as a function of the emission angle. A HgCdTe detector placed at the focal length of a ZnSe lens was used. The sample was mounted on a rotation stage. The theoretical result was obtained by computing the reflectivity of the sample and using Kirchhoff's law ( $\epsilon = \alpha = 1 - R$ ). To fit the data, we took into account the spectral resolution ( $0.22 \mu\text{m}$ ) and the angular resolution ( $3^\circ$ ) of the measurements. The disagreement is due to the fact that for the calculation, the index at room temperature is used whereas emission data were taken with a sample in a local thermal equilibrium situation at a temperature of 773 K. Comparison between the two curves illustrates the validity of Kirchhoff's law for polarized monochromatic directional quantities. The surrounding medium was at 300 K and the background signal was subtracted. The emissivity for  $s$ -polarization (not shown) does not show any peak and is very close to its value for a flat surface. Note that most of the emitted light is emitted in the narrow lobe (that is, coherently).

for random sources<sup>3,4</sup>. This effect has been demonstrated experimentally for artificial secondary sources<sup>4,5</sup> but has never been observed for direct thermal emission. In addition, the emissivity of the source is enhanced by a factor of 20 compared to the emissivity of a flat surface.

Using theories developed recently<sup>6</sup> to interpret the emission data by gratings, we have designed and optimized a periodic surface profile that produces a strong peak of the emissivity around a wavelength  $\lambda = 11.36 \mu\text{m}$ . The grating (see Fig. 1) has been ruled on a SiC substrate. A similar grating of doped silicon with very deep depths has been investigated<sup>7</sup>. The authors attributed the particular properties observed to organ pipe modes in the microstructure<sup>7</sup>. However, the role of coherence induced by surface waves and the exact mechanism were not understood at that time<sup>2,6</sup>.

The measurements of the thermal emission in a plane perpendicular to the lines of the grating are shown in Fig. 2. Emission is highly directional and looks very similar to the angular pattern of an antenna. We have also plotted (see Fig. 2) the calculated emission pattern. The qualitative discussion of the introduction suggests that the small angular width of the emission pattern is a signature of the local spatial coherence of the source. A proof of this stems from the fact that the source has a width  $L = 5 \text{ mm}$  and a spectral coherence length  $l \ll L$  and that its temperature is uniform. Hence, we can assume that the source is a quasi-homogeneous source<sup>8</sup>. With this assumption, it is known that the radiant intensity and the spectral degree of coherence in the plane of the source are related by a Fourier transform relationship<sup>8</sup>. Therefore, the angular width  $\theta$  of the lobe emission varies qualitatively with  $\lambda/l$  for this locally coherent source instead of with  $\lambda/L$ , as for a globally coherent source. Thus a small angular aperture of the far-field radiation is the signature of a spectral coherence length in the source much larger than the wavelength. To overcome the experimental resolution limit of our direct emissivity measurement, we measured the reflectivity  $R$ . From Kirchhoff's law<sup>9</sup>, we know that the polarized directional spectral emissivity  $\epsilon$  is given by  $\epsilon = \alpha = 1 - R$  where  $\alpha$  is the absorptivity and  $R$  is the reflectivity of the grating. Results are plotted in Fig. 3. There is a remarkable quantitative agreement between the data taken at room temperature and theoretical calculations. We note that the peak at  $60^\circ$  has an angular width  $\theta$  as narrow as  $1^\circ$  so that the corresponding spectral coherence length is as large as  $\lambda/\theta \approx 60\lambda \approx 0.6 \text{ mm}$ . This suggests that a thermal source with a size  $L$  on the order of the spectral coherence length  $l$ , namely a globally coherent thermal source, could be achieved.

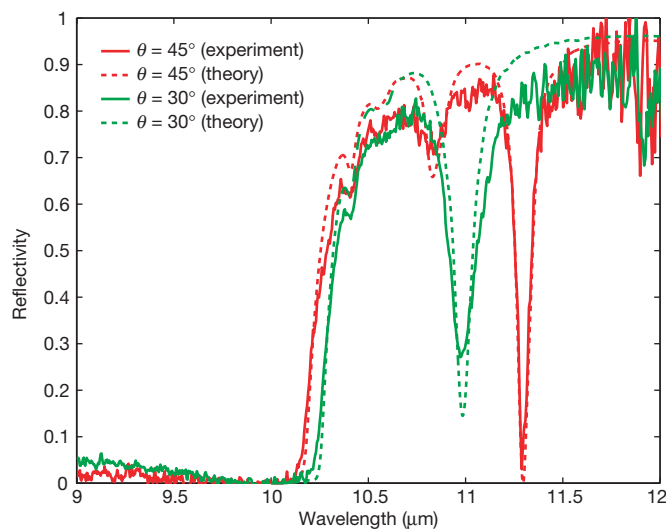


**Figure 3** Emissivity of a SiC grating in  $p$ -polarization. Blue,  $\lambda = 11.04 \mu\text{m}$ ; red,  $\lambda = 11.36 \mu\text{m}$ ; green,  $\lambda = 11.86 \mu\text{m}$ . The emissivity was deduced from measurements of the specular reflectivity  $R$  using Kirchhoff's law. The data have been taken at ambient temperature using a Fourier transform infrared (FTIR) spectrometer as a source and a detector mounted on a rotating arm. The angular acceptance of the spectrometer was reduced to a value lower than the angular width of the dip. The experimental data are indicated by circles; the lines show the theoretical results. An excellent agreement is obtained when the data are taken at ambient temperature, which supports our interpretation of the slight disagreement in Fig. 1 being due to the variation of index with temperature.

A surprising property of the emissivity patterns of Fig. 3 is that they depend strongly on the wavelength. This suggests that the emission spectra depend on the emission direction. This would be a manifestation of the Wolf effect<sup>3,4</sup>. To observe this effect, we have taken several spectra of the reflectivity of the surface at different angles. Fig. 4 shows experimental and theoretical spectra for different observation angles. The position of the peaks of emissivity (dips of reflectivity) depends strongly on the observation angle. It is important to emphasize that this property is not merely a scattering effect but is a consequence of the partial spatial coherence of the source. The value of the reflectivity is also remarkable. By ruling a grating onto a material which is essentially a mirror, we were able to produce a perfect absorber. This behaviour has already been observed for metallic gratings and attributed to the resonant excitation of surface plasmons. This is the first time, to our knowledge, that total absorption in the infrared owing to excitation of surface-phonon polaritons has been reported.

In order to prove experimentally the role of the surface wave, we have done spectral measurements of the emissivity for *s*- and *p*-polarization. The peaks are never observed for *s*-polarization nor for *p*-polarization in the spectral region where surface waves cannot exist. In order to characterize quantitatively the role of the surface waves, we have obtained the dispersion relation from reflectivity measurements<sup>6</sup>. The results are displayed in Fig. 5 and compared with theory. We note that the interaction of the surface wave with the grating produces the aperture of a gap close to the band edge. Figure 5 shows that our experiment allows us to directly see surface-phonon polaritons. It also yields additional insight into the Wolf-effect<sup>3,4</sup> mechanism. Emission of infrared light has already been used to study surface excitations, but using prisms to couple the surface waves to propagating modes<sup>10</sup>.

We now discuss the physical origin of coherent thermal emission. We wish to understand how random thermal motion can generate a coherent current along the interface. The key lies in the coherent properties of surface waves (either surface-plasmon polaritons or surface-phonon polaritons) demonstrated in refs 1 and 2. Both are mechanical delocalized collective excitations involving charges. Surface-phonon polaritons are phonons in a polar material, whereas surface-plasmon polaritons are acoustical-type waves in an electron gas. In both cases, these waves have the following two



**Figure 4** Comparison between measured and calculated spectral reflectivities of a SiC grating at room temperature. The incident light is *p*-polarized. The dip observed at 45° and  $\lambda = 11.36 \mu\text{m}$  coincides with the emission peak observed in Figs 2 and 3. The figure shows clearly that the reflectivity spectra depend on the observation angle. Using Kirchhoff's law, it follows that the emission spectra depend on the observation angle. This is a manifestation of the Wolf effect<sup>3,4</sup>.

properties: (1) they are mechanical modes of the system that can be resonantly excited; (2) they are charge-density waves, that is, they generate electromagnetic fields. Because these excitations are delocalized, so are the corresponding electromagnetic fields.

