Reviews of Electromagnetics, Vol. I, 2022
Overview paper
DOI: 10.53792/roe/2022.1/21011

Reviews of Electromagnetics
Overview paper

Overview of the Near-Ground Propagation Mechanisms in Wireless Communication: Is the Norton Surface Wave Significant?

Krzysztof A. Michalski[1]* and Juan R. Mosig[2]

Dedicated to the memory of Professor Heinrich Ott

Abstract
In this paper we re-examine the solution of the century-old Sommerfeld problem of a vertical Hertzian dipole radiating over planar ground, in the context of high-frequency applications, where the media losses are relatively small. To gain more insight into the near-ground wireless propagation mechanisms, we derive the asymptotic field of a dipole above a half-space by the modified saddle-point method carried out to the second order in the inverse radial distance, paying particular attention to the Norton surface wave and its significance in the total field of the dipole. We illustrate and validate the theoretical developments by numerical results involving various transmitter-receiver configurations of interest and different lower half-space media, including seawater and urban ground. We also briefly review the history of this problem and the most pertinent literature, addressing certain lingering controversies and correcting some recent misconceptions.

Key terms
Hertzian dipole; Half-space problem; Sommerfeld pole; Norton surface wave; Zenneck wave; Surface plasmon polariton; Near-ground wireless communication

[1] *Texas A&M University, College Station, TX 77843-3128 USA*
[2] *École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland*
*__Corresponding author__: k-michalski@tamu.edu

Revised: 30/08/2021, Accepted: 08/03/2022, Published: 01/04/2022

1. Introduction

The electromagnetic field of an antenna above a half-space has been the subject of many studies and several monographs have been devoted to this topic [1–4]. This problem was first rigorously formulated by Sommerfeld [5], who considered a vertical, z-oriented Hertzian dipole in the planar interface between air and ground, and obtained a rigorous solution expressed in terms of the eponymous integral along the real axis in the complex transverse wavenumber plane. Sommerfeld then deformed the integration path around two hyperbolic branch cuts and the pole of the integrand, now known as the Sommerfeld pole, and expressed the z component of the vector potential in the upper half-space as $\Pi = P + Q_1 + Q_2$, where he identified the pole residue term P as a cylindrical variant of the Zenneck [6] wave, and the branch cut integral contribution Q_1, associated with the upper medium (air), as space wave. The contribution Q_2 from the other branch cut he subsequently neglected, assuming a highly lossy lower half-space. The Zenneck wave term was intriguing, in view of its slow $\rho^{-1/2}$ amplitude decay, where ρ is the cylindrical radial distance along the surface. But is the Zenneck wave a dominant component of the total surface field far from the transmitter? To answer this question, Sommerfeld proceeded to the far-zone approximation of Q_1, since this integral was not amenable to a closed-form solution, and after a lengthy and intricate derivation, which repeatedly invoked the high-contrast approximation, arrived at an expression of $P + Q_1$ in

terms of the error function of a complex argument involving a square root of a parameter called "numerical distance," which is essentially a product of ρ and the complex "distance" of the pole from the branch point associated with the upper half-space. The asymptotic expansion of the error function for a large numerical distance then yielded the Zenneck wave as the dominant part of the surface field of the dipole, and thus it appeared that this slowly-decaying wave, not the space wave, was the main mechanism of near-ground radio wave propagation. However, as it later turned out, this was an erroneous conclusion, resulting from some elusive error(s) in Sommerfeld's derivation, which manifested itself by a wrong sign of the square root of the numerical distance in the argument of the error function. In conclusion of his paper, Sommerfeld also included a brief discussion of an alternative representation of the solution as a double integral over the spectrum of homogeneous plane waves propagating in all directions in space, as well as inhomogeneous plane waves propagating along the interface. The latter approach was used a decade later by Weyl [7] as the point of departure for his asymptotic solution. According to Weyl, the two-fold integral formulation afforded more freedom in the application of Cauchy's integral theorem, but his method was rather intricate and had very few followers. However, Weyl derived the first known correct asymptotic solution for the surface field of a dipole, since—as mentioned above—the Sommerfeld's original asymptotic analysis was flawed and led to a wrong conclusion that the Zenneck wave was the predominant component of the field radiated by a vertical dipole over a finitely conducting ground. Sommerfeld later revisited this problem in a paper under the same title [8], where he used a still different asymptotic procedure and this time obtained a correct solution, in agreement with Weyl.

The approximations used by Sommerfeld and his followers were ingenious but mostly ad hoc and they relied on the high-contrast assumption—which was fully justified, given the low frequency range then of interest. The derived field expressions were applicable in a restricted elevation range near the interface, with the surface field expressed in terms of the error function complement of complex argument, as exemplified by the early work of van der Pol and Niessen [9] and the much later contribution of Kockel [10]. The latter, which was written to commemorate the 50th anniversary of the Sommerfeld's [5] original paper, may be considered an epilogue to the early era in the history of the problem. A major milestone in that era was the work of Norton [11], who combined the earlier results of van der Pol [12] and Wise [13] to arrive at a field representation applicable for all elevation angles, and which became the de facto standard for many years to follow.[1] Perhaps the first systematic asymptotic analysis of the Sommerfeld problem was due to Ott [14], who adapted the modified saddle-point method of Pauli [15].[2] Similar multiplicative pole treatment was subsequently developed by Clemmow [16–18], and was later extensively employed by Wait [19, 20]. However, the multiplicative method has been largely superseded by the additive method of van der Waerden [21], which was elaborated by Felsen and Marcuvitz [22, Ch. 4], and employed early on by Bernard and Ishimaru [23] and Collin [24].[3] The multiplicative and additive variants of the modified saddle-point-method were more recently studied by Michalski and Lin [25], who showed them to be theoretically equivalent. However, since the residue at the pole of the integrand is explicit in the additive method, but it only arises as a sum of an infinite number of terms in the multiplicative variant, the latter is less accurate with the same number of terms included.

Any asymptotic solution of the Sommerfeld half-space problem comprises, at a minimum, a geometrical-optics field of the first order in r^{-1}, where r is the radial distance from the dipole to the field point, and a surface field associated with the Sommerfeld pole, nowadays usually referred to as the Norton surface wave [26]. The geometrical-optics field, which consists of the direct and the specularly reflected waves, may also be legitimately called a space wave, since it vanishes at the interface. The Norton wave comprises the so-called attenuation function of a complex parameter generally known as a numerical distance in homage to Sommerfeld, which depends on the location of the pole with respect to the wavenumber of the upper half-space, the radius, and the elevation angle. A remarkable feature of the attenuation function is that it has a different asymptotic form, depending on the argument of the numerical distance: if the argument is positive, the asymptotic form comprises a trapped pole wave akin to the Zenneck surface wave, which is then the dominant part of the far-zone surface field. However, for a conventional homogeneous ground, the argument of the numerical distance is always negative, and thus the Zenneck wave does not appear in the asymptotic surface field. In the absence of the trapped surface wave, the far-zone behavior of the surface field is r^{-2}, typical of a lateral wave.

Over the past decade there has been a renewed interest in the wave propagation mechanisms over the ground or sea surface in the context of surface wave radar, cellular wireless communications, near-ground sensors, and on-body networks [26–40].[4] In these applications the canonical half-space problem of Sommerfeld is still highly relevant, but with the trend of increasing operating frequencies some of the assumptions in the earlier solutions may have to be re-examined. The effect of the higher frequencies is to reduce the loss tangent of the media, which decreases the material contrast between the two half-spaces and increases the numerical distance parameter. As a result, the role of the Norton surface wave tends to be de-emphasized, but the hitherto always neglected branch cut wave may have to be taken into account [40, 41].

The purpose of this paper is to re-examine the solution of the Sommerfeld half-space problem in light of the recent high-frequency applications—with references to the most pertinent literature and addressing some of the associated controversies and misconceptions, and finally to answer the question posed in the title. To better understand the wireless propagation

mechanisms over flat and smooth ground or water surface, we derive the asymptotic field of a vertical dipole over a half-space medium by the modified saddle-point method carried out to the second order in the inverse radial distance, paying special attention to the Norton surface wave and its significance in the total field of the dipole. We illustrate and validate the theoretical developments by numerical examples involving various transmitter-receiver configurations of interest and different lower half-space media, including seawater and urban ground. For easy reference, we include supplementary mathematical details in the appendices. We also provide copious notes to elucidate various aspects of the problem and its solution.

2. Problem geometry and notation

Consider a z-oriented time-harmonic Hertzian dipole with the current moment $I\ell$ [Am], located on the z axis at a distance h above a dielectric half-space, as illustrated in Fig. 1. The

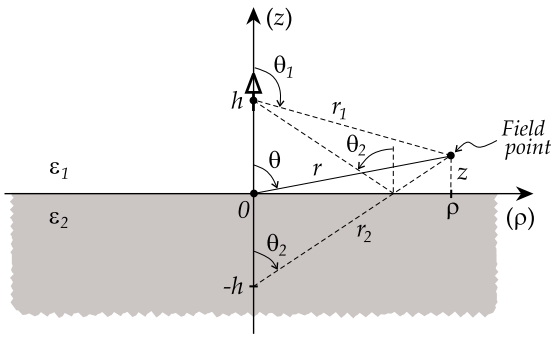

Figure 1: Vertical Hertzian dipole at a height h above a dielectric half-space.

permeability and permittivity of medium n are denoted by μ_n and ε_n, so that the corresponding intrinsic impedance and wavenumber are $\eta_n = \sqrt{\mu_n/\varepsilon_n}$ and $k_n = \omega\sqrt{\mu_n\varepsilon_n}$. The parameters of free space will be distinguished by a subscript zero, with $k_0 = 2\pi/\lambda_0$, where λ_0 is the corresponding wavelength. We limit attention to non-magnetic media with $\mu_n = \mu_0$ and the upper half-space medium will be taken as air, with $\varepsilon_1 = \varepsilon_0$, whereas the lower half-space will be assumed lossy, with complex-valued parameters. A lossy medium with the conductivity σ will be characterized by the complex-valued relative permittivity $\varepsilon_r = \varepsilon_r' - j\sigma/(\omega\varepsilon_0) = \varepsilon_r'(1 - j\tan\delta)$, where $\tan\delta$ is the loss tangent. The $e^{j\omega t}$ time convention is implied. The material contrast of the two half-spaces is given by the relative dielectric constant $\epsilon = \varepsilon_2/\varepsilon_1 = |\epsilon|e^{-j\phi}$, where $\phi \in [0, \pi/2)$, assuming non-plasmonic media [42]. The coordinates associated with the dipole are also illustrated in Fig. 1 and are self-explanatory. We note, in particular, that the wave specularly reflected at an angle θ_2, which travels the distance r_2 from the dipole to the field point, may also be interpreted as originating at the geometrical image of the source at $z = -h$.

