

Facoltà di Scienze Matematiche Fisiche e Naturali

Ph.D. School in Physics and Nanoscience

Role of the vacuum fluctuation forces in microscopic systems

Author
Andrea Benassi

Supervisor
Prof. Carlo Calandra
Buonaura

To my parents

Introduction

In the last few years an increasing attention has been addressed to the nature and intensity of the forces arising from the electromagnetic vacuum fluctuations (Casimir and van der Waals or dispersion forces) [1, 2]. The interest in this subject comes from some basic developments: (i) the increased experimental effort, that has led to accurate measurements of the force intensity for specific configurations [3, 4, 5, 6, 7, 8, 9, 10], (ii) the understanding of the relationship between the force and the elementary excitations of the interacting bodies, that has allowed to generalization to real materials of results originally obtained using idealized boundary conditions [11, 12, 13, 14, 15, 16, 17, 18], (iii) the role played by these forces in nano- and microdevices [19, 20, 21, 22, 23], (iv) the possibility of exploiting different geometrical configurations to achieve the desired properties of the forces [24, 25].

The purpose of the present thesis is to study the role that electromagnetic vacuum fluctuation forces may play in a number of issues that are relevant at the nano- microscopic scale. The study is devoted to the following problems:

i) **Stability of deposited metal films:** investigation of the importance of vacuum fluctuation force in determining the critical thickness of metal overlayer with respect to the transition between a uniform two-dimensional phase and a corrugated one [26, 27, 28]. The attention is focused first on the force acting on the film boundaries in the case of a free standing film and then on the modifications of this force when the film is deposited on a metal substrate. It is shown that, while the force does not contribute to the surface stability in the case of an isolated film, being too weak in comparison with the surface stress, it can be crucially important in the case of deposition, where vacuum electromagnetic energy may be of the same order of magnitude of the difference between the elastic energy caused by the lattice mismatch, which favours the surface corrugation to release the excess elastic energy, and the surface energy, which tends to favour the planar configuration. The study provide the conditions under which the force on the film can sustain the stability of the planar configuration, given in terms of the parameters entering into the expression of the dielectric function.

ii) **Size effects in the Casimir interaction between ultrathin films:** study of the effects of the quantization of the energy levels caused by quantum confinement on the intensity of the force between metal films. The use of bulk dielectric functions in the force calculation within the framework of the Lifshitz theory, while appropriate for thick enough films [29], is expected to give unrealistic results when the film thickness is of the order of few nanometers. The study has been carried out by determining the dielectric function of a confined system of free electrons using the RPA approximation to calculate the dielectric function [30]. This turns out to have a tensorial character, consistent with the strong anisotropy of the electron distribution, and to depend upon the film thickness through the electron energies and the dipole matrix elements. In view of the importance of the confinement potential different models have been considered: the *particle in a box* model [31] and models where the laterally averaged one electron potential is represented by a finite well. The modifications with respect to calculations with the bulk dielectric function have been analyzed as a function of the potential depth and of the electron density.

iii) **First principles calculation of the force between ultrathin silicon films:** a study aiming at understanding how modifications in the dielectric function induced by the reduced film size and by the occurrence of surface states may cause changes in the vacuum fluctuation force between silicon slabs. Density functional theory has been used to determine the film dielectric tensor starting from the one electron energies and wavefunctions calculated self-consistently for the film, using both simple RPA [30] and the RPA corrected by local field effects to determine the macroscopic dielectric tensor at vanishing wavevector from the inverse dielectric matrix [32, 33]. The force calculated using the film dielectric tensor is compared with the calculation starting from the bulk dielectric function obtained with the same theory. Evidence is provided that the presence of surface states can affect the force over a large range of film separation distances. The macroscopic dielectric tensor turns out to be very sensitive to local field effects. As a consequence the force can be significantly modified by their inclusion.

iv) **Influence of metal insulator transition in device actuated by the electromagnetic vacuum force:** the proposal of a device that exploits the phase transition to extend the distance and energy ranges over which it can be operated. The basic device components are films of GeTe [34, 35, 36] (or other compounds with similar behaviour [37, 38, 39]) , a material that undergoes a rapid transitions between polycrystalline phases, which show a metallic behaviour with a p-type conductivity due to vacant Ge sites, and amorphous phases with typical semiconductor properties [40]. Starting from empirically derived dielectric function one can obtain device potential profile and bifurcation diagrams illustrating how the device properties can be tuned optically.

The plan of the thesis is the following. The first two chapters are devoted to the presentation of the basic theory. In particular the first one summarize basic concepts of electrodynamics of linear media and presents a discussion of model dielectric functions to be used in the applications. The second chapter is a presentation of the Lifshitz theory of the electromagnetic vacuum fluctuation forces and shortly review recent developments. Each of the following chapters is devoted to one of the above outlined issues and presents a detailed discussion of the subject under consideration, a complete illustration of the results and the appropriate conclusions.

Bibliography

- [1] M. Bordag, U. Mohideen, and V.M. Mostepanenko. *Phys. Rep.*, 353:1, 2001.
- [2] S.K. Lamoreaux. *Rep. Prog. Phys.*, 68:201, 2005.
- [3] S.K. Lamoreaux. *Phys.Rev.Lett.*, 78:5, 1997.
- [4] U. Mohideen and A. Roy. *Phys. Rev. Lett.*, 81:4549, 1998.
- [5] B.W. Harris, F. Chen, and U. Mohideen. *Phys. Rev. A*, 62:052109, 2000.
- [6] T. Ederth. *Phys. Rev. A*, 62:062104, 2000.
- [7] H.B. Chan, V.A. Aksyuk, R.N. Kleiman, D.J. Bishop, and F. Capasso. *Science*, 291:1941, 2001.
- [8] G. Bressi, G. Carugno, R. Onofrio, and G. Ruoso. *Phys. Rev. Lett.*, 88:041804, 2002.
- [9] R.S. Decca, E. Fischbach, G.L. Klimchitskaya, D.E. Krause, D. Lopez, and V.M. Mostepanenko. *Phys. Rev. D*, 68:116003, 2003.
- [10] R.S. Decca, D. Lopez, E. Fischbach, G.L. Klimchitskaya, D.E. Krause, and V.M. Mostepanenko. *Ann.Phys.*, 318:37, 2005.
- [11] E.M. Lifshitz. *Sov. Phys. JEPT*, 2:73, 1956.
- [12] I.E. Dzyaloshinskii, E.M. Lifshitz, and L.P. Pitaevskii. *Adv. Phys.*, 10:165, 1958.
- [13] H. Krupp. *Advan. Colloid Interface*, 1:111, 1967.
- [14] L. Bergström. *Adv. Colloid Interface Sci.*, 70:125, 1987.
- [15] B.E. Sernelius. *Surface Modes in Physics*. Wiley-VCH, Berlin, 2001.
- [16] A. Lambrecht and S. Reynaud. *Eur. Phys. D*, 8:309, 2000.
- [17] F. Intravaia and A. Lambrecht. *Phys. Rev. Lett.*, 94:110404, 2005.
- [18] K. Joulain, J.P. Mulet, F. Marquier, R. Carminati, and J.J. Greffet. *Surf. Sci. Rep.*, 57:59, 2005.
- [19] F.M. Serry, D. Walliser, and G.J. Maclay. *J. Appl. Phys.*, 84:2501, 1998.
- [20] E. Buks and M.L. Roukes. *Phys. Rev. B*, 63:033402, 2001.
- [21] H.B. Chan, V.A. Aksyuk, R.N. Kleiman, D.J. Bishop, and F. Capasso. *Phys.Rev.Lett.*, 87:211801, 2001.
- [22] D. Iannuzzi, M. Lisanti, J.N. Munday, and F. Capasso. *Solid State Comm.*, 135:618, 2005.
- [23] J. Barcenas, L. Reyes, and R. Esquivel-Sirvent. *Appl. Phys. Lett.*, 87:263106, 2005.

-
- [24] H.B. Chan, Y. Bao, J. Zou, R. A. Cirelli, F. Klemens, W.M. Mansfield, and C.S. Pai. *Phys. Rev. Lett.*, 101:030401, 2008.
- [25] A. Lambrecht and V.N. Marachevsky. *Phys. Rev. Lett.*, 101:160403, 2008.
- [26] R. Asaro and W. Tiller. *Met.Trans.*, 3:1789, 1971.
- [27] M. Grinfield. *Sov.Phys.Dokl.*, 31:831, 1986.
- [28] Z.Y. Zhang, Q. Niu, and C.K. Shih. *Phys.Rev.Lett.*, 85:5158, 1998.
- [29] P.A. Maia Neto, A. Lambrecht, and S. Reynaud. *Phys. Rev. A*, 78:012115, 2008.
- [30] H. Ehrenreich and M.H. Cohen. *Phys.Rev.*, 115:786, 1959.
- [31] D.M. Wood and N.W. Ashcroft. *Phys.Rev.B*, 25:6255, 1982.
- [32] S.L. Adler. *Phys.Rev.*, 126:413, 1962.
- [33] N. Wiser. *Phys.Rev.*, 129:62, 1963.
- [34] K.L. Chopra and S.K. Bahl. *J.Appl.Phys.*, 40:4171, 1969.
- [35] S.K. Bahl and K.L. Chopra. *J.Appl.Phys.*, 40:4940, 1969.
- [36] S.K. Bahl and K.L. Chopra. *J.Appl.Phys.*, 41:2196, 1970.
- [37] M. Wuttig. *Nature Mat.*, 4:267, 2005.
- [38] M. Wuttig, D. Lusebrink, D. Wamwangi, W. Welnic, M. Gilleben, and R. Dreonskowski. *Nature Mat.*, 6:122, 2007.
- [39] W. Welnic, A. Pamungkas, R. Detemple, C. Steimer, S. Blugel, and M. Wuttig. *Nature Mat.*, 5:56, 2005.
- [40] W. Welnic, S. Botti, L. Reining, and M. Wuttig. *Phys.Rev.Lett.*, 98:236403, 2007.

Acknowledgments

I would like to express my special thanks to professor Carlo Calandra Buonauro for his invaluable guidance, help and support over these years.

Another special thanks to professor Elisa Molinari who hosted me in *S3 national research center* giving me the opportunity to attend schools, workshop and seminars all over the world.

I want to acknowledge *CINECA consorzio interuniversitario* for funding my Ph.D. fellowship. I am grateful to Carlo Cavazzoni from the High Performance Computing Department, for the huge amount of computational time and for the precious technical help he provided me with.

I am grateful to Dr. Andrea Ferretti, Dr. Daniele Varsano and Dr. Stefano Corni for their interest in my activity and for their important contributions.

Many thanks to Dr. Layla Martin-Samos, Dr. Agostino Migliore, Dr. Gianmarco Bramanti, Dr. Massimo Rontani for the helpful material they address me to and for the useful discussions about the theoretical concepts reviewed in the first part of this thesis.

List of publications

- *A study of the electromagnetic-fluctuation-induced forces on thin metallic films*
A. Benassi, C. Calandra
J.Phys.A:Math.Theor. **40** 13453 (2007)
- *Role of vacuum fluctuation forces in thin metal film stability*
A. Benassi, C. Calandra
J.Phys.A:Math.Theor. **41** 175401 (2008)
- *Vacuum fluctuation forces between ultra-thin films*
A. Benassi, C. Calandra
Eur.Phys.Lett. **82** 61002 (2008)
- *Effects of phase transitions in devices actuated by the electromagnetic vacuum force*
A. Benassi, C. Calandra
Eur.Phys.Lett. **84** 11002 (2008)
- *Size quantization effects in thin film Casimir interaction*
A. Benassi, C. Calandra
J.Phys. Conference Series, to be published
- *Optical Properties of Emeraldine Salt Polymers from Ab Initio Calculations: Comparison with Recent Experimental Data*
R. Colle , P. Parruccini, A. Benassi and C. Cavazzoni
J.Phys.Chem.B **111** 2800 (2007)

Contents

Theoretical background	15
1 Electrodynamics of continuous media	17
2 Macroscopic dispersion forces	67
New developments and applications	101
3 Thin films stability	103
4 Quantum confinement effects	149
5 First principle calculation of the force between silicon films	181
6 Optically tunable nano-devices	193

I

Theoretical background

1

Electrodynamics of continuous media

Contents

1.1	Maxwell's equations	17
1.2	Maxwell's equations in a medium	19
1.3	On the tensorial nature of the response functions	24
1.4	Reducible and irreducible response	26
1.5	Macroscopic and microscopic response	28
1.6	Dielectric tensor and magnetic permittivity	32
1.7	Medium excitations and dissipation	34
1.8	Dielectric function: outline of the different models	36
1.9	Sum rules	42
1.10	Response functions and many body interaction	43
1.11	The Wood and Ashcroft model	48
	Appendix A: Linear response theory	53
	Appendix B: Decomposition of a vector field into transverse and longitudinal parts	56
	Appendix C: Dispersion relations, causality and stability	57
	Appendix D: Difficulties in the dielectric function measurement	62

This first chapter contains an overview of the classical and quantum theories of the interaction between electromagnetic fields and matter. Starting from the classical theory, the concept of response function will be introduced, its physical meaning will be discussed together with its general properties. Quantum models for dielectric functions, both within the independent particle and the many body frameworks, will be analyzed and compared to the classical ones. A deep and detailed discussion about the topics of this chapter is given in references [1] and [2].

1.1 Maxwell's equations

Any electromagnetic phenomenon can be completely described by the Maxwell's equations. They relate the electric field \mathbf{E} and the magnetic induction \mathbf{B} with their sources, the charge

density ρ and the current density \mathbf{J} . Using the CGS Gauss unit system, Maxwell's equations are

$$\begin{cases} \nabla \cdot \mathbf{E} = 4\pi\rho \\ \nabla \times \mathbf{B} - \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t} = \frac{4\pi}{c} \mathbf{J} \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \nabla \cdot \mathbf{B} = 0 \end{cases} \quad (1.1)$$

These equations contain implicitly the continuity equation:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{J} = 0 \quad (1.2)$$

that can be derived combining the first equation with the divergence of the second one. Maxwell's equations are invariant under Lorentz transform so they are intrinsically relativistic. If a probe charge q is placed into the region containing fields and sources, it is subject to the force:

$$\mathbf{F} = q\left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B}\right) \quad (1.3)$$

This expression shows that the concept of field replaces, throughout the charge, the concept of force. The use of fields makes the probe charge q independent from the field sources ρ e \mathbf{J} : two different density of charge and current can give rise to the same field in the same region, since the probe charge feels the same force in both cases. The force field or the field itself does not univocally determine the charge and current distributions generating them.

Maxwell's equations in reciprocal space look like:

$$\begin{cases} i\mathbf{q} \cdot \mathbf{E} = 4\pi\rho \\ i\mathbf{q} \times \mathbf{B} + i\frac{\omega}{c} \mathbf{E} = \frac{4\pi}{c} \mathbf{J} \\ \mathbf{q} \times \mathbf{E} - \frac{\omega}{c} \mathbf{B} = 0 \\ \mathbf{q} \cdot \mathbf{B} = 0 \end{cases} \quad (1.4)$$

and they can also be written in terms of the potentials, using the Helmholtz theorem:

$$\begin{cases} \nabla^2 \phi + \frac{1}{c} \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = -4\pi\rho & \mathbf{E} = -\nabla\phi + \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} = -i\mathbf{q}\phi - \frac{i\omega}{c} \mathbf{A} \\ \square \mathbf{A} - \nabla \left(\nabla \cdot \mathbf{A} + \frac{1}{c} \frac{\partial \phi}{\partial t} \right) = -\frac{4\pi}{c} \mathbf{J} & \mathbf{B} = \nabla \times \mathbf{A} = i\mathbf{q} \times \mathbf{A} \end{cases} \quad (1.5)$$

where ϕ is the scalar potential, \mathbf{A} the vector potential and $\square = \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}$ is the D'Alembert's operator. Notice that equations (1.4) must of course satisfy the continuity equation:

$$-\omega\rho + \mathbf{q} \cdot \mathbf{J} = 0 \quad (1.6)$$

The above scalar and vector potentials are not univocally defined, it is always possible to define an arbitrary scalar function $\Lambda(\mathbf{x}, t)$ such that:

$$\phi'(\mathbf{x}, t) = \phi(\mathbf{x}, t) - \frac{1}{c} \frac{\partial \Lambda(\mathbf{x}, t)}{\partial t} \quad \mathbf{A}'(\mathbf{x}, t) = \mathbf{A}(\mathbf{x}, t) - \nabla \Lambda(\mathbf{x}, t) \quad (1.7)$$

even if ϕ and \mathbf{A} are different from the new potentials ϕ' and \mathbf{A}' , they provide the same fields. This means that one can describe the same physical system with different choices of the potentials, i.e. within different *gauge representations*. For a deeper discussion see reference [3].

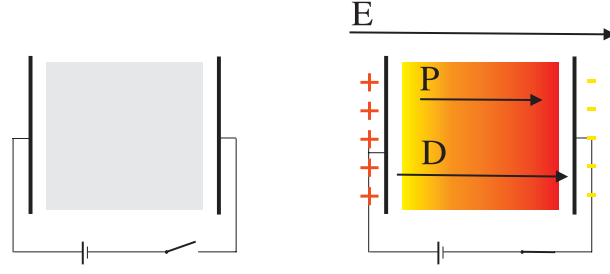


Figure 1.1: A medium placed in a capacitor. Applying a potential a field appear, between the capacitor plates, as an external perturbation \mathbf{D} for the medium. As a response, the medium polarizes giving raise to an induced field \mathbf{P} . The direction of \mathbf{P} is taken following the standard definition, the total field \mathbf{E} is defined as the difference between the two fields, see equation (1.13).

1.2 Maxwell's equations in a medium

A medium is nothing but a collection of rest or moving charges, and once its charge and current distributions are given, it is in principle possible to describe its physics uniquely by the solution of Maxwell's equations and dynamic equations (Hamilton's equations). Suppose to have a medium composed by n charges and m currents, a definition for the total charge and current densities can be given as follow:

$$\rho_{\text{MED}} = \sum_{i=1}^n \rho_i \quad \mathbf{J}_{\text{MED}} = \sum_{i=1}^m \mathbf{j}_i \quad (1.8)$$

First of all one needs to proceed from a microscopic description to a macroscopic one through some sort of averages on fields and sources:

$$\begin{aligned} \mathbf{E}_{\text{MED}}(\mathbf{q}) &= \langle \mathbf{e}_{\text{MED}}(\mathbf{q}) \rangle = \frac{1}{V} \int \mathbf{e}_{\text{MED}}(\mathbf{q}') f(\mathbf{q}', \mathbf{q}) d\mathbf{q}' \\ \eta_{\text{MED}}(\mathbf{q}) &= \langle \rho_{\text{MED}}(\mathbf{q}) \rangle = \frac{1}{V} \int \rho_{\text{MED}}(\mathbf{q}') f(\mathbf{q}', \mathbf{q}) d\mathbf{q}' \end{aligned} \quad (1.9)$$

here capital letters indicate macroscopic fields whereas the others indicate microscopic ones, $f(\mathbf{q}', \mathbf{q})$ is a generic convolution function and V is the volume occupied by the medium. The same averages hold for the magnetic induction and the current density. In the absence of external perturbations the average macroscopic electric field of a medium is zero, some ferromagnetic materials may have a non vanishing macroscopic magnetic field. If an external perturbation is switched on (see figure 1.1), the microscopic constituents of the medium are no more at equilibrium: they feel a new field and, interacting among them, they try to reach the new equilibrium configuration. This behaviour determines a no more vanishing value for the microscopic averages (1.9), the whole medium is able to raise a response against the external perturbation, the value of \mathbf{E}_{MED} takes the name of *polarization* \mathbf{P} and the value \mathbf{B}_{MED} is named *magnetization* \mathbf{M} . These macroscopic averaged fields are related to the corresponding induced charge and current by the Maxwell's equations:

$$\eta_{\text{MED}} = -\nabla \cdot \mathbf{P} \quad \mathbf{J}_{\text{MED}} = \frac{\partial \mathbf{P}}{\partial t} + c \nabla \times \mathbf{M} \quad (1.10)$$

The first term in the RHS of the current equation is the current due to the polarization of the medium, whereas the second is the current contribution due to the magnetization of the medium. Generally speaking one can distinguish between three different fields and sources:

Field and Source	Name	Symbol
\mathbf{E}_{EXT} and η_{EXT}	electric displacement, external field	\mathbf{D}
\mathbf{E}_{MED} and η_{MED}	polarization, induced field	\mathbf{P}
\mathbf{E}_{TOT} and η_{TOT}	electric field, total field	\mathbf{E}
\mathbf{B}_{EXT} and \mathbf{J}_{EXT}	magnetic displacement, external field	\mathbf{H}
\mathbf{B}_{MED} and \mathbf{J}_{MED}	magnetization, induced field	\mathbf{M}
\mathbf{B}_{TOT} and \mathbf{J}_{TOT}	magnetic induction, total field	\mathbf{B}

Even in this macroscopic formulation, the medium maintains the self-interaction, this is because microscopically, each charge feels both the external perturbation field and the induced field played on it by the rest of the medium, it reacts on its part, modifying the induced and the total field. Now the self-consistency is simply hidden inside \mathbf{P} and \mathbf{H} . Using definitions (1.10) and:

$$\eta_{\text{TOT}} = \eta_{\text{MED}} + \eta_{\text{EXT}} \quad \mathbf{J}_{\text{TOT}} = \mathbf{J}_{\text{MED}} + \mathbf{J}_{\text{EXT}} \quad (1.11)$$

one can write macroscopic Maxwell's equations for the total fields:

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{E}_{\text{TOT}} = 4\pi(\eta_{\text{EXT}} + \eta_{\text{MED}}) \\ \nabla \times \mathbf{B}_{\text{TOT}} - \frac{1}{c} \frac{\partial \mathbf{E}_{\text{TOT}}}{\partial t} = \frac{4\pi}{c} (\mathbf{J}_{\text{EXT}} + \mathbf{J}_{\text{MED}}) \\ \nabla \times \mathbf{E}_{\text{TOT}} + \frac{1}{c} \frac{\partial \mathbf{B}_{\text{TOT}}}{\partial t} = 0 \\ \nabla \cdot \mathbf{B}_{\text{TOT}} = 0 \end{array} \right.$$

$$\left\{ \begin{array}{l} \nabla \cdot (\mathbf{E} + 4\pi\mathbf{P}) = 4\pi\eta_{\text{EXT}} \\ \nabla \times (\mathbf{B} - 4\pi\mathbf{M}) - \frac{1}{c} \frac{\partial (\mathbf{E} + 4\pi\mathbf{P})}{\partial t} = \frac{4\pi}{c} \mathbf{J}_{\text{EXT}} \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \nabla \cdot \mathbf{B} = 0 \end{array} \right.$$

$$\left\{ \begin{array}{l} \nabla \cdot \mathbf{D} = 4\pi\eta_{\text{EXT}} \\ \nabla \times \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c} \mathbf{J}_{\text{EXT}} \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \nabla \cdot \mathbf{B} = 0 \end{array} \right. \quad (1.12)$$

where the trivial definitions:

$$\mathbf{E} = \mathbf{D} - 4\pi\mathbf{P} \quad \mathbf{H} = \mathbf{B} - 4\pi\mathbf{M} \quad (1.13)$$

have been introduced.

While a time dependent electric field can exist even without a magnetic field (it just can not be an electromagnetic wave), a time dependent magnetic field is always coupled with a time dependent electric field. This asymmetry in the behaviour of the two fields, allows to redefine (1.10): the response of the medium to the external magnetic field \mathbf{H} can be expressed in term of a current induced by the electric field \mathbf{Q} associated to the magnetization \mathbf{M} . In this way the magnetization disappear, \mathbf{H} is always equal to \mathbf{B} but the total induced current must be redefined:

$$\mathbf{J}_{\text{MED}} = \frac{\partial \mathbf{P}}{\partial t} + \frac{\partial \mathbf{Q}}{\partial t} \rightarrow \frac{\partial \mathbf{P}}{\partial t} \quad (1.14)$$

this implicitly gives a new definition for \mathbf{D} . Now equations (1.12) turn out to be:

$$\begin{cases} \nabla \cdot \mathbf{D} = 4\pi\eta_{\text{EXT}} \\ \nabla \times \mathbf{H} - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c} \mathbf{J}_{\text{EXT}} \\ \nabla \times \mathbf{E} + \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} = 0 \\ \nabla \cdot \mathbf{H} = 0 \end{cases} \quad (1.15)$$

where the total magnetic field and the magnetization disappeared. Systems (1.12) and (1.15) can be written in terms of \mathbf{P} and \mathbf{D} or \mathbf{P} and \mathbf{E} .

By means of the linear response theory (described in appendix A), the response \mathbf{P} of the medium can be related to the external perturbations \mathbf{D} or \mathbf{E} gaining a system of four equations and four incognita. To this aim, notation (1.15) is more convenient than (1.12): the former needs just one response function to relate the electric perturbation to the electric response whereas the latter needs also a relation between the magnetic perturbation and the magnetic response. The hamiltonian of a system with total density $\eta(q, t)_{\text{TOT}}$ and current density \mathbf{J}_{TOT} perturbed by an external field is given by:

$$\mathcal{H} = \mathcal{H}_0 + \int d\mathbf{x} \eta(\mathbf{x}, t)_{\text{TOT}} \phi(\mathbf{x}, t)_{\text{EXT}} - \frac{1}{c} \int d\mathbf{x} \mathbf{J}(\mathbf{x}, t)_{\text{TOT}} \cdot \mathbf{A}(\mathbf{x}, t)_{\text{EXT}} \quad (1.16)$$

where ϕ_{EXT} is the scalar potential associated to \mathbf{D} and \mathbf{A}_{EXT} is the vector potential associated to \mathbf{H} . \mathcal{H}_0 contains the many body interaction between the electrons through the microscopic Coulomb's potential and the Coulomb interaction with the ions of the periodic lattice of a crystal. If the perturbation is a static field or a time dependent electric field, only the scalar or vector potential must be used. If the perturbation is a time dependent magnetic field or an electromagnetic wave, both potentials must be included in the hamiltonian.

Choosing the induced current as the response observable, in the linear response theory formula (1.170) of appendix A $B(\mathbf{x}, t) = \mathbf{J}_{\text{MED}}(\mathbf{x}, t)$, under the assumption that in the unperturbed system $\mathbf{J}_{\text{MED}}(\mathbf{x}, t) = 0$, one gets:

$$\begin{aligned} \mathbf{J}_{\text{MED}}(\mathbf{x}, t) &= \frac{1}{c} \int d\mathbf{x}' \int_{-\infty}^t dt' \mathbf{A}_{\text{EXT}}(\mathbf{x}', t') \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') - \\ &\quad - \int d\mathbf{x}' \int_{-\infty}^t dt' \phi_{\text{EXT}}(\mathbf{x}', t') \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', t - t') \\ \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') &= \langle \{ \mathbf{J}_{\text{TOT}}(\mathbf{x}, t'), \mathbf{J}_{\text{MED}}(\mathbf{x}, t - t') \} \rangle \\ \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', t - t') &= \langle \{ \eta_{\text{TOT}}(\mathbf{x}, t'), \mathbf{J}_{\text{MED}}(\mathbf{x}, t - t') \} \rangle \end{aligned} \quad (1.17)$$

Here the response functions are the *current-current* correlation function and the *charge-current* correlation function, the first is a tensor, the latter is a vector, exactly as in definitions (1.175).

Using the gauge invariance it is possible to find a relation between $\phi_{\mathbf{J}\mathbf{J}}$ and $\phi_{\rho\mathbf{J}}$ in order to rewrite the induced charge in term of the current-current response function only. One can define the new potentials \mathbf{A}'_{EXT} and ϕ'_{EXT} by means of equations (1.7), however the new fields must give an induced current densities \mathbf{J}'_{MED} equal to \mathbf{J}_{MED} of equation (1.17), from this equality the following condition results:

$$- \int d\mathbf{x}' \int_{-\infty}^t dt' \nabla \Lambda(\mathbf{x}', t') \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') + \int d\mathbf{x}' \int_{-\infty}^t dt' \frac{\partial \Lambda(\mathbf{x}', t')}{\partial t'} \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', t - t') = 0 \quad (1.18)$$

Now one has to integrate by parts in the t' variable the second integral in the LHS, using also the properties (1.171). While for the first integral one must use the relation:

$$\nabla \cdot (\phi_{\text{EXT}} \phi_{\mathbf{J}\rho}) = \phi_{\mathbf{J}\rho} \cdot \nabla \phi_{\text{EXT}} + \phi_{\text{EXT}} \nabla \cdot \phi_{\mathbf{J}\rho} \quad (1.19)$$

together with the divergence theorem:

$$\int d\mathbf{x} \nabla \cdot (\phi_{\text{EXT}} \phi_{\mathbf{J}\rho}) = \int \phi_{\text{EXT}} \phi_{\mathbf{J}\rho} \cdot \mathbf{n} d\Sigma = 0 \quad (1.20)$$

where for the last step the properties (1.171) have been used again. Now equation (1.18) looks like:

$$\int d\mathbf{x}' \int_{-\infty}^t dt' \left(\nabla \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') - \frac{\partial \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', t - t')}{\partial t'} \right) \Lambda(\mathbf{x}', t') = 0 \quad (1.21)$$

that must be valid for any arbitrary Λ function, this leads to:

$$\begin{aligned} \nabla \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') &= \frac{\partial \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', t - t')}{\partial t'} \\ \nabla \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) &= -i\omega \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', \omega) \end{aligned} \quad (1.22)$$

Thanks to the $t - t'$ time dependency, due to the causality principle, Fourier's transform of equation (1.17) is suitable for the usage of condition (1.22):

$$\begin{aligned} \mathbf{J}_{\text{MED}}(\mathbf{x}, \omega) &= \frac{1}{c} \int d\mathbf{x}' \mathbf{A}_{\text{EXT}}(\mathbf{x}', \omega) \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) - \int d\mathbf{x}' \phi_{\text{EXT}}(\mathbf{x}', \omega) \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', \omega) = \\ &= \frac{1}{c} \int d\mathbf{x}' \mathbf{A}_{\text{EXT}}(\mathbf{x}', \omega) \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) + \int d\mathbf{x}' \phi_{\text{EXT}}(\mathbf{x}', \omega) \frac{\nabla \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega)}{i\omega} \end{aligned} \quad (1.23)$$

now one has to use again equations (1.19) and (1.20) to eliminate the divergence of the response function:

$$\begin{aligned} \mathbf{J}_{\text{MED}}(\mathbf{x}, \omega) &= \frac{1}{c} \int d\mathbf{x}' \mathbf{A}_{\text{EXT}}(\mathbf{x}', \omega) \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) + \\ &+ \frac{1}{i\omega} \int d\mathbf{x}' \nabla \phi_{\text{EXT}}(\mathbf{x}', \omega) \cdot \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) = \\ &= \frac{1}{i\omega} \int d\mathbf{x}' \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) \cdot \left(\frac{i\omega}{c} \mathbf{A}_{\text{EXT}}(\mathbf{x}', \omega) + \nabla \phi_{\text{EXT}}(\mathbf{x}', \omega) \right) = \\ &= -\frac{1}{i\omega} \int d\mathbf{x}' \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \end{aligned} \quad (1.24)$$

In the last step the general definition of electric field (1.5) has been used. The response function σ , relating the induced current and the external field, is named *conductivity* tensor, and is defined has:

$$\mathbf{J}_{\text{MED}}(\mathbf{x}, \omega) = \int d\mathbf{x}' \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \quad (1.25)$$

Now also the induced current can be expressed in terms of the field, one has to use the generalized definition (1.14):

$$\begin{aligned} -i\omega \mathbf{P}(\mathbf{x}, \omega) &= \frac{i}{\omega} \int d\mathbf{x}' \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \\ \mathbf{P}(\mathbf{x}, \omega) &= -\frac{1}{\omega^2} \int d\mathbf{x}' \phi_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \\ \mathbf{P}(\mathbf{x}, \omega) &= \int d\mathbf{x}' \chi(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \end{aligned} \quad (1.26)$$

where the response function χ relates the induced and the external fields, its name is *electric susceptibility* tensor. Both σ and χ are tensors, the latter is dimensionless because it relates two quantities with the same dimensions, from their definitions:

$$\chi(\mathbf{x}, \mathbf{x}', \omega) = \frac{i}{\omega} \sigma(\mathbf{x}, \mathbf{x}', \omega) \quad (1.27)$$

Another possibility is to link the total and the external fields, one has just to combine the two general definitions (1.13) and (1.14):

$$4\pi \mathbf{J}_{\text{MED}} = \frac{\partial \mathbf{D}}{\partial t} - \frac{\partial \mathbf{E}}{\partial t} \quad (1.28)$$

and to use this new relation into equation (1.25):

$$\begin{aligned} -i\omega \left(\mathbf{D}(\mathbf{x}, \omega) - \mathbf{E}(\mathbf{x}, \omega) \right) &= 4\pi \int d\mathbf{x}' \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \\ \mathbf{E}(\mathbf{x}, \omega) &= \mathbf{D}(\mathbf{x}, \omega) + \frac{4\pi}{i\omega} \int d\mathbf{x}' \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \\ \mathbf{E}(\mathbf{x}, \omega) &= \int d\mathbf{x}' \left(\mathbf{1}\delta(x - x') + \frac{4\pi}{i\omega} \sigma(\mathbf{x}, \mathbf{x}', \omega) \right) \cdot \mathbf{D}(\mathbf{x}', \omega) \\ \mathbf{E}(\mathbf{x}, \omega) &= \int d\mathbf{x}' \epsilon^{-1}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) \end{aligned} \quad (1.29)$$

Where the new *inverse dielectric tensor* has been introduced:

$$\epsilon^{-1}(\mathbf{x}, \mathbf{x}', \omega) = \mathbf{1}\delta(x - x') - \frac{i4\pi}{\omega} \sigma(\mathbf{x}, \mathbf{x}', \omega) = \mathbf{1}\delta(x - x') - 4\pi\chi(\mathbf{x}, \mathbf{x}', \omega) \quad (1.30)$$

the $\mathbf{1}$ symbol denotes the identity matrix. Each one of the above response functions can be measured with a different kind of experiment, probing each time the desired response observable and tuning the perturbation observable.

With respect to equation (1.170), it is possible to choose $B(\mathbf{x}, t) = \eta(\mathbf{x}, t)_{\text{MED}}$, the unperturbed condition is $\eta_{\text{MED}}(\mathbf{x}, t) = 0$, this means that $\langle \Delta B(\mathbf{x}, t) \rangle = \eta_{\text{MED}}(\mathbf{x}, t)$, the response functions are the *charge-charge* correlation function and the *current-charge* correlation function:

$$\begin{aligned} \eta_{\text{MED}}(\mathbf{x}, t) &= \frac{1}{c} \int d\mathbf{x}' \int_{-\infty}^t dt' \mathbf{A}_{\text{EXT}}(\mathbf{x}', t') \cdot \phi_{\mathbf{J}\rho}(\mathbf{x}, \mathbf{x}', t - t') - \\ &\quad - \int d\mathbf{x}' \int_{-\infty}^t dt' \phi_{\text{EXT}}(\mathbf{x}', t') \phi_{\rho\rho}(\mathbf{x}, \mathbf{x}', t - t') \\ \phi_{\rho\rho}(\mathbf{x}, \mathbf{x}', t - t') &= \langle \{ \eta_{\text{TOT}}(\mathbf{x}, t'), \eta_{\text{MED}}(\mathbf{x}, t - t') \} \rangle \\ \phi_{\rho\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') &= \langle \{ \mathbf{J}_{\text{TOT}}(\mathbf{x}, t'), \eta_{\text{MED}}(\mathbf{x}, t - t') \} \rangle \end{aligned} \quad (1.31)$$

Now the charge-charge correlation function is a scalar whereas the charge-current correlation function is a vector. Using the gauge invariance it is possible to find a relation like (1.22) between $\phi_{\rho\rho}$ and $\phi_{\rho\mathbf{J}}$:

$$\nabla \cdot \phi_{\rho\mathbf{J}}(\mathbf{x}, \mathbf{x}', t - t') = \frac{\partial \phi_{\rho\rho}(\mathbf{x}, \mathbf{x}', t - t')}{\partial t'} \quad (1.32)$$

now the charge-current correlation function, one needs to replace, appears in the first integral of equation (1.31) in a scalar product together with the vector potential, one can no longer use

relation (1.19). The substitution of equation (1.32) into (1.31), in order to express everything in terms of $\phi_{\rho\rho}$ only, is not reliable. The only possibility is to put $\mathbf{A}_{\text{EXT}} = 0$ limiting the problem to oscillating electric fields only, i.e. excluding electromagnetic waves in favor of longitudinal perturbations only (see appendix B). Within this assumption the first integral in the RHS of equation (1.31) vanishes giving:

$$\eta_{\text{MED}}(\mathbf{x}, t) = - \int d\mathbf{x}' \int_{-\infty}^t dt' \phi_{\text{EXT}}(\mathbf{x}', t') \phi_{\rho\rho}(\mathbf{x}, \mathbf{x}', t - t') \quad (1.33)$$

Notice that, thanks to definition (1.5), one can state that

$$\mathbf{D}(\mathbf{q}', \omega) = -i\mathbf{q}' \phi_{\text{EXT}}(\mathbf{q}', \omega) \quad \phi_{\text{EXT}}(\mathbf{q}', \omega) = \frac{i}{q'} \mathbf{D} \cdot \hat{\mathbf{q}}' \quad (1.34)$$

and using also the Fourier's transform of (1.14) one gets:

$$\begin{aligned} -iq\hat{\mathbf{q}} \cdot \mathbf{P}(\mathbf{q}, \omega) &= \int d\mathbf{q}' \chi_{\rho\rho}(\mathbf{q}, \mathbf{q}', \omega) \frac{-i\hat{\mathbf{q}}' \cdot \mathbf{D}(\mathbf{q}', \omega)}{q'} \\ \hat{\mathbf{q}} \cdot \mathbf{P}(\mathbf{q}, \omega) &= \int d\mathbf{q}' \frac{\chi_{\rho\rho}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \hat{\mathbf{q}}' \cdot \mathbf{D}(\mathbf{q}', \omega) \end{aligned} \quad (1.35)$$

As expected, the response function relates only the projections of perturbation and response on their momentum directions, i.e. the longitudinal components of the response and the perturbation. Notice that, the response function relates now two identical quantities, because of this it must be dimensionless, the division by q and q' satisfies this need. With a more compact notation:

$$P^{\text{L}}(\mathbf{q}, \omega) = \int d\mathbf{q}' \chi^{\text{LL}}(\mathbf{q}, \mathbf{q}', \omega) D^{\text{L}}(\mathbf{q}', \omega) \quad (1.36)$$

χ^{LL} takes the name of *longitudinal electric susceptibility* and it is just a component of the previously defined dielectric tensor (see appendix B). It is important to stress that the derivation of equations (1.36), (1.26) and (1.25) contains only the very general assumption of gauge invariance and it is valid under any gauge condition.

All the properties of a linear medium are hidden inside its response functions and a classification of different media can be done on the basis of response functions general properties:

ϕ	χ	static response	the medium is homogeneous and it behaves in the same way at all frequencies
$\phi(t - t')$	$\chi(\omega)$	dynamic response	the medium gives different response at different frequencies
$\phi(\mathbf{x} - \mathbf{x}', t - t')$	$\chi(\mathbf{q} - \mathbf{q}', \omega)$	translational invariance	the medium response is non local but again homogeneous
$\phi(\mathbf{x}, \mathbf{x}', t - t')$	$\chi(\mathbf{q}, \mathbf{q}', \omega)$	local effects	the medium response is non local and inhomogeneous

1.3 On the tensorial nature of the response functions

The tensorial character of a medium comes from its microscopic atomistic structure. In a crystalline solid, the symmetry properties of the atomic lattice determine a different behaviour along different directions of the response to external perturbations. In the previous

section a classification of the media in terms of the spatial and frequency dependences of the response function has been given. Another classification is possible, on the basis of the tensorial properties of the response functions. In the most general case, the tensor relating the perturbation to the response, is a complex tensor with symmetric real and imaginary parts. The symmetry assure the diagonalizability of the two parts of the tensor, but the principal axes¹ of the real part can be different from the ones of the imaginary part, the whole tensor is not diagonalizable. For an *anisotropic* medium in which the reference system is taken in such a way that the perturbation wave vector \mathbf{q} lies along the x -axis, the conductivity tensor appears as:

$$\sigma = \begin{pmatrix} \sigma^{LL} & \sigma^{LT_1} & \sigma^{LT_2} \\ \sigma^{LT_1} & \sigma^{T_1T_1} & \sigma^{T_1T_2} \\ \sigma^{LT_2} & \sigma^{T_1T_2} & \sigma^{T_2T_2} \end{pmatrix} \quad (1.37)$$

Diagonal components relate longitudinal perturbations to longitudinal responses, or transverse perturbations to transverse responses, two different transverse response exist because of the two independent polarizations of light. In an anisotropic medium a longitudinal perturbation can also induce a transverse response, and vice versa, this behaviour is accounted by the off-diagonal components. This can be formally shown applying the dyadic notation introduced in appendix B, multiplying for the longitudinal dyad the general definition (1.25):

$$\begin{aligned} \mathbf{1}^L \cdot \mathbf{J}_{\text{MED}}(\mathbf{x}, \omega) &= \mathbf{1}^L \cdot \int d\mathbf{x}' \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) = \\ &= \int d\mathbf{x}' \mathbf{1}^L \cdot \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot [\mathbf{1}^L \cdot \mathbf{D}(\mathbf{x}', \omega) + \mathbf{1}^T \cdot \mathbf{D}(\mathbf{x}', \omega)] = \\ &= \int d\mathbf{x}' [\mathbf{1}^L \cdot \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{1}^L \cdot \mathbf{D}(\mathbf{x}', \omega) + \\ &+ \mathbf{1}^L \cdot \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{1}^T \cdot \mathbf{D}(\mathbf{x}', \omega)] \end{aligned} \quad (1.38)$$

so even a transverse field can contribute to the longitudinal induced current. The same holds for the transverse induced current and the longitudinal perturbation, it is sufficient to multiply definition (1.25) by the transverse dyad.

Another interesting anisotropic medium is the one described by the conductivity tensor:

$$\sigma = \begin{pmatrix} \sigma^{LL} & 0 & 0 \\ 0 & \sigma^{T_1T_1} & \sigma^{T_1T_2} \\ 0 & \sigma^{T_1T_2} & \sigma^{T_2T_2} \end{pmatrix} \quad (1.39)$$

Now transverse and longitudinal components are decoupled, a transverse perturbation can induce only a transverse response, but the medium is still able to mix the two components of the transverse field. This is the typical behaviour of the so called *active* media, media able to change the polarization of the light passing through them.

For *cubic* media a set of principal axes always exist that diagonalizes both the real and imaginary parts, the conductivity tensor becomes:

$$\sigma = \begin{pmatrix} \sigma^{LL} & 0 & 0 \\ 0 & \sigma^{T_1T_1} & 0 \\ 0 & 0 & \sigma^{T_2T_2} \end{pmatrix} \quad (1.40)$$

chiral media, for instance, are properly described by a tensor like this. They have the property of giving a different response depending on the polarization of the impinging light. If $\sigma^{T_1T_1} = \sigma^{T_2T_2}$ the medium is *uniaxial*, in this case only two different kind of response

¹Basis set describing a vector space in which the tensor is diagonal.

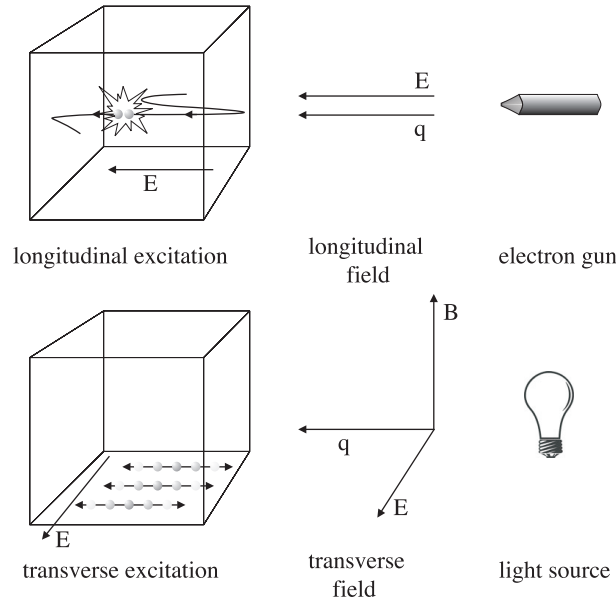


Figure 1.2: Pictorial representation of transverse and longitudinal modes of a medium.

exist. The simplest case is that of an *isotropic medium*:

$$\sigma = \begin{pmatrix} \sigma & 0 & 0 \\ 0 & \sigma & 0 \\ 0 & 0 & \sigma \end{pmatrix} \quad (1.41)$$

that shows the same behaviour for any kind of perturbation. Although the isotropic medium model seems an over simplification of the reality, in fact to small wave vector perturbations of any kind, each cubic medium appears homogeneous. In the *long wavelength limit*, also called *optical limit*, of $\mathbf{q} \rightarrow 0$ it is impossible to distinguish between transverse and longitudinal perturbations or responses, because of this, $\sigma^{\text{T}_1\text{T}_1} = \sigma^{\text{T}_2\text{T}_2} = \sigma^{\text{T}_1\text{T}_1} = \sigma^{\text{LL}}$. This is the typical situation of optical spectroscopy in which, if infrared, visible or ultraviolet light is used, the momentum of the impinging photons and the momentum transferred to the medium are negligible. In the case of cubic media, the optical response, that is a transverse one, is identical to the longitudinal response for $\mathbf{q} \rightarrow 0$.

1.4 Reducible and irreducible response

If in the hamiltonian (1.16) the external potentials ϕ_{EXT} and \mathbf{A}_{EXT} are employed, the microscopic Coulomb interaction among medium charges is hidden inside \mathcal{H}_0 . If the total fields are employed, the many-particle interaction energy is taken out from \mathcal{H}_0 and hidden inside the total potentials ϕ_{TOT} and \mathbf{A}_{TOT} that must be used in the perturbation terms instead of ϕ_{EXT} and \mathbf{A}_{EXT} . This is just a different way to describe the same physics of the medium-fields interaction, the total perturbation energy must be the same in both cases as well as the induced and total charge and current densities. The only difference is that, if the unperturbed hamiltonian \mathcal{H}_0 changes its form, the ensemble distribution function $\rho(t)$ changes. $\rho(t)$ enters the definition of statistical average of equation (1.167) determining a different definition for the correlation function i.e. for the response function. Definitions (1.17) and (1.31) for the response functions already hold, but now the statistical average is made on a different configuration space. Using this new approach involving total potentials and choos-

ing again the induced current density as the response observable, one can follow exactly the same procedure described in section 1.2 to get the analogous of equation (1.24):

$$\begin{aligned}\mathbf{J}_{\text{MED}}(\mathbf{x}, \omega) &= -\frac{1}{i\omega} \int d\mathbf{x}' \bar{\phi}_{\mathbf{J}\mathbf{J}}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) \\ \mathbf{J}_{\text{MED}}(\mathbf{x}, \omega) &= \int d\mathbf{x}' \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega)\end{aligned}\quad (1.42)$$

where $\bar{\phi}_{\mathbf{J}\mathbf{J}}$ denotes a response function defined in the same way as in (1.17) but with a different ensemble average. A relation can be found between σ and $\bar{\sigma}$. The total field contains the induced field, because of this, it is an averaged microscopic field, in equation (1.42) it accounts implicitly for the self-consistency in the response of the medium. Equation (1.25) contains the external field, that is a macroscopic field, in this case is the response function that must account for the microscopic self-interaction. χ , also called *reducible susceptibility*, does not contains the self-interaction contribution, $\bar{\chi}$, the *irreducible susceptibility*, contains the self-interaction properties of the medium. To relate the two response functions one must start equating the two definitions of the induced current density (1.25) and (1.42):

$$\int d\mathbf{x}' \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) = \int d\mathbf{x}' \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) \quad (1.43)$$

the total field \mathbf{E} can be written using equation

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{x}, \omega) - \frac{\omega^2}{c^2} \mathbf{E}(\mathbf{x}, \omega) = i\omega \frac{4\pi}{c^2} \mathbf{J}_{\text{TOT}}(\mathbf{x}, \omega) \quad (1.44)$$

derived from the combination of the second and third Maxwell's equations (first system of 1.12). In fact, using the Green approach, one can write the solution of differential equation (1.44) as:

$$\mathbf{E}(\mathbf{x}, \omega) = i\omega \frac{4\pi}{c^2} \int \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{J}_{\text{TOT}}(\mathbf{x}', \omega) d\mathbf{x}' \quad (1.45)$$

where \mathbf{G} is the Green's operator solution of the homogeneous equation:

$$\nabla \times \nabla \times \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) - \frac{\omega^2}{c^2} \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) = \delta(\mathbf{x} - \mathbf{x}') \quad (1.46)$$

Using (1.45) into (1.43) one gets:

$$\begin{aligned}\int d\mathbf{x}' \sigma(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{D}(\mathbf{x}', \omega) &= i\omega \frac{4\pi}{c^2} \iint d\mathbf{x}' \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{G}(\mathbf{x}', \mathbf{x}_1, \omega) \cdot \\ &\cdot \mathbf{J}_{\text{TOT}}(\mathbf{x}_1, \omega) d\mathbf{x}_1 = \\ &= i\omega \frac{4\pi}{c^2} \iint \left[\bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{G}(\mathbf{x}', \mathbf{x}_1, \omega) \cdot \mathbf{J}_{\text{EXT}}(\mathbf{x}_1, \omega) + \right. \\ &\left. + \int \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{G}(\mathbf{x}', \mathbf{x}_1, \omega) \cdot \sigma(\mathbf{x}_1, \mathbf{x}_2, \omega) \mathbf{D}(\mathbf{x}_2, \omega) d\mathbf{x}_2 \right] d\mathbf{x}' d\mathbf{x}_1\end{aligned}\quad (1.47)$$

in the last step decomposition (1.11) together with definition (1.42) has been used to eliminate the total current density. Now the first term in square brackets is the definition of external field \mathbf{D} through the Green's operator:

$$\begin{aligned}\nabla \times \nabla \times \mathbf{D}(\mathbf{x}, \omega) - \frac{\omega^2}{c^2} \mathbf{D}(\mathbf{x}, \omega) &= i\omega \frac{4\pi}{c^2} \mathbf{J}_{\text{EXT}}(\mathbf{x}, \omega) \\ \mathbf{D}(\mathbf{x}, \omega) &= i\omega \frac{4\pi}{c^2} \int \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{J}_{\text{EXT}}(\mathbf{x}', \omega) d\mathbf{x}'\end{aligned}\quad (1.48)$$

In in the second integral of (1.47) \mathbf{x}' , \mathbf{x}_1 and \mathbf{x}_2 are merely integration variable and one can make the exchange $\mathbf{x}_2 \leftrightarrow \mathbf{x}'$ to gain a symmetric and intuitive notation. Now it is possible to factorize the expression as follow

$$\int d\mathbf{x}' \left\{ \sigma(\mathbf{x}, \mathbf{x}', \omega) - \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) - i\omega \frac{4\pi}{c^2} \iint \bar{\sigma}(\mathbf{x}, \mathbf{x}_2, \omega) \cdot \mathbf{G}(\mathbf{x}_2, \mathbf{x}_1, \omega) \cdot \sigma(\mathbf{x}_1, \mathbf{x}', \omega) d\mathbf{x}_1 d\mathbf{x}_2 \right\} \cdot \mathbf{D}(\mathbf{x}', \omega) = 0 \quad (1.49)$$

that must be true for any arbitrary external field $\mathbf{D} \neq 0$. This leads to:

$$\sigma(\mathbf{x}, \mathbf{x}', \omega) = \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) + i\omega \frac{4\pi}{c^2} \iint \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{G}(\mathbf{x}', \mathbf{x}_1, \omega) \cdot \sigma(\mathbf{x}_1, \mathbf{x}_2, \omega) d\mathbf{x}_1 d\mathbf{x}_2 \quad (1.50)$$

The two response functions are related by a Dyson's like equation that can be solved self-consistently. Proceeding in this way the irreducible conductivity $\bar{\sigma}$ is modified, at every iteration, including a part of the self-interaction properties of the electrons of the medium. The iterative procedure converges to the reducible conductivity σ that is defined in such a way that it contains implicitly the self-interaction of the electrons of the medium. With a simple substitution of their general definitions into (1.50), one can relates χ and $\bar{\chi}$:

$$\chi(\mathbf{x}, \mathbf{x}', \omega) = \bar{\chi}(\mathbf{x}, \mathbf{x}', \omega) + \frac{4\pi\omega^2}{c^2} \iint \bar{\chi}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{G}(\mathbf{x}', \mathbf{x}_1, \omega) \cdot \chi(\mathbf{x}_1, \mathbf{x}_2, \omega) d\mathbf{x}_1 d\mathbf{x}_2 \quad (1.51)$$

This last equation can be directly derived by a many body perturbation theory, the result is exactly the same but \mathbf{G} becomes the retarded Green's function (see section 1.10).

Equation (1.51) takes a more familiar form if one restrict his interest to the longitudinal response. In this situation one deals with scalar equations only and, in reciprocal space, all the equations becomes scalar. If one starts from equation (1.33) and its equivalent in terms of the irreducible response

$$\begin{aligned} \eta_{\text{MED}}(\mathbf{q}, \omega) &= - \int d\mathbf{q}' \phi_{\text{EXT}}(\mathbf{q}', \omega) \phi_{\rho\rho}(\mathbf{q}, \mathbf{q}', \omega) \\ \eta_{\text{MED}}(\mathbf{q}, \omega) &= - \int d\mathbf{q}' \phi_{\text{TOT}}(\mathbf{q}', \omega) \bar{\phi}_{\rho\rho}(\mathbf{q}, \mathbf{q}', \omega) \end{aligned} \quad (1.52)$$

and follows the same procedure used for the case of (1.42) one gets:

$$\phi_{\rho\rho}(\mathbf{q}, \mathbf{q}', \omega) = \bar{\phi}_{\rho\rho}(\mathbf{q}, \mathbf{q}', \omega) - \int d\mathbf{q}_1 \bar{\phi}_{\rho\rho}(\mathbf{q}, \mathbf{q}_1, \omega) \frac{4\pi}{\mathbf{q}_1^2} \phi_{\rho\rho}(\mathbf{q}_1, \mathbf{q}', \omega) \quad (1.53)$$

where the trivial definition

$$\phi_{\text{TOT}} = \phi_{\text{EXT}} + \phi_{\text{MED}} \quad (1.54)$$

has been used, together with the Poisson's equation.

$$\phi_{\text{MED}}(\mathbf{q}, \omega) = \frac{4\pi}{\mathbf{q}^2} \eta_{\text{MED}}(\mathbf{q}, \omega) \quad (1.55)$$

Here the Green's operator is nothing but the bare Coulomb's potential. Figure 1.3 represent equation (1.53) in the Feynman diagrammatic representation.

1.5 Macroscopic and microscopic response

Using again the approach based on total potentials exposed in the preceding section, one can derive new susceptibility and dielectric tensors. As in section 1.2:

$$\mathbf{P}(\mathbf{x}, \omega) = \int d\mathbf{x}' \bar{\chi}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) \quad (1.56)$$

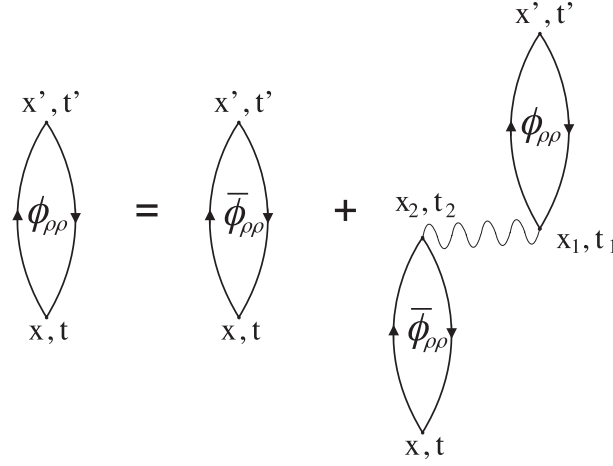


Figure 1.3: Feynman diagram for Dyson's like equation of electric susceptibility. The reducible susceptibility is expressed in terms of the irreducible one and the Coulomb potential, in a self-consistent equation.

Combining equations (1.42) and (1.28) one can define the *dielectric tensor*:

$$\mathbf{D}(\mathbf{x}, \omega) = \int d\mathbf{x}' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) \quad (1.57)$$

$$\bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) = \mathbf{1}\delta(\mathbf{x} - \mathbf{x}') + \frac{i4\pi}{\omega} \bar{\sigma}(\mathbf{x}, \mathbf{x}', \omega) = \mathbf{1}\delta(\mathbf{x} - \mathbf{x}') + 4\pi\bar{\chi}(\mathbf{x}, \mathbf{x}', \omega)$$

It must be stressed that formally speaking this quantity is the inverse of the response function defined in (1.29), even if the two response functions are defined by means of a different average procedure. Replacing (1.29) in (1.57) and factorizing the external field, one gets:

$$\int d\mathbf{x}'' \left[\mathbf{1}\delta(\mathbf{x} - \mathbf{x}'') - \int d\mathbf{x}' \epsilon^{-1}(\mathbf{x}, \mathbf{x}', \omega) \cdot \bar{\epsilon}(\mathbf{x}', \mathbf{x}'', \omega) \right] \cdot \mathbf{D}(\mathbf{x}'', \omega) = 0 \quad (1.58)$$

that must be true for any arbitrary \mathbf{D} , this means:

$$\int d\mathbf{x}' \epsilon^{-1}(\mathbf{x}, \mathbf{x}', \omega) \cdot \bar{\epsilon}(\mathbf{x}', \mathbf{x}'', \omega) = \mathbf{1}\delta(\mathbf{x} - \mathbf{x}'') \quad (1.59)$$

that is simply the definition of inverse operator. In principle measuring one of the two response functions it is always possible calculate the other one. One may think to measure the function $\epsilon^{-1}(\mathbf{q}, \mathbf{q}', \omega)$ and, inverting it, to get $\bar{\epsilon}(\mathbf{q}, \mathbf{q}', \omega)$. Practically $\bar{\epsilon}(\mathbf{q}, \mathbf{q}', \omega)$ and $\epsilon^{-1}(\mathbf{q}, \mathbf{q}', \omega)$ are microscopic, non observable quantities. One can only measure some sort of macroscopic average response functions, they will be defined in the following and their relation will be discussed.

In the specific case of crystals, thanks to the spatial periodicity one can say that $\mathbf{q} = \mathbf{K} + \mathbf{G}$ and $\mathbf{q}' = \mathbf{K} + \mathbf{G}'$ where \mathbf{K} belong to the first Brillouin's zone whereas \mathbf{G} and \mathbf{G}' are vector of the reciprocal lattice.

$$\begin{aligned} \epsilon^{-1}(\mathbf{q}, \mathbf{q}', \omega) &\longrightarrow \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}', \omega) \\ \bar{\epsilon}(\mathbf{q}, \mathbf{q}', \omega) &\longrightarrow \bar{\epsilon}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}', \omega) \end{aligned} \quad (1.60)$$

With this choice, definition (1.29) becomes:

$$\mathbf{E}(\mathbf{K} + \mathbf{G}, \omega) = \sum_{\mathbf{G}'} \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}', \omega) \cdot \mathbf{D}(\mathbf{K} + \mathbf{G}', \omega) \quad (1.61)$$

typically $\mathbf{G} \gg \mathbf{K}$, or in real space, $\lambda_{\mathbf{G}} \ll \lambda_{\mathbf{K}}$, this means that the \mathbf{K} Fourier's components account for length scales larger than the lattice one, whereas $\mathbf{K} + \mathbf{G}$ Fourier's components describe small oscillations of the same wavelength of the lattice parameter. Macroscopic quantities can not be measured with such a resolution, they can not depend upon lattice parameter scale oscillations. Because of this they will depend only upon \mathbf{K} , i.e. macroscopic fields are $\mathbf{E}(\mathbf{K})$ and $\mathbf{D}(\mathbf{K})$. Also external charge and current densities must be macroscopic, so one has $\mathbf{J}_{\text{EXT}}(\mathbf{K} + \mathbf{G}) = \eta_{\text{EXT}}(\mathbf{K} + \mathbf{G}) = 0$ for each $\mathbf{G} \neq 0$. By now the ω dependence will be omitted to keep the notation simple. Combining the second and the third Maxwell's equations of system (1.15), and using the analogous of identity (1.46) for the electric field \mathbf{E} , one gets:

$$\mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \mathbf{E}(\mathbf{K} + \mathbf{G}) - \frac{\omega^2}{c^2} \mathbf{D}(\mathbf{K} + \mathbf{G}) = \frac{i\omega 4\pi}{c^2} \mathbf{J}_{\text{EXT}}(\mathbf{K} + \mathbf{G}) \quad (1.62)$$

where the \mathbf{O} operator is given by:

$$\mathbf{O}(\mathbf{K} + \mathbf{G}) = \left[(\mathbf{K} + \mathbf{G})(\mathbf{K} + \mathbf{G}) - (\mathbf{K} + \mathbf{G})^2 \mathbf{1} \right] \quad (1.63)$$

Now using (1.61) one can write

$$\sum_{\mathbf{G}'} \left[\mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}') - \frac{\omega^2}{c^2} \mathbf{1}_{\delta_{\mathbf{G}, \mathbf{G}'}} \right] \cdot \mathbf{D}(\mathbf{K} + \mathbf{G}') = \frac{i\omega 4\pi}{c^2} \mathbf{J}_{\text{EXT}}(\mathbf{K} + \mathbf{G}) \quad (1.64)$$

writing explicitly the $\mathbf{G} = \mathbf{G}'$ term of the sum and taking $\mathbf{G} = 0$ one gets:

$$\sum_{\mathbf{G}' \neq 0} \left[\mathbf{O}(\mathbf{K}) \cdot \epsilon^{-1}(\mathbf{K}, \mathbf{K} + \mathbf{G}') \right] \cdot \mathbf{D}(\mathbf{K} + \mathbf{G}') + \left[\mathbf{O}(\mathbf{K}) \cdot \epsilon^{-1}(\mathbf{K}, \mathbf{K}) \right] \cdot \mathbf{D}(\mathbf{K}) = \frac{i\omega 4\pi}{c^2} \mathbf{J}_{\text{EXT}}(\mathbf{K}) \quad (1.65)$$

In this expression the external perturbation \mathbf{J}_{EXT} is macroscopic, in the sense that it does not depend upon the crystal microscopic structure described by the \mathbf{G} vectors. Despite of this, the external field \mathbf{D} felt locally by the charges of the medium, involves all the Fourier's components $\mathbf{K} + \mathbf{G}$, i.e. it is microscopic. This phenomenon is known as *local field effect*, the external or total field felt inside the medium are different from those caused outside it by the perturbing current or charge density. In other words the electrons of the crystal feel the external macroscopic field modified by the presence of the lattice and the rest of the medium, implicitly contained into ϵ^{-1} . Starting again from equation (1.64) and setting the term $\mathbf{G}' = 0$, one gets

$$\begin{aligned} & \sum_{\mathbf{G}' \neq 0} \left[\mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}') - \frac{\omega^2}{c^2} \delta_{\mathbf{G}, \mathbf{G}'} \right] \cdot \mathbf{D}(\mathbf{K} + \mathbf{G}') + \\ & + \left[\mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K}) \right] \cdot \mathbf{D}(\mathbf{K}) = 0 \end{aligned} \quad (1.66)$$

here the current disappeared because it is an external current and it can not contain the $\mathbf{K} + \mathbf{G}$ component, i.e. it has not a microscopic resolution and one can set $\mathbf{J}_{\text{EXT}}(\mathbf{K} + \mathbf{G}) = 0$ for each $\mathbf{G} \neq 0$. Equation (1.66) relates the macroscopic external field to the microscopic one. Inverting the preceding expression by means of the inverse operator:

$$\mathbf{T}(\mathbf{G}, \mathbf{G}') = \sum_{\mathbf{G}'' \neq 0} \left[\mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}'') - \frac{\omega^2}{c^2} \delta_{\mathbf{G}, \mathbf{G}''} \right] \quad (1.67)$$

one gets

$$\mathbf{D}(\mathbf{K} + \mathbf{G}'') = - \sum_{\mathbf{G} \neq 0} \mathbf{T}^{-1}(\mathbf{G}'', \mathbf{G}) \cdot \mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K}) \cdot \mathbf{D}(\mathbf{K}) \quad (1.68)$$

obtained multiplying both members for the inverse operator $\mathbf{T}^{-1}(\mathbf{G}'', \mathbf{G})$ to the left, and using $\sum_{\mathbf{G} \neq 0} \mathbf{T}^{-1}(\mathbf{G}'', \mathbf{G}) \cdot \mathbf{T}(\mathbf{G}, \mathbf{G}') = \delta_{\mathbf{G}'', \mathbf{G}'}$. Combining now (1.65) and (1.68), one obtains:

$$\mathbf{O}(\mathbf{K}) \cdot \left\{ \epsilon^{-1}(\mathbf{K}, \mathbf{K}) - \sum_{\mathbf{G}', \mathbf{G} \neq 0} \epsilon_{\mathbf{K}, \mathbf{K} + \mathbf{G}'}^{-1} \cdot \mathbf{T}^{-1}(\mathbf{G}', \mathbf{G}) \cdot \mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon_{\mathbf{K} + \mathbf{G}, \mathbf{K}}^{-1} \right\} \cdot \mathbf{D}(\mathbf{K}) - \frac{\omega^2}{c^2} \mathbf{D}(\mathbf{K}) = \frac{i\omega 4\pi}{c^2} \mathbf{J}_{\text{EXT}}(\mathbf{K}) \quad (1.69)$$

and if one defines a macroscopic response function in order to reach an equation with the same form of (1.64)

$$\left[\mathbf{O}(\mathbf{K}) \cdot \epsilon^{-1}(\mathbf{K}) - \frac{\omega^2}{c^2} \right] \cdot \mathbf{D}(\mathbf{K}) = \frac{i\omega 4\pi}{c^2} \mathbf{J}_{\text{EXT}}(\mathbf{K}) \quad (1.70)$$

from the comparison with (1.69) one obtains the definition of the macroscopic response functions in terms of the microscopic one:

$$\epsilon^{-1}(\mathbf{K}) = \epsilon^{-1}(\mathbf{K}, \mathbf{K}) - \sum_{\mathbf{G}', \mathbf{G} \neq 0} \epsilon^{-1}(\mathbf{K}, \mathbf{K} + \mathbf{G}') \cdot \mathbf{T}^{-1}(\mathbf{G}', \mathbf{G}) \cdot \mathbf{O}(\mathbf{K} + \mathbf{G}) \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K}) \quad (1.71)$$

If one restrict its interest only to the longitudinal part of the response function, things become easier. In this case one can simply play with the first Maxwell's equation of system (1.15) and the definition (1.57):

$$(\mathbf{K} + \mathbf{G}) \cdot \sum_{\mathbf{G}} \bar{\epsilon}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}') \cdot \mathbf{D}(\mathbf{K} + \mathbf{G}') = \eta_{\text{EXT}}(\mathbf{K} + \mathbf{G}) \quad (1.72)$$

Proceeding exactly as in the general case, one reaches the analogous of equation (1.69) and defining again a macroscopic equation like equation (1.70) one finally comes to:

$$\bar{\epsilon}(\mathbf{K}) = \bar{\epsilon}(\mathbf{K}, \mathbf{K}) - \sum_{\mathbf{G}', \mathbf{G} \neq 0} \bar{\epsilon}(\mathbf{K}, \mathbf{K} + \mathbf{G}') \cdot \epsilon^{-1}(\mathbf{K} + \mathbf{G}', \mathbf{K} + \mathbf{G}) \cdot \bar{\epsilon}(\mathbf{K} + \mathbf{G}, \mathbf{K}) \quad (1.73)$$

This expression is identical to the definition of the inverse of the first element M_{00}^{-1} of the inverse matrix \mathbf{M}^{-1} of a generic matrix \mathbf{M} :

$$\mathbf{M} = \begin{pmatrix} M_{00} & M_{0\mathbf{G}'} \\ M_{\mathbf{G}0} & M_{\mathbf{G}\mathbf{G}'} \end{pmatrix} \quad \frac{1}{M_{00}^{-1}} = M_{00} - \sum_{\mathbf{G}', \mathbf{G}} M_{0\mathbf{G}'} \cdot M_{\mathbf{G}'\mathbf{G}}^{-1} \cdot M_{\mathbf{G}0} \quad (1.74)$$

This means that microscopic response functions, in reciprocal space, must be regarded as matrixes in the indexes \mathbf{G} and \mathbf{G}' and that the macroscopic response functions are nothing but:

$$\bar{\epsilon}(\mathbf{K}, \omega) = \lim_{\mathbf{G}, \mathbf{G}' \rightarrow 0} \frac{1}{\epsilon^{-1}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}', \omega)} \quad (1.75)$$

$$\epsilon^{-1}(\mathbf{K}, \omega) = \lim_{\mathbf{G}, \mathbf{G}' \rightarrow 0} \frac{1}{\bar{\epsilon}(\mathbf{K} + \mathbf{G}, \mathbf{K} + \mathbf{G}', \omega)}$$

i.e. the inverse of the first element of the inverse microscopic response function.

By definition of Fourier's expansion coefficients, the term with zero momentum, is a spatial average of the function. In this view, the first elements of both the tensors $\epsilon^{-1}(\mathbf{K} + 0, \mathbf{K} + 0, \omega)$ and $\bar{\epsilon}(\mathbf{K} + 0, \mathbf{K} + 0, \omega)$ are a sort of macroscopic averages. To understand which is the difference between the *direct* average of the first element and the average (1.75) involving the first element of the *inverse* matrix, it is sufficient to look at equations (1.75) and (1.71):

neglecting the second addendum of the RHS, the inverse average becomes identical to the direct one, i.e. the macroscopic functions are simply given by the first term ($\mathbf{G} = \mathbf{G}' = 0$) of their matrixes because all the elements $\mathbf{G}, \mathbf{G}' \neq 0$ are zero. However to neglect the second addendum means to neglect local field effect at all. This is clearly visible from equation (1.65) in which, if all the $\mathbf{G}, \mathbf{G}' \neq 0$ elements are taken to be equal to zero, the first addendum of the LHS vanishes. Applying an external perturbation $\mathbf{J}_{\text{EXT}}(\mathbf{K})$ only the $\mathbf{D}(\mathbf{K})$ component exists, the electrons feel the same external field that is present outside the medium, no microscopic oscillations on the lattice length scale exist. This microscopic oscillations enter the macroscopic average during the elements mixing due to the matrix inversion procedure described by (1.75) and (1.71) when the second addendum is taken into account.

