Chapter 2
Theory of Wood’s Anomalies

Daniel Maystre

Abstract Discovered by Wood in 1902, grating anomalies have fascinated specialists of optics for more than one century. Long after the first interpretation given by Rayleigh, Fano has suggested that the origin of anomalies could be found in the excitation of surface waves. This chapter describes the quantitative phenomenological theory of Wood’s anomalies developed in the 1970s, based on the interpretation given by Fano and on the macroscopic laws of electromagnetics. This theory leads to a formula giving the efficiency of gratings in the region of anomaly and predicts the phenomenon of total absorption of light by a grating.

2.1 Introduction

In 1902, Wood, observing the spectrum of a continuous light source given by an optical metallic diffraction grating, noticed a surprising phenomenon: “I was astounded to find that under certain conditions, the drop from maximum illumination to minimum, a drop certainly of from 10 to 1, occurred within a range of wavelengths not greater than the distance between the sodium lines” [1]. Wood made a crucial remark: these lines were present only for p-polarized light, i.e. when the magnetic field is parallel to the grating grooves. However, he was unable to provide any interpretation to these phenomena and thus termed them “singular anomalies”, concluding that this problem was “one of the most interesting that I have ever met with”. Even though the notion of Surface Plasmon Polariton (SPP) appeared more than half a century afterwards, let us give back to Caesar what is Caesar’s: Wood must be considered as the initiator of plasmonics.

D. Maystre (✉)
Institut Fresnel, CNRS, Université d’Aix Marseille, Ecole Centrale Marseille, Avenue Escadrille Normandie-Niemen, Campus universitaire de Saint Jérôme, 13397 Marseille, France
e-mail: daniel.maystre@fresnel.fr
As it will be seen in the following, Wood’s discovery immediately raised considerable attention and the fascination of many specialists of optics for the so-called Wood’s anomalies never died. Rayleigh proposed the first explanation for the existence of the anomalies [2, 3]: an anomaly in a given spectrum occurs at a wavelength corresponding to the passing-off of a spectrum of higher order, in other words, at the wavelength for which a scattered wave emerges tangentially to the grating surface. The Rayleigh conjecture was considered as a valuable tool for the prediction of Wood’s anomalies. Indeed, the famous grating formula:

$$\sin (\theta_n) = \sin (\theta) + n\lambda/d, \quad (2.1)$$

where $\theta$ is the angle of incidence (measured anticlockwise from the normal to the grating), $\theta_n$ is the angle of diffraction (measured clockwise), $\lambda$ is the wavelength in vacuum (which can also be considered as the wavelength in the air) and $d$ is the groove period, allows one to rigorously calculate the diffraction angle of any scattered order $n$ from the grating period, the angle of incidence and the wavelength of light. The passing-off of the order $n$ occurs when $\sin (\theta_n) = \pm 1$ and thus, from Eq. (2.1), the wavelengths of a spectrum generating the passing-off of a diffracted order are given by:

$$n\lambda/d = -\sin (\theta) \pm 1, \quad n = \pm 1, \pm 2, \pm 3... \quad (2.2)$$

Figure 2.1 shows the spectra obtained by Wood for some values of the angle of incidence.

Near-normal incidence ($\theta = 4^\circ 12'$, top of the figure), a bright narrow line appeared in the yellow (about $\lambda = 610$ nm), while a larger dark line was observed in the green (near 520 nm). Decreasing the angle of incidence to $2^\circ 37'$, these lines approached one another and for angles of incidence of $0^\circ 15'$ and $0^\circ 5'$, they came in contact. Finally, at normal incidence, the lines fused and a uniform illumination was observed. With an incidence on the other side of the normal, two lines separated again, corresponding to red and orange. These lines were extremely brilliant up to a certain wavelength where the intensity very suddenly dropped to values close to zero, this fall occurring within a range not greater than the distance between the sodium lines. The paper by Wood did not mention the characteristics of the grating (period, shape, metal) but subsequently, Wood communicated to Rayleigh the period, equal to 1,760 nm. Using Eq. (2.2), Rayleigh was able to check the validity of his conjecture. The calculation predicts that, for an angle of incidence of $4^\circ 12'$, the anomalies should arise at wavelengths 543.7 nm ($n = +3$ in Eq. (2.2) with sign +) and 629.6 nm ($n = -3$ in Eq. (2.2) with sign −). The discrepancy of about 5% with the real location of anomalies in Fig. 2.1 (517 and 609 nm) seems hardly good enough. Rayleigh expressed the hypothesis that this mismatch was the consequence of an imprecise knowledge of the grating period and for 30 years, the Rayleigh conjecture remained unquestioned. The present chapter is devoted to the theory of Wood’s anomalies and it is not our purpose to mention all the experimental contributions to this phenomenon. The interested reader can find a survey of these contributions in Chap. 1. However,
at least one contribution must be cited here since it demonstrated that the Rayleigh prediction is unable to explain some vital experimental results.

Strong [4] showed Wood’s anomalies for various metallic gratings having the same period. The results implicitly evidenced the influence of the metal on the shape
of the anomalies and, much more important, on their location. Bearing in mind that the grating formula is purely geometrical and thus that the metal has no effect on the location of the passing-off, this result must be considered as the first reappraisal of the Rayleigh interpretation.

Some years later, Fano achieved the first theoretical breakthrough on Wood’s anomalies [5]. Observing the published experimental data, Fano distinguished two kinds of anomalies:

- A sharp anomaly—that is, an edge of intensity—appears along the spectrum at sharply defined wavelengths governed by the grating formula using the Rayleigh conjecture.
- A diffuse anomaly extends for a wavelength interval from the first one (the edge) to the red (i.e. higher wavelengths) and “consists generally of a minimum and a maximum of intensity (one dark band and one bright band)”.

So, Fano explained the discrepancy between the theoretical predictions by Rayleigh and the experimental data provided by Wood: the Rayleigh conjecture predicts the location of the sharp anomaly, while obviously the anomaly observed by Wood was the diffuse anomaly. Fano explained the diffuse anomaly by “a forced resonance” related to the “leaky waves supportable by the grating”. In the following, we will denote by “Rayleigh anomaly” the sharp anomaly at the passing-off of a spectrum of higher order. We will see in the following that the remarkable analysis by Fano must be considered as the starting point of the modern explanation of Wood’s anomalies, even though the connection between the leaky waves mentioned by Fano and SPPs was not stated in the paper. Hessel and Oliner were led to the same conclusions as those stated by Fano [6]. In addition, these two authors used numerical tools in order to calculate the location and shape of the anomalies. Unfortunately, the model used by Hessel and Oliner was based on the knowledge a priori of the electromagnetic impedance (ratio of the tangential components of the electric/magnetic fields) on a straight line located above the grating grooves. Nevertheless, they were able to explain some properties experimentally known, such as the possibility of anomalies for p-polarized light with very deep gratings [7, 8], or the reluctance of anomalies to merge [9].

The modern analysis of Wood’s anomalies began at the end of the 1960s and at the beginning of 1970s thanks to two revolutions, which drastically changed the experimental and theoretical tools in the study of gratings. The use of laser sources and photosist layers permitted the invention and the production of holographic gratings for scientific and industrial purpose [10]. For the first time, the holographic technology provided a rapid and accurate tool for constructing gratings with submicronic periods. At the same time, the opportunity of using the first powerful computers encouraged theoreticians to develop rigorous vector theories of gratings and to implement them on computers. These new and powerful tools allowed for the first time wide numerical studies of Wood’s anomalies and led to the first successful quantitative comparisons between experimental data and numerical calculations. As a consequence, the conditions were fulfilled for developing phenomenological
Theories of Wood's anomalies and for checking their predictions with numerical and experimental data.

The first rigorous vector theory of gratings able to predict with precision the properties of relief metallic gratings for any shape of the profile in any range of wavelength was achieved by Maystre in 1972 [11, 12]. This theory can be classified as an integral theory since it reduces the problem of scattering by a grating to the solution of an integral equation. It is amazing to notice that, at that time, this new theory was intended to provide a realistic tool for the optimization of metallic gratings embarked in satellites for spectroscopy in the ultraviolet region, where the conductivity of metals falls down. Indeed, it was considered that the grating problem was already solved in the visible and near-infrared regions, using the theory of perfectly conducting gratings [13–17]. It was accepted that the metal can be replaced by a perfectly conducting (impenetrable) material, the efficiencies in the various orders computed using this hypothesis being finally multiplied by the reflectivity of the metal. At first glance, this approximation seems to be justified as the reflectivity of metallic planes of aluminium, silver or gold in these regions exceeds 90% in general. Nevertheless, the new theory was applied to metallic gratings in the visible and near-infrared regions, in order to confirm the hypothesis of perfectly conductivity of metallic gratings in these regions.

As expected, for s-polarized light, the numerical results confirmed the usual hypothesis: the intensities scattered in the various orders can be deduced from those obtained assuming a perfect conductivity of the metals through a simple multiplication factor close to the reflectivity of the plane metallic surface. On the other hand, for p-polarized light, the first results were quite surprising since strong discrepancies appeared. As a consequence of this unexpected result, a crucial conclusion must be stated: for natural light where the intensities take the average value between both fundamental polarizations, the perfect conductivity model for gratings fails. This result was presented in 1972 by Petit et al. in a communication at the International Congress of Optics (I.C.O. IX, Santa Monica, Ca) [18]. Surprisingly for the authors, serious doubts about the validity of this result were expressed by attendants in this Congress. The criticisms were based on two remarks: is the new theory valid? Is the numerical implementation accurate? It must be recognized that these criticisms were justified: numerical tests (energy balance, reciprocity theorem [19, 20]) are not sufficient for checking the validity of the numerical results deduced from a new theory. Above all, the validity of this theory must be checked by successful comparisons of the numerical results with the experimental data.

The reader can find in Chap. 1 a detailed description of the first successful comparison between the experimental results published by Hutley and the calculations performed by Maystre and McPhedran [21] using the new integral theory and the profile of the holographic grating measured by Hutley. The agreement between the experimental and numerical results for both polarizations was all the more remarkable since on similar experimental data published by Hutley, the theory of perfectly conducting gratings previously used by McPhedran and Waterworth led to significant discrepancies for p-polarized light [22–24].
This first successful comparison between experimental and numerical results for metallic gratings in the visible region was followed by many others, using the same integral theory [25–27] or another rigorous grating theory, the differential theory of gratings [28], in such a way that the theoretical warning stated in 1972 about the non-validity of the perfect conductivity model has been fully confirmed, at least for wavelengths smaller than 10 µm.

Figure 2.2 shows the profile of one of the holographic gratings made by Hutley and Fig. 2.3 gives the experimental and theoretical efficiencies of the order $-1$ and of the total efficiency (sum of all the scattered efficiencies) for one of these gratings, covered with three different metals [27]. For all the three metals, significant absorption peaks can be observed. Two of the three energy curves at a wavelength of 476 nm (silver and aluminium) are rather similar in form, with the third (gold at 476 nm) having greatly broadened absorption regions. It must be noticed that Fig. 2.3 constitutes another indicator as to the accuracy of infinite conductivity model for metals in the visible region. In regions where the energy absorption by gratings can be of the order of 50 %, models, which cannot take this absorption into account, must be suspect.