3. Formal solution

To investigate the electromagnetic field in the upper half-space due to the vertical dipole in Fig. 1, it will suffice to consider the z component of the electric field, E_{z1}, which may readily be derived from the z component of the vector potential, A_{z1}. In Appendix A we present for easy reference the rigorous solution for A_{z1}, as well as its approximate form suitable for the asymptotic analysis. This solution comprises a Sommerfeld integral along the real axis in the k_ρ−plane, indicated as C in Fig. 14. We transform this solution by the substitutions $\rho = r_2\sin\theta_2$ and $z + h = r_2\cos\theta_2$, and also introduce a new variable of integration via

$$k_\rho = k_1\sin\xi\,, \quad k_{z1} = k_1\cos\xi \tag{1}$$

so that[5]

$$k_{z2} = k_1\kappa(\xi)\,, \quad \kappa(\xi) = \sqrt{\epsilon - \sin^2\xi} \tag{2}$$

and as a result we obtain

$$E_{z1} = -j\eta_1 k_1^2 \frac{I\ell}{4\pi}\mathcal{E} \tag{3}$$

where $\mathcal{E}$ is the normalized field given as

$$\mathcal{E} = \mathcal{E}_D + \mathcal{E}_S \tag{4}$$

in which

$$\mathcal{E}_D = \left[\sin^2\theta_1 - \left(1 - 3\cos^2\theta_1\right)\left(\frac{j}{\Omega_1} + \frac{1}{\Omega_1^2}\right)\right]\frac{e^{-j\Omega_1}}{\Omega_1} \tag{5}$$

with $\Omega_1 = k_1 r_1$ is the direct (whole-space) field of the dipole and

$$\mathcal{E}_S = \frac{1}{\sqrt{\pi\Omega_2}}\int_C \Gamma(\xi) f(\xi)\, e^{\Omega_2 q(\xi)}\, d\xi \tag{6}$$

with $\Omega_2 = k_1 r_2$ is the scattered field, which accounts for the presence of the dielectric boundary. The integrand terms in (6) are defined as

$$\Gamma(\xi) = \frac{\cos\xi - \Delta(\xi)}{\cos\xi + \Delta(\xi)}\,, \quad \Delta(\xi) = \frac{\kappa(\xi)}{\epsilon} \tag{7}$$

$$f(\xi) = \frac{\sin^2\xi}{\sqrt{2j}}\sqrt{\frac{\sin\xi}{\sin\theta_2}}\,\varpi(\xi) \tag{8}$$

$$\varpi(\xi) = 1 + \frac{j}{8\Omega_2\sin\theta_2\sin\xi} \tag{9}$$

$$q(\xi) = -j\cos(\xi - \theta_2) \tag{10}$$

where $\Gamma(\xi)$ is the reflection coefficient for parallel-polarized (transverse-magnetic) waves. The integration path C is illustrated in Fig. 2, showing the top Riemann sheet of the ξ-plane, which is the image of sheets I and II from the k_ρ-plane in the mapping (1). The Roman numerals refer to the mapped sheets and their subscripts to the quadrants, with the quadrant boundaries marked by dotted lines. The transformation (1)

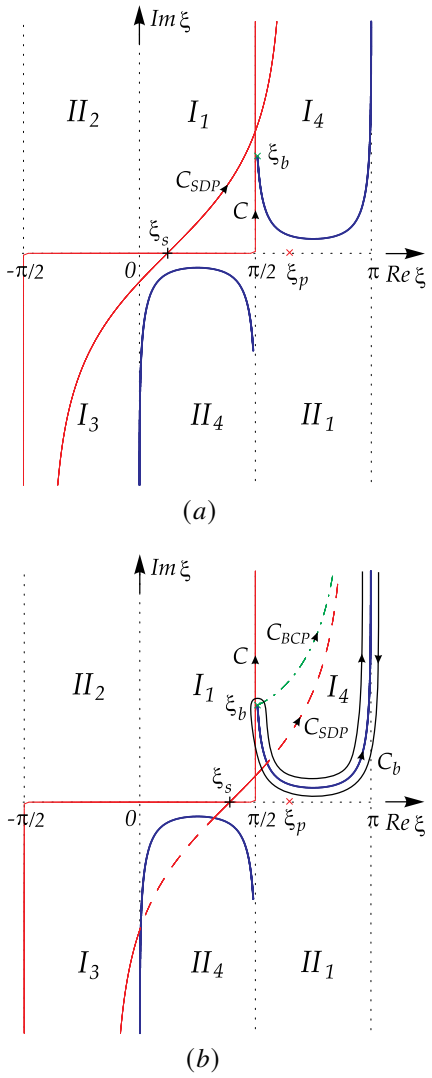

Figure 2: Top Riemann sheet of the ξ-plane for (a) $\theta_2 < \theta_{cb}$ and (b) $\theta_2 > \theta_{cb}$, assuming that $\epsilon = 4 - j0.2$ for illustration purposes.

removes the branch cuts associated with k_{z1}, but those associated with k_{z2} persist and are indicated by the solid blue lines. The condition $\Im m\, k_{z2} < 0$ holds everywhere on this sheet, with the exception of the branch cuts, where $\Im m\, k_{z2} = 0$. We also indicate by the green cross the branch point ξ_b, which is the image of k_2 from Sheet I and satisfies

$$\sin\xi_b = \sqrt{\epsilon}, \quad \cos\xi_b = -j\sqrt{\epsilon - 1} \tag{11}$$

from which we obtain

$$\xi_b = \pi/2 + j\ln\left(\sqrt{\epsilon} + \sqrt{\epsilon - 1}\right) \tag{12}$$

and by the red cross we indicate the pole ξ_p of the reflection coefficient $\Gamma(\xi)$, which is the image of the Sommerfeld pole k_p from Sheet I and satisfies

$$\sin\xi_p = \sqrt{\frac{\epsilon}{\epsilon+1}}, \quad \cos\xi_p = -\frac{1}{\sqrt{\epsilon+1}} \tag{13}$$

from which we obtain

$$\xi_p = \pi/2 + j\ln\left(\sqrt{\frac{\epsilon}{\epsilon+1}} - \frac{j}{\sqrt{\epsilon+1}}\right). \tag{14}$$

The meaning of the other paths shown in this figure will be explained in due course. Note that we do not show the bottom sheet of the ξ-plane, which is the image of sheets III and IV from the k_ρ-plane and lies directly beneath the top sheet illustrated in Fig. 2.

4. Asymptotic analysis

4.1. Saddle point contribution

Preparatory to the application of the saddle-point integration method, we next deform the original path C in (6) into the steepest-descent path passing through the saddle point ξ_s, and will denote the resulting contribution to the scattered field by $\mathcal{E}_S^{\text{SDP}}$. The saddle point is given as the root of $q'(\xi) = 0$, where the prime denotes a derivative, and in view of (10) we find that $\xi_s = \theta_2$. The steepest-descent path (SDP) is defined by the conditions $\Im m\, q(\xi) = \Im m\, q(\xi_s)$ and $\Re e\, q(\xi) \le \Re e\, q(\xi_s)$, or in our case

$$\Re e\, \cos(\xi - \theta_2) = \cos(\xi' - \theta_2)\cosh\xi'' = 1 \tag{15}$$

$$\Im m\, \cos(\xi - \theta_2) = -\sin(\xi' - \theta_2)\sinh\xi'' \le 0 \tag{16}$$

where the notation $\xi = \xi' + j\xi''$ is implied, from which we obtain

$$\xi' = \theta_2 \pm \operatorname{arccos}\operatorname{sech}\xi'', \quad -\infty < \xi'' < \infty \tag{17}$$

where the upper (lower) sign should be selected in the positive (negative) range of ξ'', so that the SDP is asymptotic to the vertical lines $\xi' = \theta_2 \pm \pi/2$. This path is denoted as C_{SDP} and indicated by a red continuous line in Fig. 2. We note that, in view of (17), the SDP passes through the saddle point θ_2 at an angle of $\pi/4$ with the real axis, and that it shifts to the right as θ_2 increases toward $\pi/2$. The limiting SDP passing through $\theta_2 = \pi/2$ is thus given as $\sin\xi'\cosh\xi'' = 1$, and since $\Re e\, k_\rho/k_1 = \Re e\, \sin\xi = \sin\xi'\cosh\xi''$, it is also a boundary of the fast wave region [43, Sec. 19.5]. The fast wave pole ξ_p, located below this boundary in the I_4 quadrant of the top sheet of the ξ plane, as illustrated in Fig. 2, will thus never be captured when C is deformed into C_{SDP}, although it may still indirectly affect the SDP integral, especially in the high-contrast case, when ξ_s may be close to ξ_p for elevation angles near the interface. The branch point ξ_b, however, can definitely be captured by the SDP, which will occur when θ_2 exceeds a certain angle θ_{cb} given as

$$\theta_{cb} = \xi_b' - \operatorname{arccos}\operatorname{sech}\xi_b'' \tag{18}$$

which in the lossless case reduces to $\theta_{cb} = \arcsin\left(1/\sqrt{\epsilon}\right)$. The SDP location for $\theta_2 < \theta_{cb}$ and $\theta_2 > \theta_{cb}$ is illustrated in Fig. 2 (a) and (b), respectively. Note that in the latter case C_{SDP} passes through the branch cuts and approaches the

vertical asymptote $\xi' = \theta_2 + \pi/2$ on the bottom sheet, hence the SDP integral should be augmented by an integral around the branch cut C_b, so that $C_{\rm SDP}$ and the original path C may be reconnected on the top sheet at infinity [2, Sec. 30.1]. The parts of the SDP on the bottom sheet are distinguished by a broken line.

To facilitate the asymptotic evaluation of $\mathcal{E}_S^{\rm SDP}$, we first change the variable of integration in (6) via the transformation

$$\cos(\xi - \theta_2) = 1 - j s^2, \quad -\infty < s < \infty \tag{19}$$

or equivalently

$$s^2 = -j\,[1 - \cos(\xi - \theta_2)] = -2j \sin^2\left(\frac{\xi - \theta_2}{2}\right) \tag{20}$$

with the Jacobian

$$\frac{d\xi}{ds} = \frac{\sqrt{2j}}{\sqrt{1 - js^2/2}} \tag{21}$$

and upon solving (19) for ξ subject to (21) we obtain

$$\xi_{\rm SDP} = \theta_2 + j \ln\left[(1 - js^2) + (1 - j)s\sqrt{1 - js^2/2}\right]. \tag{22}$$

The SDP is thus transformed into the real axis in the complex s plane, with the saddle point at the origin, and $\mathcal{E}_S^{\rm SDP}$ takes the form

$$\mathcal{E}_S^{\rm SDP} = \frac{e^{-j\Omega_2}}{\sqrt{\pi\Omega_2}} \int_{-\infty}^{\infty} G(s)\, e^{-\Omega_2 s^2} ds \tag{23}$$

where

$$G(s) = \frac{d\xi}{ds}\, \Gamma(\xi) f(\xi) \Big|_{\xi = \xi_{\rm SDP}}. \tag{24}$$

The integral in (23) is for large Ω_2 amenable to the modified saddle-point integration method, which takes onto account the influence of the Sommerfeld pole, as summarized for easy reference in Appendix B. Upon applying this method we obtain the approximation

$$\mathcal{E}_S^{\rm SDP} \approx \mathcal{E}_R^{(1)} + \mathcal{E}_R^{(2)} + \mathcal{E}_N^{(2)} \tag{25}$$

where the first two terms are the first- and second-order reflected fields given as

$$\mathcal{E}_R^{(1)} = \sin^2\theta_2\, \Gamma(\theta_2)\, \frac{e^{-j\Omega_2}}{\Omega_2} \tag{26}$$

and

$$\mathcal{E}_R^{(2)} = -j\left[(1 - 3\cos^2\theta_2)\Gamma(\theta_2) - \sin^2\theta_2\, \Upsilon(\theta_2)\right] \frac{e^{-j\Omega_2}}{\Omega_2^2} \tag{27}$$

where

$$\Upsilon(\theta_2) = \frac{5}{2}\cot\theta_2\, \Gamma'(\theta_2) + \frac{1}{2}\Gamma''(\theta_2) \tag{28}$$

with the derivatives of the reflection coefficient given as

$$\Gamma'(\xi) = -\frac{2(\epsilon - 1)}{\epsilon\, \kappa(\xi)\, \delta^2(\xi)} \sin\xi \tag{29}$$

$$\Gamma''(\xi) = \Gamma'(\xi)\left\{\frac{\epsilon\cot\xi}{\kappa^2(\xi)} + 2\frac{\sin\xi}{\delta(\xi)}\left[1 + \frac{\cos\xi}{\epsilon\,\kappa(\xi)}\right]\right\} \tag{30}$$

where $\delta(\xi) = \cos\xi + \kappa(\xi)/\epsilon$ is the denominator of $\Gamma(\xi)$.