1.6 Dielectric tensor and magnetic permittivity

In the preceding sections the possibility to relate together different perturbation and response fields has been shown. Making use of (1.57) it is possible to rewrite Maxwell's equations (1.15) in a soluble and definitive form:

$$\begin{cases} \nabla \cdot \int d\mathbf{x}' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) = 4\pi\eta_{\text{EXT}} \\ \nabla \times \mathbf{H} + \frac{i\omega}{c} \int d\mathbf{x}' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) = \frac{4\pi}{c} \mathbf{J}_{\text{EXT}} \\ \nabla \times \mathbf{E} - \frac{i\omega}{c} \mathbf{H} = 0 \\ \nabla \cdot \mathbf{H} = 0 \end{cases} \quad (1.76)$$

with four equations and four incognita. Other possible forms of the system (1.76), involving the other fields \mathbf{D} and \mathbf{P} , require the usage of the other response functions above defined. If one prefers to use the system in the (1.12) form, a new response function must be introduced in order to relate the total and external magnetic fields, the *inverse magnetic permittivity tensor* μ^{-1} . This means that the information contained inside the unique response function $\bar{\epsilon}$ appearing in (1.76) must be partitioned into a new dielectric tensor $\tilde{\epsilon}$ and μ^{-1} . The partition is almost arbitrary and the definition of μ^{-1} depends on how much information one wants to leave into $\tilde{\epsilon}$. To show how $\bar{\epsilon}$, $\tilde{\epsilon}$ and μ^{-1} are related, one must restart from equation (1.42) using the original definition (1.10):

$$ic \mathbf{q} \times \mathbf{M}(\mathbf{q}, \omega) - i\omega \mathbf{P}(\mathbf{q}, \omega) = \int d\mathbf{q}' \bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega) \cdot \mathbf{E}(\mathbf{q}', \omega) \quad (1.77)$$

multiplying both sides for $i\mathbf{q} \times$ one gets:

$$q^2 \mathbf{M}(\mathbf{q}, \omega) + \frac{\omega}{c} \mathbf{q} \times \mathbf{P}(\mathbf{q}, \omega) = \frac{i}{c} \int d\mathbf{q}' \mathbf{q} \times \bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega) \cdot \mathbf{E}(\mathbf{q}', \omega) \quad (1.78)$$

where the following vector identity has been used:

$$\nabla \times \nabla \times \mathbf{M} = \nabla(\nabla \cdot \mathbf{M}) - \nabla^2 \mathbf{M} = -\nabla^2 \mathbf{M} \quad (1.79)$$

together with the property of divergenceless magnetic fields $\nabla \cdot \mathbf{M} = 0$. The total electric field appearing in the RHS can be decomposed in its transverse and longitudinal parts (see appendix B)

$$\mathbf{E} = \mathbf{E}^T + \mathbf{E}^L = -\frac{\omega}{cq^2} \mathbf{q} \times \mathbf{B} + \hat{\mathbf{q}} \hat{\mathbf{q}} \cdot \mathbf{E} \quad (1.80)$$

the \mathbf{E}^T definition comes from the multiplication of the third Maxwell's equation of (1.191) by $i\mathbf{q}\times$ and the application of (1.79). Reorganizing the terms:

$$\begin{aligned} \mathbf{M}(\mathbf{q}, \omega) = & -\frac{\omega}{q^2 c} \mathbf{q} \times \mathbf{P}(\mathbf{q}, \omega) - \frac{i\omega}{c^2} \int d\mathbf{q}' \hat{\mathbf{q}} \times \frac{\bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \times \hat{\mathbf{q}}' \cdot \mathbf{B}(\mathbf{q}', \omega) + \\ & + \frac{i}{c} \int d\mathbf{q}' \hat{\mathbf{q}} \times \frac{\bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \cdot \hat{\mathbf{q}}' \mathbf{q}' \cdot \mathbf{E}(\mathbf{q}', \omega) \end{aligned} \quad (1.81)$$

the trivial vector identity

$$\bar{\sigma} \cdot \hat{\mathbf{q}}' \times \mathbf{B} = \bar{\sigma} \times \hat{\mathbf{q}}' \cdot \mathbf{B} \quad (1.82)$$

has been used. Now one can use definitions (1.13) and the following one for $\tilde{\epsilon}$:

$$\mathbf{D}(\mathbf{q}, \omega) = \int d\mathbf{q}' \tilde{\epsilon}(\mathbf{q}, \mathbf{q}', \omega) \cdot \mathbf{E}(\mathbf{q}', \omega) \quad (1.83)$$

to get:

$$\begin{aligned} \mathbf{H}(\mathbf{q}, \omega) = & \int d\mathbf{q}' \left[\frac{\omega}{q^2 c} \mathbf{q} \times \left(\tilde{\epsilon}(\mathbf{q}, \mathbf{q}', \omega) - \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') \right) \cdot \mathbf{E}(\mathbf{q}', \omega) + \right. \\ & + \left(\mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{i\omega 4\pi}{c^2} \hat{\mathbf{q}} \times \frac{\bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \times \hat{\mathbf{q}}' \right) \cdot \mathbf{B}(\mathbf{q}', \omega) + \\ & \left. + \frac{i}{c} \int d\mathbf{q}' \hat{\mathbf{q}} \times \frac{\bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \cdot \hat{\mathbf{q}}' \mathbf{q}' \cdot \mathbf{E}(\mathbf{q}', \omega) \right] \end{aligned} \quad (1.84)$$

$$\begin{aligned} \mathbf{H}(\mathbf{q}, \omega) = & \int d\mathbf{q}' \left[-\frac{\omega^2}{q' q c^2} \hat{\mathbf{q}} \times \left(\tilde{\epsilon}(\mathbf{q}, \mathbf{q}', \omega) - \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') \right) \times \hat{\mathbf{q}}' + \right. \\ & \left. + \left(\mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{i\omega 4\pi}{c^2} \hat{\mathbf{q}} \times \frac{\bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \times \hat{\mathbf{q}}' \right) \right] \cdot \mathbf{B}(\mathbf{q}', \omega) \end{aligned}$$

in the second step the longitudinal electric field terms cancel out, this makes visible that the relation between \mathbf{H} and \mathbf{B} depends on the choice for the response function linking \mathbf{D} and \mathbf{E} , i.e. that μ^{-1} depends on the choice of $\tilde{\epsilon}$.

$$\begin{aligned} \mu^{-1}(\mathbf{q}, \mathbf{q}', \omega) = & -\frac{\omega^2}{q' q c^2} \hat{\mathbf{q}} \times \left(\tilde{\epsilon}(\mathbf{q}, \mathbf{q}', \omega) - \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') \right) \times \hat{\mathbf{q}}' + \\ & + \left(\mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{i\omega 4\pi}{c^2} \hat{\mathbf{q}} \times \frac{\bar{\sigma}(\mathbf{q}, \mathbf{q}', \omega)}{q q'} \times \hat{\mathbf{q}}' \right) = \\ = & -\frac{\omega^2}{q' q c^2} \left(\tilde{\epsilon}^{\text{TT}} - \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') \right) + \\ & + \left(\mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{i\omega 4\pi}{q' q c^2} \bar{\sigma}^{\text{TT}} \right) \end{aligned} \quad (1.85)$$

In particular, μ^{-1} contains only transverse components and depends on the choice made for the transverse part of $\tilde{\epsilon}$, whereas the longitudinal parts of $\tilde{\epsilon}$ do not enter μ^{-1} expression. Notice that, if one chooses $\tilde{\epsilon}$ to be equal to $\bar{\epsilon}$ of equation (1.57), the difference in the previous equation cancels out and one obtains $\mathbf{H} = \mathbf{B}$, that is the assumption made in the definition (1.14) one starts from to derive $\bar{\epsilon}$. In equation (1.14) all the transverse contribution from the magnetization has been hidden inside the susceptibility, one can also decide to do the opposite, putting all the transverse part of the susceptibility inside the magnetization vector.

This means that equation (1.10) must be used where \mathbf{P} is completely longitudinal, $\mathbf{E}^T = \mathbf{D}^T$, i.e. the transverse component of $\tilde{\epsilon}$ is just:

$$\tilde{\epsilon}^{TT} = \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') \quad \tilde{\epsilon}^{LL} = \bar{\epsilon}^{LL} \quad (1.86)$$

and (1.85) becomes:

$$\mu^{-1} = \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{i\omega 4\pi}{q'qc^2} \bar{\sigma}^{TT} \quad (1.87)$$

$\tilde{\epsilon}$ takes care of all the longitudinal parts of the dielectric tensor $\bar{\epsilon}$ whereas μ^{-1} contains all the transverse parts of $\bar{\epsilon}$. Another possibility is to choose $\tilde{\epsilon}$ in such a way that it treats longitudinal and transverse components of the electric field in the same way:

$$\tilde{\epsilon}^{LL} = \tilde{\epsilon}^{TT} = \bar{\epsilon}^{LL} \quad (1.88)$$

now (1.85) becomes:

$$\mu^{-1} = \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{i\omega 4\pi}{q'qc^2} (\bar{\sigma}^{TT} - \bar{\sigma}^{LL}) = \mathbf{1}\delta(\mathbf{q} - \mathbf{q}') + \frac{\omega^2}{q'qc^2} (\bar{\epsilon}^{TT} - \bar{\epsilon}^{LL}) \quad (1.89)$$

The advantage of this new definition is that, in the limit $\mathbf{q} \rightarrow 0$, when $\bar{\epsilon}^{TT} \rightarrow \bar{\epsilon}^{TT} = \epsilon$, $\mu^{-1} \rightarrow 0$ and $\tilde{\epsilon} = \epsilon$ giving the same limit of the (1.14) choice. On the other hand, using definition 1.87, in the limit $\mathbf{q} \rightarrow 0$, μ^{-1} is non vanishing and $\tilde{\epsilon}$ is an anisotropic tensor with $\tilde{\epsilon}^{LL} = \epsilon$ and $\tilde{\epsilon}^{TT} = 1$. Independently of the above choice, the system (1.12) becomes:

$$\left\{ \begin{array}{l} \nabla \cdot \int d\mathbf{x}' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) = 4\pi\eta_{\text{EXT}} \\ \nabla \times \int d\mathbf{x}' \mu^{-1}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{B}(\mathbf{x}', \omega) + \\ + \frac{i\omega}{c} \int d\mathbf{x}' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{E}(\mathbf{x}', \omega) = \frac{4\pi}{c} \mathbf{J}_{\text{EXT}} \\ \nabla \times \mathbf{E} - \frac{i\omega}{c} \mathbf{B} = 0 \\ \nabla \cdot \mathbf{B} = 0 \end{array} \right. \quad (1.90)$$

now soluble in the variables \mathbf{E} , \mathbf{B} and the field sources.

1.7 Medium excitations and dissipation

In appendix C the role of causality in determining the response functions properties is briefly discussed. Causality forces the response functions to be analytic in the upper half-plane of complex frequencies. In the lower half-plane there can be poles and they have a well defined physical meaning: they represent, together with the zeroes, the intrinsic excitation modes of a medium. The dielectric function represents a special case of response function because also its inverse is a response function, this means that poles of the inverse dielectric functions and zeroes of the direct one are related. In the specific case of the dielectric function, causality ask not only for the absence of poles in the upper half-plane but also for the absence of zeroes, because the latter can become poles for the inverse function. Looking at the real counterpart of equation (1.57):

$$\mathbf{D}(\mathbf{x}, t) = \int d\mathbf{x}' \int dt' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', t - t') \cdot \mathbf{E}(\mathbf{x}', t') \quad (1.91)$$

one can see that if the external perturbing field is not yet switched on, $\mathbf{D} = 0$, the total field \mathbf{E} can be different from zero if and only if it is the dielectric function that vanish.

An electric field can exist before the perturbation is switched on, if and only if some of its intrinsic modes are excited. The Fourier's transform of the imaginary part vanish because of parity conditions (1.194):

$$\bar{\epsilon}(\mathbf{x}, \mathbf{x}', t - t') = \int d\omega \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) e^{-i\omega(t-t')} = \int d\omega \bar{\epsilon}_1(\mathbf{x}, \mathbf{x}', \omega) e^{-i\omega(t-t')} \quad (1.92)$$

so the zeroes of the real part of the response function describe the medium intrinsic modes. As it will be shown in the next section for some specific cases, the zeros of the real part of the dielectric function correspond to the zeroes and the poles in the lower half-plane of complex frequencies. Again, with reference to the real counterpart of equation (1.29):

$$\mathbf{E}(\mathbf{x}, t) = \int d\mathbf{x}' \int dt' \epsilon^{-1}(\mathbf{x}, \mathbf{x}', t - t') \cdot \mathbf{D}(\mathbf{x}', t') \quad (1.93)$$

when the perturbing external field is not yet switched on, the total field \mathbf{E} can be different from zero only if the inverse dielectric function diverges. Looking at the Fourier's transform of the dielectric function:

$$\frac{1}{\epsilon(\mathbf{x}, \mathbf{x}', t - t')} = \int d\omega \left(\frac{\bar{\epsilon}_1}{\bar{\epsilon}_1^2 + \bar{\epsilon}_2^2} + i \frac{\bar{\epsilon}_2}{\bar{\epsilon}_1^2 + \bar{\epsilon}_2^2} \right) e^{-i\omega(t-t')} \quad (1.94)$$

it is not possible to distinguish between an even and an odd part of the integrand, nevertheless the imaginary part must vanish, i.e. $\bar{\epsilon}_2$ must go to zero. With this assumption the real part becomes $1/\bar{\epsilon}_1$, again the zeroes of the real part of the dielectric function allow the response function to diverge, accounting for the intrinsic excitations of the medium. Some specific examples of medium excitations will be given in the next section where an outline of the different models for response functions will be treated.

Response functions rules the dispersion of the medium, i.e. the functions linking \mathbf{q} to ω for the different possible medium excitations. The simplest example is given by an electromagnetic wave propagating inside a medium:

$$E(\mathbf{x}, t) = E_0(\mathbf{x}) e^{i\mathbf{q}\cdot\mathbf{x} - i\omega t} \quad (1.95)$$

from the Maxwell's equations (1.76), one gets the well known D'Alembert's equation:

$$\nabla^2 \mathbf{E}(\mathbf{x}, t) = \frac{1}{c^2} \int d\mathbf{x}' \bar{\epsilon}(\mathbf{x}, \mathbf{x}', \omega) \frac{\partial^2 \mathbf{E}(\mathbf{x}, t)}{\partial t^2} \quad (1.96)$$

replacing (1.95) and moving to reciprocal space, under the assumption of spatial invariance:

$$\begin{aligned} \mathbf{q}^2 \mathbf{E}(\mathbf{q}, \omega) &= \frac{\omega^2}{c^2} \bar{\epsilon}(\mathbf{q}, \omega) \mathbf{E}(\mathbf{q}, \omega) \\ \mathbf{q}^2 &= \frac{\omega^2}{c^2} \bar{\epsilon} \end{aligned} \quad (1.97)$$

here the dielectric tensor is a real quantity mixing \mathbf{q} and ω of the medium excitation modes. Separating real and imaginary parts of the response function one can see that, if one vanish, the other one can diverge giving rise to unphysical behaviours: both of them must always exist to prevent divergencies. The imaginary part of the dielectric tensor is related to the dissipation capability of the medium. To show this, imagine to have an electromagnetic wave traveling into the vacuum and entering, at a certain point, into a medium. If the medium is dissipative it must take away some energy from the electromagnetic field reducing its strength, one can suppose that the decay of the wave strength is exponential:

$$E(\mathbf{x}, t) = E_0(\mathbf{x}) e^{-\mathbf{m}\cdot\mathbf{x}} e^{i\mathbf{q}\cdot\mathbf{x} - i\omega t} = E_0(\mathbf{x}) e^{i(\mathbf{q}+i\mathbf{m})\cdot\mathbf{x} - i\omega t} \quad (1.98)$$

The dissipation assumption brings to a complex wave vector $\mathbf{p} = \mathbf{q} + i\mathbf{m}$. A *refractive index* n and an *extinction coefficient* k can be defined as:

$$\mathbf{p} = \frac{\omega}{c}(n + ik) \quad (1.99)$$

Now the dispersion relation becomes:

$$\mathbf{p}^2 = \frac{\omega^2}{c^2}\bar{\epsilon} \quad \longrightarrow \quad n + ik = \sqrt{\bar{\epsilon}_1 + i\bar{\epsilon}_2} \quad (1.100)$$

a complex dielectric function can be defined in order to recover expression (1.97):

$$\bar{\epsilon}_1 = n^2 + k^2 \quad \bar{\epsilon}_2 = 2kn \quad (1.101)$$

Definition (1.57) states that the real part of the dielectric tensor is related to the imaginary part of the conductivity and vice versa, on the other hand, the power dissipated by a current is given by:

$$\frac{\partial W}{\partial t} = \mathbf{j}(\mathbf{x}, t) \cdot \mathbf{E}(\mathbf{x}, t) = \sigma(\mathbf{x}, t)\mathbf{E}^2(\mathbf{x}, t) \quad (1.102)$$

So, if $\bar{\epsilon}$ is real, like in (1.97), only the imaginary part of the conductivity $\bar{\sigma}_2$ exists, by virtue of relations (1.194) it has to be an odd function so its Fourier's transform vanish:

$$\sigma(t) = \int_{-\infty}^{+\infty} d\omega [\sigma_1(\omega) + i\sigma_2(\omega)]e^{-i\omega t} = \int_{-\infty}^{+\infty} d\omega \sigma_1(\omega)e^{-i\omega t} \quad (1.103)$$

On the other hand, if $\bar{\epsilon}$ is complex like in (1.100), also $\bar{\sigma}_1$ exists and the Fourier's transform (1.103) is non vanishing leading to a dissipative current. The existence of an imaginary part of the dielectric tensor is responsible for the rise of a current, by means of which the medium dissipates the energy subtracted at the traveling wave².

1.8 Dielectric function: outline of the different models

In this section a detailed discussion of the different models for the dielectric function will be given. First of all the classical models will be introduced, than the quantum models will be discussed showing in which conditions they reach the classical limit.

Plasma model This is the simplest classical model, it describes the dielectric function of an homogeneous, non-interacting, free electron gas. Despite of its simplicity it will be shown that all the dielectric function models, both classical and quantum, behave as the plasma model in the limit of very high frequencies. Imagine to perturb the electron gas with a total field $\mathbf{E}(t) = E_0 e^{i\omega t}$, the equation of motion for each free non- interacting electron becomes:

$$m\ddot{\mathbf{x}}(t) = -E_0 e^{i\omega t} e \quad (1.104)$$

where m is the electron mass, e the electron charge, and \mathbf{x} the electron position. With the initial conditions $\mathbf{x}(0) = 0$ and $\dot{\mathbf{x}}(0) = 0$ the solution of the differential equation is:

$$\mathbf{x}(t) = \frac{E_0 e}{m\omega^2} e^{i\omega t} \quad (1.105)$$

²One may think that, in order to know a response function in real space $\phi(t)$, only the real component $\phi_1(\omega)$ of the Fourier's transform is needed whereas the imaginary part is always useless. This is wrong because, depending on whether an imaginary part is present or not, the real part changes. This is visible also from definition (1.101) where the addition of dissipation creates both an imaginary part and a new contribution to the real one. So imaginary part does not give a direct contribution in the calculation of the dielectric function in real space, but it is implicitly present into the real part.

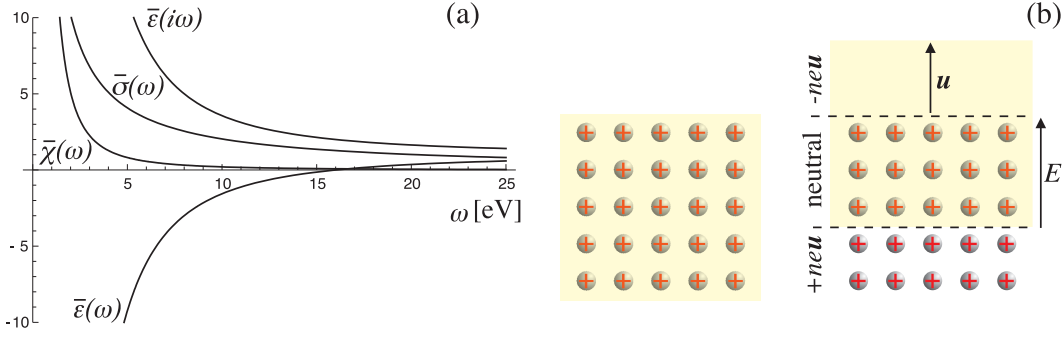


Figure 1.4: (a) Some response functions describing the plasma model. The plasma frequency is $\Omega_p = 16\text{eV}$ a typical value for metallic media. (b) Pictorial representation of a plasmon: yellow shadow represents the electrons cloud while spheres represent the ionic lattice.

from the Fourier transform of equation (1.56), neglecting the local field correction and the interaction between dipoles, one can derive an expression for the permittivity:

$$\begin{aligned} \mathbf{P} &= N\mathbf{p} = \bar{\chi}\mathbf{E} \\ \bar{\chi} &= -\frac{Ne^2}{m} \frac{1}{\omega^2} \end{aligned} \quad (1.106)$$

where the definition $\mathbf{p} = -e\mathbf{x}(t)$ has been used. Finally, using definition (1.57) one gets:

$$\bar{\epsilon}(\omega) = 1 - \frac{4\pi e^2 N}{m} \frac{1}{\omega^2} = 1 - \frac{\Omega_p^2}{\omega^2} \quad (1.107)$$

here Ω_p is the *plasma frequency* and depends on the free electron gas density. Other response functions, immediately derivable from the previous equation, are plotted in figure 1.4 (a). The dielectric function does not have in imaginary part, nevertheless imposing $\bar{\epsilon} = \bar{\epsilon}_1 = 0$ one gets the only solution $\omega = \Omega_p$. There exist only one intrinsic mode of the medium, a longitudinal collective excitation called *plasmon* that depends upon the medium electron density. At large frequencies the response function goes to 1, i.e. the medium behave as the vacuum. For large frequencies, the oscillations wave length of the total field is so small (localized in space) that the electrons can not feel it, they feel only an average electric field that, because of the periodicity, is zero, i.e. the medium is not able to polarize itself. At small frequencies the response function diverges, this is an unphysical behaviour because it leads to a divergent real response function.

The plasmon is a collective mode and can not be understood in terms of single oscillators like (1.104), a collective picture must be used. The medium itself can be regarded as an harmonic oscillator: figure 1.4 (b) shows a medium as the overlap of a positive fixed charge due to the ionic lattice and a mobile negative charge due to the electrons cloud. When the electron cloud is displaced from its equilibrium position the medium is no more neutral and a restoring electric field E take place. The behaviour of an infinitesimal volume of the electron cloud is described by the equation:

$$Nm \frac{d^2\mathbf{u}}{dt^2} = -Ne\mathbf{E} = -4\pi N^2 e^2 \mathbf{u} \quad \longrightarrow \quad \frac{d^2\mathbf{u}}{dt^2} = -\Omega_p^2 \mathbf{u} \quad (1.108)$$

where \mathbf{u} is the displacement from the volume equilibrium position and $E = 4\pi n e \mathbf{u}$ is the restoring field due to the charge induced by the displacement. The proper frequency of the harmonic oscillator (1.108) is the plasma frequency $\Omega_p = \frac{4\pi e^2 N}{m}$ of the medium. If an external perturbing field is applied with $\omega = \Omega_p$, it can resonate with the proper frequency of the

medium, switching off the perturbation the electron cloud continue to oscillate, the medium being in its intrinsic excitation mode. Any other perturbing field with $\omega \neq \Omega_p$ is unable to excite the plasmon mode.

Drude model To eliminate the unphysical behaviour of the plasma model one must add a dissipative term into differential equation (1.104):

$$\ddot{\mathbf{x}}(t) + \gamma\dot{\mathbf{x}}(t) = -\frac{E_0 e}{m} e^{i\omega t} \quad (1.109)$$

γ is the dissipation constant divided by the mass m , now the equation is no longer time reversible. The solution becomes:

$$\mathbf{x}(t) = \frac{E_0 e/m}{\omega^2 + i\gamma\omega} e^{i\omega t} \quad (1.110)$$

and the dielectric function:

$$\bar{\epsilon}(\omega) = 1 - \frac{\Omega_p^2}{\omega^2 + i\gamma\omega} \quad (1.111)$$

where γ has the dimensions of a frequency, its inverse, τ , is the *relaxation time* of the system. In the Drude transport theory, τ is the average free time between electrons collisions, its value depends on the microscopic structure of the medium. As it can be seen from figure 1.5, now the real part of the dielectric function is limited while the imaginary one diverges, this is not a problem because the latter quantity does not enters the Fourier's transform for the real dielectric function calculation. The divergence in the imaginary part is due to the fact that the electrons are free and they can absorb energy of almost every frequency dissipating it through a conduction current. Notice that, now that the real part of the dielectric function is limited, $\bar{\epsilon}_1(0) = \bar{\epsilon}(i0)$.

Imposing now the condition $\bar{\epsilon}_1 = 0$ one gets again the plasmon excitation at frequency $\omega = -\gamma^2/2 + \sqrt{(\gamma^4 + 4\Omega_p^2)/4} \simeq \Omega_p$ for γ small. In the complex plane the previous solution corresponds to $\omega = \sqrt{\Omega_p^2 - \gamma^2/4} - i\gamma/2$, this solution lies in the lower half-plane of the complex frequency, as prescribed by the causality principle. The Drude model describes quite well the response of alkali metals, where only the s orbitals are filled and the electrons behaviour is very close to the ideal free gas one. An improvement of the Drude model consists in replacing the mass m , appearing into the plasma frequency, with an effective mass m^* that accounts for the real band dispersion of the medium.

Lorentz model The Lorentz model is useful to describe the response of semiconductors and insulators for which the absorption spectrum always present a gap at low frequencies. In this model electrons are no more free, they are confined at certain positions and they can only oscillate around them. Imagine, for instance, an electron in the Bohr atomic model, while it is moving along a certain circular orbit with frequency ω_0 , an external field feels it as an harmonic oscillator with proper frequency ω_0 . In this view, the differential equation (1.109) must contain a restore elastic term:

$$\ddot{\mathbf{x}}(t) + \gamma\dot{\mathbf{x}}(t) + \omega_0^2\mathbf{x}(t) = -\frac{E_0 e}{m} e^{i\omega t} \quad (1.112)$$

here ω_0 is the restore elastic constant divided by the electron mass. Following the same procedure described above, one get:

$$\bar{\epsilon}(\omega) = 1 + \frac{\Omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega} \quad (1.113)$$

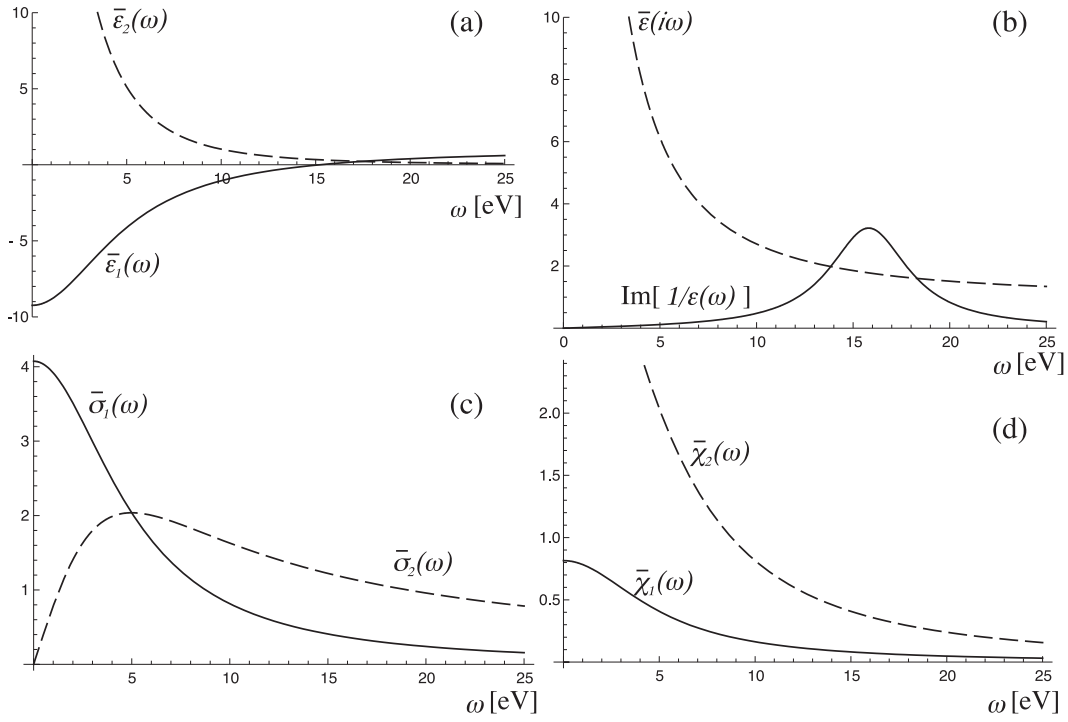


Figure 1.5: Some response functions describing the Drude model. The plasma frequency is $\Omega_p = 16\text{eV}$ whereas $\gamma = 5\text{eV}$.

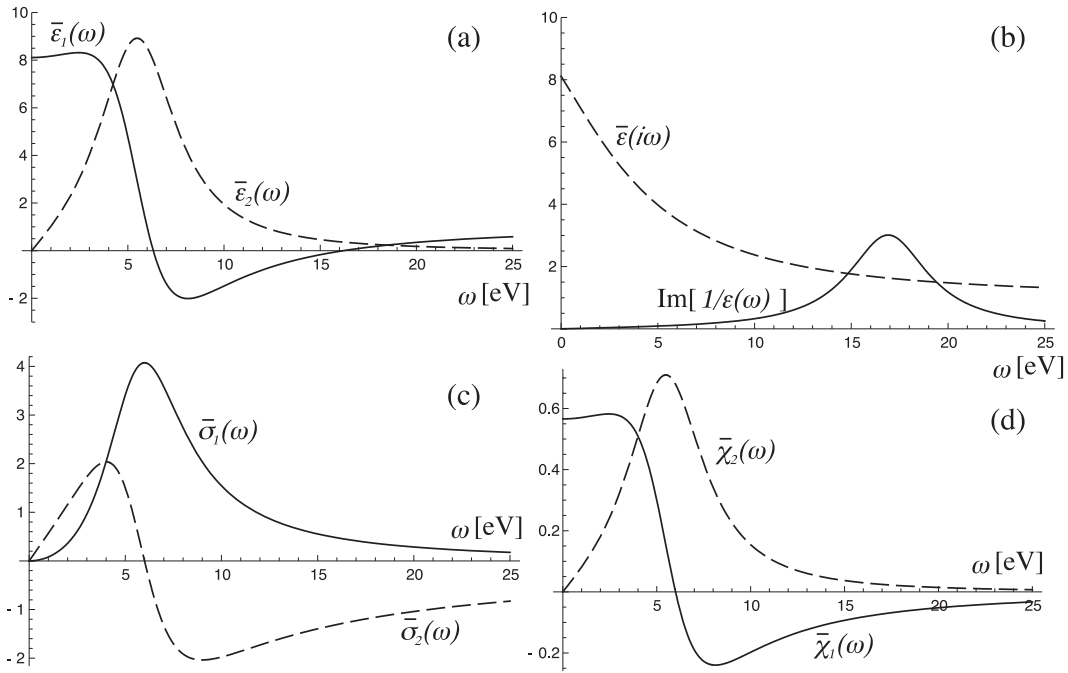


Figure 1.6: Some response functions describing the Lorentz model. The plasma frequency is $\Omega_p = 16\text{eV}$, $\gamma = 5\text{eV}$ and the transverse frequency is $\omega_0 = 6\text{eV}$.

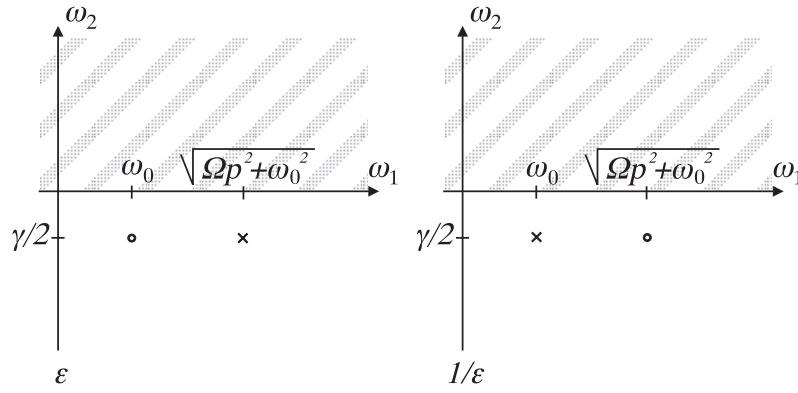


Figure 1.7: Complex frequency plane for the direct and inverse dielectric function in the Lorentz model. The upper half-plane can not contain zeroes and poles because of the causality principle, crosses indicate poles while circles indicate zeroes.

where the imaginary part is a Lorentzian curve centered at $\omega = \omega_0$ with an half height amplitude related to γ . This absorption peak, plotted in figure 1.6, represent the resonance between the perturbation field and the proper frequency of the electrons, far from ω_0 the absorption of the medium is almost negligible. ω_0 depends on the nature of the chemical bonds of the materials, many different restore terms can be included into the differential equation (1.112) corresponding to different proper frequencies ω_n of the medium, this issue will be discussed later on.

Imposing $\bar{\epsilon}_1 = 0$ one gets a new mode at $\omega \simeq \omega_0$ for γ small, together with the old plasmon excitation $\omega = \sqrt{\Omega_p^2 + \omega_0^2}$, the plasmon is shifted to higher frequencies by the presence of ω_0 . The new excitation is a transverse one, it can be related to the absorption of light. In this case the two solutions correspond to a pole and a zero in the complex plane at frequencies $\omega = \sqrt{\gamma^2/4 + \omega_0^2} - i\gamma/2$ and $\omega = \sqrt{\gamma^2/4 + \Omega_p^2 + \omega_0^2} - i\gamma/2$ respectively, they are sketched in figure 1.7 for $\bar{\epsilon}$ and $1/\epsilon$.

Generalized Lorentz model All semiconductors and insulators usually present more than a single transverse absorption frequency ω_0 , see figure 1.8 (b), even if it is always possible to determine the strongest one, an accurate description requires more than a single transverse frequency. In all the transition metals, in which d and f orbitals are partially occupied, a number of localized electrons exist together with the free electron gas, the Drude model must be integrated with the Lorentz one in order to account for the strong transverse absorption occurring at certain specific frequencies, see figure 1.8 (a). A generalization of the dielectric function

$$\bar{\epsilon}(\omega) = 1 + \sum_i \frac{f_i^2 \Omega_p^2}{\omega_i^2 - \omega^2 - i\gamma_i \omega} \quad (1.114)$$

can be built introducing the *oscillator strength* f_i , it gives a different weight to the different optical transitions occurring at frequency ω_i . To recover the Drude metallicity it is sufficient to put one of the ω_i equal to zero. As it will be shown in the next section, the oscillators strength must satisfy the constrain

$$\sum_i f_i^2 = 1 \quad (1.115)$$

in order to preserve the total electron density of the material. The generalized Lorentz model behaves exactly as the previous ones: some zeroes and poles, describing longitudinal and transverse medium excitations, can be found solving the equation $\epsilon_1 = 0$.

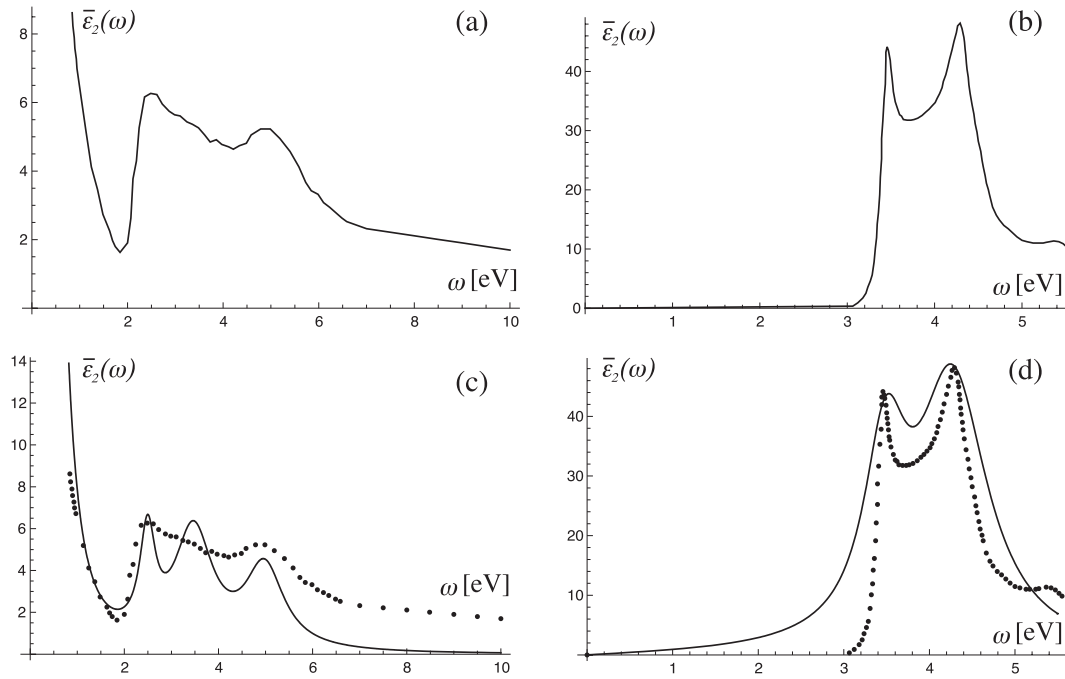


Figure 1.8: Experimental measurement of the imaginary part of the dielectric function for (a) copper and (b) silicon. (c) and (d) represent two rough attempts to use the generalized Lorentz model to fit the experimental data (circles). In the case of copper a Drude term has been used together with three different Lorentz terms, for silicon only two Lorentz terms have been used.

In all the classical models treated above, the \mathbf{q} dependence has been neglected, simple models to include the spatial dispersion can be found in references [4] and [5].

1.9 Sum rules

In all the previously introduced dielectric function models, one finds that for large frequencies i.e. $\omega \gg \omega_i$ for any i , the real part of the dielectric function follows the Plasma model. On the other hand, for large ω , the imaginary part of the response function goes rapidly to zero, becoming negligible. One can use this result together with the second Kramers-Kronig's relation (1.206) to find certain constraints that the dielectric function must satisfy. At high frequency $\omega \gg \omega_i$ so one can use the Plasma model into the LHS of the Kramers-Kronig's relation

$$1 + \frac{\Omega_p^2}{\omega^2} = 1 + \frac{2}{\pi} \wp \int_0^{\infty} \frac{\bar{\epsilon}_2(\omega')\omega'}{\omega'^2 - \omega^2} d\omega' \quad (1.116)$$

but for large frequency $\omega \gg \omega'$ and one can neglect ω' in the denominator reaching the rule:

$$\int_0^{\infty} \bar{\epsilon}_2(\omega')\omega' d\omega' = \frac{\pi}{2} \Omega_p^2 \quad (1.117)$$

this rule is a very powerful tool to check experimental spectroscopy measurements and theoretical models. A similar procedure applied to the second Kramers-Kronig's relation in (1.204) the leads to:

$$\int_0^{\infty} \epsilon_2^{-1}(\omega')\omega' d\omega' = \frac{\pi}{2} \Omega_p^2 \quad (1.118)$$

So the two functions obtained multiplying $\bar{\epsilon}_2$ and ϵ_2^{-1} by ω must subtend the same area. Another rule, valid only for semiconductors and insulator can be found combining definition (1.57) and the conductivity Kramers-Kronig's relation:

$$\begin{aligned} \bar{\sigma}_1(\omega) &= -\frac{2}{\pi} \wp \int_0^{\infty} \frac{\bar{\sigma}_2(\omega')\omega'}{\omega'^2 - \omega^2} d\omega' \\ \bar{\sigma}_1(\omega) &= -\frac{1}{2\pi^2} \wp \int_0^{\infty} \frac{[\bar{\epsilon}_1(\omega') - 1]\omega'^2}{\omega'^2 - \omega^2} d\omega' \end{aligned} \quad (1.119)$$

but in a semiconductor or in a insulator the static conductivity is zero³, i.e. $\bar{\sigma}_1(\omega \rightarrow 0) \rightarrow 0$, this means:

$$\int_0^{\infty} [\bar{\epsilon}_1(\omega') - 1] d\omega' = 0 \quad (1.120)$$

the area subtended by the real part of the dielectric function of an insulator or a semiconductor, subtracted by 1, must be zero. Finally, integrating both sides of the London's transform definition (1.207):

$$\int_0^{\infty} [\bar{\epsilon}(i\omega) - 1] d\omega = \frac{2}{\pi} \int_0^{\infty} d\omega \int_0^{\infty} dz \frac{\bar{\epsilon}_2(z)z}{z^2 + \omega^2} \quad (1.121)$$

³This can be obtained, for example, starting from the Lorentz model. The conductivity trend for the Lorentz model is plotted in figure 1.6.

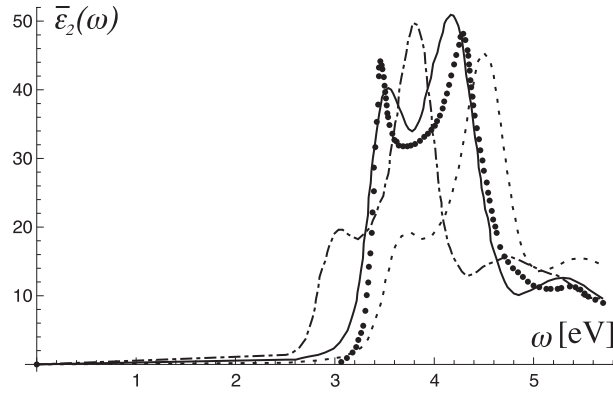


Figure 1.9: Different approximations for the imaginary part of the silicon dielectric function. Circles represent the experimental values, dot-dashed line represents the DFT calculation, dotted line represents the GW calculation and continuous line represents the BSE calculation. The GW correction enlarges the optical gap, whereas the inclusion of excitonic effects changes the distribution of spectral weight magnifying the lower energy peak. Figure adapted from reference [6].

and using the definite integral:

$$\int_0^{\infty} \frac{d\omega}{z^2 + \omega^2} = \frac{\pi}{2z} \quad (1.122)$$

one finally gets

$$\int_0^{\infty} \bar{\epsilon}_2(z) dz = \int_0^{\infty} [\bar{\epsilon}(i\omega) - 1] d\omega \quad (1.123)$$

The absorption spectrum and its London's transform must subtend the same area. This is a useful achievement because a simulated or an experimental absorption spectrum can contain very sharp spikes, background noise and many other features that make it very hard to be integrated, whereas the London's transform is a very smooth function.

The sum rule (1.118) must be satisfied by all the classical models shown in section 1.8, in particular it must be satisfied by the Drude and the Lorentz models. The generalized Lorentz model is nothing but a linear combination of Drude and Lorentz terms weighted by the oscillator strengths f_i . So, by linearity, one gets for the generalized Lorentz model:

$$\int_0^{\infty} \epsilon_2^{-1}(\omega') \omega' d\omega' = \frac{\pi}{2} \Omega_p^2 \sum_i f_i^2 \quad (1.124)$$

and it becomes clear that f_i must be subjected to the constrain $\sum_i f_i^2 = 1$ in order to let the generalized Lorentz model satisfy rule (1.118) too.

1.10 Response functions and many body interaction

Within the quantum mechanics theory, the simpler possibility to treat a free non-interacting electron gas is to describe each electron by means of a plane wave, this problem has been solved analytically by Lindhard [7] and generalized by Mermin [8]. The subsequent step is to add a periodic potential term in the hamiltonian to account for the presence of a crystal lattice. In this case the solutions of the hamiltonian are represented by Bloch's functions and

the energy dispersion of the system gives rise to an energy band structure.

In a genuine many-body approach, the independent particle wavefunctions must be replaced by linear combinations of Slater's determinants, or, in the second quantization formalism, one has to move to the Fock's space of occupation numbers sets. In this framework, the Kubo's fluctuation-dissipation theorem can be reformulated to show that the response function of a perturbed system is a contraction of the two particle Green's function called *polarization propagator*, i.e. the many-body counterpart of the electric susceptibility χ [9]. The most general many body treatment of the medium response as been proposed by Hedin [10] and is based on the self-consistent solution of a five equation system, many other many body techniques are nothing but a simplification of the Hedin's approach.

One can always map a many-body problem into an independent particle problem by means of an effective independent particle hamiltonian, the many-body interaction hidden within. The system can be described again by a set of independent particle eigenfunctions $\{\psi_j\}$ and the corresponding eigenvalues E_j solutions of the new effective hamiltonian. This is the spirit of techniques such as the Hartree-Fock (HF) self-consistent approach or the Density Functional Theory (DFT). Unfortunately HF calculations require a huge computational effort and can not be applied to large molecules or complicated bulk materials. On the other hand, DFT theory guarantee correct results only for the ground state of many-body systems. This fact, together with the approximation made for the unknown functionals, leaves the DFT unable to predict the right shape of empty bands, the correct band gap of semiconductors, the correct bond length and bond angles of certain molecules and the right dispersion interaction (i.e. van der Waals and Casimir-Polder forces) [11]. Despite of this limitation DFT is currently employed in the determination of dielectric properties of many systems with satisfactory results.

In the following, some approximate quantum model for the many-body dielectric function will be presented. These approximations are based on the solution of an hamiltonian whose eigenfunctions are linear combination of single particle wavefunctions (atomic orbitals, Bloch states ecc.) like in the Hartree case or in the Kohn and Sham scheme for the DFT. Such models are unable to account for the correlation phenomena, to this aim an antisymmetrized Slayter's determinant must be used, like in the HF theory.

Random Phase Approximation (RPA) This is a very general quantum model that requires only the capability to solve the single particle hamiltonian \mathcal{H}_0 describing the unperturbed medium in order to get its eigenvalues E_i and eigenfunctions ψ_i . Than applying the perturbation theory on the density matrix equation (1.158), with the hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{EXT}}(t)$, one is able to write the dielectric function of the medium in terms of E_i and ψ_i . The perturbation $\mathcal{H}_{\text{EXT}}(t)$ can be transverse [5], longitudinal [12, 2] or both of them [13, 14], as in equation (1.16). The dielectric function enters the problem thanks to the Maxwell's equations (1.76) that links charges and currents densities to the perturbation potentials or fields, where the charge and current densities, $\hat{\eta}$ and $\hat{\mathbf{J}}$, are carried out starting from the density matrix $\hat{\rho}$:

$$\begin{aligned}\hat{\rho} &= \sum_{i,j} \hat{\rho}_{i,j} = \sum_{i,j} |\psi_j\rangle\langle\psi_i| & \rho(\mathbf{x}, \mathbf{x}') &= \langle\mathbf{x}'|\hat{\rho}|\mathbf{x}\rangle \\ \hat{\eta} &= Tr\{\hat{\rho}\} = \sum_i \rho_{i,i} & \hat{\mathbf{J}} &= Tr\{\hat{\rho}\hat{\mathbf{J}}\} = \sum_i \sum_j \rho_{i,j} \mathbf{J}_{j,i} \\ \eta(\mathbf{x}) &= \sum_i \langle\mathbf{x}|\psi_i\rangle\langle\psi_i|\mathbf{x}\rangle = \sum_i f_i |\psi_i(\mathbf{x})|^2\end{aligned}\tag{1.125}$$

here f_i are the occupation numbers of the electrons for different eigenstates of the medium, i.e. the eigenvalues of the density operator. If one uses only the longitudinal perturbation the charge equation is enough, if one is interested in applying also a transverse perturbation,

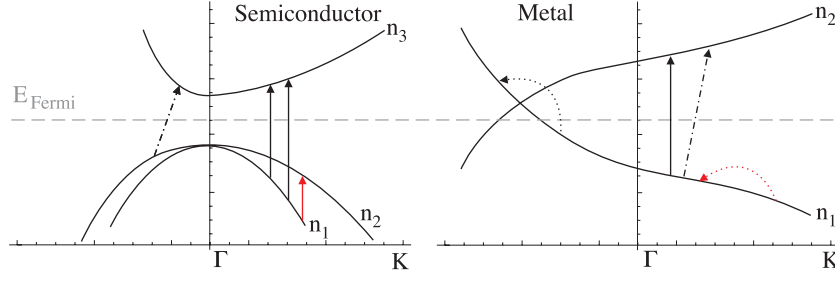


Figure 1.10: Band structure of a medium and its electron transitions. Black lines represent allowed transitions, red lines represents transitions forbidden by occupation numbers. Continuous lines represent vertical interband transitions, dot dashed lines represent non vertical interband transitions and dotted lines represent intraband transitions. Gray dashed line represent the Fermi level.

then the current equation must be used. This is because the charge equation rules only the longitudinal electric field whereas the current equation contains both transverse and longitudinal fields (see appendix B). Of course, if the medium is a cubic one, the longitudinal and transverse responses become identical in the limit of $\mathbf{q} \rightarrow 0$, as discussed in section 1.3. For the seek of clearness the transverse and longitudinal dielectric functions will be presented separately: with the scalar potential

$$\phi(\mathbf{x}, \mathbf{t}) = \sum_{\mathbf{q}} \phi(\mathbf{q}, \omega) e^{i\mathbf{q} \cdot \mathbf{x}} e^{-i\omega t + \gamma t} \quad (1.126)$$

adiabatically turned on, one gets the longitudinal dielectric function:

$$\bar{\epsilon}(\mathbf{q}, \omega) = 1 - \frac{4\pi e^2}{\mathbf{q}^2 V} \sum_{\mathbf{K}, n, n'} \frac{\langle \mathbf{K} + \mathbf{q}, n' | \mathbf{K}, n \rangle \langle \mathbf{K}, n | \mathbf{K} + \mathbf{q}, n' \rangle}{E_{\mathbf{K} + \mathbf{q}, n'} - E_{\mathbf{K}, n} - \hbar\omega - i\gamma\hbar} (f_{\mathbf{K} + \mathbf{q}, n'} - f_{\mathbf{K}, n}) \quad (1.127)$$

here the generic wavefunctions ψ_i have been replaced with Bloch's functions $|\psi_i\rangle = |\mathbf{K}, n\rangle e^{i\mathbf{K} \cdot \mathbf{x}}$ and the i index has been split into a band index n and a momentum index \mathbf{K} . γ comes from the adiabaticity of the perturbation and it is responsible for the presence of an imaginary part in the dielectric response. V is the volume of the lattice unit cell. As in the classical models zeroes and poles of this dielectric functions represent the collective or single particle excitations of the medium. A plasmon excitation exist when the second term of the RHS equals 1. Transverse (optical) excitations exist for $\hbar\omega = E_{\mathbf{K} + \mathbf{q}, n'} - E_{\mathbf{K}, n}$ with an imaginary part that shifts the poles down in the lower half-plane of complex frequencies in order to fulfill the causality principle. This excitations correspond to absorption or emission of photons by the electrons of the medium that jump from an occupied to an empty Bloch's state labeled by the quantum numbers $\mathbf{K} + \mathbf{q}, n'$ and \mathbf{K}, n . The occupation of the states is expressed by the difference $f_{\mathbf{K} + \mathbf{q}, n'} - f_{\mathbf{K}, n}$. Certain electron transitions are forbidden by the symmetry of the system, in that case, even if $f_{\mathbf{K} + \mathbf{q}, n'} - f_{\mathbf{K}, n} \neq 0$ the corresponding term of the sum vanish because $\langle \mathbf{K} + \mathbf{q}, n' | \mathbf{K}, n \rangle = 0$. The three possible kinds of optical transitions are sketched in figure 1.10, *intraband transitions* occur when $n = n'$, *interband transitions* occur when $n \neq n'$ and they can be *vertical* if $\mathbf{q} = 0$ or *non-vertical* if the photon momentum \mathbf{q} that is transferred to the electron is not negligible ($\mathbf{q} \neq 0$). Intraband transitions occur only in metals, where the Fermi level cross the valence band and, within the same band, some states are occupied and some other are empty. In semiconductors and insulators, where the Fermi level lie inside the band gap, the valence band is completely filled and the conduction band is completely empty, only interband transitions occur. From the infrared to the ultraviolet region of the electromagnetic waves spectrum, the photon momentum is negligible, $\mathbf{q} \simeq 0$,

and only vertical transition are allowed. In the small \mathbf{q} limit this model can be reconduced to a generalized Lorentz model where the intraband term is responsible for the Drude behaviour while the interband terms give rise to Lorentzian peaks. The first step is to use the $\mathbf{K} \cdot \mathbf{p}$ approximation to expand the matrix element in series for $\mathbf{q} \rightarrow 0$, it is convenient to treat separately the $n = n'$ term of intraband transitions:

$$\begin{aligned} \bar{\epsilon}(\mathbf{q}, \omega) = 1 + \frac{4\pi e^2}{\mathbf{q}^2 V} \sum_{\mathbf{K}, n} \frac{(f_{\mathbf{K}+\mathbf{q}, n} - f_{\mathbf{K}, n})}{\hbar\omega + i\gamma\hbar} - \\ - \frac{4\pi e^2 \hbar^2}{m^2 V} \sum_{\mathbf{K}, n \neq n'} \frac{|\langle \mathbf{K}, n' | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle|^2}{E_{\mathbf{K}, n'} - E_{\mathbf{K}, n} - \hbar\omega - i\gamma\hbar} \frac{f_{\mathbf{K}, n'} - f_{\mathbf{K}, n}}{(E_{\mathbf{K}, n'} - E_{\mathbf{K}, n})^2} \end{aligned} \quad (1.128)$$

where $\hat{\mathbf{q}}$ is the versor of the perturbation momentum while $\hat{\mathbf{p}}$ represent the momentum operator associated to the electrons participating to the response, i.e. involved in the medium excitation. For the interband term it has been sufficient to put $\mathbf{q} = 0$ whereas in the intraband case this causes a divergence. To overcome this problem one must play some algebraic trick but in the end

$$\begin{aligned} \bar{\epsilon}(\mathbf{q}, \omega) = 1 + \frac{4\pi e^2 \hbar^2}{m^2 V} \sum_{\mathbf{K}, n} \delta(E_{Fermi} - E_{\mathbf{K}, n}) \frac{|\langle \mathbf{K}, n | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle|^2}{\hbar^2 \omega^2 + i\gamma \hbar^2 \omega} - \\ - \frac{8\pi e^2 \hbar^2}{m^2 V} \sum_{\mathbf{K}, n \neq n'} \frac{|\langle \mathbf{K}, n' | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle|^2}{(E_{\mathbf{K}, n'} - E_{\mathbf{K}, n})^2 - \hbar^2 \omega^2 - i\gamma \hbar^2 \omega} \frac{f_{\mathbf{K}, n}}{(E_{\mathbf{K}, n'} - E_{\mathbf{K}, n})} \end{aligned} \quad (1.129)$$

a trick as been played also to the denominator of the interband term. Now using the plasma frequency definition, placing $\hbar\omega_{\mathbf{K}, n, n'} = E_{\mathbf{K}, n'} - E_{\mathbf{K}, n}$ and defining the oscillators strength as:

$$\begin{aligned} f_{\mathbf{K}, n, n'} = 2f_{\mathbf{K}, n} \frac{|\langle \mathbf{K}, n' | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle|^2}{mN\hbar\omega_{\mathbf{K}, n, n'}} \\ f_{\mathbf{K}, n, n} = \delta(E_{Fermi} - E_{\mathbf{K}, n}) \frac{|\langle \mathbf{K}, n | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle|^2}{mN} \end{aligned} \quad (1.130)$$

one obtain the classical generalized Lorentz model⁴

$$\epsilon(\omega) = 1 + \sum_{\mathbf{K}, n} \frac{\Omega_p^2}{\omega^2 + i\gamma\omega} f_{\mathbf{K}, n, n} - \sum_{\mathbf{K}, n \neq n'} \frac{\Omega_p^2}{\omega_{\mathbf{K}, n, n'}^2 - \omega^2 - i\gamma\omega} f_{\mathbf{K}, n, n'} \quad (1.131)$$

N being the number of electron in the unit cell, it guarantees the respect of the rule (1.115). Notice that, in the case of a semiconductor, the Fermi level lies inside the gap and the Dirac's delta function $\delta(E_{Fermi} - E_{\mathbf{K}, n})$ appearing in (1.130) is always zero: no intraband contribution exist for semiconductors and insulators.

A more complicated but analogous treatment can be performed in the case of a transverse perturbation

$$\mathbf{A}(\mathbf{x}, \mathbf{t}) = \sum_{\mathbf{q}} \mathbf{e} A(\mathbf{q}, \omega) e^{i\mathbf{q} \cdot \mathbf{x}} e^{-i\omega t + \gamma t} \quad (1.132)$$

this leads to an expression analogous to the longitudinal dielectric function (1.129), but now the matrix element is:

$$\langle \mathbf{K}, n | \hat{\mathbf{e}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle \quad (1.133)$$

⁴One may notice that, in the classical models, the transverse frequencies ω_i represent the orbital frequencies with which the electrons move of uniform circular motion around the nucleus, whereas, in the quantum model $\omega_{\mathbf{K}, n, n'}$ represent the energy difference between two quantized levels. The Bohr correspondence principle states that, for large quantum numbers, quantum physics meets the classical one. In the case of an isolated atom, for large quantum numbers, the frequency of each quantum orbit is almost equal to the frequency difference between two consecutive orbits, i.e. $\omega_{\mathbf{K}, n, n'} \rightarrow \omega_i$

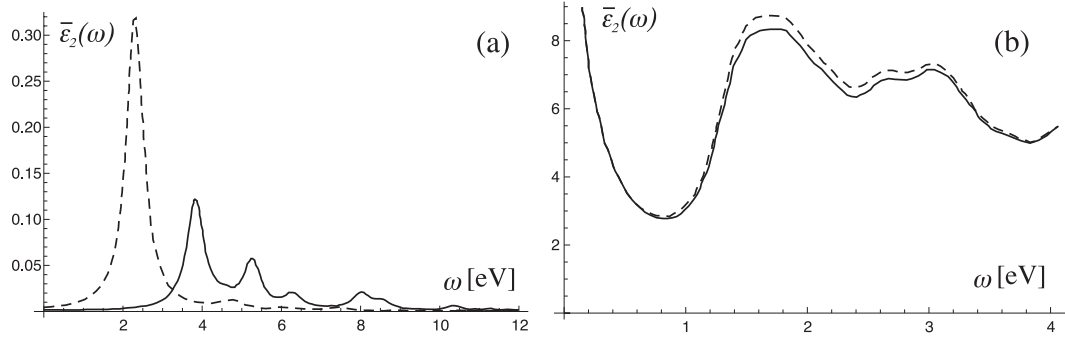


Figure 1.11: Comparison between imaginary parts of the dielectric function with (continuous line) and without (dashed line) local field effect for (a) a beryllium atom and (b) an isotropic copper bulk. Figure adapted from referece [15].

where $\hat{\mathbf{e}}$ is the polarization versor, perpendicular to $\hat{\mathbf{q}}$ by definition of electromagnetic wave. With the proper choice of the reference system, the perturbation vector $\hat{\mathbf{q}}$ can entirely lie on the x -axis, in that case the longitudinal response function (1.129) coincides with the xx component of the tensor, and, by analogy with the convention used in (1.37), it is labeled as ϵ^{LL} . Now $\hat{\mathbf{q}} \cdot \hat{\mathbf{e}} = 0$ and the transverse dielectric function enters the matrix elements $\epsilon^{T_1 T_1}$ and $\epsilon^{T_2 T_2}$ and in the mixing off diagonal term $\epsilon^{T_1 T_2} = \epsilon^{T_2 T_1}$. If one is interested in the exploitation of the other off diagonal terms, mixing the transverse and longitudinal response, one must take into account both the transverse and the longitudinal perturbations at once, following the Adler way [13].

This quantum model comes from a single particle picture and can be found as the non interacting limit of a many-body treatment, as Ehrenreich and Cohen shown [12].

When this model is applied to a free non interacting electron gas, whose wavefunctions are described by plane waves, the model takes the name of *Lindhard model* [7].

Adler and Wisner approximation This model is the simple extension of the previous one with the inclusion of the local field effect described in section 1.5, in the case of the longitudinal response, the perturbing potential is:

$$\phi(\mathbf{K}, \mathbf{t}) = \sum_{\mathbf{K}, \mathbf{G}} \phi(\mathbf{K}, \mathbf{G}, \omega) e^{i(\mathbf{K}+\mathbf{G}) \cdot \mathbf{x}} e^{-i\omega t + \gamma t} \quad (1.134)$$

and, in the limit of small \mathbf{K} , the terms of the expression (1.73) becomes:

$$\begin{aligned} \bar{\epsilon}(\mathbf{K} + \mathbf{G}, \mathbf{K}) &\rightarrow \frac{4\pi e^2}{mV\hbar} \sum_{\mathbf{K}, n \neq n'} \frac{\langle \mathbf{K}, n | e^{i\mathbf{G} \cdot \mathbf{x}} | \mathbf{K}, n' \rangle \langle \mathbf{K}, n' | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n \rangle}{\omega_{\mathbf{K}, n, n'} (\omega_{\mathbf{K}, n, n'} - \omega - i\gamma\omega)} (f_{\mathbf{K}, n'} - f_{\mathbf{K}, n}) \\ \bar{\epsilon}(\mathbf{K} + \mathbf{G}', \mathbf{K} + \mathbf{G}) &\rightarrow \delta_{\mathbf{G}', \mathbf{G}} - \frac{4\pi e^2}{V\hbar} \frac{1}{|\mathbf{G}'|^2} \sum_{\mathbf{K}, n \neq n'} \times \\ &\times \frac{\langle \mathbf{K}, n | e^{i\mathbf{G} \cdot \mathbf{x}} | \mathbf{K}, n' \rangle \langle \mathbf{K}, n' | e^{-i\mathbf{G}' \cdot \mathbf{x}} | \mathbf{K}, n \rangle}{\omega_{\mathbf{K}, n, n'} (\omega_{\mathbf{K}, n, n'} - \omega - i\gamma\omega)} (f_{\mathbf{K}, n'} - f_{\mathbf{K}, n}) \\ \bar{\epsilon}(\mathbf{K}, \mathbf{K} + \mathbf{G}') &\rightarrow \frac{4\pi e^2}{mV\hbar} \frac{1}{|\mathbf{G}'|^2} \sum_{\mathbf{K}, n \neq n'} \frac{\langle \mathbf{K}, n | \hat{\mathbf{q}} \cdot \hat{\mathbf{p}} | \mathbf{K}, n' \rangle \langle \mathbf{K}, n' | e^{-i\mathbf{G}' \cdot \mathbf{x}} | \mathbf{K}, n \rangle}{\omega_{\mathbf{K}, n, n'} (\omega_{\mathbf{K}, n, n'} - \omega - i\gamma\omega)} \times \\ &\times (f_{\mathbf{K}, n'} - f_{\mathbf{K}, n}) \end{aligned} \quad (1.135)$$

while the term $\bar{\epsilon}(\mathbf{K}, \mathbf{K})$ for $\mathbf{K} \rightarrow 0$ is nothing but the expression (1.129) derived by Ehrenreich and Cohen, so that the other terms can be seen as a correction to the previous model.

Wiser shown that such a corrections involves, at least for small \mathbf{K} , only the interband contribution to the dielectric function, the intraband term corrections cancel out [16]. The local field effect is extremely important for strongly anisotropic and inhomogeneous media or for low dimensional systems such as polymer crystal, surfaces, nanowires, molecules, or isolated atoms. Figure 1.11 shows the very different contribution of the local field corrections for a single beryllium atom and an homogeneous copper bulk, in the latter case the corrections are almost negligible.

It has been shown, in the preceding sections, that the optical properties of linear media depends upon occupied and empty states, the failure in treating empty states, makes the DFT results only a rough starting point in the optical properties calculation. The absorption and emission processes involve the promotion of a valence electron onto a conduction band and the corresponding creation of an hole, as the consequence or the starting point for the photon absorption or emission. Even if they are non-local, the approximate DFT functionals are not able to account properly neither for the relaxation of the electronic structure due to the presence of the hole, nor to account for the screening played by the medium in the electron-hole interaction. The GW and BSE techniques allow one to improve the energy dispersion structure and wavefunctions of the system, including a part of the many body interaction of the hole with the excited electrons of the system. Figure 1.9 shows a comparison between the silicon experimental absorption spectrum and the ones obtained with different theoretical techniques. The DFT absorption spectrum is the lowest level calculation, obtained directly via the DFT eigenvalues and eigenfunctions, the gap is underestimated with respect to the experimental one. GW correction reopen the gap right shifting the spectrum, in the case of bulk silicon this is the only visible effect of GW correction, for some other materials the shift can be non-rigid leading to a deformation of the spectrum and to a change of the spectral weights. Increasing the degree of accuracy, one can use a BSE approach to account for the interaction of the hole with the corresponding excited electron, the so called *exciton*. The BSE correction move spectral weight from the right to the left, to lower energy. The concepts of ab-initio optics and all the techniques that go beyond the DFT are reviewed in reference [6].

1.11 The Wood and Ashcroft model

Starting from the previous quantum model, D.M. Wood and N.W. Ashcroft derived a dielectric tensor that describes the response of a medium whose non interacting electrons are strongly confined in one or more dimensions [14]. As in the ordinary Lindhard model [7], the non interacting electrons are described by plane waves, but now these plane waves must satisfy certain boundary conditions. If the boundary conditions are applied only along one direction the model describes an ultra-thin metallic film (quantum well), if the boundary conditions are applied along all the three directions the model describes the response of a quantum dot (and it is some times named the *particle in a box* model). The electrons confined in the thin film of figure 1.12 (a) are described by the eigenfunctions of an infinitely deep square quantum well:

$$\psi_{\mathbf{k}_{\parallel},n}(\mathbf{x}) = \sqrt{\frac{2}{V}} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} \sin(\mathbf{k}_{\perp} z) = \sqrt{\frac{2}{V}} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} \sin\left(\frac{n\pi}{d} z\right) \quad (1.136)$$

with eigenvalues:

$$E_{\mathbf{k}_{\parallel},n} = E_{\parallel} + E_{\perp} = \frac{\hbar^2}{2m} (k_{\parallel}^2 + k_{\perp}^2) = \frac{\hbar^2}{2m} \left(k_{\parallel}^2 + \frac{n^2 \pi^2}{d^2} \right) \quad (1.137)$$

\mathbf{k}_{\parallel} and \mathbf{x}_{\parallel} are entirely contained in the xy plane, imposing the quantization conditions along z one gets $\mathbf{k}_{\perp} = n\pi/d$ with n positive integer.

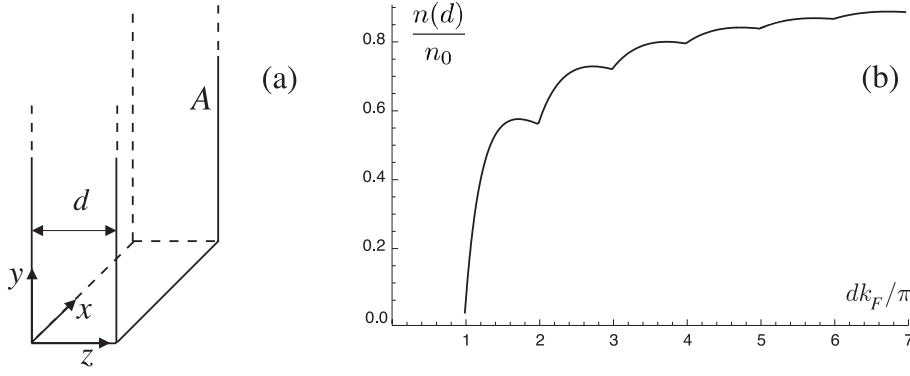


Figure 1.12: (a) Sketch of the infinitely deep quantum well used by the Wood and Ashcroft. (b) Ratio $n(d)/n_0$ as a function of dk_F/π for $k_F = 8.5 \times 10^9 1/m$, the same plot can be obtained keeping d fixed for different k_F values.

The confinement effects in a homogeneous electron gas enter the dielectric response in many ways, first of all they affect the electron density that becomes a function of the film thickness d . The electron density can be calculated starting from the density of states:

$$\begin{aligned} n(E) &= \sum_n \frac{A}{4\pi^2} \int d\mathbf{k} \delta(E - E_{\mathbf{k}_{\parallel}, n}) = \\ &= \sum_n \frac{Am}{2\pi\hbar^2} \int dE_{\parallel} \delta(E - E_{\parallel} - E_{\perp}) = \frac{Am}{2\pi\hbar^2} \sum_n \theta\left(E - \frac{\hbar^2\pi^2 n^2}{2md^2}\right) \end{aligned} \quad (1.138)$$

where A is the film surface, the integral definition of the step function θ has been used. Now the electron density can be obtained integrating the density of states over the energy, up to the Fermi's energy E_F :

$$n = \frac{2}{Ad} \cdot \int_0^{E_F} n(E) dE = \frac{m}{\pi\hbar^2 d} \sum_n \int_0^{E_F} \theta(E - n^2 E_0) dE \quad (1.139)$$

where the factor 2 appears to account for the fermionic nature of the electrons, Ad is the thin film volume and $E_0 = \hbar^2\pi^2 n^2/2md^2$. Using the step function properties the sum over n can be made explicit as:

$$\begin{aligned} n(d) &= \frac{m}{\pi\hbar^2 d} \left[\int_{E_0}^{E_F} dE + \int_{4E_0}^{E_F} dE + \dots + \int_{m_0^2 E_0}^{E_F} dE \right] = \frac{m}{\pi\hbar^2 d} \\ & (E_F - E_0 + E_F - 4E_0 + \dots) = \frac{m}{\pi\hbar^2 d} \left[m_0 E_F - E_0 \sum_{n=1}^{m_0} n^2 \right] \end{aligned} \quad (1.140)$$

where m_0 represents the integer part of $m_F = k_F d/\pi \geq m_0$. The corresponding density for a free electron gas n_0 is given by:

$$n_0 = \frac{k_F^3}{3\pi^2} = \frac{(2mE_F)^3}{3\pi^2\hbar^6} \quad (1.141)$$

The ratio $n(d)/n_0$:

$$\frac{n(d)}{n_0} = \frac{3m_0}{2m_F} \left(1 - \frac{1}{m_0 m_F^2} \sum_{n=1}^{m_0} n^2 \right) \quad (1.142)$$

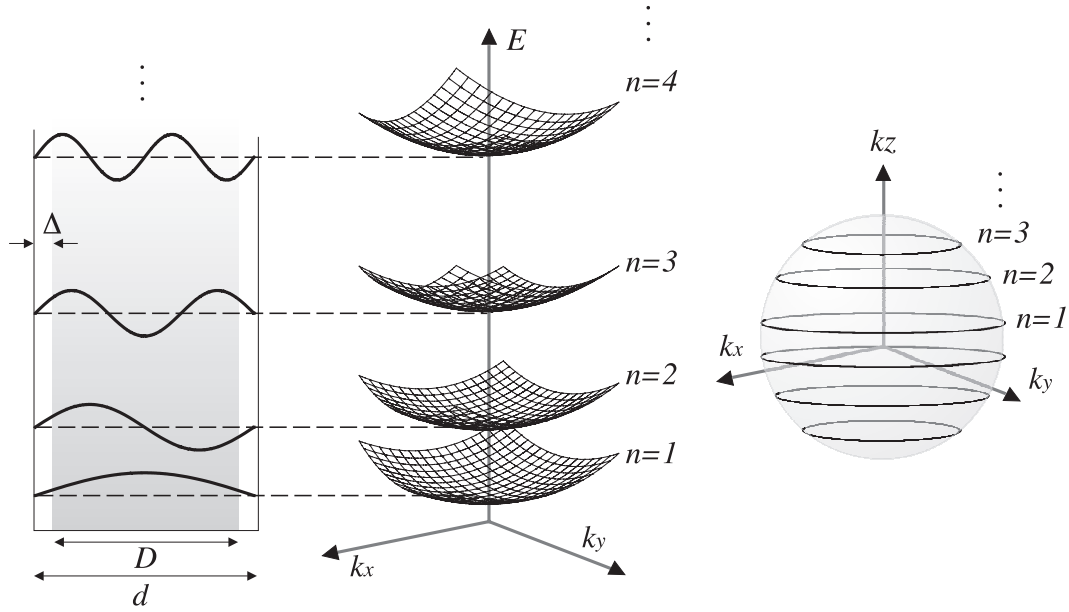


Figure 1.13: Sketch of the quantization of electrons wavefunctions for a thin film of thickness D , the electron charge can spill-out of the material of a quantity Δ . The parabolic band structure and the sliced Fermi surface of the system are also shown.

is plotted in figure 1.12 (b), cusps are present each time that a new discrete level fall below the Fermi's energy. Notice that, for large d values or large Fermi's energy, m_0 and $m_F \gg m_0 - m_F$ or in other words $m_0 \simeq m_F$ and the ratio $n(d)/n_0$ goes to one. When the well is large or when a huge number of discrete energy levels is filled, the quantization effects are small and one gets again the free electron gas behaviour.

A first problem arise: the d dependent electron density does not preserve the total charge. Imagine to have a free electron gas of density n_0 , within the jellium model of metals, a positive charge equal to n_0 must exist to balance the negative charge keeping the medium neutral. This positive charge is naturally due to the presence of the ionic lattice. Now imagine to slice away the ionic lattice until a thin film of thickness d is reached: the positive charge density is again n_0 but, due to the confinement, the electron density is now given by (1.140). Positive and negative charges are no more balanced:

$$n(d)Ad \neq n_0Ad \quad (1.143)$$

i.e. the total negative charge is different from the total positive charge of the thin film. To avoid this problem the electrons should be left free to spill-out of the film. This should be done describing the film with a finite potential well, but in this way the boundary conditions becomes too complicated and the model can not be solved analytically. Another possibility is to use again an infinitely deep squared well, but with a thickness d different from the ionic one. Naming D the ionic film thickness, d can be calculated using the relation (1.143):

$$nAd = n_0AD \quad \longrightarrow \quad d = \frac{n(d)}{n_0}D \quad (1.144)$$

and the electron spill-out can be defined as $\Delta = (d - D)/2$. This new thin film model is sketched in figure 1.13 together with its electronic band structure and the Fermi's surface. The spill-out Δ depends upon the Fermi's energy E_F and the ionic thickness D , for large ionic thicknesses ($m_0 \simeq m_F$) it reaches a constant value $\Delta = 3\pi/8k_F$ that is a bulk property

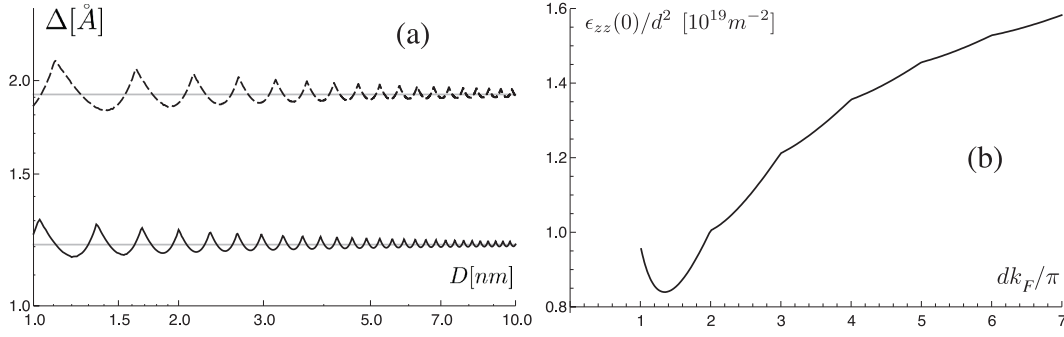


Figure 1.14: (a) Spill-out Δ as a function of the ionic thickness D , continuous line $k_F = 8.5 \times 10^91/m$, dashed line $k_F = 6 \times 10^91/m$, gray lines represent the limiting bulk values. (b) Static value of the zz component of the dielectric tensor as a function of dk_F/π for $k_F = 8.5 \times 10^91/m$.

of each material. Figure 1.14 (a) shows a plot of Δ as a function of D for thin films of different densities.

