الجمهورية الجزائرية الديمقراطية الشعبية

The People's Democratic Republic of Algeria

العالي و البحث العلمي وزارة التعليم

Ministry of Higher Education and Scientific Research

جامعة عمار ثليجي – الأغواط

University Amar Telidji of Laghouat

Faculty : ST

Departement : Electronic

Domain : ST

Field : Electronic

Speciality :Micro-electronic

Course Handout (Lectures)

Intended for students of: Master. Level: 2nd year

Photonic Crystals

Presented by: Dr. Fathi Bendelala

Grade: MCA, University of Laghouat

Email : f.bendellala@lagh-univ.dz

Academic year: 2024/2025

Preface

This handout serves as a comprehensive collection of lecture notes designed for the master's and post-graduate course on photonic crystals and metamaterials, offered during the first semester of 2025 at the Department of Electronics within Amar Tlidji University of Laghouat. The course aims to provide students with an in-depth understanding of the fundamental principles, advanced concepts, and practical applications of photonic crystals and metamaterials, which are at the forefront of modern optical and electromagnetic research. These materials have garnered significant attention due to their unique properties and potential to revolutionize various technological fields, including telecommunications, sensing, and energy harvesting. The notes are structured to guide students through the theoretical foundations, numerical modeling techniques, and experimental approaches essential for mastering this interdisciplinary subject. By the end of the course, students are expected to gain both the knowledge and analytical skills necessary to contribute to ongoing research and innovation in this rapidly evolving domain.

Contents

Contents

Contents

Contents

Chapter 01

(Introduction)

1.1. Introduction

Modern electronics, radio, and microwave techniques show fantastic device integration, miniaturization, efficiency, and speed achievements. A device containing millions of transistors can fit into a micrometre-size volume, which is extremely small in practical terms and compared to the wavelength of waves the device can control.

Regarding the miniaturization, modern electronics are coming to the limits imposed by nature. We need to develop other technologies to qualitatively improve existing telecommunication and computing systems, and optical (photonic) means are probably the most promising.

In modern telecommunication systems, the transmission of signals to long distances has been performed by optical fibres, which have primarily replaced telephone and data cables. The optical frequency range is preferred because the information capacity of a telecommunication channel based on conducting wires is fundamentally limited by the available spectrum and the physical properties of metals (see a discussion below). In contrast, modern optical fibres possess negligibly small scattering and dissipative losses in wavelength λ =1260-1675 nm, where six telecom frequency bands with the relative widths 3-5% of each are located.

The same physical limitations of electronics call for optical signal processing and storage techniques. However, optical devices are bulky, especially on the wavelength scale, and the existing technologies do not allow the integration of many functions in a single device that would be as compact as its electronic analogue.

What can we do in optics?

First, modern optics do not have nano guides or optical waveguides with a cross-section that is more petite than the guided wavelength. Second, the industry has not yet delivered optical capacitors or inductors. We do not have optically thin frequency selective sheets available for optical devices. We have to resort to multilayers of carefully selected dielectric layers, each of which has a thickness comparable to the wavelength, and illuminate them by optical beams much wider than the wavelength. Again, if we used only “optical means” to design radios, every simple filter would have a diameter of many meters! And we need many tens of such filters in every mobile phone.

We can understand that the traditional optical approaches and the signal processing so elaborated in radio and microwave engineering are incompatible. Nowadays, the only transmission of signals to long distances is performed optically. Processing of signals and computations is still done electronically. The time has come for optical and radio engineers to collaborate to create nanophotonics devices for future optical processors and fully optical telecom systems. Radio and optics converge in this area, and the breakthrough is possible only through interdisciplinary research.

The interdisciplinary field of metamaterials and nanophotonics focuses on electromagnetic devices and components that have characteristic dimensions that are either smaller or comparable (as seen in photonic crystals) to the wavelength of the radiation they manipulate. The latest advancements in this area are depicted in Fig. 1.1.

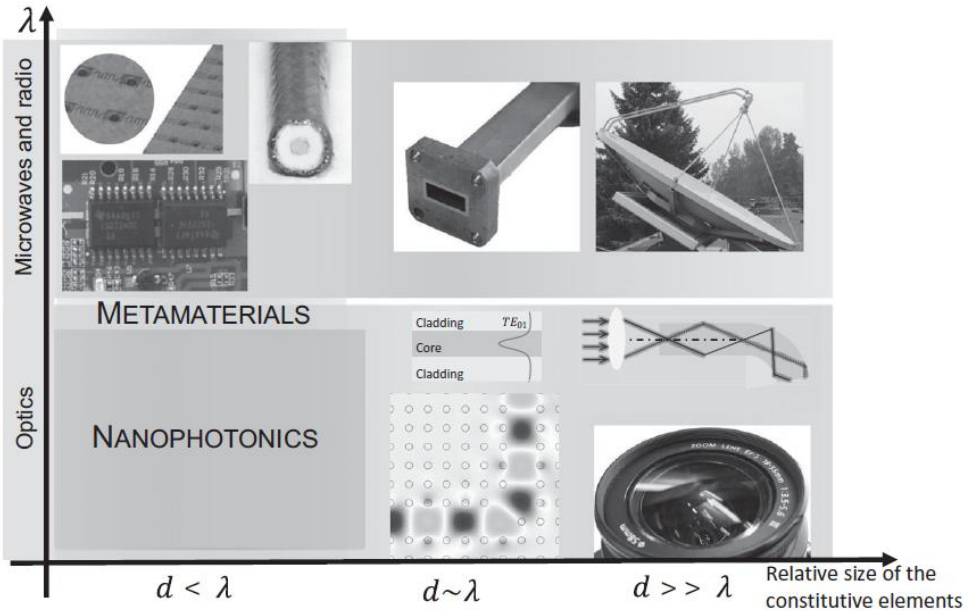

Fig 1. 1: Metamaterials and nanophotonics devices are necessary for submicrometer integration, but this field is still largely unexplored.

While for the use at microwave and radio frequencies we have a basic complete set of components (waveguides, passive, active, and nonreciprocal elements) as well as means to integrate them into extremely small devices, at the optical frequencies the corresponding box is only starting to be filled by optical nanoguides (such as plasmonic waveguides), metasurfaces, and metamaterial devices.

Only the intermediate regime where the characteristic sizes are comparable with the wavelength is well studied and developed: the photonic crystal technique is quite mature at this time. We see that any progress towards sub-micron integration of photonic devices is impossible without advances in metamaterials and nanophotonics.

One can ask why we need to develop integrated optical devices since we do have mature technology in electronics and microwaves?

Why do we really need a breakthrough if there are optoelectronic converters transposing signals from the optical band to the band of radio frequencies and back?

There are two main reasons why a breakthrough in this field is really demanded:

1. First, optoelectronic converters, even advanced ones, form a bottleneck in the modern telecommunication systems (see Chapter 8).
2. Second, there are two fundamental limitations of electronics and microwave technologies that block further qualitative progress in this field.
   - One is the thermal noise. According to the Planck law, the spectral density of thermal radiation at a certain temperature is given by Eq 1.1, where f is the frequency, T is the absolute temperature, h is the Planck constant, and k_B is the Boltzmann constant. In the optical domain,we have $h f >> k_B T$ even at room temperature, and the thermal noise is practically negligible. In contrast, thermal noise is a serious problem at radio and microwave frequencies (here, $h f << k_B T$).

$$P(f,T)=\frac{2hf^3}{c^2}\frac{1}{e^{\frac{hf}{k_B T}}-1} \qquad \text{Eq 1. 1}$$

   - The other limitation is imposed by the dispersion and power loss in interconnects between electronic chip components. The signal will be transported without distortion and information loss if the pulse shape does not significantly change upon transmission and the signal power does not strongly decay. Ideally, we would like to have a transmission line without dispersion. In optics, free-space propagation is dispersion-free, and we have low-dispersion and low-loss optical fibers. Non-dispersive waveguides have purely real characteristic impedance (in free space,

$\eta_0 = \sqrt{\mu_0 / \varepsilon_0}$). In electrical interconnects (metal transmission lines and cables), the impedance is determined by the per-unit-length parameters L,C,R:

$$Z_c = \sqrt{\frac{j\omega L + R}{j\omega C}} = \sqrt{\frac{L}{C}}\sqrt{1 + \frac{R}{j\omega L}}. \qquad \text{Eq 1. 2}$$

For propagation without pulse shape degradation we need to ensure that

$$R \ll \omega L, \qquad \text{Eq 1. 3}$$

then, $Z_c \approx \sqrt{\frac{L}{C}} = \sqrt{\mu_0 / \varepsilon_0} \times a$ geometrical factor, and we have as an ideal situation as in free-space propagation of light.

However, is this condition compatible with the submicron miniaturization of circuitry?

Let us make an estimate considering a short section of a transmission line (length *l*) formed by two metal conductors (the cross-section size *w* × *w*). The resistance *R* can be estimated as $\frac{1}{\sigma w^2}$, where σ is the metal conductivity. The inductance $L \approx \mu_0 l$ (there is also a cross-section shape-dependent factor of the order of unity, which we neglect). Substituting the room-temperature conductivity of copper at microwaves ($\sigma \approx 5.8 \times 10^7$ S/m), we see that at microwave and millimeter-wave frequencies the diameter of the connecting wire *w* should be larger than about 0.1−0.5 mkm, which is a severe limitation on device integration. In addition to pulse shape distortion, inevitable losses in metal conductors lead to signal propagation loss and device heating, which are other significant negative factors.

To summarize, one can expect that the next breakthrough in telecommunications, computing, imaging, etc., will come from advances in nanophotonics and metamaterials, which will also enable the replacement of radio and microwave signals by optical signals in signal processing and computations. We can say that nanophotonics and Metamaterials give engineers a unique chance to contribute to solving these global technological challenges.

1.2. Course content

1.2.1. Metamaterials

The electromagnetic properties of natural or chemically synthesized materials are determined mainly by their chemical composition. For realizing optical devices, the range of accessible material properties is rather limited. We have dielectrics (with moderate permittivity values, between 1 and about 12), metals (rather lossy negative-permittivity materials), and weakly chiral materials, and that is basically all. There is no natural magnetics, not to speak about more exotic and interesting media (like mu-near-zero or extreme chirality materials) that an engineer would need in designing optical nanodevices. Within the metamaterial paradigm, it becomes possible to widen the material design opportunities by engineering "meta-atoms" as constitutive elements of artificial materials using available ordinary (natural) materials.

In addition, the properties of meta-atoms can be modulated in time by some external force. The main appeal of the metamaterial concept derives from new possibilities to realize artificial materials with electromagnetic properties that are not available in any natural material and to engineer properties that are optimal for particular applications. Basic possibilities to control reflection and refraction using "usual" materials, metamaterials, and photonic crystals are illustrated in Figure 1.2.

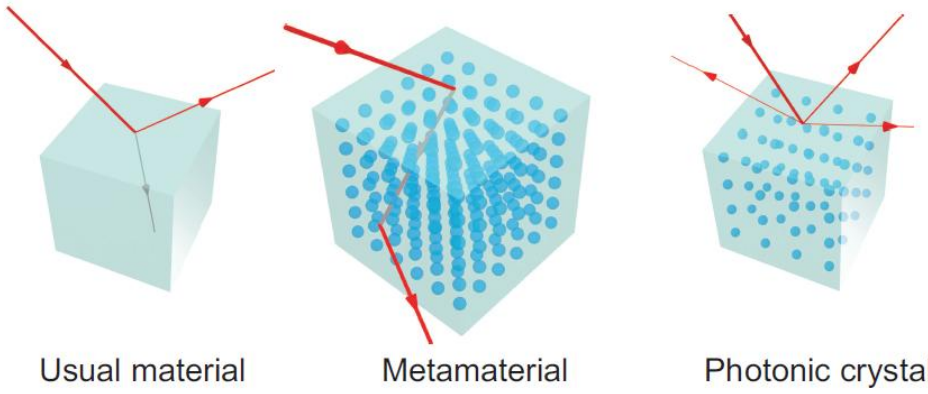

Fig 1. 2 Reflection and refraction regimes for samples of bulk materials and metamaterials.[1]

[1] Simovski, C., & Tretyakov, S. (2020). *An introduction to metamaterials and nanophotonics*. Cambridge University Press.

Research on artificial electromagnetic materials (now called metamaterials and metasurfaces) started at the end of the nineteenth century, very soon after the formulation of Maxwell's equations, but developments were slow. At present, this research field is extremely active and developing quickly. Perhaps the main current challenges are to learn how to realize materials with precisely the optimal properties for particular applications and to explore the whole physically allowed spectrum of material parameter values – in particular, including materials with extremely small or large values of the material parameters .

Currently, the main research focus is shifting from volumetric electromagnetic metamaterials to metasurfaces; see a conceptual illustration in Figure 1.3. All-dielectric metamaterials, especially all-dielectric metasurfaces, are actively studied for realisations in the visible part of the spectrum. Studies of nonlinear, bianisotropic, nonreciprocal, time-modulated, tunable, and programmable metamaterials and metasurfaces gain momentum.

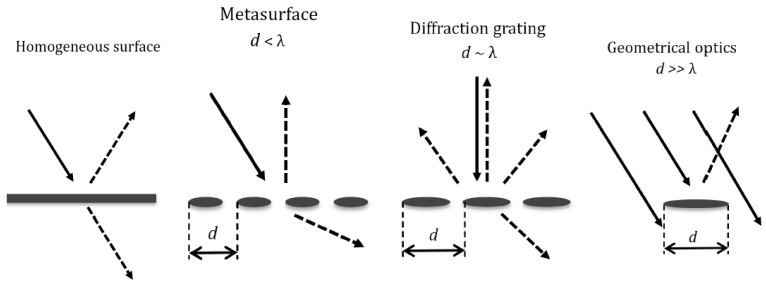

Fig 1. 3 Reflection and refraction regimes: from homogeneous sheets to optically large screens.[2]

Fundamental research on metamaterials leads to the developments of "metadevices" and systems. Last but not least, metamaterial technology comes into commercial products (e.g., products of Kymeta, Sensormetrix, Ecodyne). In the near future, one can hope to see!!

- Optimal (application-driven) designs of engineered materials and Reconfigurable, self-adapting, and software-defined metamaterials and metasurfaces
- Extreme properties and extreme-performance metamaterials
- Active, nonlinear, and parametric (time-varying) structures, including the use of non-Foster elements

[2] Simovski, C., & Tretyakov, S. (2020). *An introduction to metamaterials and nanophotonics*. Cambridge University Press.

- A full exploration of spatial dispersion (especially mesoscopic regimes between metamaterials and photonic crystals)
- Going into the quantum regime
- And more

1.2.2. Nanophotonics

Nanophotonics is the study of the behavior of light on the nanoscale and the interaction of submicron objects with light. It is a multidisciplinary scientific and technical area that comprises the most advanced parts of modern optics – including quantum optics, radio science (especially applied electromagnetics), and nanotechnology. It also concerns electrical engineering, solid-state physics, physical chemistry, biophysics, and biochemistry. In this course, we mainly concentrate on the electromagnetic part of nanophotonics where classical optics intersects with radio science. However, we will also study nanostructures for light-energy harvesting and conversion into electricity.

In its optical part, one of the most important targets of nanophotonics is the miniaturization of optical components (see also Section 1.1). For this purpose, one should learn how to squeeze macroscopic light beams – e.g., coming from an optical fiber, into a volume comparable to the wavelength or even subwavelength – and guide this concentrated light. In optoelectronic converters, guided signals should be transmitted to photodetectors: advanced photovoltaic diodes, CMOS, or charge-coupling devices that have been available with submicron sizes since the 1980s. In perspective all-optical signal processors and optical quantum computers, the squeezed signal needs to be transmitted to optical memory cells that may have a minimal size as small as 10 nm with a maximal size of the order of 500 nm. To record and erase information, such a cell needs subwavelength concentration of the field in a nanoguide with essentially a submicron cross-section. So sharp light concentration is not achieved by usual lenses or curved mirrors. The most elaborated approach to subwavelength light concentration in optics is the use of so-called plasmonic structures (see Chapter 7). These structures can be in the form of bulk plasmonic metamaterials or plasmonic metasurfaces. Thus, there is no boundary between nanophotonics and nanostructured metamaterials. This is why the present course unifies both nanophotonics and metamaterial topics.

An important target of nanophotonics is related to medical and biological applications – especially optical nanosensing and nanoimaging – sensing and imaging of submicron objects. These tools are of extreme importance for genetics, micro-, and molecular biology, and also for medical diagnostics. The challenge of a very weak interaction of light with so small objects as living cells, leucocytes, bacteria, and even single molecules is resolved using nanophotonics structures. Besides flat plasmonic surfaces used in a branch of optical nanoimaging called surface-plasmon microscopy, all other nanophotonic components used for optical nanosensing and nanoimaging have submicron structures.

They are photonic crystals, metamaterials, and even isolated plasmonic nanoparticles (in two branches of optical nanosensing called plasmon-enhanced fluorescence and plasmon-enhanced luminescence). However, the hottest nanophotonics topic nowadays is metasurfaces for moulding light at the nanoscale. Here, nanophotonics strongly intersects with a branch of optical nanosensing that had appeared long before the concept of nanophotonics and metamaterials was elaborated. This type of optical nanosensing is called surface-enhanced Raman scattering (SERS). SERS employs nanopatterned or textured plasmonic surfaces that perfectly match the basic definition of the metasurface. Nowadays, SERS is a part of nanophotonics (see Chapter 4) and all other methods of optical nanosensing and nanoimaging.

Returning to data processing, one should recall that squeezing the optical signal to the nanoguide is not enough. The signal has to be filtered and separated from other signals, preventing their cross-talks, amplification, etc. All these problems refer to nanophotonics. Photonic crystals enable very good nano guiding and separation of signals in devices based on these waveguides. The cross-section of photonic-crystal waveguides is still comparable to the wavelength: practically, of the order of one micron. However, they are traditionally subjects of nanophotonics because the majority of photonic crystals operating in the visible and telecom ranges are composed of submicron inclusions.

Filtering and amplifying squeezed signals is the subject of the so-called meta-tronics, which also covers subwavelength waveguides and some computational functionalities (see the corresponding chapter). Nanophotonics considers prospective submicron generators of optical signals called nanolasers. Next, nanostructures that enhance light conversion into electricity form an essential field of nanophotonics. It is known that solar light can be converted to electricity using photovoltaic devices called solar cells. The invisible (infrared) light produced by very hot bodies can also be converted. These devices are called thermophotovoltaic

generators. In both photovoltaic and thermophotovoltaic generating systems, nanostructures open the door to new mechanisms of light harvesting and will enable future technical breakthroughs. Therefore, the corresponding chapters of this book cover these two topics. Of course, our choice of topics reflects the scientific interests of the authors, and it is simply not possible to include all important subjects in this broad area of science and technology. In particular, we do not describe optical memory cells, nanolasers, optical transistors, switches, modulators, and logical gates, which are important nanophotonics components, mainly because these topics belong to the domain of solid-state physics and quantum physics.

Quiz & Control Questions

- Explain the meaning of the terms metamaterial and nanophotonics.
- Why are optical phenomena in nanostructures fundamentally different from phenomena in micrometre or millimetre-scale structures?
- What is the physical origin of thermal noise?
- How does the power density of thermal noise depend on the frequency and temperature? Use MATLAB to represent the spectral density of thermal radiation.
- Let us consider a communication system at room temperature in terms of thermal noise level. What is the optimal frequency range for operating it? What technology do we need to fabricate devices in this frequency range?

Chapter 02

Electromagnetic (Optical) Properties of Materials

Electromagnetic (Optical) Properties of Materials

2.1. Introduction

The study of the electromagnetic, or optical, properties of materials lies at the heart of understanding how light interacts with matter, a fundamental aspect of both classical and modern physics. This chapter delves into the behavior of materials when subjected to electromagnetic fields, particularly in the optical spectrum, which encompasses visible light and its neighboring wavelengths. The interaction of light with materials gives rise to a wide range of phenomena, including reflection, refraction, absorption, transmission, and scattering, all of which are governed by the material's intrinsic properties, such as its dielectric constant, magnetic permeability, and conductivity. These properties are not only crucial for explaining natural optical effects but also form the basis for designing and engineering advanced materials with tailored electromagnetic responses, such as photonic crystals, metamaterials, and plasmonic structures. In this chapter, we will explore the foundational principles that describe how electromagnetic waves propagate through different media, how materials respond to varying frequencies of light, and how these responses can be manipulated for practical applications. We will also examine the role of material microstructure, electronic band structure, and resonant phenomena in determining optical behavior. By understanding these concepts, readers will gain insight into the design of optical devices, such as lenses, filters, and sensors, as well as the development of cutting-edge technologies like invisibility cloaks, superlenses, and high-efficiency solar cells. This chapter serves as a bridge between theoretical electromagnetism and real-world applications, providing the tools necessary to analyze and innovate within the field of optical materials science.

2.2. Constitutive relations and material parameters

In the realm of macroscopic electromagnetics, all field vectors are subjected to averaging processes over electrically (or optically) diminutive volumes, each of which encompasses numerous atoms or molecules that constitute the material. In fact, the discourse concerning the electromagnetic properties of materials, whether they are of natural origin or artificially synthesized, is contingent upon the feasibility of such averaging. Should the macroscopic fields exhibit rapid spatial variation to the extent that the interatomic or inter-meta-atomic distances become comparable to the wavelength, one engages with photonic crystals or diffraction gratings, rendering the introduction of macroscopic material parameters to correlate macroscopic, volume-averaged, or surface-averaged field vectors impracticable. This chapter will exclusively investigate homogeneous natural and synthetic materials. The phenomena of

interference and diffraction predominantly govern the characteristics of photonic crystals, holograms, and analogous entities owing to their distinctive spatial attributes. The electromagnetic characteristics of natural materials are delineated through the application of constitutive relations, which articulate linear correlations among the four field vectors **E, H, D**, and **B**. In the prevalent instance of anisotropic linear materials, these relations are typically articulated to express the induction vectors as functions of the corresponding fields.

$$\mathbf{D} = \bar{\bar{\epsilon}}(\omega)\cdot\mathbf{E} + \bar{\bar{a}}(\omega)\cdot\mathbf{H} \qquad \mathbf{B} = \bar{\bar{\mu}}(\omega)\cdot\mathbf{H} + \bar{\bar{b}}(\omega)\cdot\mathbf{E} \qquad \text{Eq 2. 1}$$

or, equivalently, as

$$\mathbf{D} = \bar{\bar{\epsilon}}(\omega)\cdot\mathbf{E} + \bar{\bar{C}}(\omega)\cdot\mathbf{B} \qquad \mathbf{H} = \bar{\bar{\mu}}^{-1}(\omega)\cdot\mathbf{B} + \bar{\bar{D}}(\omega)\cdot\mathbf{E} \qquad \text{Eq 2. 2}$$

(the prime sign indicates that the permittivity $\bar{\bar{\varepsilon}}$ and permeability $\bar{\bar{\mu}}$ are, in general, different in the two formalisms). To address the phenomenon of frequency dispersion, all material characteristics exhibit dependence on frequency. Insignificant spatial dispersion arises as a consequence of the finite dimensions of the structural components and the interstitial distances, which is represented through the magnetoelectric coupling dyadic, $\bar{\bar{a}}$, $\bar{\bar{b}}$, and the permeability dyadic $\bar{\bar{\mu}}$.

The general bianisotropic equations (2.1) are well-suited for characterizing linear materials that can be represented on a macroscopic scale. Dielectric substances and conductors are generally represented solely by the permittivity tensor (or scalar in the case of isotropic materials). The field coupling coefficients $\bar{\bar{a}}$ and $\bar{\bar{b}}$, as well as $\bar{\bar{C}}$ and $\bar{\bar{D}}$, elucidate important characteristics like the optical activity of chiral materials and the nonreciprocal bianisotropic phenomena found in both natural ferrimagnetic crystals and metamaterials.

2.3. Frequency Dispersion

2.3.1. Linear Response and Causality

Reflect on the constitutive relationships pertaining to linear media within the time domain; specifically, the correlation of induced polarizations at a certain time t with the associated fields. In a vacuum, we can straightforwardly define:

$$\mathbf{D}(t) = \varepsilon_0 \mathbf{E}(t) \quad \text{Eq 2. 3}$$

In a vacuum, induced polarization is nonexistent. When electric fields are present in a polarizable medium, we incorporate:

$$\mathbf{D}(t) = \varepsilon_0 \mathbf{E}(t) + \mathbf{P}(t) \quad \text{Eq 2. 4}$$

where **P(*t*)** denotes the polarization vector (the electric dipole moment per unit volume). If our linear system remains stationary (indicating that its characteristics are constant over time), we can establish an integral relationship linking the polarization vector at time t with the macroscopic electric field values from all prior moments through a convolution integral:

$$\mathbf{P}(t) = \varepsilon_0 \int_{-\infty}^{t} \chi(t-t')\mathbf{E}(t')dt'. \quad \text{Eq 2. 5}$$

By restricting the integration to the interval t'∈ [−∞, t], we remove the chance of actions occurring "from the future" and guarantee that the response remains causal. Adding the free-space term $\varepsilon_0 E(t)$, for the displacement vector, we have a comparable linear integral relation.

$$\mathbf{D}(t) = \int_{-\infty}^{t} \varepsilon(t-t')\mathbf{E}(t')dt' = -\int_{\infty}^{0} \varepsilon(\tau)\mathbf{E}(t-\tau)d\tau = \int_{0}^{\infty} \varepsilon(\tau)\mathbf{E}(t-\tau)d\tau. \quad \text{Eq 2. 6}$$

Here, $\tau = t - t'$, and $\varepsilon(\tau) = \varepsilon_0\left[\delta(\tau) + \chi(\tau)\right]$, where $\delta(\tau)$ is the Dirac delta function.

The frequency-domain relationships follow from the Fourier transform.

$$f(\omega) = \int_{-\infty}^{\infty} f(t)\exp(-j\omega t)dt, \quad f(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(\omega)\exp(j\omega t)d\omega \quad \text{Eq 2. 7}$$

Among these temporal-domain connections, the Fourier transform applied to a convolution integral results in the multiplication of two transformations:

$$\mathbf{D}(\omega) = \varepsilon(\omega)\mathbf{E}(\omega), \quad \text{Eq 2. 8}$$

Where

$$\varepsilon(\omega) = \int_{0}^{\infty} \varepsilon(\tau)\exp(-j\omega\tau)d\tau \quad \text{Eq 2. 9}$$

is a complex permittivity. The integration is only relevant for positive values of the time delay τ, as dictated by causality. As we will explore in Section 2.2.2, this principle constrains the permissible frequency dispersion of permittivity and various material characteristics. Given that

all these effects are linear and causal, similar convolution integrals can be formulated for the magnetic response and magneto-electric coefficients, as shown in Eq (2.6). Frequency-dependent polarizability illustrates the material's reaction to monochromatic fields. However, since the medium's response is presumed to be linear (under weak fields), the theory is applicable to any temporal variation of fields and polarization. To examine arbitrary transient phenomena, Fourier expansions can be employed, which connect the Fourier components of fields and polarizations through the frequency-dependent permittivity.