The existence of powerful computer codes that are able to predict with a relative precision better than 1 % the grating efficiencies in the visible and near-infrared regions has permitted wide numerical studies of Wood’s anomalies. However, numerical results do not provide simple rules or formulae that are able to predict the shape and position of the anomalies in a quantitative manner. The specialist of optics likes to understand the physical origin of the observed phenomena and to use simple rules in order to select the best grating for a given application. As a consequence, attempts were made at developing a quantitative phenomenological theory [29, 30]. The following sections give a detailed description of this theory. It starts from the basic origin of Wood’s anomalies, the excitation of SPPs and then uses the theory of analytic functions of the complex variable to express the efficiency of the grating in the region of anomaly in a simple form using two complex parameters only. It will be shown that this phenomenological formula provides a remarkable precision on the efficiencies. Moreover, this formula will show that an incident wave can be absorbed by a grating in totality, due to the excitation of a SPP.
The theory of Wood’s anomalies described in this chapter is based on electromagnetics. In other words, it starts from the macroscopic properties of metals through electromagnetic parameters like permittivity and permeability and never takes into account their microscopic structure. On the other hand, a group working in the Oak Ridge National Laboratory, which was at the origin of the discovery of SPPs [31], tried to investigate the properties of SPPs using the microscopic laws of solid-state physics. Such a study has an advantage on macroscopic studies: it starts directly from the basic origin of SPPs, i.e. a collective electron resonance. Unfortunately, the complexity of the problem led the authors to use a perturbation treatment, in which the anomaly is studied in terms of an interaction between the incoming photon and the collective electron resonance. This analysis was not able to accurately describe the phenomenon, especially for deep gratings, except by tuning the optical constants of metal to fit the measured data [32–34].

One can find an impressive amount of experimental measurements on phase velocity and damping of SPPs on metallic gratings in the studies published by Raether, Pockrand, Kröger and Kretschmann. These authors illuminated the grating from
the bulk side through a prism, a device named “total attenuated reflection” by these authors. They compared their experimental measurements with an approximate electromagnetic theory [35–43].

2.2 Propagation of Surface Plasmon Polaritons on a Metallic Surface

2.2.1 Case of the Flat Surface

Problem of Scattering by a Flat Surface

In a problem of scattering, an object is illuminated by an incident electromagnetic wave and one wants to determine the total field at any point of space. This total field contains not only the incident field, which is known, but also the scattered field which has been generated by the object, which is unknown. We present in Fig. 2.4 the problem of scattering by a flat metallic surface, in which the scattering object is a metallic half-plane.

An incident plane wave propagating in a lossless dielectric material of optical index \( \nu_1 = \sqrt{\varepsilon_1} \) (with \( \varepsilon_1 \) relative permittivity) illuminates a non-magnetic metallic half-plane of complex index \( \nu_2 = \sqrt{\varepsilon_2} \). The problem of scattering is solved as soon as we know the amplitudes of the reflected and transmitted waves. Indeed, using the complex notation with a time dependence \( \exp(-i\omega t) \), an incident field \( F^i \) with unit amplitude can be written in the form:

\[
F^i = F^i \hat{z} = \exp(i k_1 (\alpha x - \beta y)) \hat{z},
\]

with \( \hat{z} \) being the unit vector of the \( z \)-axis (orthogonal to the \( xy \) plane), \( k_1 = k \nu_1 = \frac{2\pi \nu_1}{\lambda} \) wavenumber in the dielectric material (\( k \) and \( \lambda \) being, respectively, the wavenumber and the wavelength of light in vacuum), and (see Appendix 1):

\[
F^i = \begin{cases} 
\text{incident electric field } E^i \text{ for } s\text{-polarized light,} \\
\text{incident magnetic field } H^i \text{ for } p\text{-polarized light,}
\end{cases}
\]

\[
\alpha = \sin (\theta), \quad \beta = \cos (\theta),
\]

with \( \theta \) being the angle of incidence measured clockwise from the normal to the plane, \( \alpha \) and \( \beta \) being called normalized propagation constants of the wave in the following. It can be easily shown by using Helmholtz equations in both regions and boundary conditions on the interface (Eqs. (2.77) and (2.78)) that the projections on the \( z \)-axis of the reflected and transmitted fields \( F^r \) and \( F^t \) can be written as
Fig. 2.4 The problem of scattering by a flat metallic plane

\[ \begin{align*}
\text{incident wave} & \quad \theta \\
\text{reflected wave} & \\
\text{dielectric (}\nu_1\text{)} & \\
\text{transmitted wave} & \\
\text{metal (}\nu_2\text{)} & \\
\end{align*} \]

\[ F_r = r \exp (i k_1 (\alpha x + i \beta y)) , \quad (2.6) \]
\[ F_t = t \exp (i k_1 (\alpha x - i \gamma y)) , \quad (2.7) \]
\[ \gamma = \sqrt{\nu^2 - \alpha^2} , \quad \nu = \nu_2 / \nu_1 \text{ relative index of metal.} \quad (2.8) \]

It must be noticed that the relative index \( \nu \) being complex, the definition of \( \gamma \) is ambiguous. In order to choose the determination of \( \gamma \) for complex values of \( \nu \), let us put forth two remarks. First, if the metal is replaced by a lossless dielectric material, the choice of the determination of \( \gamma \) is quite clear. Indeed, the transmitted wave expressed in Eq. (2.7) must propagate downwards in order to satisfy a radiation condition, thus the real value of \( \gamma \) must be positive. Secondly, let us now assume that \( \nu \) is purely imaginary, thus \( \nu = i \nu'' \). It should be noticed that this assumption is not far from the actual values of indices of metals in the visible region (for example, the index of aluminium at 647 nm is equal to \( 1.3 + i 7.1 \)). From this assumption, the value of \( \gamma \) is given by \( \gamma = \sqrt{-\nu''^2 - \alpha^2} = \pm i \sqrt{\nu''^2 + \alpha^2} \) and the choice of determination is quite clear: since the field must decrease in modulus as \( y \to -\infty \), the imaginary part must be positive. Thus, in order to adopt a determination of \( \gamma \) which applies to any material, the authorized region of the complex plane of \( \gamma \) must include both half-lines defined by \( \text{Im} (\gamma) = 0, \text{Re} (\gamma) > 0 \) and \( \text{Re} (\gamma) = 0, \text{Im} (\gamma) > 0 \). In real life, the index of metals is neither real nor purely imaginary, but its real and imaginary parts are always positive. Thus, a natural way to adopt a general determination for \( \gamma \) is to choose in the complex plane a half-plane containing both the positive real axis (choice of \( \nu \) for a lossless dielectric material) and the positive imaginary axis (choice of \( \nu \) for a purely imaginary index). In the following, we will determine \( \gamma \) from:

\[ \text{Re} \{ \gamma \} + \text{Im} \{ \gamma \} > 0, \quad (2.9) \]

or in other words by choosing the value of \( \gamma \) in the half-plane located above the second bisector. Figure 2.5 shows the location of this half-plane in the complex plane.

From Eqs. (2.77) and (2.78), one obtains, using the expressions of the field given by Eqs. (2.3), (2.6), and (2.7):
1 + r = t for both polarizations, \hspace{1cm} (2.10)
\[ \beta (-1 + r) = -\gamma t \] for s-polarized light, \hspace{1cm} (2.11)
\[ \frac{\beta}{\varepsilon_1} (-1 + r) = -\frac{\gamma}{\varepsilon_2} t \] for p-polarized light. \hspace{1cm} (2.12)

Solving the linear system of equations expressed in Eqs. (2.10) and (2.11) provides the Fresnel coefficients for s-polarized light:

\[ r = \frac{\beta - \gamma}{\beta + \gamma}, \quad t = \frac{2\beta}{\beta + \gamma}, \hspace{1cm} (2.13) \]

and Eqs. (2.10) and (2.12) yield, for p-polarized light:

\[ r = \frac{\beta/\varepsilon_1 - \gamma/\varepsilon_2}{\beta/\varepsilon_1 + \gamma/\varepsilon_2}, \quad t = \frac{2\beta/\varepsilon_1}{\beta/\varepsilon_1 + \gamma/\varepsilon_2}. \hspace{1cm} (2.14) \]

These Fresnel coefficients allow one to solve the problem of scattering from a flat interface.

**Problem of Guiding by a Flat Surface: the SPP**

Now, we consider a different, but closely related problem: the problem of guiding. The question is to know whether a surface wave can propagate at the surface of the metal. Of course, the answer to this question requires a precise definition of a
guided wave. Classically, a guided wave is a wave propagating along the \( x \)-axis and satisfying a radiation condition at infinity. In other words:

- if \( y \to +\infty \), the field must propagate upwards or vanish,
- if \( y \to -\infty \), the field must propagate downwards or vanish.

First, let us notice that the total field in the guided wave satisfies almost the same conditions as the total field in the scattering problem: it must satisfy Maxwell equations and boundary conditions on the interface. On the other hand, it must satisfy radiation conditions at infinity on both sides of the interface. These radiation conditions make a big difference with the total field in the problem of scattering: in the problem of scattering, the field above the interface contains the incident wave, which does not satisfy the radiation condition since it propagates towards the interface. In other words, a guided wave corresponds to a problem of scattering in which the incident field does not exist. At first glance, such a guided wave should not exist. Indeed, in the scattering problem, the incident energy is shared between the scattered field and the losses inside the metal. On the other hand, in the problem of guiding, there is no incident energy at all, and of course the existence of a field generates losses in the metal. Thus it seems that such a wave cannot satisfy the energy balance. This remark is quite correct as far as the propagation constant of the guided wave is real. If we consider the expressions of the reflected and transmitted waves in the problem of scattering (Eqs. (2.6) and (2.7)), the propagation constant \( k_1 \alpha \) is real since it is imposed by the incident plane wave. This requirement does not hold in a problem of guiding and, from a heuristic point of view, it can be conjectured that a surface wave may propagate along the \( x \)-axis with an exponentially decreasing amplitude, due to the losses inside the metal. The consequence is that the imaginary part of \( k_1 \alpha \) must be positive if its real part is positive.

In conclusion, the search for a surface wave leads to the search for a solution of the field in a scattering problem, but without any incident wave, or in other words, the so-called “homogeneous solution of Maxwell equations”. We know that such a solution will have a complex (with positive imaginary part) propagation constant \( k_1 \alpha \) along the \( x \)-axis. A priori, many possibilities exist according to whether \( \alpha \), or \( k_1 \), or both have a non-null imaginary part. Here, we will consider that \( k_1 \) is real and \( \alpha \) complex. This choice entails that the amplitude of the surface wave decreases in \( x \), but that the frequency \( \omega = k_1 / \sqrt{\varepsilon_1 \mu_0} \) remains real. Other choices could be made. For example, introducing complex values for both \( \alpha \) (with positive imaginary part) and \( k_1 \) (with negative imaginary part) leads, if their product \( k_1 \alpha \) is real, to a constant amplitude along the \( x \)-axis, but an exponential decrease in \( \exp(-i\omega t) \) of this amplitude with time. The heuristic meaning of such a wave is clear: the losses in the metal generate a decrease in time of the amplitude.

The search for the solution of the guiding problem is straightforward from the Fresnel formulae. In order to cancel the incident wave, it suffices to find a pole of the reflection and transmission coefficients \( r \) and \( t \) in Eq. (2.13) for s-polarized light or in Eq. (2.14) for p-polarized light. For the s-polarized light, this pole is the root of the denominator of Eq. (2.13), i.e. the solution of \( \beta + \gamma = 0 \). However, since \( \alpha \) is complex, it is necessary to give a precise definition of the determinations of
β and γ. The determination of γ was given yet in Eq. (2.9) but now, the definition of β from the angle of incidence in the problem of scattering (Eq. (2.5)) does not hold. Since the field above the interface must satisfy the Helmholtz equation (Eq. (2.73)), we must impose that \( \alpha^2 + \beta^2 = 1 \), thus:

\[
\beta = \sqrt{1 - \alpha^2}.
\]  \hspace{1cm} (2.15)

Moreover, the field above the interface must satisfy a radiation condition; it must propagate upwards or decrease in amplitude if \( y \to +\infty \). Consequently, we are led to the same choice as that used for γ in Fig. 2.5, with

\[
\text{Re} \{\beta\} + \text{Im} \{\beta\} > 0.
\]  \hspace{1cm} (2.16)

Finally, the solution of the equation \( \beta + \gamma = 0 \) requires that \( \beta^2 = \gamma^2 \), thus from Eqs. (2.8) and (2.15), \( \nu^2 = 1 \). The solution \( \nu = +1 \) offers no interest since it means that the metal is replaced by the same dielectric material as above the interface. The solution \( \nu = -1 \) is not realistic, at least for a non-magnetic material. The conclusion to be drawn from this result is that for s-polarized light, the propagation on a flat metal surface of a guided wave is impossible.