To calculate the dielectric response one has to use the RPA, described in the previous section. In particular it can be shown that, with some algebraic tricks, the definition (1.129) can be rewritten as:

$$\epsilon(\omega) = 1 - \frac{\Omega_p^2}{\omega^2} - \frac{8\pi e^2}{m^2 A d \omega^2} \sum_{\mathbf{k}_{\parallel}, n, n'} \frac{(E_{\mathbf{k}_{\parallel}, n'} - E_{\mathbf{k}_{\parallel}, n}) |\langle \psi_{\mathbf{k}_{\parallel}, n} | \hat{\mathbf{Q}} \cdot \hat{\mathbf{P}} | \psi_{\mathbf{k}_{\parallel}, n'} \rangle|^2}{(E_{\mathbf{k}_{\parallel}, n'} - E_{\mathbf{k}_{\parallel}, n})^2 - \hbar^2 \omega^2} f_{\mathbf{k}_{\parallel}, n} \quad (1.145)$$

notice that this is not a simple separation of the Drude intraband term: the sum over the initial and final states accounts again for all the transitions between states of the same band. The imaginary part has been suppressed for the seek of simplicity, this form of the dielectric function holds even for a non zero γ value. The quantization along the z direction introduces an anisotropy in the dielectric response, in fact all the off-diagonal components of the dielectric tensor vanish, and one get $\epsilon_{xx} = \epsilon_{yy} \neq \epsilon_{zz}$. This anisotropy is due to the dipole matrix elements appearing in the previous ϵ expression, for the x and y components one has:

$$\langle \psi_{\mathbf{k}_{\parallel}, n} | p_x | \psi_{\mathbf{k}_{\parallel}, n} \rangle = \langle \psi_{\mathbf{k}_{\parallel}, n} | p_y | \psi_{\mathbf{k}_{\parallel}, n} \rangle = -i\hbar \mathbf{k}_{\parallel} \delta_{n, n'} \quad (1.146)$$

whereas in the z case one has:

$$\langle \psi_{\mathbf{k}_{\parallel}, n} | p_z | \psi_{\mathbf{k}_{\parallel}, n} \rangle = -i\hbar \frac{2}{d} \frac{n'n}{n^2 - n'^2} \left(1 - (-1)^{n'+n} \right) \quad (1.147)$$

Notice that p_x and p_y are non vanishing only for $n = n'$ whereas, in this condition, the p_z vanish: a product of two different \mathbf{p} components is always zero, i.e. all the off-diagonal components of the dielectric tensor vanish. So expression (1.145) takes the form:

$$\begin{aligned} \epsilon_{ij}(\omega) &= 0 \quad \forall \quad i \neq j \\ \epsilon_{xx}(\omega) &= \epsilon_{yy}(\omega) = 1 - \frac{\Omega_p^2}{\omega^2} \\ \epsilon_{zz}(\omega) &= 1 - \frac{\Omega_p^2}{\omega^2} - \frac{128e^2}{Adm\pi\omega^2} \sum_{\mathbf{k}_{\parallel}, n, n'} f_{\mathbf{k}_{\parallel}, n} \frac{n'^2 n^2}{n^2 - n'^2} \frac{1 - (-1)^{n+n'}}{(n^2 - n'^2)^2 - x^2} \end{aligned} \quad (1.148)$$

where the plasma frequency Ω_p contains the density of the confined gas $n(d)$ and $x = 2m\omega d^2/\pi^2\hbar$. The $\epsilon_{zz}(\omega)$ component can be adjusted further: with the assumption that

the thin film is infinitely extended along the xy plane, the sum over \mathbf{k}_{\parallel} can be replaced by a bidimensional integral and:

$$\begin{aligned} \sum_{\mathbf{k}_{\parallel}} f_{\mathbf{k}_{\parallel},n} &= \frac{A}{4\pi^2} \int d\mathbf{k}_{\parallel} f_{\mathbf{k}_{\parallel},n} = \frac{A}{2\pi} \int k dk f\left(E_F - \frac{\hbar^2 \pi^2 n^2}{2md^2} - \frac{\hbar^2 \mathbf{k}_{\parallel}^2}{2m}\right) = \\ &= \frac{Am}{\hbar^2 \pi} \int_0^{E_F - \hbar^2 \pi^2 n^2 / 2md^2} dE_{\parallel} = \frac{Am}{\hbar^2 \pi} \left(E_F - \frac{\hbar^2 \pi^2 n^2}{2md^2}\right) \theta\left(E_F - \frac{\hbar^2 \pi^2 n^2}{2md^2}\right) \end{aligned} \quad (1.149)$$

where θ is the Heaviside step function. Concluding:

$$\epsilon_{zz}(\omega) = 1 - \frac{\Omega_p^2}{\omega^2} - \frac{64e^2}{d^3 m \omega^2} \sum_{n'=1}^{\infty} \sum_{n=1}^{m_0} (m_F^2 - n^2) \frac{n'^2 n^2}{n^2 - n'^2} \frac{1 - (-1)^{n+n'}}{(n^2 - n'^2)^2 - x^2} \quad (1.150)$$

the step function disappeared but now the sum over n goes from 1 to m_0 only. While the quantized electron gas behaves as an usual Drude metal along the x and y directions, it has a semiconducting character along the z direction. The n and n' sums describe all the possible transitions between the semi-occupied bands, the Drude term cancels out with the $n = n'$ term giving rise to a finite static value of the dielectric function. In fact, at low frequency one can expand the sum over n' to get:

$$\sum_{n'=1}^{\infty} \frac{n'^2}{n^2 - n'^2} \frac{1 - (-1)^{n+n'}}{(n^2 - n'^2)^2 - x^2} \simeq -\frac{\pi^2}{32n^2} - \frac{x^2 \pi^2}{1536n^4} \left(\frac{15}{n^2} - \pi^2\right) \quad (1.151)$$

retaining only the first order term one ends up with:

$$\begin{aligned} \epsilon_{zz}(\omega \rightarrow 0) &\simeq 1 - \frac{\Omega_p^2}{\omega^2} + \frac{2\pi^2 e^2}{m\omega^2 d^3} \sum_{n=1}^{m_0} (m_F^2 - n^2) = \\ &= 1 - \frac{4\pi e^2}{m\omega^2 n(d)} + \frac{2\pi^2 e^2}{m\omega^2 d^3} (m_0 m_F^2 - \sum_{n=1}^{m_0} n^2) = 1 \end{aligned} \quad (1.152)$$

to reach the cancellation result, (1.140) has been used. Including the second order term of the expansion (1.151) the static value becomes:

$$\begin{aligned} \epsilon_{zz}(0) &= 1 + \frac{dme^2}{\pi^2 \hbar^2} \frac{1}{6} [15(S_4 m_F^2 - S_2) + \pi^2 (m_0 - S_2 m_F^2)] \\ S_q &= \sum_{n=1}^{m_0} \frac{1}{n^q} \end{aligned} \quad (1.153)$$

This expression diverges only for big d values, when the confinement is negligible, recovering the ordinary bulk metal behaviour.

Following the Wood and Ashcroft work [14], it is possible to calculate the sum of n' series:

$$\epsilon_{zz}(\omega) = 1 - \frac{\Omega_p^2}{\omega^2} - \frac{64e^2}{d^3 m \omega^2} \sum_{n=1}^{m_0} (m_F^2 - n^2) \frac{n^2}{2x^2} \Xi(n, x) \quad (1.154)$$

with:

$$\Xi(n, x) = \begin{cases} \frac{\pi}{2} \sqrt{n^2 - x} \operatorname{tg}\left(\frac{\pi}{2} \sqrt{n^2 - x}\right) + \frac{\pi}{2} \sqrt{n^2 + x} \operatorname{tg}\left(\frac{\pi}{2} \sqrt{n^2 + x}\right) & n \text{ even} \\ \frac{\pi}{2} \sqrt{n^2 + x} \operatorname{cotg}\left(\frac{\pi}{2} \sqrt{n^2 + x}\right) + \frac{\pi}{2} \sqrt{n^2 - x} \operatorname{cotg}\left(\frac{\pi}{2} \sqrt{n^2 - x}\right) & n \text{ odd} \end{cases} \quad (1.155)$$

Finally, the extension of this model to dissipative media (non zero γ value) is not so straight forward, a solution has been proposed by N. D. Mermin [8].

Appendix A: Linear response theory

This general theory it is useful in the study of the interaction of external perturbation with a medium. It provides a relation between some physically observable quantity and another observable used as a perturbation. Consider a system governed by the hamiltonian \mathcal{H}_0 , if an external force $K(t)$ is applied to the system from the infinite past ($t = -\infty$) when the system was in thermal equilibrium, the hamiltonian becomes:

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{EXT}}(t) = \mathcal{H}_0 - A(t)K(t) \quad (1.156)$$

where $A(t)$ is the conjugate quantity of the applied perturbation $K(t)$. Using classical statistical mechanics the system evolves in time driven by the Liouville's equation:

$$-\frac{\partial \rho(t)}{\partial t} = \{\mathcal{H}, \rho(t)\} = -i\mathcal{L}\rho(t) \quad (1.157)$$

the brackets in the RHS of the equation represent the Poisson's operator, $\rho(t)$ is the distribution function representing a certain statistical ensemble, \mathcal{L} is called Liouville's differential operator. In the framework of quantum mechanics, the system evolves according to the Schroedinger equation for the density matrix $\hat{\rho}(t)$:

$$-\frac{\partial \hat{\rho}(t)}{\partial t} = \frac{i}{\hbar} [\hat{\mathcal{H}}, \hat{\rho}(t)] = -i\hat{\mathcal{L}}\hat{\rho}(t) \quad (1.158)$$

now the brackets on RHS indicate a commutation operator. Thanks to the linearity of the Poisson's brackets (as well as of the commutator), equations (1.157) and (1.158) become

$$-\frac{\partial \rho(t)}{\partial t} = \{\mathcal{H}_0, \rho(t)\} + \{\mathcal{H}_{\text{EXT}}, \rho(t)\} = -i\mathcal{L}_0\rho(t) - i\mathcal{L}_{\text{EXT}}\rho(t) \quad (1.159)$$

If the applied perturbation is small a perturbative expansion can be performed as follow:

$$\rho(t) = \rho_0 + \Delta\rho(t) + O[\rho(t)^2] \quad (1.160)$$

ρ_0 is the equilibrium density, defined as:

$$\rho_0 = \rho(-\infty) = Ce^{-\beta\mathcal{H}_0} \quad (1.161)$$

Using (1.160) into (1.157) and (1.158) and neglecting second order terms, one obtains a non homogeneous, first order, ordinary differential equation in $\Delta\rho(t)$, whose solution is:

$$\Delta\rho(t) = i \int_{-\infty}^t dt' e^{i(t-t')\mathcal{L}_0} \mathcal{L}_{\text{EXT}}(t')\rho_0 \quad (1.162)$$

Both in the classical and quantum statistical mechanics, the expectation value for an observable operator is given by the trace of the product between the operator matrix and the density matrix, or equivalently, the integration of the operator function over all the phase space weighted by the distribution function. Also the variation of an observable can be expressed in the same way using the variation of the density due to the perturbation:

$$\langle \hat{\Delta B}(t) \rangle = Tr \left\{ \hat{\Delta\rho}(t) \hat{B} \right\} \quad \langle \Delta B(q, p, t) \rangle = \int d\mathbf{x} \int d\mathbf{p} \Delta\rho(t) B(\mathbf{x}, \mathbf{p}) \quad (1.163)$$

and using now the solution (1.162) one comes to:

$$\begin{aligned}
\langle \hat{\Delta B}(t) \rangle &= Tr \left\{ i \int_{-\infty}^t dt' e^{i(t-t')\mathcal{L}_0} \mathcal{L}_{\text{EXT}}(t') \rho_0 B \right\} = \\
&= i \int_{-\infty}^t dt' Tr \left\{ e^{i(t-t')\mathcal{L}_0} \mathcal{L}_{\text{EXT}}(t') \rho_0 B \right\} = \\
&= -i \int_{-\infty}^t dt' Tr \left\{ \rho_0 \mathcal{L}_{\text{EXT}}(t') e^{i(t-t')\mathcal{L}_0} B \right\} =
\end{aligned} \tag{1.164}$$

the permutations needed to change the order of the operators, in the last step of previous equation, are in odd number, this change the sign of the trace. In classical statistical mechanics the application of the operator $e^{i(t-t')\mathcal{L}_0}$ to an observable lead to the time evolution of the observable as if the system was not perturbed, in other words:

$$e^{i(t-t')\mathcal{L}_0} B(q, p, -\infty) = B(q, p, t - t') \tag{1.165}$$

In the quantum treatment this becomes:

$$e^{i(t-t')\mathcal{L}_0} \hat{B}(-\infty) = e^{i(t-t')\mathcal{H}/\hbar} \hat{B}(-\infty) e^{-i(t-t')\mathcal{H}/\hbar} \tag{1.166}$$

that is again the evolution of the operator driven by the unperturbed hamiltonian term only. Finally, using the previous definition of Liouville's operator, one reaches the expression:

$$\langle \Delta B(t) \rangle = \int_{-\infty}^t dt' K(t') Tr \left\{ \rho_0 \{A(t'), B(t-t')\} \right\} = \int_{-\infty}^t dt' K(t') \langle \{A(t'), B(t-t')\} \rangle \tag{1.167}$$

$K(t)$ is just a function and it can come out of the Poisson's brackets, the average present in the last step is an ensemble average exactly like (1.163), it takes the name of *response function*:

$$\langle \Delta B(t) \rangle = \int_{-\infty}^t dt' K(t') \phi_{\text{AB}}(t-t') \tag{1.168}$$

A response function $\phi_{\text{AB}}(t-t')$ is a correlation function that relates together the perturbation observable A and the response observable variation ΔB throughout a convolution, the response variation at time t depends on the values assumed by the perturbation during past time $t' \leq t$. This causality principle gives the response function some interesting properties that will be analyzed later on. A more general theory allow to include also spatial non-locality:

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\text{EXT}}(t) = \mathcal{H}_0 - \frac{1}{V} \int d\mathbf{x} A(\mathbf{x}, t) K(\mathbf{x}, t) \tag{1.169}$$

\mathcal{L}_{EXT} becomes an integral operator, and solution (1.162) leads to:

$$\begin{aligned}
\langle \Delta B(\mathbf{x}, t) \rangle &= \int_{-\infty}^t dt' Tr \left\{ \rho_0 \left\{ \frac{1}{V} \int d\mathbf{x}' A(\mathbf{x}', t') K(\mathbf{x}', t'), B(\mathbf{x}, t-t') \right\} \right\} = \\
&= \frac{1}{V} \int_{-\infty}^t dt' \int d\mathbf{x}' K(\mathbf{x}', t') Tr \left\{ \rho_0 \{A(\mathbf{x}', t'), B(\mathbf{x}, t-t')\} \right\} = \\
&= \frac{1}{V} \int_{-\infty}^t dt' \int d\mathbf{x}' K(\mathbf{x}', t') \phi_{\text{AB}}(\mathbf{x}, \mathbf{x}', t-t')
\end{aligned} \tag{1.170}$$

now the value of the response in each single point \mathbf{x} of the volume V depends on the value assumed by the perturbation elsewhere.

Two events separated by an infinite distance or by an infinite time interval can not be causally related, this means that the response function relating them must vanish if $|t - t'| \rightarrow \infty$ or if $|\mathbf{x} - \mathbf{x}'| \rightarrow \infty$:

$$\phi_{AB}(\mathbf{x}, \mathbf{x}', t - t') \rightarrow 0 \quad (1.171)$$

Thanks to the convolution theorem, in reciprocal space equation (1.170) becomes⁵:

$$\langle \Delta B(\mathbf{q}, \omega) \rangle = \frac{1}{V} \int d\mathbf{q}' K(\mathbf{q}', \omega) \chi_{AB}(\mathbf{q}, \mathbf{q}', \omega) \quad (1.172)$$

Under the further assumption of spatial translational invariance, i.e. $\phi_{AB}(\mathbf{x}, \mathbf{x}', t - t') = \phi_{AB}(\mathbf{x} - \mathbf{x}', t - t')$, one gets a simple algebraic equation:

$$\langle \Delta B(\mathbf{q}, \omega) \rangle = \chi_{AB}(\mathbf{q}, \omega) K(\mathbf{q}, \omega) \quad (1.173)$$

where:

$$\chi(\mathbf{q}, \omega)_{AB} = \frac{1}{V} \int d\mathbf{x} \int dt \phi_{AB}(\mathbf{x}, t) e^{i\mathbf{q}\cdot\mathbf{x} - i\omega t} \quad (1.174)$$

is the Fourier transform of the response function. Notice that for the time variable, the translational invariance is required by the causality principle and every physical response function can only be a function of $t - t'$.

The theory can be improved including also higher terms in the expansion (1.160), this will lead to non linear contributions needed in the case of strong magnitude of the perturbation. Equation (1.170) is extremely general and the average contained in the response function can be interpreted in many different ways, it can be regarded as a simple macroscopic average, as an ensemble average, as a quantum expectation value on a single well defined quantum state or as a statistic expectation value on an admixture of quantum states.

Since the response function relates together two observables, it must be a real function. Its Fourier's transform has therefore both real and imaginary components. The imaginary part is related to dissipation phenomena, as generally described by the Kubo's fluctuation-dissipation theorem. Some specific example will be given in the following sections, for a detailed analysis of the problem see references [17, 18].

A final remark concerns the scalar or tensorial nature of the response function, if it relates two scalar quantities it is a scalar function, if it relates a scalar and a vector or two vector, it can be either a scalar, a vector or a tensor:

$$\begin{aligned} \langle \Delta B(\mathbf{x}, t) \rangle &= \frac{1}{V} \int_{-\infty}^t dt' \int d\mathbf{x}' K_{\alpha}(\mathbf{x}', t') \phi_{\alpha}(\mathbf{x}, \mathbf{x}', t - t') \\ \langle \Delta B(\mathbf{x}, t) \rangle_{\alpha} &= \frac{1}{V} \int_{-\infty}^t dt' \int d\mathbf{x}' K_{\alpha}(\mathbf{x}', t') \phi(\mathbf{x}, \mathbf{x}', t - t') \\ \langle \Delta B(\mathbf{x}, t) \rangle_{\alpha} &= \frac{1}{V} \int_{-\infty}^t dt' \int d\mathbf{x}' \phi_{\alpha, \beta}(\mathbf{x}, \mathbf{x}', t - t') K_{\beta}(\mathbf{x}', t') \end{aligned} \quad (1.175)$$

⁵Moving to reciprocal space, by means of Fourier transform, requires further considerations about time causality, see appendix C.

where a sum over repeated indexes is subtended and where the response functions are defined has:

$$\begin{aligned}\phi_\alpha(\mathbf{x}, \mathbf{x}', t - t') &= Tr \left\{ \rho_0 \{ A_\alpha(\mathbf{x}', t'), B(\mathbf{x}, t - t') \} \right\} \\ \phi_{\alpha,\beta}(\mathbf{x}, \mathbf{x}', t - t') &= Tr \left\{ \rho_0 \{ A_\beta(\mathbf{x}', t'), B_\alpha(\mathbf{x}, t - t') \} \right\}\end{aligned}\quad (1.176)$$

Appendix B: Decomposition of a vector field into transverse and longitudinal parts

Any vector field that describes a real physical quantity can be uniquely decomposed into two vector fields, one of which is irrotational and the other divergenceless. The uniqueness of this decomposition is guaranteed by the Helmholtz's theorem. Consider a vector field \mathbf{F} that satisfies the equation:

$$\nabla^2 \mathbf{W} = -\mathbf{F} \quad (1.177)$$

\mathbf{W} being a vector potential of \mathbf{F} . The general solution of this equation is:

$$\mathbf{W}(\mathbf{x}) = \frac{1}{4\pi} \int d\mathbf{x}' \frac{\mathbf{F}(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \quad (1.178)$$

Now using the vector identity:

$$\nabla^2 \mathbf{W} = \nabla(\nabla \cdot \mathbf{W}) - \nabla \times (\nabla \mathbf{W}) \quad (1.179)$$

and defining a scalar function U and a vector field \mathbf{A} as follow:

$$\nabla \cdot \mathbf{W} = -U \quad \nabla \times \mathbf{W} = \mathbf{A} \quad (1.180)$$

one gets:

$$\mathbf{F} = \nabla U + \nabla \times \mathbf{A} \quad (1.181)$$

i.e. the well known Helmholtz's result, notice that U and \mathbf{A} have nothing to do with the scalar and vector potentials. Defining a longitudinal component \mathbf{F}^L and a transverse component \mathbf{F}^T :

$$\mathbf{F}^L = \nabla U \quad \mathbf{F}^T = \nabla \times \mathbf{A} \quad (1.182)$$

the field \mathbf{F} can be written as:

$$\mathbf{F} = \mathbf{F}^L + \mathbf{F}^T \quad (1.183)$$

As a consequence of their definitions, the new transverse and longitudinal fields obey the following relations:

$$\nabla \times \mathbf{F}^L = 0 \quad \nabla \cdot \mathbf{F}^T = 0 \quad (1.184)$$

The names transverse and longitudinal become more clear switching to reciprocal space:

$$\mathbf{q} \times \mathbf{F}^L = 0 \quad \mathbf{q} \cdot \mathbf{F}^T = 0 \quad (1.185)$$

By definition, an electromagnetic wave is made of transverse fields only and obey the second relation, a plasmon excitation can be represented with a longitudinal field only obeying the first relation. Now it is convenient to introduce the *dyad* concept, suppose to have a generic

vector, its longitudinal part is nothing but its projection along the \mathbf{q} direction, to get it, one should use the \mathbf{q} versor:

$$F^L = \mathbf{F} \cdot \hat{\mathbf{q}} \quad (1.186)$$

the projection is no more a vector, to recover the vectorial nature one must use again the \mathbf{q} versor

$$\mathbf{F}^L = F^L \hat{\mathbf{q}} = \mathbf{F} \cdot \hat{\mathbf{q}} \hat{\mathbf{q}} \quad (1.187)$$

the direct product of the two \mathbf{q} versors is the longitudinal dyadic product $\mathbf{1}^L$. The direct product of two vectors is a nine components matrix. One can also define a transverse dyadic product as follow:

$$\mathbf{1}^T = \mathbf{1} - \mathbf{1}^L \quad (1.188)$$

here $\mathbf{1}$ is the unit matrix. So a vector can be decomposed by mean of dyads into two parts

$$\mathbf{F} = \mathbf{F} \cdot \mathbf{1}^L + \mathbf{F} \cdot \mathbf{1}^T = \mathbf{F}^L + \mathbf{F}^T \quad (1.189)$$

A tensor can be decomposed in four parts:

$$\begin{aligned} \mathbf{T}^{LL} &= \mathbf{1}^L \cdot \mathbf{F} \cdot \mathbf{1}^L & \mathbf{T}^{TT} &= \mathbf{1}^T \cdot \mathbf{F} \cdot \mathbf{1}^T \\ \mathbf{T}^{LT} &= \mathbf{1}^L \cdot \mathbf{F} \cdot \mathbf{1}^T & \mathbf{T}^{TL} &= \mathbf{1}^T \cdot \mathbf{F} \cdot \mathbf{1}^L \end{aligned} \quad (1.190)$$

If one applies the decomposition rules defined above together with (1.185) to the fields appearing into Maxwell's equations one gets:

$$\begin{cases} \nabla \cdot \mathbf{D}^L = 4\pi\eta_{\text{EXT}} \\ \nabla \times \mathbf{H}^T - \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c} \mathbf{J}_{\text{EXT}} \\ \nabla \times \mathbf{E}^T + \frac{1}{c} \frac{\partial \mathbf{B}}{\partial t} = 0 \\ \nabla \cdot \mathbf{B}^L = 0 \end{cases} \quad (1.191)$$

Notice that, because of its divergenceless, every magnetic field is always transverse.

Appendix C: Dispersion relations, causality and stability

The causality principle states that *the cause always precedes the effect* or, equivalently, *the future has no effect on the past*. This universal statement, applied to definition (1.170), allows to relate together real and imaginary parts of the response functions and to infer some general rules on their behaviour. In practice, if the perturbation A is impulsively switched on at the time $t = t_0$ the variation of the response function, ΔB , can not be different from zero for any time $t < t_0$. In this case $K(\mathbf{x}, t) = K(\mathbf{x})\delta(t - t_0)$ and, replacing into (1.170) one gets:

$$\begin{aligned} \langle \Delta B(\mathbf{x}, t) \rangle &= \frac{1}{V} \int_{-\infty}^t dt' \int d\mathbf{x}' K(\mathbf{x}') \delta(t' - t_0) \phi_{AB}(\mathbf{x}, \mathbf{x}', t - t') = \\ &= \frac{1}{V} \int d\mathbf{x}' K(\mathbf{x}') \phi_{AB}(\mathbf{x}, \mathbf{x}', t - t_0) \end{aligned} \quad (1.192)$$

but ΔB must be zero for $t < t_0$ and this results in $\phi_{AB}(\mathbf{x}, \mathbf{x}', t - t_0) = 0$ for any $t < t_0$ or, more generally $\phi_{AB}(\mathbf{x}, \mathbf{x}', t - t') = 0$ for $t < t'$. It is only under this assumption that the

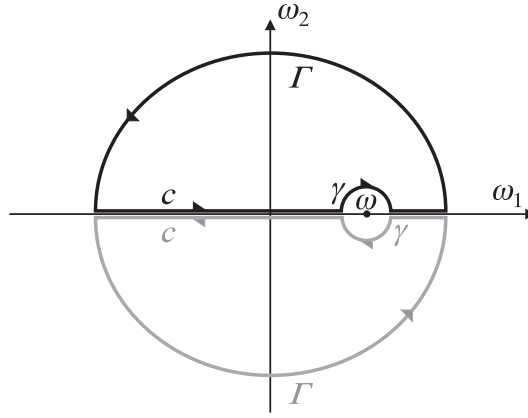


Figure 1.15: Possible integration contours for equation (1.195). No poles are allowed inside c , otherwise the RHS of equation (1.195) must contains the residues calculated on each pole.

integral over time in (1.170) can be turned into

$$\langle \Delta B(\mathbf{x}, t) \rangle = \frac{1}{V} \int_{-\infty}^{+\infty} dt' \int d\mathbf{x}' K(\mathbf{x}', t') \phi_{AB}(\mathbf{x}, \mathbf{x}', t - t') \quad (1.193)$$

allowing for the usage of Fourier's transform.

Response functions are observable and must be real, because of this, the real part ϕ_1 and imaginary part ϕ_2 of their Fourier's transform, must satisfy some parity properties following from the imposition that (1.174) and its complex conjugate must be equal:

$$\phi_1(\mathbf{q}, \omega) = \phi_1(\mathbf{q}, -\omega) \quad -\phi_2(\mathbf{q}, \omega) = \phi_2(\mathbf{q}, -\omega) \quad (1.194)$$

these properties turns out to be useful in determining the relation between the real and the imaginary part of the response function, the so called *Kramers-Kronig's relations*. They come directly from the application of the residues theorem:

$$\frac{1}{i2\pi} \oint_c \frac{\phi(\mathbf{q}, \omega')}{\omega' - \omega} d\omega' = \sum_{m=1}^j Res[p_j] \quad (1.195)$$

on the integration contours c of figure (1.15) containing m poles p_j . If the response function is limited over the whole integration half plane, the RHS of equation (1.195) vanish and the integral over the semi circle Γ vanishes too by virtue of Jordan's lemma: one gains a real integral. One can infer something about the response function properties in the complex frequency plane by looking at its Fourier's transform (1.174) generalized to complex frequency $\omega = \omega_1 + i\omega_2$ (dropping the spatial dependence):

$$\phi(\omega_1 + i\omega_2) = \int_{-\infty}^{+\infty} \phi(t - t') e^{i\omega_1(t-t')} e^{-\omega_2(t-t')} d(t - t') \quad (1.196)$$

independently from the sign of $(t - t')$, ω_1 and ω_2 , the imaginary exponential is always limited but the real exponential can be both negative or positive, i.e. limited or unlimited. So the only way for $\phi_{AB}(\mathbf{q}, \omega)$ to be limited is for positive ω_2 if $t > t'$ (upper half plane) or for negative ω_2 if $t' > t$ (lower half plane). The latter possibility is forbidden by the causality

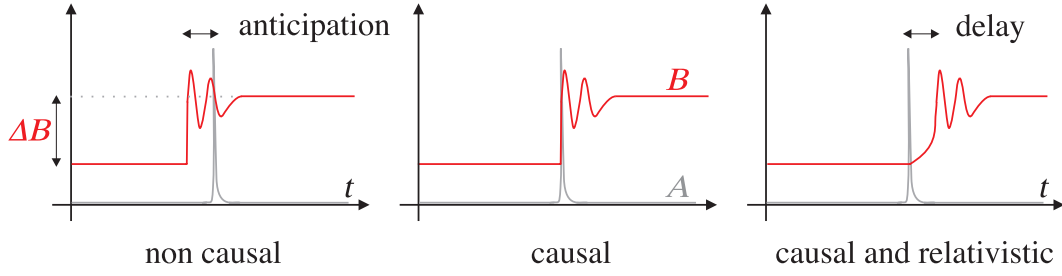


Figure 1.16: Different kinds of response as a function of the time. Gray curve represent the perturbation observable A , red curve represent the response observable B .

principle so the only allowed integration path c for the contour integral to become real is the upper half plane. The integral over Γ vanish, the one over γ gives half the residue calculated at ω and finally:

$$\phi(\mathbf{q}, \omega) = \frac{1}{i\pi} \wp \int_{-\infty}^{+\infty} \frac{\phi(\mathbf{q}, \omega')}{\omega' - \omega} d\omega' \quad (1.197)$$

here \wp stands for the *principal value* of the integral. Changing the sign convention for the Fourier's transforms one must use the opposite half plane but the result (1.197) remains unchanged. Now the response function can be split into its real and imaginary components leading to the Kramers-Kronig's relations:

$$\begin{aligned} \phi_1(\mathbf{q}, \omega) + i\phi_2(\mathbf{q}, \omega) &= \frac{1}{i\pi} \wp \int_{-\infty}^{+\infty} \frac{\phi_1(\mathbf{q}, \omega') + i\phi_2(\mathbf{q}, \omega')}{\omega' - \omega} d\omega' \\ \phi_1(\mathbf{q}, \omega) &= \frac{1}{\pi} \wp \int_{-\infty}^{+\infty} \frac{\phi_2(\mathbf{q}, \omega')}{\omega' - \omega} d\omega' \quad \phi_2(\mathbf{q}, \omega) = -\frac{1}{\pi} \wp \int_{-\infty}^{+\infty} \frac{\phi_1(\mathbf{q}, \omega')}{\omega' - \omega} d\omega' \end{aligned} \quad (1.198)$$

or, multiplying end dividing by $\omega' + \omega$ and making use of properties (1.194)

$$\phi_1(\mathbf{q}, \omega) = \frac{2}{\pi} \wp \int_0^{+\infty} \frac{\phi_2(\mathbf{q}, \omega')\omega'}{\omega'^2 - \omega^2} d\omega' \quad \phi_2(\mathbf{q}, \omega) = -\frac{2}{\pi} \wp \int_0^{+\infty} \frac{\phi_1(\mathbf{q}, \omega')\omega}{\omega'^2 - \omega^2} d\omega' \quad (1.199)$$

another useful relation is the *London's transform*, a rotation of the response function over the imaginary frequency axe. It comes from the Kramers-Kronig's relation for the real part of ϕ by simply perform the substitution $\omega \rightarrow i\omega$:

$$\phi(\mathbf{q}, i\omega) = \frac{2}{\pi} \int_0^{+\infty} \frac{\phi_2(\mathbf{q}, \omega')\omega'}{\omega'^2 + \omega^2} d\omega' \quad (1.200)$$

the advantage of this expression is that it is always real and, containing no poles on the real frequency axe, is a very smooth function easy to integrate. As it will be show in section 1.9 there are many connection between $\phi(i\omega)$ and $\phi_2(\omega)$.

The Kramers-Kronig's relations are causal but not relativistic, they allow the response to arise exactly in at the same instant in which the perturbation is switched on. If one takes into account the finite propagation velocity of the interaction, a delay occur between a change in the perturbation and the consequent change in the response, see figure 1.16. The Kramers-Kronig's relations have been generalized to the relativistic framework by Leontovich [1].

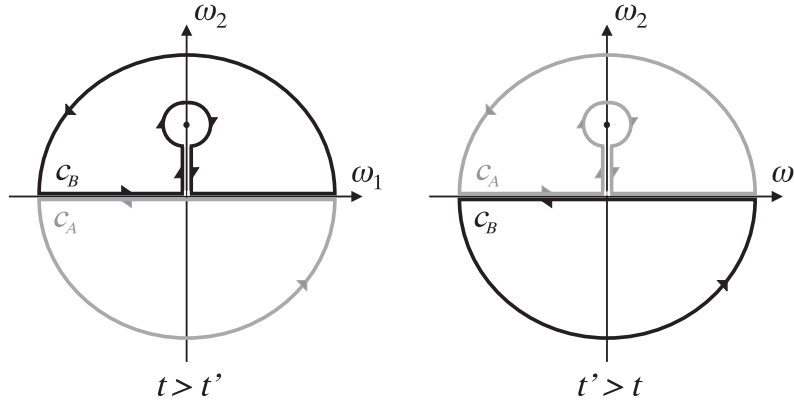


Figure 1.17: Possible contours for the solution of integral (1.201). The gray one gives a stable but non causal response, the black one gives a causal but non stable response. To get both a stable and causal behaviour the singularity at $\omega = i\beta$ must vanish.

The Kramers-Kronig's relations does not come by the causality imposition only, there is another implicit requirement, the *stability* of the medium. A medium is stable if, given any limited arbitrary perturbation, the response remains, on its part, limited, i.e. if the response function does not have poles in the upper half plane. Stability and causality separately, does not give a physical relation between perturbation and response. To show this, suppose that ϕ has a pole in the upper half plane at $\omega = i\beta$, one can roughly replace $\phi(\omega)$ with $A/(\omega - i\beta)$ were A is a generic constant, given an arbitrarily small perturbation of frequency $\omega = \beta$ the real part of $\phi(\omega)$ diverges. Replacing into the inverse Fourier's transform

$$\begin{aligned}\phi(t-t') &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(\omega_1 + i\omega_2) e^{(-i\omega_1 + \omega_2)(t-t')} d(\omega_1 + i\omega_2) \\ \phi(t-t') &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{A}{\omega_1 + i\omega_2 - i\beta} e^{(-i\omega_1 + \omega_2)(t-t')} d(\omega_1 + i\omega_2)\end{aligned}\quad (1.201)$$

it easy to see that only two contours exist to avoid the singularity, they are depicted in figure 1.17. Depending on the sign of the difference $t - t'$, the request for the real exponential to be limited, makes the two contours alternatively allowed. Using the Heaviside's step function integral definition:

$$\lim_{\tau \rightarrow 0} \theta(t-t') e^{\tau(t-t')} = \frac{1}{i2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega(t-t')}}{\omega - i\tau} d\omega \quad (1.202)$$

two results are possible for $\phi(t-t')$, integration over c_A gives a stable but non causal behaviour $\phi(t-t') \propto A\theta(t'-t)e^{\beta t-t'}$, while integration over c_B gives a causal but instable behaviour $\phi(t-t') \propto -A\theta(t-t')e^{\beta t-t'}$. In the latter case the step function is non vanishing for $t' > t$, giving rise to a non causal behaviour, the real exponent is negative and the response function, together with the response, remains limited. In the first case the step function is non vanishing for $t > t'$, now the response function has a causal behaviour, but the real exponential remains positive, i.e. the response function and the response can diverge. To conclude Kramers-Kronig's relations come from the joint imposition of causality and stability. Now it is useful to write the general definitions (1.199) and (1.200) for some specific response

functions, starting from the conductivity:

$$\begin{aligned}\sigma_1(\mathbf{q}, \omega) &= \frac{2}{\pi} \wp \int_0^{\infty} \frac{\sigma_2(\mathbf{q}, \omega') \omega'}{\omega'^2 - \omega^2} d\omega' \\ \sigma_2(\mathbf{q}, \omega) &= -\frac{2\omega}{\pi} \wp \int_0^{\infty} \frac{\sigma_1(\mathbf{q}, \omega')}{\omega'^2 - \omega^2} d\omega'\end{aligned}\tag{1.203}$$

now, using definition (1.30) into (1.198) one can switch to the susceptibility or to the inverse dielectric tensor⁶:

$$\begin{aligned}\epsilon_2^{-1}(\mathbf{q}, \omega) &= -\frac{2\omega}{\pi} \wp \int_0^{\infty} \frac{\epsilon_1^{-1}(\mathbf{q}, \omega') - 1}{\omega'^2 - \omega^2} d\omega' \\ \epsilon_1^{-1}(\mathbf{q}, \omega) &= 1 + \frac{2}{\pi} \wp \int_0^{\infty} \frac{\epsilon_2^{-1}(\mathbf{q}, \omega') \omega'}{\omega'^2 - \omega^2} d\omega'\end{aligned}\tag{1.204}$$

and to the corresponding London's transform:

$$\epsilon^{-1}(\mathbf{q}, i\omega) = 1 + \frac{2}{\pi} \int_0^{\infty} \frac{\epsilon_2^{-1}(\mathbf{q}, \omega') \omega'}{\omega'^2 + \omega^2} d\omega'\tag{1.205}$$

One can also write Kramers-Kronig's relations for the direct dielectric tensor:

$$\begin{aligned}\bar{\epsilon}_2(\mathbf{q}, \omega) &= -\frac{2\omega}{\pi} \wp \int_0^{\infty} \frac{\bar{\epsilon}_1(\mathbf{q}, \omega') - 1}{\omega'^2 - \omega^2} d\omega' \\ \bar{\epsilon}_1(\mathbf{q}, \omega) &= 1 + \frac{2}{\pi} \wp \int_0^{\infty} \frac{\bar{\epsilon}_2(\mathbf{q}, \omega') \omega'}{\omega'^2 - \omega^2} d\omega'\end{aligned}\tag{1.206}$$

with the corresponding London's transform:

$$\bar{\epsilon}(\mathbf{q}, i\omega) = 1 + \frac{2}{\pi} \int_0^{\infty} \frac{\bar{\epsilon}_2(\mathbf{q}, \omega') \omega'}{\omega'^2 + \omega^2} d\omega'\tag{1.207}$$

It is easy to show that the average energy dissipated by the perturbed medium is proportional to $\bar{\epsilon}_2$, analogously the energy lost by fast electrons traveling in a medium is proportional $-\epsilon_2^{-1}$ [1]. Both these energies are positive defined, it follows immediately that:

$$\bar{\epsilon}_2(\mathbf{q}, \omega) \geq 0 \quad \epsilon_2^{-1}(\mathbf{q}, \omega) \leq 0\tag{1.208}$$

and replacing into Kramers-Kronig's relations (1.206) (1.207), in the limit $\omega \rightarrow 0$ one gets:

$$\bar{\epsilon}_1(\mathbf{q}, 0) \geq 1 \quad \bar{\epsilon}(\mathbf{q}, i0) \geq 1\tag{1.209}$$

whereas, from (1.204) and (1.205) one gets:

$$\epsilon_1^{-1}(\mathbf{q}, 0) \leq 1 \quad \epsilon^{-1}(\mathbf{q}, i0) \leq 1\tag{1.210}$$

⁶Notice that ϵ_1^{-1} stands for $Re[\epsilon^{-1}] = \frac{\epsilon_1}{\epsilon_1^2 + \epsilon_2^2}$ that is different from the inverse of the real part of the dielectric tensor $(\epsilon_1)^{-1} = \frac{1}{Re[\epsilon]} = \frac{1}{\epsilon_1}$.

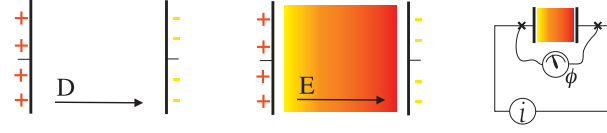


Figure 1.18: Set-up for the measurement of $\epsilon^{-1}(\mathbf{q}, \omega)$.

or equivalently:

$$\begin{aligned} \bar{\epsilon}_1(\mathbf{q}, 0) &\geq 1 & \bar{\epsilon}_1(\mathbf{q}, 0) &< 0 \\ \epsilon^{-1}(\mathbf{q}, i0) &\geq 1 & \epsilon^{-1}(\mathbf{q}, i0) &< 0 \end{aligned} \quad (1.211)$$

Not all the previously introduced response functions must obey the Kramers-Kronig's relations. While $\epsilon^{-1}(\mathbf{q}, \omega)$ can be measured experimentally for every \mathbf{q} and ω , $\bar{\epsilon}(\mathbf{q}, \omega)$ can be measured only for $\mathbf{q} \rightarrow 0$. This is due to the impossibility to manage the Fourier components $\mathbf{q} \neq 0$ of the total field \mathbf{E} when it is used as a perturbation, a detailed discussion is given in appendix D. If $\epsilon^{-1}(\mathbf{q}, \omega)$ can be measured for any \mathbf{q} and ω it must be causal for any \mathbf{q} and ω , i.e. it is always constrained to follow Kramers-Kronig's relations. In the case of $\bar{\epsilon}(\mathbf{q}, \omega)$, Kramers-Kronig's relations must be imposed only for $\mathbf{q} = 0$, thus rules (1.211) must be retained only for $\mathbf{q} = 0$:

$$\begin{aligned} \bar{\epsilon}_1(0, 0) &\geq 1 & \bar{\epsilon}_1(0, 0) &< 0 \\ \bar{\epsilon}(0, i0) &\geq 1 & \bar{\epsilon}(0, i0) &< 0 \end{aligned} \quad (1.212)$$

A deeper study about the allowed sign of the static dielectric function has been performed by Dolgov et al. discussing both general aspects and specific examples for homogeneous media, anisotropic crystals and electrons plasma [19].

Appendix D: Difficulties in the dielectric function measurement

The aim of this appendix is to show why the inverse dielectric function is always measurable whereas the direct dielectric function can be measured only when the perturbing wavevector \mathbf{q} goes to zero. Every perturbation of a given wavevector \mathbf{q} can be considered to be constant in space on a length scale $\ell \ll 1/\mathbf{q}$, in a macroscopic device, where ℓ is large, the perturbation is constant in space only if $\mathbf{q} \rightarrow 0$. Consider the capacitor in figure 1.18, if a certain charge is placed on its plates, it gives rise, on its interior, to an external field \mathbf{D} . A medium placed inside the capacitor feels the external field \mathbf{D} and polarizes giving rise to the polarization field \mathbf{P} . The total field inside the capacitor changes and, because of this, also the potential measurable between the capacitor plates changes. Changing the charge placed on the plates by means of a current generator, i.e. changing the external field, and measuring the potential between them, i.e. the total field, it is possible to measure the medium response function $\epsilon^{-1}(\mathbf{q}, \omega)$ for $\mathbf{q} \simeq 0$. Now, if the potential on the capacitor plates is kept fixed by means of a battery, when the medium is inserted into the capacitor, the total field changes but new charge is added on the plates in order to restore its old value. This means that one can also keep control on the total field \mathbf{E} measuring the external field \mathbf{D} through the current. In this way one can measure $\bar{\epsilon}(0, \omega)$, see figure 1.19. One can think to extend this method for any arbitrary $\mathbf{q} \neq 0$ by simply reducing the capacitor dimension in order to always fulfill the $\ell \ll 1/\mathbf{q}$ condition. In doing this the medium dimension becomes rapidly microscopic: defects, surfaces and microscopic structure come rapidly into play modifying the medium properties. In order to do not alter the medium properties the measurement device can be embedded into the medium, as shown in figure 1.20. Placing some charge on the conducting sphere one is again able to manage every single component \mathbf{q} of the Fourier's series of the external field \mathbf{D} , measuring the potential acting on the conducting sphere surface, one can

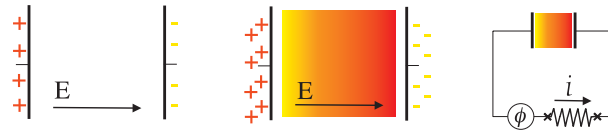
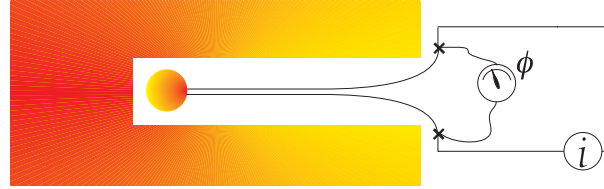
Figure 1.19: Set-up for the measurement of $\epsilon(0, \omega)$.

Figure 1.20: Measurement device embedded in the medium.

measure the value of \mathbf{E} for any arbitrary \mathbf{q} . These are different in the case of the total field. Keeping fixed the potential acting on the sphere, it is possible to know the total field only on the sphere surface, moving far away from the sphere the total field changes according to the medium properties that one needs to measure: there is no control on the Fourier components of the total field. To regain control on the total field \mathbf{E} inside the medium one should keep it constant inserting in the medium a number of conducting spheres like the one of figure 1.20, the larger their number the smaller \mathbf{q} , this solution is of course useless because the insertion of an infinite number of conductors in the sample alters its natural properties. A detailed survey on most common experimental techniques can be found in reference [20].

Bibliography

- [1] L.V. Keldysh, D.A. Kirzhnits, and A.A. Maradudin editors. Modern problems in condensed matter sciences. *The dielectric function of condensed systems*, volume 24. North-Holland Physics, 1989.
- [2] F. Wooten. *Optical properties of solids*. Academic Press, Davis, California, 1988.
- [3] J.D. Jackson. *Classical Electrodynamics*. John Wiley and Sons, 1975.
- [4] S. Lundqvist, N.H. March editors. Physics of solids, and liquids. *Theory of the inhomogeneous electron gas*. Plenum press, London, 1983.
- [5] G. Grosso and G. Pastori Parravicini. *Solid State Physics*. Academic press, London, 2003.
- [6] G. Onida, L. Reining, and A. Rubio. *Rev.Mod.Phys.*, 74:601, 2002.
- [7] J. Lindhard. *Kgl.Danske videnskab selskab mat.-fys.medd.*, 28:8, 1954.
- [8] N.D. Mermin. *Phys. Rev. B*, 1:1019, 1970.
- [9] A.L. Fetter and J.D.Walecka. *Quantum theory of many-particle system*. Dover, Mineola, New York, 2003.
- [10] L. Hedin. *Phys.Rev.*, 139:A709, 1965.
- [11] W. Kohn. *Rev.Mod.Phys.*, 71:1253, 1999.
- [12] H. Ehrenreich and M.H. Cohen. *Phys.Rev.*, 115:786, 1959.
- [13] S.L. Adler. *Phys.Rev.*, 126:413, 1962.
- [14] D.M. Wood and N.W. Ashcroft. *Phys.Rev.B*, 25:6255, 1982.
- [15] F. Sottile Ph.D. thesis. *Response functions of semiconductors and insulators: from the Bethe-Salpeter equation to time-dependent density functional theory*. CNRS-CEA/DSM, Ecole polytechnique, Palaiseau, France., 2003.
- [16] N. Wiser. *Phys.Rev.*, 129:62, 1963.
- [17] R. Kubo. *J. Phys. Soc. Japan*, 12:570, 1957.
- [18] R. Kubo. *Rep. Prog. Phys.*, 29:255, 1966.
- [19] O.V. Dolgov, D.A. Kirzhnits, and E.G. Maksimov. *Rev.Mod.Phys.*, 53:81, 1981.
- [20] E.D. Palik editor. *Handbook of optical constants of solids*. Academic press, London, 1985.

2

Macroscopic dispersion forces

Contents

2.1	The force from nothing	67
2.2	The Casimir's model	71
2.3	The Lifshitz's theory	75
2.3.1	Interface modes	75
2.3.2	Cavity modes	78
2.3.3	Finite temperature formalism	85
2.3.4	Extension to more than three layers	87
2.3.5	Anisotropic media	88
2.4	Different approximations	89
2.5	Finite temperature issues	91
2.6	Concluding notes	94
	Appendix A: Interface fields conditions	95
	Appendix B: Perfectly reflecting 3D cavity	97

This chapter is the most important among the theoretical ones, it deals with the theory used through the second part of the thesis. A brief historical overview will be given in the first part of the chapter, followed by very general considerations on the nature of the dispersion interactions. Then the earlier Casimir's model will be discussed as an introduction to the more complicated and general Lifshitz's theory, which will be widely illustrated and different approximations and generalizations will be discussed. The chapter concludes with a brief overview of the finite temperature problem. Further details and generalizations can be found in the review works [1, 2, 3, 4] and in the Milonni book [5].

2.1 The force from nothing

The Casimir force is one of the few macroscopic manifestations of the quantum nature of matter, other examples are the superconductivity, the superfluidity and the quantum Hall effect. This force acts between neutral polarizable macroscopic objects, it depends on their shapes, on their distance, on their dielectric properties, on their thermodynamic properties

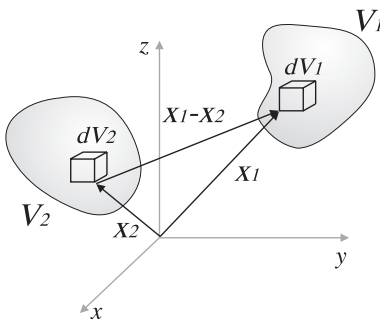


Figure 2.1: Reference framework for the Hamaker additivity assumption expressed by equation (2.4).

and on the effective interaction surface. The latter property makes this force a *surface force*, like friction. The surface forces become more relevant on decreasing the length scale of the interacting systems or reducing their dimensionality. This is the reason why, in the miniaturization era, such a force, known from the 40's, has risen a renewed interest in the condensed matter physics community. For instance, in micro and nano mechanical systems, the small length scale makes the *bulk forces* comparable with the surface ones, determining rather different mechanical behaviours with respect to the macroscopic scale. Phenomena like friction, adhesion and stiction play a crucial role at miniaturization scale.

To understand the behaviour of these macroscopic forces it is better to start from their microscopic origin. The force acting between neutral polarizable atoms or molecules have been introduced by J.D. van der Waals to account for the thermodynamics of real gases and liquids. In 1930 these forces have been properly justified within the framework of quantum mechanics by F. London [6, 7, 8]. He used a time independent perturbation theory to account for the interaction between two neutral atoms reaching the famous result:

$$F(d) = -\frac{18\hbar}{\pi d^7} \int_0^{\infty} \alpha_1(\omega)\alpha_2(\omega)d\omega \quad (2.1)$$

where d is the interatomic distance, $\alpha_1(\omega)$ and $\alpha_2(\omega)$ are the *polarizabilities* of the two interacting atoms¹. Notice that this force is always attractive (negative) and its dependence upon the atoms properties is entirely contained in the polarizability. It can be regarded as the average force between the fluctuating spontaneous dipole, that arises in one atom because of the electron motion, and the dipole induced on the other atom. In 1940s T. Overbeek and E. Verwey, studying the molecular forces in colloid solutions at Philips Research Labs, found a strange behaviour of the van der Waals force: the strength of the attractive force between colloidal particles decay more rapidly with respect to the prediction of equation (2.1). As correctly suspected by Overbeek, the nanometric length scale of colloids, much larger than the atomic length scale, revealed the need for a relativistic extension of the theory. This generalization of the van der Waals theory has been provided by H.B.G. Casimir and D. Polder in 1948 [10], taking into account the finite velocity of the electromagnetic interaction. This velocity comes into play when the distance among interacting atoms becomes comparable with the characteristic emission wave length of the atoms $\lambda_0 = c/\omega_0$, where ω_0 is the

¹The polarizability α of an atom can be defined as the susceptibility χ of a medium using the analogous of (1.106) in which the macroscopic response \mathbf{P} coincides with the single microscopic atomic dipole \mathbf{p} :

$$\mathbf{P} = \mathbf{p} \equiv \alpha \mathbf{D}$$

for a deep quantum treatment of the polarizability of atoms and molecules and their dispersion interactions see reference [9].

principal emission frequency of an atom or a molecule. In the limiting case of $d \gg \lambda_0$ one gets:

$$F(d) = -\frac{161\hbar c}{4\pi d^8} \alpha_1(0) \alpha_2(0) \quad (2.2)$$

whereas for a generic d value the force goes as a linear combination of the two inverse powers of the distance:

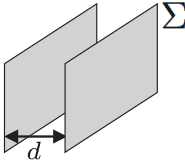
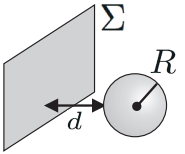
$$F(d) = -\frac{C_{vdW}}{d^7} - \frac{C_{CP}}{d^8} \quad (2.3)$$

where C_{vdW} and C_{CP} are the van der Waals and Casimir-Polder coefficients. They are positive and contain the dielectric response of the atoms. In the opposite limit of $d \ll \lambda_0$, $C_{CP} \rightarrow 0$ and one recovers the old London expression (2.1).

Once that the microscopic interaction between two neutral atoms has been understood, one may think that it is sufficient to sum up over all atom pairs of neutral bodies to obtain the total force acting between them. In 1936 J. De Boer [11] and H. Hamaker [12], assuming the simple additivity of the microscopic forces, performed the calculation of the force acting between macroscopic neutral bodies of different shape. The idea is very simple: with reference to figure 2.1, if the densities $\rho_1(\mathbf{x}_1)$ and $\rho_2(\mathbf{x}_2)$ of the two interacting bodies are known, the total macroscopic force acting between them is:

$$F = \int_{V_1} \int_{V_2} F(|\mathbf{x}_1 - \mathbf{x}_2|) \rho_1(\mathbf{x}_1) \rho_2(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \quad (2.4)$$

here $F(|\mathbf{x}_1 - \mathbf{x}_2|)$ is the force (2.3) acting between two atoms, and V_1 and V_2 are the body volumes. Hamaker calculations showed a first important point: the macroscopic force depends upon the shape and the geometry of the interacting bodies. He found different distance power laws of the force for the plane-plane and the sphere-plane configurations. Moreover he showed that the force strength depends upon the surface of the interacting planes or upon the radius of the interacting spheres:

Geometry	Non Relativistic	Relativistic
	$\frac{H_{nr}}{6\pi d^3} \Sigma$	$\frac{H_r}{d^4} \Sigma$
	$\frac{H_{nr}}{6d^2} R$	$\frac{2\pi H_r}{3d^3} R$

where H_{nr} and H_r are the *Hamaker constants* obtained summing over the microscopic van der Waals interaction or over the microscopic Casimir-Polder interaction. Unfortunately the De Boer and Hamaker results turned out to be useful only for rarefied gases, where the many body effects can be neglected. For condensed matter bodies the strong role played by many body effects makes the additivity assumption inadequate. For more than a decade, the lack of a theory able to treat properly the many body effects, left the things unchanged. Only in few special cases the force between macroscopic neutral bodies has been correctly worked out. Lennard-Jones in 1932 and Bardeen in 1940 described the van der Waals interaction of a neutral atom or molecule with a metal surface using the method of images [13, 14]. Casimir and Polder generalized this result including the relativistic effects, and Casimir alone, in



Figure 2.2: From left to right, first row: van der Waals, London, Overbeek, Casimir; second row: Polder, Lennard-Jones, Bardeen, Lifshitz.

1948 proposed a new way to look at the problem [15, 16]: the macroscopic dispersion force between two perfectly reflecting neutral planes can be seen as the change in the zero point energy of the quantized electromagnetic field trapped between them and responsible for their interaction. This problem will be analyzed in section 2.2. Only after the formulation of the fluctuation dissipation theorem and the modern theory of perturbation and response, E.M. Lifshitz succeeded in the generalization of the Casimir theory, this will be the object of section 2.3. The only way to include many body interaction in condensed matter macroscopic effects seems to be through the macroscopic response functions, as described in section 1.10, so it has been natural to continue to treat the media as a continuous distribution of charges quantizing the electromagnetic fields as originally suggested by Casimir.

In the very first microscopic approach by London the charges were treated quantum-mechanically whereas the fields among them were considered classical continuous fields, the same assumption is present in the Lennard-Jones and Bardeen theories. Casimir and Lifshitz followed a complementary approach in which the bodies were treated as continuous dielectric materials and the electromagnetic fields, describing their interaction, are quantized. In both cases an expression of the force can be carried out, that depends on the shape and geometry of the interacting bodies in which the medium properties enter the problem through the response functions. Response functions rule the dispersion of media, because of this, all the van der Waals, Hamaker, Casimir-Polder or Casimir forces, whether they are macroscopic or not, are usually called *dispersion forces*. Another frequently used name is *fluctuations forces*. In the approach of continuous fields the name refers to the fluctuations of the spontaneously arising dipoles on neutral atoms of the media, while in the approach of quantized fields the name refers to the fluctuations of the vacuum energy of the fields, due to a change in the boundary conditions.

The equivalence of the two approaches has been demonstrated within the modern quantum electrodynamics theory (QED), in which both charges and fields must be quantized [17, 18, 19]. In the QED theory, the energy normalization problem has been solved by G.C. Wick that introduced the *normal ordering operator*: since an infinite vacuum energy exist but it is not directly observable, starting to measure all the other energies from the vacuum energy one is always able to get finite values. In presence of boundary conditions the standard Wick's theorem becomes useless, an infinite number of different vacuum states exists, corresponding to different values of the geometric parameters entering the energy due to the boundary conditions. In such a situation it is incorrect to pre-assign the role of reference state to a specific vacuum state, but the finite difference between two vacuum states become observable giving

rise to the *Casimir effect* of which the Casimir force is a macroscopic observable consequence.

2.2 The Casimir's model

Following the Casimir approach, the macroscopic dispersion force between two parallel perfectly reflecting metal plates, can be reconduced to the dependence of the electromagnetic zero point energy on the distance between the two plates. Changing the plates distance causes a change in the electromagnetic zero point energy which manifests itself as a macroscopic force by virtue of the fluctuation-dissipation theorem. To understand how the electromagnetic zero point energy can depend upon a geometric parameter, imagine to have a massless scalar field confined into the onedimensional cavity of figure 2.3 (a) by means of certain boundary conditions, its energy levels result to be quantized and, from the very basic quantum theory, one gets:

$$E(\ell) = \sum_{\mathbf{k}} \hbar \omega(\mathbf{k}, \ell) \left(\frac{1}{2} + n_{\mathbf{k}} \right) \quad (2.5)$$

where ℓ is the cavity amplitude. If none of the electromagnetic modes is excited, i.e. there are no photons in the cavity $n_{\mathbf{k}} = 0 \forall \mathbf{k}$, only the zero point energy remains:

$$E(\ell) = \frac{\hbar}{2} \sum_{\mathbf{k}} \omega(\mathbf{k}, \ell) \quad (2.6)$$

To get the free space zero point energy one must let $\ell \rightarrow \infty$. Notice that the energy of each mode must depend upon the cavity geometry and that the summation over the modes is divergent. Now, if another reflecting plate is inserted into the cavity, its presence modifies the boundary conditions that have to be satisfied by the electromagnetic modes, see figure 2.3 (b). As a consequence, the zero point energy is modified. This change in energy can be calculated as:

$$\Delta E(\ell, a) = \frac{\hbar}{2} \sum_{\mathbf{k}} \omega(\mathbf{k}, \ell) - \left[\frac{\hbar}{2} \sum_{\mathbf{k}} \omega(\mathbf{k}, \ell - a) + \frac{\hbar}{2} \sum_{\mathbf{k}} \omega(\mathbf{k}, a) \right] \quad (2.7)$$

where $\omega(\mathbf{k}, \ell, 0)$ represents the empty cavity modes whereas $\omega(\mathbf{k}, \ell - a)$ and $\omega(\mathbf{k}, a)$ represent the modes of the cavity in presence of the reflecting plate. The previous difference is not well defined because it is the difference between two divergent quantities, one can introduce a *regularization function* $\mathbf{R}[\omega(\mathbf{k}), \lambda]$ in such a way that the quantity:

$$\lim_{\lambda \rightarrow 0} \sum_{\mathbf{k}} \omega(\mathbf{k}) \mathbf{R}[\omega(\mathbf{k}), \lambda] \quad (2.8)$$

remains finite. In this view equation (2.7) becomes

$$\Delta E(a) = \lim_{\ell \rightarrow \infty} \lim_{\lambda \rightarrow 0} \frac{\hbar}{2} \sum_{\mathbf{k}} \left[\omega(\mathbf{k}, \ell) - \omega(\mathbf{k}, \ell - a) - \omega(\mathbf{k}, a) \right] \mathbf{R}[\omega(\mathbf{k}), \lambda] \quad (2.9)$$

Clearly this regularization must be justified from a physical point of view: looking at the dielectric response functions it is known that, for very large frequencies, all materials become transparent ($\epsilon(\omega) \rightarrow 1$). So one can imagine that a cut-off frequency ω_c exist below which $\mathbf{R}[\omega(\mathbf{k}), \lambda]$ is finite whereas, for $\omega(\mathbf{k}) \gg \omega_c$, $\mathbf{R}[\omega(\mathbf{k}), \lambda] \rightarrow 0$, in other words: for large frequency modes the cavity and the plate are almost transparent so that large frequency modes do not contribute to the cavity energy.

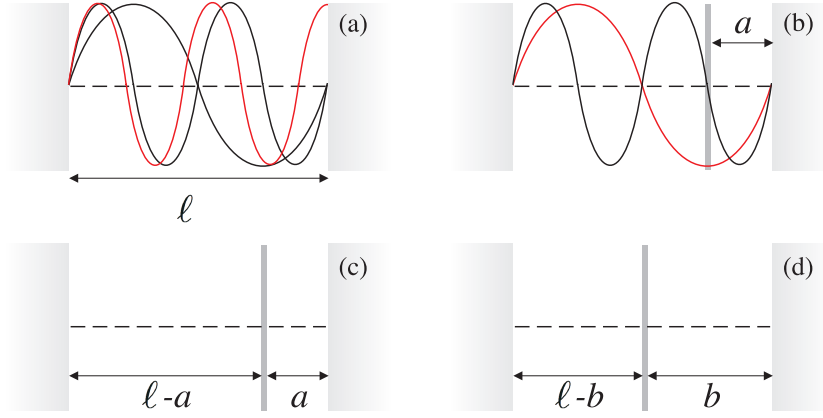


Figure 2.3: Casimir monodimensional cavity. (a) represents a cavity of amplitude ℓ , the red mode is not allowed by the boundary conditions. (b) represents the insertion of a perfectly reflecting plate, the red mode, previously allowed becomes forbidden. (c) and (d) represents a variation of the plate position, responsible for a change in the Casimir force between the plate and the cavity walls.

Now one can think that the Casimir energy of the cavity in figure 2.3 (c) can be calculated as:

$$\begin{aligned}
 E(a) &= \lim_{\ell \rightarrow \infty} \lim_{\lambda \rightarrow 0} \frac{\hbar}{2} \sum_{\mathbf{k}} \left\{ \omega(\mathbf{k}, \ell - a) + \omega(\mathbf{k}, a) \right\} \mathcal{R}[\omega(\mathbf{k}), \lambda] = \\
 &= \lim_{\ell \rightarrow \infty} \lim_{\lambda \rightarrow 0} \frac{\hbar c}{2} \sum_{n=0}^{\infty} \left(\frac{n\pi}{\ell - a} + \frac{n\pi}{a} \right) \mathcal{R}[\omega(\mathbf{k}), \lambda]
 \end{aligned} \tag{2.10}$$

where the quantization of the electromagnetic modes gives the dispersion relation $\omega(\mathbf{k}) = c|\mathbf{k}| = c\frac{n\pi}{a}$. Choosing the following regularization function for $\omega(\mathbf{k}, a)$:

$$\mathcal{R}[\omega(\mathbf{k}), \lambda] = e^{-\lambda \frac{n\pi}{a}} \tag{2.11}$$

one gets:

$$E(a) = \lim_{\ell \rightarrow \infty} \lim_{\lambda \rightarrow 0} \frac{\hbar c}{2} \sum_{n=0}^{\infty} \left(\frac{n\pi}{\ell - a} e^{-\lambda \frac{n\pi}{\ell - a}} + \frac{n\pi}{a} e^{-\lambda \frac{n\pi}{a}} \right) \tag{2.12}$$

using the property

$$\sum_{n=0}^{\infty} \frac{n\pi}{a} e^{-\lambda \frac{n\pi}{a}} = -\frac{\partial}{\partial \lambda} \left(\frac{1}{1 - e^{-\lambda \frac{\pi}{a}}} \right) = \frac{\pi}{a} \frac{1}{(1 - e^{-\lambda \frac{\pi}{a}})^2} e^{-\lambda \frac{\pi}{a}} \tag{2.13}$$

in the limit of small λ a Taylor's expansion as been performed:

$$\frac{\pi}{a} \frac{1}{(1 - e^{-\lambda \frac{\pi}{a}})^2} e^{-\lambda \frac{\pi}{a}} \simeq \frac{a}{\pi \lambda^2} - \frac{\pi}{12a} \tag{2.14}$$

replacing into (2.12) the final result is achieved:

$$E(a) = \lim_{\ell \rightarrow \infty} \lim_{\lambda \rightarrow 0} \left(\frac{\ell}{\pi \lambda^2} - \frac{\pi}{12a(\ell - a)} \right) \tag{2.15}$$

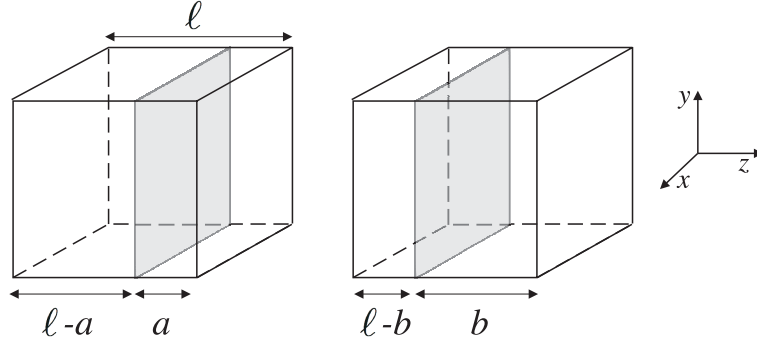


Figure 2.4: Casimir cavity in three dimensions.

This is a divergent result, the regularization function alone is not enough, also a *normalization* is needed. The Casimir energy of the cavity in figure 2.3 (c) can be calculated with respect to a reference initial configuration configuration, for example the one of figure 2.3 (d):

$$E(a) = \lim_{\ell \rightarrow \infty} \lim_{\lambda \rightarrow 0} \frac{\hbar}{2} \sum_{\mathbf{k}} \left\{ \omega(\mathbf{k}, \ell - a) + \omega(\mathbf{k}, a) - \omega(\mathbf{k}, \ell - b) + \omega(\mathbf{k}, b) \right\} \mathcal{R}[\omega(\mathbf{k}), \lambda] \quad (2.16)$$

choosing for instance $b = \ell/2$ one gets:

$$E(a) = -\frac{\pi}{12a} \quad (2.17)$$

and differentiating with respect to a one can calculate the Casimir force:

$$F(a) = -\frac{\pi}{12a^2} \quad (2.18)$$

It is possible to show that, for the specific configuration of plane cavity, the previous results are independent of the choice of b and of the regularization function. For other kinds of geometries the regularization and normalization procedures are not unique, leading to different results, a correct definition of the Casimir force and energy is not possible [4].

The result presented by Casimir in his original work [15] is nothing but the extension of the previous case to a three dimensional cavity formed by parallel perfectly reflecting (ideal metal) plates, see figure 2.4. Now the \mathbf{k} vector is in three dimensions:

$$k_x = \frac{\pi}{\ell} n_x \quad k_y = \frac{\pi}{\ell} n_y \quad k_z = \frac{\pi}{\ell} n_z \quad (2.19)$$

so the energy expression, equivalent to (2.6), becomes:

$$E(\ell) = \hbar c \frac{L^2}{\pi^2} \int_0^\infty \int_0^\infty \left[\frac{1}{2} \sqrt{k_x^2 + k_y^2} + \sum_{n=1}^\infty \sqrt{n^2 \frac{\pi^2}{\ell^2} + k_x^2 + k_y^2} \right] dk_x dk_y \quad (2.20)$$

here the cavity dimension L , along x and y directions, is supposed to be large enough to neglect the quantization along that directions. In this way an integration can be performed over k_x and k_y , introducing the L/π factor, and only for k_z the discrete sum must be retained. The $n = 0$ term has been kept separated from the others while the $n \neq 0$ terms have been multiplied by a factor 2: this is due to the fact that, in a perfectly reflecting three dimensional cavity, two equivalent modes exist for each $k_z \neq 0$ whereas only one is allowed if $k_z = 0$ (see appendix B). Using polar coordinates in the $k_x k_y$ plane:

$$E(\ell) = \hbar c \frac{L^2}{2\pi} \sum_{n=1}^\infty \int_0^\infty k dk \sqrt{n^2 \frac{\pi^2}{\ell^2} + k^2} \quad (2.21)$$

where the prime over the sum symbol indicates that the $n = 0$ term must given half weight. This integral is divergent so it must be multiplied by a regularization function $\mathbf{R}[\mathbf{k}/k_c]$ and normalized, here k_c represent a generic cut-off wavevector. To follow the Casimir derivation one must start from the definition (2.16) with $a = \ell$ and $b = 0$:

$$E(a) = \hbar c \frac{L^2}{2\pi} \left\{ \sum_{n=1}^{\infty'} \int_0^{\infty} k dk \sqrt{n^2 \frac{\pi^2}{a^2} + k^2} \mathbf{R} \left[\sqrt{n^2 \frac{\pi^2}{a^2} + k^2} / k_c \right] - \right. \\ \left. - \frac{a}{\pi} \int_0^{\infty} \int_0^{\infty} k dk dk_z \sqrt{k_z^2 + k^2} \mathbf{R} \left[\sqrt{k_z^2 + k^2} / k_c \right] \right\} \quad (2.22)$$

this equation can be simplified through the substitution $u = a^2 k^2 / \pi^2$:

$$\frac{E(a)}{L^2} = \hbar c \frac{\pi^2}{2a^3} \left\{ \sum_{n=1}^{\infty'} \int_0^{\infty} du \sqrt{n^2 + u} \mathbf{R} \left[\frac{\pi}{ak_c} \sqrt{n^2 + u} \right] - \right. \\ \left. - \int_0^{\infty} \int_0^{\infty} dudn \sqrt{n^2 + u} \mathbf{R} \left[\frac{\pi}{ak_c} \sqrt{n^2 + u} \right] \right\} \quad (2.23)$$

Now it is possible to use the Euler-Maclaurin's formula:

$$\sum_{n=1}^{\infty'} F(n) - \int_0^{\infty} F(n) dn = -\frac{1}{12} F'(0) + \frac{1}{720} F'''(0) + \dots \quad (2.24)$$

in this case:

$$F(n) dn = \int_{n^2}^{\infty} \sqrt{w} \mathbf{R} [w\pi/ak_c] dw \quad (2.25)$$

after the change of variable $w = u + n^2$, thus:

$$F'(0) = -2n^2 \mathbf{R} [w\pi/ak_c] = 0 \quad F'''(0) = -4 \quad (2.26)$$

so the total Casimir energy per unit surface is given by:

$$\frac{E(a)}{L^2} = -\frac{\hbar c \pi^2}{720 a^3} \quad (2.27)$$

this formula results to be independent of the regularization function \mathbf{R} . The higher terms of the expansion (2.24) contain powers of π/ak_c so that the Casimir result holds as long as $ak_c \gg 1$. Differentiating with respect to a one get the Casimir Force for unit area:

$$\frac{F(a)}{L^2} = -\frac{\hbar c \pi^2}{240 a^4} \quad (2.28)$$

and, to use the Casimir's words:

One is thus led to the following conclusions. There exist an attractive force between two metal plates which is independent of the material of the plates as long as the distance is so large that for wave lengths comparable with that distance the penetration depth is small compared with the distance. This force may be interpreted as a zero point pressure of the electromagnetic waves.

References [4] and [2] offer a general discussion about different renormalization techniques and the different cavity geometries.

2.3 The Lifshitz's theory

This section is devoted to the Lifshitz's generalization of the previously derived Casimir result. In his original work [20], Lifshitz made use of the Green's function technique, so far other simpler derivations have been proposed. In this section the formulation of N.G. van Kampen will be discussed [21, 22, 23]. Further extensions and generalizations of the Lifshitz's theory have been performed by I.E. Dzyaloshinskii and L.P. Pitaevskii [24, 25, 26, 27]. The first section is devoted to the introduction of the surface modes. These concepts will be employed in the next sections for the derivation of the zero temperature and finite temperature expressions of the dispersion energy and dispersion force. Finally some useful generalizations of the theory will be discussed. In any case only the geometry of parallel interacting plates (slabs) will be discussed, as in the original Lifshitz's work.

2.3.1 Interface modes

The solutions of the Maxwell's equation for a field in an infinite homogeneous, isotropic and linear medium are simple plane waves. In presence of an interface between two media more solutions become possible: some of them are again delocalized modes, i.e. plane waves that satisfy properly the interface conditions described in appendix A, some of them are localized modes, i.e. fields that rapidly vanish moving far away from the interface, they are some time called *interface modes*. Imagine to have a flat surface of a body, the surface itself is the interface between the body medium and the vacuum or a second medium. A good reference system to work with, is the one whose origin is located at the interface and the z axis is perpendicular to it, i.e. the interface coincides with the xy plane. One has to look for solution of the Maxwell's equation which behaves as:

$$\mathbf{E}(\mathbf{x}) = \mathbf{E}_0(z)e^{i\mathbf{k}_{\parallel}\cdot\mathbf{x}_{\parallel}}e^{-i\omega t} \quad \mathbf{H}(\mathbf{x}) = \mathbf{H}_0(z)e^{i\mathbf{k}_{\parallel}\cdot\mathbf{x}_{\parallel}}e^{-i\omega t} \quad (2.29)$$

$\mathbf{x} = (\mathbf{x}_{\parallel}, z)$ is the position vector, $\mathbf{k} = (\mathbf{k}_{\parallel}, k_z)$ is the wave vector of the plane wave that propagates parallel to the surface, whereas $\mathbf{E}_0(z)$ e $\mathbf{H}_0(z)$ describe the behaviour of the field in the direction perpendicular to the surface. Without loss of generality, one can choose the \mathbf{k}_{\parallel} vector along the x direction only:

$$\mathbf{E}(\mathbf{x}) = \mathbf{E}_0(z)e^{ikx} \quad \mathbf{H}(\mathbf{x}) = \mathbf{H}_0(z)e^{ikx} \quad (2.30)$$

to find an explicit form for $\mathbf{E}_0(z)$ and $\mathbf{H}_0(z)$ one must replace the solution (2.30) into the Maxwell's equations (1.76), assuming a neutral, uniform and isotropic medium. With the previous choice for \mathbf{k} , one get eight different equations, some of them are independent and one can define two different systems:

$$\left\{ \begin{array}{l} \frac{\partial E_{0z}}{\partial z} + ikE_{0x} = 0 \\ -ikE_{0z} + \frac{\partial E_{0x}}{\partial z} = \frac{i\omega}{c}H_{0y} \\ \frac{\partial H_{0y}}{\partial z} = \frac{i\epsilon(\omega)\omega}{c}E_{0x} \\ ikH_{0y} = -\frac{i\epsilon(\omega)\omega}{c}E_{0z} \end{array} \right. \quad \left\{ \begin{array}{l} \frac{\partial H_{0z}}{\partial z} + ikH_{0x} = 0 \\ -ikH_{0z} + \frac{\partial H_{0x}}{\partial z} = -\frac{i\epsilon(\omega)\omega}{c}E_{0y} \\ \frac{\partial E_{0y}}{\partial z} = -\frac{i\omega}{c}H_{0x} \\ ikE_{0y} = \frac{i\omega}{c}H_{0z} \end{array} \right. \quad (2.31)$$

they can be separately obtained under the assumption of $\mathbf{B} \perp \mathbf{E} \perp \mathbf{k}$ (transverse modes) respectively for ($H_{0z} = 0$) or for ($E_{0z} = 0$). Because of this their solution are called *transverse magnetic* (TM, first system) and *transverse electric modes* (TE, second system).