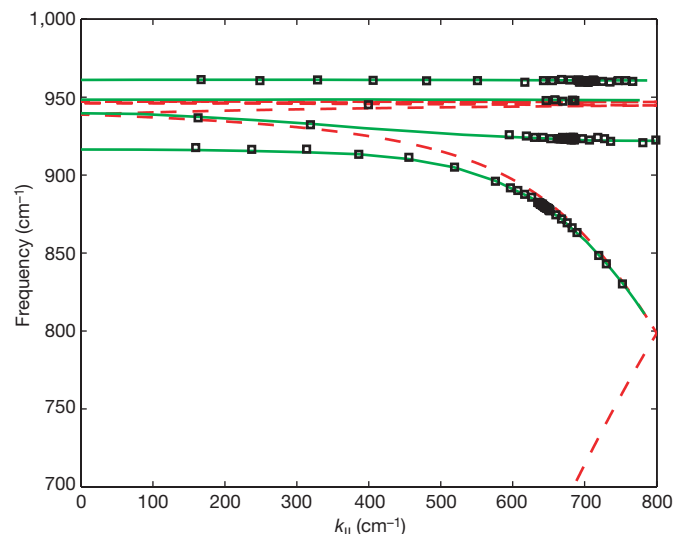
From a classical point of view, each volume element of the thermal source can be modelled by a random point-like source that excites an extended mode: the surface wave. This is similar to some extent to the emission of sound by a string of a piano. The source is a hammer that strikes the string at a particular point. Then the modes of the string are excited producing a vibration along the full length of the string. At that point of the analogy, as anyone can hear the vibrations of a piano string, we may wonder why the coherent electromagnetic surface waves are not usually observed. The reason is that surface modes have a wavevector larger than  $2\pi/\lambda$  so that they are evanescent. Their effect is not seen in the far-field.

However, by ruling a grating on the interface, we are able to couple these surface modes to propagating modes. The relationship between the emission angle  $\theta$  and the wavelength  $\lambda$  is simply given by the usual grating law

$$\frac{2\pi}{\lambda} \sin\theta = k_{\parallel} + p \frac{2\pi}{d}$$

where  $p$  is an integer and  $k_{\parallel}$  is the wavevector of the surface wave. Thus, by modifying the characteristics of the surface profile, it is possible to modify the direction and the value of the emissivity of the surface at a given wavelength. It is also possible to modify the emission spectrum in a given direction. Such gratings can be used to design infrared sources with specific properties.

This may also have interesting applications such as modifying the radiative heat transfer for a given material. Indeed, we have demonstrated that a reflectivity of 94% can be reduced to almost zero in the infrared for SiC. This could also be done for glass, which



**Figure 5** Dispersion relation wavevector,  $\omega(k_{\parallel})$ , of surface-phonon polaritons. Data points, experimental dispersion relation. Solid green curve, theoretical dispersion relation for the grating. Dotted red curve, theoretical dispersion relation for the flat surface. This figure explains the mechanism of the Wolf effect<sup>3,4</sup> for this particular source. The spatial coherence in the plane of the source is due to the presence of a surface wave. For a fixed frequency  $\omega$ , it can be seen that there is only one possible wavevector  $k_{\parallel}(\omega)$ . Thus the spectral degree of coherence at  $\omega$  oscillates<sup>2</sup> with a particular wavelength  $2\pi/k_{\parallel}(\omega)$ . When observing in the far field at an angle  $\theta$  such that  $ck_{\parallel}(\omega)/\omega = \sin\theta$  there is a strong contribution of the surface wave at frequency  $\omega$ . By varying the observation angle, the frequency varies according to the dispersion relation of the surface wave. It is seen that the strong Wolf effect produced by this source is due to (1) the thermal excitation of surface waves which produce the spatial coherence and (2) the propagation in vacuum which selects one particular wavevector.



is a polar material that has a large reflectivity in the infrared owing to the presence of resonances. This would allow us to increase the radiative cooling of the material if the emission is enhanced in a region where absorption is low, because the atmosphere does not emit. Another promising application of our results is the possibility of modifying the heat transfer in the near-field. Materials that are separated by distances smaller than the typical wavelength exchange radiative energy through evanescent waves. When surface waves are resonantly excited, they provide the leading contribution<sup>11</sup>. Thus, the heat transfer is almost monochromatic. This may be used to enhance the efficiency of infrared photovoltaic cells<sup>12</sup>. □

Received 25 June; accepted 20 December 2001.

1. Shchegrov, A., Joulain, K., Carminati, R. & Greffet, J. J. Near-field spectral effects due to electromagnetic surface excitations. *Phys. Rev. Lett.* **85**, 1548–1551 (2000).
2. Carminati, R. & Greffet, J. J. Near-field effects in spatial coherence of thermal sources. *Phys. Rev. Lett.* **82**, 1660–1663 (1999).
3. Wolf, E. Non-cosmological red-shifts of spectral lines. *Nature* **326**, 363–365 (1987).
4. Wolf, E. & James, D. F. Correlation-induced spectral changes. *Rep. Prog. Phys.* **59**, 771–818 (1996).
5. Morris, G. M. & Faklis, D. Effects of source correlation on the spectrum of light. *Opt. Commun.* **62**, 5–11 (1987).
6. Le Gall, J., Olivier, M. & Greffet, J. J. Experimental and theoretical study of reflection and coherent thermal emission by a SiC grating supporting a surface photon polariton. *Phys. Rev. B* **55**, 10105–10114 (1997).
7. Hesketh, P. J., Zemel, J. N. & Gebhart, B. Organ pipe radiant modes of periodic micromachined silicon surfaces. *Nature* **325**, 549–551 (1986).
8. Mandel, L. & Wolf, E. *Optical Coherence and Quantum Optics*. Sec. 5.3 (Cambridge Univ. Press, New York, 1995).
9. Greffet, J. J. & Nieto-Vesperinas, M. Field theory for the generalized bidirectional reflectivity: derivation of Helmholtz's reciprocity principle and Kirchhoff's law. *J. Opt. Soc. Am. A* **10**, 2735–2744 (1998).
10. Zhizhin, G. N., Vinogradov, E. A., Moskalova, M. A. & Yakovlev, V. A. Applications of surface polaritons for vibrational spectroscopic studies of thin and very thin films. *Appl. Spectrosc. Rev.* **18**, 171–263 (1982).
11. Mulet, J. P., Joulain, K., Carminati, R. & Greffet, J. J. Nanoscale radiative heat transfer between a small particle and a plane surface. *Appl. Phys. Lett.* **78**, 2931–2933 (2001).
12. Whale, M. D. in *Proc. Conf. on 'Heat Transfer and Transport Phenomena in Microscale'* (ed. Celata, G. P.) 339–346 (Begell House, New York, 2000).

## Competing interests statement

The authors declare that they have no competing financial interests.

Correspondence and requests for materials should be addressed to J.-J.G. (e-mail: greffet@em2c.ecp.fr).

## A general process for *in situ* formation of functional surface layers on ceramics

Toshihiro Ishikawa, Hiroyuki Yamaoka, Yoshikatsu Harada, Teruaki Fujii & Toshio Nagasawa

Ube Research Laboratory, Corporate Research & Development, Ube Industries Ltd, 1978-5 Kogushi, Ube City, Yamaguchi 755-8633, Japan

Ceramics are often prepared with surface layers of different composition from the bulk<sup>1,2</sup>, in order to impart a specific functionality to the surface or to act as a protective layer for the bulk material<sup>3,4</sup>. Here we describe a general process by which functional surface layers with a nanometre-scale compositional gradient can be readily formed during the production of bulk ceramic components. The basis of our approach is to incorporate selected low-molecular-mass additives into either the precursor polymer from which the ceramic forms, or the binder polymer used to prepare bulk components from ceramic powders. Thermal treatment of the resulting bodies leads to controlled phase separation ('bleed out') of the additives, analogous to the normally

undesirable outward loss of low-molecular-mass components from some plastics<sup>5–9</sup>; subsequent calcination stabilizes the compositionally changed surface region, generating a functional surface layer. This approach is applicable to a wide range of materials and morphologies, and should find use in catalysts, composites and environmental barrier coatings.

To avoid the concentration of thermomechanical stress at the interface between the surface layer and the bulk material, many materials have been developed that have gradually varying properties as the distance into the material increases<sup>10</sup>. Such materials can contain gradients in morphology or in composition. Gradients in morphology can, for example, result in materials that have a graded distribution of pore sizes on a monolith of silica aerogel, and a type of integral plastic. These materials have been created by strictly controlling the vaporization of the volatile during the production process<sup>11,12</sup>. Gradients in chemical composition have been achieved, for example: (1) chemical vapour deposition<sup>13,14</sup>, (2) powder methods such as slip cast or dry processing<sup>15</sup>, (3) various coating methods<sup>16</sup>, and (4) thermal chemical reaction<sup>2,17</sup>. Of these, (1) and (4) are relatively expensive, complicated and result in damage to bulk substrates. (2) and (3) produce stepped gradient structures, and it is difficult to control the thickness of each layer to less than 100 nm. Furthermore, most of these processes are not easily adapted to coating samples in the form of fibre bundles, fine powders or other materials with complicated shapes.