2.3.2. Main Dispersion Mechanisms and Models

In this discussion, we concentrate on dielectrics and metals, exploring how permittivity changes with frequency (known as frequency dispersion). This phenomenon can be elucidated through a classical model of electron dynamics in the presence of an external electric field. We presume that the response of the material is linear, indicating that the applied field is sufficiently weak to induce polarization without modifying the intrinsic properties of the material. Consider polar dielectrics, which maintain electric moments even in the absence of an electric field (for example, water). When the substance is in a gaseous or liquid state, the average dipole moment equals zero without a polarizing field, due to the random thermal motion of the molecules. External fields align the molecules, yielding an average electric dipole that aligns with the external electric field (refer to Fig 2.1). If we remove the external field (assuming t = 0), the polarization vector will decline exponentially as a result of thermalization:

$$\mathbf{P}(t) = \mathbf{P}_0 e^{-t/T}, \qquad \text{Eq 2. 10}$$

where T is the relaxation time, which depends on the material properties and the temperature. It means that the kernel function in (Eq 2.6) and (Eq 2.9) is an exponential function:

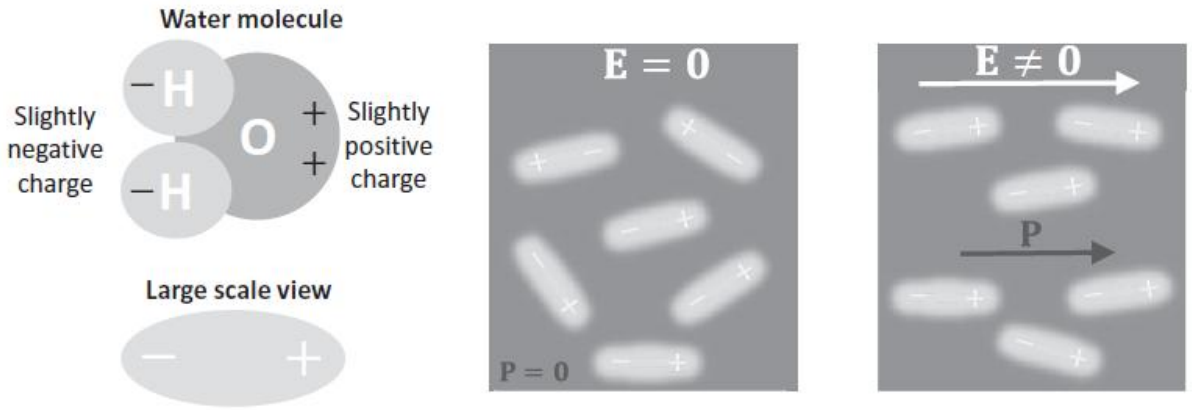

Fig.2. 1 Illustration of the polarization processes in polar dielectrics. Courtesy of A. Dyaz-Rubio.

$$\epsilon(\tau) = \epsilon_0 \left(\delta(\tau) + Ae^{-\tau/T} \right), \qquad \text{Eq 2. 11}$$

where A is a material-specific constant. The Fourier transform (Eq 2.9) yields

$$\epsilon(\omega) = \epsilon_0 + \frac{AT}{1 + j\omega T}. \qquad \text{Eq 2. 12}$$

More commonly, this result is written as

$$\epsilon = \epsilon_0 + \frac{\epsilon_s - \epsilon_0}{1 + j\omega T}, \qquad \text{Eq 2. 13}$$

Where ε_s is the permittivity at zero frequency (static permittivity). This expression is called the *Debye model*. Figure 2.2 depicts the usual relationship with frequency. Furthermore, let us examine the scenario where molecules become polarized solely upon the application of an external electric field. Take into account the oscillations of an electron around its equilibrium position r = 0 while an external electric field **E** is present (refer to Figure 2.3). The following forces act on the electron: the Lorentz force **F** = *e***E** (*e* is the elementary charge), a "return force"**Q** = −*q***r**, and a "friction force" $S = -\Gamma \frac{\partial r}{\partial t}$. When an electron is confined (as seen in dielectrics), a restoring force is present. The classical equation of motion is interpreted.

$$m\frac{\partial^2 \mathbf{r}}{\partial t^2} + \Gamma \frac{\partial \mathbf{r}}{\partial t} + q\mathbf{r} = e\mathbf{E} \qquad \text{Eq 2. 14}$$

(*m* is the electron mass). Its time-harmonic (assuming exp(*jωt*) time dependence) solution is

$$\mathbf{r} = \frac{\frac{e}{m}}{\frac{q}{m} - \omega^2 + j\omega\frac{\Gamma}{m}}\mathbf{E} = \frac{\frac{e}{m}}{\omega_0^2 - \omega^2 + j\omega\nu}\mathbf{E}. \qquad \text{Eq 2. 15}$$

According to its definition, the dipole moment is represented as **P = er**. If we consider N electrons for each unit volume (with N referred to as concentration), we can derive the dipole moment per unit volume (known as polarization): **P ≈ Np**. The displacement vector is characterized as:

$$\mathbf{D} = \epsilon_0\mathbf{E} + \mathbf{P} = \epsilon\mathbf{E}. \qquad \text{Eq 2. 16}$$

Thus, the permittivity reads

$$\epsilon = \epsilon_0\left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 + j\omega\nu}\right), \qquad \text{Eq 2. 17}$$

Where $\omega_p = \frac{Ne^2}{m\varepsilon_0}$. The parameter ω_p is called ***plasma frequency***.

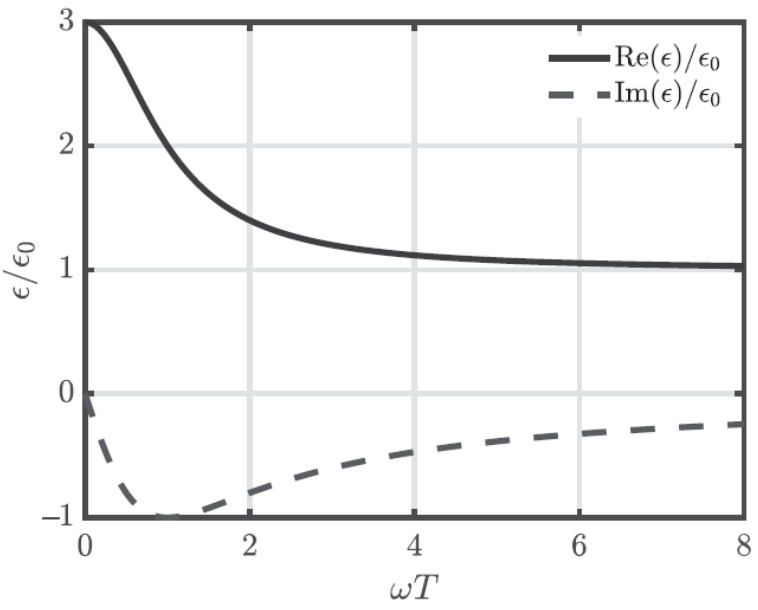

Fig.2. 2 Generic frequency dependence of permittivity on the frequency, Debye model. In this example, $\varepsilon_s = 3\varepsilon_0$.

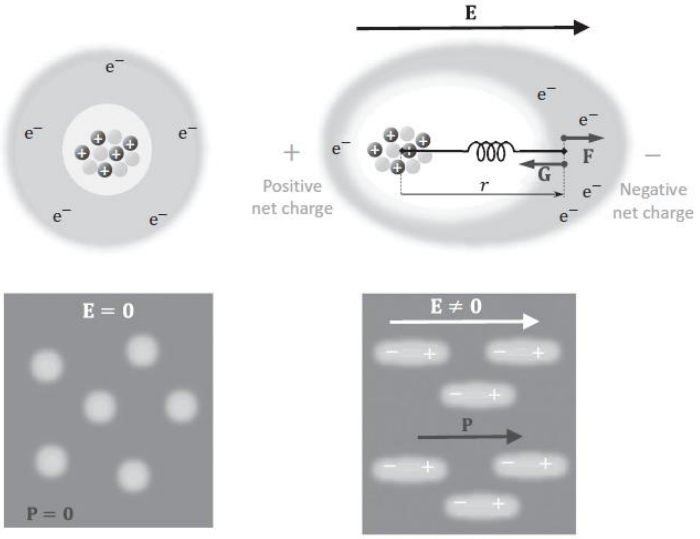

Fig.2. 3 Illustration of polarization processes in non-polar dielectrics. Courtesy of A.

This equation is referred to as the ***Lorentz*** dispersion model. Figure 2.4 illustrates a standard frequency response. This represents a typical dispersion characteristic observed in dielectric materials. More broadly, the dispersion in media containing bound electrons can be characterized as a summation of numerous components of this nature, which correspond to resonances at different frequencies.Let us now study metals, where conduction electrons are free to flow. This means that there is no return force, so we should set the parameter q in the preceding equations to zero. This is the same as setting the resonance frequency for freely moving electrons to zero, because $\omega_0^2 = q / m$. The permittivity function becomes

$$\epsilon = \epsilon_0 \left(1 - \frac{\omega_p^2}{\omega^2 - j\omega\nu} \right). \qquad \text{Eq 2. 18}$$

This expression is called the ***Drude model***. An illustration of the typical frequency dispersion can be seen in Figure 2.5.

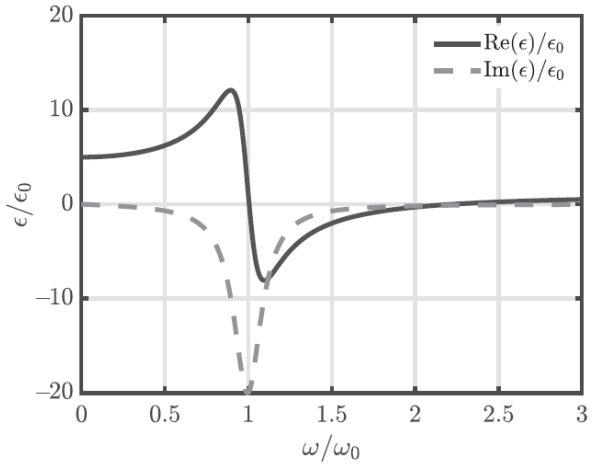

Fig.2. 4 Generic frequency dependence of permittivity on the frequency, ***Lorentz model***. In this example, $\omega_p/\omega_0 = 2$ and $\nu/\omega_0 = 0.2$

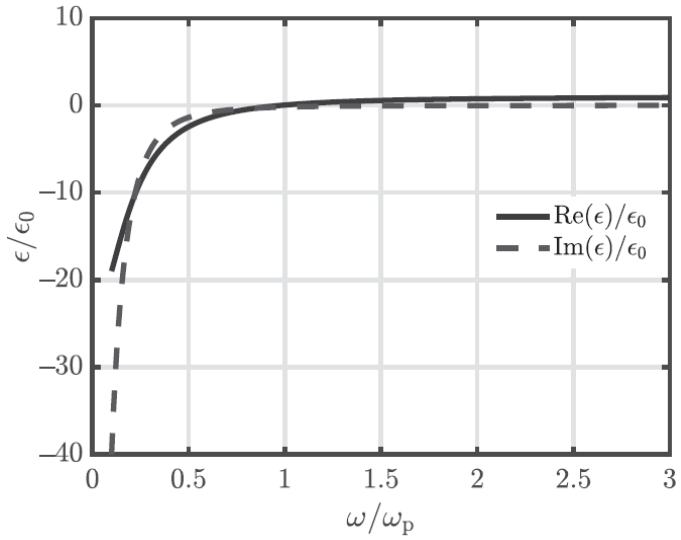

Fig.2. 5 Generic frequency dependence of permittivity on the frequency, ***Drude model***. In this example, $\nu/\omega_p = 0.2$.

Most often, various physical mechanisms of dispersion prevail in distinct frequency ranges. A prime example is seawater. Figure 2.6 illustrates the permittivity of seawater across a wide frequency spectrum, from static to ultraviolet. At lower frequencies, the behavior is primarily influenced by conductivity (according to the Drude model).

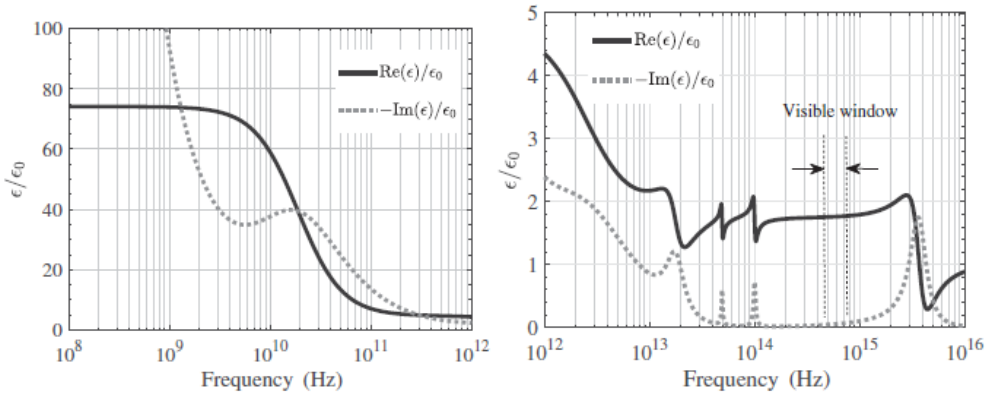

Fig.2. 6 Permittivity of seawater.

At elevated frequencies, we observe the relaxation resonance occurring in the microwave spectrum (*Debye* model) along with multiple *Lorentz* resonances present in the infrared and ultraviolet spectra. At significantly high frequencies, the permittivity approaches unity.

2.3.3. Kramers-Kronig Relations

Let us consider the frequency-domain permittivity $\varepsilon(\omega)$ (Eq 2.9) as a function of the complex variable ω, denoting $\omega = \omega' + j\omega''$ and $\varepsilon(\omega) = \varepsilon'(\omega) + j\,\varepsilon''(\omega)$. From the definition (Eq 2.9), it is clear that $\varepsilon(-\omega) = \varepsilon^*(\omega)$, $\varepsilon'(\omega) = \varepsilon'(-\omega)$, $-\varepsilon''(\omega) = \varepsilon''(-\omega)$. In passive media, $\varepsilon''(\omega) < 0$ at $\omega > 0$, which ensures that energy can be only dissipated but not generated by the medium.

The function $\varepsilon(\tau)$ the concept of "material memory" highlights that the polarization at a specific moment is influenced not only by the current field (τ = 0) but also by the values from the past (τ > 0). As τ increases, this function typically diminishes, illustrating the phenomenon of "fading memory." Regardless, in linear materials, we can infer that this function has a limit for all τ > 0. Under this assumption, we see from the definition (Eq 2.9) that $\varepsilon(\omega)$ is an analytical function in the lower halfplane of ω because the integrand has no singularities and exponentially decays at infinity, provided that $\omega'' < 0$.

It is crucial to understand that the underlying physical reason for this characteristic lies in the *causality* of the medium's response. This indicates that the polarization at any specific moment t can only be influenced by the fields from the past, which implies that τ must be greater than or equal to 0. Clearly, if we were to also expand the integration range in (Eq 2.9) to include

negative values of τ, the Fourier product would lose this property of analyticity. If the medium experiences loss, the function remains analytical along the real axis (in the case of perfectly lossless media, it may exhibit poles on the real axis). Note that the properties $1/\varepsilon$ are complementary: this function has no nulls in the lower half-space. Let us consider lossy dielectrics except for conductors. For the susceptibility $\chi(\omega)=\frac{\varepsilon(\omega)}{\varepsilon_0}-1=\varepsilon_r(\omega)-1$, we consider the function

$$\frac{\chi(\xi)}{\xi-\omega} \qquad \text{Eq 2. 19}$$

(real ω > 0). This function is analytic throughout the entire lower half-plane ξ (including the real axis), with the exception of a single simple pole located at ξ = ω. We can perform an integration of this function along a closed contour that runs along the real axis and then closes over the half-circle at infinity. The integral along the half-circle equals 0, as the function exhibits exponential decay at infinity. Consequently,

$$\mathrm{PV}\int_{-\infty}^{\infty}\frac{\chi(\xi)}{\xi-\omega}d\xi=-j\pi\chi(\omega). \qquad \text{Eq 2. 20}$$

In this context, ξ represents a real variable, and PV signifies the principal value of the integral. Splitting $\chi=\chi'+j\chi''$ into the real and imaginary parts, we get

$$\chi'(\omega)=-\frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty}\frac{\chi''(\xi)}{\xi-\omega}d\xi, \qquad \text{Eq 2. 21}$$

$$\chi''(\omega)=\frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty}\frac{\chi'(\xi)}{\xi-\omega}d\xi. \qquad \text{Eq 2. 22}$$

In the scenario of conductors, the permittivity exhibits a pole at ω = 0, resulting in an extra term in (Eq 2.22), which corresponds to the traditional expression for the conductivity contribution $\left[-j\sigma/(\omega\varepsilon_0)\right]$. These two relationships are credited to R. de L. Kronig (1926) and H. A. Kramers (1927). Using the property $\chi''(-\omega)=-\chi''(\omega)$, we can also write

$$\chi'(\omega)=-\frac{1}{\pi}\mathrm{PV}\int_{-\infty}^{\infty}\frac{\chi''(\xi)}{\xi-\omega}d\xi=-\frac{2}{\pi}\mathrm{PV}\int_{0}^{\infty}\frac{\xi\chi''(\xi)}{\xi^2-\omega^2}d\xi. \qquad \text{Eq 2. 23}$$

For nonconductors (assuming that $\chi''(\xi)$ is regular at ξ = 0), we can let ω = 0 and come to the relation

$$\chi'(0) = -\frac{2}{\pi}\int_0^\infty \frac{\chi''(\xi)}{\xi} d\xi > 0. \qquad \text{Eq 2. 24}$$

This relationship is referred to as the ***sum rule*** (multiple sum rules can be established in a similar manner). This outcome indicates, specifically, that in all passive (lossy) materials, the static value of susceptibility remains positive and the permittivity exceeds one: $\chi'(0) > 0$ and $\varepsilon'_r(0) > 1$.

Let us explore the regions of transparent frequency (where losses are minimal around ω). In this scenario, the pole contribution in (Eq 2.23) becomes null (since $\chi''(\xi) = 0$ at $\xi = \omega$). The relation for the real part becomes

$$\chi'(\omega) = \epsilon_{r'} - 1 = -\frac{2}{\pi}\int_0^\infty \frac{\xi\chi''(\xi)}{\xi^2 - \omega^2} d\xi \qquad \text{Eq 2. 25}$$

$\chi'' = \varepsilon_r''$, which is a standard, well-behaved integral. Therefore, we can apply differentiation under the integral sign. The outcome is expressed as:

$$\frac{d\epsilon_{r'}(\omega)}{d\omega} = -\frac{4\omega}{\pi}\int_0^\infty \frac{\xi\epsilon_{r''}(\xi)}{(\xi^2 - \omega^2)^2} d\xi > 0. \qquad \text{Eq 2. 26}$$

It indicates that in areas with minimal loss, the dispersion consistently remains "normal": the real component of susceptibility or permittivity increases alongside the frequency. This finding is known as the Foster theorem (since it was initially proposed by R. M. Foster for the reactances in electrical circuits).

2.4. Electromagnetic Waves in Materials

2.4.1. Dispersion Equation

Let us examine a dielectric material that exhibits anisotropic properties, defined by a permittivity matrix (dyadic) $\overline{\overline{\varepsilon}}$. The permeability is considered to be identical to that in a vacuum (μ0). Let us search for solutions to Maxwell's equations in a uniform medium devoid of sources, expressed as plane waves $e^{-jk.r}$. For the intricate amplitudes of plane-wave fields, the equations are expressed (substituting ∇ with −j*k*)

$$\mathbf{k}\times\mathbf{E}=\omega\mu_0\mathbf{H},\quad \mathbf{k}\times\mathbf{H}=-\omega\bar{\epsilon}\cdot\mathbf{E}. \qquad \text{Eq 2. 27}$$

The divergence equation tells that

$$\mathbf{k}\cdot\mathbf{D}=\mathbf{k}\cdot\bar{\epsilon}\cdot\mathbf{E}=0. \qquad \text{Eq 2. 28}$$

Eliminating vector $\mathbf{H}=\frac{1}{\omega\mu_0}\mathbf{k}\times\mathbf{E}$, we get

$$\mathbf{k}\times(\mathbf{k}\times\mathbf{E})=-\omega^2\mu_0\bar{\epsilon}\cdot\mathbf{E}. \qquad \text{Eq 2. 29}$$

By applying the concept of the double cross product and the relation **k × (k × E) = k(k · E) − E(k · k)**, the identical equation can also be expressed as:

$$\mathbf{k}(\mathbf{k}\cdot\mathbf{E})-k^2\mathbf{E}=-\omega^2\mu_0\bar{\epsilon}\cdot\mathbf{E}. \qquad \text{Eq 2. 30}$$

These equations can only be fulfilled for non-zero vectors E if the determinant of the matrix (dyadic) applied to the field vector is equal to zero:

$$\det\left[\mathbf{k}\times(\mathbf{k}\times\bar{I})+\omega^2\mu_0\bar{\epsilon}\right]=0. \qquad \text{Eq 2. 31}$$

where $\bar{\bar{I}}$ represents the identity matrix (dyadic). This expression is referred to as the dispersion equation, and it specifies the propagation vector as a function of frequency. An alternative form can be derived from (Eq2.30):

$$\det\left[\mathbf{kk}-k^2\bar{I}+\omega^2\mu_0\bar{\epsilon}\right]=0. \qquad \text{Eq 2. 32}$$

For isotropic dielectrics, where $\bar{\bar{\varepsilon}}=\varepsilon\bar{\bar{I}}$, the double cross product can be simplified to **k × (k × E) = k(k · E) − E(k · k) = −k²E** (given that, in this scenario, according to (Eq 2.28), it follows that k · E = 0), and the wave equation (Eq 2.29) is expressed as:

$$\left(-k^2+\omega^2\mu_0\epsilon\right)\cdot\mathbf{E}=0. \qquad \text{Eq 2. 33}$$

For the equation of dispersion, we have merely

$$k^2=\omega^2\mu_0\epsilon \qquad \text{Eq 2. 34}$$

and

$$k=\pm\omega\sqrt{\mu_0\epsilon}=\pm\frac{\omega}{v} \qquad \text{Eq 2. 35}$$

Here, $v = \frac{1}{\sqrt{\mu_0 \varepsilon}} = c\sqrt{\varepsilon_0 / \varepsilon}$ the phase velocity (refer to Section 2.3.3) of plane waves in a medium characterized by the permittivity ε. The two solutions, which differ by their sign, indicate that waves can propagate (at the same speed) in opposite directions along any straight line within an isotropic space. In reciprocal materials, the permittivity matrix is symmetric and can be diagonalized and fully characterized by its three eigenvectors and three eigenvalues: the permittivities along the eigenvector directions, called ***optical axes***. If all three eigenvalues are different, the medium is called biaxial (biaxial crystal). It is usually convenient to work in the coordinate basis formed by the eigenvectors of $\overline{\overline{\varepsilon}}$ (although, in lossy media, the eigenvectors can be different for the real and imaginary parts of $\overline{\overline{\varepsilon}}$).

Let us now turn our attention to perhaps the most intriguing and frequently encountered scenario in applications, uniaxial media, where two of the eigenvalues coincide. This physically signifies the existence of a single preferred direction (one optical axis), while the material remains isotropic in the plane that is perpendicular to that axis. Such substances are defined by a uniaxial permittivity dyadic:

$$\overline{\overline{\epsilon}} = \epsilon_t \overline{\overline{I}}_t + \epsilon_z \mathbf{Z}_0 \mathbf{Z}_0, \qquad \text{Eq 2. 36}$$

where z_0 represents the unit vector aligned with the optical axis (we consider it to be a real-valued vector) and It = $\mathbf{I} - \mathbf{z_0 z_0} = \mathbf{x0 x_0} + \mathbf{y_0 y_0}$ denotes the unit (identity) dyadic operating in the transverse plane. In matrix notation, the relationship (Eq 2.36) is expressed as follows:

$$\overline{\overline{\epsilon}} = \begin{pmatrix} \epsilon_t & 0 & 0 \\ 0 & \epsilon_t & 0 \\ 0 & 0 & \epsilon_z \end{pmatrix} \qquad \text{Eq 2. 37}$$

The dyadic representation (Eq 2.36) is often more practical as it expresses the permittivity through vectors that possess distinct physical significance (in this instance, the unit vector aligned with the optical axis $\mathbf{z_0}$), regardless of any specific coordinate framework. To investigate eigenwaves and the dispersion equation, it is most practical to utilize a coordinate system aligned with the direction of the optical axis (where one of the Cartesian axes is designated as $\mathbf{z_0}$) and the wavevector k of the wave under examination. Consequently, we choose the remaining two axes to align with the projection of the wavevector **k** onto the plane

that is perpendicular to z_0 (the vector $\mathbf{kt = k - (z_0 \cdot k)z_0}$) and along the vector $\mathbf{z_0 \times kt}$, which is perpendicular to both $\mathbf{z_0}$ and $\mathbf{k_t}$.