Combining the equations of the first of the systems (2.31) one get an equation that rules E_{0z}

behaviour only:

$$\begin{cases} \frac{\partial^2 E_{0z}}{\partial z^2} - k^2 E_{0z} + \frac{\omega^2 \epsilon_1(\omega)}{c^2} E_{0z} = 0 & z > 0 \\ \frac{\partial^2 E_{0z}}{\partial z^2} - k^2 E_{0z} + \frac{\omega^2 \epsilon_2(\omega)}{c^2} E_{0z} = 0 & z < 0 \end{cases} \quad (2.32)$$

1 and 2 label the two different media on the two sides of the interface plane. The solution of the system is given by:

$$E_{0z} = E_j e^{-\gamma_j z} \quad (2.33)$$

where j is the medium label and

$$\gamma_j = \sqrt{k^2 - \epsilon_j(\omega) \frac{\omega^2}{c^2}} \quad (2.34)$$

can be either real or complex giving rise to modes localized at the interface or delocalized modes respectively. Concluding the z component results:

$$E_z = E_j e^{-\gamma_j z + ikx} e^{-i\omega t} \quad (2.35)$$

Using now the other equations of (2.31) the remaining field components, E_x and E_y , can be obtained:

$$\begin{cases} E_x = \frac{i}{k} \frac{\partial E_z}{\partial z} = -\frac{i}{k} \gamma_1 E_1 e^{-\gamma_1 z + ikx} e^{-i\omega t} & z > 0 \\ E_x = \frac{i}{k} \gamma_2 E_2 e^{\gamma_2 z + ikx} e^{-i\omega t} & z < 0 \end{cases} \quad (2.36)$$

and, for the magnetic field:

$$\begin{cases} H_y = \frac{\omega \epsilon(\omega)}{ck} E_z = -\frac{\omega \epsilon_1(\omega)}{ck} E_1 e^{-\gamma_1 z + ikx} e^{-i\omega t} & z > 0 \\ H_y = -\frac{\omega \epsilon_2(\omega)}{ck} E_2 e^{\gamma_2 z + ikx} e^{-i\omega t} & z < 0 \end{cases} \quad (2.37)$$

The dispersion relation of the system modes can be found imposing the continuity condition of the fields as described in appendix A, in particular, equating the fields (2.36) at $z = 0$ one gets

$$\gamma_1 E_1 = -\gamma_2 E_2 \quad (2.38)$$

whereas for the magnetic field (2.37):

$$\epsilon_1(\omega) E_1 = \epsilon_2(\omega) E_2 \quad (2.39)$$

From the ratio between (2.38) and (2.39) one finally achieves the result:

$$k^2 = \frac{\omega^2}{c^2} \left(\frac{\epsilon_1 \epsilon_2}{\epsilon_1 + \epsilon_2} \right) \quad (2.40)$$

showing that the dispersion relation of the modes depends upon the dielectric properties of the media. In the non relativistic regime one has $k \gg \omega/c$, $\gamma_1 = \gamma_2$ and the dispersion relation is simply given by $\epsilon_1 = -\epsilon_2$.

Suppose to have an interface between a medium, whose dielectric properties are described by the Plasma model (see section 1.8), and the vacuum, equation (2.40) becomes:

$$k^2 = \frac{\omega^2}{c^2} \left(\frac{1 - \Omega_p^2/\omega^2}{2 - \Omega_p^2/\omega^2} \right) \quad (2.41)$$

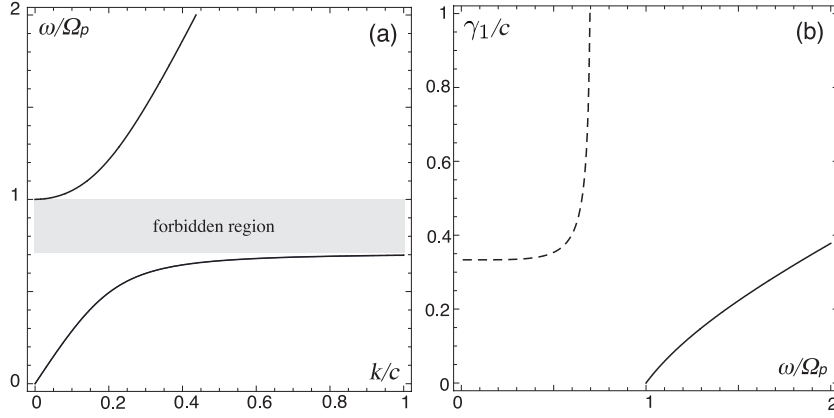


Figure 2.5: (a) represents the $\omega(k)$ relation for an isotropic plasma metal surface. (b) represents $\gamma_1(\omega)$ for an isotropic plasma metal surface, both localized (dashed line) and delocalized (continuous line) modes are shown.

such a function is plotted in figure 2.5 (a). In the frequency region between Ω_p and $\Omega_p/\sqrt{2}$ one has $\epsilon_1\epsilon_2 < 0$ and k should be imaginary, this is of course a forbidden frequency region because solution (2.33) should diverge for large k values. For large k values only one mode exists with frequency $\omega = \Omega_p/\sqrt{2}$, the so called *surface plasmon*. Notice also that in the case of a perfectly reflecting interface ($\epsilon_1 \rightarrow \infty$ and $\epsilon_2 = 0$), equation (2.40) becomes $k = \omega/c$, that is the simple plane waves dispersion relation. This justifies the choice made by Casimir to take into account only transverse modes in his perfectly reflecting cavity described in section 2.2. By Replacing the dispersion relation inside the γ_j definition (2.34) one can find, for the allowed k values, the function plotted in figure 2.5 (b): γ_j can be both real or imaginary, i.e. the real exponential of (2.33) determines an oscillating behaviour even along the z direction or a localization on the surface. Naturally, introducing more complicated models for the dielectric function, the dispersion relation can strongly change.

If the two media are anisotropic, the tensorial nature of the response function comes into play (see section 1.3), and the previous derivation must be generalized. Inserting the general solution (2.30) into the Maxwell's system (1.76) one now get:

$$\begin{cases} \epsilon_{zz} \frac{\partial E_{0z}}{\partial z} + ik\epsilon_{xx}E_{0x} = 0 \\ -ikE_{0z} + \frac{\partial E_{0x}}{\partial z} = \frac{i\omega}{c}H_{0y} \\ \frac{\partial H_{0y}}{\partial z} = \frac{i\epsilon_{xx}\omega}{c}E_{0x} \\ ikH_{0y} = -\frac{i\epsilon_{zz}\omega}{c}E_{0z} \end{cases} \quad \begin{cases} \frac{\partial H_{0z}}{\partial z} + ikH_{0x} = 0 \\ -ikH_{0z} + \frac{\partial H_{0x}}{\partial z} = -\frac{i\epsilon_{yy}\omega}{c}E_{0y} \\ \frac{\partial E_{0y}}{\partial z} = -\frac{i\omega}{c}H_{0x} \\ ikE_{0y} = \frac{i\omega}{c}H_{0z} \end{cases} \quad (2.42)$$

for the TM and TE modes. Notice that it has been assumed $\epsilon_{xx} \neq \epsilon_{yy} \neq \epsilon_{zz}$. Combining the equations of the TM system one gets the analogous of (2.32):

$$\begin{cases} \frac{\partial^2 E_{0z}}{\partial z^2} \frac{\epsilon_{zz1}}{\epsilon_{xx1}} - k^2 E_{0z} + \frac{\omega^2 \epsilon_{zz1}}{c^2} E_{0z} = 0 & z > 0 \\ \frac{\partial^2 E_{0z}}{\partial z^2} \frac{\epsilon_{zz2}}{\epsilon_{xx2}} - k^2 E_{0z} + \frac{\omega^2 \epsilon_{zz2}}{c^2} E_{0z} = 0 & z < 0 \end{cases} \quad (2.43)$$

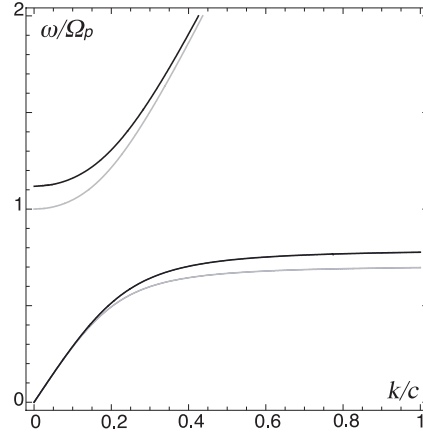


Figure 2.6: Dispersion relation of the surface described by the dielectric tensor (2.48), the TE modes (gray line) remain unchanged with respect to the isotropic case, whereas TM modes (black line) are affected by the anisotropy.

that, imposing the same solution (2.33) gives now:

$$\gamma_j^{TM} = \sqrt{\left(\frac{k^2}{\epsilon_{zzj}} - \frac{\omega^2}{c^2}\right) \epsilon_{xxj}} \quad (2.44)$$

whereas, from the TE modes one gets the result:

$$\gamma_j^{TE} = \sqrt{k^2 - \epsilon_{yyj} \frac{\omega^2}{c^2}} \quad (2.45)$$

The other components of the fields can be derived exactly as before and, by imposing the boundary conditions at the interface, one obtains the dispersion relations:

$$k_{TE}^2 = \frac{\omega^2}{c^2} \left(\frac{\epsilon_{yy1} \epsilon_{yy2}}{\epsilon_{yy1} + \epsilon_{yy2}} \right) \quad k_{TM}^2 = \frac{\omega^2}{c^2} \left(\frac{(\epsilon_{xx2} - \epsilon_{xx1}) \epsilon_{zz1} \epsilon_{zz2}}{\epsilon_{zz2} \epsilon_{xx2} - \epsilon_{zz1} \epsilon_{xx1}} \right) \quad (2.46)$$

the TE and TM modes have now different dispersion relations. In the non relativistic regime ($k \gg \omega/c$) the dispersion relation is simply given by:

$$\epsilon_{zz2} \epsilon_{xx2} = \epsilon_{zz1} \epsilon_{xx1} \quad (2.47)$$

If one of the two media is the vacuum and the other is an uniaxial medium with the Drude and Lorentz models:

$$\epsilon_{xx} = \epsilon_{yy} = 1 - \frac{\Omega_p^2}{\omega^2} \quad \epsilon_{zz} = 1 + \frac{\Omega_p^2}{\omega_0^2 - \omega^2} \quad (2.48)$$

a comparison is possible with the homogeneous surface treated before. It is shown in figure 2.6. The main effect of the anisotropy is to change the surface plasmon frequency shifting both the localized and delocalized modes.

2.3.2 Cavity modes

In the preceding section it has been discussed the problem of the electromagnetic modes and their dispersion relation in presence of an interface between two media. The starting point

to show how the Lifshitz's theory generalize the Casimir's original result is to move from the Casimir's perfectly reflecting cavity to a real dielectric cavity: the dispersion relation of the electromagnetic modes between two interfaces (between three mediums) must be studied. The procedure is exactly the same described in the preceding section but now two boundary conditions must be satisfied simultaneously.

Consider the cavity of thickness d represented in figure 2.8 (a), it is formed by the two interfaces between three different media described through their dielectric functions ϵ_1 , ϵ_2 and ϵ_3 . The field (2.30) will be again the solution for the systems (2.31) but now three different regions must be considered:

$$E_0(z) = \begin{cases} Ae^{\gamma_1 z} & z < 0 \\ Be^{-\gamma_3 z} + Ce^{\gamma_3 z} & 0 < z < d \\ De^{-\gamma_2 z} & z > d \end{cases} \quad (2.49)$$

the four constants will be determined imposing the field conditions (see appendix A) at the two interfaces.

$$\begin{aligned} E_{0x}^{(1)}(z=0) &= E_{0x}^{(3)}(z=0) & H_{0y}^{(1)}(z=0) &= H_{0y}^{(3)}(z=0) \\ E_{0x}^{(3)}(z=d) &= E_{0x}^{(2)}(z=d) & H_{0y}^{(3)}(z=d) &= H_{0y}^{(2)}(z=d) \end{aligned} \quad (2.50)$$

where the E_0 and H_0 components come from the Maxwell's equations (2.31):

$$\begin{aligned} E_{0x}^{(1)} &= \left. \frac{i}{k} \frac{\partial E_{0z}}{\partial z} \right|_1 = \frac{i}{k} \gamma_1 A e^{\gamma_1 z} \\ E_{0x}^{(3)} &= \left. \frac{i}{k} \frac{\partial E_{0z}}{\partial z} \right|_3 = -\frac{i}{k} \gamma_3 (B e^{-\gamma_3 z} + C e^{\gamma_3 z}) \\ E_{0x}^{(2)} &= \left. \frac{i}{k} \frac{\partial E_{0z}}{\partial z} \right|_2 = -\frac{i}{k} \gamma_2 D e^{-\gamma_2 z} \\ H_{0y}^{(1)} &= -\frac{\omega \epsilon(\omega)}{c} E_{0z}|_1 = -\frac{\omega}{c} \epsilon_1 A e^{\gamma_1 z} \\ H_{0y}^{(3)} &= -\frac{\omega \epsilon(\omega)}{c} E_{0z}|_3 = -\frac{\omega}{c} \epsilon_3 (B e^{-\gamma_3 z} + C e^{\gamma_3 z}) \\ H_{0y}^{(2)} &= -\frac{\omega \epsilon(\omega)}{c} E_{0z}|_2 = -\frac{\omega}{c} \epsilon_2 D e^{-\gamma_2 z} \end{aligned} \quad (2.51)$$

and, through the logarithmic derivatives:

$$\begin{aligned} \gamma_1 A &= -\gamma_3 (B - C) & \epsilon_1 A &= \epsilon_3 (B + C) \\ \epsilon_3 (B e^{-\gamma_3 d} + C e^{\gamma_3 d}) &= \epsilon_2 D e^{-\gamma_2 d} & \epsilon_3 (B e^{-\gamma_3 d} + C e^{\gamma_3 d}) &= \epsilon_2 D e^{-\gamma_2 d} \end{aligned} \quad (2.52)$$

solving this algebraic system in the variables A , B , C and D one can completely determine the fields behaviour. To calculate the dispersion relation of the electromagnetic modes one has to impose that the determinant vanishes:

$$g_{TM}(\omega, d) = \begin{vmatrix} \epsilon_2 & -\epsilon_3 & -\epsilon_3 & 0 \\ 0 & \epsilon_3 e^{-\gamma_3 d} & \epsilon_3 e^{\gamma_3 d} & \epsilon_2 e^{-\gamma_2 d} \\ \gamma_1 & \gamma_3 & -\gamma_3 & 0 \\ 0 & \gamma_3 e^{-\gamma_3 d} & \gamma_3 e^{\gamma_3 d} & -\gamma_2 e^{-\gamma_2 d} \end{vmatrix} = 0 \quad (2.53)$$

$$g_{TM}(\omega, d) = e^{d(\gamma_3 - \gamma_2)} (\epsilon_1 \gamma_3 + \epsilon_3 \gamma_1) (\epsilon_3 \gamma_2 + \epsilon_2 \gamma_3) - e^{-d(\gamma_3 + \gamma_2)} (\epsilon_3 \gamma_1 - \epsilon_1 \gamma_3) (\epsilon_2 \gamma_3 - \epsilon_3 \gamma_2) = 0$$

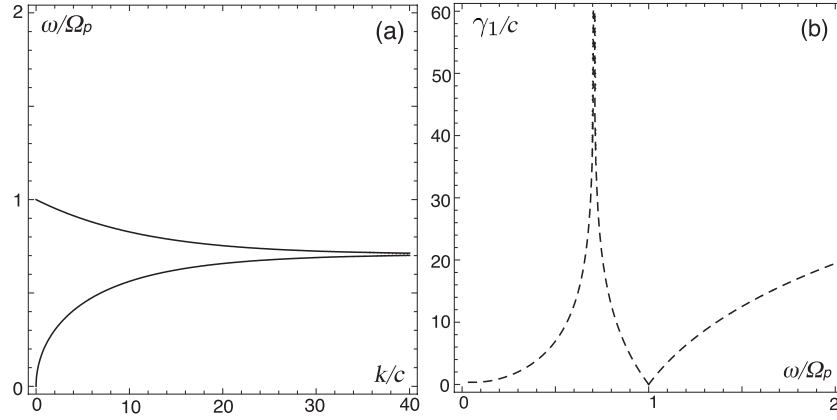


Figure 2.7: (a) represents the $\omega(k)$ relation for two interacting isotropic plasma metal surfaces at a distance d . (b) represents $\gamma_1(\omega)$ for the same two interacting isotropic plasma metal surfaces, only localized modes exist (dashed line). Both plots have been obtained in the non relativistic limit.

The zeroes of the determinant give the dispersion relations $\omega(k, d)$ for the frequencies that satisfy the field conditions at the interfaces.

The $\omega(k, d)$ relation for the TM modes of a cavity of dimension d , in the non relativistic approximation, is plotted in figure 2.7 (a): the plasma model has been used to describe the media properties. The forbidden region, present in the case of a single interface (see figure 2.5 (a)), vanishes thanks to the interaction between the two interfaces. For this configuration no modes are present with frequency higher than the bulk plasma frequency. In the case of large d values, the interaction between the two interfaces becomes negligible and the dispersion relation should converge to the single surface one of figure 2.5 (a). Figure 2.7 (b) shows the γ_1 values in the case of a cavity, in the non relativistic limit, only surface modes are present:

$$g_{TM}(\omega, d) = (\epsilon_1 + \epsilon_3)(\epsilon_3 + \epsilon_2) - e^{-2dk}(\epsilon_3 - \epsilon_1)(\epsilon_2 - \epsilon_3) = 0 \quad g_{TE}(\omega, d) = 0 \quad (2.54)$$

A detailed study of the modes of a single film standing alone in the vacuum, can be found in reference [28].

The same procedure holds also for the TE modes leading to:

$$g_{TE} = e^{d(\gamma_3 - \gamma_2)}(\gamma_3 + \gamma_1)(\gamma_2 + \gamma_3) - e^{-d(\gamma_3 + \gamma_2)}(\gamma_1 - \gamma_3)(\gamma_3 - \gamma_2) = 0 \quad (2.55)$$

If one is able to solve explicitly the two determinants, according to the definition (2.6) the zero point energy can be calculated as:

$$E(d) = \frac{\hbar}{2} \sum_{\alpha, k} \omega_{\alpha}(k, d) = \frac{\hbar}{2} \sum_{\alpha} \frac{L^2}{4\pi^2} \int dk \omega_{\alpha}(k, d) = \sum_{\alpha} \frac{L^2 \hbar}{4\pi} \int k dk \omega_{\alpha}(k, d) \quad (2.56)$$

where the α index runs over TE and TM modes whereas L^2 is the layers surface. The media properties enter the energy expression through the dependence of $\omega_{\alpha}(k, d)$ from the dielectric function. Unfortunately the direct solution of the determinants is not possible except for a small number of limiting cases. Van Kampen proposed to overcome the problem of finding the zeroes of the determinants with the help of complex contour integrals, in particular using the *logarithmic counter* theorem: *Let $f(z)$ be a meromorphic² function in a region D and*

²A complex function is said to be meromorphic in a given region of the complex plane D if it is analytic in the whole D except for a set of isolated poles.

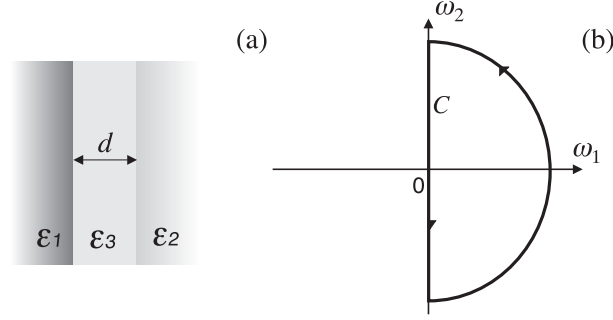


Figure 2.8: (a) sketch of the three layers cavity used by Lifshitz. (b) contour C in the complex frequency plane used for the van Kampen integral representation (2.58).

$g(z)$ an analytic function in D . Let C be a closed contour in D on which $f(z)$ is both analytic and nowhere zero. If $f(z)$ has, within C , w zeroes at $z = a_i$ of order n_i and p poles at $z = b_j$ of order m_j , then:

$$\frac{1}{i2\pi} \oint_C \frac{g(z)}{f(z)} \frac{df(z)}{dz} dz = \sum_{i=1}^w n_i g(a_i) - \sum_{j=1}^p m_j g(b_j) \quad (2.57)$$

Choosing $f(z)$ to be one of the determinants (2.55) or (2.53) and $g(z) = \omega$, and provided that these determinants have no poles, the previous integral gives all the ω values that make the determinants vanish, i.e. it gives all the dispersion relations $\omega(k, d)$ one was looking for. In other words:

$$\sum_{\alpha} \omega_{\alpha}(k, d) = \frac{1}{i2\pi} \oint_C \omega \left(\frac{1}{g_{TM}} \frac{\partial g_{TM}}{\partial \omega} + \frac{1}{g_{TE}} \frac{\partial g_{TE}}{\partial \omega} \right) d\omega \quad (2.58)$$

the C contour is reported in figure 2.8 (b). To guarantee the pertinence of (2.58), g_{TM} and g_{TE} must not vanish on the contour C : on the C semicircle, where $|\omega| \rightarrow \infty$, the medium response is negligible, $\epsilon(\omega) \rightarrow 1$ and $g_{TM} = g_{TE} \neq 0$; on the imaginary axe, for all the physical response functions described in section 1.8, if $Re[\omega] = 0$ one has $g_{TM,TE} \neq 0 \forall Im[\omega]$. The integral (2.58) over the contour C can be decomposed along the imaginary axe and the semicircle Γ :

$$\lim_{|\omega| \rightarrow \infty} \left[\frac{1}{i2\pi} \int_{i\infty}^{-i\infty} \omega \frac{1}{g_{TM}} \frac{\partial g_{TM}}{\partial \omega} d\omega + \frac{1}{i2\pi} \int_{\Gamma} \omega \frac{1}{g_{TM}} \frac{\partial g_{TM}}{\partial \omega} d\omega + \frac{1}{i2\pi} \int_{i\infty}^{-i\infty} \omega \frac{1}{g_{TE}} \frac{\partial g_{TE}}{\partial \omega} d\omega + \frac{1}{i2\pi} \int_{\Gamma} \omega \frac{1}{g_{TE}} \frac{\partial g_{TE}}{\partial \omega} d\omega \right] \quad (2.59)$$

this quantity is divergent, exactly as in section 2.2, the vacuum energy must be renormalized. The divergence can be made more explicit: when $|\omega| \rightarrow \infty$ the dielectric functions go to 1 and (2.34) becomes:

$$\gamma_j = \sqrt{k^2 - \epsilon_j(\omega) \frac{\omega^2}{c^2}} \simeq \sqrt{-\frac{\omega^2}{c^2}} \quad (2.60)$$

It should be noticed that a local treatment (without the inclusion of local field effects, i.e. a complicated k dependence in the dielectric functions) is possible only if the modes frequencies

are such that $\frac{c}{\omega} \gg d$. So, while ω diverges, d must tend to zero more rapidly and one get $e^{d(\gamma_3 - \gamma_2)} \simeq 1$. By this considerations:

$$\begin{aligned} g_{TM}(|\omega| \rightarrow \infty) &= g_{TE}(|\omega| \rightarrow \infty) = -4\frac{\omega^2}{c^2} \\ \frac{\partial g_{TM}(|\omega| \rightarrow \infty)}{\partial \omega} &= \frac{\partial g_{TE}(|\omega| \rightarrow \infty)}{\partial \omega} = -8\frac{\omega}{c^2} \end{aligned} \quad (2.61)$$

and the integral (2.59) becomes:

$$\lim_{|\omega| \rightarrow \infty} \left[\frac{1}{i2\pi} \int_{i\infty}^{-i\infty} \omega \frac{1}{g_{TM}} \frac{\partial g_{TM}}{\partial \omega} d\omega + \frac{1}{i2\pi} \int_{i\infty}^{-i\infty} \omega \frac{1}{g_{TE}} \frac{\partial g_{TE}}{\partial \omega} d\omega - \frac{2i}{\pi} \int_{\Gamma} d\omega \right] \quad (2.62)$$

the last integral is clearly divergent. Before proceeding to the renormalization of the vacuum energy, it is convenient to rewrite the first two integrals of (2.62) in a different way with an integration by parts:

$$\begin{aligned} &\frac{1}{i2\pi} \int_{i\infty}^{-i\infty} \omega \frac{1}{g_{TM}} \frac{\partial g_{TM}}{\partial \omega} d\omega + \frac{1}{i2\pi} \int_{i\infty}^{-i\infty} \omega \frac{1}{g_{TE}} \frac{\partial g_{TE}}{\partial \omega} d\omega = \\ &= \frac{1}{i2\pi} \left[\omega \ln(g_{TM}) \Big|_{i\infty}^{-i\infty} - \int_{i\infty}^{-i\infty} \ln(g_{TM}) d\omega + \omega \ln(g_{TE}) \Big|_{i\infty}^{-i\infty} - \right. \\ &\left. - \int_{i\infty}^{-i\infty} \ln(g_{TE}) d\omega \right] = \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln(g_{TM}) d\omega + \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln(g_{TE}) d\omega \end{aligned} \quad (2.63)$$

in the last step two terms disappeared: all the physical dielectric functions described in section 1.8 have the property:

$$\lim_{\omega \rightarrow \pm\infty} \epsilon(\omega) = \lim_{\omega \rightarrow \pm\infty} \epsilon(i\omega) = 1 \quad (2.64)$$

so that $g_{TM}(i\infty) = g_{TM}(-i\infty) = g_{TE}(i\infty) = g_{TE}(-i\infty)$.

To renormalize the energy one must follow the Casimir's procedure, subtracting the vacuum energy of the cavity with $d \rightarrow \infty$ to the vacuum energy for the cavity of finite thickness. For large cavity thicknesses the determinants become:

$$\begin{aligned} g_{TM}^{(\infty)} &= e^{d(\gamma_3 - \gamma_2)} (\epsilon_1 \gamma_3 + \epsilon_3 \gamma_1) (\epsilon_3 \gamma_2 + \epsilon_2 \gamma_3) \\ g_{TE}^{(\infty)} &= e^{d(\gamma_3 - \gamma_2)} (\gamma_3 + \gamma_1) (\gamma_2 + \gamma_3) \end{aligned} \quad (2.65)$$

and the renormalized dispersion relations can be calculated:

$$\begin{aligned} \sum_{n,\alpha} \omega_{n,\alpha}(k)_{ren.} &= \sum_{n,\alpha} \omega_{n,\alpha}(k) - \lim_{d \rightarrow \infty} \sum_{n,\alpha} \omega_{n,\alpha}(k) = \\ &= \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln(g_{TM}) d\omega + \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln(g_{TE}) d\omega - \frac{2i}{\pi} \int_{\Gamma} d\omega - \\ &- \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln(g_{TM}^{(\infty)}) d\omega - \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln(g_{TE}^{(\infty)}) d\omega + \frac{2i}{\pi} \int_{\Gamma} d\omega = \\ &= \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln \left(\frac{g_{TM}}{g_{TM}^{(\infty)}} \right) d\omega + \frac{1}{i2\pi} \int_{-i\infty}^{i\infty} \ln \left(\frac{g_{TE}}{g_{TE}^{(\infty)}} \right) d\omega \end{aligned}$$

where the divergent parts of the contour integrals cancel out. Notice that in the Casimir derivation, before renormalizing the energy, a regularization function was used. Here the regularization function is implicitly contained in the integrand that is not divergent except for the two terms that cancel out. For interacting objects of different shape the cancellation of divergencies is not so obvious. Redefining the ratio $g_\alpha/g_\alpha^{(\infty)}$ as Q_α :

$$\begin{aligned} Q_{TM} &= 1 - e^{-2d\gamma_3} \frac{(\epsilon_1\gamma_3 - \epsilon_3\gamma_1)(\epsilon_2\gamma_3 - \epsilon_3\gamma_2)}{(\epsilon_1\gamma_3 + \epsilon_3\gamma_1)(\epsilon_3\gamma_2 + \epsilon_2\gamma_3)} \\ Q_{TE} &= 1 - e^{-2d\gamma_3} \frac{(\gamma_3 - \gamma_1)(\gamma_3 - \gamma_2)}{(\gamma_3 + \gamma_1)(\gamma_2 + \gamma_3)} \end{aligned} \quad (2.66)$$

with the change of variable $\xi = -i\omega$ the energy becomes

$$\begin{aligned} E(d)_{ren.} &= \frac{\hbar L^2}{2} \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \sum_{\alpha} \omega_{\alpha}(k) d\mathbf{k} = \\ &= -\frac{\hbar L^2}{2} \frac{1}{4\pi^2} \frac{1}{2\pi} \left(\int_{-\infty}^{\infty} d\mathbf{k} \int_{-\infty}^{\infty} \ln(Q_{TM}(i\xi)) d\xi + \int_{-\infty}^{\infty} \ln(Q_{TE}(i\xi)) d\xi \right) \end{aligned} \quad (2.67)$$

Now the integrand is a real function because $\epsilon(i\omega)$ is a real function (see section 1.8). Using polar coordinates:

$$E(d)_{ren.} = -\frac{\hbar L^2}{4\pi^2} \int_0^{\infty} k dk \int_0^{\infty} [\ln(Q_{TM}(i\xi)) + \ln(Q_{TE}(i\xi))] d\xi \quad (2.68)$$

The frequency integration domain has been reduced only to positive values because the integrand is an even function. This integrand property comes directly from the even parity of the London's transform of the dielectric function (see appendix C of the previous chapter) in particular, from the definition (1.207), it is easy to see that under the change of variable $\omega \rightarrow -\omega$, $\epsilon(i\omega)$ remains unchanged. Another frequently form to express the same energy is through the change of variable $k^2 = \frac{\xi^2}{c^2}(p^2 - 1)$:

$$E(d)_{ren.} = -\frac{\hbar L^2}{4\pi^2 c^2} \int_1^{\infty} p dp \int_0^{\infty} \xi^2 [\ln(Q_{TM}(i\xi)) + \ln(Q_{TE}(i\xi))] d\xi \quad (2.69)$$

now Q_{TM} and Q_{TE} have different expressions due to the modification of γ_i :

$$\begin{aligned} \gamma_i &= \sqrt{k^2 - \frac{\omega^2}{c^2} \epsilon_i(\omega)} = \sqrt{k^2 + \frac{\xi^2}{c^2} \epsilon_i(i\xi)} = \sqrt{k^2 + \frac{k^2}{p^2-1} \epsilon_i(\omega)} = \\ &= \frac{k}{\sqrt{p^2-1}} \sqrt{p^2 - 1 + \epsilon_i(i\xi)} = \frac{\xi}{c} K_i \end{aligned} \quad (2.70)$$

where the definition $K_i = \sqrt{p^2 - 1 + \epsilon_i(i\xi)}$ has been adopted

$$\begin{aligned} Q_{TM} &= 1 - e^{-2d\frac{\xi}{c} K_3} \frac{(\epsilon_1 K_3 - \epsilon_3 K_1)(\epsilon_2 K_3 - \epsilon_3 K_2)}{(\epsilon_1 K_3 + \epsilon_3 K_1)(\epsilon_3 K_2 + \epsilon_2 K_3)} \\ Q_{TE} &= 1 - e^{-2d\frac{\xi}{c} K_3} \frac{(K_3 - K_1)(K_3 - K_2)}{(K_3 + K_1)(K_2 + K_3)} \end{aligned} \quad (2.71)$$

Finally, differentiating with respect to d one get the macroscopic dispersion force per unit surface:

$$F(d) = -\frac{\partial E(d)_{ren.}}{\partial d} \frac{1}{L^2} \quad (2.72)$$

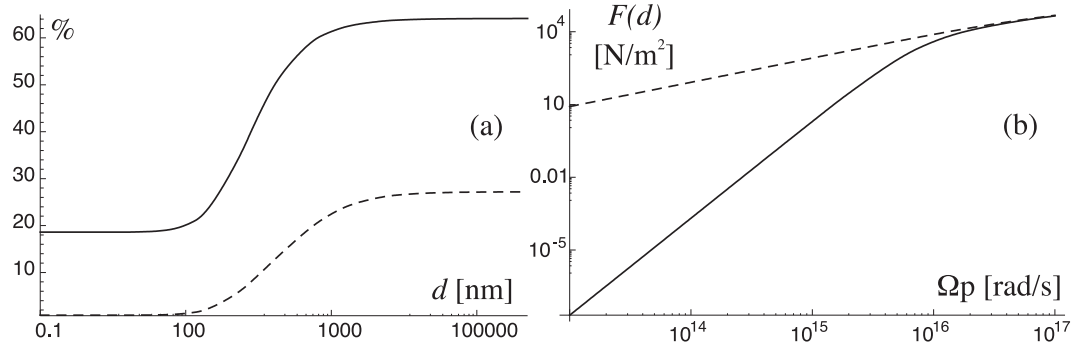


Figure 2.9: (a) Relative percent difference of the force in a plasma cavity and in a Lorentz cavity as a function of the cavity dimension d for $\Omega_p = 2 \cdot 10^{16}$ rad/s and $\omega_0 = 10^{15}$ rad/s (dashed line) or $\omega_0 = 5 \cdot 10^{15}$ rad/s (continuous line). (b) Force as a function of the cavity plasma frequency for a cavity of 10 nm length described by the plasma model (dashed line) or the Lorentz model (continuous line) with $\omega_0 = 5 \cdot 10^{15}$ rad/s.

or, explicitly:

$$F(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty k dk \int_0^\infty \gamma_3 \left[\frac{1 - Q_{TM}(i\xi)}{Q_{TM}(i\xi)} + \frac{1 - Q_{TE}(i\xi)}{Q_{TE}(i\xi)} \right] d\xi \quad (2.73)$$

$$F(d) = -\frac{\hbar}{2\pi^2 c^3} \int_1^\infty p dp \int_0^\infty K_3 \xi^3 \left[\frac{1 - Q_{TM}(i\xi)}{Q_{TM}(i\xi)} + \frac{1 - Q_{TE}(i\xi)}{Q_{TE}(i\xi)} \right] d\xi \quad (2.74)$$

Many information about the dispersion forces behaviour can be obtained from this formula: (i) first of all it is well known [29] that, if $\epsilon_3 = 1$ the force can take only negative values, i.e. two interacting plates separated by the vacuum can only attract each other. If $\epsilon_1 \leq \epsilon_3 \leq \epsilon_2$ the force can also change sign becoming repulsive. This behaviour can be easily understood looking at the interface modes only, considering surface plasmons in terms of macroscopic oscillations of the negative charge (see section 1.8). A positive force has not yet been observed but the first preliminar step has been accomplished making the first measurement of dispersion forces in liquids [30]. (ii) The force between dielectrics is small compared with the one between metals (see figure 2.9 (a)), dielectrics are more permeable than metals, less modes energy remains trapped inside the cavity and the resulting macroscopic force is smaller. (iii) Figure 2.9 (b) shows how sensitive the force can be to a change in the media parameters: moving the plasma frequency of a plasma or a Lorentz cavity in the range $10^{13} \div 10^{17}$ rad/s the force can change of several orders of magnitude. This is a very important point also when a comparison between theory and experiments is needed: depending on how the sample is prepared, the parameters for the model dielectric functions to be used in (2.73) can change significantly, producing a significant modification in the dispersion force values [31, 32]. Notice that, for large d values, both the dielectric function models lead to the Casimir's force limit.

In the limiting case of two perfectly reflecting slab separated by the vacuum ($\epsilon_1 = \epsilon_2 \rightarrow -\infty$ and $\epsilon_3 = 1$) one gets

$$Q_{TM} = Q_{TE} = 1 - e^{-2d\xi p/c}$$

$$F(d) = -\frac{\hbar}{\pi^2 c^3} \int_1^\infty p^2 dp \int_0^\infty \xi^3 \frac{e^{-2d\xi p/c}}{1 - e^{-2d\xi p/c}} d\xi \quad (2.75)$$

using the known integral:

$$\int_0^{\infty} \frac{x^3 e^{-ax}}{1 - e^{-ax}} dx = \frac{\pi^4}{15a^4} \quad (2.76)$$

the Casimir result (2.28) is achieved.

The theory as been generalized by B.E. Sernelius to account for non local effects, for those materials in which the \mathbf{k} dependence of the dielectric functions is not negligible [33].

2.3.3 Finite temperature formalism

The finite temperature extension of the dispersion forces theory must be carried out starting from thermodynamics considerations. From the fields point of view, at finite temperature, the cavity is no more empty: a gas of photons exists whose number and energy distribution depends upon T according to the black body radiation laws. From the media point of view one can say that the media modes are no more in their ground states, the modes spectrum is differently populated depending on the temperature values. In any case the cavity walls are again macroscopically neutral and the force acting between them pertains again to the dispersion forces family. The partition function of an ephemeral bosons gas is given by:

$$Q = \sum_{\{n_j\}} e^{-\frac{E_{n_j}(N,V)}{k_B T}} = \sum_{\{n_j\}} e^{-\sum_{\mathbf{k}} \frac{(1/2+n_j)\hbar\omega_{\mathbf{k}}}{k_B T}} = \prod_{\mathbf{k}} \sum_{n=0}^{\infty} e^{-\frac{(1/2+n)\hbar\omega_{\mathbf{k}}}{k_B T}} \quad (2.77)$$

where the sum over $\{n_j\}$ represents the sum over all the occupation numbers sets (i.e. all the possible microscopic configurations of the system), E_{n_j} is the energy of the j -th occupation numbers sets and k_B the Boltzmann constant. Using the usual definition for the Helmholtz's free energy:

$$A = -k_B T \ln Q = -k_B T \sum_{\mathbf{k}} \ln \left[\sum_{n=0}^{\infty} e^{-\frac{(1/2+n)\hbar\omega_{\mathbf{k}}}{k_B T}} \right] \quad (2.78)$$

Summing the known series and playing with the Eulero's formula, one obtains:

$$A = k_B T \sum_{\mathbf{k}} \ln \left[2 \sinh \left(\frac{1}{2} \frac{\hbar\omega_{\mathbf{k}}}{k_B T} \right) \right] \quad (2.79)$$

Following the zero temperature derivation, the energy of the cavity of amplitude d must be normalized subtracting the energy of the infinitely large cavity.

$$A(d)_{ren.} = k_B T \sum_{\mathbf{k}} \left\{ \ln \left[2 \sinh \left(\frac{1}{2} \frac{\hbar\omega_{\mathbf{k}}(d)}{k_B T} \right) \right] - \ln \left[2 \sinh \left(\frac{1}{2} \frac{\hbar\omega_{\mathbf{k}}(\infty)}{k_B T} \right) \right] \right\} \quad (2.80)$$

here $\omega_{\mathbf{k}}(d)$ and $\omega_{\mathbf{k}}(\infty)$ are all the possible dispersion relations of the cavity obtainable from the solution of the determinants (2.55) and (2.53) for the three layers configuration. Using the logarithmic counter theorem (2.57) with $f(z)$ given by the determinants (2.55) and (2.53) and $g(z)$ given by:

$$g(z) = k_B T \ln \left[2 \sinh \left(\frac{\hbar z}{2k_B T} \right) \right] \quad (2.81)$$

the free energy (2.79) becomes:

$$\begin{aligned} A &= \frac{k_B T}{2\pi i} \sum_{\mathbf{k}, \alpha} \int_{-\infty}^{\infty} d\omega \ln \left[2 \sinh \left(\frac{\hbar\omega}{2k_B T} \right) \right] \frac{d}{d\omega} \ln [g_{\alpha}(\omega)] = \\ &= -\frac{k_B T}{2\pi i} \sum_{\mathbf{k}, \alpha} \int_{-\infty}^{\infty} d\xi \ln \left[2 \sinh \left(\frac{\hbar i\omega}{2k_B T} \right) \right] \frac{d}{d\omega} \ln [g_{\alpha}(i\omega)] = \end{aligned} \quad (2.82)$$

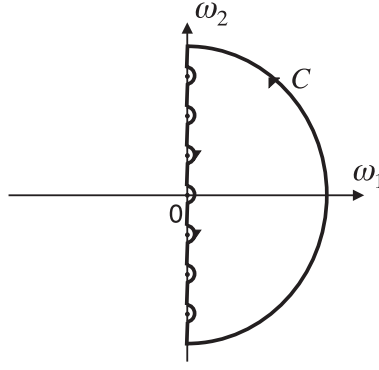


Figure 2.10: Contour used for the integration (2.84) to avoid the singularities located at the Matsubara frequencies.

the same assumption on g_α of the zero temperature derivation hold even in this case, in the second step the variable $\xi = -i\omega$ as been introduced. Integrating by parts one can simplify the integrand expression:

$$A = -\frac{k_B T}{2\pi i} \sum_{\mathbf{k}, \alpha} \left\{ \ln \left[2 \sinh \left(\frac{\hbar i \omega}{2k_B T} \right) \right] \ln [g_\alpha(i\omega)] \Big|_{-\infty}^{\infty} - \right. \\ \left. - \frac{i\hbar}{2k_B T} \int_{-\infty}^{\infty} d\xi \frac{\cosh \left(\frac{\hbar i \omega}{2k_B T} \right)}{\sinh \left(\frac{\hbar i \omega}{2k_B T} \right)} \ln [g_\alpha(i\omega)] \right\} \quad (2.83)$$

the first term in the RHS cancels out with the $g_\alpha^{(\infty)}$ in the renormalization procedure (2.80) and will be omitted. The second term in the RHS contains an hyperbolic cotangent and, because of this, the integrand has poles at frequencies $z_n = i2\pi n k_B T / \hbar$, the so called *Matsubara frequencies*. The integral over the frequency ξ can be transformed into a discrete sum over the Matsubara frequencies by means of the residues theorem, integrating over the contour of figure 2.10 the free energy becomes:

$$A = k_B T \sum_{\mathbf{k}, \alpha} \sum_{n=0}^{\infty'} \ln [g_\alpha(i\Omega_n)] \quad (2.84)$$

where $\Omega_n = 2\pi n k_B T / \hbar$. Notice that a factor 2 appear because of the parity of $\epsilon(i\omega)$ that leads to the property $\ln g_\alpha(\Omega_n) = \ln g_\alpha(\Omega_{-n})$, only the $n = 0$ term must be taken once, this is why a prime symbol appear in the n sum. Finally, coming back to the renormalized energy expression one ends up with:

$$A(d)_{ren.} = k_B T \sum_{\mathbf{k}, \alpha} \sum_{n=0}^{\infty'} \ln \left[\frac{g_\alpha(i\Omega_n)}{g_\alpha^{(\infty)}(i\Omega_n)} \right] \quad (2.85)$$

or in the case of a continuum \mathbf{k} space

$$A(d)_{ren.} = \frac{k_B T L^2}{2\pi} \sum_{\alpha} \sum_{n=0}^{\infty'} \int k dk \ln \left[\frac{g_\alpha(i\Omega_n)}{g_\alpha^{(\infty)}(i\Omega_n)} \right] \quad (2.86)$$

using the determinants of the three layers system and differentiating with respect to the cavity amplitude d one gets the force per unit area:

$$F(d) = -\frac{k_B T}{\pi} \sum_{n=0}^{\infty'} \int_0^{\infty} k dk \gamma_3 \left[\frac{1 - Q_{TM}(i\Omega_n)}{Q_{TM}(i\Omega_n)} + \frac{1 - Q_{TE}(i\Omega_n)}{Q_{TE}(i\Omega_n)} \right] \quad (2.87)$$

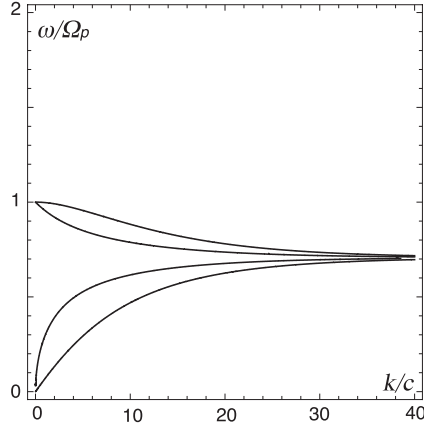


Figure 2.11: Black lines represent the dispersion relation $\omega(k)$ for two interacting plasma films of thickness $d_1 = d_2$ separated by a distance d . The calculation have been performed in the non relativistic limit.

This result was originally derived by Lifshitz and presented in its earlier work together with the zero temperature result, however it is known to suffer from many problems. The finite temperature extension of the Casimir's limit can be derived starting from the general formula (2.86) and, in the small temperature limit, the original Casimir's result can be obtained. However this derivation brings to light some known problems of the finite temperature theory and it is postponed at section 2.5, where the theory pathologies are largely discussed.

The finite temperature formalism accounts only for the force corrections due to the presence of a photon gas inside the interacting media. Another correction comes from the temperature modification of the dielectric response functions describing the media, this problem has been discussed by N. Inui [34], Høye et al. [35] and Yampolskii et al. [36] who introduced the temperature dependence in a model dielectric function $\epsilon(\omega, T)$. The joined effect of the temperature dependence and the non locality of the force has been investigated by Bo E. Sernelius [37].

2.3.4 Extension to more than three layers

A first extension of the three layers theory, described in section 2.3.2, has been provided by F. Zhou and L. Spruch [38]. Following the same procedure previously adopted, one can evaluate the energy and the force for the five layer system of figure 2.12 (a). One has simply to add two new interfaces, four new equations appear for the field conditions and (2.53) and (2.55) become 8×8 determinants. Now one deals with three geometrical parameters, d being again the thickness of the inner layer, d_1 and d_2 being the thicknesses of the outer finite layers. The dispersion relation becomes now more complicated, it is plotted in figure 2.11 for the case of two identical interacting plasma films of thickness $d_1 = d_2$ separated by a distance d . Doubling the number of interfaces two new branches appear. While the film thickness $d_1 = d_2$ increase the dispersion relation should converge to the one of a plasma cavity of dimension d (see figure 2.7 (a)), i.e. two branches must disappear. The logarithmic counter theorem can be used to transform the seek of the zeroes of the determinants into an integral over the complex frequencies plane. To normalize the theory, the energy at the separation $d \rightarrow \infty$ must be subtracted to the energy at a finite d value. Finally one gets:

$$E(d)_{ren.} = \frac{\hbar L^2}{4\pi^2} \int_0^\infty k dk \int_0^\infty [\ln(1 - Q_{TM}^1 Q_{TM}^2) + \ln(1 - Q_{TE}^1 Q_{TE}^2)] d\xi \quad (2.88)$$

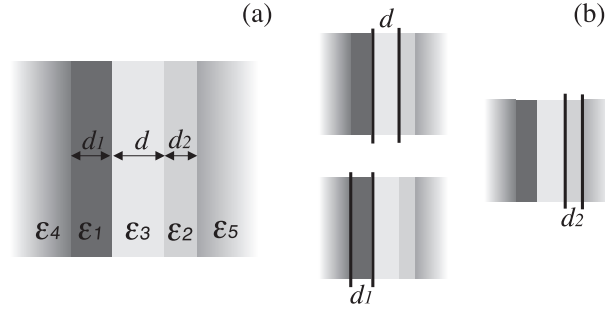


Figure 2.12: (a) sketch of the five layers cavity used by Zhou and Spruch. (b) black lines indicate the two interfaces between which the force is calculated differentiating the energy with respect to the different geometric parameters, namely $F(d)$, $F(d_1)$ and $F(d_2)$.

where:

$$Q_\alpha^1 = \frac{\rho_{13}^\alpha - \rho_{14}^\alpha e^{-2\gamma_1 d_1}}{1 - \rho_{13}^\alpha \rho_{14}^\alpha e^{-2\gamma_1 d_1}} e^{-\gamma_3 d} \quad Q_\alpha^2 = \frac{\rho_{23}^\alpha - \rho_{25}^\alpha e^{-2\gamma_2 d_2}}{1 - \rho_{23}^\alpha \rho_{25}^\alpha e^{-2\gamma_2 d_2}} e^{-\gamma_3 d} \quad (2.89)$$

and:

$$\rho_{i,j}^{TM} = \frac{\epsilon_j \gamma_i - \epsilon_i \gamma_j}{\epsilon_j \gamma_i + \epsilon_i \gamma_j} \quad \rho_{i,j}^{TE} = \frac{\gamma_i - \gamma_j}{\gamma_i + \gamma_j} \quad (2.90)$$

Notice that also d_1 or d_2 can be used as normalization parameters but the final energy result remains unchanged. To obtain the force one can differentiate the renormalized energy in the d , d_1 or d_2 variables. In each case one is deriving different forces, the three possibility are illustrated in figure 2.12 (b). Here only the force as a function of d is presented, for a comparison with expression (2.73):

$$F(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty k dk \int_0^\infty \gamma_3 \left[\frac{Q_{TM}^1 Q_{TM}^2}{1 - Q_{TM}^1 Q_{TM}^2} + \frac{Q_{TE}^1 Q_{TE}^2}{1 - Q_{TE}^1 Q_{TE}^2} \right] d\xi \quad (2.91)$$

In the limiting case of $d_1, d_2 \rightarrow \infty$ one obtains exactly the expression (2.73). If the media 1 and 2 are perfect metals ($\epsilon_1, \epsilon_2 \rightarrow -\infty$), the existence of the semi-infinite layers 4 and 5 is irrelevant, all the electromagnetic modes are trapped between the impenetrable layers 1 and 2, i.e. into the medium 3. In the latter conditions, if $\epsilon_3 = 1$ the Casimir result (2.28) is achieved.

The extension of the three and five layers models to a general multilayer system has been given by M.S. Tomáš and C. Raabe et al. [39, 40].

2.3.5 Anisotropic media

The same derivation of section 2.3.2 can be performed in the case of uniaxial media, i.e. media whose dielectric tensor is diagonal with $\epsilon_{xx} = \epsilon_{yy} \neq \epsilon_{zz}$. From the considerations of section 2.3.1 on the interface field conditions for anisotropic media, it is easy to understand why the TE modes remain unchanged and only the TM modes are affected by the anisotropic nature of the media. The uniaxial media generalization of the force (2.73) require the usage of (2.44) and (2.45):

$$Q_{TM} = 1 - e^{-2d\gamma_3^{TM}} \frac{(\epsilon_{xx1}\gamma_3^{TM} - \epsilon_{xx3}\gamma_1^{TM})(\epsilon_{xx2}\gamma_3^{TM} - \epsilon_{xx3}\gamma_2^{TM})}{(\epsilon_{xx1}\gamma_3^{TM} + \epsilon_{xx3}\gamma_1^{TM})(\epsilon_{xx3}\gamma_2^{TM} + \epsilon_{xx2}\gamma_3^{TM})} \quad (2.92)$$

$$Q_{TE} = 1 - e^{-2d\gamma_3^{TE}} \frac{(\gamma_3^{TE} - \gamma_1^{TE})(\gamma_3^{TE} - \gamma_2^{TE})}{(\gamma_3^{TE} + \gamma_1^{TE})(\gamma_2^{TE} + \gamma_3^{TE})}$$

and (2.73) becomes:

$$F(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty kdk \int_0^\infty \left[\gamma_3^{TM} \frac{1 - Q_{TM}(i\xi)}{Q_{TM}(i\xi)} + \gamma_3^{TE} \frac{1 - Q_{TE}(i\xi)}{Q_{TE}(i\xi)} \right] d\xi \quad (2.93)$$

Naturally, imposing $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz}$ one gets back the isotropic result. The same generalization applies for the five layer system of section 2.3.4, again only the TM modes are interested by the anisotropy:

$$Q_\alpha^1 = \frac{\rho_{13}^\alpha - \rho_{14}^\alpha e^{-2\gamma_1^\alpha d_1}}{1 - \rho_{13}^\alpha \rho_{14}^\alpha e^{-2\gamma_1^{\text{alpha}} d_1}} e^{-\gamma_3^\alpha d} \quad Q_\alpha^2 = \frac{\rho_{23}^\alpha - \rho_{25}^\alpha e^{-2\gamma_2^\alpha d_2}}{1 - \rho_{23}^\alpha \rho_{25}^\alpha e^{-2\gamma_2^\alpha d_2}} e^{-\gamma_3^\alpha d} \quad (2.94)$$

with:

$$\rho_{i,j}^{TM} = \frac{\epsilon_{xxj} \gamma_i^{TM} - \epsilon_{xxi} \gamma_j^{TM}}{\epsilon_{xxj} \gamma_i^{TM} + \epsilon_{xxi} \gamma_j^{TM}} \quad \rho_{i,j}^{TE} = \frac{\gamma_i^{TE} - \gamma_j^{TE}}{\gamma_i^{TE} + \gamma_j^{TE}} \quad (2.95)$$

and (2.91) becomes:

$$F(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty kdk \int_0^\infty \left[\gamma_3^{TM} \frac{Q_{TM}^1 Q_{TM}^2}{1 - Q_{TM}^1 Q_{TM}^2} + \gamma_3^{TE} \frac{Q_{TE}^1 Q_{TE}^2}{1 - Q_{TE}^1 Q_{TE}^2} \right] d\xi \quad (2.96)$$

Naturally other generalizations are possible, one can treat non uniaxial diagonal dielectric tensors or the more general case of non diagonal tensors, the expressions for γ_j^α and Q_α^j become more and more complicated.

2.4 Different approximations

As already stated in the Casimir's derivation of section 2.2 the dimension of the cavity is related to the dielectric properties of the cavity walls: the cavity walls can be thought to be perfectly reflecting mirrors only if the dimension of the cavity is large compared to the penetration depth of the field inside the walls. So it is expected that, for large d , only the static values of the dielectric functions contribute to the force. This can be formalized through the exponent of expression (2.71), if $d \rightarrow \infty$ the exponential goes to zero and the force vanish, except for the case in which also $\xi \rightarrow 0$, in that case the exponent remains small. In the large d limit the force expression remains the same but the dielectric functions $\epsilon_i(\omega)$ must be replaced by their static values $\epsilon_i(0)$. With such an assumption the frequency dependence of the integrand is extremely simple and the change of variable $x = 2d\xi K_3/c$ can be made in (2.74) to explicitate the d dependence:

$$F(d) = -\frac{\hbar c}{32\pi^2 d^4} \int_1^\infty p dp \int_0^\infty \frac{x^3}{K_3} \left[\frac{1 - Q_{TM}(0)}{Q_{TM}(0)} + \frac{1 - Q_{TE}(0)}{Q_{TE}(0)} \right] dx \quad (2.97)$$

In the large d regime the force goes as $1/d^4$ as in the Casimir's limit, in fact if $\epsilon_1 = \epsilon_2$ are plasma or Drude metals and $\epsilon_3 = 1$, the Casimir's result is recovered thanks to the $\epsilon_1(0)$ and $\epsilon_2(0)$ divergence.

In the opposite limit of small d a simpler expression of the force can be derived following Lifshitz. If $d \rightarrow 0$ the exponent of (2.71) can be set equal 1 except for the case in which also $K_3 \rightarrow \infty$, in such a situation:

$$K_3 = \sqrt{p^2 - 1 + \epsilon_1} \simeq p \quad (2.98)$$

and the force (2.74) becomes

$$F(d) = -\frac{\hbar}{2\pi^2 c^3} \int_1^\infty dp \int_0^\infty \frac{p^2 \xi^3 (\epsilon_1 - \epsilon_3)(\epsilon_2 - \epsilon_3)}{(\epsilon_1 + \epsilon_3)(\epsilon_2 + \epsilon_3) e^{2d \frac{\xi}{c} p} - (\epsilon_1 - \epsilon_3)(\epsilon_2 - \epsilon_3)} d\xi \quad (2.99)$$

where the TE modes disappeared. Now with the substitution $x = 2d\xi p/c$:

$$F(d) = -\frac{\hbar}{16\pi^2 d^3} \int_{2d\xi_c}^\infty x^2 dx \int_0^\infty \frac{d\xi}{\left(\frac{\epsilon_1 + \epsilon_3}{\epsilon_1 - \epsilon_3}\right) \left(\frac{\epsilon_2 + \epsilon_3}{\epsilon_2 - \epsilon_3}\right) e^x - 1} \quad (2.100)$$

in the small d limit $x = 2d\xi_c \simeq 0$ and one can write

$$\begin{aligned} \frac{1}{\left(\frac{\epsilon_1 + \epsilon_3}{\epsilon_1 - \epsilon_3}\right) \left(\frac{\epsilon_2 + \epsilon_3}{\epsilon_2 - \epsilon_3}\right) e^x - 1} &= \frac{\left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x}}{\left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x} \left[\left(\frac{\epsilon_1 + \epsilon_3}{\epsilon_1 - \epsilon_3}\right) \left(\frac{\epsilon_2 + \epsilon_3}{\epsilon_2 - \epsilon_3}\right) e^x - 1 \right]} = \\ &= \frac{\left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x}}{1 - \left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x}} \end{aligned} \quad (2.101)$$

Considering the geometric series:

$$\sum_{n=0}^{\infty} \left[\left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x} \right]^n = \frac{1}{1 - \left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x}} \quad (2.102)$$

of argument less than 1 for each $x > 0$, the integral becomes

$$\begin{aligned} F(d) &= -\frac{\hbar}{16\pi^2 d^3} \int_0^\infty x^2 dx \int_0^\infty d\xi \left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x} \sum_{n=0}^{\infty} \left[\left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right) \times \right. \\ &\times \left. \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right) e^{-x} \right]^n = -\frac{\hbar}{16\pi^2 d^3} \sum_{n=1}^{\infty} \int_0^\infty x^2 e^{-nx} dx \int_0^\infty \left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right)^n \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right)^n d\xi \end{aligned} \quad (2.103)$$

the integration over x is particularly simple and takes the d dependence out of the integral:

$$\int_0^\infty x^2 e^{-nx} dx = -\frac{1}{n^3} [(n^2 x^2 + 2nx + 2)e^{-nx}]_0^\infty = \frac{2}{n^3} \quad (2.104)$$

and finally the force is:

$$F(d) = -\frac{\hbar}{8\pi^2 d^3} \sum_{n=1}^{\infty} \frac{1}{n^3} \int_0^\infty \left(\frac{\epsilon_1 - \epsilon_3}{\epsilon_1 + \epsilon_3}\right)^n \left(\frac{\epsilon_2 - \epsilon_3}{\epsilon_2 + \epsilon_3}\right)^n d\xi \quad (2.105)$$

At short distances the force goes as $1/d^3$ multiplied by a factor that depends only upon the dielectric properties of the media. Notice that if ϵ_1 and ϵ_2 are both larger or smaller than ϵ_3 the integrand is always positive and the series too. But if this condition is not satisfied the series has alternate sign and the convergence of the force is slower. The second term of the series is $1/8$ smaller than the first, retaining only the $n = 1$ term one has:

$$F(d) = -\frac{\hbar}{8\pi^2 d^3} \int_0^\infty \frac{(\epsilon_1 - \epsilon_3)(\epsilon_2 - \epsilon_3)}{(\epsilon_1 + \epsilon_3)(\epsilon_2 + \epsilon_3)} d\xi \quad (2.106)$$

Figure 2.13 (a) shows the behaviour of the exact force for a plasma cavity calculated with (2.73), for large and small d values it approaches the Casimir result (2.28) and the small d series (2.105) respectively. Figure 2.13 (b) shows the relative percent difference between the the exact calculation and the approximate results of Casimir and Lifshitz.

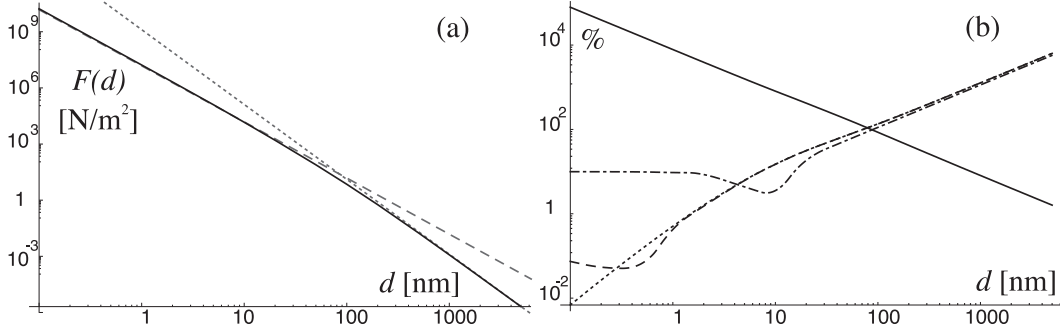


Figure 2.13: (a) Black line represents the exact dispersion force for a plasma cavity ($\Omega_p = 10^{16}$ rad/s) as a function of the cavity dimension d , gray lines are the small d series (dashes) and Casimir's (dots) approximations. (b) is the relative percent difference between the exact results and various approximations: Casimir (continuous line), small d series with $n = 1$ (dot-dashes) $n = 10$ (dashes) $n = 100$ (dots)

2.5 Finite temperature issues

If one try to recover the original Casimir's result from the finite temperature theory in the limiting case of $T \rightarrow 0$, a number of problems and questions arise. The first step of this derivation is the calculation of the force at finite temperature between two perfectly reflecting walls, i.e. the Casimir force at finite temperature. The $n = 0$ term in the force (2.87) can be calculated by two different procedures, that lead to different results:

- First the limit $\epsilon_1 = \epsilon_2 \rightarrow \infty$ is performed, than Ω_0 is set equal 0. In this case the TE and TM modes give the same contribution and the $n = 0$ term becomes:

$$F^0(d) = -\frac{k_B T}{\pi} \frac{1}{2} \int_0^{\infty} \frac{2k^2}{e^{2dk} - 1} dk = -\frac{k_B T}{4\pi} \frac{\xi(3)}{d^3} \quad (2.107)$$

where ξ is the Riemann's zeta function [41].

- First Ω_0 is set equal 0, than one let $\epsilon_1 = \epsilon_2 \rightarrow \infty$. In this case the TE modes of the $n = 0$ term vanish whereas the TM modes give the term:

$$F^0(d) = -\frac{k_B T}{\pi} \frac{1}{2} \int_0^{\infty} \frac{k^2}{e^{2dk} - 1} dk = -\frac{k_B T}{8\pi} \frac{\xi(3)}{d^3} \quad (2.108)$$

that is half of the force obtained with the previous prescription [42, 43, 44].

For all the $n > 0$ terms of the Matsubara sum, the two approaches give the same result. In the small temperature limit the Euler-Maclaurin's formula can be used giving [35]:

$$F(d) = -\frac{\hbar c \pi^2}{240 d^4} \left[1 + \frac{16}{3} \left(\frac{k_B T d}{\hbar c} \right)^4 \right] \quad (2.109)$$

with the first prescription and:

$$F(d) = -\frac{\hbar c \pi^2}{240 d^4} \left[1 + \frac{16}{3} \left(\frac{k_B T d}{\hbar c} \right)^4 \right] + \frac{k_B T}{8\pi} \frac{\xi(3)}{d^3} \quad (2.110)$$

in the case of vanishing TE modes. Both approximations work if $Td \ll 1$. In the vanishing TE case the temperature correction to the Casimir's force goes linearly with T , whereas in the first case it goes as T^4 . The same applies for the free energy:

$$A(d) = -\frac{\hbar c \pi^2}{720 d^3} \left[1 + 16 \left(\frac{k_B T d}{\hbar c} \right)^4 \right] \quad (2.111)$$

or:

$$A(d) = -\frac{\hbar c \pi^2}{720 d^3} \left[1 + 16 \left(\frac{k_B T d}{\hbar c} \right)^4 \right] + \frac{k_B T}{16\pi} \frac{\xi(3)}{d^2} \quad (2.112)$$

The entropy S is defined as the temperature derivative of the free energy A :

$$S = -\frac{\partial A}{\partial T} \quad (2.113)$$

and it goes to zero with the first prescription whereas it remains constant, violating the third thermodynamics principle, in the vanishing TE assumption. From this consideration it seems quite natural to discard the vanishing TE case in favour of the first one, in which one must replace the dielectric functions of the media and subsequently let the frequency go to zero. This suggestion seems to be supported by the fact that the zero temperature Casimir's force can be correctly obtained by the finite temperature Casimir's force only with the first prescription. By the exact solution of (2.87), reported in [35], with the first prescription one gets:

$$\begin{aligned} F &= -\frac{k_B T}{8\pi d^3} \sum_{k=1}^{\infty} \frac{1}{k^3} [s_2(\eta k) + 2s_1(\eta k) + 2s_0(\eta k)] \\ \eta &= \frac{2\pi d}{\hbar c} k_B T & s_0 &= \coth(\eta) \\ s_1 &= \frac{\eta}{\sinh^2(\eta)} & s_2 &= \frac{2\eta^2 \cosh(\eta)}{\sinh^3(\eta)} \end{aligned} \quad (2.114)$$

letting $T \rightarrow 0$, also η goes to zero recovering the original result:

$$F = -\frac{k_B T}{8\pi d^3} \sum_{k=1}^{\infty} \frac{6}{k^3 \eta^k} = -\frac{\hbar c \pi^2}{240 d^4} \quad (2.115)$$

whereas, using the vanishing TE prescription one achieves:

$$F = -\frac{k_B T}{8\pi d^3} \sum_{k=1}^{\infty} \frac{1}{k^3} [s_2(\eta k) + 2s_1(\eta k) + 2s_0(\eta k) - \xi(3)] \quad (2.116)$$

that is a smaller force due to the lack of the zero frequency TE contribution.

From the exact force between two ideal metals (2.114) one can recover the limit (2.109) when $Td \ll 1$ or the expression:

$$F = -\frac{k_B T}{4\pi d^3} \xi(3) \quad (2.117)$$

in the opposite situation of $Td \gg 1$. Letting the TE modes vanish, one gets one half the last result. Both $Td \ll 1$ and $Td \gg 1$ limits, together with the exact calculation (2.114), are shown in figure 2.14 (b) for a $T = 300^\circ$ cavity. Notice that, independently of the temperature, the force is always larger than the one calculated at zero temperature, the difference becoming larger on increasing the temperature or the cavity amplitude, see figure 2.14 (a).

Moving to a more realistic description of the dielectric properties of the cavity walls (finite Ω_p), one can introduce the Drude or the Lorentz models. In both cases the $n = 0$ term

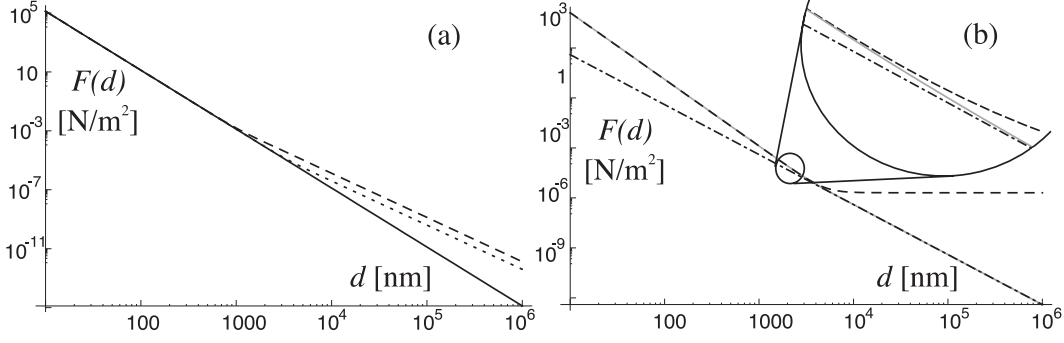


Figure 2.14: (a) Casimir's force as a function of the cavity amplitude at zero temperature (continuous line), at $T = 300^\circ\text{K}$ (dotted line) and at $T = 1000^\circ\text{K}$ (dashed line). (b) Casimir's force as a function of the cavity amplitude at $T = 300^\circ\text{K}$, (gray continuous line) and its small (dashed line) and large (dash-dotted line) Td approximations.

behaves according to the vanishing TE prescription when the Matsubara frequency Ω_n is set equal zero:

$$Q_{TM}^n = 1 - e^{-2d\sqrt{k^2 + \frac{\Omega_n^2}{c^2}}} \left[\frac{\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \beta\Omega_n}\right)\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} - \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \beta\Omega_n}\right)}}{\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \beta\Omega_n}\right)\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} + \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \beta\Omega_n}\right)}} \right]^2$$

$$\rightarrow 1 - e^{-2dk} \quad (2.118)$$

$$Q_{TE}^n = 1 - e^{-2d\sqrt{k^2 + \frac{\Omega_n^2}{c^2}}} \left[\frac{\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} - \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \beta\Omega_n}\right)}}{\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} + \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \beta\Omega_n}\right)}} \right]^2 \rightarrow 0$$

for the Drude model, and:

$$Q_{TM}^n = 1 - e^{-2d\sqrt{k^2 + \frac{\Omega_n^2}{c^2}}} \left[\frac{\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \omega_0^2}\right)\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} - \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \omega_0^2}\right)}}{\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \omega_0^2}\right)\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} + \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \omega_0^2}\right)}} \right]^2$$

$$\rightarrow 1 - e^{-2dk} \quad (2.119)$$

$$Q_{TE}^n = 1 - e^{-2d\sqrt{k^2 + \frac{\Omega_n^2}{c^2}}} \left[\frac{\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} - \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \omega_0^2}\right)}}{\sqrt{k^2 + \frac{\Omega_n^2}{c^2}} + \sqrt{k^2 + \frac{\Omega_n^2}{c^2}\left(1 + \frac{\Omega_p^2}{\Omega_n^2 + \omega_0^2}\right)}} \right]^2 \rightarrow 0$$

for the Lorentz model. Is this behaviour in contradiction with the third thermodynamics law? Figure 2.15 (a) shows the Lifshitz force as a function of the temperature for the case of a Drude cavity. The small T limits expressions, (2.109) and (2.110), are also reported for a comparison: at small temperatures the Drude force has a flat behaviour as the limit (2.109), fulfilling the third thermodynamics law. The vanishing TE prescription (2.110) as a linear behaviour as a function of T and, as previously emphasized, it violates the third thermodynamics law. The explanation for the Drude correct behaviour is discussed in reference [35]. From the figure it is also visible that, for realistic materials, the temperature can play both an increasing or decreasing effect on the force with respect to the zero temperature prediction.

To conclude there is a last issue that must be lighted: independently of the model dielectric

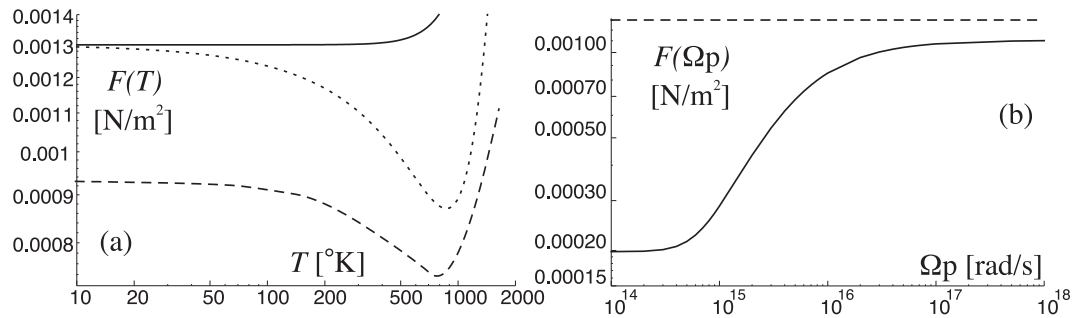


Figure 2.15: (a) Force as a function of the temperature for a Drude cavity having $\Omega_p = 10^{16}$ rad/sec and $\beta = 10^{15}$ rad/sec (dashed line), the (2.109) approximation (continuous line) and the (2.110) approximation (dotted line). The cavity amplitude is $d = 1\mu\text{m}$. (b) Force as a function of the plasma frequency Ω_p for the previous cavity, calculated at $T = 300^\circ\text{K}$. The zero temperature Casimir's result is also plotted (dashed line).

function used, the $n = 0$ term of the Matsubara sum brings to a force contribution that does not depend upon the dielectric parameters. Figure 2.15 (b) shows the force as a function of the plasma frequency for a fixed cavity dimension: an unphysical behaviour is evident for $\Omega_p \rightarrow 0$ where the force is expected to vanish because the medium is almost transparent. Another problem arises in the opposite limit $\Omega_p \rightarrow \infty$, where, due to the vanishing behaviour of the zero TE modes, the exact Casimir result can not be achieved. All these problems remain still unresolved but they vanish in the $T \rightarrow 0$ limit.

2.6 Concluding notes

A detailed survey on the different experimental techniques for the dispersion forces measurement is beyond the scope of this thesis, a rather complete review can be found in references [4] and [1]. There are many possible sources of error that can affect an experimental measurement of the dispersion forces preventing the theory to properly reproduce the experimental data. Among them the surface roughness of the interacting objects seems to play a major role together with the plate misalignment, possible corrections to the theoretical force expression are discussed in reference [4], while experimental evidence is provided by the recent measurements reported in [45].

A simple and unique way to renormalize the zero point energy can be found only for interacting systems of high symmetry, like two interacting planes, a sphere, two cylinders. Among the other geometries with a lower degree of symmetry, the configuration of a sphere (lens) above a disk (plane) is of great importance to calculate the force between a tip and a substrate in atomic force microscopy measurements. Unfortunately for this configuration an exact analytical solution is not possible, however a number of approximate techniques are available to calculate the dispersion energy for arbitrary shapes and geometries of the interacting objects. The *proximity force theorem* allows to use the result of two interacting planes of sections 2.3.2 and 2.3.3 to calculate the force between curved surfaces. The idea is to integrate the interaction between two infinitesimal plane surfaces over a surface with large curvature radius. The proximity force theorem can be used also to calculate roughness corrections. Limitations of the proximity force approximation have been shown by Chan et al. [46]. Another possibility is to calculate the force between two macroscopic objects summing pairwise over the Casimir-Polder force between single atoms: due to the non additivity of the force a correction factor must be used and there are several techniques by means of which it can be calculated. All these methods are reviewed and their limitations discussed

in reference [4]. More recent techniques are discussed in references [47] and [48].