The third term in (25), which may be considered a second-order Norton surface wave, is given as

$$\mathcal{E}_N^{(2)} = \frac{r_p}{s_p}\, \mathcal{F}(p)\, \frac{e^{-j\Omega_2}}{\Omega_2} \tag{31}$$

where we have introduced the function [25]

$$\mathcal{F}(p) \equiv F(p) + \frac{1}{2p} \tag{32}$$

which modifies the Sommerfeld-Norton attenuation factor $F(p)$ defined in Appendix B. For convenience, where there is no danger of confusion, we will refer to the modified attenuation function (32) as simply the attenuation function, and to the second-order Norton surface wave (31) as simply the Norton wave. The parameter p is the numerical distance introduced in (70), which takes the form $p = \Omega_2 s_p^2$, where we recall that s_p is the pole of $G(s)$. Since this pole is an image of ξ_p under the transformation (20), we find

$$\begin{aligned} s_p^2 &= -j\left[1 - \cos(\xi_p - \theta_2)\right] \\ &= -j\left[1 - \sin\xi_p \sin\theta_2 - \cos\xi_p \cos\theta_2\right] \end{aligned} \tag{33}$$

and upon invoking (13) we obtain

$$p = -j\Omega_2\left(1 - \sqrt{\frac{\epsilon}{\epsilon + 1}}\sin\theta_2 + \frac{\cos\theta_2}{\sqrt{\epsilon + 1}}\right). \tag{34}$$

Finally, r_p is the residue of $G(s)$ at s_p, which is identical to the residue of $\Gamma(\xi) f(\xi)$ at ξ_p, and since $f(\xi)$ is regular at ξ_p, we obtain

$$r_p = \frac{2\epsilon\sqrt{\epsilon}}{\epsilon^2 - 1} f(\xi_p) = \sin^2\theta_2\, r_p' \tag{35}$$

where

$$r_p' = \sqrt{-2j}\, \frac{\epsilon\sqrt{\epsilon}}{\epsilon^2 - 1}\left(\frac{\sin\xi_p}{\sin\theta_2}\right)^{5/2} \varpi(\xi_p). \tag{36}$$

It should be noted here that $\varpi(\xi_p)$ comprises an additive term with Ω_2^{-1} dependence, which contributes to the higher-order terms in the Norton wave, but its effect is negligible when $\Omega_2 \gg 1$.

The numerical distance p plays an important role in the asymptotic field representation, as discussed in Appendix C, where we also illustrate the dependence of p on θ_2 and the media parameters. The second-order Norton wave $\mathcal{E}_N^{(2)}$ is related to the original Norton [11] surface wave, which appears as (80) in Appendix D. In Appendix E we show that for conventional lower half-space media the Zenneck wave does not appear in the asymptotic form of $\mathcal{E}_N^{(2)}$. Furthermore, in Appendix F we demonstrate that the Norton wave may be omitted in (25) when $|p| \gtrapprox 10$.

4.2. Branch cut contribution

We next turn attention to the branch cut integral, which has to be added to the SDP integral when θ_2 exceeds θ_{cb}, as illustrated in Fig. 2(*b*). Note that on the path around the branch cut C_b, the condition $\Re e\, k_{z2} > 0$ holds on the right side of C_b and the condition $\Re e\, k_{z2} < 0$ on the left side, as may be inferred from the properties of k_{zn} illustrated in Fig. 12. We will subsequently deform C_b to the steepest-descent path $C_{\rm BCP}$ emanating from the branch point ξ_b, which is indicated by a green dash-dot line in Fig. 2(*b*). Such a deformation is permissible as there are no singularities in the region of the complex plane swept in this process [44]. It follows from the condition $\Im m\, q(\xi) = \Im m\, q(\xi_b)$ that this new BCP is given as

$$\xi' = \theta_2 + \arccos\left[\cos(\xi_b' - \theta_2)\cosh\xi_b'' \operatorname{sech}\xi''\right] \tag{37}$$

with $\xi_b'' \le \xi'' < \infty$, and we note that it approaches the vertical asymptote $\xi' = \theta_2 + \pi/2$ for $\xi'' \to \infty$. The contribution to $\mathcal{E}^S$ from the integral around the deformed branch cut path is found as

$$\mathcal{E}_B = \frac{1}{\sqrt{\pi\Omega_2}} \int_{C_{\rm BCP}} [\![\Gamma(\xi)]\!] f(\xi)\, e^{\Omega_2 q(\xi)} d\xi\,, \quad \theta_2 > \theta_{cb} \tag{38}$$

where

$$[\![\Gamma(\xi)]\!] = \frac{4\cos\xi\,\kappa(\xi)/\epsilon}{\cos^2\xi - \kappa^2(\xi)/\epsilon^2}\,, \quad \Re e\,\kappa(\xi) > 0 \tag{39}$$

represents the jump in $\Gamma(\xi)$ between the top and bottom sheets along the integration path.

To facilitate the asymptotic evaluation of $\mathcal{E}_B$, we introduce in (38) a new variable of integration via the substitution

$$\cos(\xi - \theta_2) = \cos(\xi_b - \theta_2) - js^2 \equiv \zeta(s)\,, \quad 0 \le s < \infty \tag{40}$$

with the Jacobian

$$\frac{d\xi}{ds} = \frac{2js}{\sqrt{1-\zeta^2(s)}} \tag{41}$$

and upon solving (40) for ξ we find

$$\xi_{\rm BCP} = \theta_2 + j\ln\left[\zeta(s) - j\sqrt{1-\zeta^2(s)}\right]. \tag{42}$$

After this transformation $\mathcal{E}_B$ takes the form

$$\mathcal{E}_B = \frac{e^{-j\Omega_2\cos(\xi_b-\theta_2)}}{\sqrt{\pi\Omega_2}} \int_0^\infty B(s)\, e^{-\Omega_2 s^2} s\, ds\,, \quad \theta_2 > \theta_{cb} \tag{43}$$

where

$$B(s) = \frac{2j}{\sqrt{1-\zeta^2(s)}} [\![\Gamma(\xi)]\!] f(\xi)\Big|_{\xi=\xi_{\rm BCP}}. \tag{44}$$

We next approximate $B(s)$ in (43) by the two-term Maclaurin expansion

$$B(s) \approx B(0) + B'(0)s \tag{45}$$

where we note that $B(0) = 0$, since $s = 0$ corresponds to $\xi = \xi_b$ and $[\![\Gamma(\xi_b)]\!] = 0$, and consequently the coefficient $B'(0)$ may be found as

$$B'(0) = \lim_{s\to 0}\frac{B(s)}{s} = \frac{4\sqrt{2j}\,\Lambda}{\cos\xi_b \sin(\xi_b - \theta_2)}\sqrt{\frac{\sin\xi_b}{\sin\theta_2}} \equiv \Phi(\theta_2) \tag{46}$$

where

$$\Lambda = \lim_{s\to 0}\frac{\kappa(\xi)}{s} \overset{\rm H}{=} \sqrt{\frac{-2j\sin\xi_b\cos\xi_b}{\sin(\xi_b-\theta_2)}}\,. \tag{47}$$

After this approximation, the integral in (43) may be evaluated in a closed form, which yields

$$\mathcal{E}_B^{(2)} \approx \frac{\Phi(\theta_2)}{4\Omega_2^2}\, e^{-k_1\sqrt{\epsilon-1}(h+z)}\, e^{-jk_1\sqrt{\epsilon}\rho} \tag{48}$$

where in the exponent we have invoked (11). The branch cut contribution is thus a second-order effect with Ω_2^{-2} range dependence, which we emphasize by the superscript $^{(2)}$ added in (48). This wave propagates along the interface with the phase velocity corresponding to the medium of the lower half-space, but it also decreases exponentially with the distance if the medium is lossy. Furthermore, its amplitude has a strong exponential decay in the direction normal to the interface with a rate proportional to $(h+z)/\lambda_0$.

4.3. Complete asymptotic field

If we denote by $\mathcal{E}_D^{(1)}$ and $\mathcal{E}_D^{(2)}$ the first- and second-order components of the direct field (5), we may express the complete asymptotic field of the dipole as

$$\mathcal{E} \approx \mathcal{E}_{GO} + \mathcal{E}_L + \mathcal{U}(\theta_2 - \theta_{cb})\,\mathcal{E}_B^{(2)} + \mathcal{E}_N^{(2)} \tag{49}$$

where we have introduced the geometrical-optics field

$$\begin{aligned}\mathcal{E}_{GO} &\equiv \mathcal{E}_D^{(1)} + \mathcal{E}_R^{(1)} \\ &= \sin^2\theta_1\,\frac{e^{-j\Omega_1}}{\Omega_1} + \sin^2\theta_2\,\Gamma(\theta_2)\,\frac{e^{-j\Omega_2}}{\Omega_2}\end{aligned} \tag{50}$$

which comprises the first-order direct (line-of-sight) field of the dipole and a wave specularly reflected from the half-space medium below, and the lateral wave[6]

$$\mathcal{E}_L \equiv \mathcal{E}_D^{(2)} + \mathcal{E}_R^{(2)} \tag{51}$$

which comprises the second-order direct and reflected waves. Since $\mathcal{E}_{GO} = 0$ in the on-surface configuration, where $\theta_1 = \theta_2 = \pi/2$, $r_1 = r_2 = \rho$ and $\Gamma(\theta_2) = -1$, we may also call it the space wave [45], and the remaining field components in (49) may be referred to as the ground wave [38]. In the on-surface configuration this ground wave simplifies to

$$\mathcal{E} \approx -2j\,\frac{\epsilon^2}{\epsilon-1}\left[\frac{e^{-jk_1\rho}}{(k_1\rho)^2} - \frac{1}{\epsilon\sqrt{\epsilon}}\,\frac{e^{-jk_1\sqrt{\epsilon}\rho}}{(k_1\rho)^2}\right] + \mathcal{E}_N^{(2)} \tag{52}$$

where the first and second terms in the brackets arise from the lateral and branch cut waves, respectively. The latter is always present when $\theta_2 = \pi/2$, but is subject to exponential decay in the presence of losses.