As regards p-polarized light, according to Eq. (2.14) we have to solve the equation:

\[
\beta/\varepsilon_1 = -\gamma/\varepsilon_2.
\]  \hspace{1cm} (2.17)

Taking the root of both members yields, after simplifications:

\[
\alpha = \tilde{\alpha}^{\text{plane}} = \pm \nu/\sqrt{1 + \nu^2},
\]  \hspace{1cm} (2.18)

which entails, according to Eqs. (2.8) and (2.15):

\[
\tilde{\beta}^{\text{plane}} = -1/\sqrt{1 + \nu^2}, \quad \tilde{\gamma}^{\text{plane}} = \nu^2/\sqrt{1 + \nu^2}.
\]  \hspace{1cm} (2.19)

In the symbol \( \tilde{\alpha}^{\text{plane}} \), the hat “tilde” means that it concerns a guided wave, while the superscript “plane” indicates that the interface is a plane. It can be verified that the values of \( \tilde{\beta}^{\text{plane}} \) and \( \tilde{\gamma}^{\text{plane}} \) obey the determination expressed by Eqs. (2.16) and (2.9) for usual metals (Al, Au, Ag...) in the visible or infrared regions, as well as the initial equation (we must recall that we have taken the square of both members of Eq. (2.17) to find the solution), provided that the square root of an arbitrary complex number \( \rho \exp (i\phi) \), with \( 0 \leq \phi < 2\pi \), \( \rho > 0 \), is, by definition, equal to \( \sqrt{\rho} \exp (i\phi/2) \).

The two opposite values of \( \alpha \) given by Eq. (2.18) represent the constants of propagation of two waves propagating in opposite directions. This kind of wave is called SPP or sometimes surface plasmon or sometimes surface plasmon oscillation. This wave can be written in the normalized form:
Fig. 2.6 Constants of propagation along the \(x\)-axis of a plane wave \((k_1 \sin(\theta))\), of a SPP at the interface between a metal of index \(\nu_2\) and a dielectric of index \(\nu_1(\nu_1 \sqrt{1+\nu_2^2}, \nu = \nu_2/\nu_1)\), and of a SPP at the interface between a metal of index \(\nu_2\) and vacuum \((k\nu_2/\sqrt{1+\nu_2^2})\).

\[
H = \begin{cases} 
\exp \left( ik_1 \left( \tilde{\alpha}_{\text{plane}} x + \tilde{\beta}_{\text{plane}} y \right) \right) & \text{in the dielectric,} \\
\exp \left( ik_1 \left( \tilde{\alpha}_{\text{plane}} x - \tilde{\gamma}_{\text{plane}} y \right) \right) & \text{in the metal.}
\end{cases}
\tag{2.20}
\]

For aluminium at 647 nm, the optical index \(\nu_2\) is equal to \(1.3 + i7.1\) and using fused silica as a dielectric material (index 1.45), the relative index \(\nu\) is equal to \(0.89 + i4.9\), then:

\[
\tilde{\alpha}_{\text{plane}} = 1.019 + i 7.6 \times 10^{-3}.
\tag{2.21}
\]

The real part of \(\tilde{\alpha}_{\text{plane}}\) is slightly greater than unity, while its imaginary part is very small. This is a general result for metals in the visible and infrared regions and it explains why SPPs cannot be excited by a plane wave. Indeed, as shown in Fig. 2.6, the propagation constant on the \(x\)-axis of a plane wave is equal to \(k_1 \sin(\theta)\) (circle), thus it is always smaller than \(k_1\) in modulus, in contrast with the real part of \(k_1 \tilde{\alpha}_{\text{plane}}\). Since the surface is flat, a plane wave cannot excite a field having a significantly different propagation constant and then it cannot excite the SPP.

**SPP and Brewster Effect**

In this section, it is shown that the SPP propagating at a metal–dielectric interface may be deduced by continuity from the well-known phenomenon of total transmission of light between two dielectric materials (Brewster effect). Figure 2.7 shows the trajectory of \(\tilde{\alpha}_{\text{plane}} = \nu_2/\sqrt{1+\nu_2^2}\), normalized propagation constant of the SPP on an air–metal interface, when the imaginary part of the optical index of the metal is varied from zero to infinity, the real part being equal to 1.3, i.e. that of aluminium at 647 nm.

The trajectory, which starts at \(\tilde{\alpha}_{\text{plane}} = 0\) for \(q = +\infty\) (perfectly conducting metal), reaches the point corresponding to aluminium for \(q = 7.1\). Its imaginary part first increases as long as \(q\) remains greater than unity, then it decreases and for \(q = 0\),
Fig. 2.7 Trajectory of $\tilde{\alpha}_\text{plane} = \nu_2/\sqrt{1 + \nu_2^2}$ in the complex plane when $\nu_2 = 1.3 + iq$, $q$ being varied from $+\infty$ to 0. The cross (point 2) represents the value of $\tilde{\alpha}_\text{plane}$ corresponding to the index of aluminium at $\lambda = 647$ nm ($q = 7.1$), points 1 and 6 correspond, respectively, to the cases of a perfectly conducting metal ($q = +\infty$) and a lossless dielectric ($q = 0$). Reprinted by permission from Blackwell Publishing: [30]

$\tilde{\alpha}_\text{plane}$ becomes real and equal to 0.793, while $\tilde{\beta}_\text{plane} = -1/\sqrt{1 + \nu_2^2} = -0.61$ and $\tilde{\gamma}_\text{plane} = \nu_2^2/\sqrt{1 + \nu_2^2} = 1.03$. As a consequence, the plasmon waves propagating in air and in dielectric are homogeneous plane waves propagating downwards, according to Eq. (2.20). In conclusion, the wave propagating in air is incident and generates a transmitted wave in the lossless dielectric, without any reflection in the air. Obviously, this limit is nothing else than the field in the Brewster effect. Besides, it should be noticed that in that case, since the fields are represented by plane waves, $\tilde{\alpha}_\text{plane} = \nu_2/\sqrt{1 + \nu_2^2}$ is nothing else than $\sin(\theta)$, the sine of the angle of incidence, which entails that $\tan(\theta) = \sin(\theta)\sqrt{1 - \sin^2(\theta)} = \nu_2$, which is the Brewster formula. At first glance, the strong link between SPP propagation and Brewster effect seems to be surprising. Indeed, by definition, a SPP is a homogeneous solution of Maxwell equations, i.e. a field scattered on both sides of the interface, without any incident wave. By contrast, in the Brewster effect, the field in the air is an incident wave. The explanation of this apparently paradoxical result is given in Fig. 2.8, where the corresponding trajectory of $\tilde{\beta}_\text{plane}$ is represented. It is worth noting that this trajectory, which starts from 0 and ends at $-0.61$, crosses the second bisector of the complex plane (dashed line). Bearing in mind that a scattered wave is defined as a wave satisfying Eq. (2.16), it emerges that the plasmon wave in the air, which is a scattered wave in the air, continuously becomes an incident wave when the trajectory of $\tilde{\beta}_\text{plane}$ crosses the second bisector.

In order to show the continuous evolution of the wave from the SPP of a perfectly conducting metal to the Brewster effect, we have expressed the field in air (top of Fig. 2.9) and in metal (bottom of Fig. 2.9) in the form $H = \exp[ik(\mathbf{u}_r + i\mathbf{u}_i)\rho]$ with $\rho = (x, y)$. In a lossless dielectric material, the two vectors are orthogonal, a simple consequence of Maxwell equations. The first point corresponds to a perfectly
Fig. 2.8  Trajectory of $\tilde{\beta}_{\text{plane}} = -1/\sqrt{1 + \nu_2^2}$ in the complex plane when $\nu_2 = 1.3 + iq$, $q$ being varied from 0 to $+\infty$. The dashed line shows the second bisector in the complex plane.

Fig. 2.9  Graph representation in air (top) and metal (bottom) of the normalized complex wavevectors $\mathbf{u}_r + i\mathbf{u}_i$ corresponding to the six points of Fig. 2.7. The real part $\mathbf{u}_r$ and the imaginary part $\mathbf{u}_i$ are represented by solid and dashed arrows, respectively. The direction of propagation of the wave is given by $\mathbf{u}_r$, while $\mathbf{u}_i$ indicates the direction of maximum decrease. Reprinted by permission from Blackwell Publishing: [30]

Conducting metal: the wave in air is a homogeneous plane wave propagating parallel to the interface, while in the metal $\mathbf{u}_i$ is infinite, as it will be shown in the following. The evolution from point 1 to point 6 is essentially characterized in the metal by the decrease of $\|\mathbf{u}_i\|$ and in the air by the transformation of the surface wave into an incident homogeneous plane wave.
Excitation of a SPP

The question which arises is to know how the SPP can be excited. It has been shown that it can be provoked by an electron beam on a metallic thin film [44]. Powell and Swan illuminated an aluminium thin film in normal incidence and observed peaks of absorption in the transmitted beam. One of these peaks was attributed to an excitation of a SPP. Clearly, the explanation of this phenomenon should be made in the frame of solid-state physics. However, electromagnetic theory can provide a heuristic interpretation of it. First, let us analyse the structure of the electric field, of the charges and of the currents of the SPP. From Eqs. (2.63) and (2.68) of Appendix 1,

\[
\begin{align*}
\text{in the metal,} & \quad \mathbf{j}_t = i \omega \varepsilon_0 (1 - \varepsilon_2) \mathbf{E}, \\
\text{in the dielectric,} & \quad \mathbf{j}_t = i \omega \varepsilon_0 (1 - \varepsilon_1) \mathbf{E},
\end{align*}
\]

where \( \mathbf{j}_t \) denotes the total current density, which includes both conduction current and the bound current inside the metal. The charge balance can be written in harmonic regime as:

\[
\nabla \cdot \mathbf{j}_t = i \omega \rho_t,
\]

with \( \rho_t \) being the volume density of total charges (including both free and bound charges). In both dielectric and metal, \( \Delta \cdot \mathbf{E} = 0 \), and since \( \mathbf{j}_t \) depend linearly on \( \mathbf{E} \), \( \rho_t \) vanishes: charges are located on the surface but for simplicity we will retain the same symbol \( \rho_t \) to denote the surface charge density (mathematically, \( \rho_t \) is the coefficient of a delta distribution located on the surface). Finally, the volume current density and the surface charge density can be derived from the electric field. Using Eq. (2.67), we deduce that:

\[
\begin{align*}
\text{in the dielectric,} & \quad \mathbf{E} = \frac{ik_1}{\omega \varepsilon_1 \varepsilon_0} \nabla \times \mathbf{H} = \frac{k_1}{\omega \varepsilon_1 \varepsilon_0} \left( -\beta \hat{x} + \alpha \hat{y} \right) \exp \left( ik_1 (\alpha x + \beta y) \right), \\
\text{in the metal,} & \quad \mathbf{E} = \frac{ik_1}{\omega \varepsilon_2 \varepsilon_0} \nabla \times \mathbf{H} = \frac{k_1}{\omega \varepsilon_2 \varepsilon_0} \left( \gamma \hat{x} + \alpha \hat{y} \right) \exp \left( ik_1 (\alpha x - \gamma y) \right).
\end{align*}
\]

Equation (2.80) gives the total surface density of charges:

\[
\rho_t = \varepsilon_0 \left( E_{y^+} - E_{y^-} \right) = \frac{\alpha k_1}{\omega} \left( \frac{1}{\varepsilon_1} - \frac{1}{\varepsilon_2} \right) \exp \left( ik_1 \alpha x \right).
\]

In order to analyse the space distribution of the electric field, the volume current density and the surface charge density, we simplify the expressions of these quantities by assuming that \( \nu \simeq i \nu'' \), with \( \nu'' \) real and \( \nu'' \gg 1, \nu \). Thus the relative index of metal is close to a purely imaginary number large in modulus, which is the case for the actual values of optical indices in the visible and near-infrared regions. From Eqs. (2.18) and (2.19), we find:
The electric field and the total current density at \( t = 0 \) are represented by *solid* and *dashed lines*, respectively. The *signs* represent the total surface charge density

\[
\varepsilon_2 \simeq -\nu''\varepsilon_1, \quad \hat{\alpha}^{\text{plane}} \simeq 1, \quad \hat{\beta}^{\text{plane}} \simeq i/\nu'', \quad \hat{\gamma}^{\text{plane}} \simeq i\nu'', \quad (2.28)
\]

and we can deduce that:

\[
\rho_t \simeq \frac{k_1}{\omega\varepsilon_1} \left( 1 + \frac{1}{\nu''^2} \right) \exp(ik_1x), \quad (2.29)
\]

in the metal, \( \mathbf{E} \simeq -\frac{k_1}{\omega\nu''\varepsilon_0} (i\hat{x} + \frac{\hat{y}}{\nu''}) \exp(ik_1x) \exp(+k_1\nu''y), \quad (2.30) \)

in the dielectric, \( \mathbf{E} \simeq -\frac{k_1}{\omega\nu''\varepsilon_0} (i\hat{x} - \nu''\hat{y}) \exp(ik_1x) \exp\left( -\frac{k_1}{\nu''}y \right). \quad (2.31) \)

The volume current densities can be deduced from Eqs. (2.22), (2.23), (2.30) and (2.31).