We have addressed the issue of establishing an inexpensive and widely applicable process for creating a material with a compositional gradient and excellent functionality. A schematic representative of our new *in situ* formation process for functional surface layers, which have a gradient-like structure towards the surface, is shown in Fig. 1. The important feature of our method is that the surface layer of the ceramic is not deposited on the substrate, but is formed during the production of the bulk ceramic. We confirmed that our process is applicable to any type of system as long as, in the green-body (that is, not-calcined) state, the system contains a resin and a low-molecular-mass additive that can be converted into a functional ceramic at high temperatures. Here, the resin is a type of precursor polymer (polycarbosilane, polycarbosilazane, polysilastyrene, methylchloropolysilane, and so on) or binder polymer used for preparing green bodies from ceramic powders<sup>18</sup>. Although the former case (using precursor polymers) is explained in detail in this Letter, the latter case using binder polymers was also confirmed by treating a Si<sub>3</sub>N<sub>4</sub> body with a TiN surface layer. Si<sub>3</sub>N<sub>4</sub> can exhibit excellent thermal stability and wear resistance in the high-speed machining of cast iron, but shows poor chemical wear resistance in the machining of steel<sup>19</sup>. In order to avoid this problem, TiN coating, by means of expensive chemical vapour deposition, has often been performed on previously prepared Si<sub>3</sub>N<sub>4</sub> substrates. But if our process is appropriately applied, formation of the TiN surface layer could be achieved during the sintering process of the Si<sub>3</sub>N<sub>4</sub> green body. In this case, titanium(IV) butoxide and polystyrene are used as the low-molecular-mass additive and binder polymer, respectively. By a combination of sufficient maturation (in air at 100 °C) and subsequent sintering (in NH<sub>3</sub>+H<sub>2</sub>+N<sub>2</sub> at 1,200 °C), Si<sub>3</sub>N<sub>4</sub> covered with TiN is successfully produced. This technology would be very useful for producing ceramic materials with complicated shapes and various coating layers. Moreover, our process is advantageous for preparing precursor ceramics (particularly fine particles, thin fibrous ceramics and films). The systems to which our concept is applicable are shown in Fig. 1.

Here we give a detailed account of the results for the precursor ceramic obtained using polycarbosilane. Polycarbosilane (–SiH(CH<sub>3</sub>)–CH<sub>2</sub>–)<sub>n</sub> is a representative pre-ceramic polymer for preparing SiC ceramics—for example, Hi-Nicalon fibre<sup>20</sup> and Tyranno SA fibre<sup>21</sup>. Furthermore, oxide or nitride can also be produced from the polycarbosilane by firing in air or ammonia, respectively. Our new technology makes full use of the bleed-out

## Nanoscale radiative heat transfer between a small particle and a plane surface

Jean-Philippe Mulet,<sup>a)</sup> Karl Joulain, Rémi Carminati, and Jean-Jacques Greffet  
*Laboratoire d'Énergétique Moléculaire et Macroscopique, Combustion, Ecole Centrale Paris,  
 Centre National de la Recherche Scientifique, 92295, Châtenay-Malabry Cedex, France*

(Received 20 December 2000; accepted for publication 7 March 2001)

We study the radiative heat transfer between a small dielectric particle, considered as a point-like dipole, and a surface. In the framework of electrodynamics and using the fluctuation-dissipation theorem, we can evaluate the energy exchange in the near field, which is dominated by the contribution of tunneling waves. The transfer is enhanced by several orders of magnitude if the surface or the particle can support resonant surface waves. An application to local heating is discussed. © 2001 American Institute of Physics. [DOI: 10.1063/1.1370118]

Questions about radiative heat transfer (RHT) at nanoscales have been raised by recent developments of nanotechnology.<sup>1</sup> Design of micro- and nanostructures requires a thorough understanding of physical phenomena involved in radiative energy exchange, when their sizes become comparable to the thermal mean free path or the thermal radiation wavelength.<sup>2,3</sup> Modeling RHT between two semi-infinite bodies<sup>4-7</sup> or between a tip and a substrate<sup>8</sup> is a challenging problem for all near-field microscopes (scanning tunneling microscope, atomic force microscope) or for scanning thermal microscopes.<sup>9</sup> Our work deals with the problem of RHT between a small particle—considered as a point-like dipole—and a very close plane interface. This particle could be a single molecule, a dust particle, or a model for the tip of a microscope probe. Using an electromagnetic approach, in the dipolar approximation, we have derived the expression of the radiative heat power exchanged between the particle and the semi-infinite medium. We show that the transfer increases at small distances and can be enhanced by several orders of magnitude (in comparison with the transfer at large distances) if the particle or the bulk support resonant surface waves. Results of numerical simulations are presented and an application to local heating is discussed.

In this part, we focus on the derivation of the radiative power exchanged between a small particle (of spherical shape) and a semi-infinite medium. The geometry of the problem is presented in Fig. 1: the upper medium  $z > 0$  is vacuum ( $\epsilon = 1$ ). A particle ( $P$ ) of radius  $a$  and dielectric constant (frequency dependent)  $\epsilon_P(\omega) = \epsilon'_P(\omega) + i\epsilon''_P(\omega)$  is held at temperature  $T_P$ . The lower medium  $z < 0$ , is filled by a homogeneous, isotropic material (bulk) of dielectric constant  $\epsilon_B(\omega) = \epsilon'_B(\omega) + i\epsilon''_B(\omega)$  and held at temperature  $T_B$ . The center of the particle is at a distance  $d$  above the interface.

We first calculate the mean power  $P_{\text{abs}}^{B \rightarrow P}(\omega)$  radiated by the bulk at a given frequency  $\omega$  and absorbed by the particle. We assume that the bulk is in local thermodynamic equilibrium at uniform temperature  $T_B \neq 0$ , so that there are fluctuating currents inside the bulk due to thermal fluctuations. These currents inside the bulk radiate an electromagnetic

field that illuminates the particle. An elementary fluctuating current  $\mathbf{j}_f(\mathbf{r}', \omega)$  at frequency  $\omega$ , radiates at point  $\mathbf{r}$  inside the particle, an incident electric field  $\mathbf{E}_{\text{inc}}(\mathbf{r}, \omega)$  given by

$$\mathbf{E}_{\text{inc}}(\mathbf{r}, \omega) = (i\omega\mu_0) \int_B \tilde{\mathbf{G}}(\mathbf{r}, \mathbf{r}', \omega) \mathbf{j}_f(\mathbf{r}', \omega) d^3\mathbf{r}', \quad (1)$$

where  $\mu_0$  is the magnetic permeability of vacuum and  $\tilde{\mathbf{G}}$  is the Green tensor<sup>10</sup> of a system constituted by two semi-infinite media whose dielectric constants are either 1 if  $z \geq 0$  or  $\epsilon_B(\omega)$  if  $z < 0$ . Let us now assume that  $\mathbf{E}_{\text{inc}}(\mathbf{r}, \omega)$  is uniform inside the particle. This amounts to use a dipolar approximation, whose validity in the near field has already been discussed.<sup>11</sup> Provided that this condition is satisfied, the small particle behaves as a point-like dipole  $\mathbf{p}_{\text{ind}}(\mathbf{r}_P, \omega)$  induced by the incident field  $\mathbf{E}_{\text{inc}}(\mathbf{r}_P, \omega)$ . These quantities are related by:  $\mathbf{p}_{\text{ind}}(\mathbf{r}_P, \omega) = \epsilon_0 \alpha(\omega) \mathbf{E}_{\text{inc}}(\mathbf{r}_P, \omega)$ , where  $\epsilon_0$  is the dielectric permittivity of vacuum and  $\alpha(\omega)$  is the particle polarizability. For a spherical particle of dielectric constant  $\epsilon_P(\omega)$ , we used the Clausius–Mossotti polarizability<sup>12</sup>

$$\alpha(\omega) = 4\pi a^3 \frac{\epsilon_P(\omega) - 1}{\epsilon_P(\omega) + 2}. \quad (2)$$

A more precise model,<sup>13</sup> taking into account the interaction between the dipole and its image through the interface, leads to the introduction of an effective polarizability. We have verified that, when the distance  $d$  is larger than the particle radius  $a$ , the correction to Eq. (2) is negligible. Since

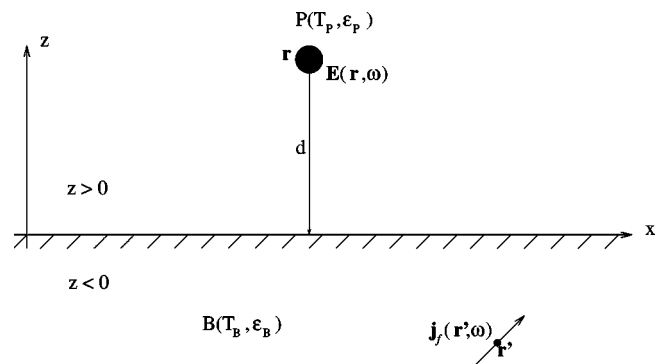


FIG. 1. Geometry of the system.

<sup>a)</sup>Electronic mail: jpmulet@em2c.ecp.fr

we did our calculation for a distance  $d$  larger than  $2a$ , we will consider that Eq. (2) is a good approximation for the polarizability in our problem.