Appendix A: Interface fields conditions

The conditions that electromagnetic fields must satisfy at an interface between two different media, can be derived starting from the Maxwell's equations. Making use of the divergence and Stokes' theorems, Maxwell's equations can be written in an integral form. Integrating over a certain volume V the divergence equations of the system (1.15) one gets:

$$\oint_V \nabla \cdot \mathbf{D} d\mathbf{x} = 4\pi \oint_V \eta_{\text{EXT}} d\mathbf{x} \quad (2.120)$$

$$\oint_V \nabla \cdot \mathbf{H} d\mathbf{x} = 0 \quad (2.121)$$

and, by the divergence theorem:

$$\oint_{\Sigma} \mathbf{D} \cdot \hat{\mathbf{n}} d\Sigma = 4\pi \oint_V \eta_{\text{EXT}} d\mathbf{x} \quad (2.122)$$

$$\oint_{\Sigma} \mathbf{H} \cdot \hat{\mathbf{n}} d\Sigma = 0 \quad (2.123)$$

where $\hat{\mathbf{n}}$ is the versor normal to the surface Σ enclosing the volume V . This is nothing but the Gauss theorem: the flux of the \mathbf{D} field through a closed surface Σ is proportional to the total charge enclosed within the surface. For \mathbf{H} one gets the same result, but a magnetic charge does not exist, so the right hand side of (2.123) vanishes. In other words the flux lines of a magnetic field are always closed and the flux through any closed surface is always zero. The two remaining equations of the system (1.15) must be integrated over an open surface Φ :

$$\oint_{\Phi} \nabla \times \mathbf{H} \cdot \hat{\mathbf{t}} d\Phi = \oint_{\Phi} \left[\frac{4\pi}{c} \mathbf{J}_{\text{EXT}} + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} \right] \cdot \hat{\mathbf{t}} d\Phi \quad (2.124)$$

$$\oint_{\Phi} \nabla \times \mathbf{E} \cdot \hat{\mathbf{t}} d\Phi = - \oint_{\Phi} \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} \cdot \hat{\mathbf{t}} d\Phi \quad (2.125)$$

and, using the Stokes' theorem:

$$\oint_C \mathbf{H} \cdot \hat{\mathbf{t}} dl = \oint_{\Phi} \left[\frac{4\pi}{c} \mathbf{J}_{\text{EXT}} + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} \right] \cdot \hat{\mathbf{t}} d\Phi \quad (2.126)$$

$$\oint_C \mathbf{E} \cdot \hat{\mathbf{t}} dl = - \oint_{\Phi} \frac{1}{c} \frac{\partial \mathbf{H}}{\partial t} \cdot \hat{\mathbf{t}} d\Phi \quad (2.127)$$

where C is the contour enclosing the surface Φ and $\hat{\mathbf{t}}$ is the versor normal to Φ . These two expressions represent the Ampere and Faraday laws respectively.

With reference to figure 2.16 one can choose: the contour C with two infinitesimal sides and the other two of length Δl enclosing the surface Φ having normal versor \mathbf{t} ; the volume V to be a cylinder of infinitesimal height and basis of surface Δa . Both C and V lie across the interface region between two media labeled with numbers 1 and 2, a surface charge σ and a surface current density \mathbf{K} can be present onto the separation surface. The nature of

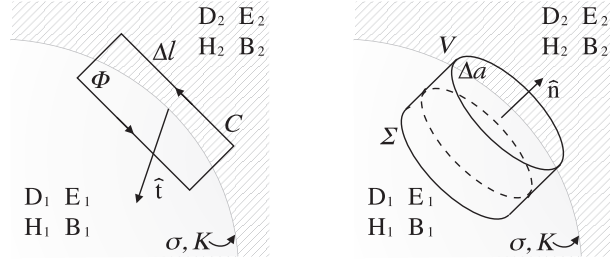


Figure 2.16: Volume, surfaces and contours for the divergence and Stokes' theorems application.

this charges and currents must be sought in the microscopic structure of the interface, in particular in the modification of the electronic structure of the media that occurs close to the separation surface. Neglecting the vanishing contribution from the infinitesimal lateral surface of the cylinder, the left hand side of equation (2.122) becomes

$$\oint_{\Sigma} \mathbf{D} \cdot \hat{\mathbf{n}} d\Sigma = (\mathbf{D}_2 - \mathbf{D}_1) \cdot \hat{\mathbf{n}} \Delta a \quad (2.128)$$

while the right hand side is:

$$4\pi \oint_V \eta_{\text{EXT}} d\mathbf{x} = 4\pi\sigma\Delta a \quad (2.129)$$

and equation (2.122) gives the condition:

$$(\mathbf{D}_2 - \mathbf{D}_1) \cdot \hat{\mathbf{n}} = 4\pi\sigma \quad (2.130)$$

If a surface charge σ is present at the interface, the component of the electric displacement \mathbf{D} perpendicular to the surface, is not conserved, it changes discontinuously moving from one medium to the other. Reasoning in the same way for the equation (2.123) one gets:

$$(\mathbf{H}_2 - \mathbf{H}_1) \cdot \hat{\mathbf{n}} = 0 \quad (2.131)$$

this condition states that, crossing the interface, the component of the magnetic field perpendicular to the surface must be the same. The same thing can be done for the other two equations. Neglecting the contribution from the the two infinitesimal sides of C , the left hand side of equation (2.127) becomes:

$$\oint_C \mathbf{E} \cdot \hat{\mathbf{t}} dl = (\hat{\mathbf{n}} \times \hat{\mathbf{t}}) \cdot (\mathbf{E}_2 - \mathbf{E}_1) \Delta l \quad (2.132)$$

while the right hand side vanish because the surface Φ is infinitesimal, and so:

$$(\mathbf{E}_2 - \mathbf{E}_1) \times \hat{\mathbf{n}} = 0 \quad (2.133)$$

Crossing the interface, the electric field component parallel to the separation surface must be the same. And finally, in the case of equation (2.126):

$$\oint_{\Phi} \left[\frac{4\pi}{c} \mathbf{J}_{\text{EXT}} + \frac{1}{c} \frac{\partial \mathbf{D}}{\partial t} \right] \cdot \hat{\mathbf{t}} d\Phi = \frac{4\pi}{c} \mathbf{K} \cdot \hat{\mathbf{t}} \Delta l \quad (2.134)$$

and one gets the last condition:

$$(\mathbf{H}_2 - \mathbf{H}_1) \times \hat{\mathbf{n}} = \frac{4\pi}{c} \mathbf{K} \quad (2.135)$$

In presence of a surface current, the magnetic displacement component parallel to the separation surface, is not expected to remain constant crossing the interface. For further details concerning optical and electronic properties of surfaces see references [49] and [50].

Appendix B: Perfectly reflecting 3D cavity

The aim of this appendix is to show that, in a three dimensional cavity with perfectly reflecting walls perpendicular to the z axis, two transverse modes with the same dispersion law exist for wave vector \mathbf{k} with $k_z \neq 0$ but, if \mathbf{k} lie in the xy plane, only one mode is allowed. The transverse modes of an empty cavity are solution of the D'Alembert's equation, in three dimensions there exist two linearly independent solutions for each \mathbf{k} vector, corresponding to the two polarizations:

$$\mathbf{E} = E_0 \hat{\mathbf{e}}_1 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} e^{ik_z z} \quad \mathbf{E} = E_0 \hat{\mathbf{e}}_2 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} e^{ik_z z} \quad (2.136)$$

here the scalar product $\mathbf{k} \cdot \mathbf{x}$ has been split to isolate the z component. $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are the polarization vectors defined as:

$$\hat{\mathbf{e}}_1 = \frac{\hat{\mathbf{n}} \times \mathbf{k}}{|\mathbf{k}|} = \frac{k_y}{|\mathbf{k}|} \hat{\mathbf{i}} - \frac{k_x}{|\mathbf{k}|} \hat{\mathbf{j}} \quad \hat{\mathbf{e}}_2 = \frac{\hat{\mathbf{e}}_1 \times \mathbf{k}}{|\mathbf{k}|} = \frac{k_x k_z}{|\mathbf{k}|^2} \hat{\mathbf{i}} + \frac{k_y k_z}{|\mathbf{k}|^2} \hat{\mathbf{j}} + \frac{k_x^2 + k_y^2}{|\mathbf{k}|^2} \hat{\mathbf{k}} \quad (2.137)$$

where $\hat{\mathbf{n}}$ is the versor normal to the cavity wall. Now it is convenient to switch between oscillating exponentials and trigonometric functions, a linear combination of solutions (2.136) must be used with wave vectors \mathbf{k} and \mathbf{k}' defined has:

$$\mathbf{k}' = k_x \hat{\mathbf{i}} + k_y \hat{\mathbf{j}} - k_z \hat{\mathbf{k}} \quad (2.138)$$

together with the Euler's formula. Naturally, changing \mathbf{k} the polarization versors change:

$$\hat{\mathbf{e}}_1' = \hat{\mathbf{e}}_1 \quad \hat{\mathbf{e}}_2' = -\frac{k_x k_z}{|\mathbf{k}|^2} \hat{\mathbf{i}} - \frac{k_y k_z}{|\mathbf{k}|^2} \hat{\mathbf{j}} + \frac{k_x^2 + k_y^2}{|\mathbf{k}|^2} \hat{\mathbf{k}} \quad (2.139)$$

Now, two linearly independent solution must be found, with the linear combination:

$$\mathbf{E} = E_0 (\hat{\mathbf{e}}_1 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} e^{ik_z z} - \hat{\mathbf{e}}_1' e^{i\mathbf{k}'_{\parallel} \cdot \mathbf{x}_{\parallel}} e^{ik_z z}) \quad (2.140)$$

one gets the following electric field:

$$E_x, E_y, E_z \propto E_0 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} \sin(k_z z) \quad (2.141)$$

or, with the other combination:

$$\mathbf{E} = E_0 (\hat{\mathbf{e}}_2 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} e^{ik_z z} + \hat{\mathbf{e}}_2' e^{i\mathbf{k}'_{\parallel} \cdot \mathbf{x}_{\parallel}} e^{ik_z z}) \quad (2.142)$$

the result is:

$$\begin{aligned} E_x, E_y &\propto E_0 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} \sin(k_z z) \\ E_z &\propto E_0 e^{i\mathbf{k}_{\parallel} \cdot \mathbf{x}_{\parallel}} \cos(k_z z) \end{aligned} \quad (2.143)$$

By virtue of the interface field conditions described in appendix A, the components E_x and E_y , parallel to the cavity walls, must be the same on both sides of the cavity walls, if the cavity walls are perfectly reflecting the field in the outer side is zero so it must be zero even on the inner one, as a result one can state the following boundary conditions:

$$E_x(x, y, 0) = E_y(x, y, 0) = 0 \quad E_x(x, y, \ell) = E_y(x, y, \ell) = 0 \quad (2.144)$$

For both the fields (2.141) and (2.143) these requirements gives the quantization condition $k_z = n\pi/\ell$, and the dispersion relation $\omega(k)$ is the same. When $n = 0$, $k_z = 0$, the wave vector is entirely contained in the xy plane: in the case of (2.141) this results in an everywhere vanishing field; for (2.143) the $n = 0$ condition gives a field uniquely oriented along the z direction. To summarize for $n \neq 0$ two modes with the same dispersion relation exist whereas for $n = 0$ a unique mode exist.

Bibliography

- [1] S.K. Lamoreaux. *Rep. Prog. Phys.*, 68:201, 2005.
- [2] G. Plunien, B. Mueller, and W. Greiner. *Phys.Rep.*, 134:87, 1986.
- [3] K.A. Milton. *J. Phys. A: Math. Gen.*, 37:R209, 2004.
- [4] M. Bordag, U. Mohideen, and V.M. Mostepanenko. *Phys. Rep.*, 353:1, 2001.
- [5] P.W. Milonni. *The quantum vacuum*. Academic Press, New York, 1994.
- [6] R. Eisenschits and F. London. *Z.Phys.*, 60:491, 1930.
- [7] F. London. *Z.Phys.*, 63:245, 1930.
- [8] F. London. *Z.Phys.Chem.B*, 11:222, 1930.
- [9] L.D. Landau and E.M. Lifshitz. *Quantum mechanics*. Mir, Mosca, 1979.
- [10] H.B.G. Casimir and D. Polder. *Phys.Rev.*, 73:360, 1948.
- [11] J. De Boer. *Trans. Faraday Soc.*, 32:10, 1936.
- [12] H. Hamaker. *Physica*, 4:1058, 1937.
- [13] J.E. Lennard-Jones. *Trans. Faraday Soc.*, 28:333, 1932.
- [14] J. Bardeen. *Phys.Rev.*, 58:717, 1940.
- [15] H.B.G. Casimir. *Proc. Kon. Ned. Akad. Wet.*, 51:793, 1948.
- [16] H.B.G. Casimir. *J.Chim.Phys.*, 46:407, 1949.
- [17] K.H. Schatten. *J.Franklin Inst.*, 755:326, 1989.
- [18] S. Scheel, L. Knoell, and D.-G. Welsh. *Phys. Rev. A*, 58:700, 1998.
- [19] B. Huttner, J.J. Baumberg, and S.M. Barnett. *Europhys.Lett.*, 16:177, 1998.
- [20] E.M. Lifshitz. *Sov. Phys. JEPT*, 2:73, 1956.
- [21] N.G. van Kampen, B.R.A. Nijboer, and K. Schram. *Phys. Lett.*, 26A:307, 1968.
- [22] E. Gerlach. *Phys. Rev. B*, 4:393, 1971.
- [23] K. Schram. *Phys.Lett. A*, 43:283, 1973.
- [24] I.E. Dzyaloshinskii and L.P. Pitaevskii. *Sov.Phys. JEPT*, 9:1282, 1959.
- [25] I.E. Dzyaloshinskii, E.M. Lifshitz, and L.P. Pitaevskii. *Sov.Phys. JEPT*, 10:161, 1960.
- [26] I.E. Dzyaloshinskii, E.M. Lifshitz, and L.P. Pitaevskii. *Adv. Phys.*, 10:165, 1958.

- [27] M. Antezza, L.P. Pitaevskii, S. Stringari, and V.B. Svetovoy. *Phys. Rev. A*, 77:022901, 2008.
- [28] Z. Lenac and M.S. Tomas. *Phys. Rev. A*, 78:023834, 2008.
- [29] D. Iannuzzi and F. Capasso. *Phys. Rev. Lett.*, 91:029101, 2003.
- [30] J.N. Munday and F. Capasso. *Phys. Rev. A*, 75:060102, 2007.
- [31] I. Pirozhenko, A. Lambrecht, and V.B. Svetovoy. *New J. Phys.*, 8:238, 2006.
- [32] V.B. Svetovoy, P.J. van Zwol, G. Palasantzas, and J.Th.M. De Hosson. *Phys.Rev.B*, 77:035439, 2008.
- [33] B.E. Sernelius. *Phys. Rev. B*, 71:235114, 2005.
- [34] N. Inui. *J. Phys. Soc. Japan*, 72:2198, 2003.
- [35] J.S. Høye, I. Brevik, J.B. Aarseth, and K.A. Milton. *Phys. Rev. E*, 67:056116, 2003.
- [36] V. A. Yampolskii, S. Savelev, Z.A. Mayselis, S.S. Apostolov, and F. Nori. *Phys. Rev. Lett.*, 101:096803, 2008.
- [37] B.E. Sernelius. *J. Phys. A: Math. Gen.*, 39:6741, 2006.
- [38] F. Zhou and L. Spruch. *Phys. Rev. A*, 52:297, 1995.
- [39] M.S. Tomaš. *Phys. Rev. A*, 66:052103, 2002.
- [40] C. Raabe, L. Knöll, and D.-G. Welsh. *Phys. Rev. A*, 68:033810, 2003.
- [41] B. Geyer, G.L. Klimchitskaya, and V.M. Mostepanenko. *Phys. Rev. A*, 67:062102, 2003.
- [42] M. Boström and B.E. Sernelius. *Phys. Rev. Lett.*, 84:4757, 2000.
- [43] I. Brevik, S.A. Ellingsen, and K.A. Milton. *New J. Phys.*, 8:236, 2006.
- [44] V.S. Bentsen, R. Herikstad, S. Skriudalen, I. Brevik, and J.S.Hoye. *J.Phys.A: Math.Gen.*, 38:9575, 2005.
- [45] P.J. van Zwol, G. Palasantzas, and J.Th.M. de Hosson. *Phys. Rev. B*, 77:075412, 2008.
- [46] H.B. Chan, Y. Bao, J. Zou, R. A. Cirelli, F. Klemens, W.M. Mansfield, and C.S. Pai. *Phys. Rev. Lett.*, 101:030401, 2008.
- [47] P.A. Maia Neto, A. Lambrecht, and S. Reynaud. *Phys. Rev. A*, 78:012115, 2008.
- [48] T. Emig, N. Graham, L. Jaffe, and M. Kardar. *Phys. Rev. Lett.*, 99:170403, 2007.
- [49] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, London, 1959.
- [50] J.D. Jackson. *Classical Electrodynamics*. John Wiley and Sons, 1975.

II

New developments and applications

3

Thin films stability

Contents

3.1	Introduction	104
3.2	Plasma model calculations	104
3.2.1	Force on isolated metallic films	105
3.2.2	The film-ideal metal substrate interaction	108
3.2.3	Bimetallic interfaces	111
3.3	Drude model calculations	114
3.3.1	Force on isolated metallic films	114
3.3.2	The film-ideal metal substrate interaction	118
3.3.3	The bimetallic interfaces	119
3.4	Stability of isolated films	119
3.4.1	Surfaces	121
3.4.2	Stability model of an isolated film	123
3.5	Stability of deposited films	125
3.5.1	Stability model for a deposited film	126
3.5.2	Results	129
3.6	Conclusions	135
	Appendix A: Basics on the elasticity theory	136
A.1	Strain and Stress tensors	136
A.2	Thermodynamics of deformations	139
A.3	Generalized Hooke's law	140
A.4	Homogeneous media and homogeneous deformations	141

In the first part of the chapter, a detailed investigation of the properties of the dispersion force acting onto films is given. Many configurations will be considered, such as an isolated film or a film deposited (grown) onto a substrate. The second part of the chapter is devoted to the description of the role that the dispersion force can have in determining the stability and the morphology of thin films. The work described here has been published in two papers, references [1, 2].

3.1 Introduction

Electromagnetic fluctuation induced forces have been the subject of several investigations both in the non retarded small distance limit (van der Waals forces) and in the retarded large distance case (Casimir forces) [3, 4, 5, 6]. Since the basic work from Lifshitz, described in section 2.3, studies have been focused mainly on the determination of the forces between two semi-infinite planar media [7] or between a sphere and a planar medium [8]. Even if model calculations have been performed for special geometries [5, 6, 9], forces on realistic systems have been obtained starting from the above mentioned configurations.

For several technological applications, multilayer systems, obtained by depositing thin films of different species onto a given substrate, are important and theoretical approaches have been devised to determine van der Waals forces between laminated media [10, 11]. Casimir forces between moving parts have been considered as possible source of instabilities in micro and nano-devices, where the components are in close proximity [12, 13, 14]. In many of these systems the situations of interest correspond to interaction between parallel interfaces of films and plates with micro or submicro-size and submicro-distances. While some of these studies have been performed using simplified models for the interaction, like the assumption that the interaction is correctly represented by the force between ideal metallic plates, it has been pointed out that a realistic description of adhesion or stiction phenomena has to account for the dependence of the force upon the shape and optical properties of the components [11, 15]. In spite of this large amount of work, a detailed study of the behaviour of the electromagnetic fluctuation induced forces in unsupported or deposited conductive films, as a function of their size and optical parameters has not been published. The interest has been focused mainly on the interaction between two-dimensional films, for which forces are supposed to show a peculiar dependence upon the film distance [16, 17, 18]. Less interest has been given to the study of the forces on the film boundaries due to vacuum fluctuations, which are present even in an isolated film and depend upon its size and properties.

One can formulate the problem as follows: suppose to have a simple metal, whose dielectric function can be expressed by the plasma model (1.107) or by the Drude model (1.111). If one considers an isolated metal film of thickness d , at $T = 0^\circ K$ it is known that, in the limit of vanishing plasma frequency, $\Omega_p \rightarrow 0$, the force on the film (the electromagnetic pressure on the film boundaries or the force between them) vanishes and the same happens in the limit of infinite plasma frequency (perfect metal case). The vanishing of the force is due to the peculiar values of film reflectivity in the two limits, it cannot stay identically zero for physical values of the reflectivity¹. The problem of what sort of behaviour has the force between these two limits has not been investigated: clearly, for a given film thickness, it must reach at least one maximum of intensity. The questions to be answered are: (i) what is the behaviour of the force as a function of Ω_p , in particular at which plasma frequencies are force maxima obtained, (ii) how do such maxima depend upon the film thickness, (iii) how is the behaviour of the force modified when the film is deposited onto a substrate, (iv) how does the electromagnetic force on the free standing film compare with the film-substrate interaction, that may be responsible of adhesion and stiction phenomena.

3.2 Plasma model calculations

The force acting on metallic films can be investigated through the Lifshitz approach described in chapter 2. In the following the plasma model dielectric functions will be used, in fact the aim here is not to achieve a precise description of the force for specific systems, since an

¹This can be easily understood by noticing that in the non-retarded regime the force on the free standing metal film is the same as the force between two semi-infinite plates of the same metal separated by a distance equal to the film thickness (see equations (3.3) and (3.4) in the text), which is obviously attractive and different from zero.

accurate evaluation with an intrinsic force uncertainty of few per cent requires a precise determination of the Drude parameters, which are very sensitive to the sample condition [19, 20]. Rather it is important to illustrate some general trends that can be understood using a model description of the dielectric properties.

Metallic films of thickness d ranging from $10nm$ to a few hundred nm will be considered. For typical metallic densities the electronic distribution deviates significantly from the bulk behaviour when the size of the film is less than about ten times the Fermi wavelength λ_F . Taking $\lambda_F \simeq 5 \text{ \AA}$, the bulk description is expected to become inaccurate when d is of the order of 50 \AA . For such ultrathin film quantum size effects are known to be important [21, 22, 23, 24, 25] and will be investigated in the chapter 4.

In the following calculations a local description of the electromagnetic properties of the metal based on a dielectric function $\epsilon(\omega)$ will be used, i.e. the wavenumber dependence of the dielectric response will be neglected (see section 1.5). This local theory is expected to be less accurate in thin films than for half space or bulk systems. Recent calculations have shown that non local corrections to electromagnetic induced forces for typical metallic densities are of the order of a few tenth of a per cent, suggesting that the local theory can be appropriate in the interpretation of the experimental data [26].

3.2.1 Force on isolated metallic films

The force per unit area F is calculated using expression (2.87) at the absolute temperature $T = 300^\circ K$. In the configuration of a standing alone film $\epsilon_1(\omega) = \epsilon_2(\omega) = 1$, therefore Q_{TM} and Q_{TE} are given by:

$$\begin{aligned} Q_{TM} &= 1 - e^{-2d\gamma_3} \frac{(\gamma_3 - \epsilon_3(i\Omega_n) \sqrt{k^2 + \frac{\Omega_n^2}{c^2} \epsilon_3(i\Omega_n)})^2}{(\epsilon_1 \gamma_3 + \epsilon_3(i\Omega_n) \sqrt{k^2 + \frac{\Omega_n^2}{c^2} \epsilon_3(i\Omega_n)})^2} \\ Q_{TE} &= 1 - e^{-2d\gamma_3} \frac{(\gamma_3 - \sqrt{k^2 + \frac{\Omega_n^2}{c^2} \epsilon_3(i\Omega_n)})^2}{(\gamma_3 + \sqrt{k^2 + \frac{\Omega_n^2}{c^2} \epsilon_3(i\Omega_n)})^2} \end{aligned} \quad (3.1)$$

where:

$$\gamma_3^2 = k^2 + \frac{\Omega_n^2}{c^2} \epsilon_3(i\Omega_n) \quad \epsilon_3 = 1 - \Omega_3^2/\omega^2 \quad (3.2)$$

Relaxation time effects will be neglected. Although they are important in determining the infrared response of metals and the following calculation can be extended to a complex dielectric function, the focus here is mainly on general trends in the behaviour of electromagnetic fluctuation induced forces. For realistic calculations on specific materials relaxation time effects have to be included and in metal films they can affect the force intensity. This effects will be investigated in section 3.3.

Notice that there is a basic difference between the electromagnetic fluctuation induced forces between two semi-infinite metals and those on the boundaries of a film. This is clearly seen if one considers an ideal (perfectly reflecting) metal, corresponding to an infinite plasma frequency: the interaction between two semi-infinite systems is expressed by the Casimir force (2.28). while for a film of finite thickness, the force on the boundaries vanishes.

Figure 3.1 displays the calculated behaviour of F as a function of the plasma frequency for films of different thickness ranging from 10 to $100nm$. Notice that at finite temperatures the force vanishes in the large plasma frequency limit, while for $\Omega_p \rightarrow 0$ there is a finite contribution from transverse magnetic modes. This contribution is due to the $m = 0$ term of the sum over Matsubara frequencies and it depends linearly upon T (see section 2.5). It is seen that the force is attractive (it tends to contract the film) and it shows a maximum and a tail at high plasma frequency. As expected from the general behavior of the electromagnetic

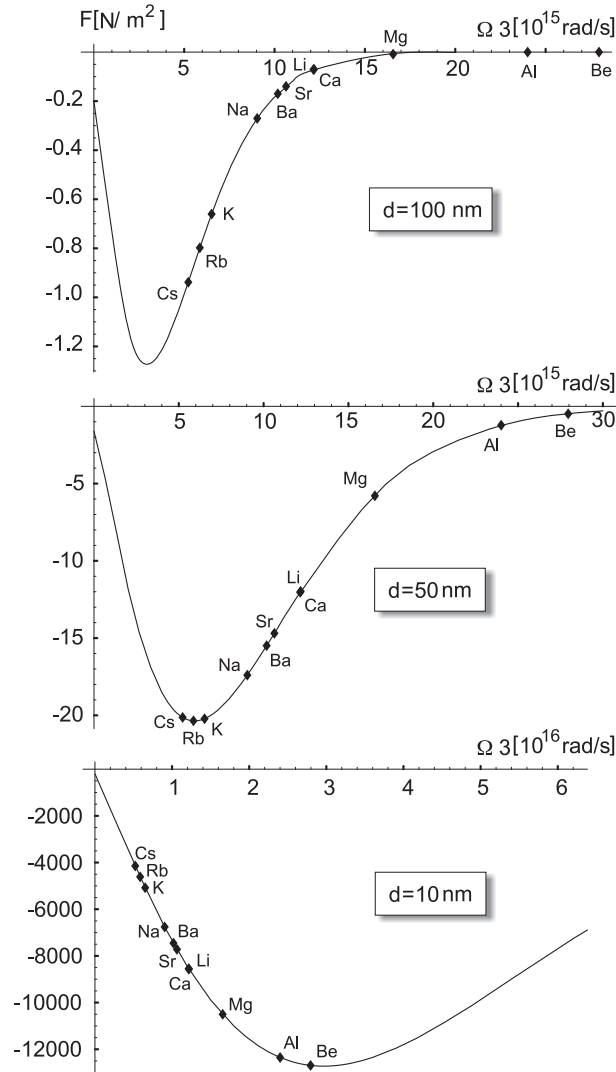


Figure 3.1: Force as a function of film plasma frequency: the force increases on decreasing the film thickness d . The calculation were performed summing the first thousand Matsubara frequencies. Alkali metals plasma frequencies are shown for comparison.

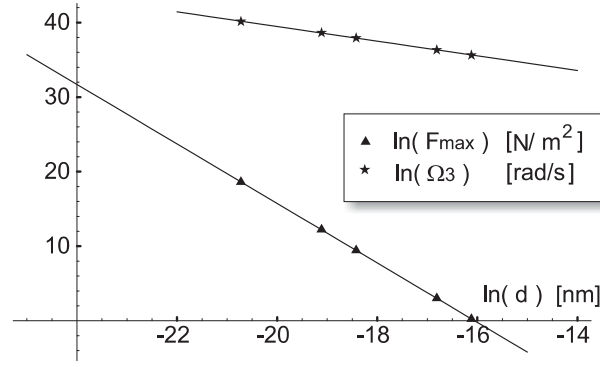


Figure 3.2: Maximum value of the force and plasma frequency at which it occurs, as a function of distance: the fitting functions are $-64.05 - 3.98\ln(d)$ (triangles) and $19.83 - 0.98\ln(d)$ (stars).

induced forces as a function of the distance, the maximum intensity reduces as a function of d , while its frequency moves to higher values. The dots in the figure correspond to the free electron plasma frequency for sp-bonded simple metals. This should not be seen as an accurate prediction of the force value for metals. It indicates only that the force on real metal films may fall on both sides of the maximum, depending upon the film thickness.

Notice that retardation effects are essential to obtain the maximum in the theoretical curve. This can be understood by a simple calculation of the force on a free standing metal film in the van der Waals regime at $T = 0^\circ K$ (see section 2.4). In this case equation (2.106) becomes:

$$F = -\frac{\hbar}{8\pi^2 d^3} \int_0^\infty \frac{(\epsilon_3(i\xi) - 1)^2}{(\epsilon_3(i\xi) + 1)^2} d\xi \quad (3.3)$$

which leads to

$$F = -\frac{\hbar\Omega_s}{32\pi d^3} \quad (3.4)$$

with $\Omega_s = \Omega_3/\sqrt{2}$ frequency of the surface plasmon that can be obtained imposing $\epsilon_3(\omega) = -1$, according to what has been found in section 2.3.1. Equation (3.4) does not show any maximum as a function of Ω_3 . This is not surprising since the above expression is valid under the condition that d is much smaller than the plasma wavelength, therefore is appropriate in the small plasma frequency regime only.

The behaviour of the maximum frequency as a function of d is given in figure 3.2: it is shown that in the range of thickness here considered, the maximum frequency falls like d^{-1} , while the intensity maximum falls as d^{-4} , as expected for the interaction in the retarded regime. The behaviour of the force maximum, that is displaced to larger values for smaller thicknesses, can be understood by noticing that the attraction arises from the interaction between the surface plasmons at the two film boundaries [27, 28, 29]. At a given film thickness the interaction is screened by the electron gas with increasing efficiency as the plasma frequency increases. For large electron density $\Omega_3 \rightarrow \infty$, the force goes to zero and one surface does not feel the presence of the other. The maximum in the force results from the balance between the surface plasmons interaction and the screening effects. In particular for small d a higher electron density is required to screen the attractive force.

Some interesting comments can be made on these data. The first concerns the unsupported film stability: the force tends to shrink the film and it has to be equilibrated by some repulsive interaction, most likely provided by the force built up by the valence electron rearrangement at the surfaces. Furthermore the force can be tuned significantly by changing the electron

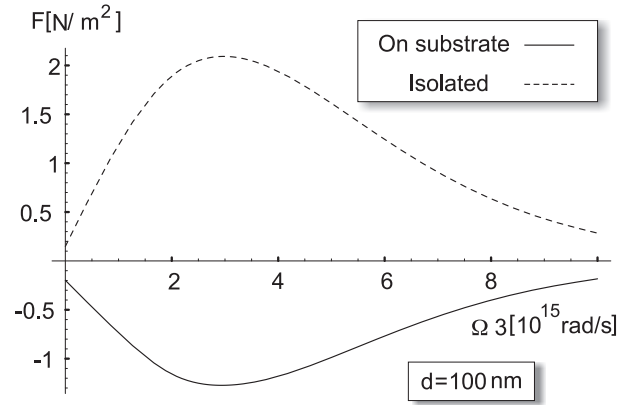


Figure 3.3: Force as a function of film plasma frequency: a change in sign occurs when the isolated metallic film is placed on a perfectly reflecting substrate (ideal metal).

density of the metal: this effect could be useful in engineering the film properties for specific applications.

3.2.2 The film-ideal metal substrate interaction

To show how the previous conclusions are modified when the metal film is interacting with a substrate, figure 3.3 displays the behaviour of the force per unit area on a film of $d = 100\text{nm}$ thickness deposited onto a perfectly reflecting substrate, (corresponding to the configuration with $\epsilon_2 = 1$ and ϵ_1 equal to infinity), as a function of film plasma frequency. This is a very simplified description of a bi-metallic interface, based on the assumption of the validity of the continuum model, that neglects all the details of the interactions between the atoms at the interface. It is expected to hold when the size of the film is large compared to the interface region (typically a few angstroms) so that the interface plays a minor role in determining the electromagnetic force. Notice that the force becomes repulsive and nearly double in intensity, although it shows the same qualitative behaviour with a maximum and a long asymmetric tail at large frequency values. It comes from the difference between the electromagnetic force per unit area on the substrate side and that on the vacuum side the behaviour of the force can be understood by noticing that at $T = 0^\circ\text{K}$ the exact calculation in the non-retarded limit gives the simple result:

$$F(d) = \frac{\hbar\Omega_p}{32\pi d^3}\sqrt{2} \quad (3.5)$$

showing the change of sign and the increased force value. This result is consistent with the behaviour of the London dispersion forces between dissimilar materials separated by a gap, that has been reported since many years [3, 30, 31]. In this case the force is known to be repulsive when $\epsilon_1 \geq \epsilon_2 \geq \epsilon_3$ and attractive when $\epsilon_1 \geq \epsilon_2 \leq \epsilon_3$ within a wide frequency range. It is interesting to understand how the force between film boundaries in a multilayer system is modified as a function of the film-substrate distance. For the case of a perfectly reflecting substrate, one can determine the range of distances over which the sign of the force changes. To this aim one has to extend equation (2.87) to a configuration with more than two planar media. In practice this amounts to replace the functions Q_{TM} and Q_{TE} by those appropriate to a multi-layer configuration. For a five layer system the appropriate expressions were derived by Zhou and Spruch and it is described in section 2.3.4. For the study of substrate-metal film interaction, in the formula (2.91) one has to take ϵ_4 equal to infinity, $\epsilon_1 = \epsilon_2 = \epsilon_5 = 1$ while ϵ_3 is the metallic film dielectric function (3.2). Since the configuration depends upon two parameters, the size d of the film and the film-substrate distance d_1 , one can define

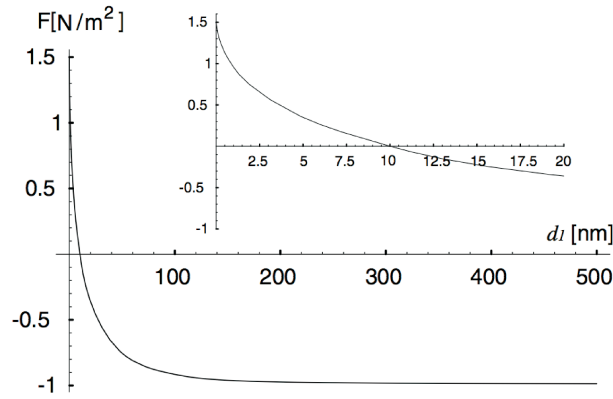


Figure 3.4: Force on the film boundaries as a function of the film-substrate distance, the film plasma frequency is $\Omega_3 = 5 \cdot 10^{15} \text{rad/s}$.

the force F between the film boundaries, given by the derivative of the free energy with respect to d , and the force F' , obtained by deriving the free energy with respect to d_1 , giving the interaction between the film and the substrate. Figure 3.4 shows the behaviour of the force F on the film boundaries as a function of the film-substrate distance for a 100nm film. It is seen that the force remains constant and attractive if the distance d_1 is larger than the film thickness d ; at lower distances the force decreases until it becomes repulsive. In other words, the film starts feeling a difference between the pressure from the metal substrate side and the external vacuum pressure, when the film-substrate distance is comparable with its thickness.

Discussions on device stability refer usually to the interaction between film and substrate (here the word substrate is used to indicate a structure of much larger size than the film, it could be a plate in a device), which gives rise to an attractive force F' . To show how this interaction behaves as a function of the ideal film-substrate distance, F' has been calculated using equations (2.91). It turns out to be attractive for any value of the film plasma frequency and, at distances smaller than the film size, it is considerably more intense than the force F on the film. This force is responsible of the change in sign observed in figure 3.3: if the film is close to the substrate, the difference between the attractive force on the film boundaries tends to stretch the film, causing a repulsive force between them.

The behaviour of the film-substrate force F' in the range of distances d_1 below the film thickness, where the substrate effect is more significant, is illustrated by the results shown in figure 3.7 for a 100nm film with plasma frequency $\Omega_3 = 5 \cdot 10^{15} \text{rad/s}$ and a perfectly reflecting substrate.

Notice that in this range of distances the force F' increases like d_1^{-x} with $3 < x < 4$, (the simple d_1^{-4} behaviour at all distances is characteristic of the interaction between ideal metal plates only and it is appropriate for real metals only at large distances). At 100nm distance this force is approximately -4.8N/m^2 , (the Casimir force between ideal metals at the same distance is of the order of -10N/m^2), while the force on the film boundaries is approximately -1N/m^2 . The gray curve in the figure displays the calculated force per unit area for a semi-infinite metal interacting with a perfectly reflecting semi-infinite substrate. It can be seen that it does not deviate significantly from the curve for the 100nm film. At higher distances the attractive force decreases while the force on the film remains approximately constant. The calculated F' for a 10nm film is reported in the same figure: in this case the force versus distance behaviour is rather different, showing a significantly higher exponent than in the 100nm case (3.52 rather than 3.29). Clearly this behaviour cannot be understood using arguments based on results for semi-infinite systems: for a semi infinite metal interacting with

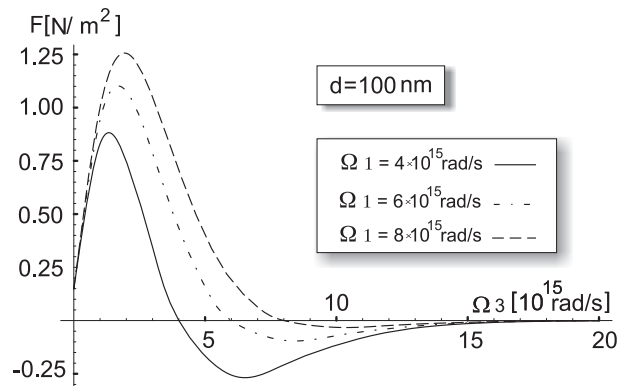


Figure 3.5: Force on the film boundaries as a function of film plasma frequency calculated for different substrate frequencies, in the calculation about two thousands Matsubara frequencies have been used.

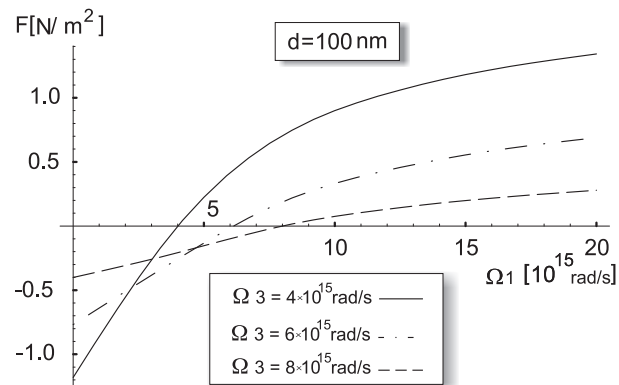


Figure 3.6: Force on the film boundaries as a function of substrate plasma frequency calculated for different film plasma frequencies, in the calculation about two thousands Matsubara frequencies have been used.

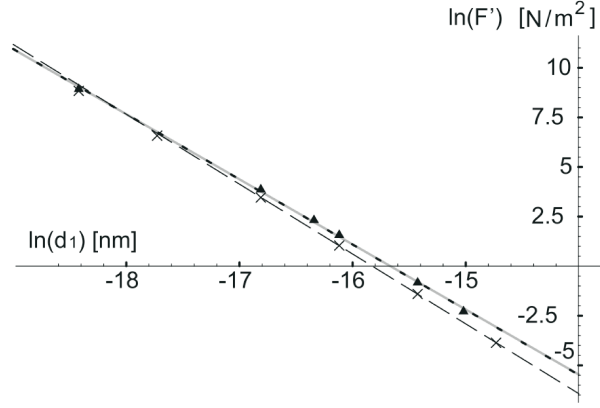


Figure 3.7: Film-substrate force as a function of the film-substrate distance between ideal metal substrate and real metal film, film thickness $100nm$ (triangles) and $10nm$ (crosses). The fitting functions are $-51.5 - 3.29\ln(d_1)$ (dotted line) and $-55.7 - 3.52\ln(d_1)$ (dashed line). The gray curve is the force between two semi-infinite systems, an ideal metal and a real metal with the same plasma frequency of the film.

an ideal substrate one would expect the exponent x to become closer to 3 upon decreasing the distance. The fact that it results to be significantly higher is a direct consequence of the finite thickness of the film. Indeed, as first pointed out by Zhou and Sprunch [32] higher negative exponents characterize the interaction in the presence of film of very small thickness. An important consequence of this behaviour is that the calculated F' at $10nm$ distance (approximately $-7511.7N/m^2$) is considerably higher than the force F on the film boundaries (approximately $0.001N/m^2$).

To conclude, the interaction of a metal film with a perfectly reflecting substrate leads to an attractive film-substrate force and, at short distances, to a repulsive force on the film boundaries. For $50 - 100nm$ thick films these forces are approximately of the same order when the film-substrate distance is comparable with the film size. In the low distance range ($1 - 10nm$) the force on the film can be neglected and the attractive film-substrate interaction prevails in intensity. These considerations are expected to be important for systems, like microswitches, that consist of two conducting electrodes, where one is fixed and the other one is able to move, being suspended by a mechanical spring. The stability of the system may depend upon the electromagnetic induced force acting on the mobile film [33, 34].

3.2.3 Bimetallic interfaces

The situation changes if one consider a more realistic description of the substrate. Referring to equation (2.87) this corresponds to take $\epsilon_1 = 1 - \Omega_1^2/\omega^2$. Figure 3.6 shows the behaviour of the force per unit area on a $100nm$ metal film deposited onto various metal substrates as a function of the substrate plasma frequency. Notice that the force is attractive when $\Omega_1 < \Omega_3$ and is repulsive in the opposite case. For $\Omega_1 \gg \Omega_3$ one gets the repulsive force corresponding to a perfectly reflecting substrate. The change in the sign it can be easily understood by considering the force in the small d limit, i.e in the non retarded regime. At $T = 0^\circ K$ the force calculated from equation (2.106) is simply given by

$$F = \frac{\hbar}{32\pi d^3} \frac{\Omega_s(\Omega_1^2 - \Omega_3^2)}{\bar{\Omega}(\bar{\Omega} + \Omega_s)} \quad (3.6)$$

where

$$\bar{\Omega} = \sqrt{(\Omega_1^2 + \Omega_3^2)/2} \quad (3.7)$$

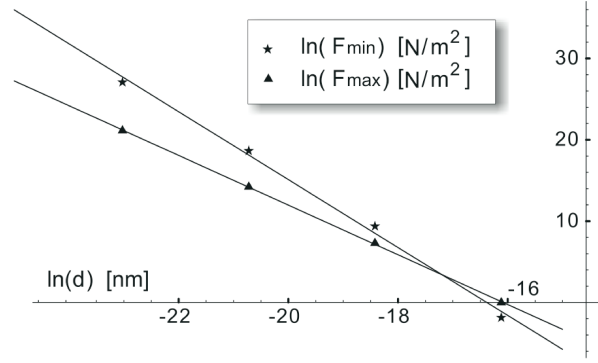


Figure 3.8: $F(\Omega_3)$ minimum and $F(\Omega_3)$ maximum as a function of distance d , the fitting functions are respectively $-68.36 - 4.17\ln(d)$ (stars) and $-49.10 - 3.05\ln(d)$ (triangles), calculation were performed at fixed $\Omega_1 = 5 \times 10^{15} \text{rad/s}$

is the interface plasmon frequency obtained from the relation $\epsilon_1(\omega) = -\epsilon_3(\omega)$, according to what has been found in section 2.3.1. Note that F shows the expected change from the repulsive to the attractive behaviour.

Figure 3.5 shows curves of the force on films deposited onto different substrates as a function of the film plasma frequencies. The curves show two extrema: on the repulsive side a maximum, that increases in intensity and moves to higher frequency upon increasing the substrate plasma frequency; on the attractive side a minimum which decreases upon increasing Ω_1 and shifts to higher frequency values. This behaviour is consistent with the previous conclusions concerning the ideal substrate: as the plasma frequency Ω_1 increases the repulsive force on the film becomes dominant.

It is interesting to see how the extrema behave upon varying the film thickness. As shown in figure 3.8, the intensity of the repulsive maximum falls like d^{-3} , in the range of distances here considered, while for the attractive minimum it falls approximatively as d^{-4} . Indeed the occurrence of the maximum can be understood on the basis of the short distance formula (3.6), which gives a d^{-3} dependence of the force, while the behaviour of the attractive part is mainly due to retarded interactions. These results lead to the conclusion that the electromagnetic fluctuation induced forces can give contribute of opposite sign, and with different dependence upon the film size, to the deposited film stability.

As in the case of the ideal substrate, one can study the electromagnetic fluctuation induced force F' between the film and the substrate as a function of the film-substrate distance. Based on the previous analysis one expects the film-substrate force to be attractive and to lead to a repulsive or attractive force between the film boundaries depending upon the difference between the plasma frequencies: if $\Omega_1 \gg \Omega_3$ the situation is similar to the ideal substrate case, while for $\Omega_1 \ll \Omega_3$ the force on the film is only weakly modified by the interaction. The various cases are illustrated in figure 3.9.

To see how these results depend upon the film thickness one has to compare the calculated curves for film-substrate force as a function of the distance d_1 with the electromagnetic induced bulk-bulk interaction. This comparison has been done in figure 3.10 for typical values of the plasma frequency. As shown the results seem not to depend significantly upon the film thickness for 100nm films, while size effects become important for d of the order of 10nm .

It is clear, from these calculations that, in the nanometric distances range, the adoption of the simple force expression appropriate to ideal plates is not correct. Both the sign and the intensity of the force may result wrong, if material properties and thickness effects are not

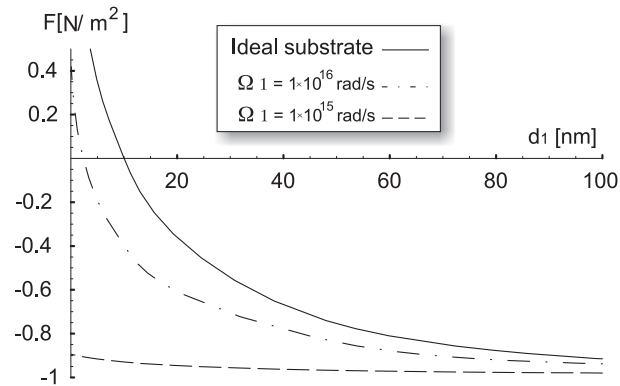


Figure 3.9: Film boundaries force as a function of film-substrate distance, comparison between an ideal substrate (continuous line), a real metal substrate with plasma frequency 10^{16} rad/s (dot-dash line) and 10^{15} rad/s (dashed line). The film plasma frequency is $5 \cdot 10^{15} \text{ rad/s}$.

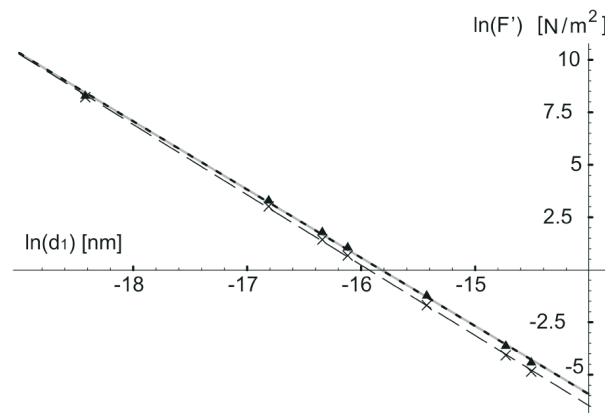


Figure 3.10: Film-substrate force as a function of the film-substrate distance for real metals with $5 \cdot 10^{15} \text{ rad/s}$, film thickness 100 nm (triangles) and 10 nm (crosses). The fitting functions are $-51.4 - 3.25 \ln(d_1)$ (dotted line) and $-53.3 - 3.35 \ln(d_1)$ (dashed line). The gray curve is the force between two semi-infinite bulks, two real metals with the same plasma frequency.

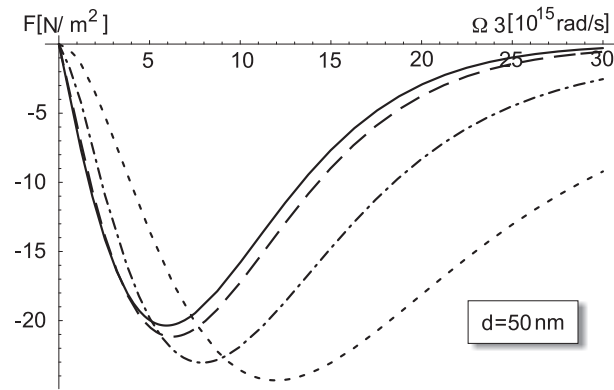


Figure 3.11: Free standing film. Force on film boundaries as a function of film plasma frequency for different dissipation coefficients. $\gamma = 0$ continuous line, $\gamma = 10^{14} \text{rad/sec}$ long dashed line, $\gamma = 10^{15} \text{rad/sec}$ dot dashed line, $\gamma = 5 \cdot 10^{15} \text{rad/sec}$ short dashed line.

properly accounted in the theory.

3.3 Drude model calculations

In the preceding section it has been shown that the force may change in sign and intensity depending upon the film size and the nature of the substrate. However the plasma model does not provide an accurate description of the real metal dielectric properties, since it neglects both the intraband absorption of the free electrons and the effects on the interband transitions in the relevant frequency range.

A more accurate description is given by the Drude model (1.111) which includes the parameters Ω_p , the plasma frequency, and γ the dissipation coefficient. For $\gamma = 0$, corresponding to infinite relaxation time, one recovers the plasma model. While in the plasma model Ω_p is fixed by the valence electron density, the Drude model allows to extract the parameters from the experimental data, thus taking into account both the real electron density and the effective mass of the electron in the metal under consideration. Recently a careful study of the fitting procedure for Au films, focused on the frequency range that is relevant in the determination of the forces due to vacuum electromagnetic field fluctuations, has shown that the Drude model can be useful to reproduce optical experimental data for films prepared in various experimental conditions [20], thus allowing for a more accurate evaluation of the force.

3.3.1 Force on isolated metallic films

It has been shown in the previous sections that a free standing metal film is subject to a negative pressure i.e. an attractive force per unit area F between the boundaries caused by zero point oscillations of the quantized electromagnetic field. The force per unit area F is calculated at zero temperature using expression (2.73). One can easily verify that also the Drude model leads to a vanishing force in the two limits $\Omega_3 \rightarrow 0$ (infinitely dilute metal) and $\Omega_3 \rightarrow \infty$ (ideal metal). Since the force cannot be identically zero for finite plasma frequency values, extrema are expected between these two limits.

Figure 3.11 displays the calculated force versus plasma frequency curves for a 50 nm film. γ has been chosen comparable with those obtained by fitting optical properties and from sample metal resistivity data [35]. For comparison the curve obtained using the plasma model is given in the same figure. It is seen that the general shape of the curve is the same:

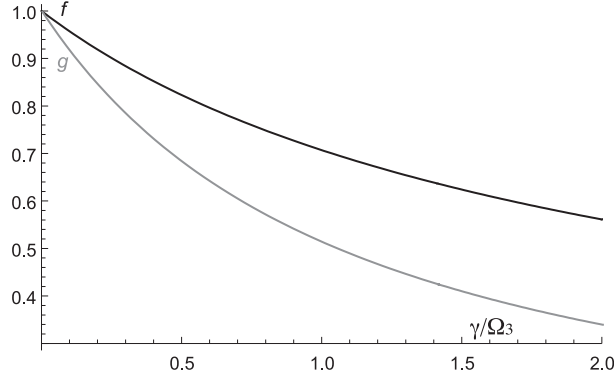


Figure 3.12: g factor of equation (3.10), gray line; f factor of equation (3.13), black line.

the force is attractive and shows a maximum of intensity (a minimum in the curve) and a long tail at high frequencies. However the inclusion of a finite dissipation coefficient leads to significant modifications in the force intensity both at low and high frequency and causes a shift in the position of the maximum.

In particular notice that, on increasing the dissipation coefficient, the force decrease at low Ω_3 while it becomes significantly higher at high plasma frequency. The weakening of the force in the low plasma frequency range, can be understood using the small distance approximation (2.106) to the Lifshitz formula, that can be applied when the film thickness d is much smaller than both the plasma wave length $\lambda_p = 2\pi c/\Omega_3$ and the relaxation wavelength $\lambda_\tau = 2\pi c/\gamma$. In such conditions the force is due to TM modes only and is simply given by:

$$F(d) = -\frac{\hbar}{8\pi^2 d^3} \int_0^\infty \frac{(\epsilon_3(i\xi) - 1)^2}{(\epsilon_3(i\xi) + 1)^2} d\xi \quad (3.8)$$

showing the typical d^{-3} behaviour of the van der Waals dispersion forces. Inserting the Drude dielectric function into this equation one gets:

$$F = F_{P1} g\left(\frac{\gamma}{\Omega_3}\right) \quad (3.9)$$

where

$$g(x) = \frac{\sqrt{2}}{\pi} \left[\frac{2(1-x^2)-2}{x(1-x^2)} - \frac{1}{(1-x^2)^{3/2}} \left[\text{ArcTan}\left(\frac{x}{\sqrt{2-x^2}}\right) - \frac{\pi}{2} \right] \right] \quad (3.10)$$

and

$$F_{P1} = -\frac{\hbar\Omega_s}{32\pi d^3} \quad (3.11)$$

is the force calculated for $\gamma = 0$ (plasma model) in the low thickness regime. $\Omega_s = \Omega_3/\sqrt{2}$ is the frequency of the free electron surface plasmon. Here $\gamma < \sqrt{2}\Omega_3$ is assumed, a condition that is commonly verified in real metals. The plot of the g function, presented in figure 3.12 for a large interval of x values, shows that the correction to the force calculated in the ideal plasma model is always negative and gets more significant on increasing γ .

At high Ω_3 values one do not have a simple formula like equation (3.9) that allows to calculate the force analytically. The plasma wavelength is small and the condition $d \ll \lambda_p$ is not satisfied. However it is possible to understand the reason of the increase in the force intensity compared with the plasma model using the following simple argument. The vanishing of the

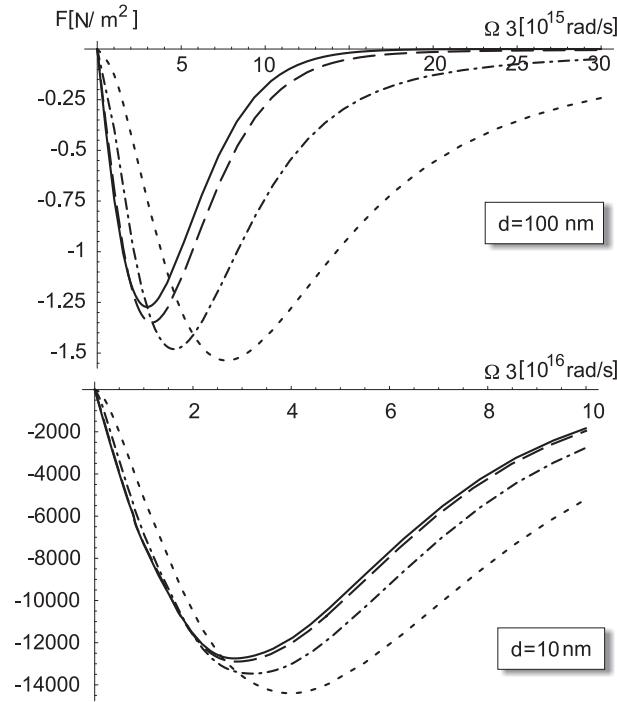


Figure 3.13: Free standing film. Force on film boundaries as a function of film plasma frequency for different dissipation coefficients. The notation and the γ values are the same of figure 3.11

force as Ω_3 goes to infinity is due to the exponential factor $e^{-2\xi K_3 d/c}$ in Q_{TM} and Q_{TE} (see equations (2.66)), since the fraction that multiplies the exponential goes to one both in the plasma and Drude model (this corresponds to the condition of reflectivity equal to one). One can write the exponential as $e^{-2\gamma_3 d}$, for high Ω_3 values one has $\gamma_3^2 \simeq \xi^2 \frac{\Omega_3}{\xi^2 + \xi\gamma}$, which shows that for the plasma model the exponential goes to zero faster than for finite γ values. This implies that both Q_{TM} and Q_{TE} go to one more slowly as γ increases, leading to a higher intensity of the force in the Drude model for high Ω_3 values. As previously observed the occurrence of a maximum is a direct consequence of the retarded nature of the interaction. Indeed formula (3.8), which is obtained by neglecting retardation effects, does not give any maximum as a function of Ω_3 .

Figure 3.13 shows the calculated force per unit area in films of different thickness. The shape of the curves and the behaviour as a function of γ are basically the same of figure 3.11 and follow the same trend that has been reported in the previous section: the force decreases in intensity and the maximum shifts at lower frequencies on increasing d . Again notice that the inclusion of the dissipation coefficient can cause a significant enhancement of the force in the high Ω_3 regime. Figure 3.14 displays the plot of the force as a function of d for films of different plasma frequency. It is seen that the force, for relatively low plasma values (figure 3.14 (a)) behaves as d^{-3} over a large thickness range. Significant deviations from this behaviour appear at very high dissipation coefficients only. At high plasma frequency (figure 3.14 (b)) different inverse power laws are present in the same interval of thicknesses, in agreement with previous findings.

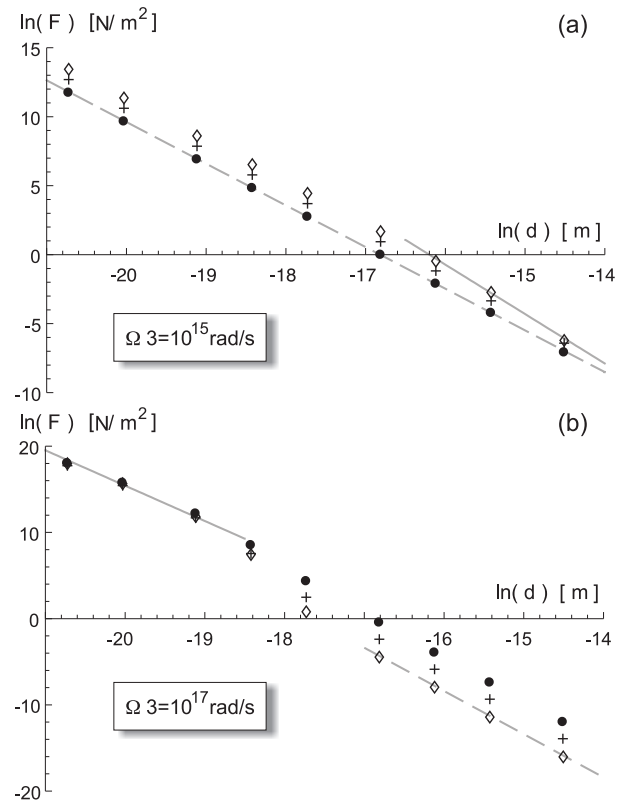


Figure 3.14: Free standing film. Force as a function of the thickness for different plasma frequency and dissipation coefficient values. $\gamma = 10^{14}$ rad/sec diamonds, $\gamma = 10^{15}$ rad/sec crosses, $\gamma = 5 \cdot 10^{15}$ rad/sec dots. Continuous line of figure (a) is $\sim 1/d^{3.6}$, dashed line of figure (a) is $\sim 1/d^3$, continuous line of figure (b) is $\sim 1/d^4$ and dashed line of figure (b) is $\sim 1/d^5$.

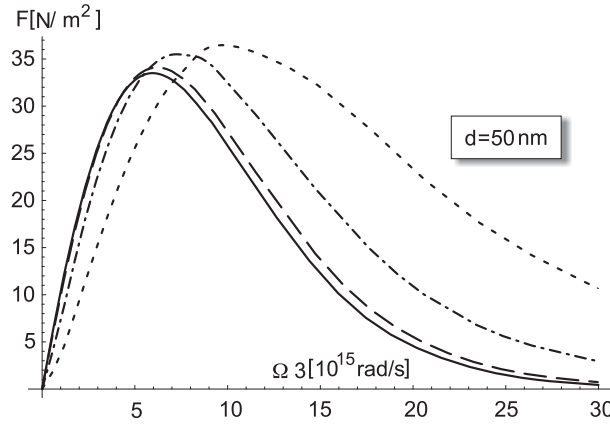


Figure 3.15: Film deposited on perfectly reflecting substrate. Force on film boundaries as a function of film plasma frequency for different dissipation coefficients. γ values are the same of figure 3.11

3.3.2 The film-ideal metal substrate interaction

The situation can be significantly modified if one consider the case of deposited films. Using the plasma model it has shown that a metal deposited onto a perfectly reflecting substrate is subject to a positive pressure between the boundaries, that tends to increase the film thickness. This force turns out to be approximately double in intensity compared to the force on the free standing film.

Figure 3.15 reports the results of the calculation of the force acting on a 50nm film deposited on a perfectly reflecting substrate as a function of Ω_3 for different γ values. For comparison, the curve corresponding to the plasma model is also given. One can notice that the force is repulsive and its intensity is nearly doubled in comparison with the isolated film case. However the shape of the curves is basically the same: the force shows a maximum and a long tail at high frequencies. With respect to the plasma model, the inclusion of relaxation leads to an increase of the force in the high Ω_3 region and a decrease at low frequencies. This behaviour, illustrated in figure 3.15, can be understood using the same arguments presented in the previous section. In particular, notice that in the small d regime the force is simply given by:

$$F = F_{P2} f\left(\frac{\gamma}{\Omega_3}\right) \quad (3.12)$$

where

$$f(x) = \frac{1}{\sqrt{1-x^2/2}} \left[1 - \frac{2}{\pi} \text{ArcTan}\left(\frac{x}{\sqrt{2-x^2}}\right) \right] \quad (3.13)$$

and

$$F_{P2} = 2 \frac{\hbar \Omega_s}{32\pi d^3} = -2F_{P1} \quad (3.14)$$

is the force calculated for the plasma model. The behaviour of the $f(x)$ factor, reported in figure 3.12, shows that the force is weakened compared to the plasma case. The strengthening of the force at high Ω_3 is basically a consequence of the slower decay of the exponential in equations (2.66), as in the case of the free standing film.

The d^{-3} dependence is characteristic of the small d regime only and it is appropriate for small plasma frequencies. As Ω_3 increase this approximation does not hold any more and, as in the case of the free standing film, different inverse power dependence are needed to represent the force behaviour.

3.3.3 The bimetallic interfaces

A more realistic description is achieved when the substrate dielectric function is described by a Drude model. Figure 3.16 displays the calculated force per unit area for films deposited onto a metallic substrate of fixed plasma frequency Ω_1 as a function of Ω_3 for different combinations of dissipation coefficient values. The thickness of the film is $50nm$. The curves are typical of the behaviour that can be obtained from the calculations for this kind of systems. The behaviour is similar to the previously reported results on plasma model. The force is repulsive in a limited range of frequencies and becomes negative when Ω_1 is smaller than the film plasma frequencies. This is strictly true when the dissipation coefficients of the substrate and the film are the same. If the substrate relaxation frequency is lower than the one of the film, the range of frequencies that corresponds to a repulsive force is larger and the force intensity increases, while in the opposite case the force curve crosses the horizontal axis at lower Ω_3 values. To understand these findings notice that the force calculated in the van der Waals, small d , regime, when the dissipation coefficient is the same for the film and the substrate is given by:

$$F(d) = \frac{\hbar}{16\pi^2 d^3} \frac{\Omega_3^2(\Omega_1^2 - \Omega_3^2)}{\Omega_s^2 - \Omega^2} \left[\frac{1}{\sqrt{4\bar{\Omega}^2 - \gamma^2}} \left(\frac{\pi}{2} - \text{ArcTan}\left(\frac{\gamma}{\sqrt{4\bar{\Omega}^2 - \gamma^2}}\right) \right) - \frac{1}{\sqrt{4\Omega_s^2 - \gamma^2}} \left(\frac{\pi}{2} - \text{ArcTan}\left(\frac{\gamma}{\sqrt{4\Omega_s^2 - \gamma^2}}\right) \right) \right] \quad (3.15)$$

where $\Omega_s = \Omega_3/\sqrt{2}$ and $\bar{\Omega} = \sqrt{(\Omega_1^2 + \Omega_3^2)/2}$. Clearly this approximation predicts a repulsive force if $\Omega_3 < \Omega_1$ and an attractive one if $\Omega_3 > \Omega_1$ in agreement with the result of the full calculation. This conclusion appears to be quite general and independent of the particular model of dielectric function adopted to describe substrate and film dielectric functions. It is consistent with the behaviour of the London dispersion force between dissimilar materials that has been reported in the literature, mainly with reference to ceramic or dielectric materials [30]. It has to be noticed that the force may change even by an order of magnitude when the dissipation coefficient of the film is modified. This frequency is expected to be different in a film compared to the one measured in the bulk optical properties, due to the contribution of surface scattering and defects. This indicates that the use of the dielectric function of the bulk metal to represent the optical properties of the film may lead to serious errors in the determination of the vacuum force. Figure 3.17 illustrates the behaviour of the force as a function of the substrate plasma frequency for a film with $\Omega_3 = 10^{16}rad/s$, when the dissipation coefficient is modified. The curves allow to understand how the force of the film is affected by gradually changing the properties of the substrate. These results like those of figure 3.16 are quite general since similar curves can be obtained by changing the film thickness and the other parameters. The examples reported in the figure show that the shape and intensity of the force can be significantly modified by varying the optical properties of film and/or substrate. It should be observed that in the real situation of film deposition, the optical properties of the film may be very sensitive to the experimental conditions: both Ω_3 and γ can vary significantly, well above the experimental errors, for films of the same metal prepared by different experimental techniques [20].

To discuss the stability problem notice that, unlike the perfect metal substrate case illustrated in the previous section, deposition onto a real metal does not necessarily give a repulsive force.

3.4 Stability of isolated films

The problem of the stability of a free standing film is commonly discussed in terms of competition between surface energy and surface stress and the volume strain energy [36, 37, 38]. As

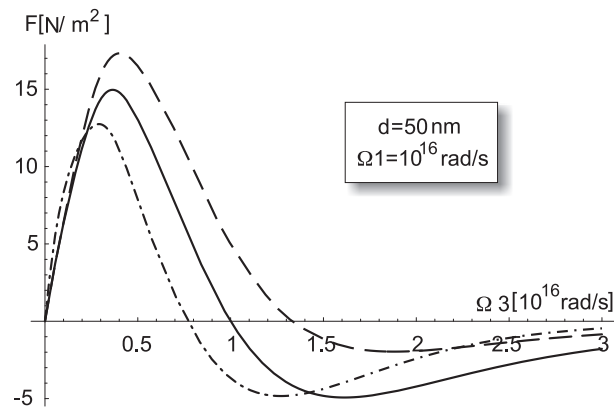


Figure 3.16: Film deposited onto a Drude substrate. Continuous line, $\gamma_3 = \gamma_1 = 10^{15} \text{ rad/s}$; dashed line, $\gamma_3 = 10^{15} \text{ rad/s}$ and $\gamma_1 = 10^{14} \text{ rad/s}$; dash dot $\gamma_3 = 10^{14} \text{ rad/s}$ and $\gamma_1 = 10^{15} \text{ rad/s}$.

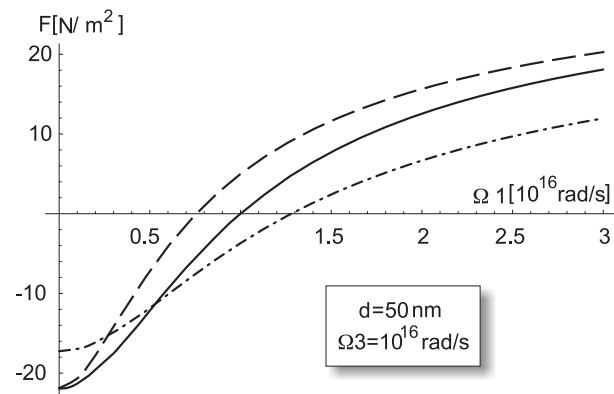


Figure 3.17: Film deposited onto a Drude substrate. The notation and the γ values are the same of figure 3.16

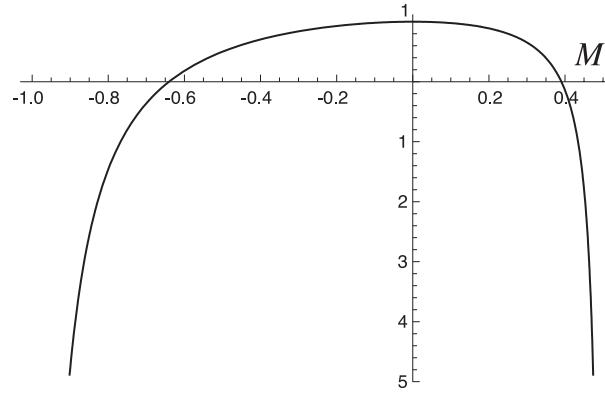


Figure 3.18: Plot of the second factor of inequality (3.34) as a function of M , by definition $-1 \leq M \leq 1/2$

discussed in the next section, surface stresses can arise as a consequence of the different local environment of surface atoms with respect to those in the bulk: if they are tensile stresses they can originate a compressive stress within the film. If the compression is sufficiently large it can cause buckling of the surface with changes in the lattice parameters of the outermost of atomic planes. A brief introduction to elasticity theory is given in appendix A.

3.4.1 Surfaces

In nature atoms aggregates together to compensate their need of stability, a bounded atom is more stable than a free one. An atom in the bulk of a defect free material is supposed to be as much bounded as possible, being surrounded by the maximum number of other atoms. The creation of a surface requires certain atoms to be surrounded by less atoms than in the bulk. An atom that constitutes the surface of a body is less stable than a bulk one, it follows that a surface is somehow unstable and an energy must be paid to keep it open. In a more formal way one should say that a surface stress exists that tends to minimize the total surface of a body and that some other force must exist in order to compensate the surface stress. This force is provided by the bulk atoms that, being in an energetically favourable configuration, work against any compression or stretch caused by the surface forces to change the surface morphology.

As a simple system to play with, in order to understand the surface effects, one can consider a thin film standing alone in the vacuum. In this case one deals with two identical flat surfaces infinitely extended along the x and y direction and separated by a distance d along the z direction. The stability condition for such a system can be worked out through the elasticity theory: the surface stress must cause some deformation on the film, this gives rise to an elastic restoring force. In the equilibrium condition the elastic force must be equal to the surface one. In the simplest situation the film deformation is homegeneous and no shear deformation exist: $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{\parallel}$, $\epsilon_{zz} = \epsilon_{\perp}$ and $\epsilon_{ij} = 0 \quad \forall \quad i \neq j$. Defining Γ_0 as the surface energy density, the total energy of the two surfaces of the film is given by:

$$F^S = 2\Gamma_0 A_0 \quad (3.16)$$

where A_0 is the film surface and Γ_0 is a measurable quantity tabulated for various materials and various possible surface orientations. If the surface stress produces a deformation on the film, also the surface energy changes, this is due both to the change in the total surface:

$$A = A_0(1 + \epsilon_{xx})(1 + \epsilon_{yy}) = A_0(1 + \epsilon_{\parallel})^2 \quad (3.17)$$

and to the change of the energy density Γ :

$$\begin{aligned}\Gamma &= \Gamma_0 + \frac{d\Gamma}{d\epsilon_{xx}}\epsilon_{xx} + \frac{d\Gamma}{d\epsilon_{yy}}\epsilon_{yy} + \frac{1}{2} \left(\frac{d^2\Gamma}{d\epsilon_{xx}^2}\epsilon_{xx}^2 + \frac{d^2\Gamma}{d\epsilon_{yy}^2}\epsilon_{yy}^2 + 2\frac{d^2\Gamma}{d\epsilon_{xx}d\epsilon_{yy}}\epsilon_{xx}\epsilon_{yy} \right) = \\ &= \Gamma_0 + 2\frac{d\Gamma}{d\epsilon_{\parallel}}\epsilon_{\parallel} + 2\frac{d^2\Gamma}{d\epsilon_{\parallel}^2}\epsilon_{\parallel}^2\end{aligned}\quad (3.18)$$

The deformed surface energy is therefore:

$$F^S = 2A_0(\Gamma_0 + 2f\epsilon_{\parallel} + 4\alpha\epsilon_{\parallel}^2) \quad (3.19)$$

where:

$$f = \Gamma_0 + \frac{d\Gamma}{d\epsilon_{\parallel}} \quad \alpha = \frac{\Gamma_0}{4} + \frac{d\Gamma}{d\epsilon_{\parallel}} + \frac{1}{2}\frac{d^2\Gamma}{d\epsilon_{\parallel}^2} \quad (3.20)$$

On the other hand, the elastic energy stored in the film to compensate the surface modifications is given by:

$$F^B = FV = FA_0d_0(1 + \epsilon_{\parallel})^2(1 + \epsilon_{\perp}) \simeq FA_0d_0 \quad (3.21)$$

where F is the bulk free energy per unit volume defined by (3.94) in appendix B and $V = Ad$ is the volume of the deformed film. In the last step only second order terms in the strain have been retained. Replacing the F expression one gets:

$$\begin{aligned}F^B &= \frac{A_0d_0}{2} \left[(\lambda + 2\mu) \left(2\epsilon_{\parallel}^2 + \epsilon_{\perp}^2 \right) + 2\lambda \left(\epsilon_{\parallel}^2 + 2\epsilon_{\parallel}\epsilon_{\perp} \right) \right] = \\ &= A_0d_0\mu \left[2\epsilon_{\parallel}^2 + \epsilon_{\perp}^2 + \frac{\lambda}{2\mu} \left(\epsilon_{\perp} + 2\epsilon_{\parallel} \right)^2 \right]\end{aligned}\quad (3.22)$$

or, in terms of the Poisson's coefficient M and Young's modulus E :

$$F^B = \frac{A_0d_0}{2} \frac{E}{1+M} \left[2\epsilon_{\parallel}^2 + \epsilon_{\perp}^2 + \frac{M}{1-2M} \left(\epsilon_{\perp} + 2\epsilon_{\parallel} \right)^2 \right] \quad (3.23)$$

To find the equilibrium conditions for the deformed film one must search for the minima of the total energy $F^{tot} = F^S + F^B$:

$$\begin{cases} \frac{\partial F^S}{\partial \epsilon_{\parallel}} + \frac{\partial F^B}{\partial \epsilon_{\parallel}} = 0 \\ \frac{\partial F^S}{\partial \epsilon_{\perp}} + \frac{\partial F^B}{\partial \epsilon_{\perp}} = 0 \end{cases} \quad (3.24)$$

from the second equation one gets a relation between the strain components:

$$\epsilon_{\perp} = -\frac{2M}{1-M}\epsilon_{\parallel} \quad (3.25)$$

whereas from the first equation one gets the condition:

$$2f + 8\alpha\epsilon_{\parallel} = -\frac{Ed}{1-M}\epsilon_{\parallel} \quad (3.26)$$

retaining only the linear terms in the strain ($\alpha = 0$) one obtains the simple condition:

$$2f = -\frac{Ed}{1-M}\epsilon_{\parallel} \quad (3.27)$$

using now equation (3.100) to calculate $\sigma_{xx} = \sigma_{yy} = \sigma_{\parallel}$ one obtains

$$\sigma_{\parallel} = K \left(2\epsilon_{\parallel} + \epsilon_{\perp} \right) + \frac{2}{3}\mu \left(\epsilon_{\parallel} - \epsilon_{\perp} \right) \quad (3.28)$$

but using (3.25):

$$\sigma_{\parallel} = \frac{E}{1-M}\epsilon_{\parallel} \quad (3.29)$$

so the condition (3.27) becomes

$$f = -\frac{d}{2}\sigma_{\parallel} \quad (3.30)$$

for the two surfaces to be in equilibrium, the surface stress must be $-d/2$ the bulk stress. With the first derivative one is able to find the extrema of the energy, to be sure of dealing with a minimum the second derivative sign must be investigated. In particular the extremum stability requires the determinant of the hessian to have positive eigenvalues, with a 2×2 matrix this is achieved when both the trace and the determinant are positive².

$$\begin{aligned} & \frac{\partial^2 F}{\partial \epsilon_{\parallel}^2} + \frac{\partial^2 F}{\partial \epsilon_{\perp}^2} > 0 \\ & \begin{vmatrix} \frac{\partial^2 F}{\partial \epsilon_{\parallel}^2} & \frac{\partial^2 F}{\partial \epsilon_{\parallel} \partial \epsilon_{\perp}} \\ \frac{\partial^2 F}{\partial \epsilon_{\perp} \partial \epsilon_{\parallel}} & \frac{\partial^2 F}{\partial \epsilon_{\perp}^2} \end{vmatrix} = \frac{\partial^2 F}{\partial \epsilon_{\parallel}^2} \times \frac{\partial^2 F}{\partial \epsilon_{\perp}^2} - \left(\frac{\partial^2 F}{\partial \epsilon_{\parallel} \partial \epsilon_{\perp}} \right)^2 > 0 \end{aligned} \quad (3.31)$$

From the first condition on gets:

$$8\alpha(1-M) + Ed > 0 \quad (3.32)$$

and defining the critical value α_c one gets:

$$\alpha > \alpha_c = \frac{Ed}{8(1-M)} \quad (3.33)$$

From the second condition (3.31) one gets:

$$\alpha > \alpha_c \frac{1-M-4M^2}{1-M-2M^2} \quad (3.34)$$

the plot of the second factor of the RHS is given in figure 3.18 for the allowed M values. From the figure it is seen that the factor is always smaller than 1, it follows that condition (3.34) is automatically fulfilled once that (3.33) is fulfilled.