Figure 2.10 shows the electric field lines (solid lines), the total current density (dashed lines) and the total surface charge density at \( t = 0 \). Here, we abandon the use of complex amplitudes and the electric field is the real part of the product of its complex amplitude by \( \exp(-i\omega t) \). It is worth noting from Eqs. (2.30) and (2.31) that the modulus of the fields decreases exponentially on both sides of the interface, the decrease being much larger in the metal than in the dielectric, while the phase remains constant in the range \(-\infty < y < +\infty\) as \( x \) is fixed. The electric field lines go from positive to negative surface charge densities. It can be noticed that the total volume current densities in the metal and in the air have opposite contributions to the total surface charge density, but the current density in the air is much smaller than that in the metal and can be neglected.
From Fig. 2.10, we can understand the mechanism of SPP propagation. Due to the current lines, charges are shifted from the region close to \( x = -\lambda_1/4 \) (or \( x = 3\lambda_1/4 \)) towards the region close to \( x = +\lambda_1/4 \) (or \( x = 5\lambda_1/4 \)). As a consequence, the charge in the vicinity of \( x = -\lambda_1/4 \) (or \( x = 3\lambda_1/4 \)) becomes negative and that located around \( x = +\lambda_1/4 \) (or \( x = 5\lambda_1/4 \)) becomes positive. The maximum charge density, located at the origin (or \( x = \lambda_1 \)), is shifted towards \( x = \lambda_1/4 \) (or \( x = 5\lambda_1/4 \)), while the minimum charge density, located at \( x = -\lambda_1/2 \) (or \( x = +\lambda_1/2 \)), is shifted to the right as well. As a consequence, the positive and negative charge densities propagate to the right, as well as the electric field and the volume charge density.

This result allows one to understand why an electron beam can excite a SPP. The electron beam creates local charges and electric fields on the metal surface, these charges and fields generate new charges and electric fields in the vicinity, and so on: as a result, SPPs propagate on both sides of the beam. Furthermore, Fig. 2.10 shows that the field penetrates more deeply in the dielectric than in the metal. More precisely, Eqs. (2.20) and (2.28) show that when the imaginary part of the index is increased, the attenuation of the field in the dielectric decreases, while that in the metal increases. At the limit, when this imaginary part goes to infinity (which corresponds to a perfect conductivity according to Eq. (2.65)), the attenuation does not exist anymore in the dielectric and the field does not penetrate at all in the metal.

In that case, \( \alpha = 1 \) and the field in the dielectric is given by:

\[
H = \exp \left( ik_1 x \right).
\]  

(2.32)

This field satisfies the Helmholtz equation and the boundary condition on the perfectly conducting metal since the electric field is parallel to the \( y \)-axis, thus its tangential component on the metal vanishes. On the other hand, such a field obviously cannot be classified as a surface wave. However, we will see that it becomes an actual surface wave as soon as the metal interface is corrugated, in such a way that it is not completely incorrect to claim that a SPP can propagate on a perfectly conducting metallic surface, even though in that case, the model of electron resonance and the name of SPP seem quite unrealistic.

Another way to excite a SPP on a flat metallic surface is to use a finite width of metal separating the dielectric material from vacuum (Fig. 2.11).

According to Eq. (2.18), and noticing that the relative index of metal with respect to vacuum is equal to \( \nu_2 \), the propagation constant \( k\tilde{\alpha}_{\text{plane}} \) at the metal–vacuum interface is given by:

\[
k\tilde{\alpha}_{\text{plane}} = k\nu_2 \sqrt{1 + \nu_2^2},
\]

and in that case, for aluminium at 647 nm,

\[
\tilde{\alpha}_{\text{plane}} = 1.009 + i \ 3.5 \ 10^{-3}.
\]

(2.34)

The real part of \( \tilde{\alpha}_{\text{plane}} \) is very close to unity, but \( k = k_1/\nu_1 \) is smaller than \( k_1 \), in such a way that, if \( \nu_1 \) if significantly greater than unity, the propagation constant
of the SPP becomes smaller than $k_1$ and, provided that the width of metal is small (in practice, smaller than the skin depth) the SPP can be excited by a plane wave illuminating the thin film of metal from the dielectric side (Fig. 2.11). Let us give in Fig. 2.12 a numerical example of the consequences of this excitation using fused silica (index = 1.45) as a dielectric material and aluminium at 647 nm. The reflectivity shows a resonance phenomenon at $\theta = 44.55^\circ$. This incidence corresponds to a propagation constant of the incident field $k_1 \sin (\theta) = k_\nu_1 \sin (\theta) = 1.017k$. Thus, assuming that the excitation occurs when the propagation constant of the incident field is equal to the real part of the propagation constant of the SPP, it turns out from Eq. (2.33) that $\text{Re}\left\{\tilde{g}^{\text{plane}}\right\}$ should be equal to 1.017, a value slightly different from the value of 1.009 given by Eq. (2.34). This discrepancy is not surprising. In our theoretical calculations, the propagation constant of the SPP has been calculated assuming an infinite width of metal, which is not the case in Fig. 2.11. It is worth noticing that the device shown in Fig. 2.11 is sometimes called “prism device” and has been widely used for both plane and modulated surfaces [35–43]. Indeed, if the light is generated by a light source located in vacuum, a plane air–dielectric interface parallel to the metal–dielectric interface (as in Fig. 2.11) cannot generate inside the
2.2.2 Case of the Diffraction Grating

Scattering from a Diffraction Grating

Figure 2.13 shows a metallic diffraction grating, i.e. a cylindrical periodic interface of period $d$ and height $h$ separating a metallic and a dielectric materials, invariant by translation with respect to the $z$-axis.

The grating is illuminated by a plane wave propagating in the $xy$ plane at incidence $\theta$. Taking into account the results obtained for a flat interface, the study will be restricted to p-polarization, thus the incident magnetic field is parallel to the $z$-axis. In these conditions, it can be shown [29] from the elementary laws of electromagnetics and from theorems of existence and uniqueness of the solution of boundary-value problems that the total magnetic field $H(x, y) = H(x, y) \hat{z}$ remains independent of $z$, parallel to the $z$-axis and that it is pseudo-periodic:

$$H(x + d, y) = H(x, y) \exp(ik_1 \alpha d), \quad \alpha = \sin(\theta). \quad (2.35)$$

It follows from Eq. (2.35) that $H(x, y) \exp(-ik_1 \alpha x)$ is periodic. Expanding this function in Fourier series and introducing it in the Helmholtz equations (Eq. (2.72)), it can be easily shown that the magnetic field can be represented in the major part of space in the form of Rayleigh expansions [29, 45]. Denoting by $y_{\text{max}}$ and $y_{\text{min}}$ the ordinates of the top and of the bottom of the grooves, the Rayleigh expansions can be written:
if \( y > y_{\text{max}} \), \[
H = \exp(ik_1 \alpha x - ik_1 \beta y) + \sum_{n=-\infty}^{+\infty} b_n \exp(ik_1 \alpha_n x + ik_1 \beta_n y),
\]
Equation (2.36)

if \( y < y_{\text{min}} \), \[
H = \sum_{n=-\infty}^{+\infty} c_n \exp(ik_1 \alpha_n x - ik_1 \gamma_n y),
\]
Equation (2.37)

\[
\beta = \cos(\theta), \quad \alpha_n = \alpha + n \lambda_1 / d, \quad \beta_n = \sqrt{1 - \alpha_n^2}, \quad \gamma_n = \sqrt{\nu^2 - \alpha_n^2},
\]
Equation (2.38)

with \( b_n \) and \( c_n \) being complex coefficients called amplitudes of the plane waves scattered in the dielectric \((y > y_{\text{max}})\) and in the metal \((y < y_{\text{min}})\). The choice of the determination of the square roots contained in Eq. (2.38) will be the same as that given by Fig. 2.5 for \( \beta \) and \( \gamma \). It is important to notice that in general, the Rayleigh expansions given above cannot represent the field inside the grooves. For example, the Rayleigh expansion given by Eq. (2.36) cannot in general represent the field in the region located between the interface and \( y = y_{\text{max}} [29, 45] \). It must be remarked that the waves scattered in the dielectric region can be separated into two parts:

- a finite number of \( y \)-propagating waves corresponding to real values of \( \beta_n \),
- an infinite number of evanescent waves corresponding to imaginary values of \( \beta_n \).

Whatever the value of \( \lambda / d \) may be, there exists at least one \( y \)-propagating wave, corresponding to \( n = 0 \) (specularly reflected order).

Equations (2.36)–(2.38) permit us to understand why a plane wave illuminating the grating surface can excite a SPP. In contrast with a flat surface, a grating scatters waves having propagation constants in \( x \) equal to \( k_1 \alpha_n = k_1 \alpha + k_1 n \lambda_1 / d = k_1 (\alpha + 2\pi n / d) \). Thus, the propagation constants of the scattered waves can take values outside the range \((-k_1, +k_1)\) and the propagation constant of one of them can be close to the propagation constant in \( x \) of the SPP \( k_1 \tilde{\alpha}_{\text{plane}} = k_1 \nu / \sqrt{1 + \nu^2} \) (see Eq. (2.18)). It can be predicted that the maximum excitation of the SPP occurs when one order \( n_e \) satisfies the equation:

\[
\alpha_{n_e} = \pm \text{Re}\left\{ \tilde{\alpha}_{\text{plane}} \right\}.
\]
Equation (2.39)

Since \( \text{Re}\{ \tilde{\alpha}_{\text{plane}} \} \) is slightly greater than unity in modulus, the order \( n_e \) must be evanescent, but close to the passing-off. It is worth noting that Eq. (2.39) assumes that the propagation constant of the SPP on a grating is equal to that on a flat surface. We will see that this is not correct and thus this equation must be considered as an approximation. Let us illustrate the excitation of SPPs on the example given at the bottom of Fig. 2.14, which shows the order \(-1\) and total efficiencies of a holographic silver grating with a period of 1,205 nm illuminated by a p-polarized plane wave at a wavelength of 521 nm [21].