We now evaluate the mean radiative power  $P_{\text{abs}}^{B \rightarrow P}(\omega)$  dissipated inside the particle. Since the scattered power is negligible,<sup>14</sup> this quantity is given by:  $P_{\text{abs}}^{B \rightarrow P}(\omega) = \text{Re}\{-i\omega \mathbf{p}_{\text{ind}}(\mathbf{r}_P, \omega) \cdot \mathbf{E}_{\text{inc}}^*(\mathbf{r}_P, \omega)\}$ , where the brackets denote a statistical ensemble average over the fluctuations of the currents inside the bulk. The components  $\langle |E_{\text{inc},i}(\mathbf{r}_P, \omega)|^2 \rangle$  ( $i=x,y,z$ ) depend only on the distance  $d$  and are given by<sup>10,15,16</sup>

$$\begin{aligned} \langle |E_{\text{inc},i}(\mathbf{r}_P, \omega)|^2 \rangle &= 2 \frac{\mu_0^2 \epsilon_0 \omega^3}{\pi} \text{Im}[\epsilon_B(\omega)] \Theta(\omega, T_B) \\ &\times \sum_{j=x,y,z} \int_B |G_{i,j}(\mathbf{r}_P, \mathbf{r}', \omega)|^2 d^3 \mathbf{r}', \end{aligned} \quad (3)$$

where  $\Theta(\omega, T_B) = \hbar \omega / [\exp(\hbar \omega / k_B T_B) - 1]$  is the mean energy of a quantum oscillator in thermal equilibrium at temperature  $T_B$ ,  $2\pi\hbar$  is Planck's constant, and  $k_B$  is Boltzmann's constant. Finally, we find the expression of the mean power radiated by the bulk and absorbed by the particle at frequency  $\omega$

$$\begin{aligned} P_{\text{abs}}^{B \rightarrow P}(\mathbf{r}_P, \omega) &= \frac{2}{\pi} \frac{\omega^4}{c^4} \text{Im}[\epsilon_B(\omega)] \text{Im}[\alpha(\omega)] \Theta(\omega, T_B) \\ &\times \sum_{i,j=x,y,z} \int_B |G_{i,j}(\mathbf{r}_P, \mathbf{r}', \omega)|^2 d^3 \mathbf{r}'. \end{aligned} \quad (4)$$

We now consider the fluctuating currents inside the particle at temperature  $T_P$  that illuminate the bulk. Using the same formalism, we can calculate the power *locally* dissipated per unit volume, at a point  $\mathbf{r}$  inside the bulk, by the following relation:  $P_{\text{abs}}^{P \rightarrow B}(\mathbf{r}, \omega) = \text{Re}\{\langle \mathbf{j}_{\text{ind}}(\mathbf{r}, \omega) \cdot \mathbf{E}_{\text{inc}}^*(\mathbf{r}, \omega) \rangle\}$ . It reads

$$\begin{aligned} P_{\text{abs}}^{P \rightarrow B}(\mathbf{r}, \omega) &= \frac{2}{\pi} \frac{\omega^4}{c^4} \text{Im}[\epsilon_B(\omega)] \text{Im}[\alpha(\omega)] \Theta(\omega, T_P) \\ &\times \sum_{i,j=x,y,z} |G_{i,j}(\mathbf{r}, \mathbf{r}_P, \omega)|^2. \end{aligned} \quad (5)$$

In this part, we present some numerical results obtained with a particle and a surface of silicon carbide (SiC). The optical properties of this material can be described using an oscillator model<sup>17</sup>

$$\epsilon_B(\omega) = \epsilon_P(\omega) = \epsilon(\omega) = \epsilon_\infty \left( 1 + \frac{\omega_L^2 - \omega_T^2}{\omega_T^2 - \omega^2 - i\Gamma\omega} \right) \quad (6)$$

with  $\epsilon_\infty = 6.7$ ,  $\omega_L = 969 \text{ cm}^{-1}$ ,  $\omega_T = 793 \text{ cm}^{-1}$ , and  $\Gamma = 4.76 \text{ s}^{-1}$ . The bulk can support resonant surface waves, called surface phonon polaritons, that produce a peak in the density of states at frequency  $\omega_B$  where  $\epsilon(\omega_B)$  satisfies the relation  $\epsilon'(\omega_B) = -1$ . The spherical particle supports volume phonon polaritons at  $\omega_P$  where  $\epsilon(\omega_P)$  satisfies the relation  $\epsilon'(\omega_P) = -2$ . Using Eq. (6), we find:  $\omega_B = 178.7 \times 10^{12} \text{ rad s}^{-1}$  and  $\omega_P = 175.6 \times 10^{12} \text{ rad s}^{-1}$ .

In Fig. 2, we plot  $P_{\text{abs}}^{B \rightarrow P}(\omega)$  for a spherical particle of radius  $a = 5 \text{ nm}$  at different distances  $d$  above the surface,

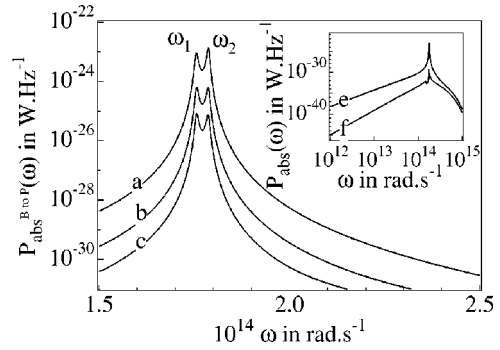


FIG. 2. Mean power radiated by the bulk (at  $T_B = 300 \text{ K}$ ) and absorbed by the particle (of radius  $a = 5 \text{ nm}$ ) vs frequency: (a)  $d = 20 \text{ nm}$ ; (b)  $d = 50 \text{ nm}$ ; (c)  $d = 100 \text{ nm}$ . The inset (log-log scale) shows the spectrum of the absorbed power between  $10^{12}$  and  $10^{15} \text{ rad s}^{-1}$ ; (e)  $d = 20 \text{ nm}$ ; and (f)  $d = 1 \text{ mm}$ .

held at temperature  $T_B = 300 \text{ K}$ . We note that the figure displays two remarkable peaks at frequency  $\omega_1 \approx 175.6 \times 10^{12} \text{ rad s}^{-1}$  and  $\omega_2 \approx 178.7 \times 10^{12} \text{ rad s}^{-1}$ . These two peaks are related to the resonant surface waves: the first corresponds to the resonance of the particle which presents a larger absorption at this frequency; the second is due to a huge increase of electromagnetic energy density close to the surface, demonstrated recently.<sup>16</sup> An asymptotic expansion of Eq. (4) for small distance  $d$  yields the radiative power spectrum

$$\begin{aligned} P_{\text{abs}}^{B \rightarrow P}(d, \omega) &\sim \frac{1}{4\pi^2 d^3} \\ &\times \underbrace{4\pi a^3}_{\text{particle}} \frac{3\epsilon''(\omega)}{|\epsilon(\omega) + 2|^2} \underbrace{\frac{\epsilon''(\omega)}{|\epsilon(\omega) + 1|^2}}_{\text{bulk}} \Theta(\omega, T_B). \end{aligned} \quad (7)$$

At this point, we must emphasize that, whereas those surface waves are evanescent waves, an energy exchange between the bulk and the particle takes place because the particle lies in the region (up to many micrometers) where the evanescent field is large, so that there is an efficient coupling between them. The inset of the figure shows (in log-log scale) the spectrum of the absorbed power between  $10^{12}$  and  $10^{15} \text{ rad s}^{-1}$  at two different distances  $d = 20 \text{ nm}$  and  $d = 1 \text{ mm}$ . It is seen that the RHT is almost monochromatic and is larger in the near field. If the particle and the bulk

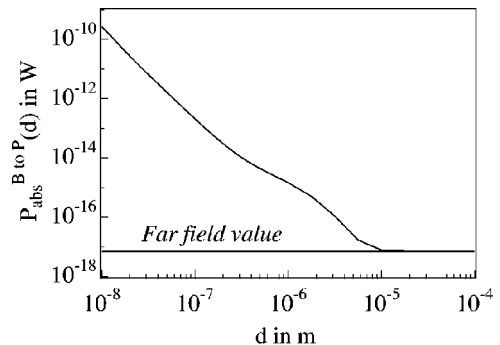


FIG. 3. Total power radiated by the bulk (at  $T_B = 300 \text{ K}$ ) and absorbed by the particle (of radius  $a = 5 \text{ nm}$ ) vs distance.

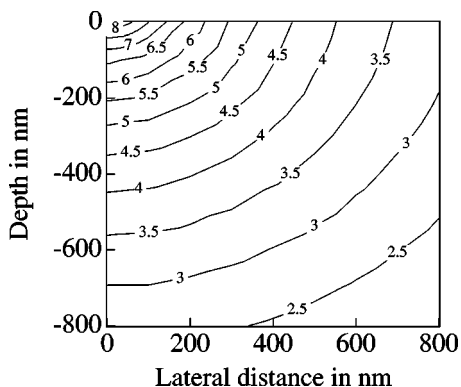


FIG. 4. Deposited power per unit volume inside the bulk. The particle has a radius  $a=5$  nm and is held at temperature  $T_p=300$  K.

were constituted by two different materials, the RHT spectrum would feature two separate sharp peaks at two different resonance frequencies.

