Quantum Electronics **49** (3) 258–265 (2019)

INTEGRATED OPTICS

https://doi.org/10.1070/QEL16747

On the nonlocality of the reflection of an electromagnetic wave from a flat surface upon excitation of surface waves

A.B. Petrin

Abstract. **Diffraction phenomena arising due to reflection of a plane electromagnetic wave from a metal film, which are caused by a limited size of the incident beam aperture, are considered in the Kretschmann scheme used to excite a surface plasmon wave. It is shown that the wave reflection occurs nonlocally, that is, the region of the surface on which the wave is incident and the region from which it is reflected do not substantially coincide. The localisation of the maximum field is investigated when a surface plasmon is excited on a silver film. The absolute values of the gain in terms of the amplitude of the field on the free surface of the film are calculated. The proposed theory is applied to explain the Goos–Hänchen effect.**

Keywords: *surface waves, surface plasmons, integrated optics, optical sensors.*

1. Introduction and problem statement

Currently, much attention is being paid to the development of multichannel optical sensor systems used in biological, chemical and physical studies and utilising surface plasmon (or plasmon-polariton) waves to diagnose changes in the propagation of surface waves caused by changes in the surface properties of the medium under study [1–4]. Such waves, for example, propagate along the metal surface, being localised near it [5, 6]. A frequently used method for the excitation of surface plasmons is the Krechmann method [7].

In the framework of the Krechmann scheme [8] (Fig. 1), a surface plasmon wave (*1*) on the surface of a metal film (*2*) and an incident wave (*3*) are matched with a glass prism (*4*).

A linearly polarised monochromatic wave $\boldsymbol{E}^{\mathrm{i}}$ is incident on the metal film from the side of the prism. The electric field vector of the incident wave lies in the plane of incidence (p-polarisation). The prism is necessary for matching the incident plane wave and the surface plasmon wave by the wavenumber, since the wavenumber of the surface plasmon wave is greater than that of the plane wave of the same frequency in free space [9, 10]. To ensure the equality of the wavenumbers of the incident (*3*), reflected (*5*) and surface plasmon (*1*) waves along the metal surface (Fig. 1), it is needed that the incident wave approaches the film from an optically denser medium (where the wavenumber is greater than that in free space). The refractive index of the prism and the angle of incidence are chosen so as to ensure the equality of the wavenumbers of the incident and surface waves. With these parameters, in the absence of a metal film (and excitation of surface plasmons), one can observe the total internal reflection of the incident wave.

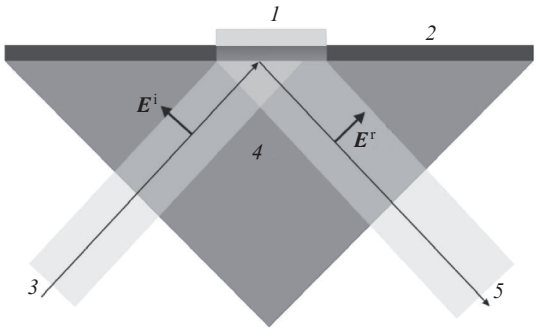

Figure 1. Excitation of a surface plasmon wave (*1*) on the surface of a metal film (*2*) according to the Kretschmann scheme: the incident wave (*3*) in the prism (*4*) generates a surface plasmon wave (*1*) and a reflected wave (*5*).

Experiments show that if the angle of incidence meets the condition of matching the wavenumbers of the incident and surface plasmon waves, then there is a sharp decrease in the reflection coefficient. For example, according to the theory outlined below, when the wavelength of the incident wave in vacuum is $\lambda = 633$ nm and the thickness of the silver film is approximately 53.8 nm (this value depends on the specific value of the refractive index of the prism and the frequency of the incident wave), there is a strong change in the reflection coefficient – almost from unity to zero in the vicinity of the angle of incidence corresponding to the matching of the free and surface plasmon waves. With a small deviation of the angle of incidence from the angle of matching, the value of the reflection coefficient changes from almost zero to unity.

Because the field of the surface plasmon wave is concentrated in a thin layer near the metal surface, the behaviour of reflection in the Krechmann scheme strongly depends on changes in the refractive index in the thin layer near the surface. This is the basis for the widespread use of the Krechmann scheme for producing various types of highly sensitive sensors [11]. Studies have shown that sensitivity is dependent on the angular width of the reflection minimum [12], which, in turn, is determined by the absorption of radiation in a metal film. The lower the absorption in the metal, the narrower the minimum.

A.B. Petrin Joint Institute for High Temperatures, Russian Academy of Sciences, ul. Izhorskaya 13, stroenie 2, 125412 Moscow, Russia; e-mail: a_petrin@mail.ru

Received 13 June 2018; revision received 18 November 2018
Kvantovaya Elektronika **49** (3) 258–265 (2019)
Translated by I.A. Ulitkin

To increase the sensitivity of the sensors by reducing the minimum angular width of the minimum, it was proposed to decrease the thickness of the metal film. It turned out that the attenuation of the surface plasmon wave in a thin film decreases with decreasing film thickness. However, to match the surface wave in the Krechmann scheme, it was necessary to introduce an additional matching layer between the prism and the metal film [13–16], and in some cases even a multilayer film [17, 18].

Note that surface waves can also propagate in dielectric films of planar waveguides [19]. In this connection, it was shown [20, 21] that the Krechmann scheme can be used to excite a surface wave (not necessarily a surface plasmon wave), which is matched in the wavenumber with the incident wave. In this case, a film (or several films) of optically dense material (with a large refractive index), separated from the prism by a film of a matching substance with a low refractive index, can be employed as a waveguide of a surface wave.

For devices of finite size (in the plane of a multilayer structure), the aperture of the incident beam has a limited cross section; therefore, reflection will not occur quite exactly as in the case of an infinite structure and ideally plane waves. In our paper [22], a narrow reflection minimum is shown to retain in the case of a limited aperture, but the amplitudes of the reflected waves have an angular distribution within the diffraction angle. Obviously, in determining the size of the sensor layer of the probe, it is necessary that the region of the film coating in which the reflected field is localised be sufficiently large. This paper is devoted to the study of this problem on the example of the excitation of surface plasmons on a metal film. In addition, the amplification of the field on the free surface of the film, important for a number of technical applications, is investigated.