The use of Eq. (2.39) allows one to predict that a SPP is excited when
Fig. 2.14 Comparison between experimental data on a holographic silver grating (by courtesy of Hutley) and theoretical results for p-polarized light. The passing-off incidences are shown by arrows and the Littrow (Bragg) position by a vertical line. Solid line experimental data, dashed line theoretical results for the efficiency in the order $-1$ of a silver grating, dashed-dotted line theoretical results for the sum of efficiencies of scattered orders, dotted line theoretical results for the efficiency in the order $-1$ of a perfectly conducting grating

\[
\alpha_{n_e} = \sin (\theta) + n_e \lambda_1 / d = \pm \text{Re} \left\{ \tilde{\alpha}_{\text{plane}} \right\},
\]

which reduces, after tedious calculations, to
\[
\sin (|\theta|) = \left| \text{Re} \left\{ \tilde{\alpha}_{\text{plane}} \right\} - |n_e| \lambda_1 / d \right|.
\]

In that case, the index of silver is $\nu_2 = \nu = 0.052 + i 3.05$ and thus $\tilde{\alpha}_{\text{plane}} = 1.06 + i 2.2 \times 10^{-3}$. Consequently, the solutions are given by $|\theta| = 39^\circ$ for $|n_e| = 1$, $|\theta| = 11^\circ$ for $|n_e| = 2$, $|\theta| = 14^\circ$ for $|n_e| = 3$, $|\theta| = 42^\circ$ for $|n_e| = 4$. If we consider the bottoms of the drops of total efficiency caused by absorption in Fig. 2.14, we notice that the anomalies predicted for $|\theta| = 39^\circ$ and $|\theta| = 42^\circ$ cannot be separated, as well as those predicted for $|\theta| = 11^\circ$ and $|\theta| = 14^\circ$, the values measured in Fig. 2.14 being $\theta = \pm 14^\circ$ and $\theta = 42^\circ$, which are close to the values predicted by theory in the range of incidence ($-20^\circ, 43^\circ$) represented in the figure.

Let us notice finally that a significant amount of energy (more than 20%) can be absorbed in the anomalous regions in Fig. 2.14. The interpretation of the absorption phenomena is obvious: due to the resonant excitation of SPPs, the field inside the metal presents local enhancements which generate strong Joule effects. A study of absorption caused by SPPs can be found in [46].

**SPP on a Grating**

The study of SPPs on a grating can be achieved using the same lines as in the case of a flat surface. A SPP is a solution of a homogeneous problem, in which a scattered field exists without any incident field. From Eqs. (2.36) and (2.37), such a wave can
be written in the form:

\[
\begin{align*}
\text{if } y > y_{\text{max}}, & \quad H = \sum_{n=-\infty}^{+\infty} b_n \exp(i k_1 \alpha_n x + i k_1 \beta_n y), \\
\text{if } y < y_{\text{min}}, & \quad H = \sum_{n=-\infty}^{+\infty} c_n \exp(i k_1 \alpha_n x - i k_1 \gamma_n y),
\end{align*}
\]

(2.41) (2.42)

\[
\alpha_n = \alpha + n \lambda_1 / d, \quad \beta_n = \sqrt{1 - \alpha_n^2}, \quad \gamma_n = \sqrt{\nu^2 - \alpha_n^2},
\]

(2.43)

the choice of the determinations of \(\beta_n\) and \(\gamma_n\) being fixed by Fig. 2.5. Of course, the energy balance entails that the values of \(\alpha_n\) must be complex with a non-null imaginary part.

Thus, it turns out that the SPP can be represented by an infinity of plane waves having normalized propagation constants in \(x\) (constants of propagation divided by \(k_1\)) spaced by multiples of \(\lambda_1 / d\). It must be noticed that the numbering of the parameters in Eqs. (2.41) and (2.42) is ambiguous since changing \(n\) into \(n + p\) (\(p\) constant integer) does not modify the sum of the series. In order to fix this determination, we can bear in mind that one of the terms of the series must tend to the SPP of the plane when the height \(h\) of the grating tends to zero. For example, if the profile is sinusoidal, we can go continuously from the grating \(y = h \cos(2\pi x / d)\) to the plane by decreasing \(h\). By definition, the term of the series corresponding to the SPP of the flat surface is numbered by 0. Thus, denoting by \(\alpha_n^{\text{grating}}\) the normalized constants of propagation, it can be written that:

\[
\lim_{h \to 0} \left\{ \hat{\alpha}_0^{\text{grating}} \right\} = \hat{\alpha}_0^{\text{plane}}.
\]

(2.44)

Figure 2.15 shows the locations of the normalized constants of propagation \(\alpha_n^{\text{grating}}\).

The different components of the SPP represented in Fig. 2.15 are very different in nature. For \(n \notin (-1, -2)\), \(\tilde{\alpha}_n^{\text{grating}}\) has a real part greater than unity in modulus, with a small imaginary part. Figure 2.16 shows the real and imaginary parts \(\beta'\) and \(\beta''\) of \(\beta = \sqrt{1 - \alpha^2}\) versus the real part \(\alpha'\) of \(\alpha = \alpha' + i \alpha''\), for \(\alpha'' = 0\) and \(\alpha'' = 5 \cdot 10^{-2}\), the determination of \(\beta\) being given by \(\beta' + i \beta'' \geq 0\). Bearing in mind that the values
Fig. 2.16  Real and imaginary parts of $\beta = \beta' + i \beta'' = \sqrt{1 - \alpha'^2}$ versus the real part of $\alpha = \alpha' + i \alpha''$ with: a $\alpha'' = 0$, b $\alpha'' = 5.10^{-2}$

of $\tilde{\beta}_n^{\text{grating}}$ deduce from those of $\tilde{\alpha}_n^{\text{grating}}$ by $\tilde{\beta}_n^{\text{grating}} = \sqrt{1 - (\tilde{\alpha}_n^{\text{grating}})^2}$ with the same determination as for $\beta$, it turns out that $\tilde{\beta}_n^{\text{grating}}$ is close to the imaginary axis and thus, in the dielectric, this kind of wave is very close to an evanescent wave. In the following, this kind of wave will be called wave of evanescent type. Let us recall that the SPP of the flat surface was of evanescent type in the dielectric. The same remark applies to the metal of any order since the values of $\tilde{\gamma}_n^{\text{grating}}$ and $\tilde{\gamma}_{\text{plane}}$ are close to the imaginary axis (see Fig. 2.17), thus the field decreases exponentially in the metal. On the other hand, the waves corresponding to $n \in (-1, -2)$ have a real part smaller than unity in modulus. Since the imaginary part is very small, $\tilde{\beta}_n^{\text{grating}}$ is close to the real axis (Fig. 2.16) and the corresponding wave in the dielectric is close to a homogeneous plane wave propagating towards $y = \infty$. It will be called wave of $y$-propagating type. In that case, $\tilde{\gamma}_n^{\text{grating}}$ remains close to the imaginary axis for any value of $n$, thus the field decreases exponentially in the metal. These properties allow us to predict that the imaginary part of $\tilde{\alpha}_n^{\text{grating}}$ is larger than that of $\tilde{\alpha}_{\text{plane}}$ since the waves of $y$-propagating type generate a supplementary loss of energy. Furthermore, this remark explains why a surface wave can propagate on a perfectly conducting grating. We have seen that the limit of the SPP of a flat perfectly conducting surface when the permittivity tends to $-\infty$ is a plane wave with wave vector parallel to the $x$-axis. The field of such a wave does not decrease on the dielectric side and thus, this wave cannot be considered as a surface wave. If a periodic modulation is introduced and if there exists at least one $y$-propagating wave in the dielectric region, this propagating wave will transfer the energy at infinity and so, a loss occurs in the propagation. Due to this loss, the value of $\tilde{\alpha}_n^{\text{grating}}$ is no more real and unitary, like $\tilde{\alpha}_{\text{plane}}$, but complex with a positive imaginary part.

Furthermore, it can be conjectured that the real part of $\tilde{\alpha}_0^{\text{grating}}$ is greater than unity. A heuristic way to explain this property is to notice that a modulation entails an increase of the interface length between two points of the surface, thus reduces
Fig. 2.17 Real and imaginary parts of $\gamma = \gamma' + i\gamma'' = \sqrt{\nu^2 - \alpha^2}$ versus the real part of $\alpha = \alpha' + i\alpha''$ for $\nu = 1.019 + i7.1$ (index of aluminium at 647 nm), with: a $\alpha'' = 0$, b $\alpha'' = 5.10^{-2}$

the projection on the $x$-axis of the propagation speed of the SPP. So, the propagation constant, which is inversely proportional to the speed, is increased. Let us notice in addition that, in the opposite case, it would be possible to excite the SPP with a plane wave. It is straightforward to show that the value of $\tilde{\beta}_{0 \text{grating}}$ corresponding to this value of $\tilde{\alpha}_{0 \text{grating}}$ has a positive imaginary part and thus the field decreases in the dielectric.

2.3 Phenomenological Study of Wood Anomalies

Until now, we have analysed the scattering and guiding properties of gratings. The aim of this section is to show from the theory of analytic functions of the complex variable that the guiding properties, i.e. the possibility of SPP propagation, have strong consequences on the scattering properties, i.e. on the amplitudes of the waves generated by the grating when it is illuminated by a plane wave. These consequences will be evaluated quantitatively.

2.3.1 Pole of the Reflection and Transmission Coefficients

The SPP of a grating is obtained by setting the scattering coefficients $b_n$ and $c_n$ equal to infinity in Eqs. (2.36) and (2.37) in order to make the incident field negligible. Let us show that all the coefficients $b_n$ and $c_n$, considered as functions of the normalized propagation constant $\alpha$ of the incident wave, have an infinity of poles located at points given by:

$$\alpha = \tilde{\alpha}_{n \text{grating}}.$$ (2.45)
If the incident wave satisfies this equation, the normalized propagation constants \( \alpha_n \) of the field scattered by the grating identify with the normalized propagation constants \( \tilde{\alpha}_{\text{grating}}^n \) of the SPP, thus the field expressed by Eqs. (2.41) and (2.42) is the solution of the scattering problem, which shows that the amplitudes of the scattered waves are infinite with respect to the amplitude of the incident wave. This remark leads us to guess that the amplitudes of the scattered waves have a pole [47] when a SPP can propagate.

It seems that this remark is useless since \( \alpha = \sin(\theta) \) is real, while the \( \tilde{\alpha}_{\text{grating}}^n \) are complex (Fig. 2.15). However, if the real part of \( \tilde{\alpha}_{\text{grating}}^n \) is less than unity in modulus (like \( \tilde{\alpha}_{-1}^\text{grating} \) and \( \tilde{\alpha}_{-2}^\text{grating} \) in Fig. 2.15), \( \alpha \) can be close to \( \tilde{\alpha}_{\text{grating}}^n \), and it can be conjectured that a resonance phenomenon will occur, provided that the imaginary part of the \( \tilde{\alpha}_{\text{grating}}^n \) is not too large. In other words, it can be considered that in real life, \( \alpha \) is real and the amplitudes \( b_n(\alpha) \) and \( c_n(\alpha) \) are complex functions of the real variable \( \alpha \). However, mathematical theorems [47] state that such a function has one and only one analytic continuation in the complex plane of \( \alpha \). In conclusion, all the \( \tilde{\alpha}_{\text{grating}}^n \) are poles of this continuation but using an actual plane wave, only some of them (\( \tilde{\alpha}_{-1}^\text{grating} \) and \( \tilde{\alpha}_{-2}^\text{grating} \) in Fig. 2.15) can be approached.