5. Numerical results and discussion

We now present sample numerical results that confirm and amplify the theoretical developments of this paper. For illustration purposes we will use three representative lower half-space media, viz., seawater, rich soil, and poor ground, with the parameters listed in Table 1. Note that poor ground has also been referred to as urban ground [46, Appendix A-6]. For these representative media, in Table 2 we list three parameters of particular interest, viz., $|\varepsilon_r|$, loss tangent $\tan\delta$, and $|k_p - k_1|/k_1$, evaluated for the three media at four sample operating frequencies, viz., 30 MHz, 900 MHz, 2.4 GHz, and 28 GHz, selected from a wide range, from HF to G5. For simplicity, we have assumed that ε_r' and σ are frequency-independent. The parameter $|k_p - k_1|/k_1$, which represents the normalized distance of the Sommerfeld pole k_p from the branch point k_1 associated with the upper half-space and which also appears in the Sommerfeld's numerical distance, plays an important role in the asymptotic behavior of the surface field. Note that the effect of the frequency increase is to reduce $\tan\delta$ with the concomitant reduction in $|\varepsilon_r|$, and to increase $|k_p - k_1|/k_1$. However, the change in the last two parameters is only significant for seawater and is almost imperceptible for rich soil and, especially, poor ground.

Table 1: Parameters of three lower half-space media.

№	Medium	ε_r'	σ (S/m)
1	seawater	80	4
2	rich soil	16	10^{-2}
3	poor ground	4	2×10^{-4}

We have validated the asymptotic field formulas against the "exact" results generated using a rigorous numerical integration of the Sommerfeld integrals [47, 48], and in some cases also against completely independent results obtained from the commercial software FEKO [49].

5.1. On-surface configuration results

We begin with the on-surface transmitter-receiver configuration, where $h = z = 0$. In Fig. 3 we plot the asymptotic field (52) and its components over a three-decade ρ range for seawater at the operating frequency of 30 MHz. We note that the total field initially follows a ρ^{-1} (−20 dB/decade) asymptote and then transitions to ρ^{-2} (−40 dB/decade) asymptote in the far zone. This behavior is the result of the interaction of the lateral wave, which has the ρ^{-2} dependence in the entire range, and the Norton wave, which initially also behaves as ρ^{-2}, but then gradually transitions to ρ^{-3} asymptote. The branch cut wave does not contribute in this case, in view of its strong exponential decay. The behavior of the Norton wave (31) changes depending on the magnitude of the numerical distance p, which is proportional to ρ/λ_0, but with a multiplicative factor, which in the on-surface case is equal to $|k_p - k_1|/k_1$ (see Appendix C). As may be gathered

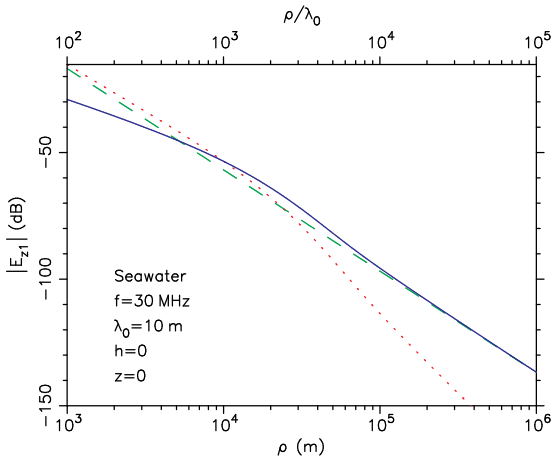

Figure 3: Plots of $|E_{z1}|$ and its components vs. ρ in the on-surface configuration ($h = z = 0$) for seawater at 30 MHz. The total asymptotic field is indicated by a full blue line, the lateral wave by a dashed green line, and the Norton wave by a dotted red line. The branch cut wave is absent due to its strong exponential decay in this case.

from Table 2, the latter parameter can be very small if the media contrast is high, which is indeed the case for seawater at 30 MHz. Therefore, a small p does not necessarily imply a small ρ. In view of (73), we note that $F(p) \approx 1$ for $|p| \ll 1$, and thus $\mathcal{F}(p) \approx 1 + 1/(2p)$, where the term $1/(2p)$ is dominant for small p, so that the Norton wave (31) initially exhibits a ρ^{-2} range dependence. However, for larger values of $|p|$ the behavior of $F(p)$ evolves towards the asymptotic form (74) with the leading term $-1/(2p)$, since the surface wave is absent in the non-plasmonic case, as discussed in Appendix E, and this leading term is cancelled by the $1/(2p)$ term that modifies $F(p)$ in (32), which results in a ρ^{-3} range dependence. We may thus take $|p| = 1$ as the demarcation point where the transition between the ρ^{-2} and ρ^{-3} asymptotes begins [37], and from this condition we obtain the "knee value" of ρ as

$$\rho_{\text{knee}}/\lambda_0 \approx \frac{1}{2\pi|k_p - k_1|/k_1} \approx \frac{|\epsilon|}{\pi} \tag{53}$$

which is seen to be inversely proportional to the normalized distance of k_p from k_1, with the second approximate expression applicable if $|\epsilon| \gg 1$. In the case of Fig. 3, we find $\rho_{\text{knee}}/\lambda_0 \approx 763$. Note that to the left of this point, where $\rho < \rho_{\text{knee}}$, there is a partial cancellation between the Norton wave and the lateral wave, which results in a ρ^{-1} initial behavior of the net total field. In fact, the lateral wave cancels the term associated with $1/(2p)$ in $\mathcal{F}(p)$. To the right of the knee point, on the other hand, there is initially a constructive interference of the Norton and lateral waves, resulting in a slight "bump" in the net field, but then the Norton wave falls below the ρ^{-2} asymptote of the lateral wave and becomes inconsequential when ρ/λ_0 exceeds the cut-off value given by (85) (see Appendix F). Hence, beyond this point, which in

Table 2: Media parameters at different frequencies.

	seawater			rich soil			poor ground		
f	$\lvert\varepsilon_r\rvert$	$\tan\delta$	$\lvert k_p-k_1\rvert/k_1$	$\lvert\varepsilon_r\rvert$	$\tan\delta$	$\lvert k_p-k_1\rvert/k_1$	$\lvert\varepsilon_r\rvert$	$\tan\delta$	$\lvert k_p-k_1\rvert/k_1$
30 MHz	2.4×10^3	30	2.1×10^{-4}	17.1	0.37	2.8×10^{-2}	4	3×10^{-2}	0.11
900 MHz	113.1	1	4.4×10^{-3}	16	1.3×10^{-2}	3×10^{-2}	4	10^{-3}	0.11
2.4 GHz	85.4	0.37	5.8×10^{-3}	16	4.7×10^{-3}	3×10^{-2}	4	3.7×10^{-4}	0.11
28 GHz	80	3.2×10^{-2}	6.2×10^{-3}	16	4×10^{-4}	3×10^{-2}	4	3.2×10^{-5}	0.11

the present case is found at $\rho/\lambda_0 \approx 2.3\times10^4$, the total field is essentially given by just the lateral wave (51).

In Fig. 4 we present similar results for poor ground at 28 GHz. Note that in this low-loss case the branch cut wave does enter the picture. Furthermore, the Norton wave knee

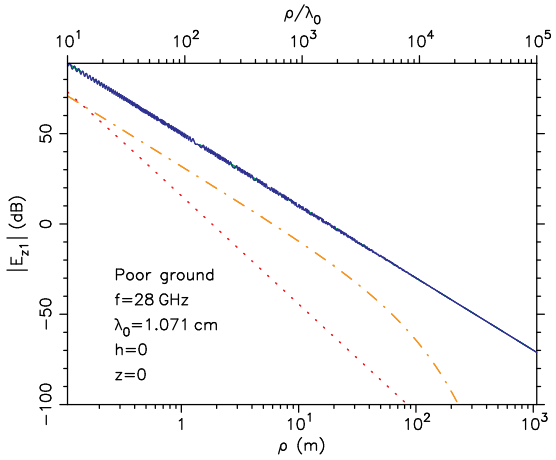

Figure 4: Plots of $|E_{z1}|$ and its components vs. ρ in the on-surface configuration ($h = z = 0$) for poor ground at 28 GHz. The total asymptotic field is indicated by a full blue line, the lateral wave by a dashed green line, and the Norton wave by a dotted red line. In contrast to the case of Fig. 3, the branch cut wave is now present and is indicated by a dash-dot orange line, but the lateral wave and the total field plots overlap and are indistinguishable, except at the beginning of the plotting range.

point is now found at $\rho_{\text{knee}}/\lambda_0 \approx 1.5$, which lies outside the plotting range, and the cut-off point at $\rho/\lambda_0 \approx 45$. Note that the net total field is essentially given by the lateral wave, which follows ρ^{-2} asymptote with superposed wiggles due to the interference from the branch cut wave, before the latter is eventually annihilated by the exponential decay. The error in the asymptotic field relative to the exact result was found as 2.4×10^{-3} %, when averaged over the range $\rho > 100\lambda_0$.

5.2. Elevated near-ground configuration results

We next turn attention to the elevated transmitter-receiver configuration with $h > z > 0$, where we compute the field at a constant level above the interface as we vary the horizontal distance from the dipole. In Fig. 5 we plot the total asymptotic field (49) and its constituents for a dipole at $h = 20$ m over seawater at 30 MHz, with $z = 2$ m and ρ varying over a four-decade range. Note that $(h+z) = 2.2\lambda_0$ in this case. The total

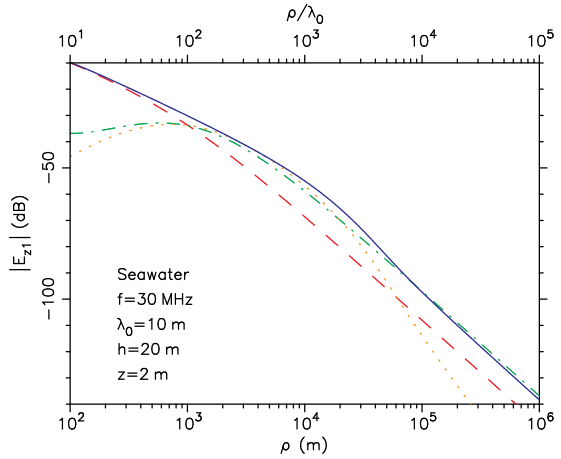

Figure 5: Plots of $|E_{z1}|$ and its components vs. ρ at a fixed $z = 2$ m for a 30 MHz dipole at $h = 20$ m above seawater. The total asymptotic field is indicated by a full blue line, the geometrical-optics field by a dashed red line, the lateral wave by a dash-dot green line, and the Norton wave by a dotted orange line.