2. Incidence of a plane electromagnetic limited-aperture beam on a plane-layered structure

Let a plane monochromatic wave with cyclic frequency ω and unit amplitude (Fig. 2) be incident from half-space (*1*) (prism) filled with a homogeneous isotropic dielectric with refractive index n_1 on a flat film with thickness d_1 at an angle α to the normal. A generalisation to an arbitrary number of layers will be given below. For definiteness, it is assumed that the film is metallic and has a complex permittivity ε_2 at the considered frequency ω, and behind the film there is a homogeneous half-space with a permittivity ε_3 (free space). Let the size D of the aperture of the wavefront of the beam in the xz plane be such that the wave is incident on a segment of the interface of length L, i.e. $D = L\cos\alpha$. In the future we will use the generalisation of ideas of work [23] (pp 377, 378).

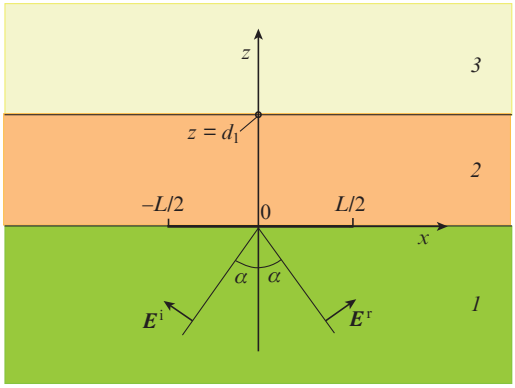

Figure 2. Geometry of the electromagnetic problem. A plane p-polarised wave is incident from the prism (*1*) on a metal film (*2*) with vacuum (*3*) behind the film.

Maxwell's equations in these regions can be written as

$$\operatorname{curl}\boldsymbol{E}_j = \mathrm{i}\omega\boldsymbol{B}_j, \tag{1}$$

$$\operatorname{curl}\boldsymbol{B}_j = -\mathrm{i}\omega\varepsilon_j\mu_j\boldsymbol{E}_j, \tag{2}$$

where $\boldsymbol{E}_j$ and $\boldsymbol{B}_j$ are the vectors of the electric field strength and magnetic field induction in the region with the number j; $j = 1$ corresponds to half-space (*1*) (for which $z < 0$); $j = 2$ corresponds to the region of the film (*2*) ($0 < z < d_1$); and $j = 3$ corresponds to free half-space ($z > d_1$). The complex representation of fields in time is described by the multiplier $\exp(-\mathrm{i}\omega t)$.

We will consider fields that are independent of the y coordinate and have only x- and z-components of the electric field strength vector. From (1) and (2) we find the wave equation for the electric field

$$\operatorname{curl}\operatorname{curl}\boldsymbol{E}_j - \omega^2\mu_j\varepsilon_j\boldsymbol{E}_j = 0. \tag{3}$$

Given that $\operatorname{div}\boldsymbol{D}_j = 0$, or $\partial_x E_{jx} + \partial_z E_{jz} = 0$, we obtain the expression

$$\operatorname{curl}\operatorname{curl}\boldsymbol{E}_j = \boldsymbol{e}_x(-\partial^2_{xx}E_{jx} - \partial^2_{zz}E_{jx}) + \boldsymbol{e}_z(-\partial^2_{zz}E_{jz} - \partial^2_{xx}E_{jz}),$$

where $\boldsymbol{e}_x$ and $\boldsymbol{e}_z$ are unit vectors in the x- and z-axis directions. As a result, from (3) we find the equations for the components E_{jx} and E_{jz}:

$$\partial^2_{xx}E_{jx} + \partial^2_{zz}E_{jx} + \omega^2\mu_j\varepsilon_j E_{jx} = 0, \tag{4}$$

$$\partial^2_{zz}E_{jz} + \partial^2_{xx}E_{jz} + \omega^2\mu_j\varepsilon_j E_{jz} = 0. \tag{5}$$

We represent the components of the electric fields in the form of Fourier expansions:

$$E_{jx}(x,z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{E}_{jx}(\xi,z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi,$$

$$E_{jz}(x,z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{E}_{jz}(\xi,z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi,$$

where the Fourier transforms are determined by the integrals

$$\tilde{E}_{jx}(\xi,z) = \int_{-\infty}^{+\infty}E_{jx}(x,z)\exp(-\mathrm{i}\xi x)\,\mathrm{d}x,$$

$$\tilde{E}_{jz}(\xi,z) = \int_{-\infty}^{+\infty}E_{jz}(x,z)\exp(-\mathrm{i}\xi x)\,\mathrm{d}x.$$

Then, for the Fourier transforms of the fields, we obtain from (4) and (5) the equations

$$\frac{\mathrm{d}^2\tilde{E}_{jx}}{\mathrm{d}z^2} + (\omega^2\mu_j\varepsilon_j - \xi^2)\tilde{E}_{jx} = 0, \tag{6}$$

$$\frac{\mathrm{d}^2\tilde{E}_{jz}}{\mathrm{d}z^2}+(\omega^2\mu_j\varepsilon_j-\xi^2)\tilde{E}_{jz}=0. \tag{7}$$

The solutions of these equations are $\tilde{E}^{\pm}_{jx}=\hat{E}^{\pm}_{jx}\exp(\pm k_{jz}z)$, $\tilde{E}^{\pm}_{jz}=\hat{E}^{\pm}_{jz}\exp(\pm k_{jz}z)$, where $\hat{E}^{+}_{jx},\hat{E}^{-}_{jx},\hat{E}^{+}_{jz}$, and $\hat{E}^{-}_{jz}$ are functions of only ξ, and $k_{jz}=k_{jz}(\xi)=(\omega^2\mu_j\varepsilon_j-\xi^2)^{1/2}$.

The general solution of equation (3) for polarisation in the plane of incidence and a given variation of the field along the x axis during propagation along the z axis and in the opposite direction can be written as

$$\boldsymbol{E}^{\pm}_j=\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}\hat{E}^{\pm}_{jx}(\xi)\\0\\\hat{E}^{\pm}_{jz}(\xi)\end{pmatrix}\exp[\pm\mathrm{i}k_{jz}(\xi)z]\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi, \tag{8}$$

where $\xi^2+k^2_{jz}(\xi)=\omega^2\mu_j\varepsilon_j$. Since $\mathrm{div}\boldsymbol{D}_j=0$, we have

$$\varepsilon_j\xi\hat{E}_{jx}\pm\varepsilon_j k_{jz}(\xi)\hat{E}_{jz}=0. \tag{9}$$

Using formula (9), expression (8) can be rewritten as

$$\boldsymbol{E}^{\pm}_j=\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\\mp\dfrac{\xi}{k_{jz}(\xi)}\end{pmatrix}\hat{E}^{\pm}_{jx}(\xi)\exp[\pm\mathrm{i}k_{jz}(\xi)z]\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi. \tag{10}$$