### 2.3.2 Zero of the Reflection Coefficient, Phenomenological Formula

Although the phenomenological approach can be generalized to more complicated cases, we now consider for simplicity that only one value of \( \tilde{\alpha}_{\text{grating}}^n \) has a real part less than unity in modulus. Bearing in mind that \( \tilde{\alpha}_{0}^\text{grating} \) is very close to unity for moderate values of the height \( h \) of the grating, Fig. 2.15 shows that \( n \) must be equal to \(-1\) and that the real part of \( \tilde{\alpha}_{-1}^\text{grating} \) must be negative. In these conditions, a resonance occurs when \( \alpha \approx \Re \{ \tilde{\alpha}_{-1}^\text{grating} \} \). In consequence, there exists only one non-evanescent reflected order since \( \alpha + n\lambda_1/d \approx \Re \{ \tilde{\alpha}_{-1}^\text{grating} + n\lambda_1/d \} = \Re \{ \tilde{\alpha}_{n}^\text{grating} \} \) and by hypothesis, all the values of \( \Re \{ \tilde{\alpha}_{n}^\text{grating} \} \) are greater than unity in modulus, except \( \Re \{ \tilde{\alpha}_{-1}^\text{grating} \} \). Thus we are led to the study of \( b_0(\alpha) \), the amplitude of the reflected order 0. Mathematically, since \( \tilde{\alpha}_{-1}^\text{grating} \) is a pole of the analytical continuation of \( b_0(\alpha) \), it will be called \( \alpha^p \) in the following. Thus, it can be deduced that, when \( \alpha \) is close to \( \Re \{ \alpha^p \} \), \( b_0(\alpha) \) can be expanded in a Laurent series [48], which can be written in the form:

\[
b_0(\alpha) \approx \frac{g_{-1}}{\alpha - \alpha^p} + g_0 + (\alpha - \alpha^p) u(\alpha),
\]

(2.46)
with $g_{-1}$ and $g_0$ being complex coefficients, and $u(\alpha)$ an entire series of $\alpha - \alpha_p$.

Neglecting the last term in the right-hand member of Eq. (2.46) yields:

$$b_0(\alpha) \simeq \frac{g_{-1}}{\alpha - \alpha_p} + g_0 \simeq \frac{g_{-1} + g_0 (\alpha - \alpha_p)}{\alpha - \alpha_p}.$$  \hspace{1cm} (2.47)

Setting

$$\alpha^z = \alpha_p - \frac{g_{-1}}{g_0},$$  \hspace{1cm} (2.48)

Eq. (2.47) can be written in the form:

$$b_0(\alpha) \simeq g_0 \frac{\alpha - \alpha^z}{\alpha - \alpha_p}.$$  \hspace{1cm} (2.49)

If the height of the grating tends to 0, there is no more resonance and then, $g_{-1}$ tends to 0, $g_0$ tends to the reflection coefficient $r$ of a flat metallic plane, and since $\tilde{\alpha}_0^{\text{grating}}$ tends to $\tilde{\alpha}_0^{\text{plane}}$:

$$\lim_{h \to 0} \{\alpha^z\} = \lim_{h \to 0} \{\alpha_p\} = \tilde{\alpha}_0^{\text{plane}} - \frac{\lambda_1}{d}.$$  \hspace{1cm} (2.50)

As a result, the effects of the pole and zero annihilate when $h$ tends to zero, according to Eq. (2.49).

### 2.3.3 Verification of the Phenomenological Formula from Numerical Results

Now, let us show that the phenomenological formula allows one to predict with great precision the efficiency of a metallic grating in the region of anomaly. We consider a gold sinusoidal grating with period 555 nm and height 60 nm, illuminated by a p-polarized plane wave of wavelength 647 nm propagating in vacuum. We compare the efficiencies obtained from a computer code based on a rigorous integral theory of gratings [21, 29] and from the phenomenological formula. This formula requires the knowledge of the parameters $g_0$, $\alpha^z$ and $\alpha_p$ contained in Eq. (2.49). We have given to $g_0$ its limit value as the height of the grating tends to 0, i.e. the reflection coefficient of a gold plane. In order to calculate $\alpha^z$ and $\alpha_p$, we have transformed the computer code based on a rigorous integral theory of gratings [11, 12]. First, this code has been extended to complex values of $\alpha = \sin(\theta)$ in order to calculate the output $b_0$ from the input $\alpha$. Then, this code has been used as a subroutine of a software able to find the zero of a complex function of a complex variable. This kind of software, based on very simple formulae like Newton’s formula, requires an estimate of the location of the zero in order to initiate the iterative process. This estimate has been deduced from Eq. (2.50). Thus we have searched for the zero of $1/b_0$ for $\alpha_p$ and of $b_0$ for
\( \alpha^z \). The results are \( \alpha^p = -0.1135 + i0.0149 \) and \( \alpha^z = -0.1126 - i0.00596 \). We have introduced these parameters in a formula deduced from Eq. (2.49) by taking the square modulus \( |b_0|^2 \) of the amplitude in the order 0 (i.e. the efficiency in the order 0), the coefficient \( |g_0|^2 \) being replaced by the reflectivity \( R = |r|^2 \) of gold, equal to 0.953:

\[
|b_0(\alpha)|^2 \approx R \left| \frac{\alpha - \alpha^z}{\alpha - \alpha^p} \right|^2.
\]

Figure 2.18 shows a comparison between the efficiencies in the order 0 deduced from this phenomenological formula and from the rigorous integral theory versus the angle of incidence.

It must be noticed that from our data, the resonance should occur for negative incidences, but of course, the symmetry of the grating profile entails that the same resonance is obtained for positive incidences. The two curves are almost identical, except in the left side of the curve where discrepancies of the order of 3\% appear. This is not surprising since a second resonance occurs for negative angles and of course, the effects of the corresponding pole and zero are not taken into account in the phenomenological formula. The obvious conclusion is that the phenomenological formula is a valuable tool for predicting the efficiency of gratings in the resonance region. It reduces the resonance phenomenon to the knowledge of two complex parameters.

Some elementary properties of the resonance curve can be easily deduced from this formula. Assuming that the real parts of the pole and zero are very close (which is in general the case), the minimum value \( e_m \) of the efficiency and the width \( w \) at half-height of the drop of efficiency are given by:

\[
e_m = R \times \left| \frac{\text{Im}\{\alpha^z\}}{\text{Im}\{\alpha^p\}} \right|^2, \tag{2.52}
\]

\[
w = 2\text{Im}\{\alpha^p\}. \tag{2.53}
\]
Furthermore, it can be noticed from the phenomenological formula that the phase of \( b_0 \) has violent variations at the vicinity of the resonance since the phase of both terms \( \alpha - \alpha^z \) and \( \alpha - \alpha^p \) in Eq. (2.49) is rapidly varying. If the pole and the zero are placed on the same side of the real axis, the total phase shift from one side of the resonance process to the other one is equal to zero, while in the opposite case, it reaches \( \pm 2\pi \).

### 2.4 Total Absorption of Light by a Diffraction Grating

#### 2.4.1 Theoretical Demonstration

The existence of a total absorption of light by a grating, which cannot scatter more than a single \( y \)-propagating wave, can be demonstrated from many arguments. First, it can be shown from numerical results. Indeed, we have seen that for a flat surface the pole and the zero are located on the same side of the real axis, with a positive imaginary part. We know that the sign of the imaginary part of the pole holds, whatever the groove depth. Thus, if numerical calculations show that for a given groove depth, the zero has a negative imaginary part, it can be deduced from a topological argument that there exists a groove depth for which the zero crosses the real axis. Since it is so given in Fig. 2.18 (drawn using \( \alpha^z = -0.1126 - i0.00596 \)) for a gold sinusoidal grating, the total absorption must occur for a groove depth smaller than that of the grating of Fig. 2.18, as it will be seen in the following paragraph.

The existence of total absorption phenomenon can also be shown through a purely theoretical demonstration. It has been shown in Sect. 2.3.2 that for a perfectly conducting structure, the pole and zero at a given wavelength are complex conjugate, whatever the groove depth may be. Thus the zero is located below the real axis. On the other hand, for a flat, lossy metallic structure, the pole and zero are identical and located above the real axis. The perfect conductivity corresponds to a permittivity \( \varepsilon_2 \) of the metal which is real, negative and infinite in modulus. Let us suppose that for a given groove depth of a sinusoidal grating, the permittivity of the metal is varied continuously from this negative and infinite value to the permittivity of an actual metal like gold. Then, let us tend the groove depth of the grating to 0. Since the zero goes from a point located below the real axis to a point located above the real axis, it must cross the real axis, at least if we assume the continuity of its trajectory. In fact, from several numerical analyses, it turns out that for a given shape (sinusoidal, triangular...) and a given period of the grating profile, with given metal and the angle of incidence, a phenomenon of total absorption occurs for p-polarized light for given wavelengths and groove depths, these two parameters depending on each other.

In Fig. 2.19, we have drawn the trajectory of the pole and zero of a sinusoidal gold grating when the groove depth is varied. The zero crosses the real axis for \( \alpha \) close to
Fig. 2.19 Trajectory of the pole $\alpha^p$ and zero $\alpha^z$ of $b_0$ for a gold sinusoidal grating with period 555 nm, illuminated by a p-polarized plane wave with wavelength 647 nm, when the height $h$ is increased. Reprinted from [50], with permission from Elsevier

$-0.115 (\theta \simeq -6.6^\circ)$ for a groove depth equal to 40 nm, thus a total absorption must occur for these parameters.

2.4.2 Experimental Verification

The experimental verification of the phenomenon of total absorption discovered in [49] from theoretical results was given in [50]. It was performed on a holographic sinusoidal gold grating having the same pitch as in Fig. 2.19 illuminated by a krypton laser beam at the wavelength 647 nm. Due to a mishap during exposure, the groove depth varied from 13 nm on one side to 75 nm on the other. In order to study the reflectance as a function of the groove depth, it was necessary to simply select the appropriate region of the grating surface. The groove depth was measured using a profilometer fitted with a chisel-shaped stylus. The experimental results are shown in Fig. 2.20b, as well as the corresponding theoretical predictions obtained from the integral theory in Fig. 2.20a. The agreement between experimental and theoretical results is excellent. The minimum recorded reflectance of 0.3% was at an angle of incidence of 6.6° and a groove depth of 37 nm, which are very close to the theoretical predictions. In the same paper, the authors obtained a total absorption for other wavelengths and were able, using a modified Mach-Zender interferometer, to show the change of $\pi$ in the phase of the reflected wave as the incident beam crossed the region of the grating where $\alpha^z$ becomes real.
This result is quite remarkable and very surprising: a very gentle modulation in the surface of a gold mirror causes the reflectance to fall dramatically from over 90% to below 1%. Figure 2.20 also shows that the width of the reflectance drop is very small, of the order of 1°. The resonance remains very selective when other parameters (wavelength, groove depth) are changed. Similar results were published by Le Perchec et al. [51] for nanometric silver lamellar gratings.
2.4.3 Some Applications

Figure 2.20 shows that the total absorption phenomenon can present a strong angular selectivity, typically less than 1°. The same selectivity can be found when, starting from total absorption, the wavelength is varied. Figure 2.21, reprinted from [50], shows the spectrum of the grating which corresponds to a total absorption in Fig. 2.20, illuminated with a collimated beam of white light from a tungsten lamp. The spectrum clearly shows a strong narrow absorption band in the red. Here, the grating absorbs in totality at 647 nm, but it is interesting to know that a nearly total absorption phenomenon holds for another wavelength when the angle of incidence is varied. Obviously, such a grating can be used as a rejection filter for eliminating a wavelength in a polychromatic light beam. A very selective absorption in incidence and frequency ranges can be used in metrology or to make selective filters like biosensors. It could prevent cross-talks between optical interconnects.

Let us describe the application to immunosensors. A direct immunosensor is an immunologically sensitized transducer that possesses the ability to observe antibody–antigen binding events in real time. An immunosensor can be made using the sensitivity of SPPs to changes in the dielectric permittivity of a dielectric-coated metallic grating [52]. It consisted of a gold or silver holographic diffraction grating on which ‘sensitizing’ immunological molecules were immobilized, realizing in some way a dielectric coating. The subsequent binding of complementary components (contained for example in human or animal serums) can be followed, in real time, by measuring changes in the reflectivity of the grating resulting from alterations in the conditions necessary for optimal SPP excitation.

Since Wood’s anomalies entail strong local enhancements of the field on the grating surface, they are also used in Raman scattering [53] or second-harmonic generation [54].