Figure 3 shows the integrated power absorbed by the same particle versus distance  $d$  (the substrate is still held at temperature  $T_B=300$  K). The near-field RHT increases as  $1/d^3$  and is larger at small distances by several orders of magnitude than the far-field one. Indeed, for  $d=10$  nm,  $P_{\text{abs}} \approx 2.6 \times 10^{-10}$  W and for  $d=10 \mu\text{m}$ ,  $P_{\text{abs}} \approx 8.9 \times 10^{-18}$  W. This enhancement comes from the contribution of evanescent waves to the energy exchange. Therefore, this contribution cannot be neglected in a near-field calculation. This phenomenon occurs when the materials can support resonant surface waves like III–V or II–VI semiconductors and because the resonant frequencies take place in the far IR region (up to  $10 \mu\text{m}$ ) where the characteristic wavelength of thermal emission at  $T \approx 300$  K is  $10 \mu\text{m}$ . A lot of materials, like oxides or glass, can also support resonant surface waves in the IR or the visible region.

Reciprocity requires that the same enhanced RHT appears when the particle illuminates the surface. This situation may help us in understanding the radiative heat exchange between a nanotip (like those used in near-field microscopy) and a sample. It is interesting to study how the energy radiated by the tip is dissipated in the sample. To answer this question, we calculated—using Eq. (5)—the total power (integrated over the frequencies) dissipated per unit volume for

different points in the sample. Figure 4 displays a map, in log scale, of the dissipation rate in the case of a 10 nm diameter sphere of SiC at  $T_p=300$  K situated at 100 nm above a sample of SiC. It is seen that the energy is dissipated on a scale comparable to the tip–sample distance. The dissipation per unit volume decreases very fast (as  $1/r^6$ ) with the distance  $r$  between the source and the point of the sample where the dissipation is considered (the isocontour labeled with a “6” corresponds to the points where the dissipation per unit volume is  $10^6 \text{ W m}^{-3}$ ). The amount of energy locally deposited is as large as  $100 \text{ MW m}^{-3}$ .

In this letter, we have shown that nanoscale RHT between a sample and a small particle is almost monochromatic and can be enhanced by several orders of magnitude when the materials involved support resonant surface waves. When illuminated by the particle, the distribution of power inside the sample extends over distances of the same order as the particle–sample separation. These results should have broad applications in near-field microscopy, in design of nanostructures and in high density storage processes by local heating.

- <sup>1</sup>A. R. Abramson and C. L. Tien, *Microscale Thermophys. Eng.* **3**, 229 (1999).
- <sup>2</sup>G. Chen, *Microscale Thermophys. Eng.* **1**, 215 (1997).
- <sup>3</sup>J. B. Pendry, *J. Mod. Opt.* **45**, 2389 (1998).
- <sup>4</sup>E. G. Cravalho, C. L. Tien, and R. P. Caren, *J. Heat Transfer* **89**, 351 (1967).
- <sup>5</sup>C. M. Hargreaves, *Phys. Rev. Lett.* **30A**, 491 (1969).
- <sup>6</sup>D. Polder and M. Van Hove, *Phys. Rev. B* **4**, 3303 (1971).
- <sup>7</sup>J. L. Loomis and H. J. Maris, *Phys. Rev. B* **50**, 18 517 (1994).
- <sup>8</sup>K. Dransfeld and J. Xu, *J. Microsc.* **152**, 35 (1988).
- <sup>9</sup>C. C. Williams and H. K. Wickramasinghe, *Appl. Phys. Lett.* **49**, 1587 (1986).
- <sup>10</sup>R. Carminati and J. J. Greffet, *Phys. Rev. Lett.* **82**, 1660 (1999).
- <sup>11</sup>P. C. Chaumet, A. Rahmani, F. de Fornel, and J. P. Dufour, *Phys. Rev. B* **58**, 2310 (1998).
- <sup>12</sup>J. D. Jackson, *Classical Electrodynamics*, 2nd Ed. (Wiley, New York, 1975), Sec. 4.4.
- <sup>13</sup>F. Pincemin, A. Sentenac, and J. J. Greffet, *J. Opt. Soc. Am. A* **11**, 1117 (1994).
- <sup>14</sup>B. T. Draine, *Astrophys. J.* **333**, 848 (1988).
- <sup>15</sup>S. M. Rytov, Yu. A. Kravtsov, and V. I. Tatarskii, *Principles of Statistical Radiophysics* (Springer, Berlin, 1989).
- <sup>16</sup>A. V. Shchegrov, K. Joulain, R. Carminati, and J. J. Greffet, *Phys. Rev. Lett.* **85**, 1548 (2000).
- <sup>17</sup>E. W. Palik, *Handbook of Optical Constants of Solids* (Academic, San Diego, CA, 1985).

# Near-field thermal-emission spectroscopy

Karl Joulain, Rémi Carminati, Jean-Philippe Mulet and  
Jean-Jacques Greffet

Laboratoire d'Energétique Moléculaire et Macroscopique,  
Combustion, CNRS, Ecole Centrale Paris, 92295  
Châtenay-Malabry Cedex, France

We propose a new method to perform local emission spectroscopy of surfaces with a subwavelength resolution, based on the measurement of the thermally emitted electromagnetic field with a scanning near-field optical microscope. The feasibility of the method is studied using a rigorous theory of scanning near-field optical microscopy and electromagnetic calculations of the near field thermally emitted from a plane surface. We show that near-field optical techniques using sharp metallic tips should provide a signal above the detection threshold. From the near-field spectrum, it is possible to recover the local dielectric constant of the material at subwavelength scale. Finally, we show that a near-field optical microscope detecting thermally emitted fields is the optical analog of the electron scanning tunneling microscope. The measurement of the local density of states of the electromagnetic field by such an instrument is discussed. ©2002 Optical Society of America

## 1. Introduction

With the development of nanotechnologies, there is an increasing interest in the determination of the properties of materials at small length scales. This is a crucial issue, for example, in microelectronics and optoelectronics, where the features of the components and their behavior have to be certified at a sub-micronic scale. Optical spectroscopy, using visible or infrared light, is a widespread and powerful tool to investigate the properties of matter. Among the various techniques, emission spectroscopy is a natural way of studying solid surfaces. Recent studies have shown that the electromagnetic field thermally emitted by a heated surface exhibits surprising coherence properties at distances much smaller than the typical wavelengths of the emission spectrum [1, 2, 3]. It was shown recently that this near-field coherence substantially modify the far-field emission properties of microstructured surfaces [4]. Concerning the spectral behavior, two major properties of the near-field thermal emission spectrum close to flat surfaces separating a solid body from a vacuum were established [2]. (1) The energy density in the near field is enhanced by several orders of magnitude compare to its far-field value. (2) The energy spectrum exhibits a sharp peak at a given frequency, so that the near-field thermal emission is quasi-monochromatic. This behavior is observed close to all materials supporting resonant surface electromagnetic modes in the infrared domain of the spectrum. Examples of such materials are polar materials, such as glass, SiC, or II-VI and III-V semi-conductors which support surface-phonon polaritons, or doped silicon which supports surface-plasmon polaritons. Surface polaritons result from the coupling between the electromagnetic field and either a phonon or a plasmon resonance in the solid. The electromagnetic wave associated to a polariton is an evanescent wave, propagating along the interface between the medium and vacuum (surface wave). Thus, the spectral and enhancement effects associated with the excitation of surface



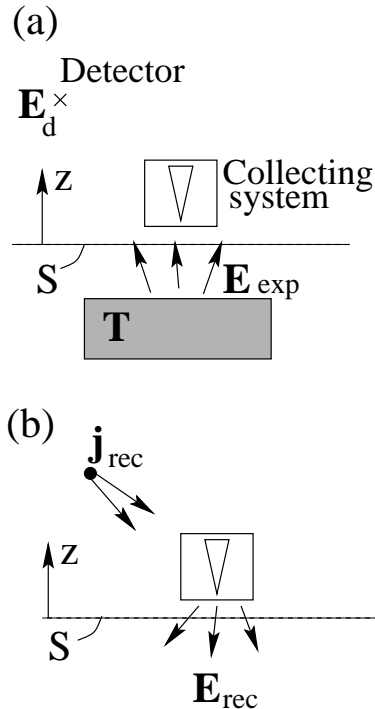


Fig. 1. Scheme of a scanning near-field optical microscope measuring a thermally emitted field. (a): experimental situation. (b): reciprocal (fictitious) situation.

polaritons are confined in the near field of the surface [1, 2].

The development of Scanning Near-field Optical Microscopy (SNOM) has made possible the measurement of confined optical near fields with a resolution reaching a few tens of nanometers [5]. SNOM techniques have been used to perform different kinds of emission spectroscopy, e.g., luminescence [6], Raman spectroscopy [7] or two-photon fluorescence [8]. For detection of infrared light, apertureless techniques [9] have shown their reliability for imaging [10] as well as for vibrational spectroscopy on molecules [11]. In this paper, we use rigorous theoretical arguments to show that the measurement of a near-field thermal-emission spectrum is possible with an apertureless SNOM using a sharp metallic tip. The predicted level of signal is above the sensitivity of usual infrared detectors. We propose such a technique to recover the dielectric constant of solid surfaces at a sub-wavelength scale. Finally, we show that an apertureless SNOM detecting thermally-emitted light is the exact optical analog of an electron scanning tunneling microscope (STM). In particular, we discuss the ability of such an instrument to measure the local density of states (LDOS) of the electromagnetic field close to a surface.