The expressions for the fields in the considered regions can be written in the form: in region 1

$$\boldsymbol{E}_1(x,z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\-\dfrac{\xi}{k_{1z}}\end{pmatrix}\hat{E}^{+}_{1x}(\xi)\exp(\mathrm{i}k_{1z}z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi$$

$$+\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\\dfrac{\xi}{k_{1z}}\end{pmatrix}\hat{E}^{-}_{1x}(\xi)\exp(-\mathrm{i}k_{1z}z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi, \tag{11}$$

in region 2

$$\boldsymbol{E}_2(x,z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\-\dfrac{\xi}{k_{2z}}\end{pmatrix}\hat{E}^{+}_{2x}(\xi)\exp(\mathrm{i}k_{2z}z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi$$

$$+\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\\dfrac{\xi}{k_{2z}}\end{pmatrix}\hat{E}^{-}_{2x}(\xi)\exp[-\mathrm{i}k_{2z}(z-d_1)]\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi, \tag{12}$$

in region 3

$$\boldsymbol{E}_3(x,z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\-\dfrac{\xi}{k_{3z}}\end{pmatrix}\hat{E}^{+}_{3x}(\xi)\exp[\mathrm{i}k_{3z}(z-d_1)]\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi. \tag{13}$$

The obtained expressions for the fields must satisfy the known boundary conditions. Consider the boundary of the first and second regions, $z=0$. The tangential components of the electric field and the normal components of electrical induction on this boundary must be continuous, that is,

$$E_{1x}(x,0)-E_{2x}(x,0)=0, \tag{14}$$

$$\varepsilon_1E_{1z}(x,0)-\varepsilon_2E_{2z}(x,0)=0, \tag{15}$$

where

$$E_{1x}(x,0)=E^{+}_{1x}(x,0)+E^{-}_{1x}(x,0),$$

$$E_{2x}(x,0)=E^{+}_{2x}(x,0)+E^{-}_{2x}(x,0),$$

$$E_{1z}(x,0)=E^{+}_{1z}(x,0)+E^{-}_{1z}(x,0),$$

$$E_{2z}(x,0)=E^{+}_{2z}(x,0)+E^{-}_{2z}(x,0).$$

Note that the condition $\varepsilon_jE_{jz}=\mathrm{const}$ in our case is equivalent to the condition of continuity of the tangential component of the magnetic field strength B_{jy}/μ_j (the magnetic field has only a component along the y axis).

From the boundary conditions (14) and (15), using expressions (11) and (12), at $z=0$ we obtain two equations:

$$\hat{E}^{+}_{1x}+\hat{E}^{-}_{1x}-\hat{E}^{+}_{2x}-\hat{E}^{-}_{2x}\exp(\mathrm{i}k_{2z}d_1)=0,$$

$$\varepsilon_1\left(-\frac{\xi}{k_{1z}}\hat{E}^{+}_{1x}+\frac{\xi}{k_{1z}}\hat{E}^{-}_{1x}\right)$$

$$-\varepsilon_2\left[-\frac{\xi}{k_{2z}}\hat{E}^{+}_{2x}+\frac{\xi}{k_{2z}}\hat{E}^{-}_{2x}\exp(\mathrm{i}k_{2z}d_1)\right]=0,$$

which can be rewritten in a matrix form

$$\begin{bmatrix}1&1\\-\varepsilon_1\dfrac{\xi}{k_{1z}}&\varepsilon_1\dfrac{\xi}{k_{1z}}\end{bmatrix}\begin{pmatrix}\hat{E}^{+}_{1x}\\\hat{E}^{-}_{1x}\end{pmatrix}=\begin{bmatrix}1&\exp(\mathrm{i}k_{2z}d_1)\\-\varepsilon_2\dfrac{\xi}{k_{2z}}&\varepsilon_2\dfrac{\xi}{k_{2z}}\exp(\mathrm{i}k_{2z}d_1)\end{bmatrix}\begin{pmatrix}\hat{E}^{+}_{2x}\\\hat{E}^{-}_{2x}\end{pmatrix}$$

or

$$\begin{pmatrix}\hat{E}^{+}_{1x}\\\hat{E}^{-}_{1x}\end{pmatrix}=\begin{bmatrix}1&1\\-\varepsilon_1\dfrac{\xi}{k_{1z}}&\varepsilon_1\dfrac{\xi}{k_{1z}}\end{bmatrix}^{-1}\begin{bmatrix}1&\exp(\mathrm{i}k_{2z}d_1)\\-\varepsilon_2\dfrac{\xi}{k_{2z}}&\varepsilon_2\dfrac{\xi}{k_{2z}}\exp(\mathrm{i}k_{2z}d_1)\end{bmatrix}\begin{pmatrix}\hat{E}^{+}_{2x}\\\hat{E}^{-}_{2x}\end{pmatrix}. \tag{16}$$

Similarly, we consider the boundary of the second and third regions ($z=d_1$). The boundary conditions are written in the form

$$E_{2x}(x,d_1)-E_{3x}(x,d_1)=0,$$

$$\varepsilon_2E_{2z}(x,d_1)-\varepsilon_3E_{3z}(x,d_1)=0.$$

Two matrix equations follow from them:

$$\begin{bmatrix}\exp(\mathrm{i}k_{2z}d_1)&1\\-\varepsilon_2\dfrac{\xi}{k_{2z}}\exp(\mathrm{i}k_{2z}d_1)&\varepsilon_2\dfrac{\xi}{k_{2z}}\end{bmatrix}\begin{pmatrix}\hat{E}^{+}_{2x}\\\hat{E}^{-}_{2x}\end{pmatrix}=\begin{bmatrix}1&0\\-\varepsilon_3\dfrac{\xi}{k_{3z}}&0\end{bmatrix}\begin{pmatrix}\hat{E}^{+}_{3x}\\0\end{pmatrix},$$

$$\begin{pmatrix}\hat{E}^{+}_{2x}\\\hat{E}^{-}_{2x}\end{pmatrix}=\begin{bmatrix}\exp(\mathrm{i}k_{2z}d_1)&1\\-\varepsilon_2\dfrac{\xi}{k_{2z}}\exp(\mathrm{i}k_{2z}d_1)&\varepsilon_2\dfrac{\xi}{k_{2z}}\end{bmatrix}^{-1}\begin{bmatrix}1&0\\-\varepsilon_3\dfrac{\xi}{k_{3z}}&0\end{bmatrix}\begin{pmatrix}\hat{E}^{+}_{3x}\\0\end{pmatrix} \tag{17}$$