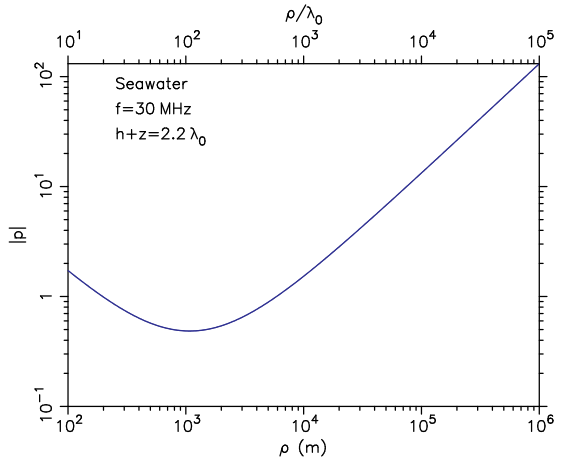

Figure 6: Plot of the magnitude of the numerical distance p vs. ρ for the case of Fig. 5.

field initially follows ρ^{-1} asymptote, and gradually transitions to ρ^{-2} asymptote in the far zone. We note that in the near zone the geometrical-optics field is the dominant component, and that this role is largely taken over by the lateral wave in the far zone. The branch cut wave is absent, due to its strong exponential decay in this case. The Norton wave is most significant in the intermediate range, where it evidently contributes a slight "bump" to the total field. This behavior is explained by the plot in Fig. 6, which shows the corresponding numerical distance $|p|$. It may be confirmed that this plot falls much below the Ω_N demarcation line (see Appendix F), which confirms the strong influence of the Norton wave in this case. We note that the dip in the $|p|$ curve corresponds to the peak of the Norton wave in Fig. 5. The error in the asymptotic field relative to the exact result was in this case found as 2.7×10^{-3} %, when averaged over the range $\rho > 100\lambda_0$. We have also applied the Norton asymptotic formula of Appendix D to this case and found that the average error increased to 4.6×10^{-2} %.

We next consider a 900 MHz dipole at $h = 3$ m above rich soil, with the field observed at $z = 1$ m and ρ varying over a three-decade range. Note that $(h+z) \approx 12\lambda_0$ in this case. We plot the asymptotic field components in Fig. 7, where we

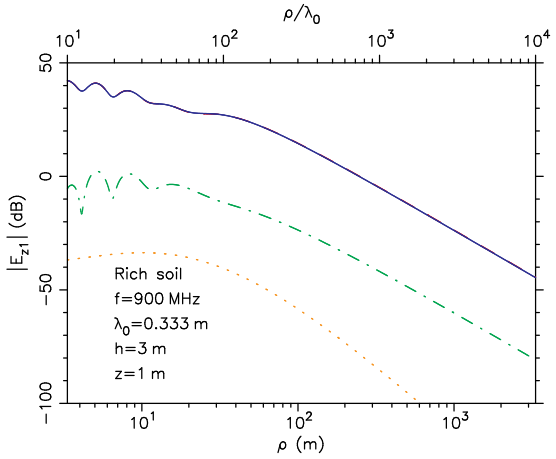

Figure 7: Plots of $|E_{z1}|$ and its components vs. ρ at a fixed level $z = 1$ m for a 900 MHz dipole at $h = 3$ m above rich soil. The total field and the geometrical-optics field plots are indistinguishable.

note that the total field is now completely determined by the geometrical-optics field, as the full blue and dashed red lines overlap and are indistinguishable. Evidently, the contributions of the lateral and Norton waves are negligible in the entire horizontal range considered. The insignificance of the Norton wave in this case could have been anticipated, based on our discussion in Appendix F, since $(h+z) > 10\lambda_0$.

As the next example, in Fig. 8 we present results for a 28 GHz dipole at $h = 20$ m over poor (urban) ground, with the field observed at the $z = 2$ m and ρ varied over a five-decade range. Note that $(h+z) \approx 2055\lambda_0$ in this case. We find again that the dipole field is fully determined by the geometrical-optics component.

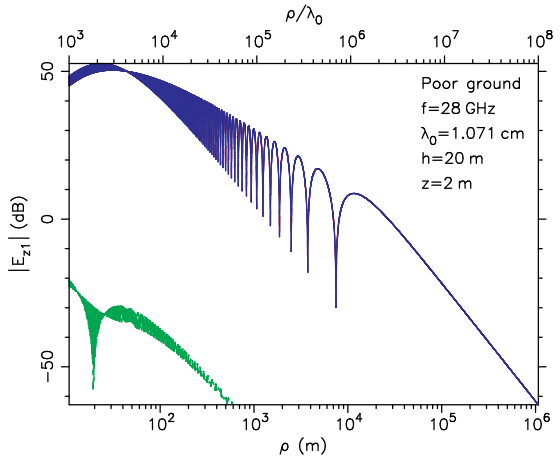

Figure 8: Plots of $|E_{z1}|$ and its components vs. ρ at a fixed $z = 2$ m for a 28 GHz dipole at $h = 20$ m above poor ground. The total asymptotic field is indicated by a full blue line, the geometrical-optics field by a dashed red line, and the lateral wave by a dash-dot green line. The total field and the geometrical-optics field plots are indistinguishable.

To further investigate the near-ground propagation mechanism in the elevated transmitter-receiver configuration, we present two additional examples in Fig. 9 and 10. These cases were previously considered by Cardoso et al. [40] and by Sarkar et al. [32, Fig. 3.15], respectively, and our asymptotic results appear to be in good visual agreement with the corresponding plots presented by these authors. As a further validation of our formulation, we also include plots of the "exact" results generated by the commercial code FEKO. Since $(h+z) \approx 24\lambda_0$ and $\approx 73\lambda_0$ in the cases of Fig. 9 and 10, respectively, here too the total field is fully determined by the geometrical-optics component.

The examples we have presented above confirm that in typical near-ground wireless communication scenarios, where $(h+z) \gtrapprox 10\lambda_0$, the dipole field is dominated by the geometrical-optics component, and thus the Norton wave, the lateral wave, and the branch cut wave are all insignificant. As regards the total field behavior, we note the initial magnitude increase for $\rho \lessapprox h$, and then oscillations—about what appears to be a ρ^{-1} asymptote, which eventually die out, as the field gradually transitions to a ρ^{-2} asymptote. Note, however, that this ρ^{-2} asymptotic behavior is not associated with the lateral wave—as was the case in the on-surface configuration, but is apparently due to the first-order geometrical-optics wave! The initial field increase may be explained by the fact that, close to the dipole where $\rho < (h-z)$, the factor $\sin^2\theta_1/\Omega_1$ in (50) is growing linearly with ρ. The oscillatory field pattern about the initial ρ^{-1} asymptote is due to the interference of the first-order direct wave with the specularly reflected wave, except near the Brewster angle. In the intermediate range we

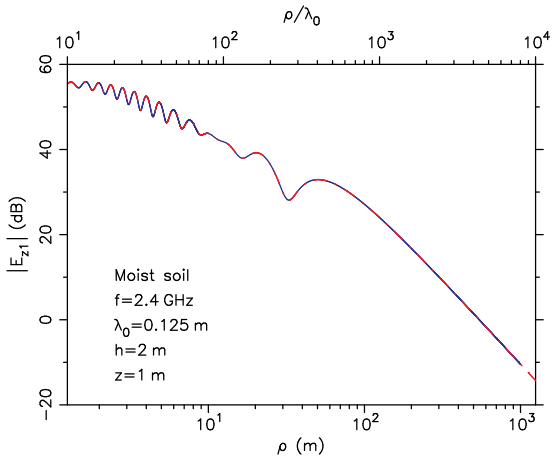

Figure 9: Plots of $|E_{z1}|$ vs. ρ at a fixed $z = 1\,\text{m}$ for a 2.4 GHz dipole at $h = 2\,\text{m}$ above moist soil $(\varepsilon_r' = 15, \sigma = 2\times10^{-3})$. The asymptotic result indicated by a dash-dot red line is superimposed on the FEKO result indicated by a full blue line. These plots are indistinguishable in the figure.

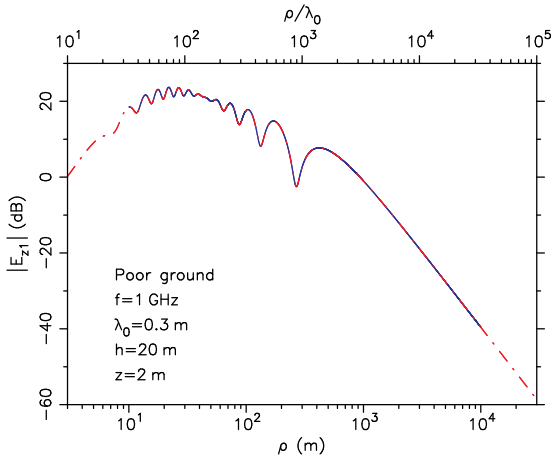

Figure 10: As in Fig. 9, but for a 1 GHz dipole at $h = 20\,\text{m}$ above poor ground, with the field observed at $z = 2\,\text{m}$ and ρ varying over a four-decade range.

find

$$|\mathcal{E}_{GO}| \approx \frac{1}{k_1\rho}\sqrt{[1+\Gamma(\theta_2)]^2 - 4\Gamma(\theta_2)\sin^2(k_1hz/\rho)} \quad (54)$$

where we have assumed for simplicity that $\rho >> (h+z)$ and that the lower half-space medium is nearly lossless, so that ϵ and $\Gamma(\theta_2)$ may be taken as real-valued. This expression explains the interference pattern with a lull where θ_2 corresponds to the pseudo-Brewster angle, at which $\Gamma(\theta_2) \approx 0$. These oscillations are superimposed on ρ^{-1} asymptote and have a "period" increasing with ρ. Their amplitude is also growing as θ_2 approaches $\pi/2$ and $\Gamma(\theta_2)$ tends toward -1. However, when ρ eventually reaches the range where $k_1hz/\rho \ll 1$, the oscillations cease and we transition to

$$|\mathcal{E}_{GO}| \approx \frac{2(h+z)}{k_1\rho^2}\sqrt{\frac{\epsilon^2}{\epsilon-1} + \left(\frac{k_1hz}{h+z}\right)^2} \quad (55)$$

which exhibits the ρ^{-2} dependence characteristic of the lateral wave $\mathcal{E}_L$. Incidentally, this expression also indicates that the far-zone field strength can increase with the distance of the receiver from the ground. The demarcation point where the ρ^{-2} asymptote begins may be estimated as $\rho/\lambda_0 = 4\pi hz/\lambda_0^2$. We should note here that neither our theoretical analysis, nor the numerical results—which have been independently confirmed—indicate the presence of a $\rho^{-1.5}$ (−30 dB/decade) asymptote in the near-ground field of a dipole, as claimed by some authors [30–32, 50].[7]

5.3. Fixed-radius radiation pattern results

We conclude this section with the results in Fig. 11, where we present radiation pattern plots at a fixed radius r for a dipole at $h = 2\,\text{m}$ over poor ground, at two different operating frequencies, viz., 900 MHz and 2.4 GHz. The excellent agreement of the asymptotic and exact results for all elevation angles should be noted. The observed pattern ripples are due predominantly to the interference of the first-order direct and specularly reflected waves, which constitute the geometrical-optics field. As was also the case in the near-ground field pattern at a constant z level, a lull in the ripples occurs near the pseudo-Brewster angle, where the reflected wave has a sharp dip and cannot interfere with the direct field. In the present case this dip occurs at $\theta \approx 64.4°$, which corresponds to the pseudo-Brewster angle $\theta_2 = \theta_B \approx 63.4°$. Note that in this low-loss case the dip location is unaffected by the more than two-fold increase in the operating frequency, which is to be expected, in view of the weak frequency dependence of the Brewster zero in a low-loss case.[8]

6. Conclusion

This paper has dealt with the problem of a vertical Hertzian dipole radiating over a half-space, often referred to as the Sommerfeld half-space problem. We have briefly reviewed its history and the most pertinent literature, beginning with Sommerfeld's seminal paper and including the most recent developments. We have also addressed some controversies associated with this problem, both old and new, clarifying any lingering confusions and correcting certain misconceptions in the recent literature. To gain more insight into the near-ground wireless propagation mechanisms in modern high-frequency applications, where the media losses are relatively small, we have derived the asymptotic field of a vertical dipole above a half-space medium by the modified saddle-point method carried out to the second order in the inverse radial distance. We have identified in the resulting expression the first-order geometrical-optics field and the second-order direct, reflected, branch-cut and Norton waves. The second-order direct and reflected waves we lump under the moniker of a lateral wave.