A deeper analysis of the elastic properties of surfaces and thin films can be found in references [37, 39].

3.4.2 Stability model of an isolated film

Treating the film as a continuum one can perform the analysis of the principal strains under the assumption that the film is in a state of biaxial stress [36]. For an homogeneous system like the one treated in section 3.4.1, taking the z -axis normal to the surface of the film, one has $\sigma_{xx} = \sigma_{yy} = \sigma_{\parallel}$ and $\sigma_{xy} = \sigma_{zx} = \sigma_{yz} = \sigma_{zz} = 0$ and, neglecting quadratic terms, the equilibrium strains are given by:

$$\epsilon_{xx} = \epsilon_{yy} = \epsilon_{\parallel} = \sigma_{\parallel} \frac{1-\nu}{Y} \quad \epsilon_{zz} = \epsilon_{\perp} = -2\sigma_{\parallel} \frac{\nu}{Y} \quad (3.35)$$

²If a matrix is diagonal the trace is the sum of its eigenvalues whereas the determinant is their product. Trace and determinant are invariant scalars so, even if the matrix is not in its diagonal form, the trace and determinant values continue to represent the sum and the product of the eigenvalues.

where E is the Young modulus and M the Poisson's ratio of the metal. In terms of the surface stress f

$$f = \Gamma + \frac{\partial \Gamma}{\partial \epsilon_{\parallel}} \quad (3.36)$$

the equilibrium condition requires

$$f = -\frac{d}{2}\sigma_{\parallel} \quad (3.37)$$

so that one can write the strains as

$$\epsilon_{\parallel} = -2f \frac{1-M}{Ed} \quad \epsilon_{\perp} = 4f \frac{M}{Ed} \quad (3.38)$$

here d is the film thickness and Γ is the surface energy per unit area. To include the vacuum fluctuation forces, one has to carry out the same analysis adding to the surface and bulk elastic energy, the vacuum electromagnetic energy of the film. As a first approximation one can think that the vacuum energy is almost constant, i.e. is not affected by the film deformation. With this assumption, the condition (3.30) derived in section 3.4.1 is still valid, and using (3.101) one has simply to include the extra stress (force per unit surface) coming from the vacuum pressure $F(d)$ described by the very general expression (2.73). In this way one is led to

$$\epsilon_{\parallel} = -2f \frac{1-M}{Ed} - M \frac{F(d)}{E} \quad \epsilon_{\perp} = 4f \frac{M}{Ed} + \frac{F(d)}{E} \quad (3.39)$$

As expected by a linear approximation, the strain induced by the surface stress and the vacuum electromagnetic force simply add to determine the strain condition of the system.

In a more general way one should remember that the dispersion force depends on the surface and the thickness of the film and that it contains the film dielectric properties, dependent on the film electron density. The film deformation produces a strain of the film along the directions parallel and perpendicular to the surface, this results in a modification of the renormalized dispersion energy $E(d)_{ren.}$ described in (2.68):

$$F^V = E(A, d)_{ren.} = E(A_0(1 + \epsilon_{\parallel})^2, d_0(1 + \epsilon_{\perp}))_{ren.} \quad (3.40)$$

here A and d are the surface area and thickness of the deformed film, whereas A_0 and d_0 are the ones of the undeformed film. In a more explicit form:

$$\begin{aligned} E(d)_{ren.} &= \frac{\hbar A_0 (1 + \epsilon_{\parallel})^2}{4\pi^2} \int_0^{\infty} k dk \int_0^{\infty} [\ln(Q_{TM}(i\xi)) + \ln(Q_{TE}(i\xi))] d\xi \\ Q_{TM} &= 1 - e^{-2d(1+\epsilon_{\perp})\gamma_3} \frac{(\gamma_3 - \epsilon_3\gamma)^2}{(\gamma_3 + \epsilon_3\gamma)^2} \quad Q_{TE} = 1 - e^{-2d(1+\epsilon_{\perp})\gamma_3} \frac{(\gamma_3 - \gamma)^2}{(\gamma_3 + \gamma)^2} \\ \gamma^2 &= k^2 + \frac{\omega^2}{c^2} \quad \gamma_3^2 = k^2 + \frac{\omega^2}{c^2} \epsilon_3 \quad \epsilon_3 = 1 - \frac{\Omega_p^2}{\omega^2} \frac{1}{(1 + \epsilon_{\parallel})^2(1 + \epsilon_{\perp})} \end{aligned} \quad (3.41)$$

Notice that also the dielectric functions are affected by the film deformation: through the plasma frequency, they depend upon the electron density, that can be modified by the change in volume induced by the film strain. One can perform the same calculation presented in section 3.4.1 considering now a total free energy given by the summation of the surface contribution, the bulk contribution and the vacuum energy contribution:

$$F^{tot} = F^S + F^B + F^V \quad (3.42)$$

of course the set of equations (3.24) is much more complicated but the final expressions will allow one to account for the changes in dispersion interaction due to the film deformation,

giving a more accurate model. Typical calculated values of the surface stress are of the order of $1 \div 5 N/m$ [38]. Taking typical values for the elastic parameters one can estimate the strain values for $d = 10 nm$ to be of the order of $10^{-3} \div 10^{-4}$. The vacuum force contributions has opposite sign, but the order of magnitude is different. For $d = 10 nm$ the maximum value of $F(d)$, as shown in figure 3.13, is of the order of $10^4 N/m^2$, leading to strains of the order of $10^{-5} \div 10^{-6}$, which are too small to compete with the surface stress effects. Therefore, under normal conditions, the electromagnetic force does not contribute significantly to the free standing film morphology. For thicknesses of the order of few monolayers, the situation is less clear, since the electromagnetic force is expected to increase at least as d^{-3} . However in such cases a dielectric continuum theory is not adequate to represent the film optical properties, since size quantization effects become important. Their inclusion leads to a very different behaviour of the dielectric function and to an expression of the force where the derivative of the film dielectric function with respect to d has to be considered (see chapter 4).

3.5 Stability of deposited films

Epitaxial deposition of thin metal films on various substrates is of crucial importance in several applications ranging from the design of systems for heterogeneous catalysis and electron emission to the technology of semiconductor devices and magnetic recording [40, 41, 42, 43]. In several cases a significant mismatch between the lattice parameters of the deposit and the substrate is present. If the mismatch is not too large, the overlayer grows in a way that the atoms of the deposit are in registry with the substrate atomic structure and the growth is pseudomorphic. As the film thickness increases, mechanical strain energy builds up in the deposited film, which eventually can cause instability unfavorable to uniform flat film growth: the formation of dislocations or the continuation of growth in the form of islands become energetically favorable. In several cases epitaxially deposited film grown layer by layer can undergo a transition to a growth mode of three dimensional dislocation-free islands, which form on top of a layer of a certain critical thickness [44, 45, 46, 47, 48]. It turns out that the critical thickness is material dependent and decreases with increasing the lattice mismatch strain [49, 50].

The structure and morphology of the deposited films is often determined by kinetic constraints rather than by the minimization of the free energy. However it is possible to work in experimental conditions that allow the system to relax toward the equilibrium configuration. In such cases equilibrium considerations may be useful to understand the physical interactions that are responsible of the film evolution and to establish the conditions for the existence of a critical thickness and its value.

One can discuss the stability issue using a continuum model [51, 52, 53, 54, 55, 39]. The value of this approach lies in the fact that one can draw qualitative and some quantitative conclusions on general aspects of the growth without the need of a microscopical description of the atomistic mechanisms present in the process. Within this model the stability of epitaxial films results primarily from the competition between the surface energy and the elastic energy, caused by lattice mismatch, since when a flat film surface is modified into a wavy shape the surface energy increases, but the elastic energy decreases. Roughening of the surface and island formation can occur when it is energetically favorable to relax the elastic energy by increasing the surface area. It has been observed that since the surface energy tends to stabilize perturbations at low wavelength, while the elastic energy amplifies the perturbation at all the wavelength, any strained film, regardless of its thickness, should be unstable. This prediction does not match with the experiments which show the existence of a critical thickness beyond which the instability occurs [49]. This suggests that some additional force has to be present that competes with elastic stress and stabilizes the film within

a range of thickness of the order of few nanometers. Long range forces, in particular the vacuum fluctuation forces between the interface and the film free surface, have been considered as a possible source of stability in the system [49, 56]. Due to their peculiar size dependence, these forces are expected to play a role only at nanometric distances and when they cause a repulsion between the films boundaries.

In this section, the conditions under which the vacuum fluctuation forces can contribute to deposited metal film stability, will be discussed. The purpose is to identify, both for the film and for the substrate, the range of optical parameters, that can determine the existence of a finite critical thickness below which the flat surface growth takes place.

3.5.1 Stability model for a deposited film

Consider the system sketched in figure 3.19 where a film of thickness d is deposited on an elastic substrate thick compared to the film. Assume that, under suitable conditions, the film can be grown as a single crystal on the substrate surface. If the film lattice constant is identical to the one of the substrate and the substrate-film interface is perfectly ordered, than the film will grow with zero macroscopic strain. On the other hand, in case of lattice mismatch, and still assuming a perfectly ordered interface, the film will grow continuing the substrate structure and will be strained in the plane of the interface, figure 3.19 (b). The film will tend to relax this elastic energy through dislocation motion or by changing the shape of the surface through mass transport, figure 3.19 (c). In the following analysis the second mechanism will be taken into account. It is generally accepted that at $T = 0^\circ K$ it is the competition between the elastic energy and the surface energy that determines the morphological stability of deposited film: indeed when a flat film surface is perturbed into a wave shape, as indicated in the figure 3.19 (a), elastic energy decreases, but surface energy increases. Surface roughening and island formation are expected to occur as a consequence of this competition, when the flat surface cannot be stable [51, 52, 53, 54, 57].

However, when the size of the epitaxial film is nanometric, as in micro- nano-scale devices, other forces come into play. Long range dispersion forces, in particular the vacuum fluctuation force between the interface and the film free surface, have been considered as a possible source of stability in the system. The analysis is carried out for a sinusoidal surface considered as a perturbation of a reference perfectly planar surface, it is based on perturbation theory and on the assumption that the elastic constant of the film and the substrate are identical and in the absence of a stabilizing influence of the surface stress. For the two dimensional model sketched in figure 3.19 (a), where x is the surface position and $z = 0$ is the reference plane and the sinusoidally wavy shape is described by a cosine function of amplitude q and wavelength λ :

$$z = d - q \cos\left(\frac{2\pi}{\lambda}x\right) \quad (3.43)$$

one can show [57] that the stress concentration (stress for unit area and unit length normal to the surface) is related to the stress concentration σ for the planar surface configuration by the equations

$$\begin{aligned} \sigma'_{xx}(x, z) &= \sigma \left[1 + \frac{4\pi q}{\lambda} \left(1 - \frac{\pi d}{\lambda} \right) e^{-2\pi d/\lambda} \cos\left(\frac{2\pi}{\lambda}x\right) \right] \\ \sigma'_{zz}(x, z) &= \sigma \frac{4\pi d q}{\lambda^2} e^{-2\pi d/\lambda} \cos\left(\frac{2\pi}{\lambda}x\right) \\ \sigma'_{xz}(x, z) &= \sigma \frac{2\pi q}{\lambda} \left(1 - \frac{2\pi d}{\lambda} \right) e^{-2\pi d/\lambda} \sin\left(\frac{2\pi}{\lambda}x\right) \end{aligned} \quad (3.44)$$

which shows that at the surface the stress concentration is minimum in the peaks of the surface wave and maximum in the valleys and that in deeper locations away from the surface

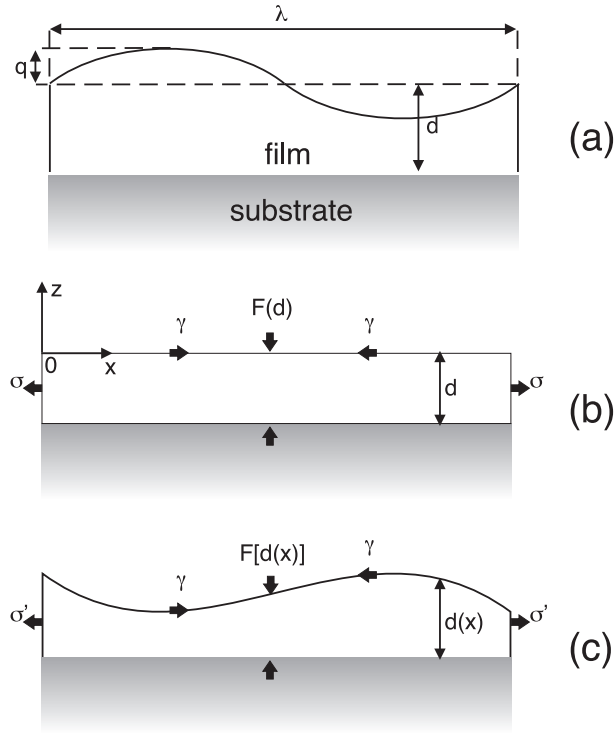


Figure 3.19: Geometric parameters for the wavy film model and different possible stable configurations.

the stress is attenuated exponentially with a characteristic length $\lambda/2\pi$. Assuming that $d \ll \lambda$ one has, to the first order, that only the tangential stress contributes to the elastic energy

$$\sigma'(x) = \sigma \left[1 + \frac{4\pi q}{\lambda} \cos\left(\frac{2\pi}{\lambda}x\right) \right] \quad (3.45)$$

Now the elastic strain energy density is given by definition (3.106):

$$F(x, q) = \frac{\sigma^2}{2Y} \left[1 + \frac{8\pi q}{\lambda} \cos\left(\frac{2\pi}{\lambda}x\right) \right] \quad (3.46)$$

where Y is the Young modulus and where the second order terms in q have been neglected. Following references [49, 58] one can calculate the elastic energy stored in the deformed film as

$$F(x, q) = \frac{dF^B(x, q)}{dzdx} \rightarrow dF^B = F(x, q) \frac{dz}{dq} dq dx \quad (3.47)$$

$$\Delta F^B = \int_0^\lambda \int_0^q F(x, q) \frac{dz}{dQ} dQ dx = \frac{\sigma^2}{2E} \int_0^\lambda \int_0^q \left[-\cos\left(\frac{2\pi}{\lambda}x\right) - \frac{8\pi Q}{\lambda} \cos^2\left(\frac{2\pi}{\lambda}x\right) \right] dQ dx$$

$$\Delta F^B = -\sigma^2 q^2 \pi / Y \quad (3.48)$$

Notice that Δf^B does not depend upon the wavelength and, being negative, gives rise to film instability for any kind of harmonic perturbation of the surface.

The change in the surface energy density can be determined by multiplying the surface tension Γ by the change in the surface length, according to definition (3.16):

$$\Delta F^S = \Gamma \Delta \ell \quad (3.49)$$

in the limit $q \ll \lambda$ the variation of the surface length is given by:

$$\Delta \ell = \int_0^\lambda \sqrt{1 + \left(\frac{dz}{dx}\right)^2} dx - \lambda \simeq \int_0^\lambda \left[1 + \frac{1}{2} \left(\frac{2\pi q}{\lambda}\right)^2 \sin^2\left(\frac{2\pi}{\lambda}x\right)\right] dx - \lambda \quad (3.50)$$

$$\Delta F^S = \Gamma q^2 \pi^2 / \lambda \quad (3.51)$$

showing that the surface energy increases as the surface is perturbed harmonically. The comparison shows that, independently on the wave amplitude and the film thickness d , the flat film surface becomes unstable when:

$$\begin{aligned} \Delta F^B + \Delta F^S &< 0 \\ \lambda &> \pi Y \Gamma / \sigma^2 \end{aligned} \quad (3.52)$$

the second member of this disequation giving a critical length below which the system is stable under a wavy perturbation. Inclusion of the vacuum fluctuation force leads to an additional contribution in the energy density that can be calculated to second order in the surface modifications under the assumption of slightly undulated surface $q \ll d$ and $q \ll \lambda$:

$$\Delta F^V = \int_0^\lambda E(z)_{ren.} dx - E(d)_{ren.} \lambda \quad (3.53)$$

but for small q :

$$\begin{aligned} E(z)_{ren.} &= E \left[d - q \cos\left(\frac{2\pi}{\lambda}x\right) \right]_{ren.} = \\ &= E(d)_{ren.} - \frac{\partial E(d)_{ren.}}{\partial d} \cos\left(\frac{2\pi}{\lambda}x\right) q + \frac{q^2}{2} \cos^2\left(\frac{2\pi}{\lambda}x\right) \frac{\partial^2 E(d)_{ren.}}{\partial d^2} + \dots \end{aligned} \quad (3.54)$$

and substituting

$$\Delta F^V = \frac{\partial^2 E(d)_{ren.}}{\partial d^2} \Big|_d \frac{q^2 \lambda}{4} \quad (3.55)$$

Adding this term to the previous ones and imposing the equilibrium condition that the total energy change be negative

$$\begin{aligned} \Delta F^B + \Delta F^S + \Delta F^V &< 0 \\ \lambda^2 \frac{\partial^2 E(d)_{ren.}}{\partial d^2} Y - \lambda 4\pi \sigma^2 + 4\pi^2 \Gamma Y &< 0 \\ \lambda_{\pm} &= 2\pi \frac{\sigma^2 \pm \sqrt{\sigma^4 - \Gamma Y^2 \frac{\partial^2 E(d)_{ren.}}{\partial d^2}}}{\frac{\partial^2 E(d)_{ren.}}{\partial d^2} Y} \end{aligned} \quad (3.56)$$

one arrives at the following conclusions:

- in order to have at least two identical roots, the second derivative has to be positive, corresponding to a repulsive force on the film;

- the film is stable under wavy perturbations provided its thickness d is lower than the critical thickness d_c defined through the relation

$$\left. \frac{\partial^2 E(d)_{ren.}}{\partial d^2} \right|_{d_c} = \sigma^4 / \Gamma Y^2 \quad (3.57)$$

- For $d > d_c$ there always exist two values λ_+ and λ_- and for any $\lambda \in [\lambda_-, \lambda_+]$ the condition (3.56) is fulfilled. The value:

$$\lambda^* = \frac{\lambda_+ + \lambda_-}{2} = \frac{2\pi\sigma^2}{Y} \left(\frac{\partial^2 E(d)_{ren.}}{\partial d^2} \right)^{-1} \quad (3.58)$$

represents the minimum of the quadratic equation (3.56), i.e. the wavelength that maximizes the energy gain of the perturbation.

For film of very small size, one can assume a van der Waals expression of the vacuum energy

$$\Delta F^V = -Hq^2\lambda/8\pi d^4 \quad (3.59)$$

where H is the Hamaker constant. One obtains the conditions

$$d_c = \left(\frac{-HY^2\Gamma}{2\pi\sigma^4} \right)^{1/4} \quad \lambda^* = -\frac{4\pi^2\sigma^2 d^4}{HY} \quad (3.60)$$

which shows that the film can be stable only for negative Hamaker constant (repulsive force) and it is inversely proportional to the stress concentration [49, 50].

Extension of the theory to three dimensions does not modify appreciably the stability condition if a single wave perturbation is considered. In particular equation (3.57) is still valid provided that one replaces Y with $Y/(1-\nu^2)$, where ν is the film Poisson's ratio [39].

3.5.2 Results

The issue of film stability is currently discussed using empirical forms of the interaction between the film boundaries based on the small d approximation [56]. The critical thicknesses that have been reported in the study of metallic growth are of the order of few nanometers. In this condition, and provided Ω_3 is not too large, the use of the small d approximation, which corresponds to the van der Waals force description, is quite reasonable. To make reference to the previous work, the results obtained under this approximation will be presented first, to be compared with those obtained including retardation effects later on.

In the case of a perfectly reflecting substrate, the energy second derivative appearing in the condition (3.57) turns out to be

$$\frac{\partial^2 E(d)_{ren.}}{\partial d^2} = \frac{3\sqrt{2}\hbar\Omega_3}{32\pi d^4} f\left(\frac{\gamma}{\Omega_3}\right) \quad (3.61)$$

Figure 3.20 presents curves of $\partial^2 E(d)_{ren.}/\partial d^2$ as a function of Ω_3 for a $6nm$ film at different dissipation coefficients in the small d limit. The plot allows to determine the range of Ω_3 where a flat surface is stable, corresponding to the values of $\partial^2 E(d)_{ren.}/\partial d^2$ above the horizontal line given by the second member of equation (3.57), which has been calculated taking typical values of the parameters representing the elastic and surface properties of the film. The values $\sigma = 500MPa$, $\Gamma = 1J/m^2$, $Y = 76GPa$, reported in [49], have been used. Changing these values simply shifts the horizontal line in the figure. The point where the second derivative curves cross the horizontal line is the minimum plasma frequency that allows for the flat

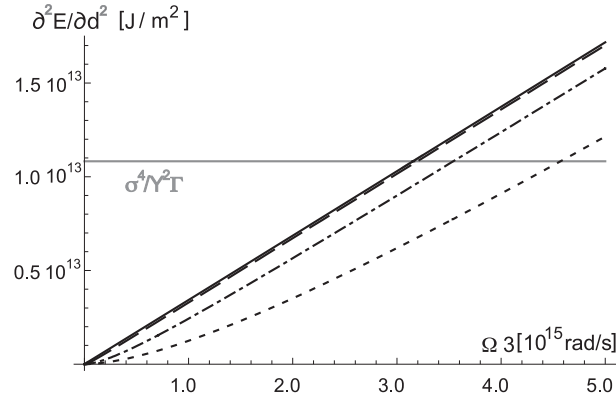


Figure 3.20: Vacuum energy second derivative plot for the Drude film deposited onto perfectly reflecting substrate. The notation and the γ values are the same of figure 3.11. The film thickness is $d = 6\text{nm}$.

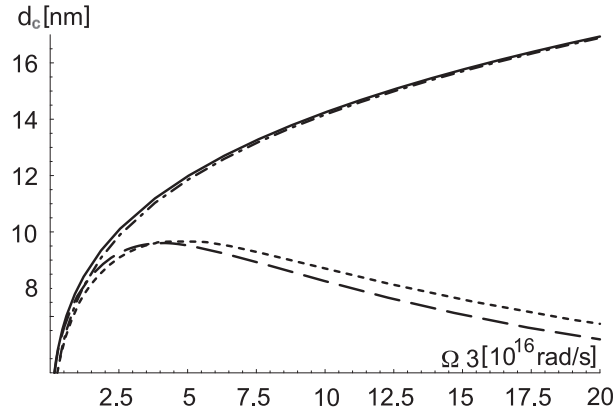


Figure 3.21: Critical thickness for a 6nm film deposited onto a perfectly reflecting substrate. The continuous and dot dashed lines represent respectively the plasma film and Drude films in the small d approximation, $\gamma = 5 \cdot 10^{15} \text{ rad/s}$. The long dashed and the dashed lines represent again the plasma and Drude films respectively, but they include also retardation effects.

configuration to be stable. It is seen that, given the film thickness, flat surface stability is possible above a threshold value. On passing from the plasma model, where $\partial^2 E(d)_{ren.} / \partial d^2$ depends linearly on Ω_3 , to the more realistic Drude description the stability threshold moves to higher Ω_3 values. This indicates that, even adopting a simplified model of the substrate, the conditions of stability depend critically upon the description of the film optical properties. It is common practice to evaluate the vacuum energy contribution using the van der Waals energy. As mentioned before, this is a good approximation only when the film size d is much less than the plasma wavelength. If this condition is not verified, one obtains a rather inaccurate estimate of the critical thickness. To give an example, figure 3.21 displays the d_c values calculated with the full theory, and those obtained with the small d approximation. The comparison shows the inadequacy of the van der Waals description in the determination of film stability for a large interval of Ω_3 values. Notice that both theories predict the existence of a finite d_c at any Ω_3 value, this result is peculiar and is a consequence of the perfect reflectivity of the substrate. Relaxation effects play a minor role and lead to differences in the critical values, more significant at large Ω_3 .

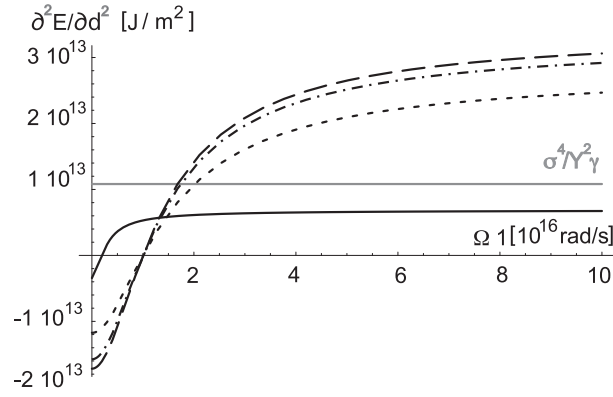


Figure 3.22: Van der Waals energy second derivative plot. In the case of plasma model the plasma frequency is $\Omega_3 = 2 \cdot 10^{15} \text{rad/s}$, continuous line and $\Omega_3 = 10^{16} \text{rad/s}$, long dashed line. Short dashed line and dot dashed line have been obtained using the Drude model with $\Omega_3 = 10^{16} \text{rad/s}$ and $\gamma_3 = 10^{15} \text{rad/s}$, $\gamma_1 = 10^{14} \text{rad/s}$ or $\gamma_3 = 10^{14} \text{rad/s}$, $\gamma_1 = 10^{15} \text{rad/s}$ respectively. The film thickness is $d = 6 \text{nm}$.

To illustrate the importance of a realistic description of the optical properties of the substrate, the outcomes of calculations using the plasma model will be discussed first, neglecting the relaxation effects for both the film and the substrate. Figure 3.22 illustrates the behaviour of $\partial^2 E(d)_{ren.} / \partial d^2$ as a function of Ω_1 for films of different plasma frequency and 6nm thickness. For $\Omega_3 = 2 \cdot 10^{15} \text{rad/s}$ no stability region is achieved since the curve never crosses the horizontal line. This conclusion is obviously dependent on the parameters chosen to represent the elastic and surface properties. On the other hand for $\Omega_3 = 10^{16} \text{rad/s}$ one finds a threshold value above which the flat film is stable. The difference is caused by the fact that for small Ω_3 values even a perfect metal does not provide a force repulsive enough to overcome the instability due to the elastic stress. Indeed for $\Omega_1 \rightarrow \infty$ the second derivative of the energy for deposition on a perfect metal in the small d approximation is simply given by:

$$\frac{\partial^2 E(d)_{ren.}}{\partial d^2} = \frac{3\sqrt{2}\hbar\Omega_3}{32\pi d^4} \quad (3.62)$$

if this quantity is less than $\frac{\sigma^4}{\gamma^2}$ one can never achieve a stability region.

In the plasma model the only relevant parameters are the plasma frequencies of the film and the substrate. It is useful to represent the results in the form of a stability diagram in the $\Omega_3\Omega_1$ plane, like those given in figure 3.23. Each curve in the diagram is obtained at fixed film thickness as the locus of the couples $\Omega_3\Omega_1$ which satisfy the equilibrium condition (3.57). The plane is separated into two areas corresponding to flat surface and rough surface condition respectively. Notice that upon increasing the film thickness the domain of $\Omega_3\Omega_1$ values that allows for the existence of a flat surface gets narrower. Figure 3.24 reports the calculated critical thickness as a function of Ω_3 for a fixed substrate. Notice that, unlike the perfect metal case, there is a finite range of film plasma values for which the flat surface growth can occur. As expected this range reduces as Ω_1 decreases. The maximum value of the critical thickness is in correspondence with the maximum intensity of the repulsive force. Turning to the more realistic Drude model one notices in the first place that the values of the second derivative curve are systematically lower than those obtained in the plasma model, as clearly shown in figure 3.22. Upon increasing the dissipation coefficient the crossing point shifts to higher Ω_3 values. Since γ depends primarily on the quality of the film, this result suggests that the stability condition can be significantly modified by the surface and defect scattering processes that take place in the film. The stability diagrams plotted in figure 3.25

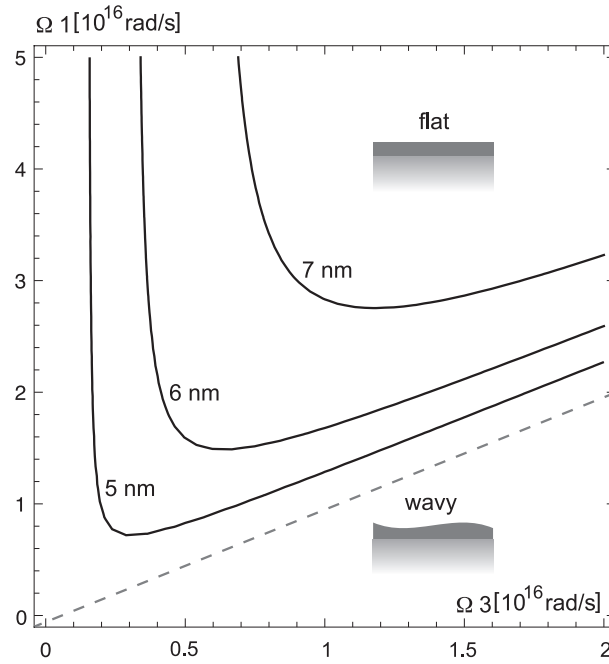


Figure 3.23: Stability diagram for the plasma film deposited onto a plasma substrate, retardation effects are not included. Different curves for different thicknesses have been represented, above them the film is stable in its wavy morphology, below them the film is stable in its flat morphology.

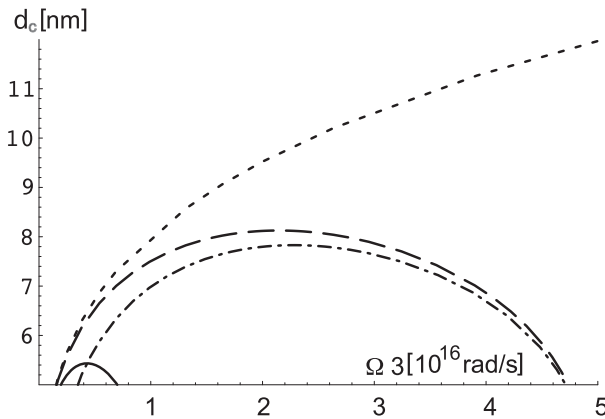


Figure 3.24: Critical thickness. For the plasma model the plasma frequency is $\Omega_1 = 10^{16} \text{ rad/s}$, continuous line and $\Omega_1 = 5 \cdot 10^{16} \text{ rad/s}$, long dashed line. Short dashed line represents the perfectly reflecting substrate $\Omega_1 \rightarrow \infty$. For the Drude model the plasma frequency value $\Omega_1 = 5 \cdot 10^{16} \text{ rad/s}$ has been used together with $\gamma_3 = \gamma_1 = 5 \cdot 10^{15} \text{ rad/s}$, dot dashed line.

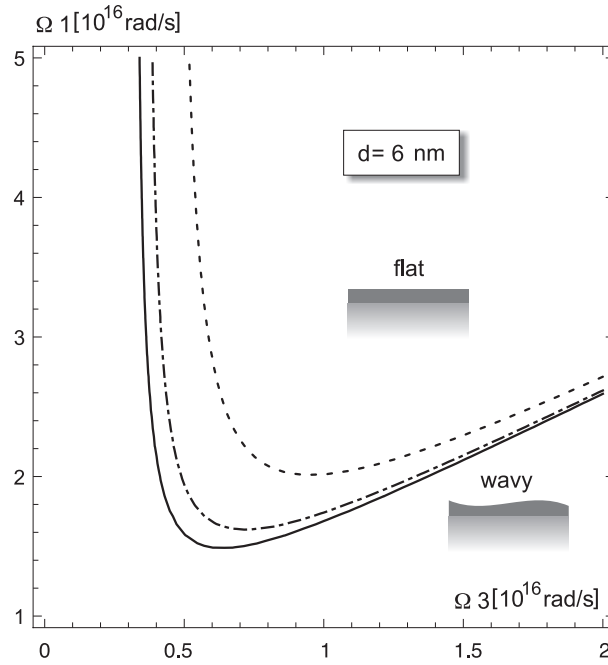


Figure 3.25: Stability diagram for the Drude film deposited onto a Drude substrate with the same dissipation coefficient of the film, retardation effects are not included. Different curves for different dissipation coefficients have been represented, $\gamma = 5 \cdot 10^{15} \text{rad/s}$ dashed line, $\gamma = 10^{15} \text{rad/s}$ dot dashed line and $\gamma = 0$ continuous line.

for a 6 nm films illustrate how the inclusion of relaxation processes can vary the range of plasma values allowing for a stable flat surface. From figure 3.24 it comes clear that the critical thickness values are reduced compared to the plasma model.

Now it is interesting to investigate the modification on the films stability due the inclusion of retardation effects in the dispersion force. The plot of the second derivative plasma frequency for the plasma model including retardation as a function of the substrate plasma frequency for a 6 nm film, presented in figure 3.26, shows that, unlike the van der Waals case, where stability is ensured for all the Ω_1 values above the crossing with the horizontal line, the full theory predicts that the flat surface condition exists for a limited range of Ω_1 values only. This result puts limitations to stability arguments based on a van der Waals description of the vacuum force, since it shows that the small d approximation is not adequate even at nanometric thicknesses, if the substrate plasma frequency is high enough. The range is smaller when relaxation effects are included into the theory. The stability plots in the $\Omega_3\Omega_1$ plane, given in figure 3.27, show the modifications occurring in the range of film-substrate parameters compatible with the flat surface growth, when both retardation and Drude effects are included in the calculations. The reduction in the values of the critical thicknesses for a given Ω_1 is illustrated in figure 3.28. It is clear from this analysis that the role of the vacuum forces in determining the stability of deposited films depends crucially upon the optical properties of both the film and the substrate. If a reasonably accurate description of the dielectric functions is not available, then it is hard to draw reliable conclusions on the expected surface morphology. Considerations based on simple empirical expressions of the force based on the van der Waals approximation may lead to overestimate the range of stability, if the conditions that make the small d approximation applicable are not met in the system under consideration.

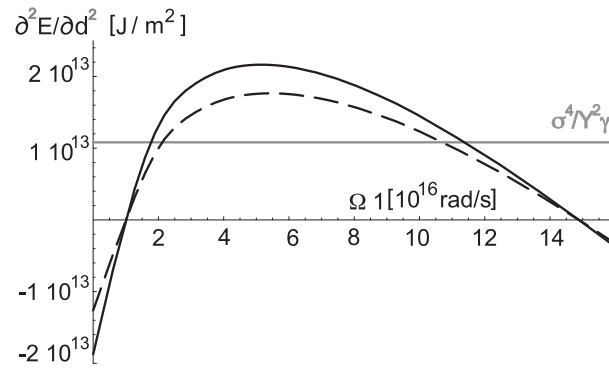


Figure 3.26: Vacuum energy second derivative plot including retardation effects. Continuous line represents the plasma model with $\Omega_3 = 10^{16} \text{rad/s}$. The Drude model, represented by the dashed line, is obtained with the same plasma frequency and $\gamma_3 = 10^{15} \text{rad/s}$ and $\gamma_1 = 10^{14} \text{rad/s}$. The critical thickness is $d = 6 \text{nm}$.

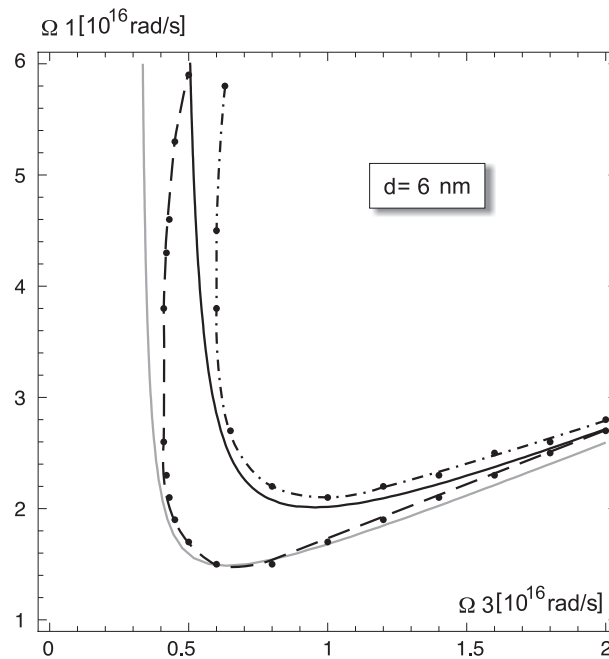


Figure 3.27: Stability diagram with retardation effects. The dashed and gray continuous lines represent the plasma film deposited onto plasma substrate with and without retardation effects respectively. The dot dashed and black continuous lines represent the Drude film deposited onto Drude substrate with the same dissipation coefficient $\gamma = 5 \cdot 10^{15} \text{rad/s}$ with and without retardation effects.

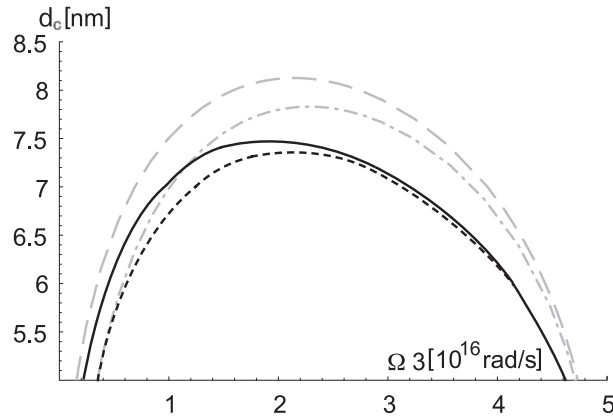


Figure 3.28: Critical thickness including retardation effects. The gray lines represent the plasma and Drude film of figure 3.24 in the small d limit. The continuous and dashed lines represent again the plasma and Drude models respectively, but including retardation effects. The substrate plasma frequency is $\Omega_1 = 5 \cdot 10^{16} \text{rad/s}$.

3.6 Conclusions

A rather complete set of results based on a continuum dielectric model has been presented, to illustrate trends in the behaviour of the electromagnetic fluctuation induced forces on free-standing and supported metal films, which allow to identify the conditions under which the force is attractive or repulsive and how it depends upon the film thickness and the interacting substrate (plate) properties.

It has been shown that both the sign and the intensity of the force between a film and a plate depend upon the difference in the plasma frequencies and can be modified upon changing the carrier density. This is in line with the recent proposal of modulating the Casimir force between a metal and a semiconductor plate by illuminating the semiconducting material, i.e. by enhancing the electron plasma and creating a hole plasma in the semiconductor plate [59]. Any experimental system that allows to change the difference in plasma frequencies can be used to modulate the electromagnetic force. The inclusion of a relaxation term in the dielectric response function has been discussed showing its effect on the dispersion force.

In the second part of the chapter the possibilities that metal films deposited on metal substrates can be stabilized by vacuum fluctuation forces has been widely investigated. It has been shown that there is a variety of situations where these forces may play a role and the constraints that have to be satisfied have been expressed, in order to allow the maintenance of flat surface condition in the film growth, in terms of the parameters entering in the description of the optical properties of both the film and the substrate.

The theory is based on a continuum description of the mismatch stress and a dielectric approach to the film and substrate optical properties. The application to a real system requires a detailed knowledge of the film optical properties. The requirement that the force be repulsive as well as its intensity may or may not be satisfied for a film of the same material depending upon the value of the relaxation frequency, the change in electron density that takes place in the pseudomorphic overlayer, the surface scattering, etc. A definite assessment on the role of the vacuum force on a specific system can be made only when these effects are properly accounted in the film dielectric function.

The present theory can be improved along two main lines of development. One can improve the continuum approach (i) by adopting a more realistic description of the elastic energy in the study of deposited film stability with the inclusion of strain modifications in the substrate and surface stress effects [39, 56], which have been neglected in the present paper; (ii) by

including interband transition effects into the dielectric function, which are appropriate when transition metal or noble metal are considered [60]. One can also extend the theory to deal with the case of metal film onto semiconducting substrate.

For very small thickness, i.e. ultra thin films with $d < 1 \div 10nm$, the theory has to be modified to include size quantization effects in the dielectric function has discussed in chapter 4. This is expected to lead to significant modifications in the vacuum force, due to the change in the film dielectric constant caused by the discretization of the electron energy levels. Moreover for ultra thin film the surface energy is expected to depend upon the thickness. Such modifications are consistent with the so called electronic growth model that has been frequently used to describe metal growth onto semiconductor substrate [61, 62, 63, 64].

Appendix A: Basics on the elasticity theory

This appendix contains an introduction to the fundamental concepts of the elasticity theory, as in the case of classical electrodynamics, one must forget about the microscopic atomistic structure of matter and regard the media as a continuum. Within the theory of elasticity, the medium is essentially treated as a spring that, subjected to a force that drive it out of equilibrium, reacts with a restoring force proportional to its displacement. The concepts of applied force and displacement must be generalized introducing the stress and the strain tensors. Then the Hooke's law can be reformulated in order to describe the medium as a complicated set of springs. The thermodynamics of the elastic deformation is also discussed and the appendix concludes with some simple applications to homogeneous media. A deeper analysis of this theory, together with the most common applications can be found in references [65, 66].

A.1 Strain and Stress tensors

A solid body subjected to external forces can undergo deformations, changing its shape or its volume. This process can be described knowing the initial coordinates \mathbf{x} of each point of the body and the corresponding coordinates after the deformation \mathbf{x}' . A *displacement vector* can be defined as:

$$\mathbf{u} = \mathbf{x}' - \mathbf{x} \quad u_i = x'_i - x_i \quad (3.63)$$

where u_i is the i -th component of the vector \mathbf{u} . Naturally the final coordinates \mathbf{x}' are a function of the initial coordinates $\mathbf{x}' = \mathbf{x}'(\mathbf{x})$, the knowledge of this function allows one to completely describe the body deformation. During the deformation of the body, the distance between its points can change. If one considers two points, 1 and 2, extremely close to each other, the i -th component of the difference vector $\mathbf{x}_1 - \mathbf{x}_2$ is dx_i whereas after the deformation the difference vector $\mathbf{x}'_1 - \mathbf{x}'_2$ has components $dx'_i = dx_i + du_i$. The distances l and l' between the two points, i.e. the moduli of the difference vectors are:

$$dl = \sqrt{\sum_i dx_i^2} \quad dl' = \sqrt{\sum_i (dx_i + du_i)^2} \quad (3.64)$$

using now the differential definition

$$du_i = \sum_k \frac{\partial u_i}{\partial x_k} dx_k \quad (3.65)$$

one gets:

$$dl'^2 = dl^2 + 2 \sum_{i,k} \frac{\partial u_i}{\partial x_k} dx_i dx_k + \sum_{i,k,l} \frac{\partial u_i}{\partial x_k} \frac{\partial u_i}{\partial x_l} dx_k dx_l \quad (3.66)$$

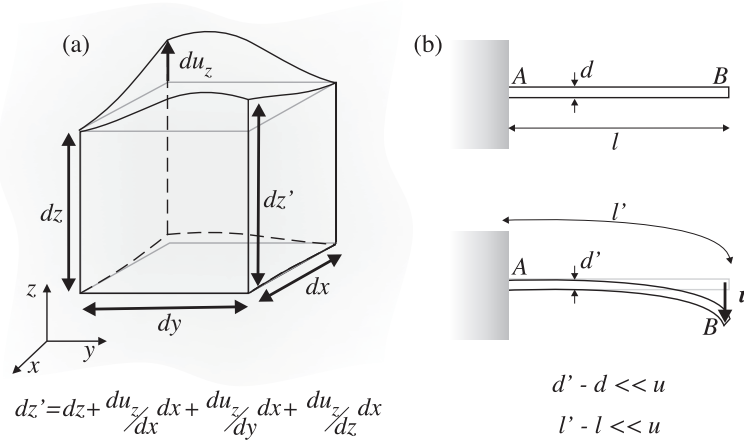


Figure 3.29: Pictorial representation of (a) the displacement in an infinitesimal unit volume inside a medium and (b) in a flexed rod.

i and l are dummy indexes and can be exchanged to give

$$dl'^2 = dl^2 + 2 \sum_{ik} \epsilon_{ik} dx_i dx_k = (\delta_{ik} + 2 \sum_{ik} \epsilon_{ik}) dx_i dx_k \quad (3.67)$$

where the *strain tensor*

$$\epsilon_{ik} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} + \sum_l \frac{\partial u_l}{\partial x_k} \frac{\partial u_l}{\partial x_i} \right) \quad (3.68)$$

as been introduced. This tensor is dimensionless and expresses the variation of the infinitesimal length in a certain direction due to the body deformation along another direction, see figure 3.29 (a). Definition (3.68) shows that the strain tensor is symmetric $\epsilon_{ik} = \epsilon_{ki}$, this means that for each point of the space there exist a set of principal axes that allows to diagonalize it. In the diagonal case only ϵ_{ii} are non vanishing and expression (3.67) becomes

$$dl'^2 = \sum_i (1 + 2\epsilon_{ii}) dx_i^2 \quad (3.69)$$

one can define the relative stain along a certain direction as:

$$\frac{dx'_i - dx_i}{dx_i} = \sqrt{1 + 2\epsilon_{ii}} - 1 \quad (3.70)$$

In most of the applications of the elasticity theory one deals only with small strains, i.e. ϵ_{ik} is always small. This does not means that even the body deformation is small. Take for example the thin elastic rod of figure 3.29 (b): even in the case of large flexions, its length and thickness values are slightly modified with respect to the rest position. Nevertheless for bulk systems (where none of the three dimensions is small with respect to the others) a small strain always correspond to a small body deformation and the second order term of the tensor (3.68) can be neglected

$$\epsilon_{ik} \simeq \frac{1}{2} \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right) \quad (3.71)$$

for the same reason only the first term in ϵ_{ii} of the Taylor's expansion of (3.70) must be retained, $(dx'_i - dx_i)/dx_i \simeq \epsilon_{ii}$, the components of the diagonalized tensor becomes directly

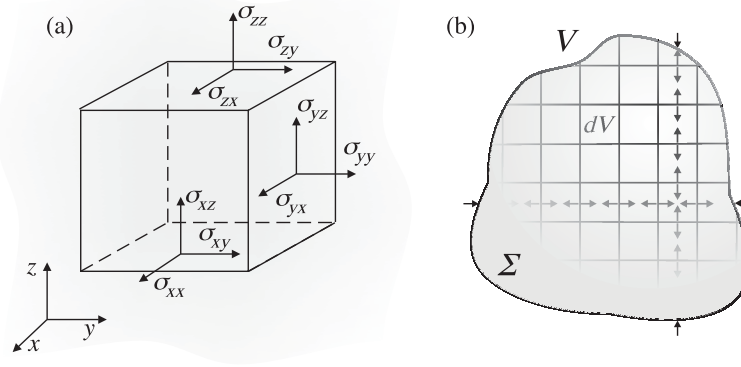


Figure 3.30: (a) Pictorial representation of the stress tensor, σ_{ii} components represent the *normal stresses* while σ_{ik} components represent the *shear stresses*. (b) Representation of the internal forces, the dV nearest neighbors interaction annihilates all the opposite forces, only those applied on the bounding surface Σ are relevant.

the relative strains.

It is also possible to calculate the effects of the body deformation on the unit volume $dV = \prod_i dx_i$:

$$dV' = \prod_i dx'_i = \prod_i (1 + \epsilon_{ii}) dx_i = dV \prod_i (1 + \epsilon_{ii}) \quad (3.72)$$

retaining only first order terms in ϵ_{ii} one get:

$$dV' = dV \left(1 + \sum_i \epsilon_{ii} \right) \quad (3.73)$$

so finally the relative volume variation is $(dV' - dV)/dV = \sum_i \epsilon_{ii}$, i.e. the trace of the strain tensor. The trace of a tensor is an invariant so, even if the principal axes change point by point, this last result is independent of the choice of the coordinate system.

When a body is deformed by external forces it goes out of its equilibrium state, an *internal force* rises to restore a new equilibrium condition, these are typically elastic restoring forces. A new equilibrium condition is reached when the external forces are balanced by the internal ones. The total force \mathbf{R} acting on the deformed medium volume V inside the deformed body, can be obtained summing all the forces acting to each infinitesimal volume dV :

$$\mathbf{R} = \int_V \mathbf{F} dV \quad (3.74)$$

where \mathbf{F} is the force per unit volume, i.e. a force density. Part of the force $\mathbf{F}dV$ applied to each infinitesimal volume dV is compensated by the force applied on the infinitesimal volumes nearest neighbors, as shown in figure 3.30 (b). The only uncompensated forces are the one applied on the external boundary surface of the body, because of this, the volume integral (3.74) can be transformed into a surface integral by means of the generalized³ divergence theorem:

$$\mathbf{R} = \int_V \mathbf{F} dV = \int_V \nabla \cdot \sigma dV = \oint_{\Sigma} \sigma \cdot d\hat{\mathbf{n}} \quad (3.75)$$

³The standard divergence theorem states that the volume integral of a scalar function can be transformed into a surface integral of a vector field if the scalar function can be expressed through the divergence of the vector field. Here one deals with the volume integral of a vector that can be transformed into a surface integral if the vector field can be expressed through the divergence of a rank two tensor.

here Σ is the surface bounding the volume V , $\hat{\mathbf{n}}$ is the versor normal to the surface Σ and σ is the *stress tensor*. For a single component one obtains:

$$R_i = \int_V F_i dV = \sum_k \int \frac{\partial \sigma_{ik}}{\partial x_k} dV = \sum_k \oint_{\Sigma} \sigma_{ik} d\hat{n}_k \quad (3.76)$$

from a dimensional point of view the stress tensor is a pressure, $\sum_i \sigma_{ik} d\hat{n}_k$ represents the force acting on the infinitesimal surface element $d\hat{n}_k$, as shown in figure 3.30 (a). Using the simple argument that even the momentum of the force $\mathbf{F} \times \mathbf{x}$ must be expressed by a surface integral it comes clear that the stress tensor, as well as the strain one, must be symmetric, i.e. $\sigma_{ik} = \sigma_{ki}$ [65]. In an equilibrium configuration the total force acting on the medium must be zero, so the total force \mathbf{R} acting on each possible volume V of the medium must be zero too, the equilibrium condition is given by:

$$\sum_k \frac{\partial \sigma_{ik}}{\partial x_k} = 0 \quad \forall i \quad (3.77)$$

this is analogous to the elementary dynamics statement that the resultant of all the forces acting on a pointy mass at equilibrium must be zero. Some force fields, as the gravitational one, act on the whole body, through each single infinitesimal volume dV of the medium, these *body forces* must be considered as all the other forces acting on the volume V and included in the previous condition

$$\sum_k \frac{\partial \sigma_{ik}}{\partial x_k} + F_i = 0 \quad \forall i \quad (3.78)$$

Finally mechanical forces (bending forces, torsional forces, mechanical pressures) are always applied to the boundary of the bodies, they can not enter the stability conditions (3.78) and (3.77). These *surface forces* rule the problem by means of boundary conditions: if a pressure \mathbf{P} is applied to the body surface, $\mathbf{P}d\hat{\mathbf{n}}$ is the force applied on the infinitesimal surface element $\hat{\mathbf{n}}$, in an equilibrium condition this force must balance all the other forces \mathbf{R} acting on a single surface element of the medium:

$$\begin{aligned} P_i d\hat{\mathbf{n}} - \sum_k \sigma_{ik} d\hat{n}_k &= 0 \\ \sum_k \sigma_{ik} \hat{n}_k &= P_i \quad \forall i \end{aligned} \quad (3.79)$$

A.2 Thermodynamics of deformations

Suppose to apply a certain force on a medium in order to deform it, once that the force is no more acting on it, the medium can behave in different ways, a *perfectly elastic medium* is able to recover the original volume and shape it had before the deformation. In *plastic media* this is not possible, part of the deformation energy is dissipated modifying the medium microscopic structure, the original volume and shape can not be recovered. A wide range of materials present an elastic behaviour for small deformation whereas they enter a plastic regime when the deformation are very strong. In the following only the elastic regime will be treated. If one considers a local infinitesimal deformation of a medium, described by the displacement $\delta \mathbf{u}$, it is possible to calculate the work W needed to keep the deformation, i.e. the same work made by the internal forces against the deformation:

$$W = \int_V dw dV = \sum_i \int_V F_i \delta u_i dV = \sum_{ik} \int_V \frac{\partial \sigma_{ik}}{\partial x_k} \delta u_i dV \quad (3.80)$$

where dw is the work density, integrating by parts with the generalized formula:

$$\int_V \nabla \cdot \beta \cdot \mathbf{Y} dv = \oint_{\Sigma} \beta \cdot \mathbf{Y} \cdot d\hat{\mathbf{n}} - \int_V \beta \nabla \cdot \mathbf{Y} dv \quad (3.81)$$

where β is a rank two tensor and \mathbf{Y} is a vector, one gets:

$$W = \sum_{ik} \oint_{\Sigma} \sigma_{ik} \delta u_i d\hat{n}_k - \sum_{ik} \int_V \sigma_{ik} \frac{\partial \delta u_i}{\partial x_k} dV \quad (3.82)$$

the deformation is supposed to be local so, on a large surface Σ , including a portion of medium much larger than the one in which the deformation occurs, the surface integral in the RHS vanishes. Now using the symmetry properties of the stress tensor, the last expression can be rewritten to introduce the strain tensor (3.71):

$$W = \sum_{ik} \int_V \sigma_{ik} \frac{\partial \delta u_i}{\partial x_k} dV = -\frac{1}{2} \sum_{ik} \int_V \sigma_{ik} \delta \left(\frac{\partial u_i}{\partial x_k} + \frac{\partial u_k}{\partial x_i} \right) dV \simeq - \sum_{ik} \int_V \sigma_{ik} \delta \epsilon_{ik} dV \quad (3.83)$$

To conclude the work density is given by

$$dw = - \sum_{ik} \sigma_{ik} \delta \epsilon_{ik} \quad (3.84)$$

Now using the first thermodynamics principle, with all the thermodynamics variables expressed per unit volume, one gets:

$$dE = TdS - dw = TdS + \sum_{ik} \sigma_{ik} d\epsilon_{ik} \quad (3.85)$$

or, in the case of Helmholtz's free energy (per unit volume):

$$dF = -SdT + \sum_{ik} \sigma_{ik} d\epsilon_{ik} \quad (3.86)$$

or finally, in the case of the Gibbs's free energy (per unit volume):

$$dG = -SdT - \sum_{ik} \epsilon_{ik} d\sigma_{ik} \quad (3.87)$$

Now some general expressions can be derived:

$$\sigma_{ik} = \left(\frac{\partial E}{\partial \epsilon_{ik}} \right)_S = \left(\frac{\partial F}{\partial \epsilon_{ik}} \right)_T \quad \epsilon_{ik} = \left(\frac{\partial G}{\partial \sigma_{ik}} \right)_T \quad (3.88)$$

A.3 Generalized Hooke's law

In the preceding section it has been shown that the free energy F is a function of the strain tensor, one can think to expand F in Taylor's series. If the expansion variable is a tensor, the coefficients of the expansion are tensors too:

$$F = F_0 + \sum_{ik} \lambda_{ik} \epsilon_{ik} + \frac{1}{2} \sum_{iklm} \lambda_{iklm} \epsilon_{ik} \epsilon_{lm} + \dots \quad (3.89)$$

using the definition (3.88) one gets:

$$\sigma_{ik} = \left(\frac{\partial F}{\partial \epsilon_{ik}} \right)_T = \lambda_{ik} + \sum_{lm} \lambda_{iklm} \epsilon_{lm} + \dots \quad (3.90)$$

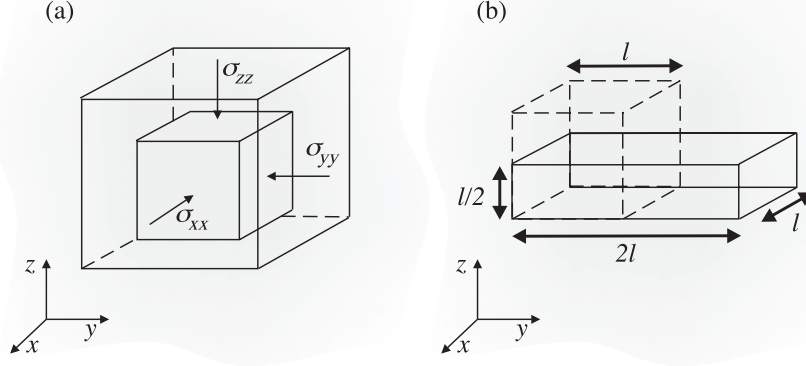


Figure 3.31: Uniform compression of a cubic volume (a) and pure shear deformation example (b).

but if there is no deformation, $\epsilon_{ij} = 0$, the stress vanish, $\sigma_{ij} = 0$, this means that $\lambda_{ik} = 0$: the first term of the series expansion of the free energy F must be the second order one:

$$F \simeq F_0 + \frac{1}{2} \sum_{iklm} \lambda_{iklm} \epsilon_{ik} \epsilon_{lm} \quad (3.91)$$

as a consequence, the stress tensor takes the form:

$$\sigma_{ik} = \sum_{lm} \lambda_{iklm} \epsilon_{lm} \quad (3.92)$$

this is the equivalent of the classical Hooke's law, the elastic force replaced by the stress tensor, the displacement replaced by the strain tensor and the Hooke's constant replaced by a rank four tensor with 81 components. λ_{iklm} it is usually called *elasticity modulus tensor*, it must be said that, the strain tensor being symmetric, only 21 of the total 81 components are linearly independent, in fact $\epsilon_{ik} \epsilon_{lm}$ must be invariant for indexes permutation, and it follows that

$$\lambda_{iklm} = \lambda_{kilm} = \lambda_{ikml} = \lambda_{lmik} \quad (3.93)$$

Through the elasticity modulus tensor, it is possible to recover informations about the microscopic character of a medium, in fact, due to the peculiar microscopic symmetry of the crystal lattice, the elastic response to a certain perturbation along different direction can be the same. The higher the symmetry of the lattice, the lower the number of linearly independent λ components. For a cubic system the number of linearly independent components decrease from 21 to 3. In the case of a quite homogeneous amorphous or a polycrystalline medium with small grains size, one can use the approximation of homogeneous medium in which only two λ components are independent.

A.4 Homogeneous media and homogeneous deformations

This last part of the appendix is devoted to the study of homogeneous media acted on by homogeneous deformations. In an homogeneous medium the general expression (3.91) contains a lot of non-independent coefficients, it can be simplified using the linear algebra theorems and reconduced to the linear combination of the two independent scalars built up with the tensor elements, the squared trace and the sum of the squared elements:

$$F = F_0 + \frac{\lambda}{2} \left(\sum_i \epsilon_{ii} \right)^2 + \mu \sum_{ik} \epsilon_{ik}^2 \quad (3.94)$$

where λ and μ are called *Lamé coefficients*. Equation (3.73) shows that the sum over the diagonal elements of the strain tensor gives the relative volume variation, if this quantity is zero, it means that the deformation modifies the body shape leaving the volume unchanged, such a deformation is called *pure shear deformation*, see figure 3.31 (b). The opposite situation is that of a *uniform compression (expansion)*, that changes the volume leaving unchanged the body shape, e.g. to perform a uniform compression on the cube of figure 3.31 (a), the stress tensor must be diagonal: $\sigma_{ik} = \text{const} \cdot \delta_{ik}$. Each deformation can be decomposed in a pure shear deformation and a uniform compression (expansion), this can be seen by adding and subtracting the trace $1/3 \sum_l \delta_{ik} \epsilon_{ll}$ to the generic strain tensor component

$$\epsilon_{ik} = \left(\epsilon_{ik} - \frac{1}{3} \sum_l \delta_{ik} \epsilon_{ll} \right) + \frac{1}{3} \sum_l \delta_{ik} \epsilon_{ll} \quad (3.95)$$

the term in brackets is a pure shear term: it vanishes summing the three diagonal terms ϵ_{ii} , the last term in the RHS is a uniform compression (expansion term) because it is in the form $\text{const} \cdot \delta_{ik}$. Replacing this last expression into (3.94), after some simplifications, one gets

$$F = F_0 + \mu \sum_{ik} \left(\epsilon_{ik} - \frac{1}{3} \delta_{ik} \sum_l \epsilon_{ll} \right)^2 + \frac{K}{2} \left(\sum_l \epsilon_{ll} \right)^2 \quad (3.96)$$

where the *compression modulus* K is given by:

$$K = \lambda + \frac{2}{3} \mu \quad (3.97)$$

and μ is the same as before but in this context it takes the name of *shear modulus*. When $\epsilon_{ik} = 0$ no forces are acting on the medium, F must have a minimum so the quadratic form (3.96) must be positive, this means that both μ and K must be positive defined. The general definitions (3.88) can be used to calculate the stress tensor, the total energy differential is:

$$dF = K \sum_l \epsilon_{ll} d\epsilon_{ll} + 2\mu \sum_{ik} \left(\epsilon_{ik} - \frac{1}{3} \delta_{ik} \sum_l \epsilon_{ll} \right) d\epsilon_{ik} \quad (3.98)$$

writing $d\epsilon_{ll} = \delta_{ik} d\epsilon_{ik}$ it becomes

$$dF = \sum_{ik} \left[K \delta_{ik} \sum_l \epsilon_{ll} + 2\mu \left(\epsilon_{ik} - \frac{1}{3} \delta_{ik} \sum_l \epsilon_{ll} \right) \right] d\epsilon_{ik} \quad (3.99)$$

and finally by virtue of (3.88) one reach the Hooke's law for an homogeneous medium

$$\sigma_{ik} = K \delta_{ik} \sum_l \epsilon_{ll} + 2\mu \left(\epsilon_{ik} - \frac{1}{3} \delta_{ik} \sum_l \epsilon_{ll} \right) \quad (3.100)$$

The sum of the three diagonal components gives $\sum_i \sigma_{ii} = 3K \sum_i \epsilon_{ii}$ and is the scalar quantity that allows to reverse the Hooke's law

$$\epsilon_{ik} = \frac{1}{9K} \delta_{ik} \sum_l \sigma_{ll} + \frac{1}{2\mu} \left(\sigma_{ik} - \frac{1}{3} \delta_{ik} \sum_l \sigma_{ll} \right) \quad (3.101)$$

F is a quadratic form in the strain ϵ_{ik} , from the Euler's theorem

$$\sum_{ik} \epsilon_{ik} \frac{\partial F}{\partial \epsilon_{ik}} = 2F \quad (3.102)$$

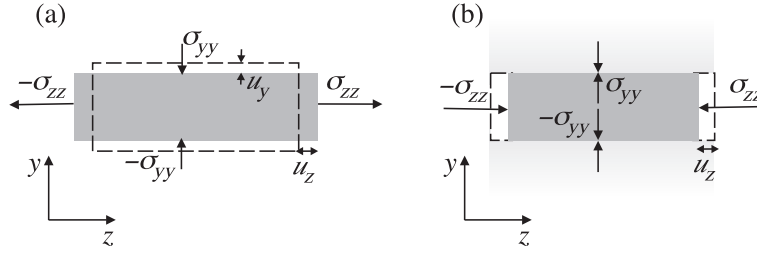


Figure 3.32: Stretching of a free rod (a) and compression of a constrained rod (b). Dashed contours indicate the rod shape before the pressure is applied.

and, using again (3.88) one gets:

$$F = \sum_{ik} \frac{\sigma_{ik}\epsilon_{ik}}{2} \quad (3.103)$$

An *homogeneous deformation* is a deformation that leaves the strain tensor ϵ_{ik} constant in all the medium volume. Imagine to pull the rod of figure 3.32 (a) along the z axis with the uniform pressure p , the stress tensor can be determined imposing the boundary conditions (3.79): no forces are applied to the lateral surfaces of the rod so $\sum_k \sigma_{ik}n_k = 0$ for $i = x, y$ whereas $\sum_k \sigma_{zk}n_k = p$ gives the result $\sigma_{zz} = p$ because the \mathbf{n} vector is oriented along the z direction and $n_x = n_y = 0$. Under the assumption that the medium is a uniform one, this result can be used into the inverse Hooke's law (3.101) to calculate the strains:

$$\begin{aligned} \epsilon_{xx} = \epsilon_{yy} &= -\frac{1}{3} \left(\frac{1}{2\mu} - \frac{1}{3K} \right) p \\ \epsilon_{zz} &= \frac{1}{3} \left(\frac{1}{3K} + \frac{1}{\mu} \right) p = \frac{p}{E} \end{aligned} \quad (3.104)$$

E is known as the *Young modulus* and it is always positive. ϵ_{zz} and ϵ_{xx} have opposite sign, pulling the rod along the z direction it squeezes along the other two. The ratio between compression and elongation of the rod is called *Poisson's coefficient*

$$M = -\frac{\epsilon_{xx}}{\epsilon_{zz}} = -\frac{\epsilon_{yy}}{\epsilon_{zz}} = \frac{1}{2} \frac{3K - 2\mu}{3K + \mu} \quad (3.105)$$

and it ranges between -1 and $1/2$. From the energy (3.103) comes the result

$$F = \frac{p^2}{2E} \quad (3.106)$$

Finally consider the rod of figure 3.32 (b), it is laterally constrained: under compression no deformation can take place along x and y directions. The only non vanishing strain component is ϵ_{zz} , one can use the Hooke's law (3.100) to calculate the stresses

$$\begin{aligned} \sigma_{xx} = \sigma_{yy} &= \left(K - \frac{2}{3}\mu \right) \epsilon_{zz} = \frac{EM}{(1+M)(1-2M)} \epsilon_{zz} \\ \sigma_{zz} &= \left(K + \frac{4}{3}\mu \right) \epsilon_{zz} = \frac{E(1-M)}{(1+M)(1-2M)} \epsilon_{zz} \end{aligned} \quad (3.107)$$

combining the two expressions one can calculate the stress induced on x and y directions from the constrains

$$\sigma_{xx} = \sigma_{yy} = -\frac{K - \frac{2}{3}M}{K + \frac{4}{3}M} p = -\frac{M}{1-M} p \quad (3.108)$$

where $\sigma_{zz} = -p$, a compression force per unit area. And to conclude the energy results:

$$F = \frac{(1+M)(1-2M)}{2E(1-M)}p^2 < \frac{p^2}{2E} \quad (3.109)$$

the constrained rod is less strained with respect to the free one, because of this it contains less elastic energy.

Bibliography

- [1] A. Benassi and C. Calandra. *J.Phys.A: Math.Theor.*, 40:13453, 2007.
- [2] A. Benassi and C. Calandra. *J.Phys.A: Math.Theor.*, 41:175401, 2008.
- [3] J. Mahanty and B.N. Ninham. *Dispersion forces*. Academic Press, New York, 1976.
- [4] P.W. Milonni. *The quantum vacuum*. Academic Press, New York, 1994.
- [5] M. Bordag, U. Mohideen, and V.M. Mostepanenko. *Phys. Rep.*, 353:1, 2001.
- [6] S.K. Lamoreaux. *Rep. Prog. Phys.*, 68:201, 2005.
- [7] I.E. Dzyaloshinskii, E.M. Lifshitz, and L.P. Pitaevskii. *Adv. Phys.*, 10:165, 1958.
- [8] B.V. Derjaguin and I.I. Abrikosova. *Sov. Phys. JEPT*, 3:819, 1957.
- [9] G. Barton. *J. Phys. A: Math. Gen.*, 34:4083, 2001.
- [10] R. Podgornik and V.A. Parsegian. *J. Chem. Phys.*, 120:3410, 2004.
- [11] L.R. White, R.R. Dagastine, P.M. Jones, and Yiao-Tee Hsia. *J. Appl. Phys.*, 97:104503, 2005.
- [12] F.M. Serry, D. Walliser, and G.J. Maclay. *J. Appl. Phys.*, 84:2501, 1998.
- [13] E. Buks and M.L. Roukes. *Phys. Rev. B*, 63:033402, 2001.
- [14] Y.-P. Zhao, L.S. Wang, and T.X. Yu. *J. Adhesion Sci. Technol.*, 17:519, 2003.
- [15] J. Barcenas, L. Reyes, and R. Esquivel-Sirvent. *Appl. Phys. Lett.*, 87:263106, 2005.
- [16] B.E. Sernelius and P. Bjork. *Phys. Rev. B*, 57:6592, 1998.
- [17] M. Boström and B.E. Sernelius. *Phys. Rev. B*, 61:2204, 2000.
- [18] M. Boström and B.E. Sernelius. *Phys. Rev. B*, 62:7523, 2000.
- [19] A. Lambrecht and S. Reynaud. *Eur. Phys. D*, 8:309, 2000.
- [20] I. Pirozhenko, A. Lambrecht, and V.B. Svetovoy. *New J. Phys.*, 8:238, 2006.
- [21] J.P. Rogers III, P.H. Cutler, T.R. Feuchtwang, N. Miskovski, and A.A. Lucas. *Surf. Sci.*, 141:61, 1984.
- [22] J.P. Rogers III, P.H. Cutler, T.R. Feuchtwang, and A.A. Lucas. *Surf. Sci.*, 181:436, 1987.
- [23] P.D. Loly and J.B. Pendry. *J. Phys. C*, 16:423, 1986.
- [24] S.A. Lindgren and L. Wallden. *Phys. Rev. Lett.*, 59:3003, 1987.
- [25] S.A. Lindgren and L. Wallden. *Phys. Rev. Lett.*, 61:2894, 1988.