From matrix equations (16), (17) we obtain the expression

$$\begin{pmatrix}\hat{E}^{+}_{1x}\\\hat{E}^{-}_{1x}\end{pmatrix}=\hat{M}\begin{pmatrix}\hat{E}^{+}_{3x}\\0\end{pmatrix}, \tag{18}$$

where the matrixx $\hat{M}$ is represented as a product of three matrices: $\hat{M} = \hat{T}_1\hat{T}_2\hat{T}_3$, in which

$$\hat{T}_1 = \begin{bmatrix} 1 & 1 \\ -\varepsilon_1\dfrac{\xi}{k_{1z}} & \varepsilon_1\dfrac{\xi}{k_{1z}} \end{bmatrix}^{-1} = \frac{1}{2}\begin{bmatrix} 1 & -\dfrac{k_{1z}}{\varepsilon_1\xi} \\ 1 & \dfrac{k_{1z}}{\varepsilon_1\xi} \end{bmatrix},$$

$$\hat{T}_2 = \begin{bmatrix} 1 & \exp(\mathrm{i}k_{2z}d_1) \\ -\varepsilon_2\dfrac{\xi}{k_{2z}} & \varepsilon_2\dfrac{\xi}{k_{2z}}\exp(\mathrm{i}k_{2z}d_1) \end{bmatrix}\begin{bmatrix} \exp(\mathrm{i}k_{2z}d_1) & 1 \\ -\varepsilon_2\dfrac{\xi}{k_{2z}}\exp(\mathrm{i}k_{2z}d_1) & \varepsilon_2\dfrac{\xi}{k_{2z}} \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} \cos(k_{2z}d_1) & \mathrm{i}\dfrac{k_{2z}}{\varepsilon_2\xi}\sin(k_{2z}d_1) \\ \mathrm{i}\dfrac{\varepsilon_2\xi}{k_{2z}}\sin(k_{2z}d_1) & \cos(k_{2z}d_1) \end{bmatrix},$$

$$\hat{T}_3 = \begin{bmatrix} 1 & 0 \\ -\varepsilon_3\dfrac{\xi}{k_{3z}} & 0 \end{bmatrix}.$$

From (18) we obtain that $\hat{E}^+_{1x} = M_{11}\hat{E}^+_{3x}$ and $\hat{E}^-_{1x} = M_{21}\hat{E}^+_{3x}$. By introducing the function $R(\xi) = -M_{21}/M_{11}$, we can express the reflected wave as $\hat{E}^-_{1x} = R(\xi)\hat{E}^+_{1x}$. In addition, we introduce the function $\mathfrak{T}(\xi) = 1/M_{11}$ with the help of which we can write $\hat{E}^+_{3x} = \mathfrak{T}(\xi)\hat{E}^+_{1x}$.

Note that formula (18) can be generalised to the case of a structure of N layers by induction, which leads to the expression:

$$\begin{pmatrix} E^+_{1x} \\ E^-_{1x} \end{pmatrix} = \hat{M}\begin{pmatrix} E^+_{N+2\,x} \\ 0 \end{pmatrix};$$

moreover, the matrix $\hat{M}$ has the form

$$\hat{M} = \hat{T}_1\left(\prod_{m=2}^{N+1}\hat{T}_m\right)\hat{T}_{N+2},$$

where

$$\hat{T}_1 = \frac{1}{2}\begin{bmatrix} 1 & -\dfrac{k_{1z}}{\varepsilon_1\xi} \\ 1 & \dfrac{k_{1z}}{\varepsilon_1\xi} \end{bmatrix},$$

$$\hat{T}_m = \begin{bmatrix} \cos(k_{mz}d_{m-1}) & \mathrm{i}\dfrac{k_{mz}}{\varepsilon_m\xi}\sin(k_{mz}d_{m-1}) \\ \mathrm{i}\dfrac{\varepsilon_m\xi}{k_{mz}}\sin(k_{mz}d_{m-1}) & \cos(k_{mz}d_{m-1}) \end{bmatrix},$$

$$\hat{T}_{N+2} = \begin{bmatrix} 1 & 0 \\ -\varepsilon_{N+2}\dfrac{\xi}{k_{N+2\,z}} & 0 \end{bmatrix}.$$

In this paper, we will be interested in incident, $\boldsymbol{E}^{\mathrm{i}}_1(x,z)$, and reflected, $\boldsymbol{E}^{\mathrm{r}}_1(x,z)$, waves in the first medium ($z \leqslant 0$) and in a wave behind the film in the third medium $\boldsymbol{E}_3(x,z)$ ($z \geqslant d_1$). These fields, taking into account (11) and (13), can be rewritten as

$$\boldsymbol{E}^{\mathrm{i}}_1(x,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix} 1 \\ 0 \\ -\dfrac{\xi}{k_{1z}} \end{pmatrix}\hat{E}^+_{1x}(\xi)\exp(\mathrm{i}k_{1z}z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi, \quad (19)$$

$$\boldsymbol{E}^{\mathrm{r}}_1(x,z) = -\frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix} 1 \\ 0 \\ \dfrac{\xi}{k_{1z}} \end{pmatrix}R(\xi)\hat{E}^+_{1x}(\xi)\exp(-\mathrm{i}k_{1z}z)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi, \quad (20)$$

$$\boldsymbol{E}_3(x,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix} 1 \\ 0 \\ -\dfrac{\xi}{k_{3z}(\xi)} \end{pmatrix}\mathfrak{T}(\xi)\hat{E}^+_{1x}(\xi)$$

$$\times\exp[\mathrm{i}k_{3z}(z-d_1)\exp(\mathrm{i}\xi x)\,\mathrm{d}\xi. \quad (21)$$

Formulas (19) and (20) show that the incident field can be decomposed into plane waves with different harmonic variation along the x axis [i.e. with dependence on x in the form of $\exp(\mathrm{i}\xi x)$] and each such plane wave will be reflected from the flat-layered structure with the reflection coefficient $R(\xi)$ at an angle α corresponding to the given ξ. To prove this, we first consider the incident of a plane infinite wave of unit amplitude onto a film. The incident field at a plane boundary has a variation along the x axis, proportional to $\exp(\mathrm{i}k_{1x}x) = \exp(\mathrm{i}k_1x\sin\alpha)$, where $k_1 = n_1\omega/c$ is the wavenumber in the first medium, c is the speed of light in a vacuum, and $n_1 = \sqrt{\varepsilon_1}$ is the refractive index of the first medium.