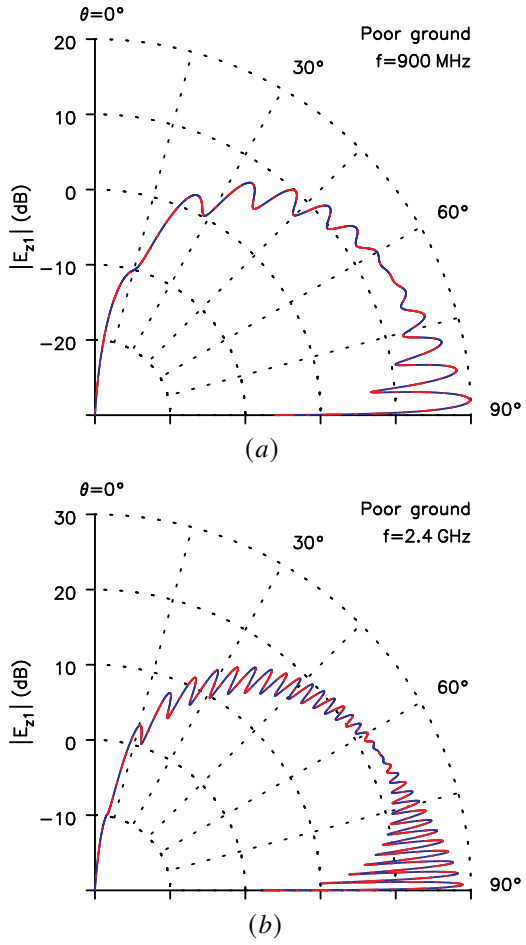

Figure 11: Radiation pattern at $r = 100\,\text{m}$ of a dipole located at $h = 2\,\text{m}$ over poor ground at (a) 900 MHz and (b) 2.4 GHz. Asymptotic results indicated by a red dash-dot line are superposed on the exact results indicated by a full blue line.

The geometrical-optics field vanishes when the dipole and the observation points are on the interface, and may thus be referred to as space wave. On the other hand, the Norton wave, which is associated with the pole of the reflection coefficient, is most significant at the interface, in the horizontal range where a complex parameter called numerical distance is small to moderate. For a sufficiently large numerical distance the Norton wave is dominated by the lateral wave and the surface field behaves asymptotically as ρ^{-2}. The branch cut wave contribution is only visible in cases of vanishingly-small losses. In the elevated configuration and a typical near-ground wireless communication scenario with the transmitting antenna located more than approximately ten wavelengths above the surface, the Norton wave is insignificant and the radiation field is for all practical purposes solely determined by the geometrical-optics space wave. Furthermore, the field at a constant z level is initially oscillatory about ρ^{-1} asymptote and it gradually transitions to a smooth ρ^{-2} asymptotic behavior as ρ increases. It should be emphasized here that in this case the ρ^{-2} lateral wave-like asymptote results from a partial cancellation of the first-order direct and specularly reflected waves in the geometrical-optics field, and not from the Norton wave, which has negligible effect in this scenario. In the intermediate zone between the ρ^{-1} asymptote and the final ρ^{-2} asymptote the slope varies continuously and there is no discernible $\rho^{-1.5}$ asymptote, contrary to what has been claimed in the recent literature. We have illustrated these theoretical findings by numerical results for various transmitter-receiver configurations of interest and different lower half-space media, including seawater and urban ground, and have validated them against commercial code results.

Appendices

A. Solution of the Sommerfeld half-space problem

The z component of the electric field radiated in the upper half-space by the vertical Hertzian dipole of Fig. 1 may be found as [37]

$$E_{z1} = -j\omega\left(1+\frac{1}{k_1^2}\frac{\partial^2}{\partial z^2}\right)A_{z1} = \frac{j\omega}{k_1^2}\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial A_{z1}}{\partial\rho}\right) \tag{56}$$

where A_{z1} is the z component of the magnetic vector potential above the surface, which may be expressed as

$$A_{z1} = \mu_0 k_1 \frac{I\ell}{4\pi}\left(\frac{e^{-jk_1 r_1}}{k_1 r_1} + \mathcal{I}\right) \tag{57}$$

with

$$\mathcal{I} = \int_{-\infty}^{\infty}\Gamma(k_\rho)\frac{e^{-jk_{z1}(z+h)}}{2jk_{z1}}H_0^{(2)}(k_\rho\rho)\frac{k_\rho dk_\rho}{k_1} \tag{58}$$

where k_ρ is the transverse wavenumber, $H_0^{(2)}(\cdot)$ is the zero-order Hankel function of the second kind, and Γ is the half-space reflection coefficient given as

$$\Gamma(k_\rho) = \frac{k_{z1}-k_{z2}/\epsilon}{k_{z1}+k_{z2}/\epsilon}, \quad k_{zn} = \sqrt{k_n^2 - k_\rho^2}. \tag{59}$$

Note that in (57) the first term in the parentheses represents the direct whole-space potential of the dipole. Although it has been a common practice to also extract in a closed form an image dipole term—positive or negative, the resulting total asymptotic potential remains the same, provided that the additive modified saddle-point method is employed [25,51].

To facilitate the asymptotic evaluation of (58) by the saddle-point method, we approximate the Hankel function in the integrand by its large-argument form

$$H_0^{(2)}(k_\rho\rho) \approx \sqrt{\frac{2j}{\pi k_\rho\rho}}\,\varpi(k_\rho\rho)\,e^{-jk_\rho\rho} \tag{60}$$

with

$$\varpi(k_\rho\rho) = 1 + \frac{j}{8k_\rho\rho} \tag{61}$$

where we have retained the first two terms of the expansion, and as a result, we obtain

$$\mathcal{I} \approx \int_{-\infty}^{\infty} \frac{\Gamma(k_\rho \rho)}{\sqrt{2j\pi k_1 \rho}} \sqrt{\frac{k_\rho}{k_1}}\, \varpi(k_\rho \rho)\, e^{-j\left[k_{z1}(z+h)+k_\rho \rho\right]} \frac{dk_\rho}{k_{z1}}\,. \tag{62}$$

Since the approximation (60) is applicable for $|k_\rho \rho| \gg 1$, its use in (62) may be questioned on the grounds that this condition is clearly violated on the integration path near the origin of the k_ρ plane. However, since most of the contribution to the integral arises from the vicinity of the saddle point, which approaches k_1 for elevation angles near the surface, the use of this approximation will still be justified, as long as $k_1 \rho \gg 1$ [52, Ch. 12]. We note that the effect of the differential operators in (56) is to introduce a factor $-j\omega k_\rho^2/k_1^2$ in the integrand of (62).

The longitudinal propagation constants k_{zn} in (59) are double-valued in the complex k_ρ plane, which necessitates the introduction of branch cuts emanating from the branch points k_n, so that k_{zn} becomes uniquely defined over a two-sheeted Riemann surface. It is convenient to choose hyperbolic cuts given by the condition $\Im m\, k_{zn} = 0$, so that the sign of $\Im m\, k_{zn}$ remains unchanged on an entire Riemann sheet. We will

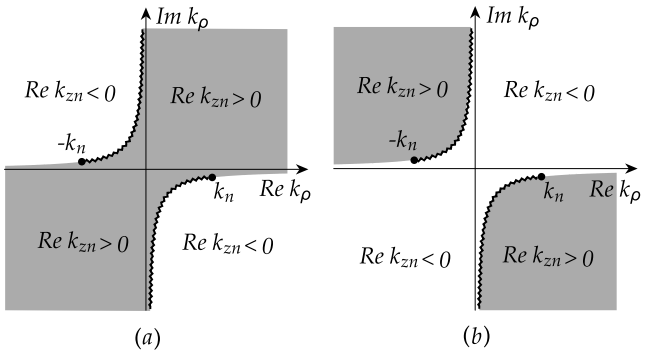

Figure 12: (a) Proper $\left(\Im m\, k_{zn} < 0\right)$ and (b) improper $\left(\Im m\, k_{zn} > 0\right)$ Riemann sheets in the complex k_ρ plane for the case of small losses, with the branch cuts indicated by wiggly lines. In the losses case the branch points $\pm k_n$ move to the real axis and the hyperbolic branch cuts fall on the real and imaginary axes.

then designate the sheet where $\Im m\, k_{zn} < 0$ as proper, and the sheet where $\Im m\, k_{zn} > 0$ as improper. These sheets are also referred to as the top and bottom sheets, respectively, and their properties are illustrated in Fig. 12, where in the shaded regions $\Re e\, k_{zn} > 0$. The integrand functions in (58) and (62) involve both k_{z1} and k_{z2} and may thus be defined over a four-sheeted Riemann surface illustrated in Fig. 13, and the integration path follows the real axis on Sheet I, as indicated in Fig. 14. The Hankel function introduces another branch cut along the negative-real axis, which has no physical significance and will be of no further consequence. The path C should skirt this branch cut and the second-quadrant branch points below, and the fourth-quadrant singularities above. In

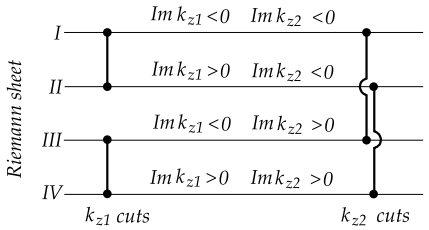

Figure 13: Schematic representation of the four-sheeted Riemann surface associated with k_{z1} and k_{z2}. A continuous transition between two sheets can be effected by crossing a branch cut joining these sheets.