In this coordinate system, the electric field vector can be expressed as $\mathbf{E} = \mathbf{E}_t + \mathbf{E}\times + E_n\mathbf{z}_0$, where Et and E× represent the projections of the electric field along the directions of kt and z0 × kt, respectively. We can now reformulate the wave equation (2.30) in this coordinate system by substituting the electric field vector and the wavevector k, which is represented in these coordinates as $\mathbf{k = k_t + k_z z_0}$. It is clear that $\mathbf{k^2 = k_t^2 + k_z^2}$ and $\mathbf{k \cdot E = k_t E_t + k_z E_n}$. The vector equation (2.30) divides into three scalar equations: projections along the three axes. By writing down its determinant (the three rows in (Eq 2.38) correspond to the projections of (Eq 2.32) onto the directions of $\mathbf{k_t}$, $\mathbf{z_0 \times kt}$, and $\mathbf{z_0}$, respectively) and setting it to zero, we can transform (Eq 2.32) into:

$$\begin{vmatrix} -k_z^2 + \omega^2\mu_0\epsilon_t & 0 & k_z k_t \\ 0 & -k_z^2 + \omega^2\mu_0\epsilon_t & 0 \\ k_z k_t & 0 & -k_t^2 + \omega^2\mu_0\epsilon_z \end{vmatrix}$$

$$= (k_z^2 - \omega^2\mu_0\epsilon_t)[(-k_z^2 + \omega^2\mu_0\epsilon_t)(-k_t^2 + \omega^2\mu_0\epsilon_z) - k_z^2 k_t^2] = 0. \qquad \text{Eq 2. 38}$$

This dispersion equation seemingly divides into two separate equations:

$$k_z^2 + k_t^2 = k^2 = \omega^2\mu_0\epsilon_t \qquad \text{Eq 2. 39}$$

and

$$k_z^2\epsilon_z + k_t^2\epsilon_t = \omega^2\mu_0\epsilon_t\epsilon_z. \qquad \text{Eq 2. 40}$$

It indicates that two distinct modes can travel through the medium. For one of these modes, which follows (Eq 2.39), the wavenumber k remains entirely independent of ε_z, and its dispersion equation is, in fact, identical to that found in simple isotropic media (Eq 2.34). From a physical perspective, this implies that the wave lacks an electric field component along z_0, and consequently, it does not perceive that particular component of the permittivity matrix as being different. This mode is referred to as the ordinary wave. The second mode, which adheres

to (Eq 2.40), is known as the extraordinary wave. Its dispersion equation is frequently expressed in this manner (derived by dividing (Eq 2.40) by $\varepsilon_t\,\varepsilon_z$):

$$\frac{k_z^2}{\epsilon_t}+\frac{k_t^2}{\epsilon_z}=\omega^2\mu_0. \qquad \text{Eq 2. 41}$$

2.4.2. Dispersion Diagrams and Constant-Frequency Contours

Let's begin with the solution of the dispersion equation (Eq 2.35) for an isotropic dielectric (or free space), expressed as $\omega = vk$, where k represents the wavenumber and v denotes the speed of light within it. If we limit our analysis to light propagating in a single direction—specifically, along the x-axis—then ω as a function of the wavevector component k_x can be written as $\omega = v|k_x|$.

By allowing light to propagate within the (x−y) plane, we derive the function $\omega = v\sqrt{k_x{}^2 + k_y{}^2}$, and the graphical representation of this function is known as the light cone; refer to Figure 2.7(a). This cone encompasses all plane waves with $k_z = 0$. For any location within this cone, $k_x{}^2 + k_y{}^2 < k^2$. Consequently, $k_z = \sqrt{k^2 - k_x{}^2 - k_y{}^2} > 0$; meaning that the wavevector is real, and the area inside the light cone includes all possible propagating plane waves in the medium at any given frequency. For any location outside this cone, $k_x{}^2 + k_y{}^2 > k^2$. Thus, $k_z = \sqrt{k^2 - k_x{}^2 - k_y{}^2}$ is imaginary, and the area outside the light cone is associated with evanescent waves within the medium. When a plane ***ω = const*** intersects the light cone, the result is a circle, and the collection of these circles is referred to as the constant-frequency contours of the medium; see Figure 2.7(b).

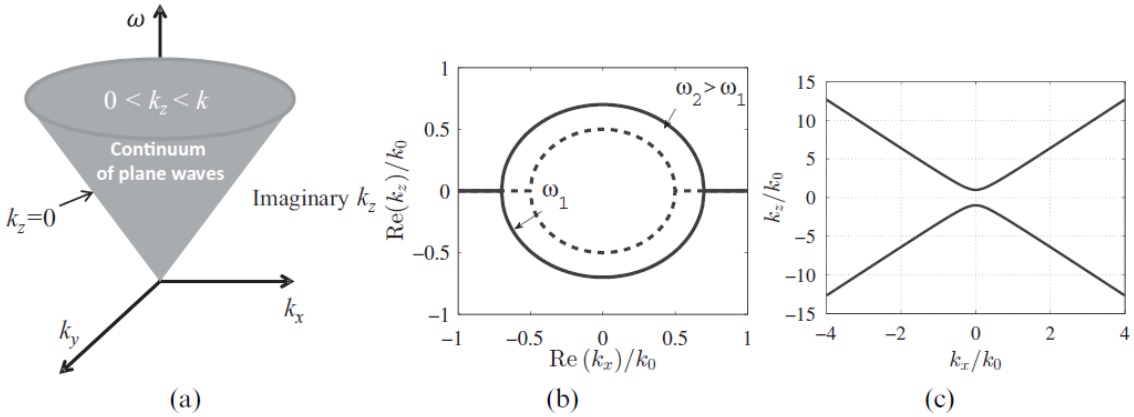

Fig.2. 7 (a) Light cone: Dispersion diagram of isotropic dielectric space (here, $k = \sqrt{\varepsilon\mu}$). The continuum of plane waves (real k_z) exists within the light cone, while the continuum of evanescent waves (imaginary k_z) lies beyond it. (b) Contours of constant frequency for an isotropic dielectric medium. (c) Plots of constant frequency for a hyperbolic material. In this example, $z = -0.1\varepsilon_0$ and $\varepsilon_t = \varepsilon_x = \varepsilon_y = \varepsilon_0$.

Constant-frequency contours are plots in the plane (k_x–k_y). For isotropic media, constant-frequency surfaces are spheres $\omega = v\sqrt{k_x^{\,2} + k_y^{\,2} + k_z^{\,2}}$ in the space (k_x–k_y–k_y).

This realm – that of the wavevectors associated with waves moving through the medium or crystal lattice – is referred to as the ***reciprocal space***. It's worth noting that both constant-frequency contours and constant-frequency surfaces are frequently referred to as isofrequencies. Understanding that ***isofrequencies*** are meaningful solely for waves traveling through the medium or the lattice (eigenwaves) is crucial.

For uniaxial media characterized by positive eigenvalues of the permittivity dyadic ε_t and ε_z , the constant-frequency surfaces are ellipsoids (more specifically, spheroids), as demonstrated in (Eq 2.41). In cases where one eigenvalue is negative while the other is positive, the forms indicated by (2.41) become hyperboloids; refer to Figure 2.7(c). Such materials are referred to as hyperbolic media, which exhibit remarkably intriguing and practically significant properties. It is important to note that the propagation constants of the propagating modes can attain large values, potentially much greater (theoretically infinite) than the propagation constant in free space. This characteristic presents opportunities for far-field imaging with subwavelength resolution.

Certain natural crystals exhibit permittivity characterized by the form (in the coordinate system where the symmetric matrix $\overline{\overline{\varepsilon}}$ is diagonal):

$$\overline{\overline{\epsilon}} = \begin{pmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{pmatrix} \qquad \text{Eq 2. 42}$$

In naturally occurring low-loss substances, all three eigenvalues $\varepsilon_{1,2,3}$ exhibit positivity, and the isofrequency surfaces take the form of non-spheroidal ellipsoids. Such materials are referred to as ***biaxial*** crystals.

2.4.3. Phase and Group Velocities

The relationship of electromagnetic fields in a plane wave characterized by complex amplitude $Ae^{-jk.r}$ with spatial coordinates (position vector r) and time t is expressed by $\mathrm{Re}\left[Ae^{-j(\omega t - k.r)}\right]$. Specifically, it is characterized by the wave ***phase*** $\phi(r,t) = \omega t - k.r$. At any given instant, surfaces of constant phase manifest as planes, described by the equation:

$$\mathbf{k}\cdot\mathbf{r} - \omega t = \text{const.} \qquad \text{Eq 2. 43}$$

As time progresses, the locations where the phase assumes a specific value move through space following the direction of vector ***k***. From (Eq 2.43), we observe that the phase propagation velocity is:

$$\frac{dr}{dt} = v = \frac{\omega}{k}, \qquad \text{Eq 2. 44}$$

where ***k*** represents the magnitude of vector k. This particular speed is referred to as ***phase velocity***. In diagrams depicting constant-frequency contours, the phase velocity vectors are aligned along the vectors of ***k***. The concept of phase velocity is established for plane waves at constant frequencies. If the medium exhibits dispersion (in dielectrics, where permittivity varies with frequency), the phase velocity will differ for waves of varying frequencies. Consequently, when we investigate the propagation of wave packets consisting of plane waves (for instance, finite-duration pulses transformed into Fourier spectra of plane waves), the concept of group velocity becomes essential, as it indicates the speed of the wave packet (the pulse). The peak

amplitude of a pulse occurs at the location where the plane-wave components of the pulse are in phase, which corresponds to the maximum value of phase (***ωt−k·r***) in relation to the wavevector ***k***. We can determine the extremum condition (such that, in a small neighborhood of this specific point, the phase remains stationary with respect to the wavevector) by equating the

derivative of the phase to zero:

$$\frac{\partial}{\partial \mathbf{k}}(\omega t - \mathbf{k}\cdot\mathbf{r}) = \frac{\partial \omega}{\partial \mathbf{k}} t - \mathbf{r} = 0. \qquad \text{Eq 2. 45}$$

This derivative represents the rate of change of the phase in the direction of vector ***k***; that is,

$$\frac{\partial}{\partial \mathbf{k}} = \frac{\partial}{\partial k_x}\mathbf{x}_0 + \frac{\partial}{\partial k_y}\mathbf{y}_0 + \frac{\partial}{\partial k_z}\mathbf{z}_0. \qquad \text{Eq 2. 46}$$

We observe that the location in space where, at a specific moment, the plane-wave components all combine constructively (the pulse maximum) advances through space at the speed of:

$$v_g = \frac{\partial \omega}{\partial \mathbf{k}}. \qquad \text{Eq 2. 47}$$

In practical computations, it is often more straightforward to determine the inverse quantity by differentiating the components of vector ***k*** concerning the frequency. Along the constant-frequency contours, the group velocity vectors are oriented perpendicular to these contours. Clearly, in anisotropic and dispersive media, the directions of group velocity and phase velocity typically differ.

2.5. Conclusion

In conclusion, the study of the electromagnetic (optical) properties of materials provides a fundamental framework for understanding how light interacts with matter, enabling the design and optimization of materials for a wide range of applications. From natural optical phenomena to advanced engineered structures like photonic crystals and metamaterials, the principles discussed in this chapter highlight the critical role of material properties in shaping electromagnetic responses. By mastering these concepts, researchers and engineers can develop innovative technologies that harness light in novel ways, paving the way for breakthroughs in fields such as telecommunications, energy, and sensing. This chapter underscores the importance of interdisciplinary approaches in materials science, combining physics, engineering, and technology to push the boundaries of what is possible with light and matter.

Quiz & Control Questions

- What is the physical interpretation of the Kramers-Kronig relations?
- For which materials are these relations inapplicable?
- In which types of materials can dispersion be accurately described by the Lorentz or Debye models?
- Vacuum is non-dispersive and maintains a constant relative permittivity of 1 across all frequencies. Consequently, according to the Kramers-Kronig relations, vacuum is devoid of losses – meaning its permittivity has no imaginary component. Is it feasible to create a non-dispersive material with a relative permittivity of 2?
- Examine the permittivity of good conductors (metals) at microwave frequencies utilizing the Drude model.

$$\epsilon = \epsilon_0 - \frac{\omega_p^2}{\omega^2 - j\omega\nu}. \qquad \text{Eq 2. 48}$$

In this case, $\omega^2 << \omega\nu$ (justify this assumption for typical metals like silver or copper from literature data on their parameters). Based on this premise, obtain approximate expressions for both the imaginary and real components of the permittivity (applying the Taylor expansion). Analyze the findings. In the literature, metals in the microwave range are typically represented by their conductivity σ, while the permittivity is generally considered to be roughly:

$$\epsilon = \epsilon_0 - j\frac{\sigma}{\omega}. \qquad \text{Eq 2. 49}$$

Does this estimation align with your findings derived from the Drude model? The real components of the permittivity are evidently (quite) distinct.

Examine which of the two models aligns more closely with reality. Seek out experimental data regarding the real component of the permittivity of metals in the microwave range (as documented in the literature). Be aware that it is occasionally represented by postulating that the conductivity σ is a complex number (while the permittivity is equal to ε_0).

Draw comprehensive conclusions from your research.

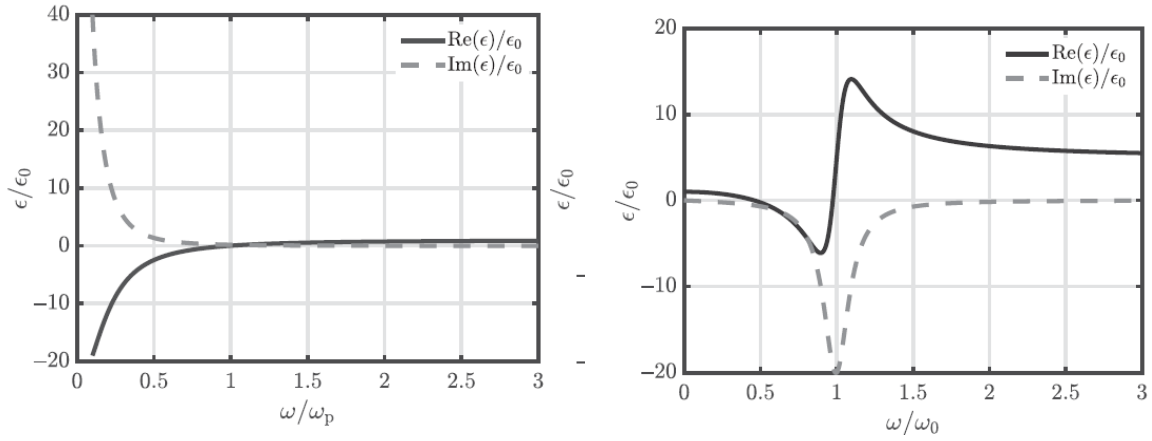

Fig.2. 8 Examples of nonphysical behavior of material parameters.

- Figure 2.8 illustrates the dependence of permittivity on frequency, resembling the dispersion plots of Drude and Lorentz. Nevertheless, these dependencies cannot occur in any passive material. Please clarify which fundamental physical principles prevent such behavior.[3]

[3] The curve for the real part of the permittivity on the right plot shows a typical example of the so-called *antiresonant* behavior, which can result from the use of inappropriate homogenization models for metamaterials with resonant unit cells;

Chapter 03

Photonic crystals

3.1. Introduction

3.1.1. What are photonic crystals

Photonic crystals can be defined as composites with a space-periodical internal structure operating at frequencies where the wavelength is comparable to the period of the structure. This is probably the most general definition, encompassing the most complex structures operating in various frequency ranges. Alternative names are ***electromagnetic crystals*** and ***bandgap structures***. The latter name is given because the frequency bands in which the electromagnetic waves cannot propagate (stopbands) are called bandgaps if they refer to photonic crystals. When such regular composites are designed to operate at optical frequencies (from the far infrared range to the visible and ultraviolet light), they are usually called photonic crystals. Building blocks of photonic crystals are made of usual (effectively homogeneous) materials, and since homogeneous materials in the optical range lose magnetic susceptibility, the internal periodicity of photonic crystals is often defined as a periodic coordinate dependence of the permittivity $\varepsilon(\mathbf{r}) = \varepsilon(\mathbf{r}+\mathbf{a})$, where $\mathbf{a} = (a_x, a_y, a_z)$ is the lattice unit vector. For isotropic dielectric media, the relative permittivity ε and the refractive index n are related as $\varepsilon = n^2$; thus, in this case, photonic crystals can be defined as artificial media with a space-periodic contrast of refractive index.

However, it is important to remember that the definitions in terms of the permittivity or refractive index are limited to specific (although very wide) classes of electromagnetic crystals. For example, periodical arrangements of magnetic or magneto-optical inclusions, the so-called ***magnetophotonic*** crystals, are very important nonreciprocal photonic crystals. Lattices of magnetically biased ferrimagnetic crystals (like yttrium iron garnet) or self-biased hexaferrite inclusions are conventionally used. In the microwave and millimeter ranges, these materials have tensorial magnetic permittivity with a non-zero antisymmetric part, responsible for nonreciprocal effects. In the optical range, these materials exhibit nonreciprocal magneto-optical effects due non-zero antisymmetric part of the permittivity tensor. In photonic crystals made of biased magnetic and magneto-optical materials, the nonreciprocal effects can be dramatically enhanced due to light localization. Stopband (bandgap) frequencies for the two orthogonal polarizations can be electrically controlled.

The wavenumber in magnetic and magneto-optical media depends not only on the frequency but also on the propagation direction. The refractive index for such material is simply meaningless, and the definition is not applicable. Furthermore, the definition in terms of

refractive index is not suitable for metallic photonic crystals which are also quite important for applications. Though metals in optics are reciprocal and isotropic materials, they do not possess refractive properties because of $\mathrm{Re}(\varepsilon_{metal}) < 0$.

Photonic or electromagnetic crystals can be alternatively considered as metamaterials with periodically arranged meta-atoms in cases when the metamaterial unit cell becomes comparable to the wavelength and spatial dispersion effects become very significant. These regular metamaterials are not effectively continuous media.

Similarly, metasurfaces in the frequency range where the period is comparable to λ behave as two-dimensional analogies of photonic crystals. These two-dimensional arrays exhibit stopbands for surface waves, similarly to stopbands for plane waves in volumetric photonic crystals. Obviously, the definition in terms of periodically varying permittivity does not apply to these planar stopband structures. Interestingly, it is possible to combine the notions of photonic crystals and metamaterials, considering photonic crystals are formed by complex inclusions performed of a metamaterial. Such photonic crystals are sometimes called ***metametamaterials***.

In this course, we will consider photonic crystals in the form of inhomogeneous media whose complex permittivity tensor is spatially periodic. This class covers all nonmagnetic classes of photonic crystals operating at optical frequencies, including anisotropic and metal lattices. Within this definition, we do not exclude the low-frequency region where the period is small on the wavelength scale so that the photonic crystal behaves as an effectively homogeneous medium. However, in this chapter, we do not consider interesting phenomena that may arise at low frequencies. The low-frequency region is included in our general consideration mainly to show a complete ***dispersion diagram*** of a photonic crystal – a plot whose ordinate axis starts from zero frequency.

Geometrically, volumetric photonic crystals split on to three classes: onedimensional (single-period lattices of layers), two-dimensional (doubly periodic crystals of cylinders, not necessarily circular ones), and three-dimensional (triply periodic lattices). Examples of such lattices are shown in Figure 3.1.

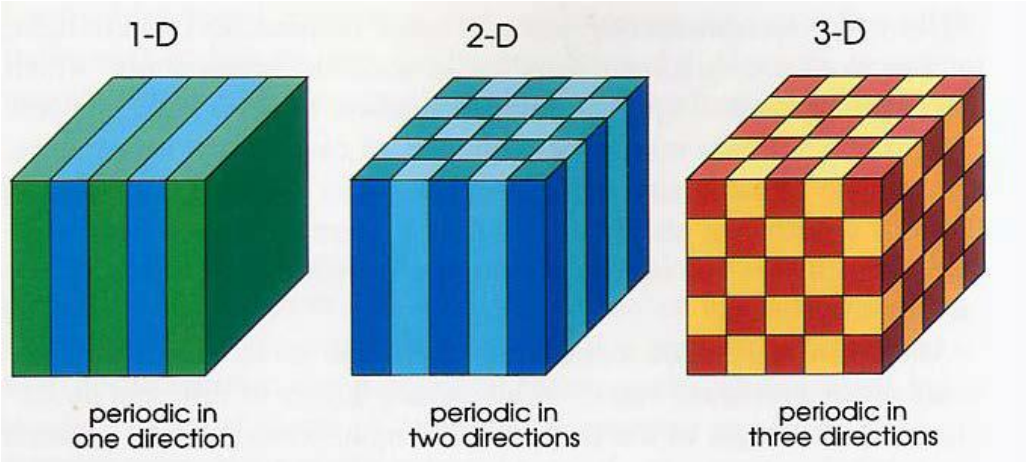

Fig.3. 1 Schematic illustration of the idea of **(a)** one-dimensional, **(b)** two-dimensional, and **(c)** three-dimensional photonic crystals.

One-dimensional and even two-dimensional photonic crystals had been known before this terminology appeared. The term "photonic crystal" arose about 30 years ago when the interest in spatially periodic optical composites sharply grew. The reason for this keen interest was the publication of two pioneering works in the same month of 1987: that by E. Yablonovitch and that by S. John[4]. These experimental papers reported two exciting effects: an amazing local enhancement of light intensity in photonic crystals with broken periodicity and inhibited spontaneous emission of quantum emitters in strictly periodic three-dimensional photonic crystals. Both these effects were theoretically predicted earlier[5]. However, these theoretical works have not produced a breakthrough. What has increased both theoretical and experimental investigation of photonic crystals was their fabrication and experimental confirmation of the effects predicted long ago.

[4] S. John, "Strong localization of photons in certain disordered dielectric superlattices," *Physical Review Letters* **58**, 2486–2489 (1987).

[5] V. P. Bykov, "Spontaneous emission in a periodic structure," Soviet Physics Journal of Experimental and Theoretical Physics 35, 269–273 (1972). V. P. Bykov, "Spontaneous emission from a medium with a band spectrum," Soviet Journal of Quantum Electronics 4, 861–871 (1975).

3.1.2. Natural Photonic Crystals

Among natural crystalline media, the only known photonic crystal is opal. Opal has a larger lattice unit cell than those of other solid crystals. Though in the visible and IR ranges of frequencies it also behaves as a continuous medium, it becomes a photonic crystal in ultraviolet light. However, natural photonic crystals operating in the visible range are also known. They are not natural solid crystals but parts of living systems. The two most famous examples of living photonic crystals are illustrated in Figure 3.2. The first one is the sea mouse hair in Figure 3.2, left panel. The sea mouse is a marine worm with a 15–20 cm long body living on the sea shelf at rather shallow depths. It is covered with thin, internally structured hairs also called tubular, spines. Each hair is a two-dimensional photonic crystal of submicron cylindrical holes oriented along the hair axis in the matrix of tissue whose refractive index is close to 1.5. When light is incident perpendicular to the axis of a hair, it exhibits red coloration. For off-axis incidences, green and blue shades appear. This optical effect is due to the so-called ***directional photonic bandgap***. For the normal incidence, the bandgap corresponds to the red light. The term bandgap means that the light whose frequencies lie in a certain band (in the present case, red light) cannot propagate inside the structure and, therefore, is totally reflected from it. The ***directional bandgap*** means that the band of frequencies for which this total reflection holds varies versus the incident light direction. When the incidence angle increases, the band of light totally reflected from the hair becomes yellow, green, and finally – for large angles – blue. This effect visually manifests in the angle-dependent coloration of the hair observed in the sunlight.

Another example of a photonic crystal in a living system is the moth-eye. The photonic crystal (a hexagonal lattice) is formed by identical dielectric protrusions with lengths of about 600 nm and the lattice step 400 nm (Figure 3.2, right panel). It operates as a super prism – a device with extremely high dispersion of light. This is an important nanophotonic device, and we discuss it in the next chapter. Here, we only mention that this super prism in the eye allows a butterfly to distinguish any more colors and shades that humans can do. Descriptions of some other examples of living photonic crystals can be found[6].

[6] P. Fratzl, J.W. C. Dunlop, and R.Weinkamer, *Materials Design Inspired by Nature*, RSC Publishing, 2013.

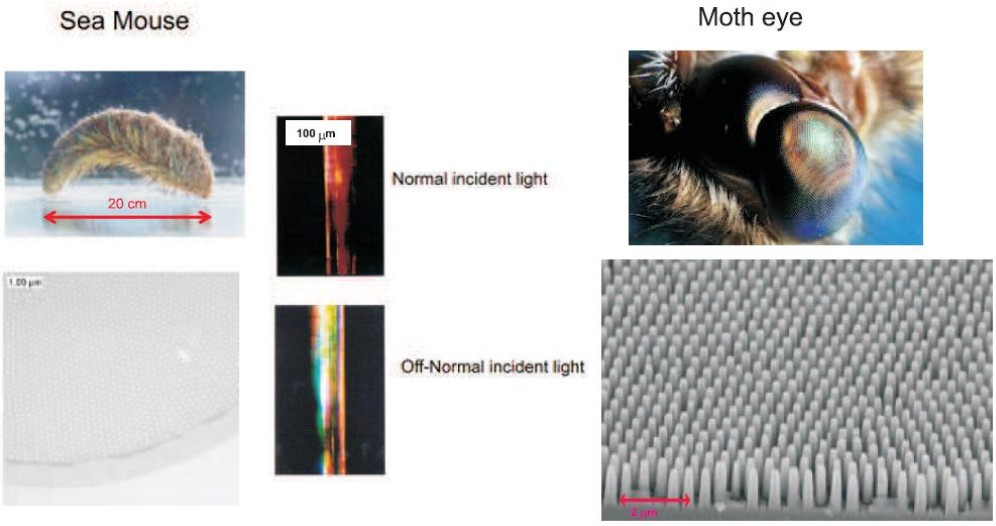

Fig.3. 2 Left panel: Sea mouse (general view); scanning electron microscope picture of the cross-section of its hair; its hair under normally (top) and obliquely (bottom) incident light. Right panel: Month eye (general view) and its nanostructure, obtained by transmission electron microscope.

3.2. Photonic Band Structures

3.2.1. Bragg Phenomenon as a Reason of Bandgaps

The key phenomenon inherent to every photonic crystal is the so-called ***Bragg phenomenon*** or Bragg's scattering. This phenomenon results from the in-phase interference of waves reflected from a set of parallel crystal planes. The phenomenon was explained theoretically by Lord Rayleigh[7] after W. L. Bragg and W. H. Bragg – had revealed a complete reflection of X-rays impinging natural crystals at a set frequency specific for a given crystal [20]. The theory by Rayleigh not only predicted this phenomenon for any periodic arrays of electromagnetic scatterers; it was further used in solid-state physics, where it was developed by F. Bloch and other scientists for electron wave functions in dielectric and semiconductor crystals. In natural crystals, an electron propagates through the crystal as a package of the so-called ***de Broglie waves***, and these particles are atoms (ions). In photonic crystals, constitutive particles are macroscopic inclusions in the host matrix, and propagating waves are electromagnetic–light – waves.