Let the incident wave be polarised in the plane of incidence. The electric field of this wave can be then represented (Fig. 2) in the form

$$\boldsymbol{E}^{\mathrm{i}}_1(x,z) = \boldsymbol{E}^+_1(x,z) = \exp[\mathrm{i}(k_1x\sin\alpha + k_1z\cos\alpha)]$$

$$\times(\boldsymbol{e}_x\cos\alpha - \boldsymbol{e}_z\sin\alpha).$$

At $z = 0$ we obtain

$$\hat{E}^{\mathrm{i}}_{1x} = \int_{-\infty}^{+\infty}E^+_x(x)\exp(-\mathrm{i}\xi x)\,\mathrm{d}x$$

$$= \int_{-\infty}^{+\infty}\exp(\mathrm{i}k_1x\sin\alpha)\exp(-\mathrm{i}\xi x)\cos\alpha\,\mathrm{d}x$$

$$= 2\pi\delta(k_1\sin\alpha - \xi)\cos\alpha.$$

From formulas (19) and (20) we find the field in the first medium (taking into account the fact that $k_{1z}|_{\xi = k_1\sin\alpha} = k_1\cos\alpha$):

$$\boldsymbol{E}_1(x,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\begin{pmatrix} 1 \\ 0 \\ -\dfrac{\xi}{k_{1z}} \end{pmatrix}\exp(\mathrm{i}k_{1z}z) - \begin{pmatrix} 1 \\ 0 \\ \dfrac{\xi}{k_{1z}} \end{pmatrix}R(\xi)\exp(-\mathrm{i}k_{1z}z)\right]$$

$$\times 2\pi\delta(k_1\sin\alpha - \xi)\exp(\mathrm{i}\xi x)\cos\alpha\,\mathrm{d}\xi = \begin{pmatrix} \cos\alpha \\ 0 \\ -\sin\alpha \end{pmatrix}$$

$$\times\exp[\mathrm{i}(k_1x\sin\alpha + k_1z\cos\alpha)] - R(k_1\sin\alpha)$$

$$\times\begin{pmatrix} \cos\alpha \\ 0 \\ \sin\alpha \end{pmatrix}\exp[\mathrm{i}(k_1x\sin\alpha - k_1z\cos\alpha)].$$

As a result, we obtain that the full field is the sum of the fields of the incident and reflected plane waves, with the reflection coefficient $R = R(k_1\sin\alpha)$.

Suppose now that a bounded plane wave is incident on the film, and at $z = 0$, it (Fig. 2) is only incident on the segment $x \in [-L/2, L/2]$. Then,

$$\hat{E}^{\mathrm{i}}_{1x}(\xi) = \int_{-L/2}^{L/2} E^{+}_{x}(x)\exp(-\mathrm{i}\xi x)\,\mathrm{d}x$$

$$= \int_{-L/2}^{L/2} \exp(\mathrm{i}k_1 x\sin\alpha)\exp(-\mathrm{i}\xi x)\cos\alpha\,\mathrm{d}x$$

$$= 2\cos\alpha\frac{\sin[(k_1\sin\alpha-\xi)L/2]}{k_1\sin\alpha-\xi}. \qquad (22)$$

Substituting this expression into formulas (19) and (20), we find the full field in the first medium

$$\boldsymbol{E}_1(x,z) = \frac{L\cos\alpha}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\-\dfrac{\xi}{k_{1z}}\end{pmatrix}\frac{\sin[(k_1\sin\alpha-\xi)L/2]}{(k_1\sin\alpha-\xi)L/2}$$

$$\times\exp[\mathrm{i}(\xi x + k_{1z}z)]\,\mathrm{d}\xi - \frac{L\cos\alpha}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\\dfrac{\xi}{k_{1z}}\end{pmatrix}$$

$$\times\frac{R(\xi)\sin[(k_1\sin\alpha-\xi)L/2]}{(k_1\sin\alpha-\xi)L/2}\exp[\mathrm{i}(\xi x - k_{1z}z)]\,\mathrm{d}\xi. \qquad (23)$$

The first integral in (23) is the sum of plane waves into which the incident beam with a limited aperture can be decomposed. The second integral is the corresponding sum of the reflected plane waves. The influence of the aperture on the reflected wave leads to the distribution of the incident wave energy over waves in different directions in the vicinity of the angle α.

Thus, the full field in the first medium can be represented as the sum of the incident and reflected fields: $\boldsymbol{E}_1(x,z) = \boldsymbol{E}^{\mathrm{i}}_1(x,z) + \boldsymbol{E}^{\mathrm{r}}_1(x,z)$. The expressions for the incident and reflected fields in the region $z \leqslant 0$ and for the field passed to the third medium in the region $z \geqslant d_1$ can be written in the form

$$\boldsymbol{E}^{\mathrm{i}}_1(x,z) = \frac{L\cos\alpha}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\-\dfrac{\xi}{k_{1z}}\end{pmatrix}\frac{\sin[(k_1\sin\alpha-\xi)L/2]}{(k_1\sin\alpha-\xi)L/2}$$

$$\times\exp[\mathrm{i}(\xi x + k_{1z}z)]\,\mathrm{d}\xi, \qquad (24)$$

$$\boldsymbol{E}^{\mathrm{r}}_1(x,z) = -\frac{L\cos\alpha}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\\dfrac{\xi}{k_{1z}}\end{pmatrix}\frac{R(\xi)\sin[(k_1\sin\alpha-\xi)L/2]}{(k_1\sin\alpha-\xi)L/2}\times$$

$$\times\exp[\mathrm{i}(\xi x - k_{1z}z)]\,\mathrm{d}\xi, \qquad (25)$$

$$\boldsymbol{E}_3(x,z) = \frac{L\cos\alpha}{2\pi}\int_{-\infty}^{\infty}\begin{pmatrix}1\\0\\-\dfrac{\xi}{k_{3z}}\end{pmatrix}\frac{\sin[(k_1\sin\alpha-\xi)L/2]}{(k_1\sin\alpha-\xi)L/2}$$

$$\times\mathfrak{T}(\xi)\exp\{\mathrm{i}[\xi x + k_{3z}(z-d_1)]\}\,\mathrm{d}\xi. \qquad (26)$$

It follows from (25) that the amplitude of the field reflected at an angle β within the small angle $\mathrm{d}\beta$ (given that $\xi = k_1\sin\beta$ and $\mathrm{d}\xi = k_1\cos\beta\mathrm{d}\beta$) is

$$\frac{\mathrm{d}E^{\mathrm{r}}_{1\mathrm{a}}}{\mathrm{d}\beta} = \frac{k_1 L\cos\alpha}{2\pi}\left|\frac{\sin[(\sin\alpha-\sin\beta)k_1L/2]}{(\sin\alpha-\sin\beta)k_1L/2}R(k_1\sin\beta)\right|. \qquad (27)$$

Expression (27) determines the angular distribution of the amplitude of the reflected field when a wave with a unit amplitude is incident from the first medium on the boundary $z = 0$. The incident beam is limited by the aperture $D = L\cos\alpha$.