## 2. Theory of near-field detection of thermal light

We consider a scanning near-field optical microscope (SNOM) working in the detection mode, and detecting the electromagnetic field thermally emitted by a sample held at a temperature  $T$ . The system is depicted in Fig. 1. The microscope tip is scanned at close proximity of the interface separating the solid body from a vacuum. The signal is measured in the far field, by a point detector sensitive to the energy flux carried by the electromagnetic field. We assume that a analyzer is placed

in front of the detector (polarized detection). The direction of polarization of the analyzer is along the direction of the vector  $\mathbf{j}_{rec}$ . If the solid angle  $d\Omega$  under which the detector is seen from the tip is small (a condition we assume for simplicity), the signal  $\S$  at the detector, at a given frequency  $\omega$ , writes

$$\langle S(\omega) \rangle = \frac{\epsilon_0 c}{2} |\mathbf{E}_d(\omega)|^2 r^2 d\Omega \quad (1)$$

where  $c$  is the speed of light in vacuum,  $r$  is the distance between the tip and the detector, and  $\mathbf{E}_d$  is the electric field at the position of the detector. Let us denote by  $\mathbf{E}_{exp}$  (experimental field) the thermal field, emitted by the sample, in the gap region between the tip and the sample. This field can be, in principle, calculated following the approach recently used in refs.[1, 2]. For simplicity, and because this seems realistic in practice, we shall neglect the thermal emission from the tip itself (which is assumed to be cold) compared to that of the heated sample. But we do not need, at this stage, to assume a weak coupling between the tip and the sample. In particular, in the expressions derived in this section, the experimental field  $\mathbf{E}_{exp}$  is the field emitted by the sample alone, in the presence of the detecting tip. Following the approach of ref.[12], based on the reciprocity theorem of electromagnetism [13], an exact relationship between the signal  $\langle S(\omega) \rangle$  and the experimental field  $\mathbf{E}_{exp}$  can be established. To proceed, one considers a fictitious situation in which the sample is removed, and a point source, represented by a monochromatic current  $\mathbf{j}_{rec}$  oscillating at frequency  $\omega$ , is placed at the position of the detector (see Fig. 1(b)). The orientation of this reciprocal source is chosen along the direction of polarization of the analyzer used in the experimental situation. The field created around the tip in this reciprocal situation is denoted by  $\mathbf{E}_{rec}$ . Using the reciprocity theorem, the field at the detector can be written [12]:

$$\mathbf{E}_d(\omega) \cdot \mathbf{j}_{rec} = \frac{-2i}{\mu_0 \omega} \int_S \frac{\partial \mathbf{E}_{rec}(\mathbf{R}, z, \omega)}{\partial z} \cdot \mathbf{E}_{exp}(\mathbf{R}, z, \omega) d\mathbf{R} \quad (2)$$

where the integration is performed in a plane  $z = cte$  between the tip and the sample and  $\mathbf{R} = (x, y)$  are the coordinates along this plane. Equation (2) connects the field above the surface  $\mathbf{E}_{exp}$  to the field in the detector  $\mathbf{E}_d$  along the direction of the analyzer. Note that the reciprocal field  $\mathbf{E}_{rec}$  encodes all the informations about the detection system (tip and collection optics). Reporting the expression of the field at the detector (2) in (1), one finds the expression for the measured signal:

$$\langle S(\omega) \rangle = \frac{\epsilon_0 c}{8\pi^2} \int_S \int_S H_{ij}(\mathbf{R}, \mathbf{R}', z, \omega) W_{ij}(\mathbf{R}, \mathbf{R}', z, \omega) d\mathbf{R} d\mathbf{R}' . \quad (3)$$

Equation (3) is the main result of this paper. It establishes a linear relationship between the signal and The cross-spectral density tensor  $W_{ij}$  of the electric field defined by

$$\langle E_{exp,i}(\mathbf{R}, z, \omega) E_{exp,j}^*(\mathbf{R}', z, \omega') \rangle = W_{ij}(\mathbf{R}, \mathbf{R}', z, \omega) \delta(\omega - \omega') . \quad (4)$$

The response function  $H_{ij}$  only depends on the detection system (in particular the tip geometry and composition), and is given by

$$H_{ij}(\mathbf{R}, \mathbf{R}', \omega) = \frac{\partial \mathbf{E}_{rec,i}(\mathbf{R}, z, \omega)}{\partial z} \frac{\partial \mathbf{E}_{rec,j}(\mathbf{R}', z, \omega)}{\partial z} . \quad (5)$$

The cross-spectral density tensor  $W_{ij}$  describes the electric-field spatial correlation at a given frequency  $\omega$ . For the thermal emission situation considered here, it depends only on the dielectric constant, on the geometry and on the temperature of the sample. The linear relationship described in Eq. (3) suggested that the optical properties of the sample (frequency-dependent dielectric constant) might be recovered from the SNOM signal. Furthermore, the behavior of  $W_{ij}$  has been

studied recently in the near field of flat surfaces supporting resonant surface electromagnetic modes, such as SPP [1, 2]. It has been shown that at nanometric distances,  $W_{ij}(\mathbf{R}, \mathbf{R}', z, \omega)$  is quasi-monochromatic and is enhanced by several orders of magnitude compare to its far field value. Near such materials, we shall see below show that the signal detected by an apertureless SNOM using a sharp metallic tip should be above the detection treshold, and have a spectral shape allowing to determine the dielectric constant of the material.

### 3. Principle of the method of near-field spectroscopy

The principle of the method can be established by considering a model sample made of a semi-infinite homogenous material filling the half-space  $z < 0$ , with a flat surface in the plane  $z = 0$ . For this geometry, the cross-spectral density tensor can be expanded asymptotically for small distances  $z$  between the observation point and the surface [2, 3, 14]. The result may be written as the product of a frequency-dependent function and a function which depends only on the observation point:

$$W_{ij}(\mathbf{R}, \mathbf{R}', z, \omega) \approx \frac{\hbar}{4\pi^2\epsilon_0} \frac{Im[\epsilon(\omega)]}{|\epsilon(\omega) + 1|^2} \frac{1}{\exp(\hbar\omega/k_B T) - 1} w_{ij}(\mathbf{R}, \mathbf{R}', z) \quad (6)$$

Here  $\hbar$  is the reduced Planck constant,  $k_B$  is the Boltzmann constant,  $T$  is the temperature of the sample and  $w_{ij}(z, \mathbf{R}, \mathbf{R}')$  is a known function of the spatial coordinates which does not need to be explicitly given here. As shown in refs.[2, 3],  $W_{ij}(z, \mathbf{R}, \mathbf{R}', \omega)$  exhibits a peak when varying the frequency for a given position in the near field above materials supporting surface waves, such as SPP. Indeed, the quantity  $Im[\epsilon(\omega)]/|\epsilon(\omega) + 1|^2$  is proportional to the density of electromagnetic modes close to the interface, which is enhanced at the frequency corresponding to the resonance of the material. This resonance occurs where  $Re(\epsilon) = -1$ . An analysis of the peak of the function  $W_{ij}$  peak could lead to the measurement of the dielectric constant. More precisely, recording a spectrum of  $W_{ij}(z, \mathbf{R}, \mathbf{R}', \omega)$  for a given position allows, in principle, to retrieve the quantity  $Im(\epsilon)/|\epsilon + 1|^2 = Im[(\epsilon - 1)/(\epsilon + 1)]/2$ .  $(\epsilon - 1)/(\epsilon + 1)$  being a causal quantity,  $Re[(\epsilon - 1)/(\epsilon + 1)]$  can be calculated from  $Im[(\epsilon - 1)/(\epsilon + 1)]$  using the Kramers-Kronig relations. Once  $[\epsilon(\omega) - 1]/[\epsilon(\omega) + 1]$  is known,  $\epsilon(\omega)$  is known. Note that because the spectrum of  $Im(\epsilon)/|\epsilon + 1|^2$  is sharply peaked on materials supporting SPP, one do not need a wide-band spectral measurement to accurately use the Kramers-Kronig relations. A narrow-band spectrum around the peak frequency should be sufficient to ensure a precise computation of  $Re[(\epsilon - 1)/(\epsilon + 1)]$ .

This procedure assumes that in practice, the cross-spectral density tensor  $W_{ij}$  can be deduced from the SNOM signal. This is the case if the response function  $H_{ij}$  appearing in Eq. (3) is known. We shall see in the next section that an accurate model for  $H_{ij}$  can be developed for apertureless SNOM techniques using conical metallic tips. This model allows a quantitative estimation of the signal level, thus showing on a rigorous basis the feasibility of the method.