[7] V. P. Bykov, "Spontaneous emission in a periodic structure," *Journal of Experimental and Theoretical Physics* **35**, 269–273 (1972).

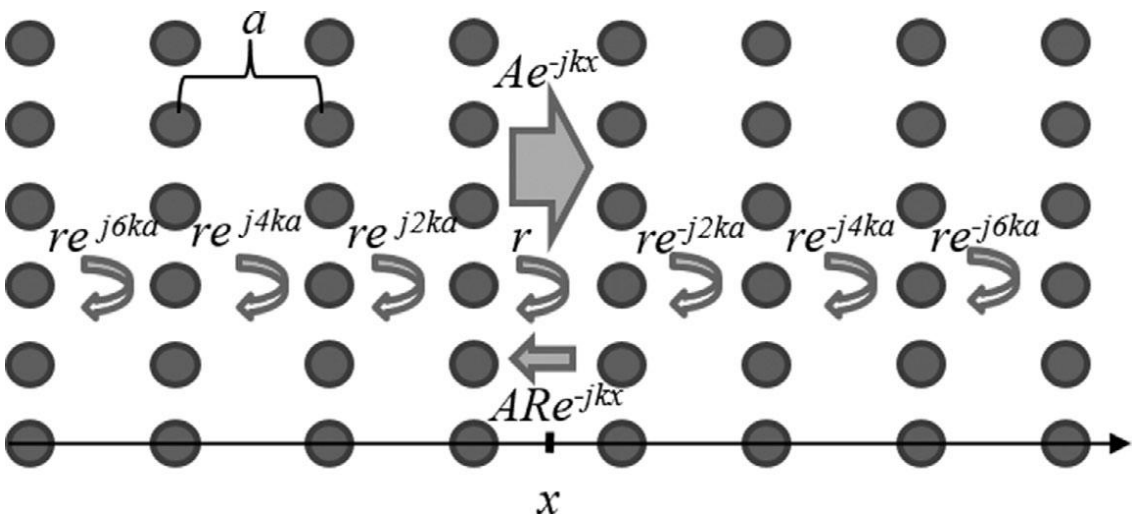

Fig.3. 3 Illustration of the Bragg phenomenon. In the uniform space between any two crystal planes, the propagating field of a lattice eigenmode can be considered as a sum of two waves traveling in opposite directions: one with the amplitude A propagating along x with the wavevector $\mathbf{k} = k\mathbf{x}_0$, and the other one propagating in the opposite direction with the amplitude RA. At any point x, this "reflected" wave is a sum of partial waves reflected by all crystal planes located behind x.

In Figure 3.3, we show the formation of the reverse-directed partial wave in the crystal as a summation of partial waves reflected from planar arrays of scattering particles. These planar arrays form a set of partially reflecting grids. If the grid period is sufficiently small, these grids can be thought of as effective semitransparent sheets. The reflection coefficient $\boldsymbol{R}$ at the reference plane x between two arbitrarily selected adjacent sheets results from multiple reflections that occur with the reflection coefficient r at each plane located behind x. Here, the reflection coefficient r is that of a single effective sheet, and $|r| < 1$ since the scatterers are passive objects. The path of the wave reflected from an arbitrary crystal plane is evidently larger by $2a$ than the path of the wave reflected from the previous crystal plane. It implies the phase shift $2ka$ for two waves reflected by adjacent crystal planes. Therefore, the result for R comprises a geometric progression with a factor $\exp(-j2ka)$:

$$R = re^{-jkx} + re^{-jkx}e^{-j2ka} + re^{-jkx}e^{-j4ka} + \cdots = re^{-jkx}\frac{1}{1-e^{-j2ka}}. \qquad \text{Eq. 3. 1}$$

The right-hand side of (Eq 3.1) diverges when $\exp(-2jka) = 1$, i.e. when $k \equiv \omega/v = m\pi/a$, where m is an integer number and v is the speed of light in the host medium. This physically meaningless divergence implies that the propagation of waves along x at frequencies $\omega_n = mv\pi/a$ is impossible. Formula $\omega_m = mv\pi/a$ (or its equivalent in terms of the wavelengths $\lambda_m = 2a/m$) is called the Bragg condition[8], which describes the frequencies (wavelengths) at which the crystal completely reflects normally incident plane waves. For the oblique propagation, the differential path $2a$ should be replaced by $2a \sin(\theta)$, and the set of frequencies for which $R \to \infty$ changes respectively.

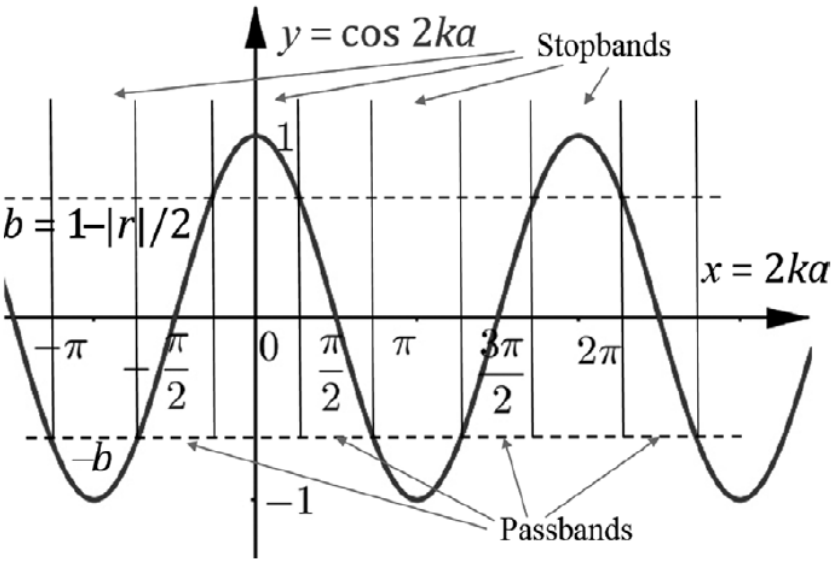

Fig.3. 4 Illustration of the concept of stopbands in a set of parallel equidistant crystal planes shown in Fig.3.3. If the absolute value of the cosine function exceeds b the absolute value of R exceeds unity.

The Bragg condition $R \to \infty$ gives a discrete set of frequencies (or wavelengths) at which the propagation is forbidden for the given direction of propagation. However, finite values of $|R|$ larger than unity are also prohibited. Really, the propagation of a wave with finite energy along x in a passive structure such as a lattice of scatterers cannot give birth to a wave with higher

[8] It makes sense if the period a is much larger than the other periods of the lattice. If they are comparable an additional set of prohibited frequencies arises due to the reections from other crystallographic planes.

energy. Imposing the condition $|R| < 1$ to the right-hand side of (Eq 3.1), we obtain after a simple algebra an inequality describing the passbands:

$$|\cos 2ka| < 1 - \frac{|r|^2}{2}. \qquad \text{Eq. 3. 2}$$

At these frequencies (recall that $k \equiv \omega/v$), the propagation along x is allowed, whereas, at frequencies for which (Eq 3.2) is not satisfied, the wave cannot propagate along x.

In Figure 3.4, we show the graphic solution of (Eq 3.2) for these frequencies and see that the lattice passbands and stopbands periodically alternate. In this example, r does not depend on ka, and our set of bandgaps includes the range of low frequencies, where $ka < \arcsin |r|/2$. This quasi-static bandgap does not arise for lattices of three-dimensional scatterers. Since the optical size of the scatterer vanishes in the quasi-static limit, for any planar array of finite-size scatterers, we have $|r| \to 0$ when $ka \to 0$. It is possible to prove that in this limit the absolute value of R in (Eq 3.1) keeps lower than unity. The same refers to the case when the lattice is formed by planar grids of dielectric cylinders.

However, a lattice formed by perfectly conducting wires (metal wires at radio frequencies can be thought perfectly conducting if we do not aim to study finer effects and only calculate the stopbands and passbands) has the low-frequency passband for its eigenmode whose electric field is parallel to the axes of the wires. In this case, $|r|$ does not vanish in the quasi-static limit, and the formula (Eq 3.2) stays relevant. If the electric field is orthogonal to the wires, $|r|$ tends to 0 when ka tends to 0. The same refers to the grids of dielectric cylinders and the grids of three-dimensional scatterers. For this mode, the low-frequency bandgap in a lattice of parallel wires does not exist, either.

If we consider an oblique propagation in the same lattice, we obtain a set of passbands and stopbands that depend on the propagation direction. By specifying a frequency and a propagation direction, we specify the wavevector; therefore, generalizing the ***Rayleigh*** formula (Eq 3.1) for arbitrary direction allows us to relate the eigenfrequencies in the passbands to the wavevector. This relationship, which is different from that established for continuous media, manifests ***spatial dispersion*** of photonic crystals. A directional bandgap is the most spectacular manifestation of spatial dispersion. It means that in a certain sheer of directions, a signal of a

given frequency cannot propagate, whereas, in all other directions, it can propagate. This is impossible in continuous media, even in anisotropic media.

In realistic photonic crystals, the reflection coefficient r is itself dependent on the frequency. Therefore, the width of passbands and stopbands varies over the frequency axis. Moreover, many photonic crystals have a lattice more complex than the orthorhombic (rectangular) lattice that is depicted in Figure 3.3. This makes the band structure quite complicated. Therefore, the Bragg conditions $\lambda_m = 2a/m$ (for normal incidence) and $\lambda_m = 2a \sin\theta/m$ (for oblique incidence), which are very important for natural crystals, are not so useful for photonic crystals. For natural crystals, the Bragg conditions determine the frequencies ω_m of X-rays that completely reflect from the crystal for the given incidence direction.

Neglecting electromagnetic (optical) losses (as we did earlier), we have to prohibit the propagation of X-rays not only at frequencies $\omega_m = mc\pi/a$ (except that with $m = 0$, for which r=0) but also in the stopbands centering these frequencies. However, high optical losses of natural crystals in the range of X-rays reduce the impact of the stopbands and make the absolute value of the reflection coefficient smaller than unity at all frequencies. It maximally approaches unity at Bragg frequencies, where the impact of the optical loss turns negligible compared to that of the interference.

It is not so for photonic crystals. Here, the constitutive particles are usually substantial in size, as compared to the lattice period. The reflection coefficient r of a crystal plane in a photonic crystal strongly depends on both frequency and incidence angle. It results in a more complicated spatial dispersion than that of X-rays in solids. Also, photonic crystals operate at the visible and infrared light frequencies, where optical losses – at least for dielectric photonic crystals – are lower than those inherent to solids in the range of X-rays. The reflection of the electromagnetic wave from such photonic crystals is practically complete over the whole stopband.

Therefore, the Bragg frequencies for photonic crystals are not very relevant (moreover, it is difficult to estimate them due to the strong frequency dispersion of r). Instead of Bragg frequencies, for photonic crystals, one calculates the set of bandgaps and passbands. The Rayleigh approach to this calculation expressed by formula (Eq 3.2) is not very helpful, even though we know the frequency dispersion of r. In many photonic crystals, this value is modified by interaction between adjacent crystal planes and cannot be calculated separately. Moreover, calculating the positions of bandgaps on the frequency axis is insufficient to describe the

electromagnetic properties of photonic crystals. In this chapter, we will discuss which characteristics of photonic crystals need to be calculated and how to calculate them.

3.2.2. Bandgap Structures in General

Since there are similarities in the theories of electromagnetic waves in photonic crystals and theories of electron waves in crystalline semiconductors, and solid-state physics was developed earlier, the terminology of photonic crystals is largely borrowed from solid-state physics. Stopbands are usually called bandgaps, and one distinguishes *complete* and *directional* bandgaps. In a complete bandgap, the propagation is forbidden in all directions and for all polarizations of waves. An optical emitter embedded into a three-dimensional photonic crystal cannot radiate photons with the frequencies inside a complete bandgap. Respectively, a complete bandgap is sometimes called ***photonic bandgap***. If the propagation is forbidden for some directions and allowed for the other ones the corresponding frequency band is called directional bandgap.

Most of three-dimensional photonic crystals have only directional bandgaps. Complete bandgaps are possible in diamond-like lattices and in so-called ***inverted opal*** lattices. As to simple lattices, such as the cubic or orthorhombic ones, a complete bandgap is possible if the inclusions are resonant. In the optical range, they can be, for example, metallic nanoparticles. Even simple-shape nanoparticles such as nanospheres of some metals experience the so-called ***plasmon resonances*** in the range of the visible light. From the applications point of view, plasmonic structures have the disadvantages of larger dissipation losses. Here, we do not consider photonic crystals formed by resonant inclusions. We concentrate on photonic crystals of non-resonant constituents.

For one-dimensional photonic crystals, the relevant propagation direction is usually fixed in the direction normal to the layers, and one does not distinguish complete and directional bandgaps. For two-dimensional photonic crystals – lattices of cylindrical inclusions parallel to a certain axis – the complete bandgap as a rule implies prohibited propagation in the plane perpendicular to the axis. The exceptions are so-called ***photonic-crystal fibers*** in which only the propagation along the axes of cylinders is allowed (these cylinders are practically air voids in the fiber glass). Thus, not only the propagation across the fiber but also the oblique propagation should be prohibited.

Atypical sample of a two-dimensional photonic crystal represents an optically thick dielectric layer with optically contrast (another dielectric or void) cylindrical inclusions stretched across the layer. In this situation, the finite length of the cylindrical inclusions plays no role: cylinders occupy the whole volume between the boundaries of the photonic crystal slab. The photonic bandgap prevents leakage of electromagnetic energy through the layer perimeter. The waves may propagate along the axes of the cylinders and obliquely to them but not orthogonally. This means that leakage of waves is possible through the top and bottom boundaries of the layer – which, in fact, can be useful for applications. One interface is usually mirrored, and the light is radiated by the second interface. In the next chapter, we will discuss this application of two-dimensional photonic crystals.

Notice that two-dimensional photonic crystals of metal rods in a dielectric host are also known and used in some applications. However, they are more practically important in the low-frequency regime than in the photonic-crystal regime. In the low-frequency regime – i.e., when the period is much smaller than the wavelength – such arrays are called ***wire media***. Wire media refer to spatially dispersive metamaterials and deserve a separate study. We refer the reader to an overview. As to the frequency range where a wire medium becomes a metallic photonic crystal described by a set of passbands and stopbands, it found applications only at microwaves[9]. In microwave electromagnetic bandgap structures, the wire diameters are in the millimeter and submillimeter ranges, whereas in the optical wire media, they are ***nanowires***. Since metals in the optical range are rather lossy, metallic photonic crystal are usually inefficient. In both passbands and stopbands,waves propagate with decay, and the difference of decay rates in a passband and in a stopband is not very large. In optics, dielectric photonic crystals are practically more important than metallic ones.

[9] C. Jin, B. Cheng, B. Man, D. Zhang, S. Ban, B. Sun, L. Li, X. Zhang, and Z. Zhang, "Two-dimensional metallodielectric photonic crystal with a large band gap," *Applied Physics Letters* **75**, 1201–1205 (1999).

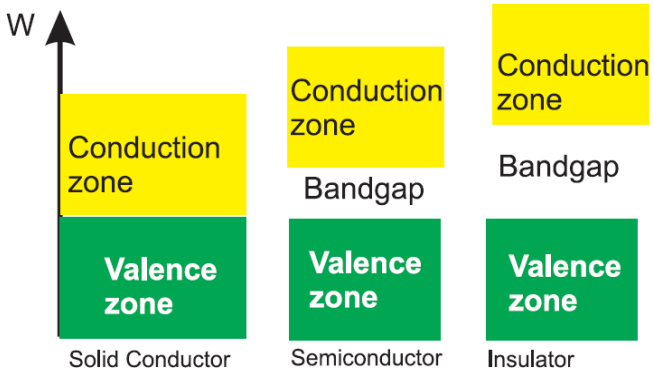

Fig.3. 5 Band structures of electron states in solid conductors, semiconductors, and insulators. Potential energyW_c of an electron in the conductive state is related to the frequency of optical transition ω as $W_c = \hbar\omega + W_v$, whereW_v is the electron potential energy in the valence zone.

As to the analogy of photonic crystals to natural ones, it is straightforward only for three-dimensional photonic crystals, because natural ones are bulk media of atoms (ions). Energy zone structures for all three solid states – conducting, semiconducting, and insulating – are illustrated in Figure 3.5. This band diagram shows only the ***main bandgap*** – that between the valence band and the conduction band. Three-dimensional photonic crystals without a complete bandgap are analogous to metals, those with a narrow complete bandgap are analogous to semiconductors, and those with a broad complete bandgap – to insulators.

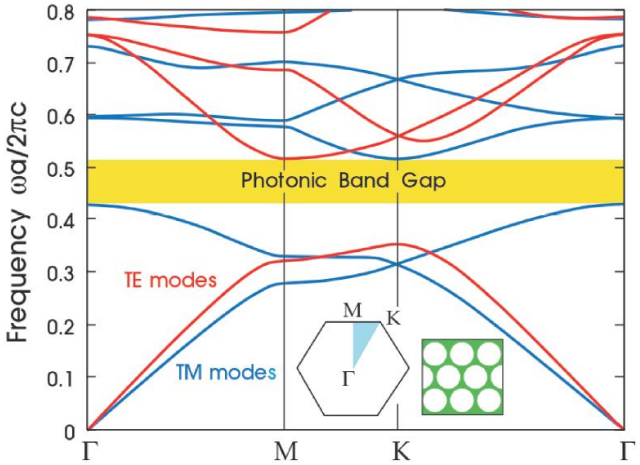

Fig.3. 6 Dispersion diagram and bandgap structure of a honeycomb lattice of holes in a dielectric matrix. The bandgap is shown in the dispersion diagram by yellow color, and the lattice structure is depicted in an inset together with the so-called *Brillouin zone*.

This simplified band diagram does not allow us to see the directional bandgaps and, in general, does not tell anything about spatial dispersion (relations between eigenfrequencies and wavevectors). The spatial dispersion and the band structure of photonic crystals are visualized in so-called ***Brillouin dispersion diagrams*** – plots of the frequency (or photon energy) versus the wavevector of the propagating eigenmode. An example of the dispersion diagram is shown in Figure 3.6. In the following paragraphs, we learn this graphic tool and relate it with another graphic tool – constant-frequency surfaces and contours, also called ***isofrequencies***.

We have briefly discussed these tools in Chapter 2. Now, let us recall that the dispersion diagram of an isotropic dielectric space corresponds to the dispersion equation $\omega = \nu k$, where $k = \sqrt{k_x^2 + k_y^2 + k_z^2}$ is the wavenumber and $\nu = c/n = c/\sqrt{\varepsilon}$ is the phase velocity in the dielectric medium (ε is its relative permittivity, and n is the refractive index). For $k_z = 0$ (in-plane propagation), the dispersion diagram is a light cone. A section of this light cone by a plane $\omega =$ const is a circle of radius $k = \omega/\nu$ in the coordinates (k_x–k_y), and this circle is an isofrequency contour of the uniformisotropic space. Higher frequencies correspond to larger circles in the plane (k_x–k_y). Thus, isofrequency contours in the plane ($k_x - k_y$) are horizontal sections of the dispersion plot $\omega(k_x, k_y)$ by the planes of constant ω. This link of the dispersion diagram and the isofrequency contours holds for photonic crystals as well.

3.2.3. Bloch's Theorem

In order to understand the dispersion diagram of an infinite periodic structure, we need to know a very important theorem concerning eigenwaves of a periodic structure. It was formulated and proven by F. Bloch for wave functions of electrons in natural crystals and further generalized to all periodic structures supporting electromagnetic waves, acoustic waves, plasma waves, etc. This became possible because electron wave functions obey the Schrödinger equation, which in the linear regime reduces to the wave equation like Maxwell's equations do in linear media. The same refers to the linearized dynamic equations for gases, plasmas, and liquids: they all reduce to the wave equation.

For simplicity, let us consider a one-dimensional photonic crystal with the period a, shown in Figure 3.7. In accordance with the ***Bloch theorem***, complex amplitudes of each monochromatic eigenmode of this structure (in our case, the eigenmode is formed by electric **E** and magnetic **H** fields) can be presented as a sum of the following waves:

$$\begin{Bmatrix}\mathbf{E}(\mathbf{r})\\ \mathbf{H}(\mathbf{r})\end{Bmatrix} = \begin{Bmatrix}\mathbf{E}_0(X,Y)\\ \mathbf{H}_0(X,Y)\end{Bmatrix} e^{-jq_z z} + \sum_{n\neq 0}\begin{Bmatrix}\mathbf{E}_n(X,Y)\\ \mathbf{H}_n(X,Y)\end{Bmatrix} e^{-j(q_z+G_n)z}, \qquad \text{Eq. 3. 3}$$

where $G_n = 2\pi n/a$, n is an integer, and the time dependence of fields is assumed in form $exp(j\omega t)$. Parameter q is called the ***Bloch wavenumber***. The Bloch theorem evidently generalizes to two-dimensional and three-dimensional photonic crystals. For a three-dimensional photonic crystal with the periods (a_x, a_y, a_z), we write $\mathbf{n} = (n_x, n_y, n_z)$ and $\mathbf{G_n} \equiv (G_x,G_y,G_z) = (2\pi n_x/a_x, 2\pi n_y/a_y, 2\pi n_z/a_z)$ and arrive at the Bloch plane-wave expansion for the phasors of electric and magnetic field vectors:

$$\begin{Bmatrix}\mathbf{E}(\mathbf{r})\\ \mathbf{H}(\mathbf{r})\end{Bmatrix} = \begin{Bmatrix}\mathbf{E}_0\\ \mathbf{H}_0\end{Bmatrix} e^{-j(\mathbf{q}\cdot\mathbf{r})} + \sum_{n\neq 0}\begin{Bmatrix}\mathbf{E}_n\\ \mathbf{H}_n\end{Bmatrix} e^{-j(\mathbf{q}\cdot\mathbf{r}+\mathbf{G}_n\cdot\mathbf{r})}. \qquad \text{Eq. 3. 4}$$

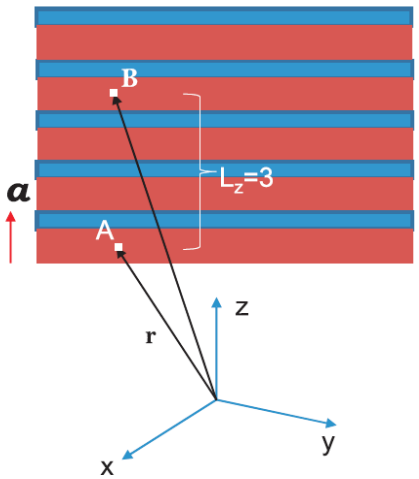

Fig.3. 7 Let point B in an infinite photonic crystal with period a in the z-direction be obtained from point A shifted by L_z periods along the coordinate axis. Then the electromagnetic field of the eigenmode at these two points differs only by the phase shift $L_z\mathbf{q} \cdot \mathbf{a}$, where L_z is an integer and **q** is called the Bloch vector of the eigenmode.

In formulas (Eq 3.3) and (Eq 3.4), the fundamental ($n = 0$) spatial harmonic whose wavevector is equal to **q** is shared out from the infinite plane-wave expansion. This plane wave is called the ***Bloch wave***. In the one-dimensional case, the value of this fundamental ***Bloch wavenumber*** $q \equiv q_z$ is restricted by the condition $-\pi/a < q < \pi/a$. All spatial harmonics with $n \neq 0$ oscillate with smaller spatial periods. If we assume that in the Bloch expansion the absolute value of the Bloch wavenumber is higher (e.g., $\pi/a < q < 2\pi/a$), this expansion term will be not distinguishable from that with the wavenumber ($q - 2\pi/a$), which enters the second term of expression (Eq 3.3):

the series of higher-order harmonics. Then, the replacement $q_{new} \rightarrow (q - 2\pi/a)$ will return the correct Bloch wavenumber with $-\pi/a < q_{new} < 0 < \pi/a$. Similarly, in the three-dimensional case, the components of the ***Bloch wavevector*** satisfy $-\pi/a_{x,y,z} < q_{x,y,z} < \pi/a_{x,y,z}$.

Bloch's theorem has two important implications. First, the problem of the infinite lattice can be reduced to the so-called cell problem. Really, from (Eq 3.3) the so-called ***Bloch quasi-periodicity conditions*** follow:

$$\begin{cases} \mathbf{E}(\mathbf{r} + L_z\mathbf{a}) \\ \mathbf{H}(\mathbf{r} + L_z\mathbf{a}) \end{cases} = \begin{cases} \mathbf{E}(\mathbf{r}) \\ \mathbf{H}(\mathbf{r}) \end{cases} e^{-jL_z\mathbf{q}\cdot\mathbf{a}}. \qquad \text{Eq. 3. 5}$$

For three-dimensional crystals, $L_z\mathbf{q}\cdot\mathbf{a}$ reads $L_xq_xa_x+L_yq_ya_y+L_zq_za_z$. The special case $L_{x,y,z} = 1$ shows that the fields at the opposite faces of the unit cell differ only by the phase shift $\mathbf{q} \cdot \mathbf{a}$. The boundary problem for the wave equation in a spatial domain of the lattice unit cell in the case when the a priori unknown wave functions on the front and back sides of the cell are related to one another in a known way is called in mathematical physics the ***Hamilton*** [3] ***cell problem***. The Hamiltonian cell problem makes sense only in the case when the cell comprises inhomogeneities with known boundary conditions on their boundaries. In the case of photonic crystals, Maxwell's boundary conditions are satisfied at the surfaces of inclusions and the quasi-periodicity condition relates the values of the wave function at the cell sides. In the general case, the Hamiltonian cell problem is solved using the so-called ***Hamilton operator***, often simply called the ***Hamiltonian***. We refer the readers to for the proof that for lossless cells the solutions are discrete sets of real-valued eigenfrequencies ω corresponding to a given Bloch wavevector $\mathbf{q}$ ($-\pi/a_{x,y,z} < q_{x,y,z} < \pi/a_{x,y,z}$).

We see that Bloch's theorem is the possibility to find all eigenmodes of the photonic crystal considering one isolated unit cell of the lattice and varying $\mathbf{q}$. The second implication is a possibility to restrict the wavevector axis in the dispersion diagram by the fundamental wavevector $\mathbf{q}$. This allowed L. Brillouin to replace the dispersion plot $\omega(\mathbf{q})$, where $\mathbf{q}$ is the wavevector with an arbitrary absolute value ($-\infty < q_{x,y,z} < \infty$) inherent to the unbounded space by the Brillouin's dispersion diagram. The idea by Brillouin is illustrated by Figure 3.8 for a one-dimensional crystal. In this figure, the Brillouin diagram is shown for a photonic crystal formed by two dielectric layers of equal thicknesses $a/2$ and relative permittivities $\varepsilon_1 = 6$ and $\varepsilon_2 = 7$.