3. Field distribution at the film boundaries. Reflection nonlocality

We consider the scheme for the excitation of Krechmann surface plasmons (Fig. 1). We assume that the film is made of silver. As was shown, for example, in [21], with an infinite aperture of the incident beam, there is an optimal silver film thickness. For a prism with a refractive index $n_1 = 1.6$, the optimum film thickness and its complex permittivity are listed in Table 1. It is assumed that the wavelength of the incident wave in vacuum is $\lambda = 633$ nm. In this case, a beam with an infinitely large aperture, incident at an angle α, generates a wave that is reflected strictly at an angle α (the distribution over the angles of reflection is a δ-function). The dependence of the reflection coefficient on the angle of incidence α for this case is shown in Fig. 3. In the case considered, the entire surface of the interfaces is excited (the aperture of the incident beam is infinite).

Table 1.

Medium number j	Layer name	Relative permittivity	Film thickness/nm
1	prism	1.6[2]	semi-infinite medium
2	silver film	$\varepsilon_2 = -18.2 + \mathrm{i}0.5$	53.8
3	free space (air)	1	semi-infinite medium

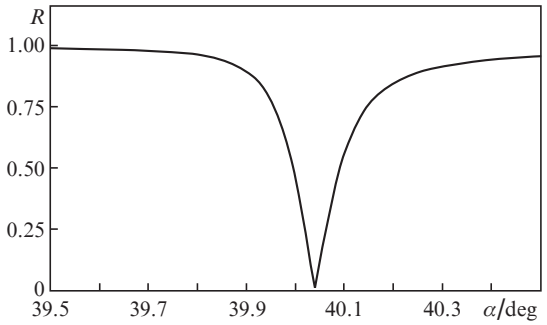

Figure 3. Field reflection coefficient upon incidence of an infinite p-polarised plane wave on a silver film as a function of the angle of incidence α.

If we now illuminate the film with a limited-aperture beam at an optimal angle $\alpha = \alpha_{\mathrm{opt}} \approx 40.041°$, then because of finiteness of L, the distribution of the reflected field over the reflection angles β will no longer be a δ-function. We introduce the function:

$$F(\beta) = \frac{2\pi}{k_1 L \cos\alpha_{\text{opt}}} \frac{\mathrm{d}E_{1\text{a}}^{\text{r}}}{\mathrm{d}\beta}$$

$$= \left| \frac{\sin[(\sin\alpha_{\text{opt}} - \sin\beta)k_1 L/2]}{(\sin\alpha_{\text{opt}} - \sin\beta)k_1 L/2} R(k_1 \sin\beta) \right|. \qquad (28)$$

The function $F(\beta)$ is proportional to the amplitude of the electric field strength of the wave reflected at an angle β and determines the angular dependence of the reflection coefficient. Figure 4 shows the dependence $F(\beta)$ for $L = 25$, 50, and 100 μm.

One can see from Fig. 4 that with an increase in L and, hence, in the apertures $D = L\cos\alpha_{\text{opt}}$, the angular spread of the spatial harmonics decreases. The angular spread up to a constant can be estimated by the function $\Delta\beta \propto \lambda_1/D$, which is exactly determined by expression (28). Considering the dependences in Fig. 4 in terms of the sensitivity of the sensors, we can draw the following conclusion. If the sensor analyses the angular spectrum of the reflected waves as in [24], then at a diffraction width $\Delta\beta$ several times greater than the angular width of the minimum associated with the surface wave (Fig. 3), the diffraction broadening of the angular spectrum caused by the finiteness of the incident beam aperture will not affect the surface plasmon sensor sensitivity. In addition, Fig. 4 shows that the larger the L, the greater part of the wave energy is in the zone of strong interaction with the surface plasmon wave and the more efficiently it is excited.

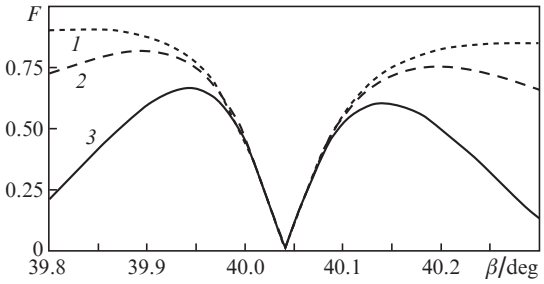

Figure 4. Angular distribution of the reflected signal amplitude $F(\beta)$ in the case of incidence of a plane p-polarised wave with a limited aperture of the beam on a silver film at $L = (1)$ 25, (2) 50 and (3) 100 μm.

We now consider how the finite size of the aperture affects the localisation of the reflected field at the boundary $z = 0$. Incident, $\boldsymbol{E}_1^{\text{i}}(x,0)$, and reflected, $\boldsymbol{E}_1^{\text{r}}(x,0)$, fields at this boundary are determined by formulas (24) and (25), respectively. Figure 5a shows the distribution of the amplitudes of the x-components of the incident, $|E_{1x}^{\text{i}}(x,0)|$, and reflected, $|E_{1x}^{\text{r}}(x,0)|$, fields for $L = 50$ μm. It is seen that the region onto which the wave is incident, and the region from which it is reflected, are different. Near the discontinuity points of wave amplitudes, oscillations arise associated with the well-known Gibbs phenomenon in the case of Fourier transforms of discontinuous functions [25] (p. 91). However, this will not affect the physical results of the problem in question.