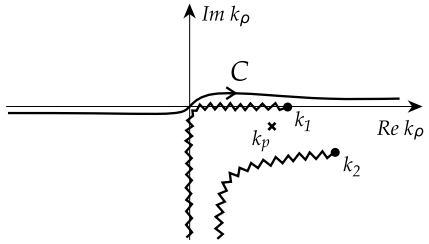

Figure 14: Integration path C along the real axis on the top Riemann sheet. The second-quadrant singularities are not shown for simplicity.

addition to the branch points, the integrand has poles $\pm k_p$ on Sheet I, which are roots of the denominator of the reflection coefficient $\Gamma(k_\rho)$, and thus satisfy the dispersion relation

$$k_{z1} + k_{z2}/\epsilon = 0\,. \tag{63}$$

The fourth-quadrant pole, first uncovered by Sommerfeld [5], is given as

$$k_p = k_1 \sqrt{\frac{\epsilon}{\epsilon+1}} \tag{64}$$

and its approximate position with respect to the branch cuts is indicated in Fig. 14. In addition to the poles $\pm k_p$ on Sheet I, there are also poles at the same locations on Sheet IV, as well as zeros (sometimes referred to as Brewster zeros) on Sheets II and III [53, Sec. 15.B]. Note that, in view of the properties of k_{zn} illustrated in Fig. 12, the Sheet I pole k_p is located in the region of the k_ρ plane where $\Re e\, k_{z1} < 0$ and $\Re e\, k_{z2} > 0$. Consequently, at this pole the longitudinal propagation constants in the two half-spaces assume the values [37]

$$k_{z1}\Big|_{k_\rho = k_p} \equiv k_{z1p} = -\frac{k_p}{\sqrt{\epsilon}} = -\frac{k_1}{\sqrt{\epsilon+1}} \tag{65}$$

$$k_{z2}\Big|_{k_\rho = k_p} \equiv k_{z2p} = \sqrt{\epsilon}\, k_p = \frac{k_1 \epsilon}{\sqrt{\epsilon+1}} \tag{66}$$

which clearly satisfy the dispersion relation (63).[9]

The integration path C in Fig. 14 may be deformed around the hyperbolic branch cuts, capturing in this process the Sommerfeld pole k_p on Sheet I [54]. Consequently, in this representation a term $-2\pi j R_p$, where R_p is the residue of the

integrand of (58), should be included together with the two branch cut integrals along the paths C_1 and C_2. This residue may readily be found as

$$\begin{aligned} R_p &= \frac{j\epsilon^2}{\epsilon^2-1}\,\frac{k_{z1p}}{k_1}\,e^{-jk_{z1p}(z+h)}\,H_0^{(2)}(k_p\rho) \\ &\approx \frac{j\epsilon^2}{\epsilon^2-1}\,\frac{k_{z1p}}{k_1}\sqrt{\frac{2j}{\pi k_p\rho}}\,\varpi(k_p\rho)\,e^{-jk_{z1p}(z+h)}e^{-jk_p\rho} \end{aligned} \tag{67}$$

where in the second expression we have approximated the Hankel function using the large-argument form (60). The leading term of this formula is recognized as a cylindrical variant of the Zenneck wave [6], and has been variously referred to as Uller-Zenneck wave, Sommerfeld-Zenneck wave, or simply Zenneck wave. This wave propagates along the interface with a slow $1/\sqrt{\rho}$ amplitude decay due to the radial spread of energy, and decays exponentially in the transverse direction, since $\Im m\, k_{z1p} < 0$ on Sheet I. In view of the Sommerfeld pole location to the left of k_1 in the non-plasmonic case, the Zenneck wave propagates along the interface with the phase velocity greater than the speed of light and is thus termed a "fast" wave.[10] The path C in Fig. 14 can also be deformed around vertical lines emanating downwards from the fourth-quadrant branch points, in which case the Sommerfeld pole on Sheet I is not intercepted and does not explicitly contribute to the resulting field representation [1, p. 55], [38].

B. Modified saddle-point integration method

Consider the integral

$$I(\Omega) = \frac{e^{-j\Omega}}{\sqrt{\pi\Omega}}\int_{-\infty}^{\infty} G(s)\,e^{-\Omega s^2}\,ds \tag{68}$$

where $G(s)$ has a pole s_p near the origin with a residue r_p and $\Omega > 0$ is a large parameter. Applying the second-order additive modified saddle-point method [22, Ch. 4], [21, 25], we obtain the approximation

$$I(\Omega) \approx \left\{\left[G(0)+\frac{G''(0)}{4\Omega}\right]+\frac{r_p}{s_p}\left[F(p)+\frac{1}{2p}\right]\right\}\frac{e^{-j\Omega}}{\Omega} \tag{69}$$

where

$$p = \Omega s_p^2 \tag{70}$$

is a parameter related to the Sommerfeld numerical distance on which we elaborate in Appendix C, and

$$F(p) = 1 - j\sqrt{\pi p}\,w\big(-\sqrt{p}\big) \tag{71}$$

is the Sommerfeld-Norton attenuation factor [20, 55], here expressed in terms of the Faddeeva function [56], which is related to the complementary error function and is given as [57, Eq. 7.2.3]

$$w(z) = e^{-z^2}\,\mathrm{erfc}(-jz) = e^{-z^2}\,\frac{2}{\sqrt{\pi}}\int_{-jz}^{\infty} e^{-t^2}\,dt\,. \tag{72}$$

The attenuation function $F(p)$ has the absolutely-convergent ascending-power series representation [20, 55]

$$F(p) \approx 1 - j\sqrt{\pi p}\,e^{-p} - 2p + \frac{(2p)^2}{1\cdot 3} - \frac{(2p)^3}{1\cdot 3\cdot 5} + \cdots \tag{73}$$

and the asymptotic expansion

$$F(p) \sim -2j\sqrt{\pi p}\,e^{-p}\,\mathcal{U}\big(\Im m\, p\big) - \frac{1}{2p} - \frac{1\cdot 3}{(2p)^2} - \frac{1\cdot 3\cdot 5}{(2p)^3} - \cdots \tag{74}$$

where $\mathcal{U}$ is the Heaviside unit-step function. The field component associated with the first term in (74) has the characteristics of a trapped surface wave,[11] but is only present when $\Im m\, p > 0$, hence the expansion (74) has a different form depending on the argument of p, which is a property known as the Stokes phenomenon [24]. Although half of this term also appears in (73), it cannot be considered in isolation, since it is "swamped" by the other terms of the power series. The minimum value of p for which the asymptotic series (74) is applicable depends on the accuracy required, but we may use the condition $|p| \gtrapprox 10$ as a rule of thumb [23, 24]. The magnitude of the error will be bounded by the first term neglected in the asymptotic expansion. We note that the term $1/(2p)$ modifying $F(p)$ in (69) asymptotically cancels the leading term of (74) in the non-plasmonic case where the surafce wave is absent. This term arises from the second-order member in the Maclaurin series expansion of $G(s)$ after the pole has been subtracted out.

C. The Sommerfeld numerical distance

A parameter that plays a crucial role in Sommerfeld's asymptotic solution[12] is the numerical distance defined as [5]

$$\varrho = -j(k_1 - k_p)\rho \approx \frac{-jk_1\rho}{2\epsilon} \tag{75}$$

where k_p is the Sommerfeld pole (64), and where the approximation applies if $|\epsilon| \gg 1$. Strictly speaking, the term "numerical distance" ("Numerische Entfernung" in German) should be applied to $|\varrho|$, rather than to the complex-valued ϱ itself, but it has been a common practice not to make this distinction. Note that ϱ is a difference in the complex phase between the Zenneck wave and the space wave after having travelled along the interface a distance ρ from the dipole. Sommerfeld observed that the behavior of the asymptotic surface field was determined not by ρ or $k_1\rho$ alone, but by the numerical distance (75), which also includes the effect of the dielectric constants of the media. Consequently, the same values of ϱ result in the asymptotic fields of similar magnitude and nature, even if the radial distance ρ from the dipole is much different, which Sommerfeld dubbed the law of similarity of wireless telegraphy [58]. Furthermore, all circumstances which decrease ϱ, increase the received field strength. For example, the fact that ϱ is much smaller for seawater (due to its relatively high conductivity) than for fresh

water or dry ground (for the same distance ρ), explains the good radio reception at sea [59, p. 256].

If we were to extend the concept of Sommerfeld's numerical distance to a dipole and field points above the interface, we would replace ϱ by [23]

$$p = -j(k_1 - \boldsymbol{k}_p \cdot \hat{\boldsymbol{r}}_2) r_2 \tag{76}$$

where $\hat{\boldsymbol{r}}_2$ is the unit vector oriented from the dipole image location to the field point and $\boldsymbol{k}_p = \hat{\boldsymbol{\rho}} k_p + \hat{\boldsymbol{z}} k_{z1p}$ with k_{z1p} given in (65), is the wavevector of the Zenneck wave. Note that r_2 is the distance travelled by a wave from the dipole antenna to the receiver after undergoing a specular reflection from the interface at an angle θ_2, as illustrated in Fig. 1. Upon evaluating the dot product in (76) we obtain

$$p = -j\Omega_2 \left(1 - \frac{k_p}{k_1} \sin\theta_2 - \frac{k_{z1p}}{k_1} \cos\theta_2\right) \tag{77}$$

and thus our generalized numerical distance is exactly the parameter (34) which arises in the asymptotic field expression obtained by the modified saddle-point method. We note that (77) correctly reduces to Sommerfeld's ϱ in (75) for $\theta_2 = \pi/2$. In the high-contrast case, where $|\epsilon| \gg 1$, we may use in (34) the approximations [19]

$$\begin{aligned} p &\approx -j\Omega_2 \left(1 - \sin\theta_2 + \frac{\cos\theta_2}{\sqrt{\epsilon}} + \frac{\sin\theta_2}{2\epsilon}\right) \\ &\approx \frac{-j\Omega_2}{2} \left(\cos\theta_2 + \frac{1}{\sqrt{\epsilon}}\right)^2 \approx \varrho \left(1 + \sqrt{\epsilon}\, \frac{h+z}{\rho}\right)^2 \end{aligned} \tag{78}$$

where in the last two expressions we further assume small elevation angles, i.e., $(h+z)/\rho \ll 1$. The rightmost expression in (78) is thus a generalization of the original Sommerfeld's numerical distance (75) to the elevated case, but is only valid near the interface. This formula was first derived by Weyl [7] and later by Sommerfeld in his second paper [8].

To illustrate the dependence of p on the elevation angle and the media, we plot in Fig. 15 its normalized magnitude and phase angle vs. θ_2 for seawater, rich soil, and urban ground, computed at the operating frequency of 900 MHz. We compare in these plots the modified saddle-point numerical distance (77) and the Norton numerical distance (81). The plots for the three media are similar in character, with the argument of p in the negative range, which has important consequences for the appearance of the Zenneck wave in the asymptotic field, as discussed in Appendix E. The magnitude of p decreases as the elevation angle approaches the interface and at $\theta_2 = 90°$ it reaches much smaller values for seawater than for urban ground. These minimum values of $|p|$ are directly related to the distance of the Sommerfeld pole k_p from the branch point k_1, listed in Table 2. The Norton numerical distance is in close agreement with the modified saddle-point numerical distance for elevation angles near the interface, especially in the high-contrast seawater and rich soil cases, but it grows more rapidly in magnitude than the latter when θ_2 decreases, approaching infinity near the dipole axis.