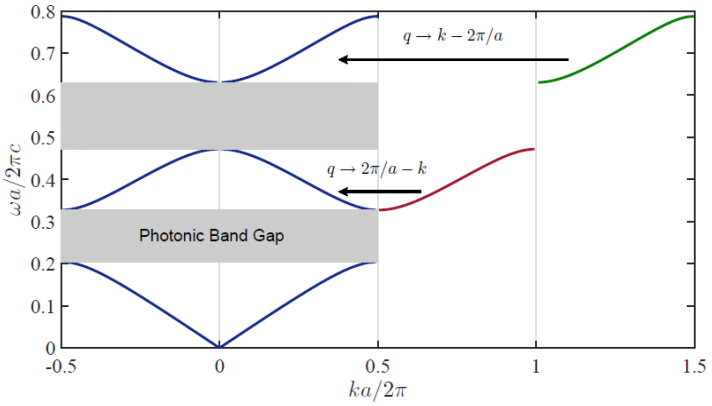

Fig.3. 8 Brillouin's band diagram of an infinite stack of dielectric bilayers oriented orthogonally to axis *z*. Layers of equal thicknesses have relative permittivities 6 and 7, respectively. In the first passband the wavenumber is equal to q. In the second passband q is replaced by $q_{new} = 2\pi/a - q$. In the third passband $q_{new} = q - 2\pi/a$.

Let us explain the physical meaning of this procedure. For a uniform medium with permittivity $\varepsilon = (\varepsilon_1 + \varepsilon_2)/2 = 6.5$, the dispersion equation is as simple as $\omega = v|k|$, where $v = c/n = c/\sqrt{\varepsilon}$ is the phase velocity. The dispersion curve of this homogeneous medium is shown in Figure 3.8 for positive k as a straight dashed line. Next, if we solve the Hamiltonian cell problem for this one-dimensional photonic crystal, we obtain the dispersion curve that in the band of normalized frequencies $\omega a/2\pi c = [0.0{-}0.8]$ comprises three passbands separated by two bandgaps. In the passbands, in the intervals $\omega a/2\pi c = [0.0{-}0.2]$, $[0.32{-}0.48]$, $[0.63{-}0.79]$, the eigenmode is quite similar to the wave in the homogeneous medium with the averaged permittivity ε, because the contrast of the two permittivity is not high. The dispersion plot $\omega(q)$ in the second passband is the thick dark curve in Figure 3.8, and that in the third passband is the lighter curve. This dispersion is also not very different from the dispersion of the homogeneous medium. However, the fundamental wavenumber q of the Bloch plane-wave expansion is restricted by π/a. Therefore, for the second passband, we replace $q \to 2\pi/a - q$. Due to this shift, the dispersion branch inverts, being transposed to the Brillouin zone. In the Brillouin diagram presented in Figure 3.8, the curves for all even-numbered passband, seemingly appear to correspond to backward waves because here, $\partial\omega/\partial q < 0$. However, we should remember that these branches have been inverted when transposed to the Brillouin zone. Thus, in this example, all waves in the full dispersion diagram are forward waves with co-directed phase and group velocities. On the other hand, if one measures the field phase at discrete points separated by the period, the

wave will indeed appear as a backward wave. This property is exploited in some devices, such as backward-wave amplifiers and generators.

Notice that the dispersion diagram for one-dimensional photonic crystals with reciprocal constituents is symmetric: identical eigenmodes can propagate in opposite directions. In Figure 3.8, we have shown the part of the diagram where $-\pi/a < q < 0$, but usually the part with $q < 0$ for such photonic crystals is omitted. It is worth noting that in a standard Brillouin diagram, the wavenumber and frequency axes are normalized, as we have done in Figure 3.8. To conclude this discussion, let us compare the optical properties of photonic crystals in their passbands with optical properties of natural anisotropic crystals (see Chapter 2). The differences follow from the fact that in natural crystals at optical frequencies, the period is small as compared to the wavelength while, in photonic crystals, these two quantities are comparable. If the incident light wave is polarized so that it does not feel the anisotropy of a natural crystal, it excites only an ordinary wave in the crystal. For such incidence, the interface of a crystal behaves as that of an isotropic dielectric. For general excitation, both ordinary and extraordinary waves are excited. It results in double refraction, also called the ray birefringence. In a photonic crystal, analogous response is allowed only for the first passband where the absolute majority of photonic crystals behave as effectively continuous media. Indeed, at low frequencies, the wavelength of incident light is large compared to the lattice period and the quickly varying higher-order terms of the Bloch expansion ray birefringence (with $n \neq 0$) are evanescent waves which exist near the inhomogeneities of the internal structure. The propagation direction inside the crystal is determined by the fundamental spatial harmonics of the Bloch expansion – the Bloch waves of two eigenpolarizations. For two-dimensional crystals – most often, TE- and TM-polarized – waves propagate in the crystal bulk. Thus, in the first passband, a photonic crystal refracts incident waves like a natural crystal does. If the internal geometry is isotropic, two polarizations degenerate, and there is one refracted wave. If is it anisotropic, birefringence of refracted waves is observed. These properties mean that the photonic crystal behaves similarly to an effectively homogeneous natural crystal.

In the other extreme of high frequencies of external excitation, when $2\pi n/a < k_0$ with $n \_ 0$ (k_0 is the wavenumber in the medium from where the incident wave comes), one or more higher-order harmonics can correspond to propagating waves in external space and also in the crystal bulk. Several propagating reflected waves can be created as in a diffraction grating. These

higher-order modes also propagate inside the photonic crystal body, corresponding to several refracted waves. Thus, at high frequencies, the photonic crystal either completely reflects the incident wave (in stopbands) or splits it into a number of transmitted waves (in passbands). This splitting is not surprising for an optician. A photonic crystal slab is the same as amultilayer diffraction grating. For a finite-thickness photonic crystal, the wavevectors of dominating spatial harmonics in the Bloch expansion are wavevectors of Fraunhofer's diffraction lobes in the theory of diffraction gratings.

3.2.4. Brillouin Diagrams of Two-Dimensional Photonic Crystals

A reciprocal lattice with the coordinate axes (q_x, q_y) that are the components of the wavevector **q** can be introduced for a two-dimensional lattice with periods (a_x, a_y). As usual, only the Brillouin zone $-\pi/a_{x,y} < q_{x,y} < \pi/a_{x,y}$ is considered. The Brillouin zone of a rectangular lattice ($a_x \equiv a$, $a_y \equiv b$) is illustrated by Figure 3.9. Since the Brillouin diagram is a plot $\omega(\mathbf{q})$, where **q** is a vector, the horizontal axis splits onto intervals in which different components of the Bloch vector vary separately. In order to make this diagram more compact, L. Brillouin has suggested the following algorithm.

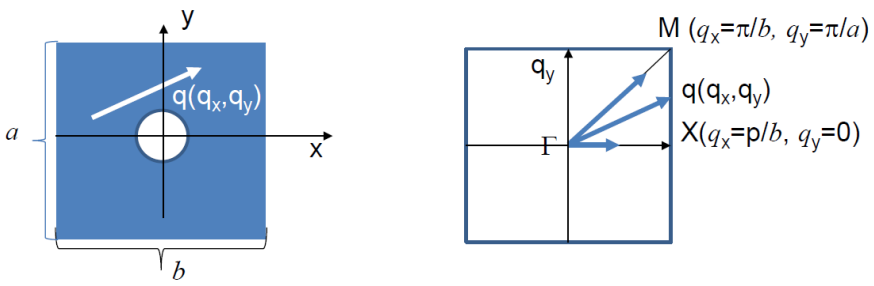

Fig.3. 9 Geometry of a unit cell of a rectangular lattice of holes with periods a and b where the eigenmode with the Bloch vector **q** propagates obliquely (left panel) and the Brillouin zone of this lattice (right panel). The Brillouin dispersion diagram is obtained by the variation of **q** from the characteristic point $\Gamma(q_x = 0, q_y = 0)$ to point X $(q_x = \pi/a, q_y = 0)$, then to point $M(q_x = \pi/a, q_y = \pi/b)$ and again to Γ.

First, for reciprocal lattices with a unit cell symmetric along axes x and y, the negative part of the Brillouin zone $-\pi/a < q_{x,y} < 0$ is removed, as containing no new information; only one quadrant of the Brillouin zone is relevant. Second, there is no need to find all ω for all possible values of the components of **q** within this quadrant. It is enough to vary this vector over a

triangle formed by the characteristic points of the Brillouin zone: points Γ, X, and M. The variation of the vector **q** on the Brillouin diagram starts from and ends at point Γ: the origin of the reciprocal space ($q_x = q_y = 0$). In the first part of the band diagram, the wavevector **q** is directed along *x*, and $q = q_x$ varies from zero to π/a (point X). In the second part, the wavevector **q** keeps the same *x*-component $q_x = \pi/a$, and its *y*-component varies from zero to π/b (point M). Finally, keeping the same diagonal direction of the wavevector one decreases its absolute value from $\sqrt{(\pi/a_x)^2 + (\pi/a_y)^2}$ to zero, returning in this way to Γ.

Wavevectors varying over any other contour onto which the Brillouin zone may split will not reveal new eigenfrequencies. The sufficiency of the triangle Γ-X-M results from the so-called ***group symmetry*** of the rectangular cell problem. In the case of more complicated lattices, such as honeycomb lattices, the Brillouine zone and the corresponding path of **q** become more complicated. We will discuss this issue in the following paragraphs.

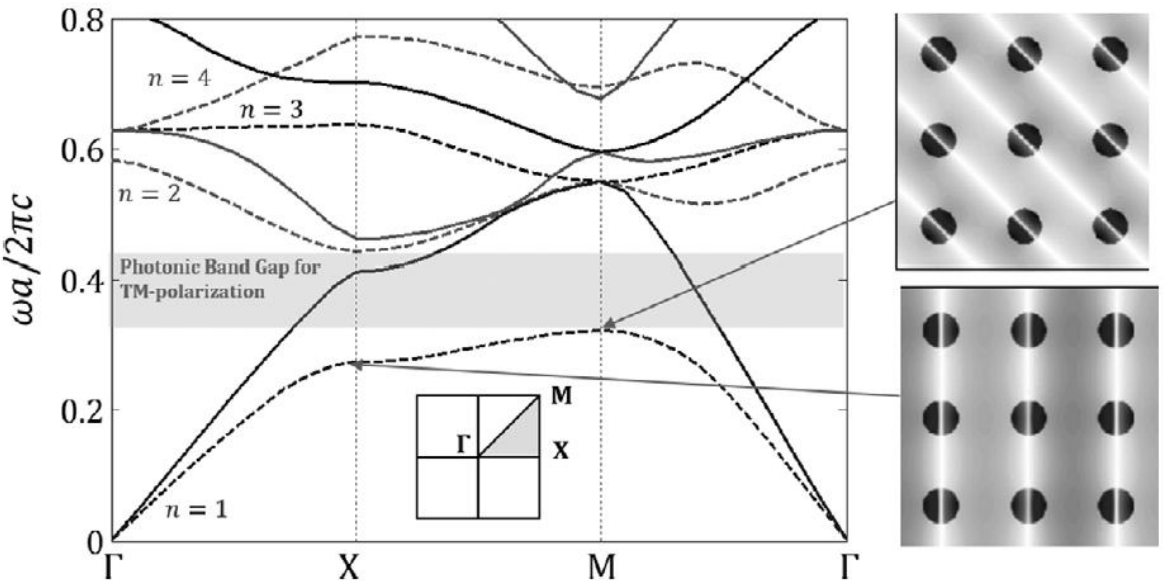

Fig.3. 10 The Brillouin diagram of a square lattice of cylinders with the period *a*, the relative permittivity ε = 10, and the diameter *a*/2 hosted by free space. Dispersion curves of the TE-polarized eigenmodes are solid; those of the TM-polarized ones are dashed. The photonic bandgap holds only for the TM-polarized modes. On the inset, the electric field phase maps are shown for two Bragg modes, corresponding to points X and M on the boundary of the Brillouin zone. Simulations are done by Ana Dیaz-Rubio.

The Brillouin diagram for a square ($a = b$) lattice of dielectric cylinders of diameter $a/2$ with permittivity $\varepsilon = 10$ located in free space is shown in Figure 3.10. From this diagram, it is clear why the photonic bandgap term is preferable for two-dimensional photonic crystals rather than the complete bandgap. Really, the stopband for all propagation directions (in the plane orthogonal to the cylinders axes) is observed only for waves whose magnetic field is polarized along the cylinders: for TM-polarized eigenwaves. For TEpolarized waves whose dispersion curves are shown as solid lines, only irectional stopbands are present. For example, for propagation along x, there is a gap $\omega a/2\pi c = 0.41 \ldots 0.44$ between the first and second dispersion curves for TEwaves. For the bisectorial propagation there is a gap $\omega a/2\pi c = 0.56 \ldots 0.59$, etc. On the insets of Figure 5.10 the color maps of simulated electric field are shown for two eigenmodes: that with $q_x = \pi/a$, $q_y = 0$ (corresponding to point X) and that with $q_x = q_y = \pi/a$ (point M). At these boundary points of the Brillouin zone any pair of inclusions neighboring along the propagation direction is polarized in the opposite phase. Strictly speaking, such propagatingwaves cannot exist because for them, the Bragg condition (5.1) is satisfied. Therefore, the modes whose **q** corresponds to the intersections of the dispersion curve with points X and M are called the Bragg modes or (more often) the *staggered modes* of the photonic crystal. It is seen in Figure 3.10 that the Bragg mode has zero group velocity ($\partial\omega/\partial q = 0$), which means absence of energy propagation. These points on the dispersion curve are assumed to be punctured.

In fact, a Bragg mode on the dispersion curve is truly prohibited shallow in a purely lossless crystal. Due to finite (even very small) optical losses, the group velocity in photonic crystals is not exactly equal to $\partial\omega/\partial q$, and Bloch's vector **q** depicted in the horizontal axis of the Brillouin diagram represents the real part of the complex wavevector $\mathbf{k} = \mathbf{q} + j\mathrm{Im}(\mathbf{k})$. This imaginary part is negative and responsible for wave decay. Therefore, in realistic photonic crystals, the Bragg mode exponentially attenuates versus the distance from the source. If it is excited by an incident plane wave at the crystal boundary, it practically decays over a certain path that may be longer than the lattice period. In the present simulations, $\mathrm{Im}(\varepsilon) = -10^{-7}$), and in this case, the decay length of the Bragg mode equals to about a dozen of lattice periods.

Modes that have $\partial\omega/\partial q = 0$ inside the Brillouin zone can exist, too. In the present case, this point is on the dispersion curve of the $2d$ passband between and Γ points. Due to non-zero

losses, this property does not mean that the light is "frozen" (as it is sometimes erroneously thought). This simply means that, at this point, the imaginary part of the wavevector prevails, and the group velocity has no physical meaning. Waves with $\mathrm{Re}(q) \neq 0$ and $|\mathrm{Im}(q)| > \mathrm{Re}(q)$ practically do not propagate into the bulk and rapidly decay. Such modes are sometimes called complex modes or staggered modes. However, most often they are called *polaritons*[10]. In realistic photonic crystals, polaritons (including the Bragg modes) can be excited either by an embedded source or by a planewave. In the last case, they are excited at the interface of the crystal. The effect of polaritons can be used in some practical applications involving relatively thin photonic crystal layers. In a thin structure of a few periods across the layer, polariton fields reach the rear interface before it decays and may even experience multiple reflections on both interfaces[11].

3.2.5. Brillouin Diagrams of Complex Three-Dimensional Photonic Crystals

For two-dimensional photonic crystals whose lattice is different from the rectangular one and for three-dimensional photonic crystals whose lattice is different from the orthorhombic one (i.e., the unit cell is not a parallelepiped) determination of the Brillouin zone is not so simple. The answer for any explicit lattice can be found applying the theory of groups. We do not explain this theory here. Recipes for building the Brillouin zone for various photonic crystals are available in books (e.g. footnote 1).

In Figure 3.11, we show two typical band diagrams for two typical photonic crystals with three-dimensional honeycomb (opal) lattices, also called bodycentered cubic lattices. The difference between Figure 3.11(a) and Figure 3.11(b) is the content of the inclusions and the matrix. For a usual opal, the spherical inclusions have a higher permittivity than that of the matrix. For the inverted opal, the matrix has a higher permittivity than the inclusions. In the inset showing the Brillouin zone, we can see that the number of characteristic points (six) is larger than that for a simple cubic lattice (four).

[10] K. Sakoda, *Optical Properties of Photonic Crystals*, Springer-Verlag, 2005.

[11] S. K. S.Rahman, T. Klein, J. Gutowski, S. Klembt, and K. Sebald, "Tunable Bragg polaritons and nonlinear emission from a hybrid metal-unfolded ZnSe-based microcavity," *Scientific Reports* **7**, 767 2017).

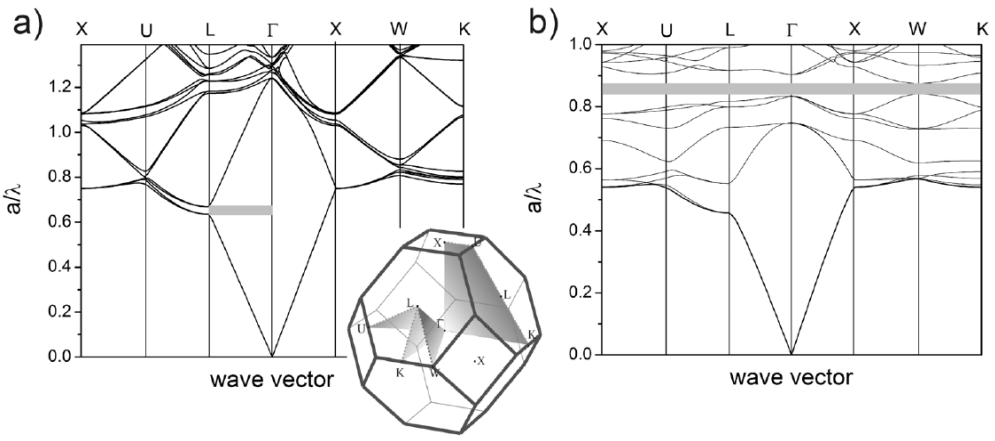

Fig.3. 11 The Brillouin band diagrams for two three-dimensional photonic crystals with complex lattices: (a) that of a three-dimensional honeycomb lattice of spherical inclusions called an artificial opal, (b) a honeycomb lattice of voids (inverted opal). (a) Gray rectangle points out a stopband along the vertical axis of the lattice. This stopband (directional bandgap) shifts upward when the propagation direction deviates from the vertical axis. The inset shows the Brillouin zone of this lattice.

Figure 3.11 is instructive for comparison of directional and complete bandgaps. The appearance of a directional bandgap is seen in Figure 3.11(a). Between points Γ and X, light is propagating along the *x*-axis. Between points X and W, q_x is fixed ($q_x = \pi/a$), and q_y grows from zero – i.e., light is obliquely propagating with an increasing angle in the *xy*-plane. The central frequency of the first directional bandgap (between the first and second passbands) clearly increases versus the propagation angle. For a sea mouse hair, this prohibited frequencies correspond to red light at point X and to blue light at point W. Similarly, the directional bandgap effect holds when the propagation angle changes in other planes, between other characteristic points of the diagram. For the inverse opal whose inclusions are voids, one observes in Figure 3.11(b) a complete bandgap.

A wider complete bandgap than that of opals is observed for artificial diamond lattices. Diamond-type photonic crystal with voids (called yablonovite since the technique of its fabrication was suggested by the group of E. Yablonovich) was theoretically and experimentally studied in works, which manifested the start of the era of nanophotonics in modern optics and electrodynamics.

3.3. Numerical Simulations of Photonic Crystals & Their Scalability

The theory of the cell problem formulated via the Hamiltonian operator is instructive and allows one to understand the main features of the lattice dispersion. However, in practical cases, one needs to calculate the parameters of eigenmodes of the photonic crystal numerically. For this purpose, one uses the socalled governing equation that provides information on both Bloch wavevector **q** and relations between the amplitudes of the spatial harmonics in the Bloch plane-wave expansion of the eigenmode. Let us derive this equation.

3.3.1. Numerical Solution via Plane-Wave Expansion

If a photonic crystal is built from isotropic nonmagnetic materials (dielectrics or metals), the photonic crystal properties are determined by a scalar periodically varying permittivity $\varepsilon(\mathbf{r})$. From Maxwell's equations

$$\nabla \times \mathbf{H}(\mathbf{r}) + j\omega\varepsilon(\mathbf{r})\varepsilon_0\mathbf{E}(\mathbf{r}) = 0,$$
$$\nabla \times \mathbf{E}(\mathbf{r}) - j\omega\mu_0\mathbf{H}(\mathbf{r}) = 0, \qquad \text{Eq. 3. 6}$$

we obtain the governing equation of a photonic crystal:

$$\nabla\times[\epsilon^{-1}(\mathbf{r})\nabla\times\mathbf{H}(\mathbf{r})] - k_0^2\mathbf{H}(\mathbf{r}) = 0. \qquad \text{Eq. 3. 7}$$

Let us restrict our consideration by the case of simple (orthorhombic) lattices, where the natural lattice axes are Cartesian ones. In this case, we have $\varepsilon(x, y, z) = \varepsilon(x + a_x, y + a_y, z + a_z)$, and the Bloch plane-wave expansion takes the form,

$$\mathbf{H} = \sum_{mps}\mathbf{H}_{mps}e^{-j(q_x+\frac{2\pi m}{a_x})x - j(q_y+\frac{2\pi p}{a_y})y - j(q_z+\frac{2\pi s}{a_z})z}, \qquad \text{Eq. 3. 8}$$

where $G_x = 2\pi m/a_x$, $G_y = 2\pi p/a_y$, $G_z = 2\pi s/a_z$ are the components of the lattice vector **G** (m, p, and s are integers). The Bloch expansion allows us to replace the operator $\nabla\times$ in the governing equation by the vector product $-j(\mathbf{q} + \mathbf{G})\times$.

Since $\varepsilon(\mathbf{r})$ is a periodic function of coordinates, the inverse permittivity $\varepsilon^{-1}(\mathbf{r})$ is also a periodic function, and it can be expanded into a Fourier series,

$$\epsilon^{-1}(\mathbf{r}) = \sum_{MPS} b_{MPSE} e^{-j(\frac{2\pi M}{a_X}x + \frac{2\pi P}{a_Y}y + \frac{2\pi S}{a_Z}z)}. \quad \text{Eq. 3. 9}$$

The Fourier coefficients of the inverse permittivity b_{MPS} can be found via the standard integration procedure since the function $\varepsilon(x, y, z)$ is known. Substituting (Eq 3.7) and (Eq 3.8) into (Eq 3.6), we obtain an infinite system of algebraic equations for unknown coefficients $\mathbf{H}_{mnp}$ with the unknown vector $\mathbf{q}$ having, in the general case, all three components q_x, q_y, and q_z. This system is source-free; i.e., to allow for nontrivial solutions, its determinantmust be equal to zero. This fact allows us to find $\mathbf{q}$ for a given frequency $\omega \equiv k_0 c$ without solving the cell problem. Assume that the wave propagates along x – i.e., $q_y = q_z = 0$. In this case we obtain the system

$$\sum_{mnp} \xi_{MmPpSs} \mathbf{H}_{mps} e^{-j(q_x + \frac{2\pi m}{a_x})x - j(q_y + \frac{2\pi p}{a_y})y - j(q_z + \frac{2\pi s}{a_z})z} = 0, \quad (M, P, S) = 0, \pm 1, \pm 2, \ldots \quad \text{Eq. 3. 10}$$

where coefficients ξ_{MmPpSs} can be easily expressed through q_x and b_{MPS} using (Eq 3.6) with the substitution of (Eq 3.8) and $\nabla = -j(q_x + 2\pi m/a_x)\mathbf{x}_0 - j2\pi p/a_y \mathbf{y}_0 - j\mathbf{2}\pi s/a_z \mathbf{z}_0$. Then we truncate the system (Eq 3.9), choosing finite $M_{\max}$, $P_{\max}$, and $S_{\max}$ so that the finite-sum approximation of the inverse permittivity initially described by the infinite sum (Eq 3.8) provides required accuracy. Most often, numbers $M_{\max}$, $P_{\max}$, and $S_{\max}$ corresponding to system (Eq 3.9) are smaller than similar numbers that would arise if we were using the alternative governing equation

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) - k_0^2 \epsilon(\mathbf{r}) \mathbf{E}(\mathbf{r}) = 0, \quad \text{Eq. 3. 11}$$

which also follows from Maxwell's equations, and applied the Fourier expansion for $\varepsilon(\mathbf{r})$ [1]. After this truncation, we obtain a linear source-free system of equations (Eq 3.9) for a finite number of unknowns $H_{x,mps}$, $H_{y,mps}$, and $H_{z,mps}$. Equating the determinant of the system to zero,we come to an algebraic equation for q_x that can be numerically solved, for example, using MATLAB.

Similarly, one can find q_y and q_z for the cases when the wave propagates along these axes. Next, one fixes the relationship between the components of the Bloch wavevector $\mathbf{q}$ that corresponds to a selected direction of oblique propagation. The needed propagation directions are determined by the corners of the Brillouin zone. For oblique propagation, we again obtain a

finite source-free linear system with a scalar q whose absolute value is found equating the determinant to zero.

Once we know q for a given propagation direction, we may set the amplitude of the component of $\mathbf{H}_{000}$ orthogonal to this propagation direction equal to unity and find the relative amplitudes of all spatial harmonics. It is what we need because, in general, the eigenmode amplitude is not defined – this is all that we can know about an eigenmode propagating in a given direction at a given frequency.

As we have already discussed, the geometry of majority of photonic crystals allows separate existence of the TE- and TM-polarized modes along all relevant directions. This property simplifies the eigenvalue problem for $\mathbf{H}_{mps}$. In this case, it becomes a scalar problem for two field components. Using Maxwell's equations, the calculation of the other components of the modal electromagnetic field is straightforward.