On the other hand, it is interesting to realize a strong absorption in the wide ranges of angle of incidence or wavelength. Equation (2.53), deduced from the phenomenological formula, shows that the angular width of the absorption peak increases with the imaginary part of the pole. Furthermore, Fig. 2.19 clearly shows that this imaginary part strongly increases with the groove depth. As a consequence, it can be deduced that a wide-range strong absorption can be obtained with deep gratings. The SPPs of such gratings are sometimes termed “localized SPPs” since they are

![Fig. 2.21 Spectrum of a white light beam after reflection from a gold grating of Fig. 2.20 with groove height 37 nm with incidence 6.6°. Reprinted [50], with permission from Elsevier](image)
very rapidly attenuated when they propagate. A nearly total absorption on a wide
range of angles of incidence or wavelengths could lead to many practical appli-
cations, for example the realization of solar absorbers [55] or, as suggested by Teperik
et al. [56], efficient photovoltaic cells, or light shielding of micro-photonic devices.
As suggested by the Kirchhoff law, this kind of grating should constitute omni-
directional black-body emitters, possibly with the narrow a spectral range. The interested
reader can find other possibilities to use SPPs in [57–73].

2.5 Further Properties of Surface Plasmon Polaritons

2.5.1 Physical Interpretation and Fundamental Properties of the
Zero of the Reflection Coefficient

It is possible to give a physical interpretation of the existence of a zero close to the
pole. With this aim, let us consider the complex conjugate $H^*$ of the expression of the
SPP given by Eqs. (2.41) and (2.42). Such a field satisfies in the metal the Helmholtz
equation

$$\nabla^2 H^* + k^2 \varepsilon_2^* H^* = 0,$$

(2.54)

while in the dielectric, the permittivity is real and thus the Helmholtz equation
remains unchanged. In the following, we will call the structure so obtained as
the adjoint structure. It is straightforward to show that $H^*$ satisfies the boundary
conditions on the interface and thus $H^*$ is a solution of the elementary laws of
electromagnetics in a structure made of a dielectric of permittivity $\varepsilon_1$ and a mate-
rial of permittivity $\varepsilon_2^*$. This permittivity corresponds to a material with gain [74].
However, $H^*$ completely differs from a SPP since some waves included in
$H^*$ do not satisfy in the dielectric the radiation condition expressed by Eq. (2.16). Indeed,
the complex conjugation changes the expression $\exp \left( i k_1 \tilde{\alpha}_n^{\text{grating}} + i k_1 \tilde{\beta}_n^{\text{grating}} \right)$ into

$$\exp \left( -i k_1 \left( \tilde{\alpha}_n^{\text{grating}} \right)^* - i k_1 \left( \tilde{\beta}_n^{\text{grating}} \right)^* \right),$$

which shows that $\tilde{\alpha}_n^{\text{grating}}$ becomes $-\left( \tilde{\alpha}_n^{\text{grating}} \right)^*$ and $\tilde{\beta}_n^{\text{grating}}$ becomes $-\left( \tilde{\beta}_n^{\text{grating}} \right)^*$. As a consequence, the real parts
of the propagation constants $-\left( \tilde{\alpha}_n^{\text{grating}} \right)^*$ of $H^*$ along the $x$-axis become the oppo-
site of those of $\tilde{\alpha}_n^{\text{grating}}$, while the imaginary part holds. This remark entails that the
direction of propagation of $H^*$ on the adjoint structure and of the SPP in the initial
structure is opposite and that the amplitude is increased in the direction of propaga-
tion, a fact which is not surprising since the metal has been replaced by a material
with gain. A simple way to restore a propagation towards $x = +\infty$ is to change $x$
into $-x$ in the expression of $H^*$, which is the equivalent of changing $-\left( \tilde{\alpha}_n^{\text{grating}} \right)^*$
into $+ \left( \tilde{\alpha}_n^{\text{grating}} \right)^*$. After this symmetry, the propagation constant along $x$ of $H^*$ and of the SPP of the initial structure are complex conjugate.

As regards the radiation condition on the dielectric side, two cases must be distinguished. If the $n$th component of the SPP is of the evanescent type, $\tilde{\beta}_n^{\text{grating}}$ is very close to the imaginary axis, with a positive imaginary part. This positive imaginary part holds for $-\left( \tilde{\beta}_n^{\text{grating}} \right)^*$ and thus the wave remains of the evanescent type and satisfies the radiation condition. On the other hand, for the waves of $y$-propagating type, $\tilde{\beta}_n^{\text{grating}}$ and $-\left( \tilde{\beta}_n^{\text{grating}} \right)^*$ are very close to the real axis and their real parts are opposite. Consequently, for the adjoint structure, the waves contained in $H^*$ do not propagate towards $y = +\infty$ anymore, but towards $y = -\infty$. Such a wave corresponds to an incident wave that propagates in the direction of the grating surface. In the metal, since the values of $\tilde{\gamma}_n^{\text{grating}}$ are close to the imaginary axis, it is easy to show that the series of waves inside $H^*$ still satisfy the radiation condition.

In conclusion, after a symmetry in $x$, $H^*$ has a propagation constant along $x$ which is the conjugate of that of the SPP on the initial structure. Moreover, all the waves included in $H^*$ still satisfy the radiation condition, except those corresponding to $y$-propagating waves in the dielectric, which become incident instead of scattered. Assuming again that the only wave of propagating type in the dielectric is the order $n = -1$, $H^*$ can be considered, after a symmetry along $x$, as the field scattered by an incident plane wave with propagation constants $\left( \tilde{\alpha}_{-1}^{\text{grating}} \right)^*$ and $-\left( \tilde{\beta}_{-1}^{\text{grating}} \right)^*$ along the $x$- and $y$-axes. Furthermore, there is no scattered wave corresponding to $n = -1$, i.e. in the direction $\left( \tilde{\alpha}_{-1}^{\text{grating}} \right)^*$, $+ \left( \tilde{\beta}_{-1}^{\text{grating}} \right)^*$. It can be deduced that after symmetry with respect to the $x$-axis, $H^*$ is the field corresponding to the zero of the adjoint structure, which is thus the conjugate of the pole of the initial structure. Following the same lines, it is easy to show that the zero of the initial structure is, after symmetry along $x$, the conjugate of the pole of the adjoint structure, i.e. the conjugate of the propagation constant of the SPP of the adjoint structure.

This property shows that there exists complete symmetry with respect to the real axis between the pole and zero of the initial structure and the pole and zero of the adjoint structure. Vital properties of the poles and zeros can be deduced. Let us notice first that the zero is an actual theoretical zero. It was not possible to state this important property from Eq. (2.47), which was obtained by neglecting the last term in Eq. (2.46). Furthermore, we have seen that it can be conjectured that the imaginary part of the pole of the initial structure should increase with the height of the grating, due to losses through propagating waves in the dielectric. For the adjoint structure, there is an exponential increase of the amplitude in the propagation. The propagating waves of the SPP of the adjoint structure will provoke the opposite effect. Since they entail losses of energy, the increase of the amplitude will be attenuated and may be cancelled. In that case, the propagating wave in the dielectric is an actual plane wave with real constants of propagation in $x$ and $y$, which propagates towards $y = +\infty$: the grating, which is not illuminated by any incident wave, scatters this plane wave to infinity. Conversely, the zero of the initial structure is real: the grating, illuminated
by a plane wave, does not scatter any plane wave. This is another way to theoretically predict the phenomenon of total absorption studied in the preceding section. Finally, for a perfectly conducting structure, \( \varepsilon_2 \) is real, negative and infinite in modulus. In that case the initial and adjoint structures identify and so, the pole and zero of the initial structure are complex conjugate, as well as those of the adjoint structure.

### 2.5.2 Analogy with Guided Waves in Dielectric Films Deposited on Metallic Surfaces

It has been shown that the basic origin of the grating anomalies must be found in the excitation of a surface wave: the SPP. The consequence is that this kind of anomaly occurs only for p-polarized light. However, it should be noticed that other kinds of surface waves can propagate on a flat metallic surface, provided that the metal is covered by a plane dielectric coating, as shown in Fig. 2.22. The wave propagates in the dielectric film following a zigzag path, with a lossy metallic reflection on one side and a lossless total reflection on the other side, provided that the angle of incidence on the air exceeds the critical angle of total reflection. In that case, the field on the air side reduces to an evanescent wave and the propagation constant of the guided wave is greater than the wavenumber of the light in vacuum, like for a SPP. Thus this guided wave can play the same role as a SPP for generating a resonance phenomenon as soon as one of the interfaces (or both) is periodically modulated. The big difference with SPPs is that anomalies can occur for both polarizations.

Let us show an example of anomaly of a metallic, dielectric-coated grating for s-polarized light, taken from [28]. Figure 2.23 shows a scanning electron micrograph of the grating profile of a 1,264 lines/mm ruled aluminium grating (with triangular
Fig. 2.23 Electron micrograph of the profile of a ruled grating made in NPL. Reprinted by permission from IOP Publishing Ltd: [28] p. 90

groove) made in the NPL (National Physical Laboratory, Teddington, U.K.). The blaze angle of the grating (angle between the large facet and the mean plane of the profile) was equal to 23° and the included angle of the ruling diamond (angle between the two facets) was 110°. Two identical aluminium gratings were realized and one of them was covered with silicon monoxide with a thickness close to 100 nm (measured using multiple-beam interferometry). The efficiency in the order $-1$ was measured in Littrow mounting for both gratings. Let us recall that in the Littrow mount, the order $-1$ and the incident wave propagate in opposite directions. Using Eq. (2.1) with $n = -1$, it turns out that since $\theta_{-1} = -\theta$, the angle of incidence and the wavelength in vacuum satisfy the relation

$$\frac{\lambda}{d} = 2 \sin (\theta).$$

(2.55)

The efficiency curves are shown in Fig. 2.24. Experimental measurements were realized in the NPL, while the theoretical results were obtained from the differential theory [28, 29]. For the uncoated grating, experimental and theoretical curves show a relatively featureless shape, the edge corresponding to a Rayleigh anomaly (passing-off of the orders $-2$ and $+1$). Both theory and experiment show that the coating introduces just after the Rayleigh anomaly a strong and sharp minimum, even though the theory gives an anomaly somewhat weaker than was observed. This discrepancy has been explained: the profile used in the calculations is perfectly triangular, which is not the case in Fig. 2.23, and calculations have shown that the strength and the width of the anomaly are very sensitive to small changes in profile form and coating thickness.
A phenomenon of total absorption has been verified for s-polarized light with a 2,400 lines/mm $10^\circ 22' \text{ blaze angle}$, aluminium grating covered with a 130 nm magnesium fluoride coating [75]. Total absorption was predicted by theory for a coating thickness of 152 nm, a wavelength of 492 nm and an angle of incidence equal to $7.5^\circ$. The measurements were performed with a 130 nm depth of magnesium fluoride, for a wavelength equal to 492 nm and at an angle of incidence close to $7^\circ$, the efficiency was found to drop sharply to just 5.5 %. Other results on total absorption for s-polarized light can be found in [76]. The total absorption can be generalized to unpolarized light using crossed gratings [77, 78] or even using classical 2D gratings in conical (off-plane) mountings [79].

These examples show that guided waves generated by dielectric coatings on top of the metal can generate the same resonance phenomena as SPPs. This remark could be generalized to other domains of plasmonics.
2.6 Conclusion

The study of Wood’s anomalies and SPP resonances, initiated more than one century ago, is a typical example of the interest of converging analyses of experimentalists, theoreticians and specialists of phenomenology.

It is amazing to bear in mind that, initially, grating anomalies were considered as strong defects of gratings. The reason is that the gratings were mainly used in spectroscopy, where efficiency drops caused by anomalies around certain wavelengths considerably reduce the possibility of analysing a light spectrum in a wide range of wavelengths with good precision. Nowadays, non-spectroscopic applications of gratings have become a crucial field of optics and the practical applications of SPP anomalies have taken a strong importance in many fields of science and technology.