One can explain the behaviour of dependences in Fig. 5a as follows. At point $x = -L/2$, the incident wave begins to interact with the film; at this point the amplitude of the surface wave, which radiates into the prism, begins to increase. This wave interferes with the wave reflected from the boundary $z = 0$ and forms a reflected wave, whose amplitude

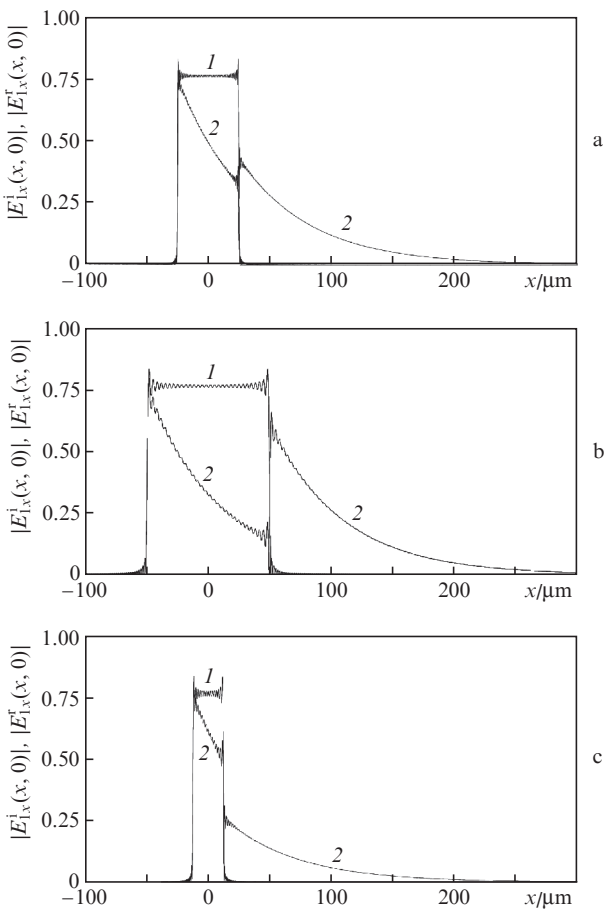

Figure 5. Distribution of the amplitudes of the x-components of the (*1*) incident, $|E_{1x}^{\text{i}}(x,0)|$ and (*2*) reflected, $|E_{1x}^{\text{r}}(x,0)|$, fields at the boundary z = 0 for $L =$ (a) 50, (b) 100 and (c) 25 μm.

decreases with increasing x. To the right of point $x = L/2$, the incident wave disappears, as well as the wave reflected from the boundary $z = 0$, and only the wave emitted by the surface wave into the prism remains, the amplitude of which decreases exponentially as the surface wave attenuates due to this re-emission.

Note that the surface wave propagates at the free boundary beyond the excitation region and at the same time radiates into the prism. The field of the surface wave decays here exponentially, and, as calculations have shown, the decay constant is independent of L.

The given explanation is confirmed by Figs 5b and 5c, which show the same dependences as in Fig. 5a, but for $L =$ 100 and 25 μm, respectively (i.e., for the twice as large and twice as small aperture of the incident beam).

Figure 5b shows that within the region of incidence of the surface wave, its amplitude increases more than in Fig. 5a. This leads to a greater decrease in the amplitude of the reflected wave at the boundary $x = L/2$ and a larger amplitude of the wave radiated at $x > L/2$ by the surface wave into the prism. Figure 5c shows the reverse changes. Note that for all three apertures, the amplitude of the surface wave at $x > L/2$ decreases exponentially with the same decay constant.

This fact can be confirmed by considering the field at $z = d_1$. This field is determined by the surface plasmon wave excited on the outer surface of the metal film.

Figure 6a shows the distributions of the amplitudes of the x-component, $|E_{3x}(x,d_1)|$ [curve (*1*)], and the z-component, $|E_{3z}(x,d_1)|$ [curve (*2*)], of the electric field at the interface of the film and free space at $z = d_1$, calculated by formula (26) at $L = 50$ μm. For comparison, the distribution of the x-component $|E^{i}_{1x}(x,0)|$ of the incident electric field at the interface of the film and the prism is also shown at $z = 0$ [curve (*3*)]. At the boundary $z = d_1$, there is only a surface plasmon wave. It can be seen that its amplitude increases to the right boundary of the excitation region and then decreases exponentially when the external feed of the surface wave stops. In the place where the incident wave disappears, the amplitudes of the components of the surface wave have maxima.

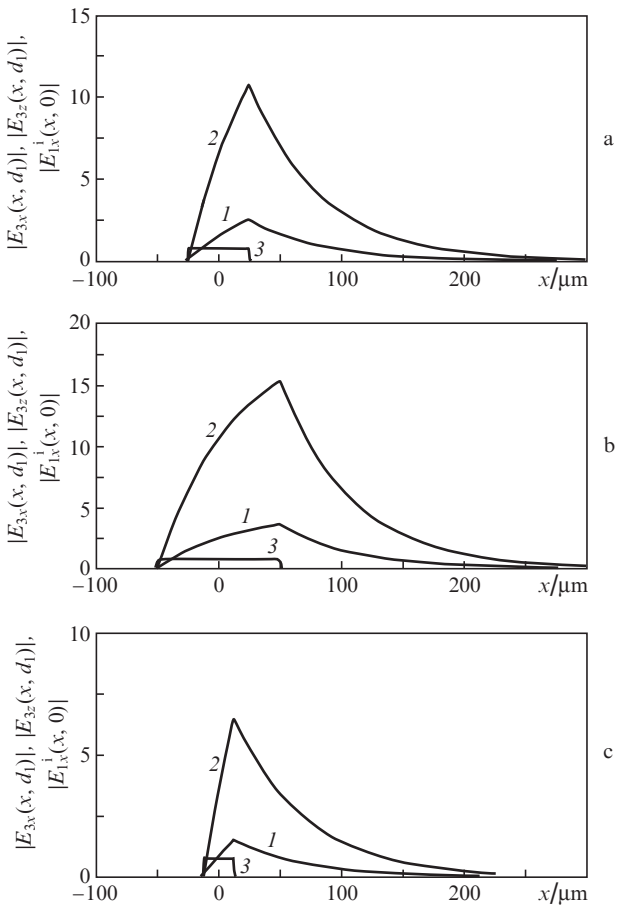

Figure 6. Distribution of the amplitudes of the (*1*) x-component, $|E_{3x}(x,d_1)|$, and (*2*) z-component, $|E_{3z}(x,d_1)|$, of the electric field at the boundary $z = d_1$ between the film and free space, and (3) the x-component $|E^{i}_{1x}(x,0)|$ of the incident electric field at the boundary $z = 0$ between the film and prism at L = (a) 50, (b) 100 and (c) 25 μm.

Figures 6b and 6c show the same dependences as in Fig. 6a, but for $L = 100$ and 25 μm, respectively. It is seen from Fig. 6 that the longer the length L, the greater the magnitude of the electric field amplitude in the maximum on the right boundary of the excitation region. The characteristic length over which the amplitude of the surface wave decreases is the same for all three cases. The amplitudes of the fields are normalised to the unit amplitude of the incident wave. Knowledge of the maximum field regions and the amplitudes of the field components in the maximum may be important for such applications as optical emission of electrons from the surface of metal films [26, 27] and generation of high optical harmonics [28].