We see that, using the governing equation and the Bloch theorem, we can find the relative amplitudes of all spatial harmonics of eigenmodes propagating in the infinite lattice at a given frequency in a given direction. The knowledge of q for all propagation directions at a given frequency ω gives us the isofrequency surface or contour, and the same data for a sufficient set of frequencies provides a family of isofrequencies and the dispersion diagram. The amplitude $\mathbf{H}_{000}$ is defined by excitation of waves by external sources (most often, by plane waves illuminating crystal surfaces). Here we do not discuss methods for numerical solutions of boundary problems for photonic crystals; we refer the reader to book[12].

3.3.2. Numerical Solution of the Cell Problem

Numerical simulations of the band diagram and isofrequencies of photonic crystals can be done using commercial software packages such as HFSS, CST Studio, and COMSOL Multiphysics. These packages comprise solvers of cell problems using Bloch's quasi-periodicity conditions applied for a unit cell having six faces and possessing translation symmetry. Such cells are called *monoclinic prisms*.

In the HFSS solver, the option for solving such cell problems is called *Optimetric Analysis*. Choosing this option, the user draws the unit cell of the lattice with its internal geometry in an HFSS project, defines in the descriptive part of the project the material parameters of all the

[12] K. Sakoda, *Optical Properties of Photonic Crystals*, Springer-Verlag, 2005.

internal parts of the cell, and selects the option *Eigenmode solution*. Next, the user suggests varying from zero to π the phase shift of the eigenwave traveling across the unit cell, determining in the solver the so-called *master and slave faces* of the cell.

Master faces are the cell sides "illuminated by the eigenwave." Slave faces are opposite to the master faces – i.e., located at the back side of the monoclinic prism. The solver varies the propagation directions and the absolute value of **q** so that **q** varies from the center (Γ) to the boundary of the Brillouin zone. Then, for the all passible values of **q** (varying with a very small step), the solver delivers the set of eigenfrequencies. If there are optical losses (i.e., the permittivity of the inclusion or the host medium is complex), the eigenfrequencies corresponding to given real-valued **q** are complex numbers.

Only real parts of these complex numbers are used in the dispersion diagram that is calculated within the framework of the optimetric analysis. Of course, photonic crystals are never prepared of very lossy constituents since their practical applications are not absorption and they must be sufficiently transparent in their passbands. Therefore, eigenfrequencies of the propagating waves are in the practical cases sufficiently close to real values; otherwise, either the solution is incorrect or losses are too high.

Similar general procedures correspond to other commercial solvers of the cell problem for photonic crystals. However, the numerical solution can be done in different ways. The HFSS solver uses the finite-element method for Maxwell's equations together with the Bloch quasi-periodicity conditions, whereas the CST Studio Suite involves two techniques: volume integral equations in the frequency domain and the finite-difference-time-domain (FDTD) approach. For reliability, one often simulates the same photonic crystal using two or more numerical techniques. Beyond commercial software, there is homemade software applying different techniques for solving the cell problem. Besides the plane-wave expansion method that we have already discussed, people also use the so-called *transfer-matrix approach* and FDTD approach . In some special cases, it is possible to solve the cell problem for simple photonic crystals analytically in closed form.

3.3.3. Scalability of Photonic Crystals

Here we discuss the scalability of photonic crystals – the invariance of the solution with respect to the magnification of the unit cell accompanied by the corresponding reduction of the operation frequency.

For the dispersion equation, this scalability follows from the fact that both axes in the dispersion diagram are dimensionless: $\omega a/2\pi c$ and $qa/2\pi$. For the electromagnetic field vectors, the scalability is not so evident, and we need to prove it. Denote the radius vector in the enlarged photonic crystal as **R**, and in the original photonic crystal, let it be **r**. Then $\nabla_r = s\nabla_R$, and we may rewrite (Eq 3.6) as

$$s^2\nabla_R \times [\epsilon^{-1}(\mathbf{R})\nabla_R \times \mathbf{H}(\mathbf{R})] = k_0^2\mathbf{H}(\mathbf{R}). \qquad \text{Eq. 3. 12}$$

Dividing both parts of this equation by s^2, we obtain $k_0^{new} = k_0/s$, i.e., $\lambda_{new} = s\lambda$. This means that optical phenomena of nanostructured photonic crystals can be modeled at microwave frequencies where the unit cell is scaled to millimeters or even centimeters. Scalability offers a great tool for experimental study and optimization of photonic crystals. The main restriction of this approach refers to photonic crystals with metal inclusions. Metals in the optical range have complex permittivity whose negative real part prevails over the imaginary part. Such materials at microwaves and other radio frequency ranges are absent, except some kinds of plasma. An inclusion of plasma needs an encapsulation. It is difficult to prepare a matrix with voids filled with plasma, and it is even more difficult to make plasmas identical in all these voids. It would be easier and cheaper to fabricate an original photonic crystal with submicron metal inclusions.

3.4. Conclusion

This chapter explored the fundamental principles and applications of photonic crystals. We began by defining photonic crystals and examining naturally occurring examples. A significant portion of the chapter focused on photonic band structures, explaining the underlying Bragg phenomenon and its role in creating band gaps. We delved into the theoretical framework using Bloch's theorem and visualizing band structures through Brillouin diagrams for both two- and three-dimensional photonic crystals. Finally, the chapter addressed the practical aspects of numerically simulating photonic crystals using methods like plane-wave expansion and the cell problem, along with a discussion of their scalability. The included quiz and control questions allow for self-assessment of the concepts covered. The next chapter will build upon this foundation to introduce the plasmonics phenomena.

Quiz & Control Questions

1. Is the explanation given by Lord Rayleigh to the phenomenon of Bragg valid for lattices of particles of substantial size? Can the reflection by a planar grid of large particles be treated as a reflection by an infinitesimally thin crystal plane? Explain your opinion.
2. Why are Bloch’s quasi-periodicity conditions called quasi-periodic? When they become periodic conditions, what is the limit case in terms of the Bloch wavevector?
3. What is the advantage of using the Brillouin zone in the dispersion diagram of lattices?
4. Let an isotropic medium with the relative permittivity $\varepsilon_1 = 4$ form a flat interface with a uniaxial medium having tangential component of the permittivity tensor $\varepsilon_{2t} = 4$ and the normal component $\varepsilon_{2n} = 1$. Can the effect of total internal reflection be observed for waves obliquely incident from the isotropic medium?
5. Why is the governing equation of the photonic crystal formulated for the magnetic field of an eigenmode and not for the electric field?
6. Why is the scalability of photonic crystals not used by researchers to simplify the experimental investigations of lattices of metal nanoparticles and optical wire media?

Chapter 04

Plasmonics

4.1. Introduction

In order to control and process light at the subwavelength (nano) scale, we need to find ways to confine electromagnetic fields in subwavelength volumes, and we need to find waveguiding structures with subwavelength cross-section dimensions. It is obvious that such conventional approaches to light concentration as dielectric lenses or parabolic mirrors do not offer a possibility of achieving this goal, because their focusing ability is limited by diffraction. One needs other means, where propagating waves would excite oscillations that vary in space much faster than propagating waves do. In radio and microwave engineering, such devices are abundant. For example, an electrically small but resonant wire dipole antenna being excited by a propagating plane wave (whose field variations are on the scale of the wavelength λ_0) collects the incident wave power from the area of the order of λ_0^2 and delivers it into a tiny gap between the two antenna arms (which is very small compared to λ_0) where the load is connected.

One of the possible approaches toward solving these problems in optics is the use of noble metals. Electromagnetic waves cannot propagate inside bulk pieces of conductors, but resonant modes at interfaces between metals (negative permittivity) and dielectrics (positive permittivity) can have rather localized distributions of electromagnetic fields in space. One can think about an analogy with the radio-frequency circuits, where a resonator formed by a capacitor filled with a dielectric (negative reactance, positive permittivity, electric stored energy due to accumulated charge) and an inductor made of a conducting wire (positive reactance, "negative permittivity," magnetic stored energy due to conduction current) has dimensions very small compared to the wavelength.

Indeed, radio-frequency antennas made of metals can effectively confine electromagnetic fields, metal coaxial cables can guide waves, and LC-resonators can be very small, so can we use metal elements to do the same in the visible and infrared spectra? This question is addressed by plasmonics.

4.2. Surface Plasmon Polaritons

Surface plasmon polaritons are surfacewaves that can propagate along interfaces between dielectrics and metals. They are basically inhomogeneous plane waves whose "propagation constant" in the direction orthogonal to the interface is imaginary so that the fields exponentially

decay away from the interface. If this decay factor is high, and the fields concentration at the surface is strong enough, such a surface-wave structure may possibly serve as a waveguide with a subwavelength cross-section size. If the propagation constant along the surface is large enough (much larger than the free-space constant), we also may be able to realize subwavelength field confinement along the surface.

4.2.1. Dispersion Equation

To understand the properties of these surface waves, we will need to introduce some preliminary notions regarding plane waves near planar interfaces. Here, we mean waves whose complex amplitudes depend on the position vector r as $exp(-jk \cdot r)$, where r can be a complex vector. In this meaning, evanescent waves are also plane waves, and their other name is *inhomogeneous plane waves*. Let us consider plane electromagnetic waves in an isotropic medium in the vicinity of a planar interface with a different material or with a planar boundary. In this situation, it is convenient to introduce a unit vector normal to the interface n and consider tangential to the surface (orthogonal to n) components of the planewave fields, because these components are continuous across the interface.

This situation is illustrated in Figure 4.1. The wavevector k is split into its normal and tangential components: $k = k_n n + k_t$, where $k_t \cdot n = 0$. Denoting the length of the wavevector $k = \omega\sqrt{\varepsilon\mu}$, we write the dispersion equation for plane waves as

$$k_n^2 + k_t^2 = k^2. \qquad \text{Eq 4. 1}$$

Thus,

$$k_n = \sqrt{k^2 - k_t^2}. \qquad \text{Eq 4. 2}$$

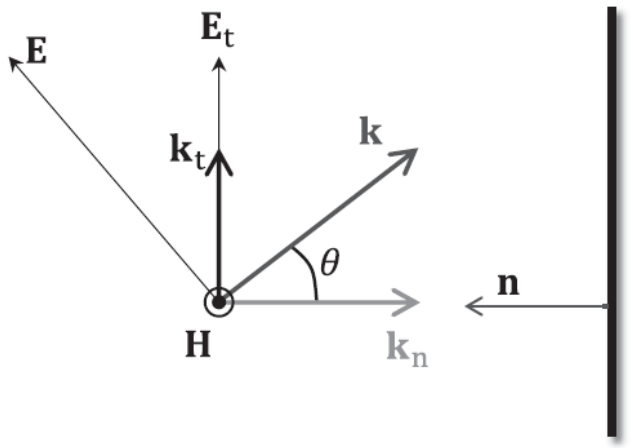

Fig. 4. 1 A transverse-magnetic (TM) polarized plane wave in a vicinity of a planar surface, shown by a vertical line.

Because both tangential field components are continuous across the interface, it is convenient to introduce wave impedance as the ratio of the *tangential to the interface* components of the plane-wave fields. For this example of transversemagnetic (TM) polarized waves, we have

$$\mathbf{E}_t = \sqrt{\frac{\mu}{\epsilon}}\cos\theta\,\mathbf{n}\times\mathbf{H} = \sqrt{\frac{\mu}{\epsilon}}\frac{k_n}{k}\mathbf{n}\times\mathbf{H} = \frac{k_n}{\omega\epsilon}\mathbf{n}\times\mathbf{H}. \qquad \text{Eq 4. 3}$$

Here, θ is the angle between vector ***k*** and the normal to the interface (the "incidence angle"). Now we are ready to find the dispersion equation for surfacewaves propagating along the interface and decaying in the normal direction. Equations (4.2) and (4.3) hold for plane-wave fields on both sides of the interface; we only need to substitute the corresponding material parameters of the two media and satisfy the boundary conditions, which tell us that the tangential fields must be continuous across the interface. Thus, we have

$$\frac{k_{n1}}{\omega\epsilon_1} = -\frac{k_{n2}}{\omega\epsilon_2} \qquad \text{Eq 4. 4}$$

(here we have taken into account that the unit vector *n* is directed oppositely on the other side of the interface)

Let us assume that the medium on one side is a lossless metal with the permittivity $\varepsilon_1 < 0$, and the other half-space is filled with a lossless isotropic dielectric with the permittivity $\varepsilon_2 > 0$. Since we want to find solutions that would exponentially decay away from the interface, the normal component of the wavevector of plane waves on both sides must be imaginary; that is, we are looking for solutions with $|k_t| > |k_{1,2}|$ (see (Eq 4.2)).

Because waves should decay in both the positive (medium 1) and the negative (medium 2) directions of the unit vector *n*, the signs $k_{n1,2}$ should be opposite to each other. Thus, this equation can have nontrivial solutions for surface waves only if the signs of ε_1 and ε_2 are opposite, which is exactly the case of our interface between lossless dielectric (positive ε) and metal (negative ε). Substituting the values of $k_{n1,2}$ from (Eq 4.2) and solving for the tangential to the interface component of the propagation constant k_t, we find the dispersion equation

$$k_t = \omega\sqrt{\mu_0}\sqrt{\frac{\epsilon_1\epsilon_2}{\epsilon_1+\epsilon_2}} = \omega\sqrt{\epsilon_0\mu_0}\sqrt{\frac{\epsilon_{r1}\epsilon_{r2}}{\epsilon_{r1}+\epsilon_{r2}}}. \qquad \text{Eq 4. 5}$$

The eigenvalue equation for the orthogonal (TE) polarization can be obtained upon replacing permittivities by the corresponding permeabilities in (Eq 4.4). For interfaces between metals and dielectrics, it has no solutions because both permeabilities are positive.

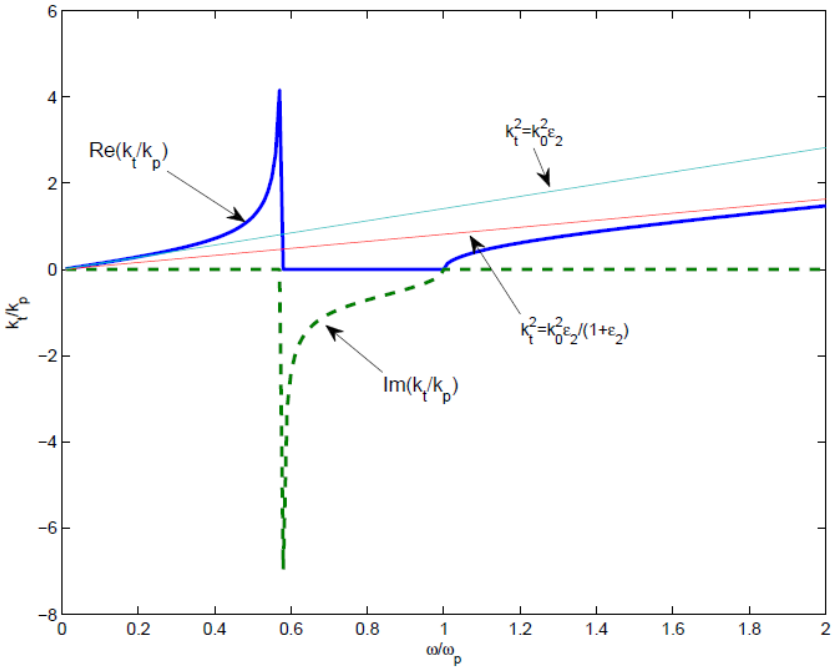

Fig. 4. 2 Example of dispersion curves for surface plasmon polaritons for the case of negligible absorption. The wavenumbers are normalized to the plasma wavenumber $k_p = \omega_p\sqrt{\varepsilon_0\mu_0}$. Plot parameters: $\varepsilon_{r2} = 2$, the Drude model for lossless conductors.

An example of dispersion curves for surface waves at interfaces between dielectrics and metals (neglecting losses in both media) is shown in Figure 4.2. Dispersion of the permittivity of the dielectric is neglected, and the dispersion of metal is modeled by the Drude model (Eq 2.18) with the loss factor $v = 0$. The resonant frequency (sometimes called the *surface plasmon frequency*)

$$\omega_{SPP} = \omega_p \sqrt{\frac{\epsilon_0}{\epsilon_0 + \epsilon_2}} \qquad \text{Eq 4. 6}$$

corresponds to the point where the denominator of (Eq 4.5) tends to zero, so the propagation constant k_t tends to infinity. In the frequency range $\omega_{SPP} < \omega < \omega_p$, there is a stop band (no solutions with real values of k_t), and at $\omega > \omega_p$, the waves are not bounded to the surface because the condition $|k_t| > |k_{1,2}|$ is not satisfied.

For lossy metals and/or dielectrics, the propagation constant is complex at all frequencies, so there is some decay in the passbands and some propagation in the stop band; see a typical plot in Figure 4.3.

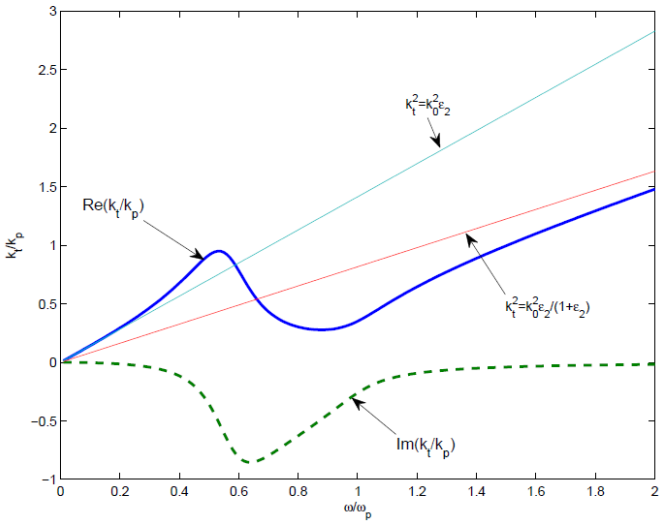

Fig. 4. 3 Example of dispersion curves for surface plasmon polaritons: absorptive metal. Plot parameters: $\varepsilon_{r2} = 2$, Drude model for lossy conductors, $v/\omega p = 0.2$.

For applications in the visible and infrared ranges, silver is the usual metal of choice because of its relatively small dissipation losses as compared to other metals. Calculations based on the measured data for silver permittivity give the curves shown in Figure 4.4. The dielectric medium is the same as in the previous examples: a lossless dielectric $\varepsilon_{r2} = 2$.We see that the

surface plasmon polariton can propagate in the range from about 200 THz ($\lambda_0 = 1.5$ μm) to about 750 THz ($\lambda_0 = 0.4$ μm = 400 nm), covering the visible (about 390–700 nm) and infrared

spectra. However, the properties of the wave are quite dissimilar at different frequencies. We note from the same plot that the decay rate, measured by the imaginary part of the propagation constant, grows with increasing frequency. It can be visualized better by plotting the so-called *propagation length*, which is equal to the inverse of the imaginary part of the propagation constant: $L = 1/\mathrm{Im}(k_t)$. This curve is shown in Figure 4.5. It is clear that at the highfrequency end of the visible spectrum, the propagation length is only a couple of microns, while in the infrared range, it can be two orders of magnitude larger than the wavelength.

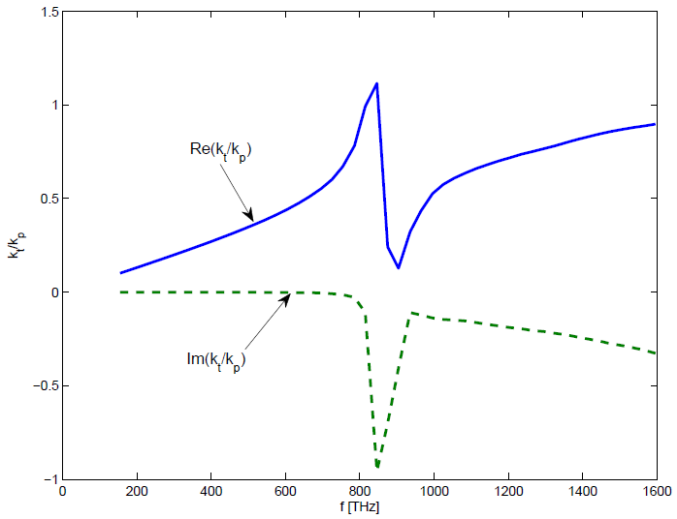

Fig. 4. 4 Dispersion curves for surface plasmon polaritons on an interface between silver and a dielectric ($\varepsilon_{r2} = 2$). Experimental data for silver properties. The propagation constant is normalized to the fitting value of the plasma wavenumber, corresponding to the plasma frequency ωp/(2π)=2175THz.

Another important parameter is the confinement of the wave fields to the surface. After all, one of the main motivations for using plasmonic structures is a possibility of confining the field in subwavelength volumes. In this case, we look for a waveguiding structure whose cross-section size would be small compared to the wavelength. The field confinement is measured by the field decay rate in The directions normal to the surface – that is, by the inverse values of the normal component of the propagation constant (its imaginary part). The corresponding plots are shown in Figure 4.6. We see that while the fields decay rather quickly into the silver volume, fields extend quite far into the dielectric, especially at low frequencies. Actually, the “width”

of the plasmonic waveguide is comparable to the free-space wavelength at the corresponding frequency.

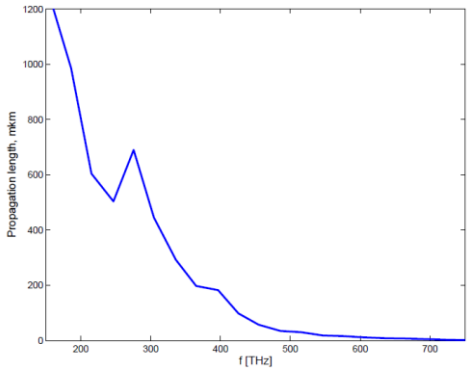

Fig. 4. 5 Propagation length of surface plasmon polaritons on an interface between silver and a dielectric ($\varepsilon_2 = 2$). Experimental data for silver properties[13].

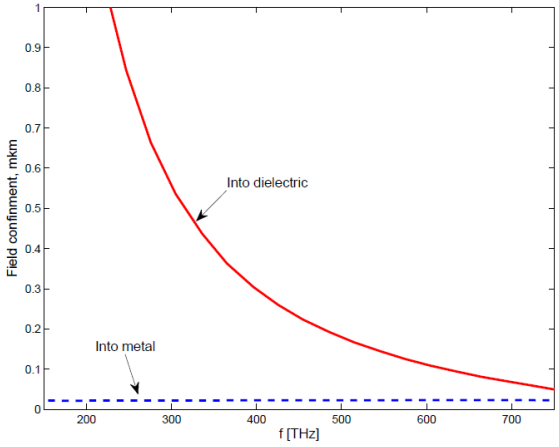

Fig. 4. 6 Confinement of fields of surface plasmon polaritons on an interface between silver and a dielectric ($\varepsilon_{r2} = 2$). Experimental data for silver properties[1].

4.2.2. Slab Waveguides

Propagating surface plasmon polaritons have small propagation length due to losses in metal. In order to reduce losses, it is preferable to use waveguide topologies where it is possible to concentrate most of the fields in low-loss dielectrics or in the air. One possibility is to use the slab topology. Let us consider a three-layer structure, where a dielectric slab (permittivity ε_2) is bounded from both sides by semi-infinite metal spaces (permittivity ε_1). We would like to find a mode whose field maximum would be located inside the dielectric so that the fields at

[13] P. B. Johnson and R. W. Christy, "Optical constants of noble metals," *Physical Review B* **6**, 4370–4379 (1972).

the interfaces with metal will be smaller than in the waveguide center. In case of plasmons on a single interface, which we considered in Section 4.2.1, the tangential fields at the metal surface are, in fact, at their maximum value (they exponentially decay in both directions away from the interface).

To get more physical insight, this layered waveguide structure can be viewed as a system of two coupled plasmons at the two metal-dielectric interfaces or as a metal parallel-plate waveguide with unideally conducting walls or as a planar dielectric waveguide with lossy cladding. The field solution can be found using the standard technique for analyzing open (in particular, planar) waveguides. Considering TM modes, the longitudinal component of the electric field in the two media $E_{z1,2}$ satisfies the Helmholtz equation

$$\nabla_t^2 E_{z1,2} + (k_{1,2}^2 - k_t^2) E_{z1,2} = 0. \qquad \text{Eq 4. 7}$$

The boundary conditions on the two interfaces read

$$\frac{\epsilon_1}{k_t^2 - k_1^2}\frac{\partial E_{z1}}{\partial n} = \frac{\epsilon_2}{k_t^2 - k_2^2}\frac{\partial E_{z2}}{\partial n}, \qquad \text{Eq 4. 8}$$

where ***n*** is the outward normal to the dielectric slab. These equations as well as the resulting dispersion equation are the same as for conventional dielectric-slab waveguides:

$$\tanh(k_{n2}d) = -\frac{k_{n1}\epsilon_2}{k_{n2}\epsilon_1} \quad \text{odd modes} \qquad \text{Eq 4. 9}$$

$$\tanh(k_{n2}d) = -\frac{k_{n2}\epsilon_1}{k_{n1}\epsilon_2} \quad \text{even modes,} \qquad \text{Eq 4. 10}$$

where d is the thickness of the dielectric slab. The solutions need to be found numerically, considering that the metal permittivity strongly depends on the frequency. The results show that, indeed, among even modes, there is a mode with a cosine-type distribution of the electric field in the dielectric slab such that the fields at the interfaces with metal are smaller than in the core center. Since, in this case, most of the power propagates in dielectric and not in metal, losses can be considerably smaller than in single-interface plasmonic waveguides. On the other hand, it is important to note that these relatively low-loss modes exist in the case when the thickness of the dielectric slab is comparable to or larger than the wavelength in the dielectric. Thus, these waveguides do not offer a compact (subwavelength cross-section) solution. Compared to conventional planar optical waveguides, the advantages come from the fact that the waveguide is closed by metal walls, isolating the guided fields from the environment. Other

modifications of plasmonic waveguides are metal strips, grooves, and hollow channels in bulk metal.

4.2.3. Surface Waves at Low Frequencies

As is obvious from Figure 4.2, if the operational frequency is much smaller than the plasma frequency of metal ($\omega << \omega_p$), the surface mode exists, but its propagation constant k_t is very close to that of plane waves in dielectric. The corresponding normal component of the propagation constant k_{n2} is close to zero. This conclusion is also seen directly from the dispersion equation (4.4): when the permittivity of metal ε_1 takes large negative values, the right-hand side is very small. Physically, this means that the wave is very weakly bound to the surface; in the limit of infinite conductivity of metal, it becomes simply a homogeneous plane wave propagating in the dielectric medium along the conducting surface. However, if the conductivity of the first medium is moderate, well-bounded surface waves can exist.