### 4. Results for apertureless near-field optical microscopy

As shown in equation (3), the signal at the detector is determined by the value of the two-point function  $H_{ij}(\mathbf{R}, \mathbf{R}', z, \omega) W_{ij}(\mathbf{R}, \mathbf{R}', z, \omega)$  in a plane  $z = cte$  between the sample and the tip. We now restrict the discussion to apertureless SNOM techniques [9], using a metallic tip of (approximately) conical shape. A model for such tips has been introduced [15], based on an asymptotic expression of the field close to the apex of a perfectly-conducting cone [16]. This model has been used recently, together with the expression (2), to describe the spectral response of an apertureless set-up [12]. The model allows to calculate the reciprocal field  $\mathbf{E}_{rec}$ , which exhibits a frequency dependence in  $\omega^\nu$ , where  $\nu$  is a real number which depends only on the cone angle. From this dependence and Eq. (2), the complete frequency dependence of the signal has been calculated, in excellent agreement

with experiments [17]. In the case of the present study, the frequency dependence of the response function  $H_{ij}$  can be determined along the same lines.

Let us take the tip apex at the origin of the coordinates. A detailed study of the function  $H_{ij}(\mathbf{R}, \mathbf{R}', z, \omega)$  in a plane  $z = cte$  shows that it takes significant values only if  $|\mathbf{R}|$  and  $|\mathbf{R}'|$  are smaller than the distance between this plane and the tip apex. We can thus infer that in the plane of integration  $z = cte$ , the function  $H_{ij}(\mathbf{R}, \mathbf{R}', z, \omega) W_{ij}(\mathbf{R}, \mathbf{R}', z, \omega)$  will take large values only on a zone situated just below the tip which lateral size is on the order of the tip-sample distance. In other words, the signal  $\langle S(\omega) \rangle$  at the detector corresponds to the near-field detection of the thermal field on a zone whose lateral size is on the order of the tip-sample distance.

Let us now turn to the spectral dependence. The use of the perfectly-conducting cone model to evaluate the response function  $H_{ij}$  leads to:

$$\langle S(\omega) \rangle = \frac{Im[\epsilon(\omega)]}{|\epsilon(\omega) + 1|^2} \frac{\omega^{2\nu}}{\exp(\hbar\omega/k_B T) - 1} I(d) \quad (7)$$

where  $d$  is the distance between the tip and the sample.  $I(d)$  is a function which depends on the set-up geometry, but not on frequency. We have evaluated  $I(d)$  numerically by calculating the integral in Eq. (3). It scales approximately as  $1/d^3$ . We see also in Eq. (7) that the signal at the detector has a strong spectral dependence. When the tip is in the near field of a material supporting surface waves, the density of electromagnetic modes is resonant near the frequency where  $Re[\epsilon(\omega)] = -1$ . The signal at the detector exhibits a sharp peak around this frequency, and the signal integrated over all frequencies is also strongly enhanced. We have evaluated the signal at the detector in the case of a silicon carbide (SiC) surface, which is known to produce a strong enhancement of the thermally emitted near field [1, 2] due to the excitation of surface-phonon polaritons. In the case of a tip situated at 10 nm above a SiC sample heated at  $T = 500$  K, the signal at the detector is greater than  $10^{-7} \text{W.sr}^{-1}$  which is above the detection threshold of usual infrared detectors used in SNOM [18]. Equation (7) show also that if the function  $I(d)$  is known, and if we are able to record a spectrum of the signal  $\langle S(\omega) \rangle$ , one should be able to retrieve the dielectric constant  $\epsilon(\omega)$  following the method described in section 3. Moreover, the measurement could be done with a lateral resolution on the order of the tip-sample distance, which is below  $10 \text{nm}$  in apertureless SNOM.

## 5. Measurement of local density of states and analogy with STM

With the development of near-field optics, the direct measurement of the local density of states (LDOS) of the electromagnetic field close to a surface seems to be possible [19, 20, 21]. The LDOS is a basic quantity allowing to compute radiative lifetimes of atoms or molecules, as well as thermal fields energy and forces. Shaping the LDOS by structuring a surface would allow, for example, to control the emission of elementary sources. This is a key issue for nanooptics. In this section, we show that measuring thermally emitted fields with a SNOM is a natural way to map the LDOS of the electromagnetic field. This results also demonstrates that a SNOM measuring thermal fields is the optical analog of an electron STM.

### A. Local density of states

The cross-spectral density defined in Eq. (4) allows to compute the electric energy density  $U(\mathbf{r}, \omega)$  at a given point  $\mathbf{r} = (\mathbf{R}, z)$  by

$$U(\mathbf{r}, \omega) = \frac{\epsilon_0}{2} W_{kk}(\mathbf{r}, \mathbf{r}, \omega) \quad (8)$$

For light thermally emitted into a vacuum by a homogeneous medium, one can write

$$W_{kk}(\mathbf{r}, \mathbf{r}, \omega) = \frac{\mu_0 \omega^3}{\pi c^2} \epsilon''(\omega) \Theta(\omega, T) \sum_{k,m} \int_V |G_{km}(\mathbf{r}, \mathbf{r}', \omega)|^2 d\mathbf{r}' \quad (9)$$

where  $V$  is the volume of the emitting media,  $\epsilon''$  the imaginary part of its dielectric constant and  $G_{km}$  the Green dyadic accounting for the geometry of the medium. For a semi-infinite medium separated from vacuum by a flat interface, its expression is given in ref.[1]. Note that Eq. (9) is valid for any geometry, provided that the Green dyadic for this geometry is used. For complicated geometries with low symmetris, this dyadic can be calculated numerically, following for example the method described in refs.[22].

In order to introduce the LDOS of the electromagnetic field, we rewrite the energy density in the form

$$U(\mathbf{r}, \omega) = \rho(\mathbf{r}, \omega) \frac{\hbar\omega}{\exp(\hbar\omega/k_B T) - 1} \quad (10)$$

For a thermal equilibrium situation, this equation is a standard result of statistical physics which tells us that the energy density at a given point  $\mathbf{r}$  and a given frequency  $\omega$  is the product of the photon energy, the mean number of photons per states (given here by the Bose-Einstein statistics) and the local density of states  $\rho(\mathbf{r}, \omega)$ . For the *non-equilibrium* situation considered here (the thermal emission by a sample in a vacuum assumed to be at  $T = 0K$ ), we can use Eq. (10) as a definition of the LDOS. Using Eqs. (8) and (9), this yields

$$\rho(\mathbf{r}, \omega) = \frac{\omega^3}{2\pi c^4} \epsilon''(\omega) \sum_{k,m} \int_V |G_{km}(\mathbf{r}, \mathbf{r}', \omega)|^2 d\mathbf{r}' \quad (11)$$

This expression deserves a comment. The LDOS defined in Eq. (11) is not given by the imaginary part of the trace of the Green dyadic, and therefore differs from the definition often used for the electromagnetic field LDOS [19, 20, 21]. The reason is that Eq. (11) describes a non-equilibrium thermal-emission problem, for which all the modes of the electromagnetic field cannot be excited. Indeed, the LDOS defined here only accounts for the modes which can be occupied in a vacuum corresponding to the half-space  $z > 0$ , after thermal emission by a medium occupying the half-space  $z < 0$ . Note that, as we shall see below, this LDOS is the relevant quantity that could be measure by a near-field optical microscope.

## B. Detection of the LDOS by an ideal probe

In this setion, we discuss the measurement of the LDOS defined above by a near-field optical microscope working in the detection mode. Equation (3) gives an exact expression of the signal measured by such an instrument. The general relationship between the signal and the cross-spectral density tensor is non-local and strongly polarization dependent. This shows that one do not measure in general a quantity which is proportional to  $W_{kk}(\mathbf{r}, \mathbf{r}, \omega)$ , and thus to the LDOS.

Let us see what would be measured by an ideal probe consisting of a single electric dipole described by a polarizability  $\alpha(\omega)$ . Note that such a probe was proposed as a model for the uncoated dielectric probe sometimes used in photon scanning tunneling microscopy (PSTM), and gives good qualitative prediction [23]. We assume that the thermally emitting medium occupies the half-space  $z < 0$ , and that the probe is placed at a point  $\mathbf{r}_t$ . As in section 2, the detector placed in the far field measures the field intensity at a given point  $\mathbf{r}_d$ , through an analyser whose polarization direction is along the vector  $\mathbf{j}_{rec}$ . In this case, Eq. (2) simplifies to write

$$\mathbf{j}_{rec} \cdot \mathbf{E}_d = \alpha(\omega) \frac{\omega^2}{4\pi c^2} \frac{\exp(ik|\mathbf{r}_d - \mathbf{r}_t|)}{|\mathbf{r}_d - \mathbf{r}_t|} \mathbf{j}_{rec} \cdot \overset{\leftrightarrow}{\mathbf{h}}(\mathbf{u}_d) \cdot \mathbf{E}_{exp}(\mathbf{r}_t, \omega) \quad (12)$$