Let us return to the reflected waves in Fig. 5 [curves (*2*)]. One wave is reflected to the right of the left boundary of the illuminated region ($x = -L/2$), and the second is reflected to the right of the right boundary of this region ($x = L/2$). Obviously, the contribution to the first wave of the surface wave is minimal (the amplitude of the surface wave is still small), and the second wave is entirely due to the radiation of the surface wave into the prism. Therefore, these waves, generally speaking, are reflected with a phase shift, which strongly depends on the influence of the medium on the surface wave. This phenomenon is usually called the Goos–Hänchen effect [29, 30]. The effect is used to design highly sensitive optical sensors [31–33]. Based on the above theory, we can use the amplitudes of the two waves reflected from the inner surface of the film ($z = 0$) to find the phase difference of the waves reflected far from the film, near the detector. As the calculations showed, the phase difference of these waves is close to π and is exactly equal to π in the far zone, where the deep minimum of interference is observed at an angle α_{opt} (see Fig. 4). From Fig. 5 it can be seen that the distance between the reflected beams in the near zone is $D = L\cos\alpha_{opt}$t.

Note that the results of calculations performed in this work are consistent with the experimental results of work [34] on the study of the Goos–Hänchen effect. However, in addition to the incident p-polarised wave, Moskalenko et al. [34] considered an s-polarised wave, which is reflected without a shift. A very narrow incident beam is also investigated. If we neglect the s-polarised wave and consider only the incident p-polarised wave, as in the present paper, one could noticed that the distance between the reflected beams is close to $D = L\cos\alpha_{opt}$.

4. Conclusions

A theoretical method is proposed for studying diffraction phenomena arising due to reflection of a plane electromagnetic beam with a limited aperture from a silver film in the Kretschmann scheme, which is easily generalised to the case of a multilayer metal–dielectric structure. It is shown that in the presence of a surface wave, the reflection of the wave occurs nonlocally, that is, the region of wave incidence on the film and the region from which the reflected wave is emitted do not coincide. The conditions for maximising the amplitude of the surface wave at the free boundary of the film and the absolute values of the field amplification are determined. The application of this theory to the description of the Goos–Hänchen effect is considered. The obtained results can be useful in the development of modern optical multichannel biological, chemical and physical sensor systems, as well as systems with optically stimulated electron emission and nonlinear optics systems.

References

1. Homola J., Yee S.S., Gauglitz G. *Sens. Actuators B*, **54**, 3 (1999).
2. Barnes W.L., Dereux A., Ebbesen T.W. *Nature*, **424**, 824 (2003).
3. Homola J. *Chem. Rev.*, **108**, 462 (2008).
4. Spoto G., Minunni M. *J. Phys. Chem. Lett.*, **3**, 2682 (2012).
5. Raether H. *Surface Plasmons* (Berlin: Springer-Verlag, 1988).
6. Barnes W.L. *J. Opt. A: Pure Appl. Opt.*, **8**, S87 (2006).
7. Kretschmann E., Raether H. *Z. Naturforsch. A*, **23**, 2135 (1968).

8. Piliarik M., Homola J. *Opt. Express*, **17** (19), 16505 (2009).
9. Liedberg B., Nylander C., Lundstrom I. *Sens. Actuators*, **4**, 299 (1983).
10. Liedberg B., Nylander C., Lundstrom I. *Biosens. Bioelectron*., **10**, i-ix (1995).
11. Garabedian R., Gonzalez C., Richards J., et al. *Sens. Actuators A*, **43**, 202 (1994).
12. Yeatman E.M. *Biosens. Bioelectron*., **11**, 635 (1996).
13. Sarid D. *Phys. Rev. Lett*., **47**, 1927 (1981).
14. Matsubara K., Kawata S., Minami S. *Opt. Lett*., **15**, 75 (1990).
15. Yang F., Bradberry G.W., Sambles J.R. *Phys. Rev. Lett*., **66**, 2030 (1991).
16. Kessler M.A., Hall E.A.H. *Thin Solid Films*, **272**, 161 (1996).
17. Nenninger G.G., Tobiska P., Homola J., Yee S.S. *Sens. Actuators B*, **74**, 145 (2001).
18. Toyama S., Doumae N., Shoji A., Ikariyama Y. *Sens. Actuators B*, **65**, 32 (2000).
19. Airoudj A., Debarnot D., Beche B., Poncin-Epaillard F. *Anal. Chem*., **80**, 9188 (2008).
20. Petrin A.B., Vol'pyan O.D., Sigov A.S. *Opt. Spectrosc*, **123** (5), 798 (2017) [*Opt. Spektrosk*., **123** (5), 786 (2017)].
21. Petrin A.B., Vol'pyan O.D., Sigov A.S. *Tech. Phys*., **63** (3), 422 (2018) [*Zh. Tekh. Fiz*., **88** (3), 433 (2018)].
22. Petrin A.B. *Opt. Spectrosc*., **125** (3), 390 (2018) [*Opt. Spektrosk*., **125** (3), 375 (2018)].
23. Novotny L., Hecht B. *Principles of Nano-Optics* (Cambridge: Cambridge Univ. Press, 2006; Moscow: Fizmatlit, 2009).
24. Löfas S., Malmqvist M., Rönnberg I., et al. *Sens. Actuators B*, **5**, 79 (1991).
25. Angot A. *Compléments de mathématiques à l'usage des ingénieurs de l'électrotechnique et des télécommunications* (Paris: Masson, 1972; Moscow: Nauka, 1967).
26. Racz P., Irvine S.E., Lenner M., et al. *Appl. Phys. Lett*., **98** (11), 111116 (2011).
27. Welsh G.H., Wynne K. *Opt. Express*, **17**, 2470 (2009).
28. Kim S., Jin J., Kim Y., et al. *Nature*, **453** (5), 757 (2008).
29. Goos F., Hänchen H. *Ann. Phys*., **1**, 333 (1947).
30. Renard R. *J. Opt. Soc. Am*., **54**, 1190 (1964).
31. Yin X., Hesselink L. *Appl. Phys. Lett*., **89**, 261108 (2006).
32. Wan Y., Zheng Z., Zhu J. *J. Opt. Soc. Am. B*, **28** (2), 314 (2011).
33. Parks A.D., Spence S.E. *Appl. Opt*., **54** (18), 5872 (2015).
34. Moskalenko B.V., Soboleva I.B., Fedyanin A.A. *JETP Lett*., **91** (8), 382 (2010) [*Pis'ma Zh. Eksp. Teor. Fiz*., **91** (8), 414 (2010)].