- [26] R. Esquivel-Sirvent and V.B. Svetovoy. *Phys. Rev. B*, 72:045443, 2005.
- [27] N.G. van Kampen, B.R.A. Nijboer, and K. Schram. *Phys. Lett.*, 26A:307, 1968.
- [28] E. Gerlach. *Phys. Rev. B*, 4:393, 1971.
- [29] F. Intravaia and A. Lambrecht. *Phys. Rev. Lett.*, 94:110404, 2005.
- [30] J.N. Israelachvili. *Intermolecular and Surface Forces*. Academic Press, London U.K., 1991.
- [31] R.H. French. *J.Am.Ceram.Soc.*, 83:2117, 2000.
- [32] F. Zhou and L. Spruch. *Phys. Rev. A*, 52:297, 1995.
- [33] G. Palasantzas. *J. Adhesion Sci. Technol.*, 20:1321, 2006.
- [34] G. Palasantzas. *J. Appl. Phys.*, 97:126104, 2005.
- [35] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Cornell University, Philadelphia, 1976.
- [36] F.F. Streitz, R.C. Cammarata, and K. Sieradzki. *Phys.Rev.B*, 49:10699, 1994.
- [37] R.C. Cammarata. *Prog. Surface Sci.*, 46:1, 1994.
- [38] H. Ibach. *Surface Sci. Reports*, 29:193, 1997.
- [39] P. Mueller and A. Saul. *Surface Sci. Reports*, 54:157, 2004.
- [40] K.L. Chopra. *Thin Film Phenomena*. McGraw-Hill Book Company, New York, 1969.
- [41] J.W. Matthews, editor. *Epitaxial Growth*. Academic Press, New York, 1975.
- [42] H. Dreyssé and C. Demangeat. *Surface.Sci.Rep.*, 28:65, 1997.
- [43] Qiang Fu and T. Wagner. *Surface.Sci.Rep.*, 62:411, 2007.
- [44] D.T. Jang, E.D. Crozier, and B. Heinrich. *Phys.Rev.B*, 44:6401, 1991.
- [45] A.M. Bittner, J. Wintterlin, and G. Ertl. *Surface.Sci.*, 376:267, 1997.
- [46] M. Dietterle, T. Will, and D.M. Kolb. *Surface.Sci.*, 396:189, 1998.
- [47] R.J. Randler, D.M. Kolb, B.M. Ocko, and I.K. Robinson. *Surface.Sci.*, 447:187, 2000.
- [48] L.B. Freund. *Int.J.Solids Structures*, 37:185, 2000.
- [49] Zhigang Suo and Zhenyu Zhang. *Phys. Rev. B*, 58:5116, 1998.
- [50] O. Kraft, M. Hommel, and E. Arzt. *Mat.Sci.Eng.A*, 288:209, 2000.
- [51] R. Asaro and W. Tiller. *Met.Trans.*, 3:1789, 1971.
- [52] M. Grinfield. *Sov.Phys.Dokl.*, 31:831, 1986.
- [53] D. Stolovitz. *Acta Metal.*, 37:621, 1989.
- [54] H. Gao. *Int.J.Solids Structures*, 28:703, 1991.
- [55] B.J. Spencer, P.W. Voorhees, and S.H. Davis. *Phys.Rev.Lett.*, 67:3696, 1991.
- [56] H.Y. Jiang and L.H. He. *J.Cryst.Growth*, 262:28, 2004.

- [57] H. Gao. *J.Mech.Phys.Solids*, 39:443, 1991.
- [58] Y.-P. Zhao. *Archive of applied mechanics*, 72:77, 2002.
- [59] G.L. Klimchitskaya, U. Mohideen, and V.M. Mostepanenko. *J. Phys. A:Math.Theor.*, 40:F841, 2007.
- [60] F. Wooten. *Optical properties of solids*. Academic Press, Davis, California, 1988.
- [61] T.C. Chiang. *Surface Sci. Rep.*, 39:181, 2000.
- [62] P. Czoshke, H. Hong, L. Basile, and T.C. Chiang. *Phys.Rev.B*, 72:035305, 2005.
- [63] Jin-Feng Jia, Shao-Chun Li, Yan-Feng Zhang, and Qi-Kun Xue. *J.Phys.Soc.Japan*, 76:082001, 2007.
- [64] Z.Y. Zhang, Q. Niu, and C.K. Shih. *Phys.Rev.Lett.*, 85:5158, 1998.
- [65] L.D. Landau and E.M. Lifshitz. *Theory of elasticity*. Mir, Mosca, 1979.
- [66] S.P. Timoshenko and J.N. Goodier. *Theory of elasticity*. McGraw-Hill, Singapore, 1970.

4

Quantum confinement effects

Contents

4.1	Introduction	149
4.2	Two interacting films	151
4.2.1	The particle in a box model	151
4.2.2	Finite depth well	157
4.2.3	The generalized particle in a box model	165
4.3	Isolated metallic film	166
4.4	Temperature dependence	171
4.5	Conclusions	173
	Appendix A: Reflection coefficient of a slabs	174

This chapter deals with the modifications of the dispersion force due to the inclusion of the quantum nature of the electron gas composing the interacting media. This quantum nature comes clearly out when the size of the interacting objects is small enough to impose a confinement to the electrons wavefunctions. For this reason thin films provide a suitable system to be investigated since their electron gas is confined along one dimension and behaves as a free classical gas along the other two. Moreover, isolated thin films or interacting thin films, constitute a system with a relatively simple geometry compared to other low dimensional structures. The dispersion interaction can be investigated with the Lifshitz theory obtaining exact results. Generalizations of the ordinary Drude and plasma dielectric function models will be presented and used to calculate the dispersion force acting on isolated films or between identical films. A detailed comparison with the force calculated using the standard dielectric function models will be given. The work described here has been published in two papers, references [1, 2]

4.1 Introduction

Since the basic work from Lifshitz and coworkers, illustrated in section 2.3, theoretical studies have been focused mainly on the determination of the forces between slabs, including semi-infinite slabs, on the basis of a continuum description of the material dielectric function

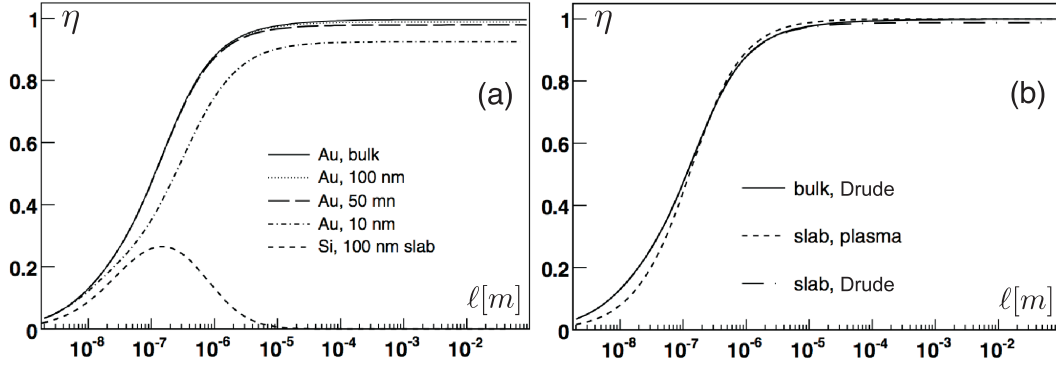


Figure 4.1: (a) reduction factor for gold and silicon films of different thickness as a function of the film separation. (b) reduction factor of gold obtained with different dielectric function models as a function of the film separation.

[3, 4, 5, 6, 7]. Both metal and semiconductor films have been considered with the use of different dielectric function models. Figure (4.1), adapted from a work of Lambrecht et al. [6], summarize all the interesting features of the interaction between identical films of thickness D separated by a distance ℓ . It displays the reduction factor η defined as the ratio between the Lifshitz force and its Casimir limit:

$$\eta = \frac{F}{F_{CAS}} \quad (4.1)$$

for gold and silicon films of different thicknesses. Independently of the film thickness, the film distance and the film dielectric properties, the force between films is always lowered with respect to the interaction between semi-infinite slabs, i.e. $\eta \leq 1$, the smaller the film thickness the larger the force attenuation. For a non dissipative metal described by the plasma model, $\eta \rightarrow 1$ at large film separation; inclusion of a dissipative behaviour by means of a Drude model, leads to a force that always remains lower than the Casimir one. In the case of a semiconductor described by a Lorentz model, the η factor goes to zero both at large and small film separation. All these properties can be understood by looking at the reflectivity of a single film, and considering that the modes that contribute the most to the dispersion interaction are those whose wavelength is comparable with the cavity amplitude ℓ . Following the derivation in appendix A, the reflection coefficient of a single film is given by:

$$r_\alpha = \rho_\alpha \frac{1 - e^{2id\beta}}{1 - \rho_\alpha^2 e^{2id\beta}} \quad \beta_j = \sqrt{\epsilon_j(\omega) \frac{\omega^2}{c^2} - k^2} \quad (4.2)$$

where α denotes the two possible polarizations (TE and TM modes), ρ_α is the single surface reflectivity. Performing the change of variable $\omega \rightarrow i\omega$ one obtains exactly the integrand of the Lifshitz formula (2.73):

$$r_\alpha = \rho_\alpha \frac{1 - e^{-2d\beta}}{1 - \rho_\alpha^2 e^{-2d\beta}} \quad \beta_j = \sqrt{\epsilon_j(i\omega) \frac{\omega^2}{c^2} + k^2} \quad (4.3)$$

for large thicknesses one recovers the single surface reflection coefficient whereas, for small D values, the film reflection coefficient vanishes. In the latter case the dispersion interaction is dramatically decreased because the cavity walls are almost transparent to the electromagnetic field. In the case of large thickness film:

- For large film separation only the large wavelength modes, i.e. the low frequency modes, contribute significantly to the dispersion interaction. If a non dissipative metal film is

thick enough, the reflectivity is simply given by the plasma model, it completely reflects the small frequency modes trapping them inside the cavity, one recovers exactly the Casimir's limit and $\eta \rightarrow 1$. In the case of the Drude model, even if the cavity walls are able to trap the low frequency modes inside the cavity, part of their energy is dissipated by the films and the dispersion force can not reach exactly the Casimir's value. On the other hand, if the thick film is described by a Lorentz dielectric function, it is transparent to the low frequencies modes, i.e. the cavity walls are transparent to the most significant cavity modes, $\eta \rightarrow 0$.

- For small film separation only the small wavelength modes, i.e. the high frequency modes, contribute significantly to the dispersion interaction. But both metals and semiconductors are transparent to the high frequency fields and again they become transparent to the cavity modes that contribute the most. η decrease rapidly with decreasing ℓ .

4.2 Two interacting films

Model dielectric functions, give a simplified description of a film, that neglects the modifications in the electronic structure related to the presence of the boundaries. These are expected to be valid when the size of the film is large, so that the surfaces play a minor role in determining the dielectric response. For metallic films the calculated electronic distributions deviate significantly from the bulk behaviour when the size of the film is less than ten times the Fermi wavelength [8, 9]. If the size of the film is of the order of few nanometers the continuum model does not provide an accurate description of the film properties and boundary effects cannot be neglected. Such effects arise as a consequence of the discretization of the energy bands due to the confinement potential, which produces the quantization of the electron energy levels in various sub-bands and affects the optical and electrical properties [10, 11, 12, 13, 14, 15, 16, 17].

4.2.1 The particle in a box model

A first rough model to account for the quantum size effects (QSE) felt by the electrons in a thin film, is provided by the Wood and Ashcroft model described in section 1.11, also known as *the particle in a box model* (PBM). Independent electrons of mass m are confined within a distance d in the direction of the surface normal [18, 19]. Assuming a jellium model and perfect planar surfaces the eigenvalue spectrum is simply given by:

$$E_{\mathbf{k}_{\parallel},n} = \frac{\hbar^2}{2m}(k_{\parallel}^2 + k_{\perp}^2) = \frac{\hbar^2}{2m}\left(k_{\parallel}^2 + \frac{n^2\pi^2}{d^2}\right) \quad (4.4)$$

i.e. described by a continuous quantum number k_{\parallel} , giving the modulus of the parallel wavevector, and a discrete sub-band index n coming from the quantization of the perpendicular wavevector k_{\perp} . The corresponding wavefunctions are:

$$\psi_{\mathbf{k}_{\parallel},n}(\mathbf{r}) = \sqrt{\frac{2}{V}} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}} \sin\left(\frac{n\pi}{d} z\right) \quad (4.5)$$

where V is the volume of the quantum well given by the product of the well surface A and the well thickness d . In this simple model the electrons behave as a two-dimensional gas along the x and y space directions and as standing waves in the z direction, with nodes on the boundaries. As a first approximation one can assume the size d of the quantum well to be the same as the size D of the ion distribution of the film. This approximation is too crude

since it does not allow the spilling of the electron density past the film boundaries given by the positive charge distribution, thus leading to a depletion of negative charge near the surfaces. One way to eliminate this inconvenience without introducing softer boundaries is to allow the electron charge to be distributed on a larger size than the positive charge. It is common practice to discuss the effects of the confining potential in deposited films using the phase accumulation model [20], according to which the condition that is satisfied by a quantum well state is

$$2k_{\perp}d + \phi_A + \phi_B = 2\pi n \quad (4.6)$$

where ϕ_A and ϕ_B are the phases of the eigenfunctions accumulated at the two film interfaces and n are integer numbers. For a free standing film, limited by two vacuum-film interfaces, $\phi_A = \phi_B$. The case of an infinite quantum well is recovered by simply imposing $\phi_A = -\pi$. For softer profiles of the confinement potential the phase shifts are expected to increase and to depend upon the energy. One can introduce a more realistic description of the film electron states, still using the infinite potential well, by allowing the effective width d to be larger than the size of the ion distribution $d = D + 2\Delta$, thus obtaining the quantization condition $D + 2\Delta_n = n\pi/k_{\perp}$, where Δ_n gives the shift in the potential well width that allows to reproduce the charge spilling out of the n -th state. This introduces an energy dependence of the shift that in principle could be obtained by fitting the quantum well energies to experimental data or to the energy level distribution resulting from a first principle calculations. Previous works [21, 22] have shown that the energy spectrum is very sensitive to the position of the barriers and relatively insensitive to the barriers height. One can take an average d value obtained by simply imposing the film to be neutral with an electron charge of size d and a positive charge of size D . This leads to the following expression for the effective film thickness:

$$d = \frac{D}{G(m_F)} \quad (4.7)$$

with:

$$G(m_F) = \frac{3m_0}{2m_F} \left[1 - \frac{(m_0 + 1)(2m_0 + 1)}{6m_F^2} \right] \quad (4.8)$$

where $m_F = k_F d / \pi$, k_F is the Fermi wavevector that is related to the ion electron density n_0 by the relation $k_F^3 = 3\pi^2 n_0$, and m_0 is the integer part of m_F . Notice that, once D and the positive charge density n_0 are given, the d value can be determined unambiguously¹. For large D , when $m_0 \simeq m_F$, one obtains $d = D + 3\pi/4k_F$, which shows that the spilling out of the charge in this limit is proportional to the Fermi wavelength.

With respect to the classical description of thin film optical and transport properties [23, 24], where surface effects are expressed as boundary conditions on the electron distribution function in terms of the mean free path and the fraction of the electrons scattered specularly at the surface, this theory has the surface effects incorporated as boundary conditions in the one-electron hamiltonian, the only parameters being the film size and the Fermi energy.

The dielectric response of the Wood and Ashcroft model film, is described in section 1.11. This new model differs from the plasma one, adopted in previous studies, in that: (i) it has a tensor character with $\epsilon_{xx} = \epsilon_{yy} \neq \epsilon_{zz}$, (ii) the plasma frequency depends upon the film density, which changes as a function of the film thickness, (iii) through the double sum in the second member it accounts for transitions between lateral sub-bands, whose probability amplitude is expressed by the momentum matrix element between the one electron wavefunctions (4.5). In section 1.11 the occurrence of kinks both in the electron density (figure 1.12) and in the dielectric function (figure 1.14) has been demonstrated to be a consequence of the quantization of the energy levels. The occurrence of these quantum size oscillations has been pointed out by several authors specially with reference to the electron density, the total electronic energy and the film electrical conductivity [19, 25, 26].

¹For the calculation of the film charge density see for example reference [19].

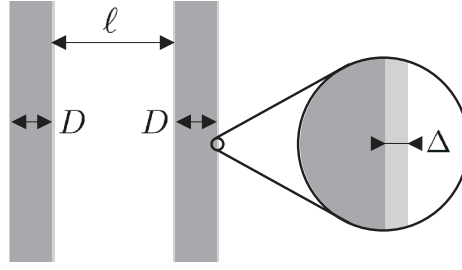


Figure 4.2: Two identical interacting thin films, definition of ℓ , D and the electronic spill-out Δ .

The expression of the force per unit area F at $T = 0^\circ$ K in the configuration illustrated in the inset of figure 4.2 can be obtained by extending the results relative to isotropic films discussed in section 2.3.4, to the case of films with anisotropic dielectric tensor illustrated in section 2.3.5. For this specific geometry, one obtains:

$$F = -\frac{\hbar}{2\pi^2} \int_0^\infty k dk \int_0^\infty d\omega \gamma(\omega) \left[\frac{Q_{TM}(i\omega)^2}{1 - Q_{TM}(i\omega)^2} + \frac{Q_{TE}(i\omega)^2}{1 - Q_{TE}(i\omega)^2} \right]$$

$$Q_{TM} = \frac{\rho_{TM}(1 - e^{-2\gamma_{TM}D})}{1 - \rho_{TM}^2 e^{-2\gamma_{TM}D}} e^{-\gamma\ell} \quad Q_{TE} = \frac{\rho_{TE}(1 - e^{-2\gamma_{TE}D})}{1 - \rho_{TE}^2 e^{-2\gamma_{TE}D}} e^{-\gamma\ell} \quad (4.9)$$

$$\rho_{TM} = \frac{\gamma_{TM}(\omega) - \gamma\epsilon_{xx}(\omega)}{\gamma_{TM}(\omega) + \gamma\epsilon_{xx}(\omega)} \quad \rho_{TE} = \frac{\gamma_{TE}(\omega) - \gamma(\omega)}{\gamma_{TE}(\omega) + \gamma(\omega)} \quad (4.10)$$

$$\begin{aligned} \gamma(\omega) &= \sqrt{k^2 - \frac{\omega^2}{c^2}} & \gamma_{TE}(\omega) &= \sqrt{k^2 - \frac{\omega^2}{c^2} \epsilon_{xx}(\omega)} \\ \gamma_{TM}(\omega) &= \sqrt{\left(\frac{k^2}{\epsilon_{zz}(\omega)} - \frac{\omega^2}{c^2} \right) \epsilon_{xx}(\omega)} \end{aligned} \quad (4.11)$$

This equation, which to the author knowledge has never been used previously for vacuum fluctuation force calculations, deserves a few comments. (i) it has been obtained by considering the zero point energies associated with the electromagnetic modes of films of finite thickness under the assumption that the dielectric permittivity is represented by an anisotropic diagonal tensor. It differs from the original Lifshitz formula both because it depends upon the film size D (i.e. has been obtained with the electromagnetic field boundary conditions appropriate to a finite size film and not for a semi-infinite system) and for the presence of the anisotropic permittivity. (ii) the force goes to zero as the film size D vanishes, since both Q_{TM} and Q_{TE} go to zero². (iii) to calculate the force one has to determine the frequency dependence of the dielectric tensor using the eigenvalues and the wavefunctions of the quantized film, coming to the result (1.148). In the following the force per unit area calculated for the dielectric tensor (1.148) appropriate to the film is denoted by F_Q , while F_P indicates the force calculated using the isotropic plasma model i.e. $\epsilon_{xx} = \epsilon_{yy} = \epsilon_{zz}$. To evaluate the importance of the size quantizations, one should compare F_Q , calculated for given film thickness D and ion density n_0 at different distances ℓ , with F_P calculated in the same configuration. Figure 4.3 (a) displays curves of F_Q calculated at fixed Ω_p and ℓ values as a function of the film thickness. The Ω_p values correspond to free carrier densities of heavily doped semiconductors. These

²If D goes to zero also $n(d)$ goes to zero, as shown by equation (1.142) because $m_0 = m_F = 1$. In this way both the transvers and longitudinal components of the dielectric tensor go to 1.

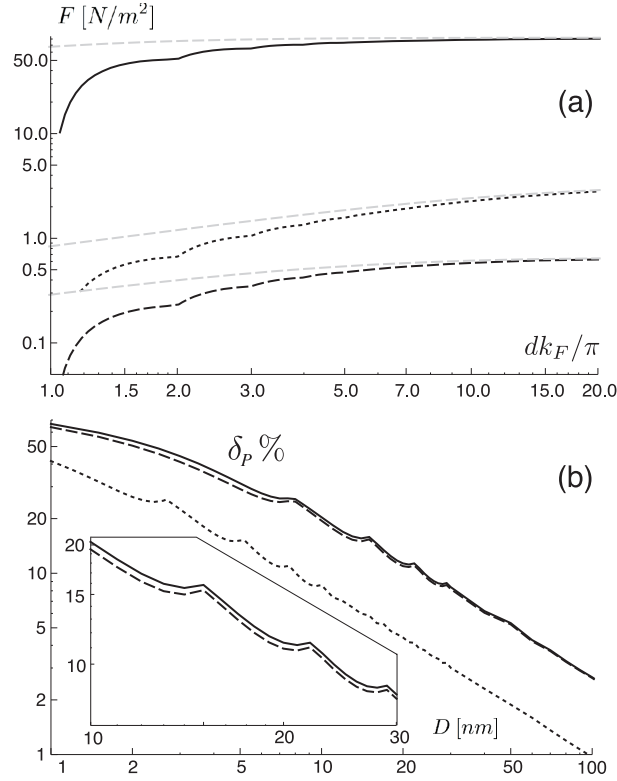


Figure 4.3: (a) Black lines represent F_Q as a function of the film thickness for different film densities and separations, $\Omega_p = 10^{14}$ rad/sec and $\ell = 10$ nm (continuous line), $\Omega_p = 5 \cdot 10^{14}$ rad/sec and $\ell = 50$ nm (dotted line), $\Omega_p = 10^{14}$ rad/sec and $\ell = 50$ nm (dashed line). The gray dashed lines represent F_P for the same parameters. (b) relative percent difference between the force with and without QSE.

systems are those which display higher modifications. Distances ℓ ranging from 10 to 50 nm have been considered. For comparison the curves of the corresponding F_P values are reported. One can see significant differences between the two models for d values that are of the order of few multiple integer of the half of the Fermi wavelength. The curves show quantum size oscillations with the expected periodicity $k_F d = n\pi$ superimposed over a regularly increasing behaviour. For thick films the results of the isotropic plasma model are recovered. The size induced modifications are better illustrated by the quantity:

$$\delta_P = \frac{F_P - F_Q}{F_P} \quad (4.12)$$

which gives the relative variation of the force with respect to the isotropic plasma model. Fig 4.3 (b) presents plots of δ_P for the cases under consideration. It is seen that the size induced modifications are very large, ranging from 50% to 10%, for films of nanometric thickness even at distances ℓ of several nanometers.

One can see similar modifications in the theoretical results obtained by keeping the film size constant and changing its density. This is illustrated in figs. 4.4 (a) and (b) which display force curves obtained as a function of plasma frequency. Notice that even for typical metallic densities (Ω_p of the order of $10^{15} \div 10^{16}$ rad/sec) the deviations from the plasma model are quite significant and can be of the order of several percents for film separation distance ℓ ranging from 10 to 50 nm. In a previous study on the Casimir effect for metal and semi-

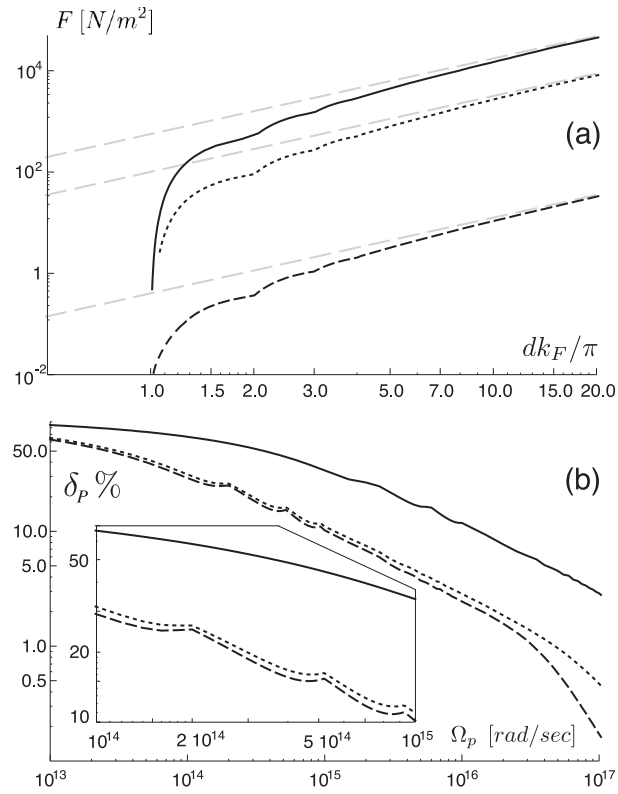


Figure 4.4: (a) Black lines represent F_Q as a function of the plasma frequency for different film thicknesses and separations, $D = 1$ nm and $\ell = 10$ nm (continuous line), $D = 5$ nm and $\ell = 10$ nm (dotted line), $D = 5$ nm and $\ell = 50$ nm (dashed line). The gray dashed lines represent F_P for the same parameters. (b) relative percent difference between the force with and without QSE.

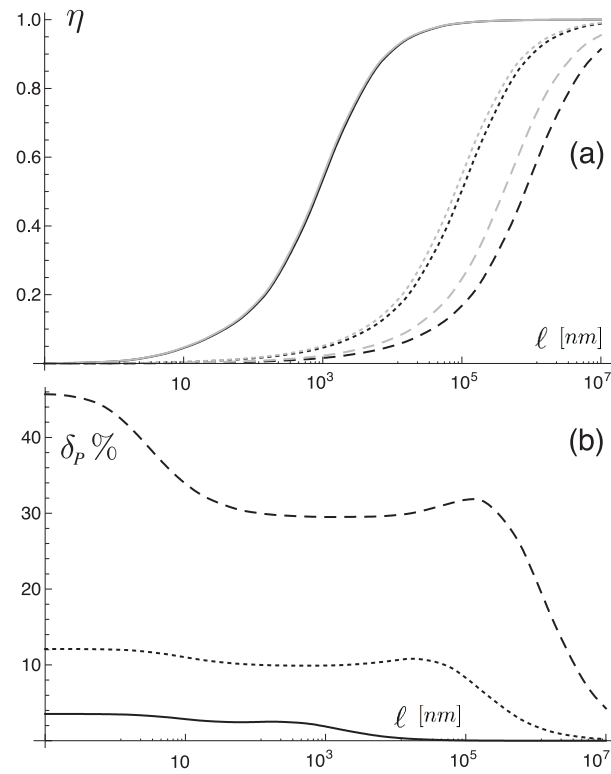


Figure 4.5: (a) Black lines represent η_Q as a function of the films distance ℓ for different densities and thicknesses: continuous line $\Omega_p = 10^{16}$ rad/sec $D = 5$ nm, dashed line $\Omega_p = 10^{15}$ rad/sec $D = 1$ nm, dotted line $\Omega_p = 10^{15}$ rad/sec $D = 5$ nm. Gray lines represent η_P as a function of ℓ for the same parameters. (b) Percent difference (4.12) for the three cases of the (a) plot.

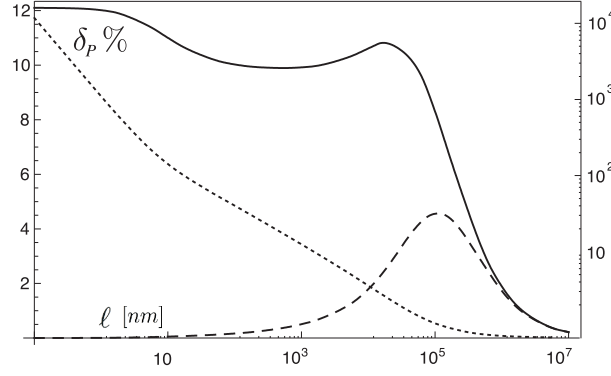


Figure 4.6: Continuous line represents the percent relative difference δ_P , dashed line represents $\eta_P - \eta_Q$, dotted line represents the ratio $1/\eta_P$ for the case $\Omega_p = 10^{15}$ rad/sec $D = 5$ nm

conductor slabs it has been pointed out that the Casimir force can be considerably reduced by decreasing the slab thickness [6]. It is interesting to see whether quantum confinement effects are important in determining the reduction at large ℓ values. To this end figure 4.5 (a) shows the reduction factors $\eta_P = F_P/F_{CAS}$ and $\eta_Q = F_Q/F_{CAS}$ [27], where

$$F_{CAS} = -\frac{\hbar c \pi^2}{240 \ell^4} \quad (4.13)$$

is the force between perfectly reflecting mirrors at separation distance ℓ [28]. The films have a nominal thickness $D = 1 \div 5$ nm and different electron densities. One can see that the effects of the quantum confinement tend to decrease the reduction factor over a large distance range. The correction depends upon both the film size and the electron density, being larger for smaller values of these quantities. Figure 4.5 (b) plots the calculated relative difference δ_P as a function of ℓ for the same cases. It shows that relative variations of the force remain significant even at very large distances thus giving sizeable reduction to the force in the large ℓ Casimir regime. Notice that δ_P does not show a monotonous behaviour as a function of ℓ : it decreases regularly at short distances, it remains constant in a large range of ℓ values and, before going to zero for very large distances, it shows a local maximum that shifts at higher ℓ values upon decreasing the thickness and/or the electron density. This behaviour is quite typical and can be reproduced in several cases. To understand this behaviour one should notice that δ_P can be written as

$$\delta_P = 1 - \frac{\eta_Q}{\eta_P} \quad (4.14)$$

Figure 4.6 shows the typical behaviour of the numerator and of $1/\eta_P$ as a function of ℓ . The first curve decreases more or less regularly upon increasing the film separation. The change in the slope arises from different inverse power behaviour of the short range (dispersion force) compared to long range Casimir force [29, 30, 4, 31]. The second curve shows a regular increase at short distances, it has a maximum and goes to zero at large ℓ values. Therefore the maximum in the δ_P corresponds to the range of ℓ values where the reduction of the Casimir force is larger compared to the ideal case, while the plateau is due to the combined effect of the decrease of $\eta_P - \eta_Q$ and the rise of the $1/\eta_P$ curve.

4.2.2 Finite depth well

In this section the PBM will be improved introducing a finite well model for the one-electron potential along the surface normal. Notice that the use of a soft confining potential like a

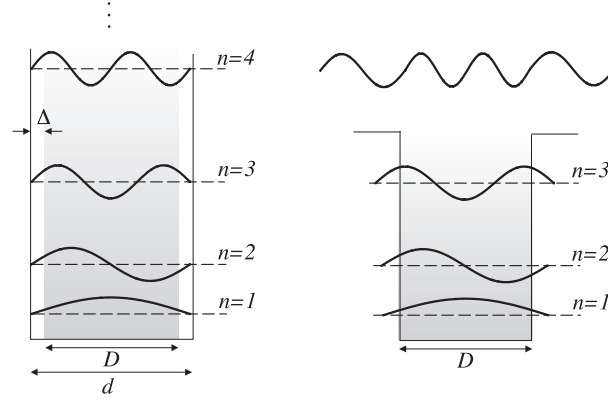


Figure 4.7: Comparison between the infinitely deep quantum well with artificial spill-out and the finite quantum well with natural spill-out.

finite well, besides being closer to the real shape of the one-electron potential as determined by first principles calculations [32, 33, 34, 35], allows for a better treatment of QSE for films of small thickness, since it does not introduce a priori a distance within which electrons are to be confined and leaves the Fermi energy free to oscillate with the film size [8, 9], a feature that is not present in the PBM model, where the Fermi energy is kept equal to the bulk value.

The electron gas will be again described by an independent particle model with the one electron potential $V(z)$. The electron energies are given by

$$E_{\mathbf{k}_{\parallel},n} = \frac{\hbar^2}{2m} \mathbf{k}_{\parallel}^2 + E_n \quad (4.15)$$

i.e. they are described by the continuous quantum number \mathbf{k}_{\parallel} and by the discrete sub-band index n coming from the quantization of the perpendicular wavevector, m being the electron mass. Notice that now E_n can not be determined analytically, a numerical evaluation of the one-dimensional well potential must be setup. The film wavefunctions are given by

$$\psi_{\mathbf{k}_{\parallel},n}(\mathbf{r}_{\parallel}, z) = \frac{1}{\sqrt{A}} e^{i\mathbf{k}_{\parallel} \cdot \mathbf{r}_{\parallel}} \phi_n(z) \quad (4.16)$$

here A is the surface area, again \mathbf{k}_{\parallel} and \mathbf{r}_{\parallel} are two dimensional wavevectors parallel to the surface. Also $\phi_n(z)$ must be evaluated numerically and normalized independently from the parallel component, in particular the functions $\phi_n(z)$ are the solutions of the equation

$$\left\{ \frac{\hbar^2}{2m} \frac{\partial^2}{\partial z^2} + V(z) \right\} \phi_n(z) = E_n \phi_n(z) \quad (4.17)$$

with the proper boundary conditions. The Fermi energy E_F is obtained through the *aufbau* procedure i.e. by arranging the eigenvalues in ascending numerical order and counting until the number of states needed to accommodate all the electrons in the film is reached. This procedure leads to a Fermi energy that depends upon the film size and is generally different from the bulk value [9]. This can be understood by noting that, to ensure charge neutrality, the number of electrons and the number of ions per unit area have to be equal. The electron density $n(z)$ can be simply obtained from the wavefunctions:

$$n(z) = \frac{1}{2\pi} \sum_{m=1}^{m_0} \left(\frac{2mE_F}{\hbar^2} - E_n \right) |\phi_m(z)|^2 \quad (4.18)$$

where m_0 is the label of the last occupied state, while the ion density is simply given by $n_0 = k_{FB}^3/3\pi^2$, where k_{FB} is the bulk Fermi wavevector. By integrating the densities along the z axis and imposing that both give the same number of charges per unit area, one gets the relation

$$\frac{1}{2\pi} \sum_{m=1}^{m_0} \left(\frac{2mE_F}{\hbar^2} - E_n \right) = n_0 D \quad (4.19)$$

which, for finite D values, is generally not satisfied if one replaces E_F with its bulk counterpart $E_{FB} = \hbar^2 k_{FB}^2/2m$. In the case of the PBM model this equation is not satisfied, since one assumes that the Fermi level is the same as in the bulk. To obtain charge neutrality one has to impose the additional condition that the electron density be confined on a length d larger than D . As discussed in the previous section, this artificially introduces the electronic charges spill-out but has the consequence that the average electronic density is lower than the ion density [8, 19, 25].

For the purpose of the present study the potential is assumed to be a finite well $V(z) = -V_0$ inside the film and zero outside. For such finite well model (FWM) the energies of the bound states can be written as

$$E_{\mathbf{k}_{\parallel},n} = \frac{\hbar^2}{2m} \mathbf{k}_{\parallel}^2 + \frac{\hbar^2}{2m} k_{zn}^2 - V_0 \quad (4.20)$$

where k_{zn} are the quantized transverse wavevectors. They are obtained from the equation giving the condition for the existence of bound states in a quantum well [36]:

$$k_{zn} = \frac{n\pi}{D} - \frac{2}{D} \sin^{-1} \left(\frac{k_{zn}}{k_0} \right) \quad (4.21)$$

with $k_0 = \sqrt{2mV_0}/\hbar$. Notice that the first term at the second member is the value of the transverse wavevector for an infinite well model of size D and the second term goes to zero as V_0 goes to infinity. This implies that, for given film size and number of electrons, the Fermi energy referred to the well bottom is higher in the infinite well model. Notice that when V_0 goes to infinity one does not recover the PBM, since the Fermi energy is varied with respect to the bulk value in order to satisfy the charge neutrality condition (4.19). Figure 4.7 illustrates the difference: in the PBM the electronic charge density is confined within a distance $d = D + \Delta$, larger than the size of the ionic charge distribution, to allow for the electronic charge spill-out and to ensure global neutrality for $E_F = E_{FB}$. In the FWM the charge spill-out results naturally from the behaviour of the single particle states while the charge neutrality is achieved by varying the Fermi energy with respect to the bulk value. Obviously in the limit of infinitely deep well (hereafter indicated as IWM) the electronic charge turns out to be entirely localized within the length D and the Fermi energy is strongly increased compared to its bulk value.

Once the electron energies and wavefunctions have been obtained, one can calculate the dielectric tensor from the expression [18]:

$$\epsilon_{\alpha\alpha}(\omega) = 1 - \frac{\omega_p^2}{\omega^2} - \frac{8\pi e^2}{Adm^2\omega^2} \sum_{\mathbf{k}_{\parallel},n} \sum_{\mathbf{k}'_{\parallel},n'} (E_{\mathbf{k}_{\parallel},n} - E_{\mathbf{k}'_{\parallel},n'}) \frac{f(E_{\mathbf{k}_{\parallel},n}) |\langle \psi_{\mathbf{k}_{\parallel},n} | \hat{p}_{\alpha} | \psi_{\mathbf{k}'_{\parallel},n'} \rangle|^2}{(E_{\mathbf{k}_{\parallel},n} - E_{\mathbf{k}'_{\parallel},n'})^2 - \hbar^2\omega^2} \quad (4.22)$$

a generalization of the Wood and Ashcroft model of section 1.11 where $\alpha = x, y, z$ labels the cartesian component of the tensor, \hat{p}_{α} indicates the component of the electron linear momentum, $\omega_p = \Omega_p n(D)/n_0$ is the plasma frequency of the quantized electron gas ($\Omega_p = \sqrt{4\pi e^2 N_0/m}$ is the free electron plasma frequency and $n(D)$ is the average electron density of the film) and $f(E_{\mathbf{k}_{\parallel},n})$ is the occupation factor of the $(\mathbf{k}_{\parallel}, n)$ state. In the IWM and the FWM $\omega_p = \Omega_p$, while in PBM $n(D)$ is smaller than n_0 , since the electronic charge is distributed over a larger distance than the ionic charge (see figure 4.7).

In order to illustrate QSE at different densities and potential depths, the elements *Al*, *Ag* and *Cs* will be taken, their radius per electron in Bohr units r_s/a_0 is equal to 2.07, 3.02 and 5.62 respectively. The value of the potential depth is obtained by summing the metal work function W with the calculated Fermi energy. Figure 4.8 reports the calculated Fermi energies as a function of the product between the Fermi wavevector and the film thickness. This allows to better point out the oscillations and the cusps arising from the crossing of the Fermi energy by new subbands upon varying the film size, as in figure 1.12 (b). The figure shows the V_0 value appropriate to the bulk and to the large D limit, the results for the IWM are also plotted. The comparison allows to illustrate the effects of the potential softening. In agreement with previously published results [32, 9, 38] one finds that:

- E_F is systematically larger than E_{F_B} and goes to the bulk value as D goes to infinity. The difference is more pronounced and the bulk limit is achieved at larger size in the low density systems, as it is clearly shown by the comparison between *Cs* and *Al* curves;
- As expected from the discussion of the previous section, the softening of the potential leads to less pronounced deviations from the bulk values. Because of the stronger electron confinement, the IWM has a larger Fermi energy than the FWM;
- The cusps correspond to integer values of half of the Fermi wavelength in the IWM case. This feature is only approximately verified for the FWM.

As a consequence of size quantization, the z -component of the dielectric tensor is expected to go to a finite value $\epsilon_{zz}(0)$ as the frequency goes to zero. This value increases proportionally to D^2 in the large size limit [18]. Figure 4.9 shows the quantity $\epsilon_{zz}(0)/D^2$ for the three cases under study. Again the results show the cusps due to the filling of new subbands as D increases, as in figure 1.14 (b). The asymptotic limit is obtained for $k_F D/\pi$ of the order of $5 \div 6$ in the three cases. Significant differences appear in the large D behaviour when the FWM results are compared with those from the IWM model: $\epsilon_{zz}(0)$ is larger for finite wells of small size, while it is smaller at high D values. The convergence to the asymptotic limit is considerably slower for the infinite well, specially in the low density metals. This behaviour reflects the differences in the distribution of the eigenvalues of equation (4.17). For the infinite well there are infinite bound states whose energy scales like n^2 , see equation (4.21), and the separation between two successive levels increases linearly with n . Such behaviour is not present in the FWM, for which equation (4.17) has a finite number of eigenvalues corresponding to bound states and a continuum spectrum at positive energies. The PBM curve takes values closer to the FWM than to the IWM. This is primarily a consequence of the fact that PBM allows for electron charge spill-out, while in IWM the electron distribution is confined within D i.e. it has the same size of the positive charge.

The changes in the force caused by the size quantization will be discussed with respect to the results obtained by using the isotropic continuum plasma model, where the z -component of the film dielectric tensor is equal to the planar components. Again the symbol F_Q indicates the force per unit area calculated for the quantized film, while F_P is the force per unit area calculated in the isotropic plasma model. To better illustrate the results, in figures 4.10 and 4.11 the quantity (4.12) has been plotted as a function of the separation distance ℓ for films of 1 and 5 nm thickness respectively. Each figure displays the results for the three cases under study and compares the finite well with the IWM at the same density. This allows one to point out the modifications caused by the potential softening. Also show the curves appropriate to the PBM are displayed. In agreement with the previous findings one observes that

- QSE tend to reduce the intensity of the force;
- the reduction is more significant at low density (*Cs*) than at high density (*Al*);

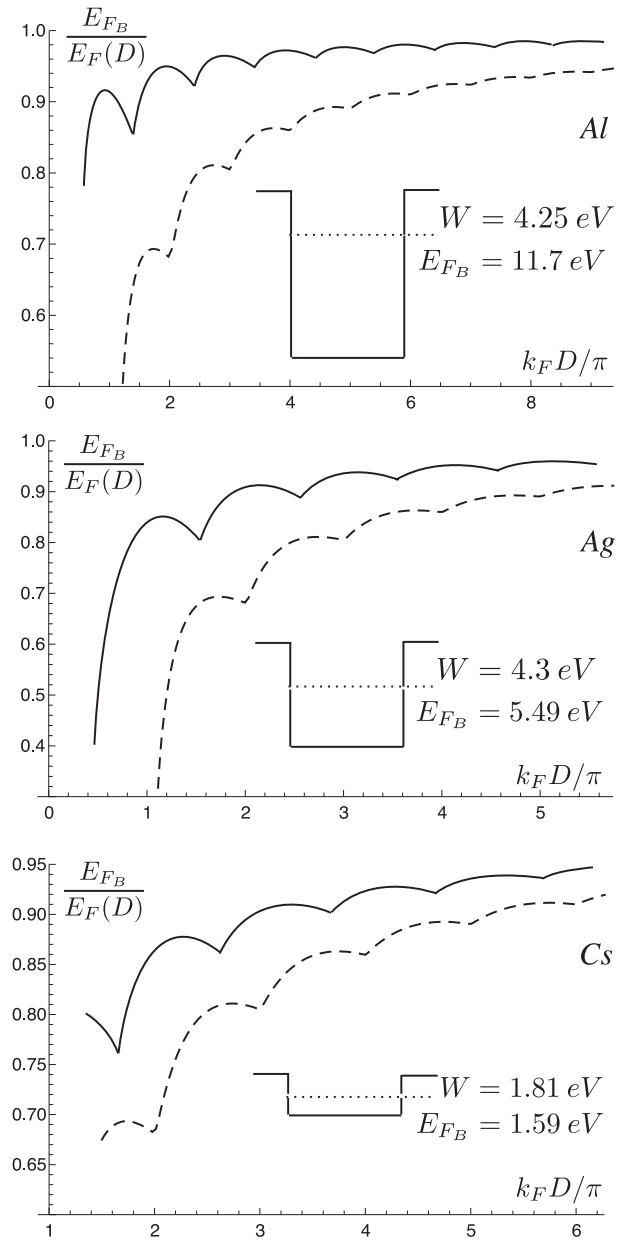


Figure 4.8: Fermi energy normalized to its bulk value for *Al*, *Ag* and *Cs* using the IWM (dashed line) and the FWM (continuous line). The bulk Fermi energies and the work functions have been taken from reference [37].

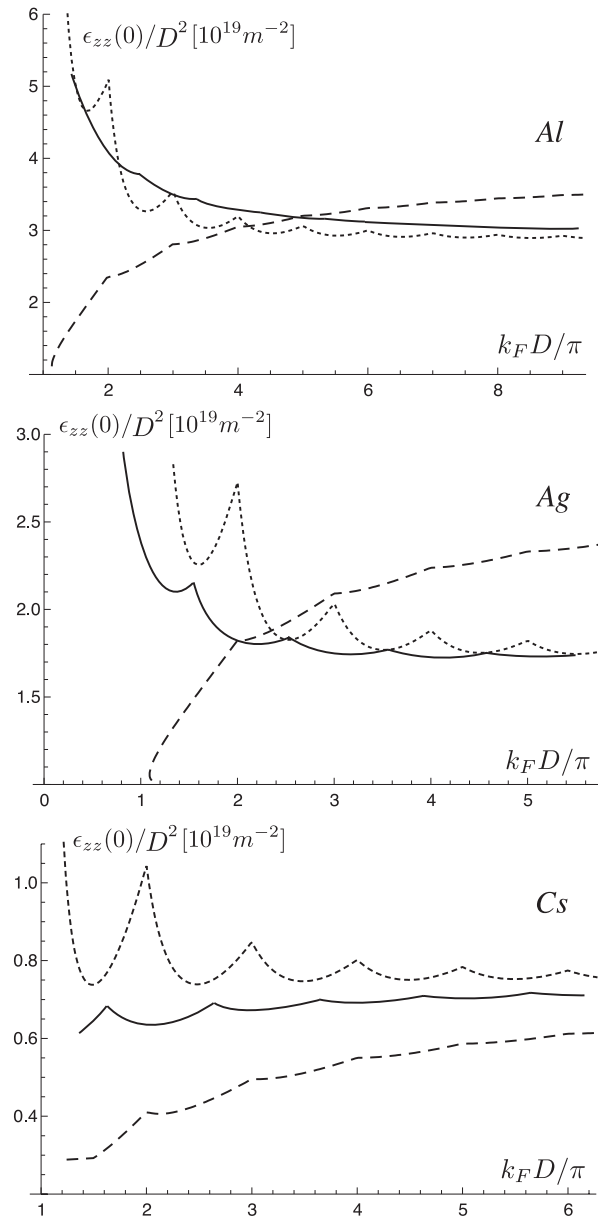


Figure 4.9: Static value of the zz component of the dielectric tensor for *Al*, *Ag* and *Cs* using the IWM (dashed lines), the PBM (dotted lines) and the FWM (continuous lines).

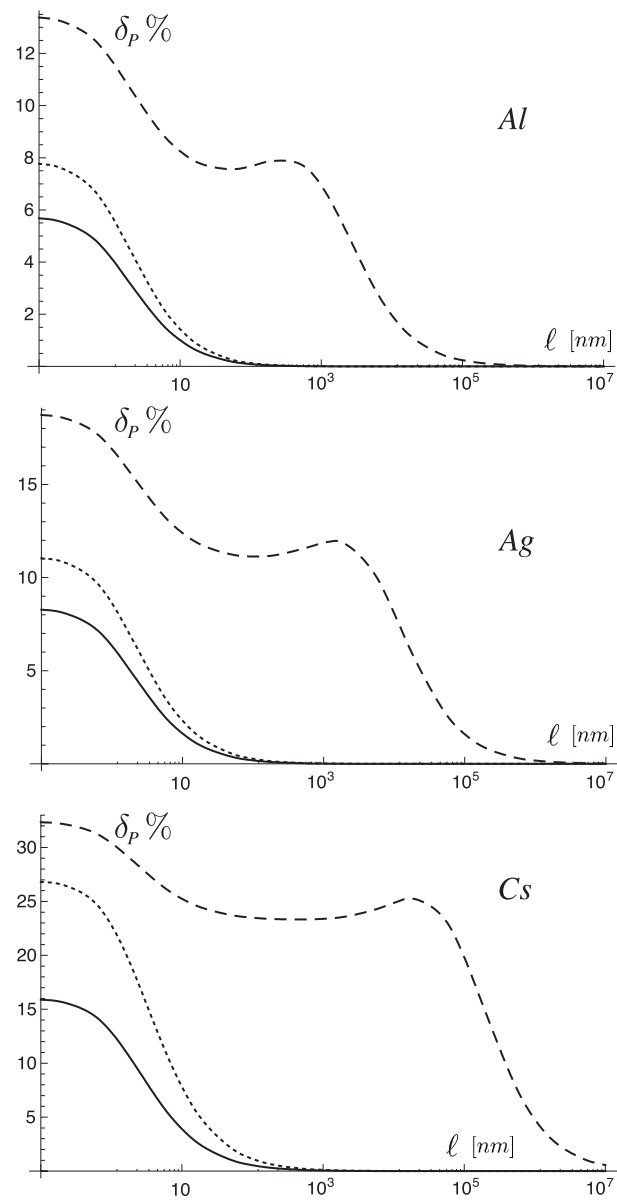


Figure 4.10: Relative percent difference for the force between two identical films of thickness $D = 1$ nm as a function of the films separation ℓ , for *Al*, *Ag* and *Cs*. Dashed lines have been obtained using the PBM, dotted lines have been obtained using the IWM and the continuous lines represents the finite well results.

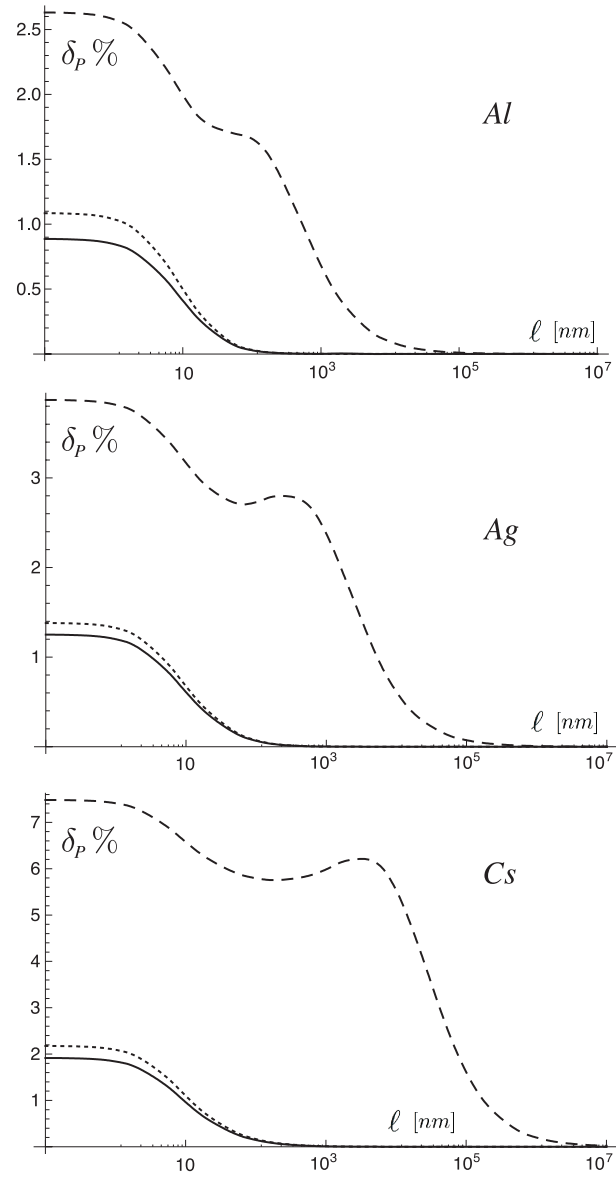


Figure 4.11: Relative percent difference for the force between two identical films of thickness $D = 5$ nm as a function of the films separation ℓ , for *Al*, *Ag* and *Cs*. Dashed lines have been obtained using the PBM, dotted lines have been obtained using the IWM and the continuous lines represents the finite well results.

- it may be considerably higher than 10% for 1 nm thickness and reduce to few per cent at 5 nm;
- it can be appreciable over a distance interval up to $10 \div 50$ nm.

The most important conclusion that can be drawn from the figures is that the potential shape is important and can lead to a substantial modifications of the quantum size effects both at small and at large distances. The models which confine the electronic charge tend to overestimate the force reduction induced by size quantization. The curves for the PBM show large force reduction (greater δ_P values) over a wide interval of distances. On passing to the IWM case one notice that the removal of the constriction that the Fermi energy be equal to the bulk value, still keeping an infinitely deep potential, leads to smaller δ_P values and to a more rapid decay of the curves at large distances. Reducing the well depth to finite values has a similar effect: it causes a general decrease of δ_P and a narrowing of the distance interval over which QSE are appreciable. This also implies that any increase of the confining potential depth at fixed ion density leads to higher δ_P values and to more significant QSE. The large values taken by δ_P in the PBM case do not arise from the charge confinement only, since, as pointed out before, the constraint on the charge distribution is weaker than in the IWM. To a large extent they are a consequence of the plasma frequency normalization caused by the decrease in the average electron charge density that it is necessary in order to achieve global charge neutrality [19, 25]. In the isotropic plasma model one takes $\omega_p = \Omega_p$. In the PBM this value is obtained only at large film thickness. Neglecting this normalization i.e. taking the free electron plasma frequency in the parallel components of the dielectric tensor (but not in ϵ_{zz}) would lead to δ_P values closer to the well potential models.

4.2.3 The generalized particle in a box model

To introduce intraband absorption one has to modify the dielectric tensor in a way that allows to include relaxation time effects in the parallel components and to recover the Drude behaviour in the large D limit. This cannot be done by simply introducing an imaginary part of the frequency ω , since this violates the continuity equation locally [39, 40]. The appropriate recipe is to replace into equation (4.22) ω^2 with $\omega(\omega + i2\pi/\tau)$, where τ is the relaxation time. For the parallel components this leads to the Drude dielectric function

$$\epsilon(\omega) = 1 + \frac{\omega_p^2}{\omega(\omega + i\gamma)} \quad (4.23)$$

where $\gamma = 2\pi/\tau$. Figure 4.12 shows the comparison between ϵ_{zz} for the PBM with and without the intraband absorption. The inclusion of a finite τ eliminates the divergencies corresponding to the transverse (optical) modes, the static value of the real part of ϵ_{zz} remains unchanged. The results of the calculation of the force per unit area with the modified dielectric tensor are displayed in Figures 4.13 and 4.14 for 1 nm and 5 nm films. The quantity

$$\delta_D(\gamma) = \frac{F_D(\gamma) - F_{QD}(\gamma)}{F_D(\gamma)} \quad (4.24)$$

is reported, where F_D is the force calculated using the bulk Drude model with a given relaxation time and F_{QD} in the one obtained by the calculation with the same relaxation time and with size quantization included. For $\tau \rightarrow \infty$ one recovers the plasma model so that $\delta_P = \delta_D(0)$. The calculations have been performed assuming a finite potential well and for two values of the relaxation frequency γ . The chosen values correspond approximately to those reported for the metals under consideration [37]. The figures show a comparison with the curves obtained by the plasma model with QSE. It is clear from these results that the main effect of the inclusion of intraband absorption is to increase δ i.e. to increase the

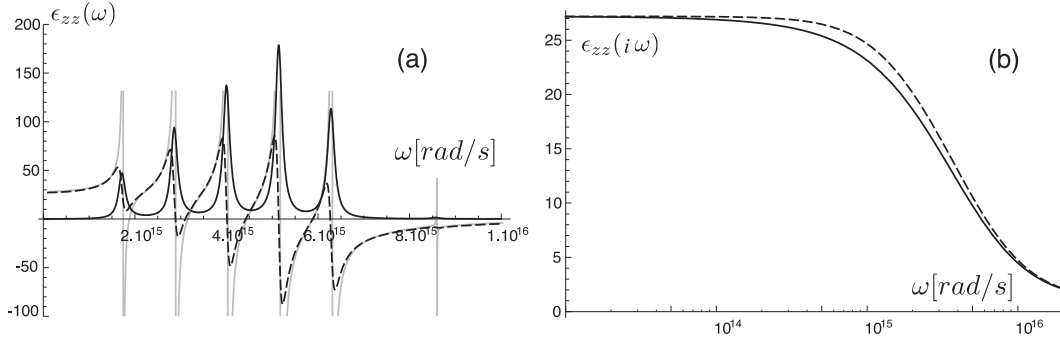


Figure 4.12: zz components of the dielectric tensor (a) and their London transforms (b) for the PBM, gray line represents the $\tau \rightarrow \infty$ case, black lines represent the real and imaginary parts in the case $\tau = 10^{-14}$ sec. The film thickness is $D = 1$ nm.

difference with respect to the calculations with the bulk dielectric function. The smaller is the relaxation time the larger is the reduction of the force. The effect is qualitatively the same in the three metals under study, but it depends upon the well depth and the film size. The influence of the shape of the potential is illustrated in figure 4.15 showing the curves of Ag films at a given relaxation frequency for the various models. Again it should be noticed that the FWM gives the lower δ_D values. The PBM results show large δ_D values over a very wide interval of distances. To a large extent this behaviour has to be imputed to the renormalization of the plasma frequency. Figure 4.16 shows typical curves of δ_D as a function of the film thickness for different values of the relaxation frequencies at a given separation distance of 5 nm. As expected δ_D decreases with D , but the slope at large thicknesses (D of the order of $10 \div 50$ nm) depends significantly upon the relaxation frequency.

4.3 Isolated metallic film

The behaviour of the dispersion force acting on an isolated metallic films has been extensively studied in chapter 3. The quantum confinement is expected to modify the dispersion force on the film boundaries in the same way it modifies the force between films. For the purpose of this study, only the PBM will be considered, but the following considerations hold for the other models too. To study the QSE on the vacuum pressure acting on the film boundaries, one has to calculate the vacuum energy with the PBM thickness dependent dielectric tensor and differentiate with respect to the thickness, obtaining an expression for the Lifshitz force very different from the usual one (2.93). Nevertheless this expression can be rearranged in the form:

$$F(d) = -\frac{\hbar}{2\pi^2} \int_0^\infty k dk \int_0^\infty \left[\gamma_3^{TM} \frac{1 - Q_{TM}(i\xi)}{Q_{TM}(i\xi)} C^{TM} + \gamma_3^{TE} \frac{1 - Q_{TE}(i\xi)}{Q_{TE}(i\xi)} C^{TE} \right] d\xi \quad (4.25)$$

this is the ordinary Lifshitz force (2.93) where two new coefficients appear, C^{TM} and C^{TE} , whose expression is:

$$C^{TE} = 1 + \frac{\epsilon'_{xx}}{2(\epsilon_{xx} - 1)} \left[d - \frac{1}{k^2 + \omega^2/c^2 \epsilon_{xx}} (\gamma^2 d + 2\gamma) \right]$$

$$C^{TM} = 1 + \frac{\frac{d}{2} \left(\frac{\gamma^{TM}}{\epsilon_{xx}} \epsilon'_{xx} - \frac{\epsilon_{xx}}{\epsilon_{zz}^2} \epsilon'_{zz} k^2 \right) - \frac{1}{2\gamma^{TMz}} \left(\frac{\gamma^{TM}}{\epsilon_{xx}} \epsilon'_{xx} - \frac{\epsilon_{xx}}{\epsilon_{zz}^2} \epsilon'_{zz} k^2 \right) (d\gamma^2 \epsilon_{xx}^2 + 2\epsilon_{xx} \gamma) + 2\epsilon'_{xx} \gamma}{\epsilon_{xx} \left(\frac{k^2}{\epsilon_{zz}} + \frac{\omega^2}{c^2} \right) - \epsilon_{xx}^2 k^2 - \frac{\omega^2}{c^2} \epsilon_{xx}^2} \quad (4.26)$$

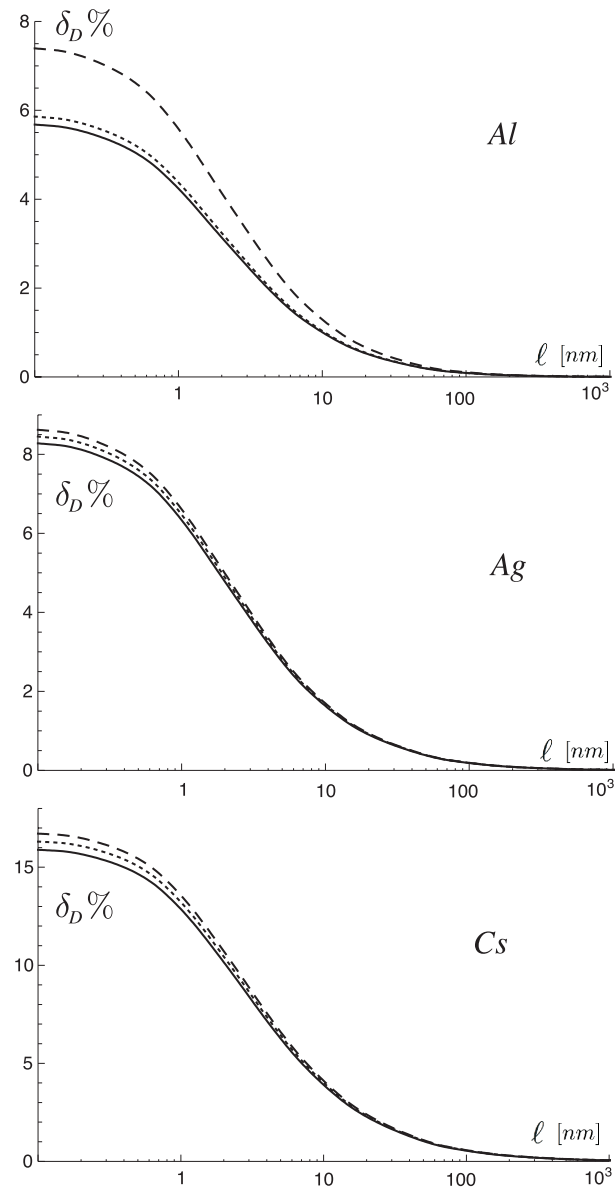


Figure 4.13: Relative percent difference for the force between two identical films of thickness $D = 1$ nm as a function of the films separation ℓ , for *Al*, *Ag* and *Cs*. Continuous lines represent the results for $\gamma = 0$. Dotted lines have been obtained using $\gamma = 5 \times 10^{13}$ rad/s for *Ag* and *Cs* and $\gamma = 10^{14}$ rad/s for *Al*. Dashed lines have been obtained using $\gamma = 10^{14}$ rad/s for *Ag* and *Cs* and $\gamma = 10^{15}$ rad/s for *Al*.

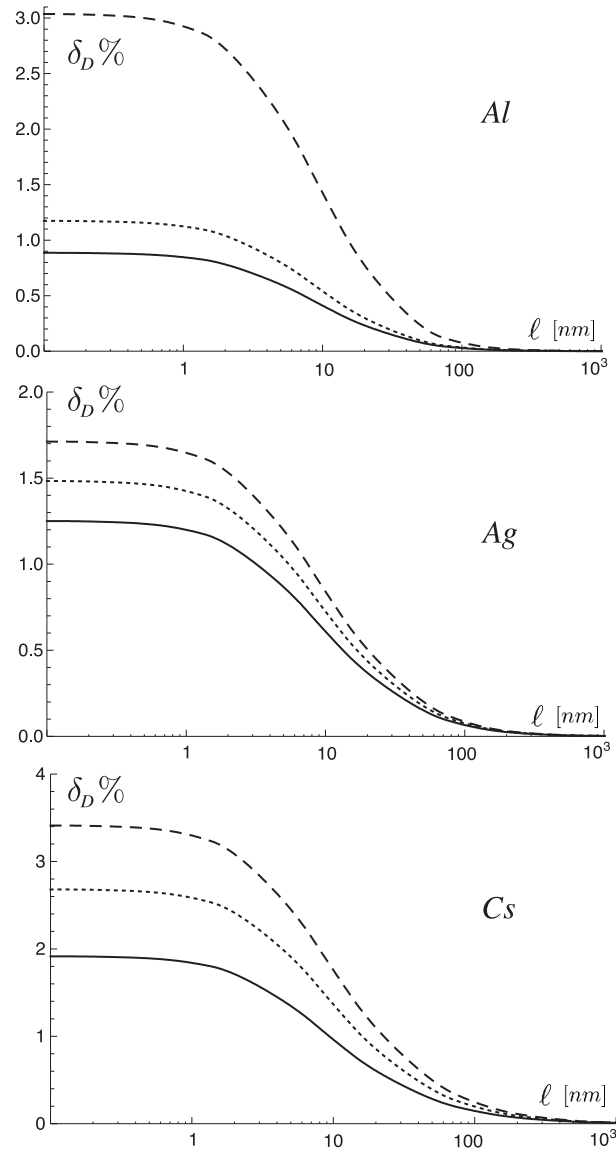


Figure 4.14: Relative percent difference for the force between two identical films of thickness $D = 5$ nm as a function of the films separation ℓ , for *Al*, *Ag* and *Cs*. Continuous lines represent the results for $\gamma = 0$. Dotted lines have been obtained using $\gamma = 5 \times 10^{13}$ rad/s for *Ag* and *Cs* and $\gamma = 10^{14}$ rad/s for *Al*. Dashed lines have been obtained using $\gamma = 10^{14}$ rad/s for *Ag* and *Cs* and $\gamma = 10^{15}$ rad/s for *Al*.

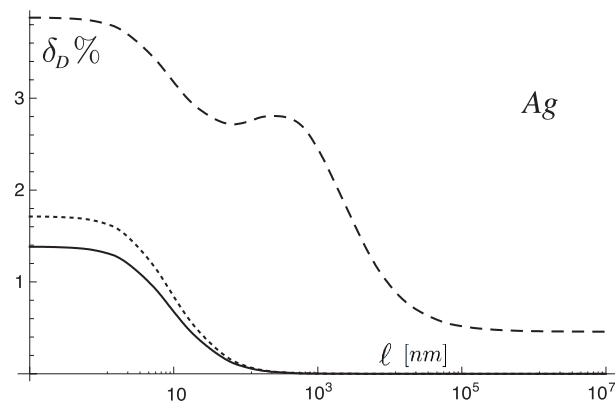


Figure 4.15: Relative percent difference for the force between two identical *Ag* films of thickness $D = 5$ nm as a function of the films separation ℓ , $\gamma = 10^{14}$ rad/s. Continuous lines represent the FWM, dotted line the IWM and dashed line the PBM.

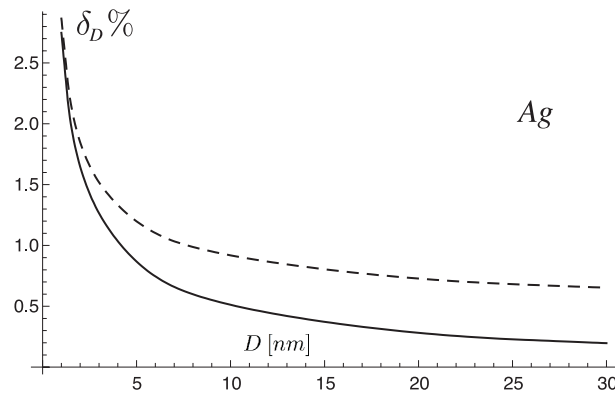


Figure 4.16: Relative percent difference for the force between two identical *Ag* films separated by a distance $\ell = 5$ nm as a function of the films thickness D . Continuous line represents the result for $\gamma = 0$ whereas the dashed line has been obtained with $\gamma = 10^{14}$ rad/s.

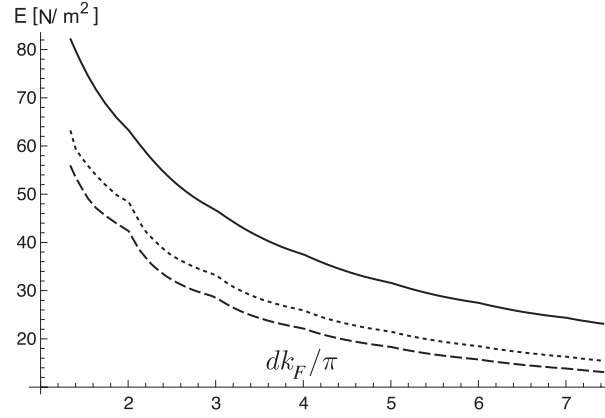


Figure 4.17: Vacuum energy as a function of the film thickness for the PBM. Continuous line $\Omega_p = 10^{16}$ rad/sec, dotted line $\Omega_p = 5 \cdot 10^{15}$ rad/sec and dashed line $\Omega_p = 10^{15}$ rad/sec

here γ^{TM} is given in (2.44), $\gamma = \sqrt{k^2 + \omega^2/c^2}$, ϵ'_{xx} and ϵ'_{zz} are the first derivatives of the dielectric components. Notice that for large thicknesses d , $\epsilon_{zz}(d)$ becomes constant, so that $\epsilon'_{zz} = \gamma^{TE'} = \gamma^{TM'} = 0$ and $C^{TM} = C^{TE} = 1$ recovering the usual force expression for a plasma homogeneous film. With the definitions:

$$C^{TM} = 1 + \tilde{C}^{TM} \quad C^{TE} = 1 + \tilde{C}^{TE} \quad (4.27)$$

the force can also be written as:

$$\begin{aligned} F(D) &= F^L(D) + F^*(D) = \\ &= F^L(D) - \frac{\hbar}{2\pi^2} \int_0^\infty k dk \int_0^\infty \left[\gamma_3^{TM} \frac{1 - Q_{TM}(i\xi)}{Q_{TM}(i\xi)} \tilde{C}^{TM} + \gamma_3^{TE} \frac{1 - Q_{TE}(i\xi)}{Q_{TE}(i\xi)} \tilde{C}^{TE} \right] d\xi \end{aligned} \quad (4.28)$$

where the first term is exactly the Lifshitz force (2.93) and the second one is a correction due to the d dependence of the dielectric tensor. For large film thicknesses, the second term must vanish whereas the first one must reach the value obtained with an isotropic plasma model. Notice also that $F^L(D)$ does not contain any dielectric tensor derivative, so it must be a continuous and non differentiable function. On the other hand $F^*(D)$ contains ϵ'_{zz} that is the derivative of a non differentiable function: in the neighborhood of any cusp $F^*(D)$ is a discontinuous function. Figure (4.17) shows the vacuum energy as a function of the film thickness d : notice that cusps appear as in the case of two interacting film (in the latter case the force was obtained differentiating with respect to the film distance ℓ). For free standing films the force is obtained differentiating with respect to d , but the energy is not differentiable in the neighbourhood of each cusp. As for the case of two interacting films, the QSE are more pronounced for low electron densities. Figure (4.18) (a) shows the absolute value force acting on the film boundaries as function of the film thickness, the separate contributions coming from the two terms of equation (4.28) are also plotted. As previously discussed, F^* is a discontinuous function. Notice also that F^L is larger than the total force, this means that F^* has opposite (positive) sign. A comparison with isotropic plasma model, described in section (3.2.1), is also given. Figure (4.18) (b) shows the force as a function of the plasma frequency. It is seen that the QSE lowers the total strength of the force.

For the IWM and the FWM only ϵ_{zz} is thickness dependent, $\epsilon'_{xx} = 0$, and the coefficients

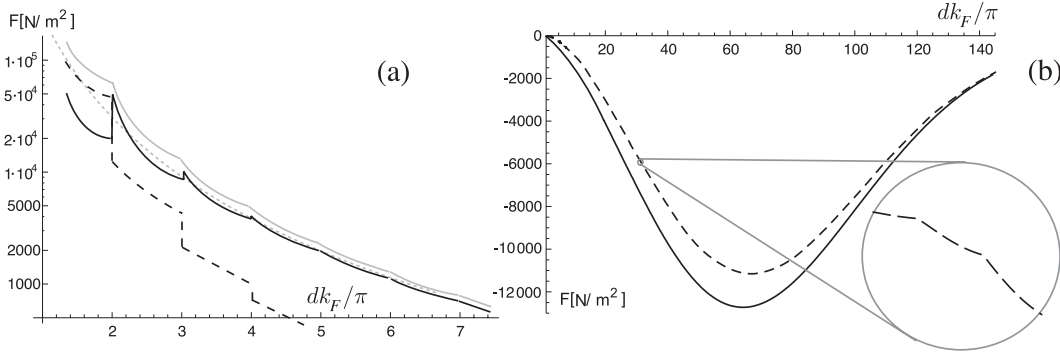


Figure 4.18: (a) Black continuous line represents the absolute value of the force as a function of the film thickness, gray line represents F^L , dashed line represents $|F^*|$ and dotted line represents the force calculated with the homogeneous plasma model. The plasma frequency is $\Omega_p = 10^{15}$ rad/sec. (b) Force as a function of the plasma frequency for a 10 nm film. Continuous line represent the result obtained with the homogeneous plasma model whereas dashed line is obtained with the PBM.

(4.26) simplifies as follow:

$$C^{TE} = 1 \quad C^{TM} = 1 + \frac{-\frac{d}{2} \frac{\epsilon_{xx}}{\epsilon_{zz}^2} \epsilon'_{zz} k^2 + \frac{1}{2\gamma_{TM2}} \frac{\epsilon_{xx}}{\epsilon_{zz}^2} \epsilon'_{zz} k^2 (d\gamma^2 \epsilon_{xx}^2 + 2\epsilon_{xx}\gamma)}{\epsilon_{xx} \left(\frac{k^2}{\epsilon_{zz}} + \frac{\omega^2}{c^2} \right) - \epsilon_{xx}^2 k^2 - \frac{\omega^2}{c^2} \epsilon_{xx}^2} \quad (4.29)$$

only the TM modes are affected by the correction F^* .

4.4 Temperature dependence

The effect of the temperature is to spread the electrons all around the Fermi energy partially occupying the empty states. This electron distribution eliminates the discontinuity at the Fermi level introduced by the *aufbau* procedure at $T = 0^\circ\text{K}$ and the cusps disappear. In the microscopic models previously used to describe the dielectric response of a thin film, the effects of the temperature can be included through the occupation factor $f(E_{\mathbf{k}_{\parallel}, \mathbf{n}})$ appearing in (4.22). In the zero temperature case the occupation factor is given by:

$$f(E_{\mathbf{k}_{\parallel}, \mathbf{n}}) = 2\theta(E_F - E_{\mathbf{k}_{\parallel}, \mathbf{n}}) \quad (4.30)$$

where θ is the Heaviside step function and the factor 2 accounts for the two possible spin polarizations. At non zero temperature, following the Fermi-Dirac distribution, one has:

$$f(E_{\mathbf{k}_{\parallel}, \mathbf{n}}) = \frac{2}{e^{\beta(E_{\mathbf{k}_{\parallel}, \mathbf{n}} - \mu)} + 1} \quad (4.31)$$

where $\beta = 1/k_B T$ and μ is the chemical potential, with the assumption of $k_B T/E_F \ll 1$ one can state that $\mu \simeq E_F$. In the simplest case of an infinite well one gets:

$$\sum_{\mathbf{k}_{\parallel}} f_{\mathbf{k}_{\parallel}, \mathbf{n}} = \frac{Am}{\pi \hbar^2} \frac{1}{\beta} \ln \left[1 + e^{-\beta \left(\frac{\hbar^2 \pi^2 n^2}{2md^2} - E_F \right)} \right] \quad (4.32)$$

For $n \gg m_0$ the exponent diverges and the exponential, together with the logarithm goes to zero, the states far from the Fermi level are again empty. For $n \ll m_0$:

$$e^{\beta E_F} \gg 1 \rightarrow \frac{1}{\beta} \ln \left(e^{\beta E_F} \right) = E_F \quad (4.33)$$

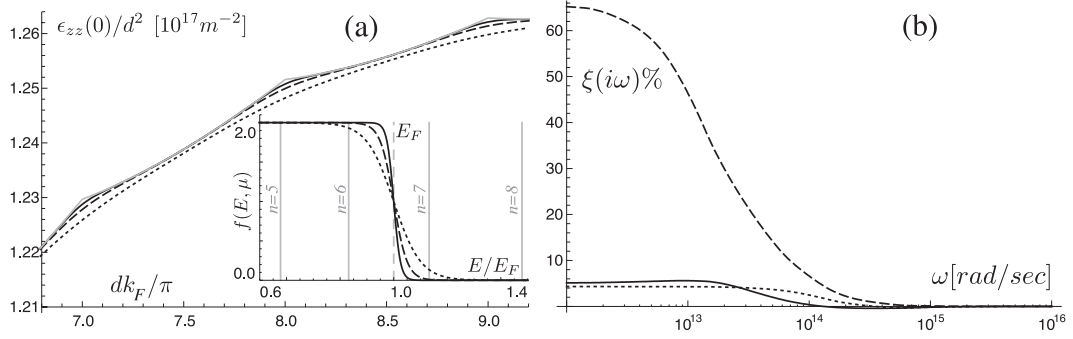


Figure 4.19: (a) Static value of the response function versus the film thickness, for $\Omega_p = 2 \times 10^{14}$ rad/sec at different temperatures, $T = 0^\circ K$ (gray line), $T = 5^\circ K$ (continuous line), $T = 10^\circ K$ (dashed line), $T = 20^\circ K$ (dotted line). The inset shows the temperature spread of the electrons through the energy levels for a 50 nm film. (b) $\xi(i\omega)$ for the PBM having $d = 50$ nm, $\Omega_p = 5 \times 10^{14}$ rad/sec (black line) $d = 50$ nm, $\Omega_p = 10^{14}$ rad/sec (dashed line) $d = 10$ nm, $\Omega_p = 5 \times 10^{14}$ rad/sec (dotted line).

independently of the temperature. So the sum over the occupation factors is again convergent for any $T \neq 0^\circ K$. For $T \rightarrow 0$ the β factor diverges and the exponential vanishes.