where  $k = \omega/c$ ,  $\mathbf{u}_d = (\mathbf{r}_d - \mathbf{r}_t)/|\mathbf{r}_d - \mathbf{r}_t|$  is the unit vector pointing from the probe towards the detector and  $\overleftrightarrow{\mathbf{h}}(\mathbf{u}_d) = \overleftrightarrow{\mathbf{I}} - \mathbf{u}_d \mathbf{u}_d$  is the dyadic operator which projects a vector on the direction transverse to  $\mathbf{u}_d$ ,  $\overleftrightarrow{\mathbf{I}}$  being the unit dyadic operator. The dyadic  $\overleftrightarrow{\mathbf{h}}(\mathbf{u}_d)$  being symmetric, the scalar product in the right-hand side in Eq. (12) can be transformed using the equality  $\mathbf{j}_{rec} \cdot \overleftrightarrow{\mathbf{h}}(\mathbf{u}_d) \cdot \mathbf{E}_{exp}(\mathbf{r}_t, \omega) = \mathbf{E}_{exp}(\mathbf{r}_t, \omega) \cdot \overleftrightarrow{\mathbf{h}}(\mathbf{u}_d) \cdot \mathbf{j}_{rec}$ . Finally, the signal at the detector writes

$$\langle S \rangle = |\alpha(\omega)|^2 \frac{\omega^4}{4\pi c^4} d\Omega \sum_{i,j} A_i A_j^* W_{ij}(\mathbf{r}_t, \mathbf{r}_t, \omega) \quad (13)$$

where  $\mathbf{A} = \overleftrightarrow{\mathbf{h}}(\mathbf{u}_d) \cdot \mathbf{j}_{rec}$  is a vector depending only on the detection conditions (direction and polarization). Note that if  $\mathbf{j}_{rec}$  is transverse with respect to the direction  $\mathbf{u}_d$ , which is approximately the case in many experimental set-ups, then one simply has  $\mathbf{A} = \mathbf{j}_{rec}$ .

Equation (13) shows that with an ideal probe consisting of a signal dipole (with an isotropic polarizability  $\alpha(\omega)$ ), one locally measures the cross-spectral density tensor at the position  $\mathbf{r}_t$  of the tip. Nevertheless, polarization properties of the detection still exists so that the energy density (i.e. the trace of  $W_{ij}$ ), and therefore the LDOS, is not directly measured. A possibility of measuring the trace would be to measure a signal  $\langle S_1 \rangle$  in the direction normal to the surface with an unpolarized detection, and a signal  $\langle S_2 \rangle$  in the direction parallel to the surface, with an analyzer in the vertical direction.  $\langle S_1 \rangle$  would be a sum of the two signals obtained with  $\mathbf{j}_{rec}$  along the  $x$ -direction and along the  $y$ -direction.  $\langle S_2 \rangle$  would correspond to the signal measured with  $\mathbf{j}_{rec}$  along the  $z$ -direction. Using Eq. (13), we see that the signal  $\langle S \rangle = \langle S_1 \rangle + \langle S_2 \rangle$  will be proportional to the trace  $W_{kk}(\mathbf{r}_t, \mathbf{r}_t, \omega)$ , and thus to the LDOS  $\rho(\mathbf{r}_t, \omega)$ .

### C. Analogy with scanning (electron) tunneling microscopy

The result in this section shows that a SNOM measuring the thermally emitted field with a dipole probe (for example a sphere much smaller than the existing wavelengths) measures the electromagnetic LDOS of the sample. As discussed above, the measured LDOS is that of the modes which can be excited in the thermal emission process in a cold vacuum. This result was obtained from Eq. (2) assuming a weak tip-sample coupling, i.e., the experimental field is assumed to be the same with or without the tip.

The same result could be obtained starting from the generalized Bardeen formula derived in ref.[24]. Using this formalism for a dipole probe, one also ends up with Eq. (13), which explicitly shows the linear relationship between the signal and the LDOS. This derivation is exactly the same as that used in the Tersoff and Haman theory of the STM [25]. This theory showed, in the weak tip-sample coupling limit, that the electron-tunneling current measured in STM was proportional to the electronic LDOS of the sample, at the tip position, and at the Fermi energy. This result, although obtained under some approximations, was a breakthrough in understanding the STM signal. In the case of near-field optics, the present discussion, together with the use of the generalized Bardeen formula [24], shows that under similar approximations, a SNOM using an ideal dipole probe and measuring the field thermally emitted by the sample is the real optical analog to the electron STM. We believe that this situation provides for SNOM a great potential for local solid-surface spectroscopy, along the directions opened by STM.

## 6. Conclusion

We have shown that a near-field thermal emission spectrum is theoretically measurable by an apertureless SNOM above a surface separating a vacuum from a material exhibiting resonant surface waves. The signal level and its frequency dependence have been determined using a rigorous theory.

A novel method to perform a local spectroscopy of a solid surface has been proposed, based on a measurement of the thermally emitted electromagnetic near field. The analogy between a SNOM measuring thermal fields and the electron STM has been established and discussed. This shows that the proposed method should have the same potential, in infrared spectroscopy, as the spectroscopic techniques developed in STM.

## References

1. R. Carminati and J.-J. Greffet, Phys. Rev. Lett. **82**, 1660 (1999).
2. A.V. Shchegrov, K. Joulain, R. Carminati and J.-J. Greffet, Phys. Rev. Lett. **85**, 1548 (2000).
3. C. Henkel, K. Joulain, R. Carminati and J.-J. Greffet, Opt. Comm. **186**, 57 (2000).
4. J.-J. Greffet, R. Carminati, K. Joulain, J.-P. Mulet, S. Mainguy and Y. Chen, Nature **416**, 61 (2002).
5. D.W. Pohl and D. Courjon (eds.), *Near-Field Optics* (Kluwer, Dordrecht, 1993); M.A. Paesler and P.J. Moyer, *Near Field Optics: Theory, Instrumentation and Applications* (Wiley-Interscience, New-York, 1996); M. Ohtsu (ed.), *Near-field Nano/Atom Optics and Technology* (Springer-Verlag, Tokyo, 1998)
6. P.J. Moyer, C.L. Jahnckle, M.A. Paesler, R.C. Reddick and R.J. Warmack, Phys. Lett. A **45**, 343 (1990).
7. C.L. Jahnckle, M.A. Paesler and H.D. Hallen, Appl. Phys. Lett. **67**, 2483 (1995); J. Grausem, B. Humbert, A. Burenau, J. Oswald, Appl. Phys. Lett. **70**, 1671 (1997).
8. J.E. Sanchez, L. Novotny and X. S. Xie, Phys. Rev. Lett. **82**, 4014 (1999).
9. F. Zenhausern, M.P. O'Boyle and H.K. Wickramasinghe, Appl. Phys. Lett. **65**, 1623 (1994); Y. Inouye and S. Kawata, Opt. Lett. **19**, 159 (1994); P. Gleyzes, A.C. Boccara and R. Bachelot, Ultramicroscopy **57**, 318 (1995).
10. A. Lahrech, R. Bachelot, P. Gleyzes and A.C. Boccara, Appl. Phys. Lett. **71**, 575 (1997).
11. B. Knoll and F. Keilmann, Nature **399**, 134 (1999).
12. J.A. Porto, R. Carminati and J.-J. Greffet, J. Appl. Phys. **88**, 4847 (2000).
13. J.-J. Greffet and R. Carminati, Prog. Surf. Sci. **56**, 133 (1997).
14. J.-P. Mulet, K. Joulain, R. Carminati and J.-J. Greffet, Appl. Phys. Lett **78**, 2931 (2001).
15. H. Cory, A.C. Boccara, J.C. Rivoal and A. Lahrech, Microwave Opt. Technol. Lett. **18**, 120 (1998).
16. J.J. Bowman, T.B.A Senior and P.L.E. Uslenghi, *Electromagnetic and acoustic scattering by simple shapes* (Hemisphere, New-York, 1987), p. 662.



17. L. Aigouy, F.X. Andreani, A.C. Boccarda, J.C. Rivoal, J.A. Porto, R. Carminati, J.-J. Greffet and R. Megy, *Appl. Phys. Lett.* **76**, 397 (2000).
18. A.C. Boccarda, L. Aigouy and Y. De Wilde, *Private Communication*.
19. C. Girard, C. Joachim and S. Gauthier, *Rep. Prog. Phys.* **63**, 893 (2000).
20. G. Colas des Francs, C. Girard, J.C. Weeber, C. Chicane, T. David, A. Dereux and D. Peyrade, *Phys. Rev. Lett.* **86**, 4950 (2001).
21. A. Dereux, C. Girard and J.C. Weeber, *J. Chem. Phys.* **112**, 7775 (2000).
22. O.J.F. Martin, C. Girard and A. Dereux, *Phys. Rev. Lett.* **74**, 526 (1995); N.B. Piller and O.J.F. Martin, *IEEE Trans. Antennas Prop.* **46**, 1126 (1998).
23. D. Van Labeke and D. Barchiesi, *J. Opt. Soc. Am. A* **10**, 2193 (1993).
24. R. Carminati and J.J. Sáenz, *Phys. Rev. Lett.* **84**, 5156 (2000).
25. J. Tersoff and D.R. Hamann, *Phys. Rev. B* **31**, 805 (1985).