Considering again the equations for the normal to the interface components of the propagation constants in the two media $k_{n1,2}$, we can notice that the propagation constants are complex numbers if one of the contacting media is a lossy dielectric (complex permittivity). In this case, even if the real parts of both permittivities are positive, the field can be also bound to the surface and propagate along it. The propagation constant along the interface is also complex in this case, so the wave exponentially decays. A classic example is the radiofrequency surface wave over the Earth's surface. Average, soil or sea water are moderately conducting media. Surface waves can indeed propagate, and they are called Zenneck waves (1907). A comparative overview of the Zenneck wave and surface plasmons can be found in paper[14].

Surface waves can also propagate over metal surfaces at low frequencies if we structure the surface at the subwavelength scale. To understand these waves, let us revisit the derivation of the dispersion equation. Considering Figure 4.1, let us assume that to the right of the interface we have a dielectric material with the permittivity ε. The normal component of the propagation constant k_n should be a negative imaginary number so that the surface-wave fields would exponentially decay away from the interface. Thus, the corresponding wave impedance (Eq 4.3) is a negative imaginary quantity, a capacitive reactance. In order to satisfy the dispersion equation (Eq 4.4), we should bring it in contact with a surface that has a positive, inductive

[14] A. Sihvola, J. Qi, and I. V. Lindell, "Bridging the gap between plasmonics and Zenneck waves," *IEEE Antennas and Propagation Magazine* **52**, 124– 136, (2010).

surface reactance. If we fill the left half-space with a metal and the frequency is well below the plasma frequency, that is impossible because the metal impedance is very close to zero. However, imagine that we corrugate a metal surface with electrically thin (and densely packed) grooves of depth *d*, orienting the grooves orthogonally to the tangential component of the electric field vector. At the bottom of the groove, the surface impedance is nearly zero (short circuit by the metal wall). But at the distance *d* from the groove bottom, we see the impedance of a short-circuited transmission line of the length *d*, which is proportional to *j tan(kd)*. Here, *k* is the propagation constant of the material that fills the grooves (air, for example). Since the tangent function can take any value from $-\infty$ to $+\infty$, we can obviously choose the depth *d* so that the impedance at the surface of the corrugated metal plate takes the desired positive imaginary value.

Such structured surfaces (there are other possible realizations in addition to corrugations) are called *high-impedance surfaces*. In microwave engineering, the use of structured surfaces to engineer dispersion of surface waves has been known for a long time. More recently, this possibility was also recognized by experts in optics, and the new name, "spoof surface plasmons," was proposed to describe surface waves over subwavelength structured surfaces.

Another possibility to adjust the surface impedance of a metal surface is to cover it with a dielectric slab. Using the simple transmission-line model, it can be seen that the nearly zero (short-circuit) impedance of metals at microwave frequencies is transformed into reactive (if dissipation is small) impedance at the surface of the dielectric slab. For some propagation constant of surface waves, it becomes possible to satisfy the surface-mode dispersion equation. Actually, the presence of surface modes on grounded dielectric layers creates problems in the design of printed-circuit components and microstrip antennas because these waves carry the electromagnetic energy away from the device components, increasing losses and creating undesired side lobes of microstrip antennas.

4.2.4. Excitation of Surface Plasmon Polaritons

Surface waves between two media characterized by wavenumbers $k_{1,2}$ are bound to the surface, meaning that the propagation constant along the interface k_t is larger than the wavenumbers in both media: $|k_t| > |k_{1,2}|$. It means that it is not possible to excite such waves by propagating plane waves incident on the interface: the exiting wave should be in phase synchronism with the surface wave, but the tangential component of the propagating wave is always *smaller* than the

length of the wavevector $|k_{1,2}|$. Thus, we need a coupling device that would transfer propagating plane waves fields into fields whose variations in space would be fast enough.

One possibility is to use a prism in the full-internal-reflection regime. When a plane wave propagating inside the prism material is fully reflected from one of the prism faces, the field outside the prism is not zero (due to continuity of the tangential field components), but it decays exponentially away from the prism (because all the incident energy is reflected). Obviously, this exponentially decaying field is, in fact, a surface wave: the tangential component of the propagation constant is larger than the length of the wavevector in the medium outside the prism. Thus, if we bring such a prism close to an interface that supports surface plasmon polaritons, phase synchronism with the eigenwave of the interface mode becomes possible.

Another possibility is to use a periodical structure (grating). Let us consider again a dielectric/metal interface that supports surface plasmons. If we periodically perturb the dielectric surface – for example, cutting small parallel groves – the surface mode between the dielectric and metal will be not just a single inhomogeneous plane wave but an infinite series of Floquet harmonics with the wavenumbers

$$k_m = k_t + n\frac{2\pi}{D}, \qquad \text{Eq 4. 11}$$

where $n = \pm 1, 2, \ldots$, and D is the perturbation period. Properly selecting the period, it is possible to ensure that one of these harmonics is in phase sync with the incident plane wave; that is, k_{tn} is equal to the tangential component of the wavevector of the incident propagating plane wave.

Furthermore, coupling can be provided by positioning one or more small particles in the vicinity of the interface supporting surface plasmon polaritons. Either the particle should be considerably smaller than the wavelength in the surrounding space or it should have shape or material inhomogeneities on the subwavelength scale. In the near vicinity of subwavelength inhomogeneities, the spatial variations of fields are fast, and coupling to plasmon wave whose wavelength is smaller than that in propagating waves becomes possible. One can also understand it from the point of view of Fourier integral expansion of fields in the vicinity of a small inhomogeneity. In order to resolve the fine spatial structure of the fields, the Fourier spectrum must contain plane-wave components with large wavenumbers, which can come to sync with the surface mode. Typical examples of such couplers are metal or dielectric

nanoparticles or tips of near-field scanning microscopes. Any inhomogeneity of the surface acts as a coupler between propagating waves and surface modes. Yet another example is the excitation of surface waves at surfaces of a finite size by illuminating the edge of the surface by propagating waves.

4.3. Plasmonic Nanoparticles

4.3.1. Polarizability

Let us consider an optically small (with respect to the wavelength in the surrounding space) metal sphere. Since the sphere is small, it is possible to model its response as that of an electric dipole, excited by external electric fields, which can be assumed to be approximately uniform over the particle volume. We denote the complex amplitude of the incident electric field at the position of the particle as $\mathbf{E}_{\text{inc}}$ and the complex amplitude of the incident-wave Poynting vector as $\mathbf{P}_{\text{inc}}$. We model the linear particle response by its polarizability α defined through

$$\mathbf{p} = \alpha \mathbf{E}_{\text{inc}}, \qquad \text{Eq 4. 12}$$

where **p** is the complex amplitude of the induced electric dipole moment. The power extracted by a dipole inclusion (dipole moment **p**) from a given incident field $\mathbf{E}_{\text{inc}}$ can be found from the classical formula

$$Extraitdecode P_{\text{ext}} = \frac{1}{2}\text{Re}\int_V \mathbf{J}^* \cdot \mathbf{E}_{\text{inc}} dV = \frac{1}{2}\text{Re}(-j\omega \mathbf{p}^* \cdot \mathbf{E}_{\text{inc}})$$
$$= -\frac{\omega}{2}\text{Im}(\alpha)\,|\,\mathbf{E}_{\text{inc}}\,|^2 = -\eta\omega \qquad \text{Eq 4. 13}$$

($\eta = \sqrt{\mu/\varepsilon}$ is the wave impedance of the surrounding space and, as usual, the harmonic time dependence is in the form $\exp(j\omega t)$). Here, **J** is the volumetric electric current density inside the particle. This general formula is valid for any dispersive and lossy particle, assuming only that it is a small dipole particle and $\mathbf{E}_{\text{inc}}$ is uniform over its volume. The result in terms of the incident power density P_{inc} holds if the particle is excited by a propagating plane wave. The corresponding value of the incident time-averaged power flow density is

$$P_{\text{inc}} = \frac{1}{2\eta}\,|\,E_{\text{inc}}\,|^2 . \qquad \text{Eq 4. 14}$$

The power that is scattered (reradiated) by the particle is the power radiated by the electric dipole **p**:

$$P_{sc} = \frac{\mu_0 \omega^4}{12\pi c} |\, p\,|^2 = \frac{\mu_0 \omega^4}{12\pi c} |\, \alpha\,|^2 |\, E_{inc}\,|^2 = \frac{k^4}{6\pi \epsilon^2} |\, \alpha\,|^2 \; P_{inc}, \qquad \text{Eq 4. 15}$$

where $k = \omega\sqrt{\mu\varepsilon}$ is the wavenumbers in the surrounding space. We can find the general expression for the scattering loss (radiation damping) factor in the dipole polarizability by equating the extracted and scattered powers for the case of a lossless particle (no absorption). This allows us to find an expression for the imaginary part of the inverse polarizability due to scattering loss

$$\frac{1}{\alpha} = \frac{\mathrm{Re}(\alpha) - j\mathrm{Im}(\alpha)}{|\,\alpha\,|^2} \qquad \text{Eq 4. 16}$$

$$\mathrm{Im}\left(\frac{1}{\alpha}\right) = -\frac{\mathrm{Im}(\alpha)}{|\,\alpha\,|^2} = \frac{\mu\omega^3}{6\pi c} = \frac{k^3}{6\pi\epsilon}. \qquad \text{Eq 4. 17}$$

This is a classical result that dates back to the work of M. Planck (1902). We see that, in the general case, the inverse polarizability of a dipole particle has the form

$$\frac{1}{\alpha} = \xi + j\frac{k^3}{6\pi\epsilon}, \qquad \text{Eq 4. 18}$$

where the complex-valued parameter $\xi = \xi' + j\xi''$ depends on the particle size, shape, and material. The imaginary part ξ'' measures absorption in the particle, if it is lossy.

For electrically small particles, parameter ξ can be found using the quasi-static approximation, since the scattering properties have been already considered by the radiation damping term. For a small sphere made of an isotropic material with the permittivity _1, the quasi-static solution gives

$$\xi = \frac{1}{3\epsilon V}\frac{\epsilon_1 + 2\epsilon}{\epsilon_1 - \epsilon}, \qquad \text{Eq 4. 19}$$

where V is the volume of the sphere. This approximation of uniform internal field is adequate for Ag and Au nanospheres of radii smaller than 30 and 50 nm, respectively. For a special case of a metal sphere modeled by the Drude permittivity

$$\epsilon_1 = \epsilon_0\left[1-\frac{\omega_p^2}{\omega(\omega - j\nu)}\right], \qquad \text{Eq 4. 20}$$

we have

$$\xi' = \frac{\omega_p^2/3-\omega^2}{\epsilon V\omega_p^2}, \quad \xi'' = \frac{\omega\nu}{\epsilon V\omega_p^2}. \qquad \text{Eq 4. 21}$$

4.3.2. Localized Surface Plasmons

At the frequency where the real part of the permittivity of the sphere satisfies

$$\varepsilon_1' = -2\varepsilon' \qquad \text{Eq 4. 22}$$

(ε' is the real part of the host permittivity ε), the absolute value of parameter ξ (Eq 4.19) has a minimum, meaning that the amplitude of the particle polarizability (Eq 4.18) has a maximum: there is a resonance of the particle response. As seen from (Eq 4.21), for metal nanoparticles modeled by the Drude formula, the resonance frequency is $\omega_0 = \omega_p/3$. Condition (Eq 4.22) is called the Frôhlich resonance condition. This resonant effect is called the plasmonic resonance, and the excited resonant mode is called the *localized surface plasmon.* The last name can be perhaps confusing, since the distribution of the polarization current (proportional to that of the internal electric field) is practically uniform, meaning that the internal field is not localized at the surface of the sphere. On the other hand, the fields in space around the particle are, in a certain sense, localized: they behave as the fields of a small electric dipole. In contrast to surface modes on planar interfaces, the fields do not decay exponentially away from the particle: reactive near fields decay as 1/distance^2 and 1/distance^3, while the radiated wave decays as 1/distance (assuming lossless background material).

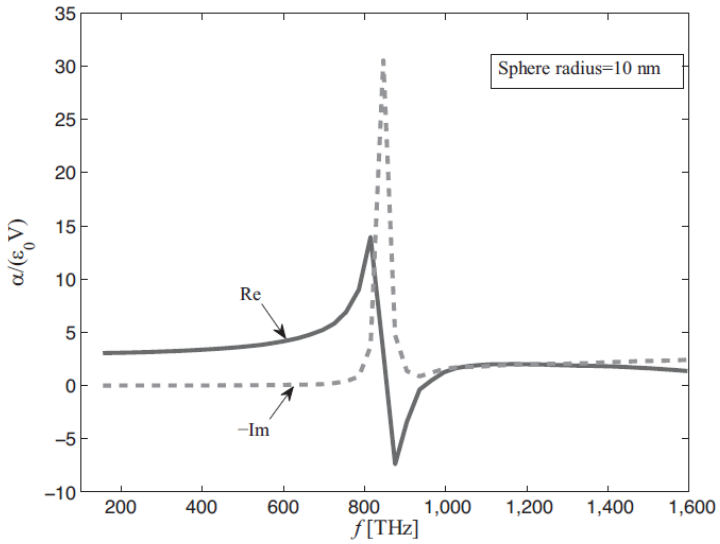

Fig. 4. 7 Normalized polarizabilities of spherical silver nanoparticles in a dielectric host $\varepsilon = 2\,\varepsilon_0$ for the sphere radius 10 nm. The radiation damping is negligible.

Figures 4.7–4.9 show the normalized polarizability $\alpha/(\varepsilon V)$ as a function of the frequency for silver nanoparticles in a dielectric host with the permittivity $\varepsilon = 2\,\varepsilon_0$. For moderate permittivities of the surrounding medium, the resonance is in the visible range, and the resonance frequency can be obviously tuned by varying the host permittivity. The dotted curves show the calculations based on the quasi-static model, neglecting the third-order scattering loss term in (Eq 4.18). In this approximation, the polarizability is proportional to the particle volume, and for the normalized polarizability, these curves are the same. We can clearly see that for very small spheres, the scattering loss is negligible as compared to the absorption in the particle (in the plot for the sphere of radius 10 nm, the curves cannot be distinguished). On the other hand, for larger spheres, the scattered power is comparable to or even dominates over absorption, as is clear from the plot for the radius equal to 50 nm.

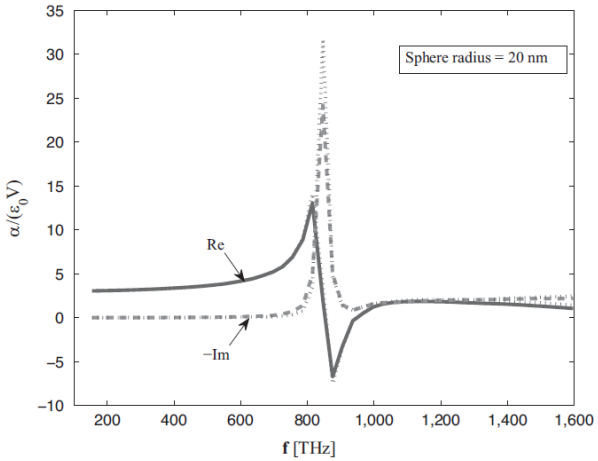

Fig. 4. 8 Normalized polarizabilities of spherical silver nanoparticles in a dielectric host $\varepsilon = 2\ \varepsilon_0$ for the sphere radius 20 nm. Dotted curves show the polarizability calculated neglecting the radiation damping term.

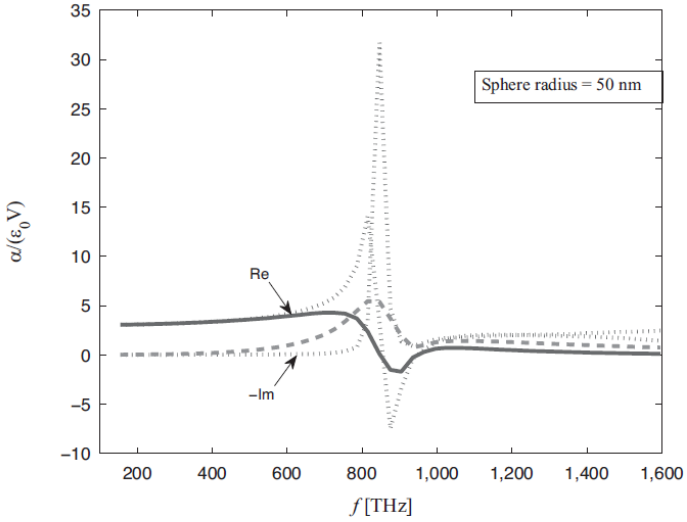

Fig. 4. 9 Normalized polarizabilities of spherical silver nanoparticles in a dielectric host $\varepsilon = 2\ \varepsilon_0$ for the sphere radius 50 nm. Dotted curves show the polarizability calculated neglecting the radiation damping term.

4.4. Chains of Plasmonic Nanoparticles

Surface plasmon polaritons can propagate along metal-dielectric interfaces, and to some extent, such structures can be used as waveguides with subwavelength cross sections. However, as we saw in Section 4.2, the field confinement is ensured only at the high-frequency end of the visible spectrum, and the propagation loss at these frequencies is rather high. Although various

plasmon-polariton waveguide topologies, such as strip guides and grove waveguides, have been proposed and studied, considerable loss and poor field localization in the surrounding dielectric limit possible applications.

There is an alternative possibility to confine and guide light in a subwavelength waveguide: using chains of resonant nanoparticles – in particular, plasmonic nanoparticles. Each nanoparticle is an optically small resonator, which can confine light in a subwavelength vicinity of the particle. Is it possible to adjust the particle interactions in a chain of plasmonic nanoparticles so that the excitation will propagate from particle to particle, as if it were a waveguide? This possibility was probably described for the first time in paper [8], where calculations for a finite-length chain were shown.

Let us consider a periodical arrangement of small resonant particles along a certain line in free space or in an isotropic dielectric (along the axis z, as shown in Figure 4.10). For simplicity, we consider particles polarizable along the axis

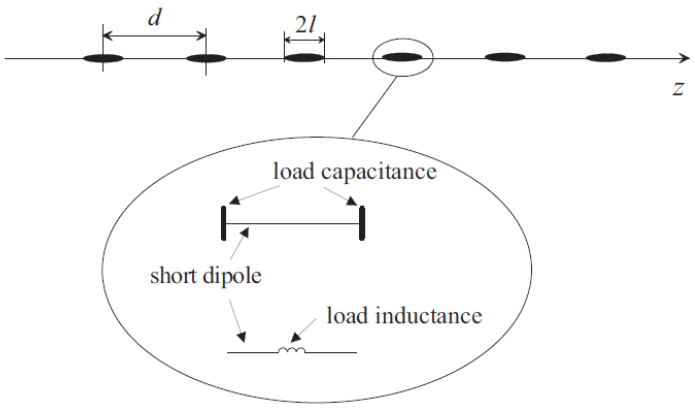

Fig. 4. 10 Geometry of a generic waveguiding structure formed by a periodical chain of small resonant electric-dipole particles.

direction so that all induced electric dipoles are oriented along this direction. For an arbitrarily chosen dipole on the line array (position $z = 0$), the electric dipole reads

$$p(0) = \alpha E_{\text{loc}}, \qquad \text{Eq 4. 23}$$

where E_{loc} is the *z*-component of the *local* field created by external sources and all the other particles. Assuming that there is no external field and looking for eigenwaves along the linear chain, we can make use of the Floquet theorem and write

$$p(nd) = e^{-jqnd} p(0), \qquad \text{Eq 4. 24}$$

where q is the propagation factor of waves traveling along the chain and d is the period. The local field is created by the other dipoles in the array:

$$E_{\text{loc}} = \sum_{n=-\infty, n\neq 0}^{\infty} \frac{1}{2\pi\epsilon_0}\left(\frac{1}{(|n|d)^3} + \frac{jk}{(nd)^2}\right) e^{-jk|n|d} e^{-jqnd} p(0), \qquad \text{Eq 4. 25}$$

where we have substituted the electric dipole fields (we use the usual notation $k = \omega\sqrt{\varepsilon_0\mu_0}$).

Let us denote by β the *interaction constant*

$$\beta = \sum_{n=-\infty, n\neq 0}^{\infty} \frac{1}{2\pi\epsilon_0}\left(\frac{1}{(|n|d)^3} + \frac{jk}{(nd)^2}\right) e^{-jk|n|d} e^{-jqnd}. \qquad \text{Eq 4. 26}$$

With this notation, at the position of the reference dipole *p*(0), the local field is

$$E_{\text{loc}} = \beta p(0). \qquad \text{Eq 4. 27}$$

If there is no incident field, then $p(0) = \alpha E_{loc} = \alpha\beta\, p(0)$. Thus, we have the eigenvalue equation $\alpha\beta = 1$, or

$$\frac{1}{\alpha} = \sum_{n=-\infty, n\neq 0}^{\infty} \frac{1}{2\pi\epsilon_0}\left[\frac{1}{(|n|d)^3} + \frac{jk}{(nd)^2}\right] e^{-jk|n|d} e^{-jqnd}. \qquad \text{Eq 4. 28}$$

For lossless scatterers, the imaginary part of the inverse polarizability is known (Eq 4.17), and we get

$$\text{Re}\left\{\frac{1}{\alpha}\right\} + j\frac{k^3}{6\pi\epsilon_0} = \sum_{n=-\infty, n\neq 0}^{\infty} \frac{1}{2\pi\epsilon_0}\left[\frac{1}{(|n|d)^3} + \frac{jk}{(nd)^2}\right] e^{-jk|n|d} e^{-jqnd} \qquad \text{Eq 4. 29}$$

The imaginary part of this series can be calculated analytically, and the result reads

$$\mathrm{Im}\{\beta\} = \frac{1}{\pi\epsilon_0 d^3}\begin{cases} \frac{(kd)^3}{6} + \frac{\pi}{4}[(qd)^2 - (kd)^2], & 0 < qd < kd \\ \frac{(kd)^3}{6}, & kd < qd < 2\pi - kd \\ \frac{(kd)^3}{6} + \frac{\pi}{4}[(2\pi - qd)^2 - (kd)^2], & 2\pi - kd < qd < 2\pi + kd \end{cases} \qquad \text{Eq 4. 30}$$

(this function is 2π-periodic with respect to *qd*).

We see that if $q > k$ (more exactly, if $kd < qd < 2\pi - kd$), the imaginary parts of this equation cancel out, which means that propagating modes are possible if we tune the real part of the polarizability to be equal to the real part of the interaction constant β. Obviously, since $q > k$, the fields of the guided mode exponentially decay away from the axis *z* as $e^{-jk_\rho \rho}$ with the decay factor $k_\rho = k^2 - q^2$. Numerically calculated examples with field distributions can be found, e.g., in[15].

4.5. Applications

Resonant phenomena in optically small metal nanoparticles can be used in a broad variety of applications, mainly exploiting possibilities to concentrate and resonantly enhance electromagnetic fields in optically small regions. Also, the ability to guide light along plasmonic structures is of practical interest.

Let us list the main application areas:

- Optical circuits (metatronics); see Chapter 6;
- Enhanced emission, fluorescence, and nonlinearities; see Chapter 6;
- Enhanced solar cells;
- Metamaterials, including superlenses; see Chapter 5
- Nanolasers
- Applications in biology and medicine (mainly in sensors);

The main practical problem: high dissipation losses in resonant metal nanoparticles

[15] S. A. Maier, P. G. Kik, and H. A. Atwater, "Optical pulse propagation in metal nanoparticle chain waveguides," *Physical Review B* **67**, 205402 (2003).

or in resonant surface modes.

4.6. Conclusion

This chapter investigated the properties and applications of surface plasmon polaritons (SPPs) and plasmonic nanoparticles. We started by deriving the dispersion equation for SPPs and explored their behavior in slab waveguides, focusing on low-frequency surface waves and excitation methods. The chapter then shifted to plasmonic nanoparticles, examining their polarizability and the phenomenon of localized surface plasmons. The collective behavior of plasmonic nanoparticles arranged in chains was also analyzed. Finally, the chapter concluded with a survey of the diverse applications of these concepts, setting the stage for further exploration in subsequent chapters. The next chapter will likely explore the fundamentals of metamaterials and their connection to all the previous mechanisms.

Quiz & Control Questions

1. Surface plasmon polaritons can exist and propagate along flat metal-dielectric interfaces. What other structures can support surface plasmon polaritons? Discuss the necessary conditions for existence of surface modes.
2. Surface plasmon polaritons on infinite interfaces cannot be excited by propagating plane waves, but localized surface plasmons can be excited this way. Explain this difference.
3. Plot the dispersion curve for surface modes at a planar interface of a lossless metal (Drude dispersion, plasma frequency ω_p) with a lossless isotropic medium having Lorentz dispersion. Assume that the resonance frequency of the dielectric response is $\omega_0 = 1.5\omega_p$. Study the properties of waves as functions of other parameters of the Lorentz material. Discuss the conditions for existence of surface plasmon polaritons.
4. The absorption cross section of a small particle is defined as the ratio of the absorbed power P_{abs} to the incident power density (the amplitude of the Poynting vector) P_{inc}:

$$\sigma_{abs} = \frac{P_{abs}}{P_{inc}}. \qquad \text{Eq 4. 31}$$

This value can be viewed as an effective area from which the particle “collects” energy, which is then dissipated into heat. Similarly, the total scattering cross section is defined as the ratio of

the power scattered (reradiated) by the particle P_{sc} into space when it is illuminated by a plane wave with the incident power density P_{inc}:

$$\sigma_{sc} = \frac{P_{sc}}{P_{inc}}. \qquad \text{Eq 4. 32}$$

Finally, the extinction cross section is the sum $\sigma_{ext} = \sigma_{sc} + \sigma_{abs}$.

Consider an arbitrary linear electric dipole scatterer – for example, a plasmonic nanoparticle – and find if the values of these three cross sections have a fundamental upper bound. If yes, determine these bounds in general and, in particular, for metal nanoparticles modeled by the Drude formula. Discuss the physical meaning and implications of your results.