Finite temperature corrections become relevant when the separation between energy levels is of the same order of $k_B T$, since only in that situation the levels above the Fermi energy can be significantly populated. For practical purposes the temperature value can not be taken much larger than the room temperature $T = 300^\circ K$ (many experimental difficulties arise at higher temperature) so the maximum energy spread induced by the temperature is of the order of $k_B T \simeq 0.026$ eV. This means that temperature corrections are relevant for films thick enough to have an energy level separation ≤ 0.025 eV. In the case of an infinite well, where the energy level separation increases linearly with the quantum number n , the temperature effects are more significant for low density films, in which only closer energy levels are occupied. For small electron densities the assumption of $k_B T/E_F \ll 1$ does not hold anymore and one has to calculate μ starting from the conservation of the electron number with increasing temperature. The electron density is given by

$$n(d) = \int_0^\infty f(\mu(T), T) n(E) dE \quad (4.34)$$

where $n(E)$ is the density of states, independent of T , and where $\mu(0) = E_F$. So, moving from zero temperature to finite temperature:

$$2 \int_0^{E_F} n(E) dE = 2 \int_0^\infty \frac{n(E)}{e^{\beta(E_{\mathbf{k}\parallel, n} - \mu)} + 1} dE \quad (4.35)$$

this expression provides an iterative way to obtain the μ value once that the Fermi energy is fixed.

Figure 4.19 (a) shows the static value of the zz component of the dielectric tensor for different temperatures, notice that cusps disappear and, increasing T , the dielectric function decreases. The latter behaviour increases the anisotropy of the dielectric tensor, enhancing the QSE. Figure 4.19 (b) shows the relative percent difference ξ between the dielectric functions at $T = 0^\circ$ and $T = 300^\circ K$:

$$\xi(i\omega) = \frac{\epsilon_{zz}(i\omega, 0^\circ) - \epsilon_{zz}(i\omega, 300^\circ)}{\epsilon_{zz}(i\omega, 0^\circ)} \quad (4.36)$$

as a function of the frequency. Notice that ξ increases by increasing the thickness and by decreasing electron density. Changes in sign for ξ can occur because of a crossing of $\epsilon_{zz}(i\omega, 0^\circ)$ and $\epsilon_{zz}(i\omega, 300^\circ)$. The introduction of a finite temperature can strongly modify the dielectric function behaviour, the static value can change significantly, producing a change in the dispersion force. A deeper investigation in this direction requires the use of the finite temperature Lifshitz theory, and the choice of a technique to solve its intrinsic problems, described in section 2.5.

4.5 Conclusions

The theoretical results that have been presented, illustrate the possible role of size quantization effects in the electromagnetic vacuum force between very thin films and show how the determination of these effects depends upon the description of the film electronic structure and upon the inclusion of intraband effects. However it is necessary to point out that this analysis is still a mere indication of the corrections to the simple picture that assumes the same dielectric function for films and bulk solids. Although the basic features of size quantization (confinement of the electronic charge, anisotropy of the dielectric tensor, presence of inter-subbands transitions in the dielectric function) are already present in the models discussed in this chapter, there is room for substantial improvements before an accurate comparison with experimental data, like those obtained by Lisanti et al. for *Pd* films [41], can be done. A more detailed treatment should include (i) band structure effects, (ii) non-locality of the dielectric response and (iii) non-local treatment of the reflectivity. A self consistent first principles calculation of the inverse dielectric matrix for a slab of the appropriate size and symmetry, from which a macroscopic dielectric function can be derived with band structure and non-local effects included, could provide the appropriate treatment of the first two issues [42, 43]. Corrections to the Fresnel optics, along the lines indicated by several authors [44, 45, 46], can lead to important modifications of the reflectivity even in the case of a free electron gas film.

Still when comparing theory with experiments for thin films one should consider the fact that measured relaxation frequencies turn out to depend upon the film size and morphology [15, 47, 48, 49, 50]. This is largely due to the so called classical size effects arising from the scattering of the electrons at the film boundaries. In view of the sensitivity of QSE to the value of the relaxation time, comparison with experimental data may be possible only if a realistic estimate of the modifications in the relaxation times caused by surface scattering is available. Moreover it has been shown that variations in the experimental optical parameters, caused by film morphology, can determine a change in the force of the order of 10% [51, 52]. A different film growth technique or different growth conditions can determine large differences in the plasma and relaxation frequencies, in view of a comparison between theory and experiment, this parameters must be measured for each specific sample.

In extending this theory to the more realistic case of deposited films, one has to account for the penetration of the electron wavefunctions into the substrate. This requires a modification of the quantization condition compared to the abrupt barrier potential adopted in this chapter. However for semiconductor substrates where the confinement is essentially due to the fundamental gap, the modifications of the potential should not alter appreciably the main conclusions. For metallic substrate the extension is less obvious since the confinement is related to the existence of symmetry or relative gaps of the projected bulk band structure, that exist for particular directions only. In this case the solid-film interface has to be simulated with a softer confinement potential [35].

In view of the sensitivity of the presented results to the film density, the observation of the effects reported in this chapter could be made possible by some experimental technique that allows to modify the carrier density in overlayers (doping [53], creating electron-hole plasma

by illumination [54] etc.). Also by changing the substrate one may affect the confinement potential, making it more abrupt or more rounded, and enhance or reduce QSE.

Appendix A: Reflection coefficient of a slabs

In section 2.3 the behaviour of the electromagnetic field in presence of surfaces and stratified systems has been investigated. The calculation of dispersion forces requires the dispersion relation of such systems to be known. Given a certain number of interfaces, one must solve the Maxwell's equations in the different media and impose the proper continuity conditions. The aim of that calculations was the determination of the intrinsic modes of the system, both localized or non localized. If one is interested in the determination of the response of the system (in this case a slab of thickness d) to an external field, the set of equations (2.49) must be modified in order to account for an incoming field:

$$E_0(z) = \begin{cases} Ae^{-\gamma_1 z} + Be^{\gamma_1 z} & z < 0 \\ Ce^{-\gamma_3 z} + De^{\gamma_3 z} & 0 < z < d \\ Ee^{-\gamma_2 z} & z > d \end{cases} \quad (4.37)$$

also in this case γ_i can be both real or imaginary, according to definition (2.34). Generally, once that an incoming direction has been established, the reflection coefficient r is defined as the ratio between the amplitude B of the field back scattered opposite to the incoming direction and the amplitude A of the incoming field. A transmission coefficient t can be defined as the ratio between the amplitude E of the field coming out of the multilayer and the amplitude A of the incoming field. Reflectivity R and transmissivity T are defined as the square of the previous ratios. One can say that, because the intrinsic modes of the system can be excited even in absence of an external field, the set (2.49) is a particular case of the set (4.37) in which $A \rightarrow 0$. Consequently the intrinsic modes of the system are given by the real poles of the reflectivity coefficient. In agreement with the convention of the scattering theory, γ_i can be redefined in order to put in evidence a complex phase:

$$\beta_j = \sqrt{\epsilon_j(\omega) \frac{\omega^2}{c^2} - k^2} \quad (4.38)$$

$$E_0(z) = \begin{cases} e^{-i\beta_1 z} + r e^{i\beta_1 z} & z < 0 \\ \frac{C}{A} e^{-i\beta_3 z} + \frac{D}{A} e^{i\beta_3 z} & 0 < z < d \\ t e^{-i\beta_2 z} & z > d \end{cases} \quad (4.39)$$

Solving the system of equation coming from the interface condition, one can easily get an expression for the coefficient r and t , both for the TE and TM modes. For the geometry of a slab of thickness d with dielectric function ϵ_ω surrounded by the vacuum, after tedious calculations one obtains the following expressions:

$$r_\alpha = \rho_\alpha \frac{1 - e^{2id\beta}}{1 - \rho_\alpha^2 e^{2id\beta}} \quad t_\alpha = \frac{\rho_\alpha^2 e^{id\beta}}{1 - \rho_\alpha^2 e^{2id\beta}} \quad (4.40)$$

where α denotes the *TE* or *TM* modes and ρ is the reflectivity of a single slab surface (vacuum-slab interface):

$$\begin{aligned}\rho_{TM} &= \frac{\sqrt{\epsilon(\omega)\frac{\omega^2}{c^2} - k^2} - \epsilon(\omega)\sqrt{\frac{\omega^2}{c^2} - k^2}}{\sqrt{\epsilon(\omega)\frac{\omega^2}{c^2} - k^2} + \epsilon(\omega)\sqrt{\frac{\omega^2}{c^2} - k^2}} \\ \rho_{TE} &= \frac{\sqrt{\epsilon(\omega)\frac{\omega^2}{c^2} - k^2} - \sqrt{\frac{\omega^2}{c^2} - k^2}}{\sqrt{\epsilon(\omega)\frac{\omega^2}{c^2} - k^2} + \sqrt{\frac{\omega^2}{c^2} - k^2}}\end{aligned}\tag{4.41}$$

Further details can be found in reference [55].

Bibliography

- [1] A. Benassi and C. Calandra. *Europhys. Lett.*, 82:61002, 2008.
- [2] A. Benassi and C. Calandra. *J.Phys.: Conf. Ser.*, 2009. To be published.
- [3] I. Pirozhenko and A. Lambrecht. *Phys.Rev.A*, 77:013811, 2008.
- [4] M. Bordag, U. Mohideen, and V.M. Mostepanenko. *Phys. Rep.*, 353:1, 2001.
- [5] V.S. Bentsen, R. Herikstad, S. Skriudalen, I. Brevik, and J.S.Hoye. *J.Phys.A: Math.Gen.*, 38:9575, 2005.
- [6] A. Lambrecht, I. Pirozhenko, L. Duraffourg, and Ph. Andreucci. *Europhys. Lett.*, 77:44006, 2007. Erratum: *Europhys. Lett.* 81:19901, 2008.
- [7] B. Geyer, G.L. Klimchitskaya, and V.M. Mostepanenko. *J.Phys.A: Math.Gen.*, 40:13485, 2007.
- [8] J.P. Rogers III, P.H. Cutler, T.R. Feuchtwang, N. Miskovski, and A.A. Lucas. *Surf. Sci.*, 141:61, 1984.
- [9] J.P. Rogers III, P.H. Cutler, T.R. Feuchtwang, and A.A. Lucas. *Surf. Sci.*, 181:436, 1987.
- [10] P.D. Loly and J.B. Pendry. *J. Phys. C*, 16:423, 1986.
- [11] S.A. Lindgren and L. Wallden. *Phys. Rev. Lett.*, 59:3003, 1987.
- [12] S.A. Lindgren and L. Wallden. *Phys. Rev. Lett.*, 61:2894, 1988.
- [13] T.C. Chiang. *Surface Sci. Rep.*, 39:181, 2000.
- [14] M. Jalochofski, E. Bauer, H. Knoppe, and G. Lilienkamp. *Phys. Rev. B*, 45:13607, 1992.
- [15] M. Jalochofski, H. Knoppe, G. Lilienkamp, and E. Bauer. *Phys. Rev. B*, 46:4693, 1992.
- [16] E.I. Rogacheva, O.N. Nashchekina, S.N. Grigorov, M.A. Us, M.S. Dresselhaus, and S.B. Cronin. *Nanotechnology*, 13:1, 2002.
- [17] Z. Hens, D. Vanmaekelbergh, E.S. Kooji, H. Wormeester, G. Allan, and C. Delerue. *Phys.Rev.Lett.*, 92:026808, 2004.
- [18] D.M. Wood and N.W. Ashcroft. *Phys.Rev.B*, 25:6255, 1982.
- [19] P. Czoshke, H. Hong, L. Basile, and T.C. Chiang. *Phys.Rev.B*, 72:035305, 2005.
- [20] P.M. Echenique and J.B. Pendry. *J.Phys.C*, 11:2065, 1978.
- [21] R. Otero, A.L. Vazquez de Parga, and R. Miranda. *Surf. Sci.*, 447:143, 2000.
- [22] R. Otero, A.L. Vazquez de Parga, and R. Miranda. *Phys.Rev.B*, 66:115401, 2002.

- [23] G.E. Reute and E.H. Sondheimer. *Proc.Roy.Soc. London Sect. A*, 195:336, 1948.
- [24] F.E. Hutchinson and W.N. Hansen. *Phys.Rev.B*, 20:4069, 1979.
- [25] P. Czoschke, Hong. Hawoong, L. Basile, and T.C.Chiang. *Phys.Rev.B*, 72:075402, 2005.
- [26] Z. Wang, S. Wang, S. Shen, and S. Zhou. *Phys.Rev.B*, 55:10863, 1997.
- [27] A. Lambrecht and S. Reynaud. *Eur. Phys. D*, 8:309, 2000.
- [28] H.B.G. Casimir. *Proc. Kon. Ned. Akad. Wet.*, 51:793, 1948.
- [29] E.M. Lifshitz. *Sov. Phys. JEPT*, 2:73, 1956.
- [30] I.E. Dzyaloshinskii, E.M. Lifshitz, and L.P. Pitaevskii. *Adv. Phys.*, 10:165, 1958.
- [31] F. Zhou and L. Spruch. *Phys. Rev. A*, 52:297, 1995.
- [32] K.F. Schuler. *Surf.Sci.*, 55:427, 1976.
- [33] P.J. Feibelman. *Phys. Rev. B*, 27:1991, 1983.
- [34] S. Ciraci and I.P. Batra. *Phys. Rev. B*, 33:4294, 1985.
- [35] E. Ogando, N. Zabala, E.V. Chilkov, and M.J. Puska. *Phys.Rev.B*, 71:205401, 2005.
- [36] A.S. Davydov. *Quantum Mechanics*. Pergamon Press, 1976.
- [37] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Cornell University, Philadelphia, 1976.
- [38] B.E. Sernelius, K.F. Berggren, M. Tomak, and C. McFadden. *J. Phys. C: Solid State Phys.*, 18:225, 1985.
- [39] N.D. Mermin. *Phys. Rev. B*, 1:1019, 1970.
- [40] P. Garik and N.W. Ashcroft. *Phys.Rev.B*, 21:391, 1980.
- [41] M. Lisanti, D. Iannuzzi, and F. Capasso. *Proc. Nat. Ac. Sci. USA*, 102:11989, 2005.
- [42] J.L. Li, J. Chun, N.S. Wingreen, R. Car, I.A. Aksay, and D.A. Saville. *Phys. Rev. B*, 71:235412, 2005.
- [43] B.E. Sernelius. *J. Phys. A: Math. Gen.*, 39:6741, 2006.
- [44] P. Appel. *Physica Scr.*, 24:795, 1981.
- [45] P.J. Feibelman. *Prog.Surf.Sci.*, 12:287, 1982.
- [46] K. Kempa and W.L. Scaich. *Phys. Rev. B*, 34:547, 1986.
- [47] G. Fahsold, A. Bertel, O. Krauth, N. Maggi, and A. Pucci. *Phys. Rev. B*, 61:1408, 2000.
- [48] J. Sotelo, J. Ederth, and G. Niklasson. *Phys. Rev. B*, 67:195106, 2003.
- [49] G.K. Pribil, B. Johs, and N.J. Ianno. *Thin Solid Films*, 455-456:443, 2004.
- [50] H. Hoffmann and J. Vancea. *Thin Solid Films*, 85:147, 1981.
- [51] I. Pirozhenko, A. Lambrecht, and V.B. Svetovoy. *New J. Phys.*, 8:238, 2006.
- [52] V.B. Svetovoy, P.J. van Zwol, G. Palasantzas, and J.Th.M. De Hosson. *Phys.Rev.B*, 77:035439, 2008.

-
- [53] F. Chen, G.L. Klimchitskaya, V.M. Mostepanenko, and U. Mohideen. *Phys.Rev.Lett.*, 97:170402, 2006.
- [54] G.L. Klimchitskaya, U. Mohideen, and V.M. Mostepanenko. *J. Phys. A:Math.Theor.*, 40:F841, 2007.
- [55] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, London, 1959.

5

First principle calculation of the force between silicon films

Contents

5.1	Introduction	181
5.2	The simulation of a silicon film	182
5.3	The film dielectric tensor	185
5.4	The force calculation	185
5.5	Local field effects	187
5.6	Conclusions	189

The corrections to the dispersion forces due to the quantum size effects have been investigated in chapter 4. It has been shown that they are quite sensitive to the microscopic model used to describe the electron confinement. In the previously used models, the film has been treated as a uniform, continuum background of positive charge, the aim of this chapter is to move to an atomistic description of the film, accounting for its ionic lattice structure. The electron-ion and electron-electron interaction will be taken into account, providing substantial improvements in the microscopic description of the film and leading to a more realistic description of dielectric properties. Besides quantum confinement effects, ab-initio calculations allow to disclose the role of surface states whose presence is expected to be relevant for ultra thin films.

5.1 Introduction

One of the most powerful tools used to describe the microscopic structure of matter, including the electron-ion interaction and, at least in part, the many-body interaction among the electrons, is the Density Functional Theory (DFT) within the Kohn and Sham scheme (KS) [1, 2]. The KS theory provides single particle eigenvalues and eigenfunctions of an effective hamiltonian for the electrons of the system once that the ion positions are fixed. Within the Born-Oppenheimer approximation one can also perform ionic relaxations. For each step of the ion dynamics, the forces among ions are calculated once that the electrons are in their ground state. It is well known that DFT provides correct results only for the ground state

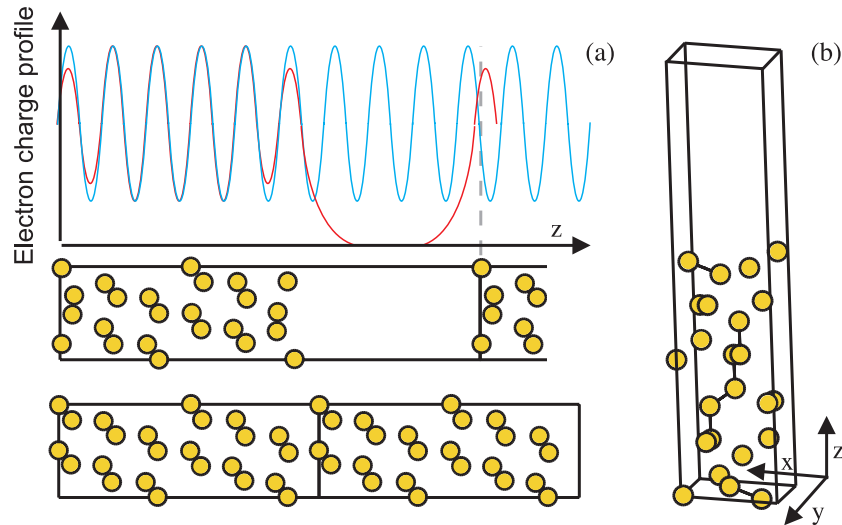


Figure 5.1: (a) sketch of the density profile along the z axes for the bulk cell (blue line) and for the repeated film cell (red line). (b) 3D view of the film simulation cell with 24 atoms and the $[111]$ surface with 2×1 reconstruction.

of the system, whereas the dielectric properties of a system depend also on its excited states. Despite of this limitation DFT is currently employed in the determination of properties of many systems with satisfactory results [1]. A number of many-body techniques exist that, starting from a KS calculation, can correct eigenvalues and eigenfunctions of the excited states to better describe the response functions [3] (see also the end of section 1.10).

In the present study the KS theory will be used to calculate the dielectric tensor of a silicon thin film, from which the dispersion force can be evaluated and compared to the results obtained with different dielectric function models. The interest has been focused on silicon for the following reasons, first silicon has a big relevance in the fabrication of micro- and nano-mechanical systems, second, semiconductors allow to investigate the modification of the dielectric properties due to the presence of surface states. Surface states arise because the atoms at a surface have dangling bonds that must be saturated in some way. In covalent semiconductors, the strength of the chemical bonds produces a reconstruction of the lattice close to the surface and *surface bands* appear inside the gap. This strong modification of the band structure of the system is expected to change the dielectric response of the medium significantly. In free electron metals the bond is isotropic and one cannot identify surface states coming from dangling bonds, even if the electron density of states may be modified at the surface.

The use of a microscopic atomistic description allows to include in the response function local field effects and to investigate their role in the dispersion force.

5.2 The simulation of a silicon film

To simulate the silicon film one has to chose a given surface, i.e. one has to cut the FCC lattice along a particular direction. In the following the $[111]$ direction has been chosen, the most stable reconstruction of this surface is the 7×7 but it requires a very huge simulation cell and it has a metallic behaviour [4]. In the following the 2×1 reconstruction has been considered [5] which is the one naturally obtained after the cleavage operation. It shows a semiconducting

behaviour [4]. KS calculations are known to underestimate the gap of semiconductors [1], they cannot account for neutral excitations (excitons) and they can suffer of many other approximations made by the specific kind of implementation (functionals, pseudopotentials ec.). For this reasons the results obtained for the silicon film cannot be directly compared with the bulk experimental dielectric function or with other bulk models. To be consistent with the approximation of the method, one has to simulate also a silicon bulk and compare the dielectric tensors of bulk and film obtained within the same method.

To calculate the electronic structure of the film the repeated slab approach has been adopted. This technique is commonly used in the study of surface electronic structure. One has to consider an artificial lattice obtained with a periodic array of silicon films intercalated by vacuum films. This allows to exploit the Bloch periodicity and to adopt band structure computational methods. The bulk case is achieved by putting the size of the vacuum film equal to zero. The unit cell, illustrated in figure 5.1 has been obtained by regularly stacking 24 silicon atomic planes along the [111] direction (total thickness 1.9 nm) intercalated by a vacuum film of 1.1 nm thickness. Allowing surface atom relaxation, which gives rise to the 2×1 reconstruction, one obtains a silicon film of 1.7 nm thickness.

Figure 5.1 shows the calculated electron charge density along the z direction normal to the film boundaries. It has the same behaviour obtained in a bulk calculation inside the film, while it shows significant modifications compared to the bulk at the two film surfaces. Such modifications arise from the surface states. Notice that the film density goes quickly to zero as one moves from the film boundaries into the vacuum part of the cell. It seems possible to conclude that periodicity along the z direction does not introduce artificial interactions between films. The surface reconstruction involves only the outer atomic layer and results in the formation of a zig-zag chain of dimers along the y direction of the cell.

Figure 5.2 (a) gives a plot of the film band structure along the high symmetry directions of the two dimensional Brillouin zone. The two arrows indicate the location of the conduction and valence band edges obtained in the calculation of the bulk band structure. Notice that the bulk band gap is 0.6 eV, lower than the experimental value 1.2 eV [6]. It is seen that two surface bands appear in the bulk forbidden energy region, which are separated by a gap of 0.1 eV. The lower band is filled and the upper one is empty, so in principle optical transitions between the two bands are possible. To illustrate the nature of the electron states in these bands figure 5.3 gives the calculated probability density at the J point of the two dimensional Brillouin zone. They appear to be localized on the two dimer chains present at the surface. Most of the bands plotted in the figure correspond to states that behave in a way similar to the bulk states, showing a regular periodicity in the inner layers and decaying exponentially into the vacuum. To illustrate the different contribution of the atoms in the inner layers and at the surface, figure 5.2 (b) shows the local density of states (LDOS) at the surface and in the middle atomic layer to be compared with the bulk DOS. The comparison indicates that the LDOS inside the film shows structures that are similar and cover the same energy intervals that the bulk DOS. However significant differences are observed in the relative weight of the structures both in the valence and in the conduction energy regions.

The calculations have been performed using the DFT KS implementation of the *PWscf* code [7]. This code works with a plane waves basis set, norm-conserving pseudopotentials have been used, and a local density approximation (LDA) functional has been chosen. Convergence tests have been carried out in order to choose a sufficient quantity of vacuum in the simulation cell, to properly sample the 2D Brillouin zone, to take into account a large enough number of empty bands and of plane waves.

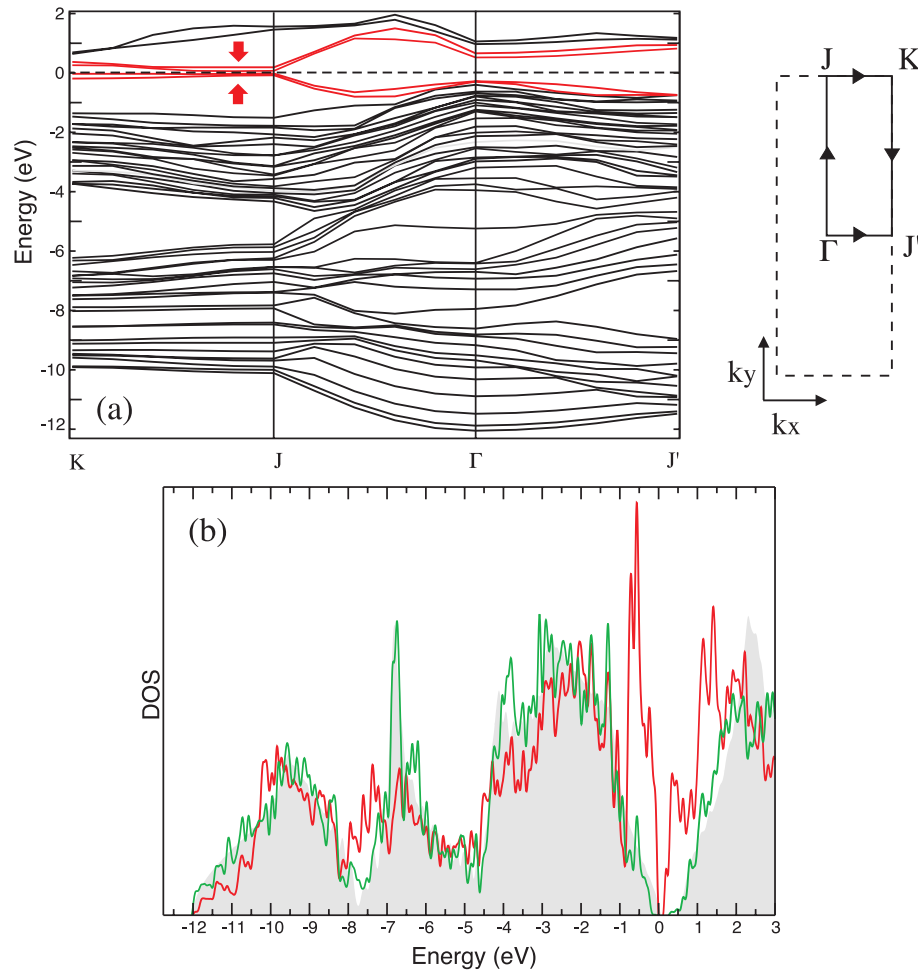


Figure 5.2: (a) film band structure on the 2D Brillouin zone, red lines represent the surface states inside the gap, dashed line represents the Fermi energy. The inset shows the high symmetry points in the 2D Brillouin zone. (b) Silicon bulk DOS (shaded area), DOS projected on the surface atomic layer (red line) and DOS projected on an inner atomic layer (green line).

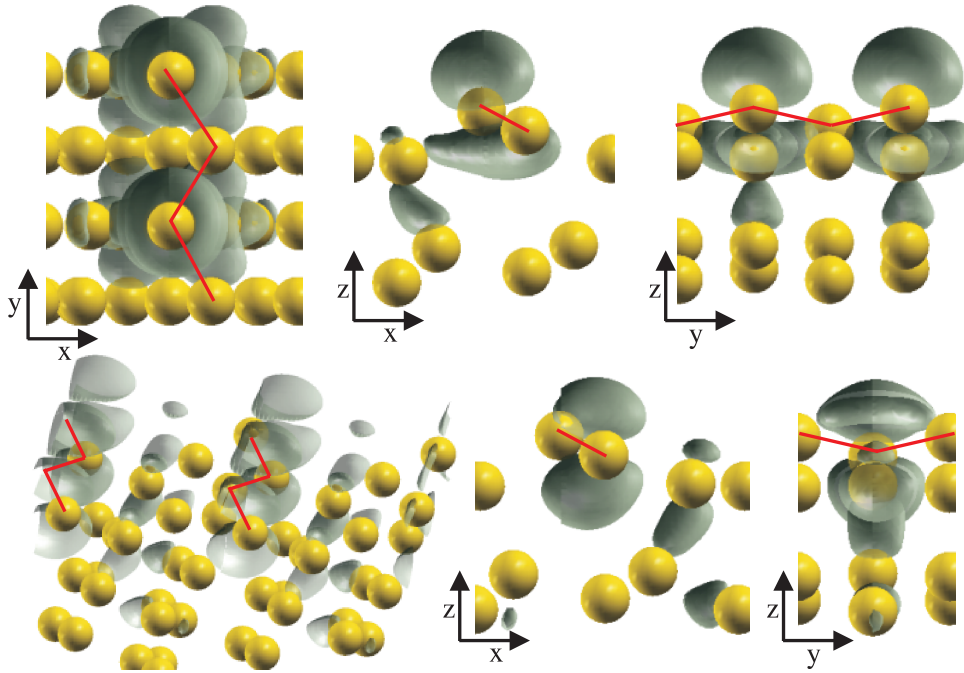


Figure 5.3: Square modulus of the surface states: the first row shows the highest occupied valence state seen from different directions, the second row shows the lowest unoccupied conduction state. The red line indicates the zig-zag surface dimers chains.

5.3 The film dielectric tensor

Once that the DFT energies and wavefunctions are known, the dielectric matrix can be calculated by means of the RPA of section 1.10. Figure 5.4 shows the diagonal components of the dielectric tensor of the film. From the band structure of figure 5.2, a peak inside the silicon gap was expected, but it appears only for the yy component. This is because the dipole matrix elements along the xx and zz directions for the surface states, vanish. One can easily see this, looking at the different parity of the lowest conduction and the highest valence state of figure 5.3. Along the yy direction, for instance, the valence and conduction states show opposite parity: the gradient of the valence state will have the same parity of the conduction one and their overlap integral will be large. On the contrary, along the xx direction the two states show the same parity: on applying the gradient operator the valence state changes his parity and the overlap becomes negligible.

The off-diagonal components of the dielectric tensor are negligible.

5.4 The force calculation

The force has to be calculated using the Lifshitz theory for a five layers system (section 2.3.4) generalized to the case of non isotropic media (section 2.3.5). The general expression is given in equation (2.96). First of all it must be noticed that the KS calculations cannot be directly compared with experimental dielectric functions. To test how far the DFT calculations are from the experimental values, one can compute the bulk dielectric function and compare it with the measured one [8]. Figure 5.5 (a) shows this comparison for the imaginary parts of the dielectric function. The absorption edge is underestimated and the spectral weight moves

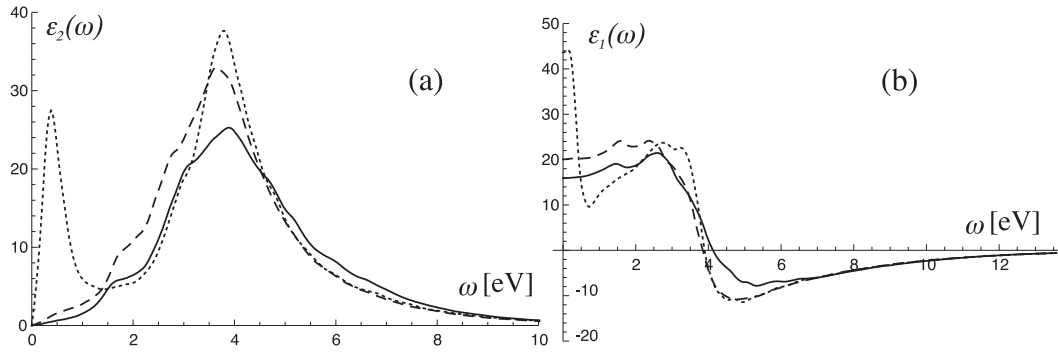


Figure 5.4: Real and imaginary parts of the diagonal components of the film dielectric tensor: ϵ_{xx} dashed line, ϵ_{yy} dotted line and ϵ_{zz} continuous line

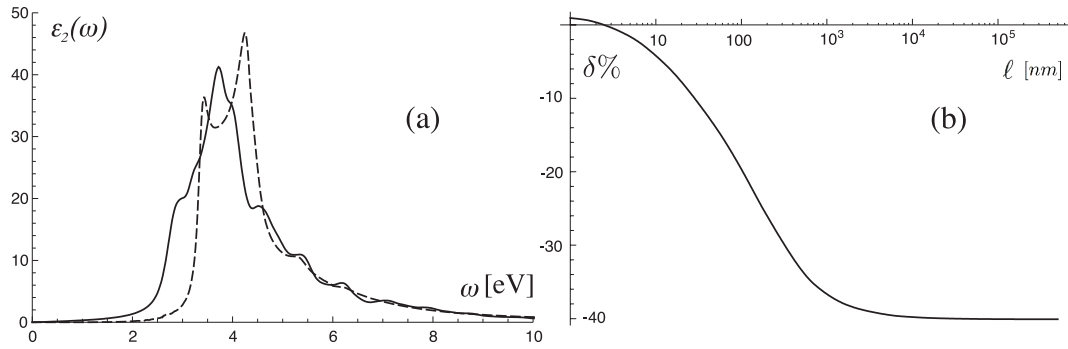


Figure 5.5: (a) DFT (continuous line) and experimental (dashed line) absorption spectrum of a silicon bulk. (b) relative percent differences (5.1).

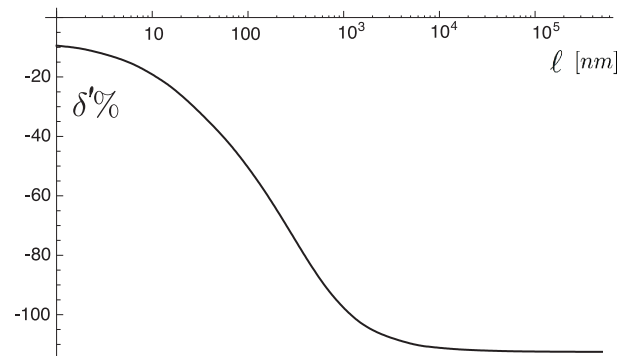


Figure 5.6: Relative percent differences (5.2).

to lower frequencies. The relative percent difference between the force calculated with the DFT dielectric tensor and the experimental one

$$\delta = \frac{F_{exp} - F_{DFTbulk}}{F_{exp}} \quad (5.1)$$

is plotted in figure 5.5 (b). Notice that, at small distances, when the contribution from all the frequency spectrum is important, the experimental force is larger than the DFT one. At 5 nm distance, only small frequencies become relevant and the underestimation of the gap is crucial, the force is larger and the relative percent difference becomes negative, reaching values up to 40%.

To study the quantum size effects on the force between films, excluding the effects of the various approximations made in the calculations, the comparison has to be performed between the theoretical silicon film and the theoretical silicon bulk. The relative percent difference:

$$\delta' = \frac{F_{DFTbulk} - F_{DFTfilm}}{F_{DFTbulk}} \quad (5.2)$$

is plotted in figure 5.6 showing a discrepancy larger than 100% at large films separation. The presence of two flat regions at large and small ℓ has been found also in chapter 4 and therein explained. The large anisotropy and the strong absorption inside the silicon gap, due to the surface states, make the force between films always larger than the force evaluated with the bulk dielectric constant, i.e. the relative percent difference is always negative.

5.5 Local field effects

The RPA dielectric function used in the previous section, contains implicitly the average electron density of the system. It is expected to hold when the local oscillations of the charge are small and close to the average value. When a surface is introduced, rapid and strong oscillations of the electron charge take place in its neighborhood, see figure 5.1. In this case a local field theory can be used to compute a dielectric tensor that, being again a macroscopic average, is more sensitive to the local oscillation of the charge (see section 1.5). To this aim one can use the Adler and Wiser dielectric tensor model described in section 1.10. The London transforms of the bulk and film dielectric tensors are plotted in figure 5.7. Notice that, in the bulk, the inclusion of local field effects decreases the static value of the response function getting closer to the experimental one. A decreasing trend can be noticed also in the film case, however each component undergoes a different reduction and the anisotropy increases.

Once the force has been calculated some interesting comparisons can be made. One can calculate the analogous of (5.2) like in figure 5.8 (a). The relative percent difference decreases by 20% with respect to the RPA. This is mainly due to the decrease of the static values of the film dielectric tensor. Without the inclusion of local field effects, all the film components are well above the bulk ones and the relative percent difference of the force is always negative, with the local field correction some of the components fall below the bulk ones. For this reason at small film separations, when the contribution of all frequencies is relevant, the force calculated with the bulk dielectric function becomes larger than the force calculated with the film dielectric tensor, leading to positive values of δ' . Another interesting comparison is the one between the force acting between films calculated with and without the local field corrections:

$$\delta'' = \frac{F_{DFTfilm}^{NLF} - F_{DFTfilm}^{LF}}{F_{DFTfilm}^{NLF}} \quad (5.3)$$

this relative difference is plotted in figure 5.8 (b) and is always positive, as expected by the decreasing effect of local field corrections. Here the relative percent difference can be higher than 25%.

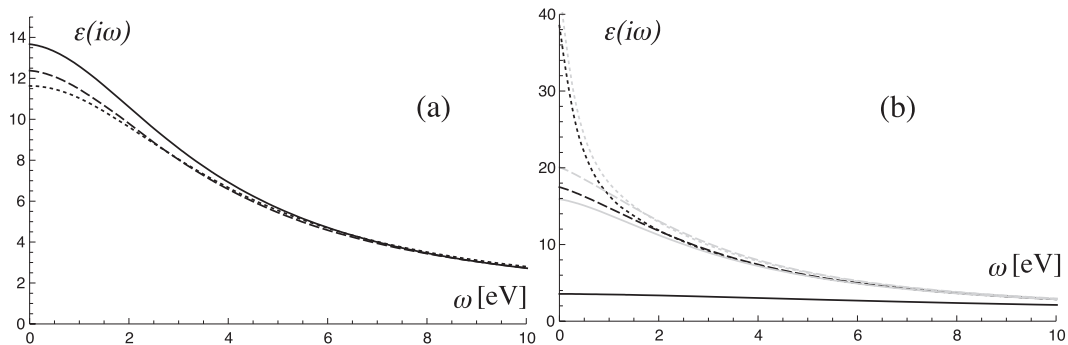


Figure 5.7: London's transforms of the dielectric tensor with local field effect. (a) represent the bulk London's transform with (dashed line) and without (continuous line), local field effects. The transform obtained from the experimental dielectric function is also given for a comparison (dotted line). (b) gives a comparison of the diagonal components of the film tensor with (black lines) and without (gray lines) local field effects. The convention for the different components is the same as in figure 5.4.

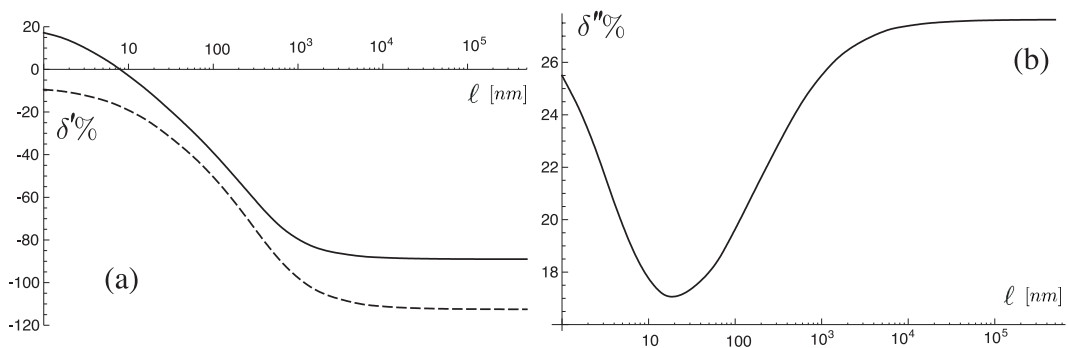


Figure 5.8: (a) relative percent difference (5.2) with (continuous line) and without (dashed line) local field effects. (b) relative percent difference (5.3).

5.6 Conclusions

The analysis of the theoretical results allows to draw the following conclusions.

- The calculation of the dispersion force with the dielectric function appropriate to a Si film of nanometric thickness leads to results that are widely different from those obtained using the bulk dielectric function. Only in the high thickness limit the two methods give the same results.
- The differences arise from two main factors. First the occurrence of both filled and empty surface states in the bulk band gap gives rise to characteristic absorption structures in the dielectric tensor, most clearly seen in the ϵ_{yy} component, at low energy, where the bulk is transparent. This effect is present both in the RPA calculation and when the non-local field effects are included. Second, the modifications occurring in the electronic states that are not confined to the surface lead to a behaviour of the film dielectric tensor considerably different with respect to the bulk case even in the range of frequencies at and above the bulk absorption edge. Both effects combine to give a strongly anisotropic tensor.
- The inclusion of non-local effects in the determination of the macroscopic RPA dielectric function causes deep modifications in its shape compared to the simple RPA. As a consequence the calculated dispersion forces show significant differences.

The conclusion about the role of the surface states seems to be at variance with the frequently reported statement that surface effects are negligible provided the separation distance is much larger than the surface region i.e. the region where surface state wavefunctions are different from zero. Since surface states are typically localized within a few angstrom distance from the outermost atomic plane, one expects their contribution to the force to be negligible when the separation distance is of the order of $5 \div 10$ nm. This is certainly true if referred to the force between two semi-infinite or very thick slabs, since the surface region has a small width compared to the film. However in the case of interacting nanometric films the surface states contribution to the electronic structure is comparable with the contribution of the more extended states and causes deep modifications in the force behaviour. The occurrence of a surface gap smaller than the bulk band gap provides a natural explanation of the discrepancies in both δ' and δ'' quantities at high separation distances. An absorption frequency affects the force at distances comparable with the wavelength. Since the frequency of the surface optical transition is smaller than the bulk band gap, its effects on the force are more evident at large distances.

One can conclude that modifications in the surface structure and morphology of thin nanometric films can produce significant changes in the dispersion force.

The estimate of the effects of non-locality in the determination of the macroscopic dielectric function seems to indicate that these effects are crucial in order to obtain an accurate evaluation of the force starting from first-principle calculations. Previous attempts of including non-locality of the dielectric function into dispersion force calculations have been based on the use of RPA approximate forms [9] or on empirical expressions [10] where the wavevector dependence of the dielectric function was explicitly given. For two semi-infinite gold films the corrections lead to a significant improvement in the agreement with the experimental data [9]. No attempt has been made up to now to calculate the macroscopic dielectric function with its full dependence upon the frequency and the wavevector starting from the inverse dielectric matrix.

In the present calculations the wavevector dependence of the macroscopic dielectric function has not been explicitly considered, which would require a considerable computational effort, only the macroscopic dielectric function in the vanishing wavevector limit has been considered. Even with this limitation the results indicate that a realistic description of the dispersion

force cannot be achieved without including the lattice periodicity in the calculation of the film response.

Bibliography

- [1] W. Kohn. *Rev.Mod.Phys.*, 71:1253, 1999.
- [2] R.M. Dreizler and E.K.U. Gross. *Density Functional Theory*. Springer-Verlag, Berlin, 1990.
- [3] G. Onida, L. Reining, and A. Rubio. *Rev.Mod.Phys.*, 74:601, 2002.
- [4] O. Madelung, W. Martienssen, and G. Chiarotti editors. Landolt-Bornstein new Seris. *Physics of Solid Surfaces: electronic and vibrational properties*, volume 24 B group III. Springer Verlag, 1994.
- [5] K.C. Pandey. *Phys.Rev.B*, 25:4338, 1982.
- [6] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Cornell University, Philadelphia, 1976.
- [7] S. Baroni, A. Dal Corso, S. de Gironcoli, and P. Giannozzi, 2001. <http://www.pwscf.org>.
- [8] E.D. Palik editor. *Handbook of optical constants of solids*. Academic press, London, 1985.
- [9] B.E. Sernelius. *Phys. Rev. B*, 71:235114, 2005.
- [10] R. Esquivel-Sirvent and V.B. Svetovoy. *Phys. Rev. B*, 72:045443, 2005.

6

Optically tunable nano-devices

Contents

6.1	Introduction	193
6.2	The model device	194
6.3	GeTe dielectric properties and the force	195
6.4	Tailoring the device performance	197
6.5	Conclusions	200

This chapter is devoted to the study of a possible application of the electromagnetic vacuum forces in the field of micro and nano-mechanical systems. Since dispersion forces between macroscopic bodies depend upon their dielectric properties, the equilibrium distance between two interacting bodies can be tuned changing their dielectric response. Polycrystalline GeTe is a suitable materials for this purpose, infact it can undergo a fast and reversible metal-insulator phase transition, changing strongly its optical reflectivity and conductivity. Using recent measurements of GeTe dielectric properties, the feasibility of a dispersion forces based device will be discussed, whose the mechanical motion of a cantilever can be tailored via crystalline-amorphous transitions. The work described here has been published in the paper of reference [1].

6.1 Introduction

With the development of modern technology towards smaller structures an increasing attention has been addressed to the role of electromagnetic vacuum fluctuation forces (dispersion or van der Waals and Casimir forces) in micro- and nano-electromechanical systems [2, 3, 4, 5, 6, 7, 8]. In chapter 2 it has been shown that these forces vary typically with the third or fourth inverse power of the distance between the surfaces of the interacting bodies and therefore can be very intense in the sub-micrometer regime. When the moving parts of a device come to such a close distance, dispersion forces affect the dynamics and the operation of the device and, under certain circumstances, they can cause adhesion or stiction between the surfaces, thus limiting the device lifetime. Devices actuated by Casimir force have been designed and demonstrated [9, 10, 11]. Since these forces depend upon the reflectivity, the

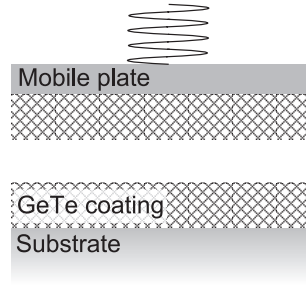


Figure 6.1: Simple representation of the dispersion forces based device.

size and the geometry of the interacting parts [12, 13], several possibilities can be explored to devise structures that exploit such dependence in order to improve the performance of existing devices or to design new types of devices [14].

This chapter explores the possibility of using materials, that undergo fast and reversible phase transitions, to extend the travel range of a device. Thermally induced phase changes between the amorphous and the polycrystalline state in a thin film have been observed since a long time in tellurides containing Ge and Sb (GST materials) [15, 16, 17, 18, 19, 20, 21, 22]. The transition is accompanied by significant changes in optical and transport properties, a feature that is exploited in optical data storage [23] and could be useful in new nanoscale memories [24, 25]. Previous theoretical studies have shown that metal insulator transitions may cause significant changes in the Casimir force behaviour [26, 27]. In the following this transition will be used to modify the force between the components of a simple device. It will be shown that the locations of the extrema in the total energy potential curve of the device are displaced by the transition in such a way that the range of distances and the energy interval over which the device can be operated can be significantly modified.

6.2 The model device

An ideal actuator [2, 14] will be considered consisting of two parallel plates separated by a gap, with one plate fixed on a substrate and the other suspended by an elastic restoring force $F = -Kx_0\delta$, where x_0 is the unactuated distance and K is the spring constant. The device is sketched in figure 6.1. The stationary plate has a flat surface at $\delta = 1$, while $\delta = 0$ denotes the equilibrium position of the movable plate in the absence of dispersion forces and corresponds to the unstretched state of the spring. The attractive interaction between the plates is provided by the electromagnetic vacuum fluctuation force, whose expression in terms of the temperature T , the plates dielectric functions ϵ_1 and ϵ_2 and the inter-plate distance $d = (1 - \delta)x_0$, is given by the Lifshitz formula (2.87):

$$F = -\frac{A}{\pi\beta} \sum_{n=0}^{\infty} \int_0^{\infty} \gamma k dk \left[\frac{1 - Q_{TM}(i\Omega_n)}{Q_{TM}(i\Omega_n)} + \frac{1 - Q_{TE}(i\Omega_n)}{Q_{TE}(i\Omega_n)} \right] \quad (6.1)$$

here A is the plate surface, $\beta = 1/k_B T$, k_B is the Boltzmann constant, $\Omega_n = 2\pi n/\hbar\beta$ is the Matsubara frequency corresponding to the n -th thermal fluctuation mode, the prime on the summation indicates that the $n = 0$ term is given half weight. Q_{TM} and Q_{TE} refer to

transverse magnetic (TM) and transverse electric (TE) modes respectively and are given by:

$$Q_{TM}(i\Omega_n) = 1 - \frac{(\epsilon_1\gamma - \gamma_1)(\epsilon_2\gamma - \gamma_2)}{(\epsilon_1\gamma + \gamma_1)(\epsilon_2\gamma + \gamma_2)} e^{-2d\gamma} \quad (6.2a)$$

$$Q_{TE}(i\Omega_n) = 1 - \frac{(\gamma - \gamma_1)(\gamma - \gamma_2)}{(\gamma + \gamma_1)(\gamma + \gamma_2)} e^{-2d\gamma} \quad (6.2b)$$

with:

$$\gamma_i^2 = k^2 + \frac{\Omega_n^2}{c^2} \epsilon_i(i\Omega_n) \quad \gamma^2 = k^2 + \frac{\Omega_n^2}{c^2} \quad (6.3)$$

and the dielectric functions are evaluated at the frequency $i\Omega_n$ by means of the London transform (see appendix C of the previous chapter). In the calculation for the metallic phase the prescription of reference [28] has been adopted, i.e. the $n = 0$ contribution of the TE modes has been set equal zero. The behaviour of the force as a function of d is non linear: at short distances it can be reproduced by a third inverse power of d (van der Waals regime), but a single inverse power term does not reproduce its behaviour for d values larger than a few tens of nanometer. As pointed out by several authors [29, 30], mainly in connection with electrostatic actuators, there is an intrinsic instability in a device of this sort, which prevents the plates to be stably positioned over a large distance. The system is bistable, having a total potential energy with a local and an absolute minimum separated by a barrier. This limits the range of motion of the device and determines stiction at short distances between the plates [2, 3, 5, 6, 7]. The configuration illustrated in figure 6.1 has been adopted, where both plates are made of telluride films deposited on proper substrates. For simplicity it has been assumed that the thickness of the deposited films is larger than the distance between the plates (if the film size is comparable to the distance, the substrate is expected to affect the intensity of the force). This assumption can be removed by extending the expression of the Lifshitz force to deal with a five layer system [31, 14].

6.3 GeTe dielectric properties and the force

To illustrate the model device, the attention has been focussed on GeTe, whose optical properties have been experimentally determined in both phases. It undergoes rapid transitions between polycrystalline and amorphous states under either optical or electrical excitations. According to the band theory in a perfect crystal rocksalt structure GeTe should exhibit a semiconducting behaviour. However the stable polycrystalline phases are characterized by large vacancy concentrations (typically 10^{20} cm^{-3}) and local distortions [32, 33]. Compared to other defects, that may be present, germanium vacancies have the lower formation energy [34]. GeTe thin films show a p -type electrical conductivity with a high carrier density of up to $5 \times 10^{20} \text{ cm}^{-3}$. This density is consistent with the concentration of Ge vacant sites, i.e. the metallic conductivity is the consequence of the holes that are formed per single vacant site. In the optical properties direct evidence of the metallic behaviour is provided by the presence of Drude-like intraband absorption in crystalline phase [35] (see figure 6.2). On the other hand the amorphous phase has typical semiconductor properties, as illustrated by the behaviour of the imaginary part of the dielectric function, reported in the same figure. Conventional semiconductors, such as Si or GaAs, do not show such large differences in the optical properties as a consequence of the amorphous-crystal transition.

Figure 6.2 displays the calculated London transform for both phases. To obtain this curve one has to reproduced the experimental data available in the range of frequencies $0.1 \div 5 \text{ eV}$ with a Drude-Lorentz model including both intra- and inter-band transitions in the metallic phase, and pure inter-band transitions in the amorphous phase. The values of the carrier density, the carrier effective mass, the relaxation time and the plasma frequency agree with those derived from various experimental studies [16, 17, 18]. It is seen that $\epsilon(i\omega)$ turns out to

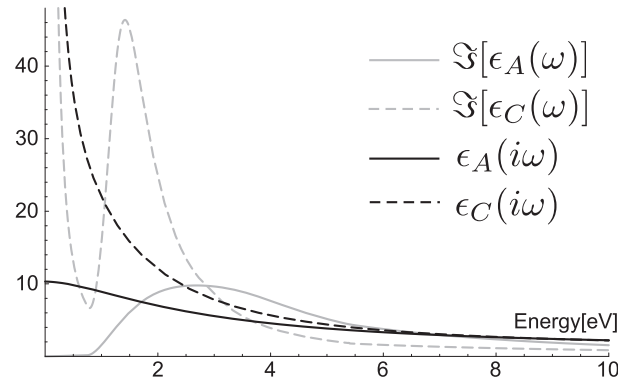


Figure 6.2: Imaginary parts of amorphous and crystalline GeTe dielectric functions and their London transforms, adapted from reference [35] as discussed in the text.

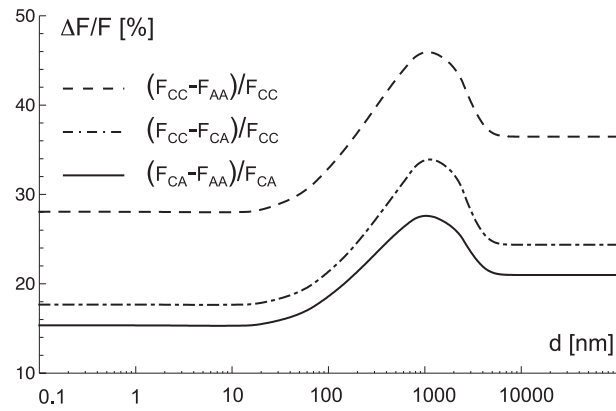


Figure 6.3: Relative difference between F_{CC} , F_{CA} and F_{AA} .

be very different in the two phases in almost all the range of frequencies. As a consequence of the intraband absorption, that is present in the metallic phase only, the curve for the polycrystalline phase for small ω takes much higher values than the one for the amorphous phase. Such differences over all the integration interval of equation (6.1) are expected to determine significant changes in the force. One can use equation (6.1) to calculate the force between two amorphous coated plates F_{AA} , two crystal coated plates F_{CC} , and the mixed configuration F_{CA} . Figure 6.3 reports, as a function of the separation distance, the relative variation of the dispersion force between the plates with respect to the configuration with both plates in the crystalline phase, when one or both plates undergo the transition from the crystalline to the amorphous phase. The curves have a similar behaviour with two ranges, at small and large distances, where the relative difference is constant, and a maximum in the range between 50 nm and 5 μm . The constant values are a consequence of the fact that the forces have identical behaviour as a function of the plate separation. At short distances the force can be approximated by [36]:

$$F = -\frac{1}{4\pi\beta d^3} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \frac{1}{m^3} \left(\frac{\epsilon_1 - 1}{\epsilon_1 + 1} \right)^m \left(\frac{\epsilon_2 - 1}{\epsilon_2 + 1} \right)^m \quad (6.4)$$

which gives a relative change of approximately 28%, that does not depend upon the distance and the temperature. Notice that for the $n = 0$ term of the Matsubara sum, the crystalline

dielectric function diverges and:

$$\frac{\epsilon - 1}{\epsilon + 1} \rightarrow 1 \quad \sum_{m=1}^{\infty} \frac{1}{m^3} = \zeta(3) \quad (6.5)$$

where $\zeta(x)$ is the Riemann zeta function. On the other hand, at large distances within the present treatment of the temperature dependence of the Casimir force, the force between polycrystalline plates is simply given by [28, 37]¹:

$$F_{CC} \simeq -\frac{\zeta(3)}{8\pi\beta d^3} \quad (6.6)$$

In the amorphous case the force for large d cannot be written in a closed form. However it can be very well approximated by the $n = 0$ term of the Matsubara sum

$$F_{AA} \simeq -\frac{1}{8\pi\beta d^3} Li_3 \left[\left(\frac{\epsilon_A(0) - 1}{\epsilon_A(0) + 1} \right)^2 \right] \quad (6.7)$$

where $Li_n[x]$ is the polylogarithmic function of n -th order in the argument x . From these expressions one obtains a relative change of the order of 36%, in agreement with the exact numerical results in figure 6.3.

6.4 Tailoring the device performance

Device instability is usually discussed in terms of a dimensionless parameter, like the ratio between the Casimir energy at the equilibrium distance and the elastic energy [2, 14, 8]. This can be done conveniently when the force has a simple inverse power dependence with the separation distance. The analysis of the bifurcation diagrams as a function of the parameter allows one to draw conclusions about changes in the critical separation with the material dielectric properties. In the present case the force as a function of the separation distance cannot be reproduced by a simple power law, so that a single dimensionless parameter cannot be used to study the device instability.

The behaviour of the device depends upon the spring constant k , the area A of the plates and the unactuated distance x_0 . The ratio $\Xi = A/k$ gives a measure of the relative importance of the vacuum and the elastic force, since they increase linearly with A and k respectively. According to the values attributed to these parameters one obtains different intervals of distances and energies of device operation at fixed x_0 . To show the effects of the phase transition figure 6.4 gives the plots of the potential energy of the device for fixed k and A at x_0 values of the order of hundred nanometers. This choice corresponds to distances where the relative difference of the force varies significantly (see figure 6.3). As expected, on passing from the CC to the AA configuration the weakening of the vacuum force shifts the stable equilibrium position towards smaller δ values and the maximum at larger δ . The displacement is more pronounced for low Ξ i.e. soft spring and small area (figure 6.4 (a)), where it can be of the order of 80% and more, than in the case of large Ξ plotted in figure 6.4 (b), where the increase of the travel range is of the order of 10%. The range of energies for which the device can be operated is modified in a similar way. These curves provide examples of the possibility of tuning in reversible way the device properties using the metal-insulator phase transition. Figure 6.5 displays bifurcation diagrams of the device as a function of Ξ for x_0 equal to 100 nm (figure 6.5 (a)) and 5 μ m (figure 6.5 (b)) [14, 38, 39] for the CC , CA and

¹Notice that this behaviour is different from the expression of the Casimir force at $T = 0^\circ K$ that for a perfect metal gives $F_{CC} = -\frac{hc\pi^2}{240d^4}$. This formula has been frequently used in previous studies of the role of electromagnetic forces in stiction phenomena.

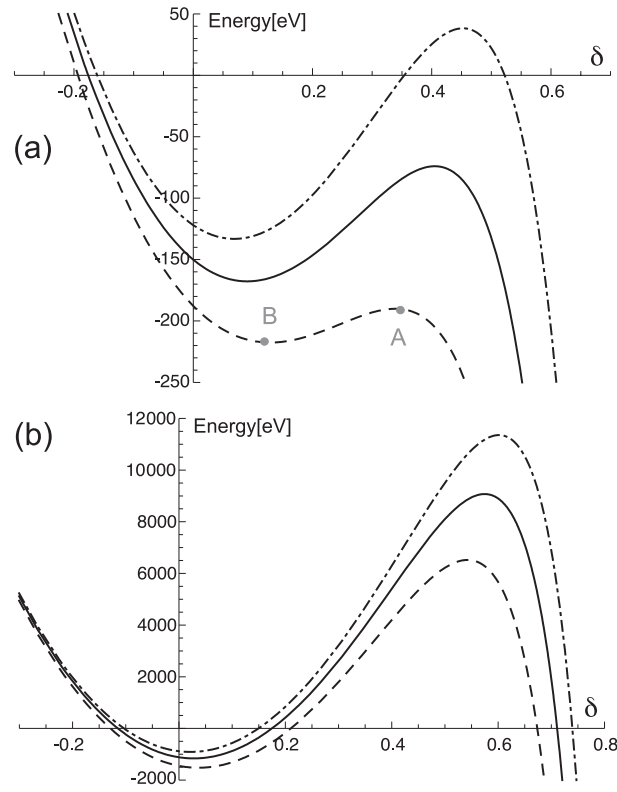


Figure 6.4: Potential profiles for *CC* configuration (dashed line), *CA* configuration (continuous line) and *AA* configuration (dot-dashed line). In the (a) plot $x_0 = 100 \text{ nm}$, $k = 0.1 \text{ N/m}$ and $A = 2 \cdot 10^{-10} \text{ m}^2$ letters A and B refer to figure 6.5 (a); in the (b) plot $x_0 = 200 \text{ nm}$, $k = 0.5 \text{ N/m}$ and $A = 10^{-8} \text{ m}^2$.

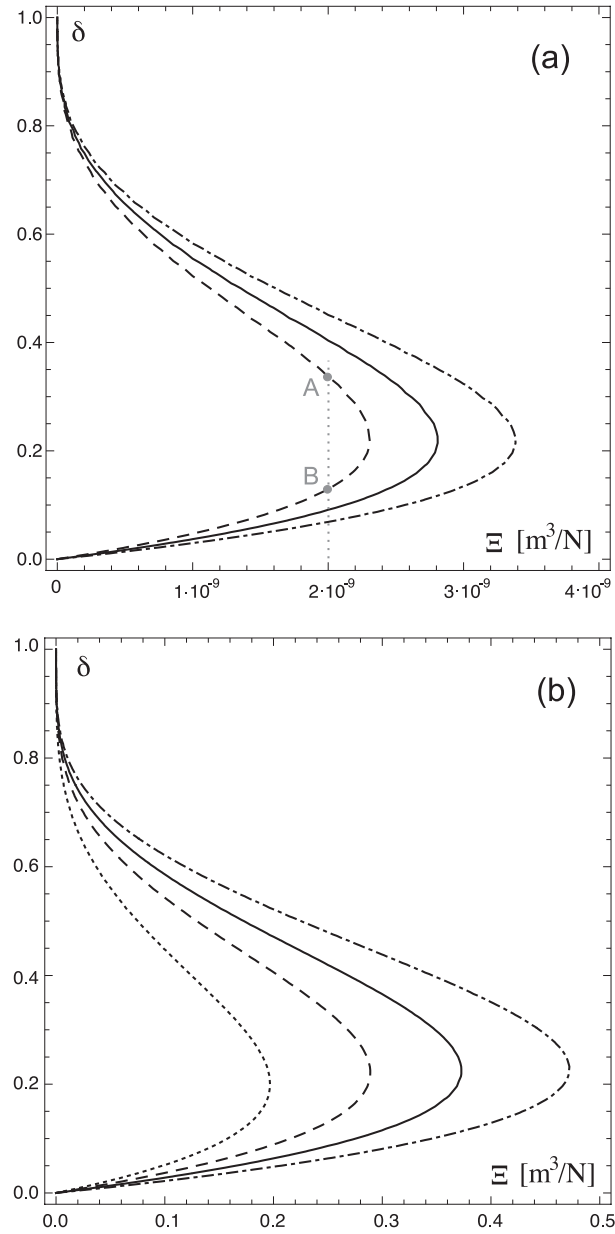


Figure 6.5: Bifurcation diagrams for *CC* configuration (dashed line), *CA* configuration (continuous line) and *AA* configuration (dot-dashed line). In the (a) plot $x_0 = 100 \text{ nm}$, the dotted line represents the Ξ value used in figure 6.4(a). In the (b) plot $x_0 = 5 \text{ } \mu\text{m}$, the dotted line represents the force calculated using the $T = 0^\circ \text{ K}$ Casimir's force.

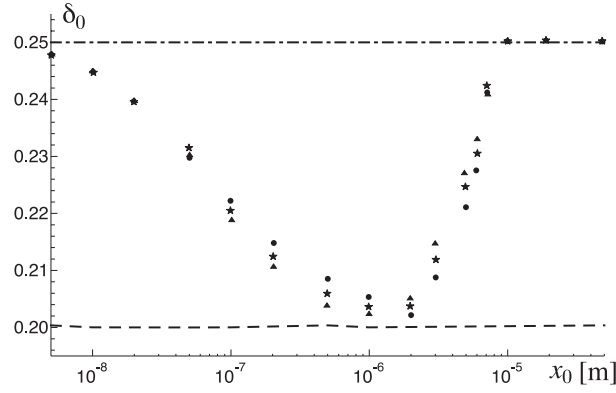


Figure 6.6: δ_0 as a function of x_0 . Dashed line corresponds to the $T = 0^\circ K$ Casimir's force, dot-dashed line is obtained in the small and large d limits of equations (6.4) and (6.6). Dots represent the exact results for F_{CC} (circles), F_{AA} (triangles) and F_{CA} (stars).

AA configurations. The curves have similar shapes, with a lower branch before the fold that corresponds to stable solutions while the upper branch indicates the unstable states. The δ values which allow to operate the device at a given value Ξ^* are those which lie within the intersections of the line $\Xi = \Xi^*$ with the upper and the lower branch of the curves (points A and B in figure 6.5 (a)). At the Ξ value corresponding to the curve fold there is only one intersection, which gives the critical separation δ_0 . As expected, the crystalline-amorphous transition changes significantly the range of Ξ parameters over which the device can work. Notice that δ_0 varies between 0.21 and 0.22. The critical value is expected to change when the functional dependence of the vacuum force with the plate separation is modified: for the d^{-4} dependence typical of the Casimir force $\delta_0 = 0.20$, while for the d^{-3} behaviour, which is characteristic of the van der Waals interaction and of the large distance force at $T > 0^\circ K$, $\delta_0 = 0.25$ [2, 14]. In the case illustrated in figure 6.5 (a), i.e. 100 nm distance, the critical value deviates significantly from the third inverse power behaviour. The curves of the 5 μm case show a similar behaviour. For the sake of comparison figure 6.5 (b) reports the curve corresponding to a pure Casimir interaction, which has been often adopted in similar device analysis. This force is appropriate to describe the case of metallic plate at large distances for $T = 0^\circ K$. Comparison with the results for the *CC* configuration shows that the ranges of Ξ and δ values are considerably underestimated by this interaction. Notice that the vacuum force at such distance cannot yet be described by a third inverse power law, as expected from equation (6.6) at $T > 0^\circ K$. This is most clearly seen by the critical δ_0 value which is smaller than 0.25.

To better illustrate how the separation distance dependence of the vacuum force can affect the device performance, figure 6.6 represents the critical value as a function of the unactuated distance. It is seen that δ_0 varies between 0.25 at very short (less than 10 nm) and very large (more than 10 μm) distances and a minimum slightly larger than 0.20. The minimum occurs at different distances in the three configurations and it is found at smaller x_0 for the *AA* case. It is clear from this curve that simulating device behaviour on the basis of simple inverse power law forces does not allow to get realistic results.

6.5 Conclusions

It has been shown that the metal-insulator transition in germanium tellurides can be exploited in a simple device to modify its performance and the conditions under which it can be used. The choice of the appropriate values for the parameters entering into the model may depend

upon a number of factors, like the film quality [40], the possibility of controlling the kinetics of the phase transition, the mechanical properties of the components, the surface roughness [41] etc. which have to be clarified in order to plan the realization of a specific device. On the theoretical side two aspects have to be further investigated. The first concerns the role of the film size in determining the change in the vacuum force: this can be carried out along the lines of previous studies [31, 14] with the purpose of determining the minimum thickness that allows to detect significant vacuum force variations. The second has to do with the effects of the finite plate area, which in the metal case may be needed to avoid inconsistencies related to the Drude description of the intraband absorption [42].

Bibliography

- [1] A. Benassi and C. Calandra. *Europhys. Lett.*, 84:11002, 2008.
- [2] F.M. Serry, D. Walliser, and G.J. Maclay. *J. Appl. Phys.*, 84:2501, 1998.
- [3] F.M. Serry, D. Walliser, and G.J. Maclay. *J. Microelectromechanical Syst.*, 4:193, 1995.
- [4] E. Buks and M.L. Roukes. *Europhys. Lett.*, 54:220, 2001.
- [5] E. Buks and M.L. Roukes. *Phys. Rev. B*, 63:033402, 2001.
- [6] H.J. De los Santos. *Proc. IEEE*, 91:1907, 2003.
- [7] F.W. Delrio, M.P. De Boer, J.A. Knapp, P.J. Clews E.D. Reedy jr, and M.L. Dunn. *Nature Mat.*, 4:629, 2005.
- [8] R.C. Batra, M. Porfiri, and D. Spinello. *Europhys.Lett.*, 77:20010, 2007.
- [9] H.B. Chan, V.A. Aksyuk, R.N. Kleiman, D.J. Bishop, and F. Capasso. *Science*, 291:1941, 2001.
- [10] H.B. Chan, V.A. Aksyuk, R.N. Kleiman, D.J. Bishop, and F. Capasso. *Phys.Rev.Lett.*, 87:211801, 2001.
- [11] A. Ashourvan, M. Miri, and R. Golestanian. *Phys.Rev.Lett.*, 98:140801, 2007.
- [12] D. Iannuzzi, M. Lisanti, and F. Capasso. *Proc. Nat. Ac. Sci. USA*, 101:4019, 2004.
- [13] F. Chen, G.L. Klimchitskaya, V.M. Mostepanenko, and U. Mohideen. *Phys.Rev.B*, 76:035338, 2007.
- [14] J. Barcenas, L. Reyes, and R. Esquivel-Sirvent. *Appl. Phys. Lett.*, 87:263106, 2005.
- [15] S.R. Ovshinsky. *Phys.Rev.Lett.*, 21:1450, 1968.
- [16] K.L. Chopra and S.K. Bahl. *J.Appl.Phys.*, 40:4171, 1969.
- [17] S.K. Bahl and K.L. Chopra. *J.Appl.Phys.*, 40:4940, 1969.
- [18] S.K. Bahl and K.L. Chopra. *J.Appl.Phys.*, 41:2196, 1970.
- [19] N. Yamada, E. Ohno, K. Nischiuchi, N. Akahira, and M. Takao. *J.Appl.Phys.*, 69:2849, 1991.
- [20] N. Yamada. *MRS Bull.*, 21:48, 1996.
- [21] M. Libera and M. Chen. *J.Appl.Phys.*, 73:2272, 1993.
- [22] B.S. Lee, J.R. Abelson, S.G. Bishop, D.H. Kang, B. Cheong, and K.B. Kim. *J.Appl.Phys.*, 97:093509, 2005.
- [23] T. Ohta. *J. Optoelectr. Adv. Mat.*, 3:609, 2001.

-
- [24] M.H.R. Lankhorst, B.W.S.M.M. Ketelaars, and R.M. Wolters. *Nature Mat.*, 4:347, 2005.
- [25] M. Wuttig. *Nature Mat.*, 4:267, 2005.
- [26] I. Pirozhenko and A. Lambrecht. *Phys.Rev.A*, 77:013811, 2008.
- [27] R. Esquivel-Sirvent. *Phys. Rev. A*, 77:042107, 2008.
- [28] K.A. Milton. *J. Phys. A: Math. Gen.*, 37:R209, 2004.
- [29] K.L. Ekinici and M.L. Roukes. *Rev. Sci. Instrum.*, 76:061101, 2005.
- [30] Wen-Ming Zhang, Guang Meng, and Di Chen. *Sensors*, 7:760, 2007.
- [31] F. Zhou and L. Spruch. *Phys. Rev. A*, 52:297, 1995.
- [32] W. Welnic, A. Pamungkas, R. Detemple, C. Steimer, S. Blugel, and M. Wuttig. *Nature Mat.*, 5:56, 2005.
- [33] M. Wuttig, D. Lusebrink, D. Wamwangi, W. Welnic, M. Gilleben, and R. Dreonskowski. *Nature Mat.*, 6:122, 2007.
- [34] A.H. Edwards, A.C. Pineda, P.A. Schultz, M.G. Martin, A.P. Tompson, H.P. Hjalmarson, and C.J. Umrigar. *Phys. Rev. B*, 73:045210, 2006.
- [35] W. Welnic, S. Botti, L. Reining, and M. Wuttig. *Phys.Rev.Lett.*, 98:236403, 2007.
- [36] U. Hartmann. *Phys. Rev. B*, 42:1541, 1990.
- [37] J.S. Høye, I. Brevik, J.B. Aarseth, and K.A. Milton. *Phys. Rev. E*, 67:056116, 2003.
- [38] A. Pelesko. *SIAM J. Appl. Math.*, 62:888, 2002.
- [39] Wen-Hui Lin and Ya-Pu Zhao. *Chaos,solitons and fractals*, 23:1777, 2005.
- [40] I. Pirozhenko, A. Lambrecht, and V.B. Svetovoy. *New J. Phys.*, 8:238, 2006.
- [41] G. Palasantzas and J.Th. DeHosson. *Phys. Rev. B*, 72:121409, 2005.
- [42] B. Geyer, G.L. Klimchitskaya, and V.M. Mostepanenko. *J.Phys.A: Math.Gen.*, 40:13485, 2007.