These modern applications of Wood’s anomalies have been permitted by the progress of process of fabrication and of numerical tools, which allow one to optimize almost all kinds of diffraction gratings in all ranges of wavelength. It can be conjectured that the extraordinary features of the resonant excitation of SPPs will encourage the specialists of modern optics and nanophotonics to develop and to extend the field of applications of grating anomalies in the future.

Appendix 1: Electromagnetic modelling in Two Dimensions

We consider an interface $\Sigma$ separating two homogeneous regions $R_1$ and $R_2$. The Maxwell equations write:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t},$$  \hspace{1cm} (2.56)

$$\nabla \times \mathbf{H} = \mathbf{j}_c + \frac{\partial \mathbf{D}}{\partial t},$$  \hspace{1cm} (2.57)

where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic fields, $\mathbf{D}$ and $\mathbf{B}$ are the electric and magnetic inductions, and $\mathbf{j}_c$ is the conduction current density. In the harmonic regime, the fields and currents have a sinusoidal behaviour in time. A function of space and time $f(\mathbf{r}, t)$ with sinusoidal behaviour in time can be written as:

$$f(\mathbf{r}, t) = a(\mathbf{r}) \cos (\omega t - \varphi(\mathbf{r})), \hspace{1cm} (2.58)$$

with $a$ and $\varphi$ being real functions of space called amplitude and phase, $\omega$ being the frequency. Such a function is classically represented by its complex amplitude $\hat{f}(\mathbf{r})$ independent of time, and Eq. (2.58) is re-written in the form:

$$f(\mathbf{r}, t) = \text{Re} \{a(\mathbf{r}) \exp(-i\omega t + i\varphi(\mathbf{r}))\} = \text{Re} \{a(\mathbf{r}) \exp(i\varphi(\mathbf{r})) \exp(-i\omega t)\}. \hspace{1cm} (2.59)$$

The expression $a(\mathbf{r}) \exp(i\varphi(\mathbf{r}))$ being called complex amplitude of $f$ and denoted by $\hat{f}$, Eq. (2.59) can be written:
Thus the function \( f \) deduces from its complex amplitude \( \tilde{f} \) by multiplying by \( \exp(-i\omega t) \) then by taking the real part of the product. It is straightforward to show that the multiplication of \( f(\mathbf{r}, t) \) by a real function \( u(\mathbf{r}) \) results in a multiplication of the complex amplitude by the same function and conversely, and that the complex amplitude of \( \partial f/\partial t \) is equal to \(-i\omega \tilde{f}\). Thus, Maxwell’s equations can be written using the complex amplitudes of the field and, for simplicity, the complex amplitudes of the fields and current are denoted using the same names and symbols as the fields and current themselves, in such a way that harmonic Maxwell equations can be written:

\[
\nabla \times \mathbf{E} = i\omega \mathbf{B},
\]

\[
\nabla \times \mathbf{H} = j_{c} - i\omega \mathbf{D}.
\]

In addition to Maxwell equations, constitutive relations allow one to express the electromagnetic properties of the materials. In contrast with Maxwell equations, they are not rigorous (except in vacuum). Assuming that a material is non-magnetic, homogeneous, isotropic and linear, these relations can be written:

\[
\mathbf{B} = \mu_{0} \mathbf{H}, \quad \mathbf{D} = \varepsilon_{0} \varepsilon \mathbf{E}, \quad j_{c} = \sigma \mathbf{E},
\]

with \( \mu_{0} = 4 \cdot \pi \cdot 10^{-7} \) being the permeability of vacuum and \( \varepsilon_{0} = 1/(36\pi 10^{9}) \) being the permittivity of vacuum. The parameters \( \varepsilon' \) and \( \sigma \) denote the relative dielectric permittivity and the conductivity of the material respectively. Using the relations \( \mathbf{D} = \varepsilon_{0} \mathbf{E} + \mathbf{P}, \quad \mathbf{P} = \varepsilon_{0} \chi \mathbf{E} \), with \( \mathbf{P} \) electric polarization density and \( \chi \) electric susceptibility, it turns out that \( \varepsilon' = 1 + \chi \). Introducing the constitutive relations in Maxwell equations yields:

\[
\nabla \times \mathbf{E} = i\omega \mu_{0} \mathbf{H},
\]

\[
\nabla \times \mathbf{H} = (\sigma - i\omega \varepsilon_{0} \varepsilon') \mathbf{E}.
\]

Defining the complex permittivity:

\[
\varepsilon = \varepsilon' + i\sigma/(\omega \varepsilon_{0}) = 1 + \chi + i\sigma/(\omega \varepsilon_{0}) .
\]

Equation (2.65) takes a form symmetrical to Eq. (2.64):

\[
\nabla \times \mathbf{H} = -i\omega \varepsilon_{0} \varepsilon \mathbf{E}.
\]

This equation can be expressed in the form:

\[
\nabla \times \mathbf{H} = -i\omega \varepsilon_{0} \mathbf{E} + j_{t}, \text{ with } j_{t} = \sigma \mathbf{E} - i\omega \varepsilon_{0} \chi \mathbf{E} = i\omega \varepsilon_{0} (1 - \varepsilon) \mathbf{E} = j_{c} + j_{b}.
\]
with \( \mathbf{j}_h \) total current density, including both the conduction current density \( \mathbf{j}_c \) and the bound current density \( \mathbf{j}_b = -i\omega \mathbf{P} = -i\omega\varepsilon_0\chi \mathbf{E} \) resulting from the electric polarization. The optical index of a material is given by \( \nu = \sqrt{\varepsilon} \).

Let us notice that by taking the divergence of Eqs. (2.64) and (2.67) and using \( \nabla \cdot (\nabla \times \mathbf{V}) = 0 \), one can get the complementary couple of Maxwell equations in harmonic regime:

\[
\nabla \cdot \mathbf{H} = 0, \quad \nabla \cdot (\varepsilon \mathbf{E}) = \nabla \cdot (\varepsilon' \mathbf{E}) = 0.
\]

By combining Eqs. (2.64) and (2.67), one can obtain partial derivative equations for each field inside a homogeneous region. Introducing the value of \( \mathbf{H} \) given by Eq. (2.64) in Eq. (2.67), we obtain:

\[
\nabla \times \nabla \times \mathbf{E} - k^2 \mathbf{E} = 0, \quad \text{with} \quad k = \omega \sqrt{\varepsilon_0\mu_0}.
\]

In a homogeneous region, \( \nabla \cdot (\varepsilon \mathbf{E}) = 0 \) entails that \( \nabla \cdot \mathbf{E} = 0 \) then using Eq. (2.69) and the vector relationship \( \nabla \times \nabla \times \mathbf{E} = \nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} \) we get:

\[
\nabla \mathbf{E} + k^2 \mathbf{E} = 0,
\]

and, following the same lines for the magnetic field:

\[
\nabla \mathbf{H} + k^2 \mathbf{H} = 0.
\]

It is worth noting that Maxwell equations (Eqs. (2.64) and (2.67)) are valid in the sense of distributions. In other words, they include the boundary conditions at the limit between two homogeneous materials. In order to express them in an explicit form, one can recall that the surface density of \( \mathbf{V} \) included in \( \Delta \times \mathbf{V} \) is equal to \( \mathbf{n} \times (\mathbf{V}_+ - \mathbf{V}_-) \), with \( \mathbf{V}_+ - \mathbf{V}_- \) being the jump of \( \mathbf{V} \) across the interface in the direction of \( \mathbf{n} \) [29]. It must be recalled that this surface term is the coefficient of a Delta distribution located on the surface. In Eqs. (2.64) and (2.67), the right-hand member contains the fields and thus they should not include distributive surface parts, thus the left-hand member satisfies the same property. We deduce that the tangential components of the electric and magnetic fields are continuous across an interface.

In the two-dimensional case, the interface is invariant by translation along the \( z \)-axis. One can distinguish the two fundamental cases of polarization: s-polarization with the electric field \( \mathbf{E} = E\mathbf{\hat{z}} \) and p-polarization with the magnetic field \( \mathbf{H} = H\mathbf{\hat{z}} \). In both cases, the boundary-value problem becomes scalar. Projecting Eqs. (2.71) and (2.72) on the \( z \)-axis, we obtain:

\[
\nabla^2 E + k^2 E = 0 \quad \text{for s-polarized light and} \quad \nabla^2 H + k^2 H = 0 \quad \text{for p-polarized light}.
\]

Thus the electric and magnetic fields satisfy a scalar Helmholtz equation.

As regards the boundary conditions on the interface, the continuity of the tangential component of the fields entails the continuity of \( E \) for s-polarization and the continuity of \( H \) for p-polarization. Since the partial derivative equation is of
the second order, a second boundary condition is needed. For s-polarization, one can express the continuity of the tangential component of the magnetic field using Eq. (2.64). Bearing in mind the vector relation \( \nabla \times (E \hat{z}) = \nabla E \times \hat{z} \), we obtain:

\[
H = \frac{1}{i\omega\mu_0} \nabla E \times \hat{z}.
\] (2.74)

Furthermore, on each side of the interface with normal vector \( \mathbf{n} \), the continuity of the tangential component of the magnetic field entails the continuity of \( \mathbf{n} \times \mathbf{H} \), and from Eq. (2.74), it turns out that

\[
\mathbf{n} \times \mathbf{H} = \frac{1}{i\omega\mu_0} \mathbf{n} \times (\nabla E \times \hat{z}) = \frac{1}{i\omega\mu_0} \nabla E (\mathbf{n} \cdot \hat{z}) - \hat{z} (\mathbf{n} \cdot \nabla E) = -\frac{1}{i\omega\mu_0} \hat{z} \frac{\partial E}{\partial n}.
\] (2.75)

It follows that the normal derivative of the electric field \( \frac{\partial E}{\partial n} \) is continuous across the interface.

For p-polarization, the electric field can be expressed from Eq. (2.67):

\[
E = \frac{1}{-i\omega\varepsilon_0} \nabla \times (H \hat{z}) = \frac{1}{-i\omega\varepsilon_0} \nabla H \times \hat{z},
\] (2.76)

and thus the continuity of \( \mathbf{n} \times \mathbf{E} \) leads to the continuity of \( \frac{1}{\varepsilon} \frac{\partial H}{\partial n} \). This second boundary condition is slightly more complicated than that obtained for s-polarization, due to the fact that, in contrast with the permeability, the permittivity is not the same on the two sides of the interface. It can be noticed that this dissymmetry disappears for magnetic materials, for which \( \frac{1}{\mu} \frac{\partial E}{\partial n} \) is continuous across the interface.

In conclusion, the boundary conditions across the interface can be written in 2D problems:

\[
\text{E and } \frac{\partial E}{\partial n} \text{ are continuous for s-polarized light,} \hspace{1cm} \text{(2.77)}
\]

\[
\text{H and } \frac{1}{\varepsilon} \frac{\partial H}{\partial n} \text{ are continuous for p-polarized light.} \hspace{1cm} \text{(2.78)}
\]

The surface charge densities can be derived by taking the divergence of Eqs. (2.62) and (2.68). Bearing in mind that the surface density included in \( \Delta \cdot \mathbf{V} \) is equal to \( \mathbf{n} \cdot (\mathbf{V}_+ - \mathbf{V}_-) \), with \( \mathbf{V}_+ - \mathbf{V}_- \) being the jump of \( \mathbf{V} \) in the direction of \( \mathbf{n} \) across the interface \[29\], it comes out that:

\[
\rho_c = \mathbf{n} \cdot (\mathbf{D}_+ - \mathbf{D}_-)
\] (2.79)

\[
\rho_t = \varepsilon_0 \mathbf{n} \cdot (\mathbf{E}_+ - \mathbf{E}_-),
\] (2.80)

with \( \rho_c \) and \( \rho_t \) being the surface densities of free and total charges.
References

2. Lord Rayleigh, Note on the remarkable case of diffraction spectra described by Prof. Wood. Philos. Mag. 14, 60–65 (1907)
16. A. Wirgin, Théorie électromagnétique de la diffraction d’une onde par une surface quasim plane, Thèse d’Etat, Université de Paris, France, 1967