Hints

The inverse polarizability of any dipole scatterer can be written as

$$\frac{1}{\alpha} = \xi' + j\xi'' + j\frac{k^3}{6\pi\epsilon_0}. \qquad \text{Eq 4. 33}$$

The imaginary part ξ'' measures absorption *loss*, so it must be non-negative for any scatterer: $\xi'' \geq 0$. The real part ξ' can take any value – positive, negative, or zero.

Using the known formulas for electric dipoles, one can express the cross sections in terms of the polarizability. The power extracted from the field is

$$P_{\text{ext}} = -\frac{\omega}{2}\text{Im}(\alpha)|E_{\text{inc}}|^2 = -\eta_0\omega\text{Im}(\alpha)P_{\text{inc}}. \qquad \text{Eq 4. 34}$$

The scattered power reads

$$P_{sc} = \frac{\mu_0\omega^4|\alpha|^2}{12\pi c}|E_{inc}|^2 = \frac{k^4}{6\pi\epsilon_0^2}|\alpha|^2 P_{inc}. \qquad \text{Eq 4. 35}$$

The absorbed power is the difference between the power extracted by the particle from the incident field (Eq 4.34) and the power scattered by the same particle into the surrounding space (Eq 4.35):

$$P_{abs} = -\frac{\omega}{2}\mathrm{Im}(\alpha)\,|E_{inc}|^2 - \frac{\mu_0\omega^4\,|\alpha|^2}{12\pi c}\,|E_{inc}|^2 .$$ Eq 4. 36

Next, we express all the cross sections in terms of the polarizability (Eq 4.33):

$$\sigma_{abs} = \frac{P_{abs}}{P_{inc}} = \frac{k}{\epsilon_0}\frac{\xi''}{\xi'^2 + \left(\xi'' + \frac{k^3}{6\pi\epsilon_0}\right)^2},$$ Eq 4. 37

$$\sigma_{sc} = \frac{P_{sc}}{P_{inc}} = \frac{k^4}{6\pi\epsilon_0^2}\frac{1}{\xi'^2 + \left(\xi'' + \frac{k^3}{6\pi\epsilon_0}\right)^2},$$ Eq 4. 38

$$\sigma_{ext} = \frac{P_{ext}}{P_{inc}} = \frac{k}{\epsilon_0}\frac{\xi'' + \frac{k^3}{6\pi\epsilon_0}}{\xi'^2 + \left(\xi'' + \frac{k^3}{6\pi\epsilon_0}\right)^2}.$$ Eq 4. 39

Now you can study the bounds, remembering that $\xi'' \geq 0$. Answers can be found in [16].

[16] S. Tretyakov, "Maximizing absorption and scattering by dipole particles," *Plasmonics* **9**, 935–944 (2014).

Chapter 05

Metamaterials

5.1. Metamaterials Concept

Electromagnetic properties of natural or chemically synthesized materials are determined mainly by their chemical composition – that is, by properties of assembled atoms and molecules. Within the metamaterial paradigm, the role of atoms or molecules is played by "meta-atoms" as constitutive elements of artificial materials. Meta-atoms are macroscopic objects made from usual materials (metals, dielectrics, etc.) As long as the meta-atom sizes and distances between them remain sufficiently small on the wavelength scale of interest, the composites formed by meta-atoms can be considered as "materials," described by effective material parameters (such as permittivity or permeability), and the same approaches to the macroscopic description of electromagnetic properties of matter in terms of material relations (Eq 2.1) can be applied to metamaterials made of meta-atoms, just like to ordinary materials formed by atoms or molecules.

The electromagnetic response of arrangements of meta-atoms can be engineered by properly choosing their shapes, internal structures, sizes, mutual orientations, etc. Moreover, the response of individual meta-atoms can be controlled and tuned by external signals or internal, programmable microprocessors, opening up possibilities to engineer tunable, time-modulated, sensing, adaptive, and software-defined artificial materials and surfaces.

The main appeal of the metamaterial concept derives from new possibilities to realize artificial materials with electromagnetic properties that are not available in any natural material and to engineer properties that are optimal for particular applications.

Metamaterial can be defined as artificial, effectively homogeneous material formed by "an arrangement of artificial structural elements, designed to achieve advantageous and unusual electromagnetic properties" (this definition is adopted by the Virtual Institute for Artificial Electromagnetic Materials and Metamaterials; see www.metamorphose-vi.org). Optically thin layers (effectively sheets) formed by engineered meta-atoms are usually called *metasurfaces*, sometimes defined as two-dimensional versions of metamaterials. Arrangements of meta-atoms along a line in space can be called *metawires*. The generality of the metamaterial concept is illustrated in Figure 5.1: The meta-atoms can have any shape and chemical composition, exhibit engineered nonlinear or active properties or even contain controllable and programmable

microcontrollers (software-defined materials). The arrangements of meta-atoms in metamaterials

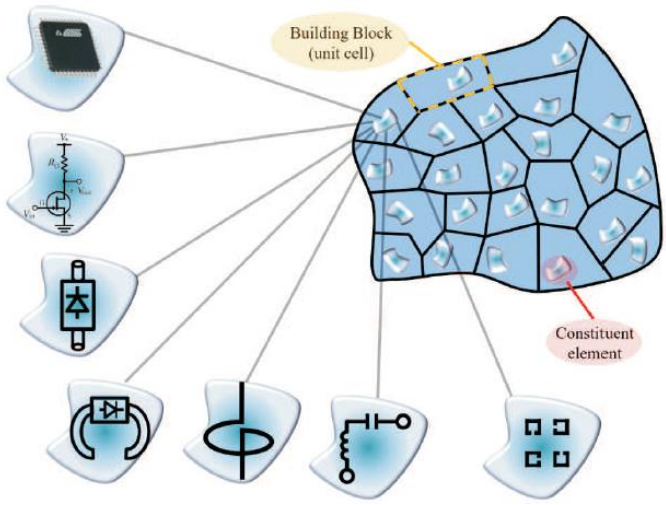

Fig. 5. 1 Illustration of the metamaterial concept. Courtesy of M. Lapine and M. Albooyeh.

and metasurfaces can be regular or random. The only limitations are on the sizes of the meta-atoms and on the (average) distances between them: they should be sufficiently small with respect to the wavelength of interest, to allow the homogenized, material description. Of course, we assume that meta-atoms are macroscopic objects, which are sufficiently large compared to the sizes of the natural atoms and molecules composing them. It is an important restriction for nanostructured metamaterials, whose meta-atoms can be as small as tens of nanometers. Particles smaller than a few nanometers are quantum rather than macroscopic objects.

If meta-atoms are arranged in a lattice whose period is comparable to the wavelength, the lattice response strongly depends on interference of waves reflected by unit cells. Dispersion curves and surfaces show stop bands in the frequency regions where the reflected waves interfere destructively. Such structures are called *photonic crystals* or *electromagnetic bandgap structures (EBG);* see Chapter 3. However, there is no sharp boundary between the notions of the metamaterial and photonic crystal. For example, in designing metasurfaces, it is sometimes required to engineer a dense distribution of small meta-atoms within a unit cell of a periodical

structure whose period is comparable to the wavelength. One can perhaps define such structures as photonic crystals made of metamaterials.

For natural materials, the wavelength is very large as compared to the distances between atoms and molecules even for the visible light and ultra-violet waves. However, working with metamaterials, it is not always easy or even possible to create desired structures of small enough dimensions (small compared to the wavelength), which is an important limitation. Fabrication of optical metamaterials always means nanofabrication.

Although the name *metamaterial* was coined only in the year 2000, when the notion of artificial materials with negative refractive index was introduced, the idea of creating artificial electromagnetic materials from macroscopic elements made from usual materials is quite old, dating back to the nineteenth century.

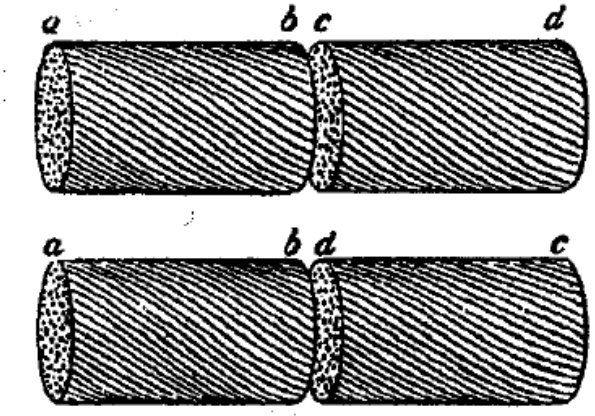

Fig. 5. 2 Chiral bundles of jute by Bose.

One of the earliest interesting examples of artificial electromagnetic materials is related to the understanding and studies of optical activity and circular dichroism in chiral media. Optical activity is observed in materials whose microstructure lacks mirror symmetry. In natural materials, the effects of chirality are very weak because the molecules are very small as compared to the wavelength of light, and effects are not resonant. Moreover, in the nineteenth century, there were no means to experimentally study the relations between the geometrical symmetry of atoms and molecules and the optical properties of materials. However, after the works of H. Hertz, it became possible to work with radio and microwave frequency waves, and it was realized that it became possible to create artificial materials whose microwave properties could emulate the optical properties of natural materials. J. Bose did very interesting pioneer

experimental studies of artificial chiral molecules and artificial chiral materials made of twisted jute bundles; see the illustrations in Figure 5.2. In the beginning of the twentieth century, K. Lindman, in Finland, experimented with artificial chiral materials whose meta-atoms were made from metal springs embedded in cotton balls. More information on the history of metamaterials research can be found, e.g., in [8]. Next, we discuss artificial materials with negative permittivity and permeability, the first metamaterial actively studied in the twenty-first century

5.2. Double-Negative Materials

Double-negative materials are, by definition, isotropic materials obeying the constitutive relations

$$\mathbf{D} = \epsilon \mathbf{E}, \qquad \mathbf{B} = \mu \mathbf{H}, \qquad \text{Eq 5. 1}$$

where the material parameters ε and μ are real and negative. Since all materials are lossy, it is more accurate to say that both $\mathrm{Re}\{\varepsilon\} < 0$ and $\mathrm{Re}\{\mu\} < 0$. These materials have many alternative names in addition to *double-negative (DNG) media*: materials with negative parameters, backward-wave media, materials with negative refraction index (NRI), left-handed materials, and Veselago media. The plurality of names perhaps reflects broad interest and interesting physical properties. The first experimental realization of double-negative materials (at microwaves, for a single polarization and in-plane propagation directions) was built using split rings and metal wires; see Figure 5.3.

Fig. 5. 3 First DNG/Veselago material.

As can be seen from Maxwell's equations for plane waves in form $e^{-j\mathbf{k}\cdot\mathbf{r}}$

$$\mathbf{k}\times\mathbf{E}=\omega\mu\mathbf{H},\qquad \mathbf{k}\times\mathbf{H}=-\omega\epsilon\mathbf{E}, \tag{Eq 5. 2}$$

planewaves in double-negative materials are *backward waves*; that is, the direction of the Poynting vector **S** is opposite to the wavevector **k**:

$$\mathbf{S}=\frac{1}{2}\mathrm{Re}(\mathbf{E}\times\mathbf{H}^{*})=\frac{|E|^{2}}{2\omega\mu}\mathbf{k}=\frac{|H|^{2}}{2\omega\epsilon}\mathbf{k}. \tag{Eq 5. 3}$$

This concept is illustrated in Figure 5.4.

Plane waves at an interface between vacuum or a usual (double-positive) isotropic medium and an isotropic double-negative half-space exhibit negative refraction, see Figure 5.5. This phenomenon can be understood from the continuity of the tangential component of the wavevector at the interface (which is necessary to make sure that the tangential field components are continuous across the interface). In the upper half-space filled with vacuum, both the Poynting vector and the wavevector are directed from the source (at infinity) to the interface. In the bottom half-space, the Poynting vector must be also directed down from the source. But the wavevector is now directed oppositely to the Poynting vector; thus, the only possible configuration that ensures continuity of the tangential component of the wavevector is the configuration of negative refraction, as shown in Figure 5.5.

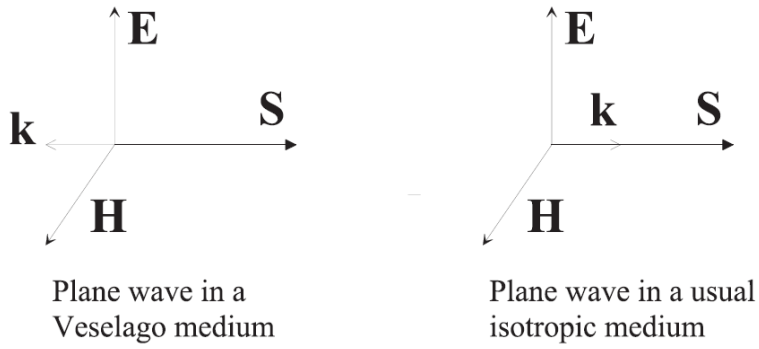

Fig. 5. 4 Forward and backward waves.

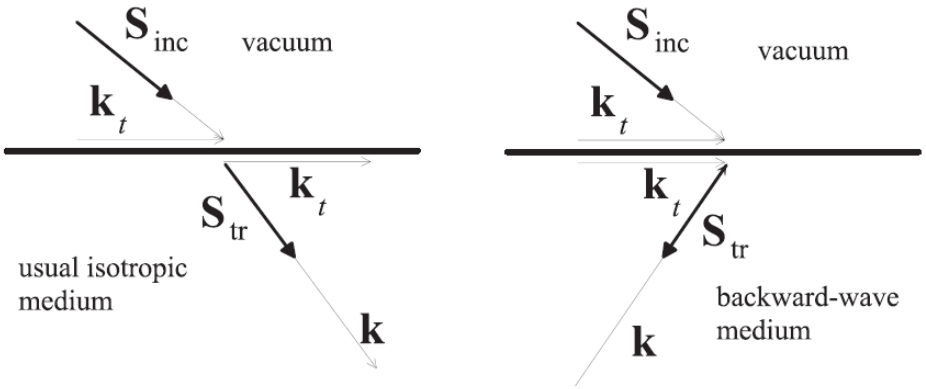

Fig. 5. 5 Negative refraction of plane waves.

While this consideration appears to be trivial, the phenomenon of negative refraction is not so simple, and the first publications on experiments on negative refraction caused an intensive discussion. To understand this effect better, it is necessary to consider the propagation of pulses, restricted both in time and space. The result is illustrated in Figure 5.6, where is it shown that the beam gets distorted. Although the Poynting vector indeed exhibits negative refraction, the phase front is bent in the usual way.

Next, we will discuss surface plasmon modes at interfaces with double-negative materials. Here, we can make use of the theory of surface plasmon polaritons of Chapter 4. Let us recall the notions of propagating and evanescent waves. From (Eq 4.2), it is obvious that if $|k_t| \leq |k|$ (here, we consider, for simplicity, waves in media with negligible losses, where k is a real number, positive or negative), k_n is a real number, and waves propagate without amplitude decay. However, if $|k_t| > |k|$, k_n is an imaginary number, and waves exponentially decay in the direction of **n**, away from the surface (we consider a single wave in homogeneous

space). Such waves are called *evanescent* waves. In problems of image transfer, we can envisage that a certain distribution of electric field, defined in a plane orthogonal to ***n***, is expanded into the spatial Fourier integral, and $\mathbf{k}_t$ is the two-dimensional Fourier variable. Each of these components will propagate as a plane wave in space, carrying information about the image. However, only propagating components with small enough $|k_t| \le |k|$ will propagate far away from the image plane. All evanescent waves will decay at distances considerably larger than the wavelength. This conclusion means that information about tiny details of the object will be lost. To resolve small details that have the size of the order a, we need to include in the Fourier spectrum harmonics of the order $k_t \approx 2\pi/a$. The smallest detail that will be retained at optically long distance from the source plane corresponds to $k_t \sim k$, meaning that the smallest $a \approx 2\pi/k = \lambda$. This limitation is called the *diffraction limit*.

Washing out image detail in free-space propagation can be considered analytically by calculating the corresponding Fourier integral

$$\mathbf{E}(x,y,z) = \frac{1}{4\pi^2}\int_{-\infty}^{+\infty} \mathbf{E}(k_x,k_y)e^{-j(k_x x + k_y y \pm \sqrt{k^2-k_x^2-k_y^2}\,z)}\,dk_x\,dk_y, \qquad \text{Eq 5. 4}$$

where x, y are the coordinates of the source and image planes, and wave propagation is along z (along ***n***). An illustration is given in Figure 5.7.

Now we are ready to consider an interface between free space and a double negative medium. We will write formulas for TM polarization. The other polarization is dual, and the results will be the same.

The reflection (R) and transmission (T) coefficients can be found in the usual way in terms of the wave impedances for tangential field components (requiring continuity). The result reads

$$R = \frac{\eta - \eta_0}{\eta + \eta_0}, \qquad T = \frac{2\eta}{\eta + \eta_0}, \qquad \text{Eq 5. 5}$$

where η_0 is the impedance in free space (see (Eq 4.3)) and η is the impedance in the double-negative medium. To find it, we should substitute negative values of the material parameters in (Eq 4.3):

$$\eta = \frac{k_{\mathrm{n}}}{\omega\epsilon}, \qquad k_{\mathrm{n}} = \sqrt{k^2 - k_{\mathrm{t}}^2}.$$ Eq 5. 6

The impedance in free space we write for the free-space parameters:

$$\eta_0 = \frac{k_{\mathrm{n}}}{\omega\epsilon_0}, \qquad k_{\mathrm{n}} = \sqrt{k_0^2 - k_{\mathrm{t}}^2}.$$ Eq 5. 7

Let us first consider propagating waves (real values of k_n). In a DNG medium, $\varepsilon < 0$ and $\mu < 0$, and $k_n < 0$ (the backward-wave medium). It is important to note that if $\varepsilon = -\varepsilon_0$ and $\mu = -\mu_0$, we have $k = -k_0$, $k^2 = k_0^2$, $\eta = \eta_0$, and

$$R = 0, \qquad T = 1.$$ Eq 5. 8

This result means that for these values of material parameters of the doublenegative medium, we have total transmission and no reflection *for all angles of incidence*. Indeed, this result is independent from the value of k_t, as long as it is smaller than k (that is, for all propagating incident waves).

But if we excite the same interface with an evanescent wave (with $|k_t| > |k|$), the result is dramatically different. For evanescent waves,

$$k_{\mathrm{n}} = \sqrt{k^2 - k_{\mathrm{t}}^2} = -j\alpha, \quad k_{\mathrm{n}} = \sqrt{k_0^2 - k_{\mathrm{t}}^2} = -j\alpha_0, \quad \alpha > 0, \quad \alpha_0 > 0.$$ Eq 5. 9

The square-root branch is chosen so that the fields in both media decay away from the source, which is located in free space. Furthermore, the impedances are also imaginary:

$$\eta = \frac{k_n}{\omega\epsilon} = \frac{-j\alpha}{\omega\epsilon}, \qquad \eta_0 = \frac{k_n}{\omega\epsilon_0} = \frac{-j\alpha_0}{\omega\epsilon_0}.$$ Eq 5. 10

We can now proceed to substitute the derived values into the expressions for the reflection and transmission coefficients, specifically those denoted by (5.5), and subsequently analyze the resulting outcomes. It is evident that when the permittivity, ϵ, is equal to $-\epsilon 0$, and the permeability, μ, is equal to $-\mu 0$, the attenuation constant, α, becomes $\alpha 0$. Furthermore, the wave

impedances are purely imaginary, leading to the characteristic impedance, η, being equal to −η0 for all values of k_t. This condition results in a resonance phenomenon.

$T = \frac{2\eta}{\eta + \eta_0} \to \infty, \qquad R = \frac{\eta - \eta_0}{\eta + \eta_0} \to \infty.$	Eq 5. 11

This resonance phenomenon indicates the presence of surface modes, specifically surface plasmon-polaritons. In typical scenarios, such as interfaces between metals and dielectrics, surface modes at a particular frequency correspond to one or a few discrete values of the propagation constant, k_t. However, in this unique case, the resonance condition is satisfied for all values of kt greater than *k*. In other words, at the specific frequency where the permittivity $\epsilon = -\epsilon 0$ and the permeability $\mu = -\mu 0$, the dispersion equation holds true for all wavenumbers associated with surface states. This distinct property carries significant implications and forms the foundation for the concept of the "perfect lens".

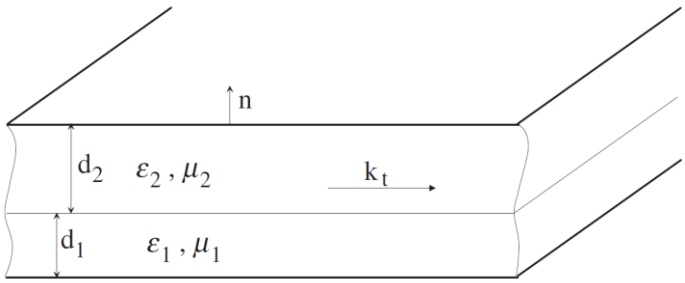

Fig. 5. 6: A pair of isotropic slabs between two perfectly conducting screens.

As an example, let us consider a planar double-layer bounded on bothsides by perfect mirrors, Fig. 5.6. The dispersion equation for waves between the two mirrors can be derived by requiring the continuity of the tangential fields at the interface between the two slabs. The result reads

$\frac{k_{n1}}{\epsilon_1}\tan k_{n1}d_1 + \frac{k_{n2}}{\epsilon_2}\tan k_{n2}d_2 = 0,$ TM modes	Eq 5. 12
$\frac{\mu_1}{k_{n1}}\tan k_{n1}d_1 + \frac{\mu_2}{k_{n2}}\tan k_{n2}d_2 = 0,$ TE modes,	Eq 5. 13

where $k_{n1,2} = \sqrt{k_{1,2}^2 - k_t^2}$ **. The indices 1, 2 refer to the parameters of the two slabs. Let us first study waves travelling along *n* (standing waves between the two metal boundaries) and let** k_t=0 **in the above equations. The eigenvalue equation simplifies to:**

$\frac{\mu_1}{k_1}\tan k_1 d_1 + \frac{\mu_2}{k_2}\tan k_2 d_2 = 0.$	Eq 5. 14

In particular, for thin layers, retaining the first terms of the Taylor expansions, we get $\mu_1 d_1 + \mu_2 d_2 = 0$. This is the resonance condition for standing waves bouncing between the two mirrors, but note that if one of the slabs is a double-negative one, the distance between the two mirrors can now be as small as we like, very much in contrast to usual resonators, whose minimum size equals half wavelength. Physically, it happens because, in one slab, we have phase advance; in the other one, we have phase delay, and they can compensate for each other.

5.3. Conclusion

This chapter introduced the concept of metamaterials, focusing specifically on double-negative materials. It explored the fundamental principles underlying metamaterials and their unique electromagnetic properties, particularly the ability to exhibit negative refractive index. The chapter provided a foundation for understanding how these artificially engineered materials can manipulate electromagnetic waves in ways not possible with naturally occurring materials. The concluding section summarized the key concepts and likely highlighted the potential applications and future research directions for metamaterials, paving the way for the next chapter, which will likely cover meta-devices (in the course presentation “ppt”).

Quiz & Control Questions

1. Explain the notions of isotropy, anisotropy, and bianisotropy.
2. Discuss the differences in the electromagnetic response of double and single split rings. Which shape is preferable if the goal is to create an artificial magnetic material?
3. Discuss the differences between omega and chiral field coupling. Do you know any materials in nature that possess such properties?
4. Give examples of natural and artificial uniaxial media.
5. Consider the excitation of an omega particle formed by a short electric dipole and a conducting loop. Explain what polarizations (electric and magnetic dipole moments) will be induced by a uniform electric field directed along the electric dipole. Answer the same question for excitation by a uniform magnetic field orthogonal to the loop plane. Will there be bianisotropic effects if the electric field is polarized in the loop plane orthogonally to the wire dipole?

Bibliography

[1] Jahani, S., & Jacob, Z. (2016). All-dielectric metamaterials. *Nature nanotechnology*, *11*(1), 23-36.

[2] Litchinitser, N. M., & Sun, J. (2015). Optical meta-atoms: Going nonlinear. Science, 350(6264), 1033-1034.

[3] Asadchy, V. S., Díaz-Rubio, A., & Tretyakov, S. A. (2018). Bianisotropic metasurfaces: physics and applications. Nanophotonics, 7(6), 1069-1094.

[4] Sounas, D. L., & Alù, A. (2017). Non-reciprocal photonics based on time modulation. Nature Photonics, 11(12), 774-783.

[5] Simovski, C. R. (2010). On electromagnetic characterization and homogenization of nanostructured metamaterials. *Journal of Optics*, *13*(1), 013001.

[6] Serdyukov, A., Semchenko, I., Tretyakov, S., & Sihvola, A. (2001). Electromagnetics of bi-anisotropic materials: Theory and applications.

[7] Sihvola, A. H. (1999). *Electromagnetic mixing formulas and applications* (No. 47). Iet.

[8] Bykov, V. P. (1975). Spontaneous emission from a medium with a band spectrum. *Soviet Journal of Quantum Electronics*, *4*(7), 861.

[9] Meade, R. D. V., Johnson, S. G., & Winn, J. N. (2008). Photonic crystals: Molding the flow of light.

[10] Sakoda, K., & Sakoda, K. (2005). *Optical properties of photonic crystals* (Vol. 2, p. 253). Berlin: Springer.

[11] Troia, B., Paolicelli, A., De Leonardis, F., & Passaro, V. M. (2013). Photonic crystals for optical sensing: A review. *Advances in photonic Crystals*, 241-295.

[12] Butt, M. A., Khonina, S. N., & Kazanskiy, N. L. (2021). Recent advances in photonic crystal optical devices: A review. *Optics & laser technology*, *142*, 107265.

[13] McGurn, A. R. (2022). *Introduction to photonic and phononic crystals and metamaterials*. Springer Nature.

Bibliography

[14] Li, L., Shi, Y., & Cui, T. J. (Eds.). (2024). *Electromagnetic metamaterials and metasurfaces: from theory to applications*. Singapore: Springer.

[15] Valipour, A., Kargozarfard, M. H., Rakhshi, M., Yaghootian, A., & Sedighi, H. M. (2022). Metamaterials and their applications: an overview. *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, *236*(11), 2171-2210.

[16] Maier, S. A. (2007). *Plasmonics: fundamentals and applications* (Vol. 1, p. 245). New York: springer.

[17] Wehrspohn, R. B., Kitzerow, H. S., & Busch, K. (Eds.). (2008). *Nanophotonic materials: photonic crystals, plasmonics, and metamaterials*. John Wiley & Sons.