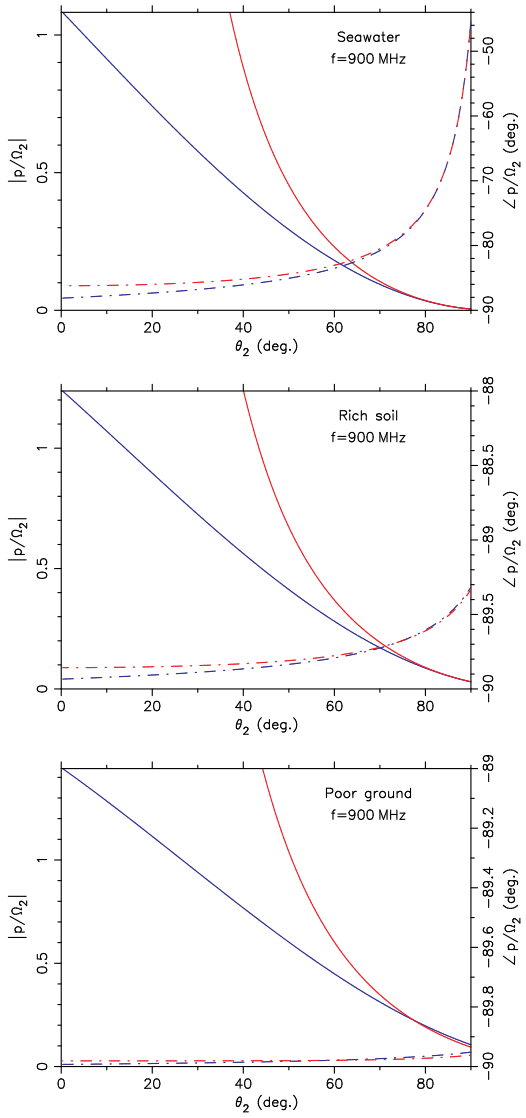

Figure 15: Plots of the normalized numerical distance p/Ω_2 vs. θ_2 for seawater, rich soil, and poor ground, computed at 900 MHz. The magnitude is indicated by full lines and the phase angle by dash-dot lines. The modified saddle-point numerical distance (77) is plotted in blue and the Norton numerical distance (81) in red.

D. The asymptotic formula of Norton

The asymptotic solution of Norton [11] may be expressed as [60, 61], [4, Sec. 3.7]

$$\mathcal{E} \approx \mathcal{E}_{GO} + \mathcal{E}_N \tag{79}$$

where $\mathcal{E}_{GO}$ is the geometrical-optics field introduced in (50) and $\mathcal{E}_N$ is the famous Norton surface wave, which is given as

$$\mathcal{E}_N = \sin^2\theta_2 \, [1 - \Gamma(\theta_2)] \, F(p) \, \frac{e^{-j\Omega_2}}{\Omega_2} \tag{80}$$

where $F(p)$ is the Sommerfeld-Norton attenuation factor introduced in (71) and p is the numerical distance parameter

given as

$$p = \frac{-j\Omega_2}{2\sin^2\theta_2}(\cos\theta_2 + \Delta)^2 . \tag{81}$$

Note that Γ and Δ are defined in (7). The asymptotic expression for E_{z1} obtained by Norton [11] had some errors [61] and also included additional second- and third-order direct and image dipole terms. Norton derived his "practical formula" from the van der Pol integral [12], and to arrive at (81) he replaced the term $1/\sqrt{\epsilon}$ with Δ, which was an ad-hoc step to make his solution consistent with that of Wise [13] for elevation angles away from the interface. We note that under the high-contrast approximation the Norton numerical distance (81) reduces to the second approximate expression in (78)—apart from the $\sin^2\theta_2$ term in the denominator, which was actually missing in the original Norton formulation, but was later restored in his wartime classified report [62].[13] The effect of this denominator term is to increase the numerical distance, and thus deemphasize the Norton surface wave, for large elevation angles.

E. Is there in the field of a transmitter a Zenneck wave?

This question already appeared in the title of a paper by Ott published 70 years ago [63]. The key to the answer lies in the first term of the asymptotic expansion (74) of the Sommerfeld-Norton attenuation factor $F(p)$, which appears in the Norton wave (31). To determine the contribution of this term to $\mathcal{E}_N^{(2)}$, we first note that

$$e^{-p}e^{-j\Omega_2} = e^{-jk_{z1p}(z+h)}\, e^{-jk_p\rho} \tag{82}$$

where we have used (77), and upon noting that $p = \Omega_2 s_p^2$ we obtain

$$-2\pi j\, \frac{r_p}{\sqrt{\pi\Omega_2}}\, e^{-jk_{z1p}(z+h)}\, e^{-jk_p\rho} \tag{83}$$

where r_p is given in (35)-(36). Note that apart from a factor $(k_p/k_1)^2$ accounting for the differential operator in (56), which converts A_{z1} into E_{z1}, the expression (83) is identical to $-2\pi j R_p$, where R_p is the Zenneck wave (67) discussed in Appendix A. However, according to (74), this wave only appears in the asymptotic field expression when $\Im m\, p > 0$. Therefore, let us examine the argument of $p = |p|e^{j\beta}$ with $\theta_2 = \pi/2$, which corresponds to the limiting SDP in Fig. 2. In this case (70) reduces to $p = \varrho = -jk_d\rho$, where $k_d \equiv k_1 - k_p = |k_d|e^{j\alpha}$, and thus $\beta = \alpha - \pi/2$ will be negative and $\Im m\, p < 0$, as long as $\alpha < \pi/2$. We now refer to Fig. 16, where k_1, k_p, and k_d are represented as vectors in the complex k_ρ plane, and which shows that the last condition is always satisfied for fast-wave poles, for which $\Re e\, k_p < k_1$. Consequently, the Zenneck wave (83) will *not* be present in the asymptotic field of the dipole over a non-plasmonic half-space with $\Re e\, \epsilon > 0$, consistent with our observation in Section 4 that a fast-wave pole located below the limiting SDP in Fig. 2 will never be captured. We should note here that in the absence of the surface wave term, the leading term of the asymptotic form of $\mathcal{F}(p)$ in (32) behaves as p^{-2}.

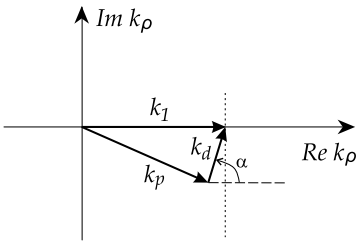

Figure 16: Demonstration that the argument of $k_d = k_1 - k_p$ is less than $\pi/2$ when the Sommerfeld pole k_p lies to the left of the vertical line passing through the branch point k_1.

The presence of the pole wave in the Norton surface wave can be tied to the character of Δ, which in the high-contrast approximation may be viewed as the surface impedance of the lower medium, normalized to η_0 [20, 61]. It follows from (81) that $\arg p = 2\arg\Delta - \pi/2$, and thus the pole wave will only occur for a "highly inductive" lower half-space, i.e., when the argument of Δ is in the range $\pi/4 \le \arg\Delta < \pi/2$. This condition holds for a plasmonic medium with $\varepsilon'_{r2} < 0$, in which case the pole wave contributes a surface plasmon polariton,[14] but it can also be satisfied when the surface is corrugated or is covered by an electrically thin low-loss dielectric overlayer, in which case the pole wave represents a "trapped" guided wave mode. The latter situation occurs, for example, when the ground is covered by asphalt or snowpack, or when water is overlayed by ice. Consequently, any surface roughness or medium inhomogeneity may significantly affect the near-ground propagation mechanism.

F. When is the Norton wave significant?

As noted in Appendix B, the asymptotic expansion of the Sommerfeld-Norton attenuation factor $F(p)$ is applicable for $|p| \gtrapprox 10$. Furthermore, as demonstrated in Appendix E, the surface wave term does not arise in (74) for the non-plasmonic media considered here. Therefore, under these conditions, the term $1/(2p)$ in the attenuation function $\mathcal{F}(p)$ will cancel the leading term of (74) and the leading term of (32) will be $-3/(2p)^2$. Consequently, in view of (70), the Norton wave $\mathcal{E}_N^{(2)}$ will be of the third order in Ω_2^{-1} and will become negligible for a sufficiently large p, as compared with $\mathcal{E}_R^{(2)}$ in (49), which behaves as Ω_2^{-2}. It may be argued that at the cut-off value of $|p|$, at which (31) becomes negligible, the term $3/|2p|^2$ should be a small fraction of $1/|2p|$, say 0.05, which leads to the condition $|p| \gtrapprox 30$, and further to[15]

$$r_2/\lambda_0 \gtrapprox \Omega_N \equiv \frac{15/\pi}{|s_p^2|} \tag{84}$$

where s_p^2 is given in (33). To illustrate how this criterion depends on the elevation angle and the media, we include in Fig. 17 the plots of Ω_N vs. θ_2 for the three media of Table 1

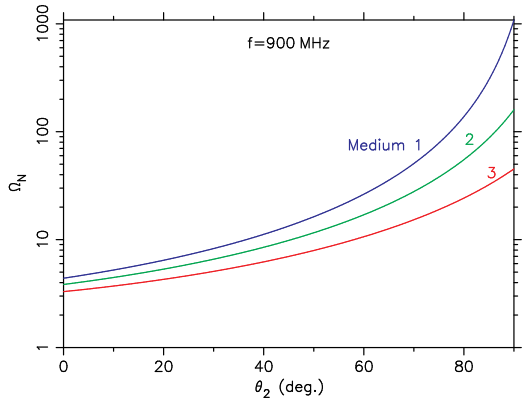

Figure 17: Plots of Ω_N vs. θ_2 for the three media of Table 1 at 900 MHz. The Norton wave $\mathcal{E}_N^{(2)}$ may be omitted if $r_2/\lambda_0 \gtrapprox \Omega_N$.

at the operating frequency of 900 MHz. The significance of these plots is that the Norton wave may be omitted whenever r_2/λ_0 falls above the line corresponding to the medium considered. We note that this requirement is most difficult to satisfy, i.e., the largest values of r_2/λ_0 are required, when θ_2 approaches 90°, which is the case of surface-to-surface propagation. Incidentally, in the latter case $|s_p^2| = |k_p - k_1|/k_1$, which is the normalized distance of the Sommerfeld pole k_p from the branch point k_1, and its values are listed in Table 2. In the on-surface configuration, where $\theta_2 = \pi/2$ and $r_2 = \rho$, the criterion (84) reduces to

$$\rho/\lambda_0 \gtrapprox \frac{15/\pi}{|k_p - k_1|/k_1} \approx \frac{30}{\pi}|\epsilon| \tag{85}$$

where the rightmost expression applies in the high-contrast case with $|\epsilon| \gg 1$.

To further investigate the contribution of the Norton wave to the near-ground asymptotic field, in Fig. 18 we plot the magnitude of the numerical distance and the corresponding magnitude of the attenuation function vs. ρ/λ_0 for the three media of Table 1 at the operating frequency of 900 MHz in the elevated configuration with $h+z = 10\lambda_0$. We note that the plots of $|p|$ for the three media are not monotonically increasing with ρ and have dips at some distance from the dipole. These dips, which are due to an interplay of the θ_2−dependent terms in (34), move farther away and get deeper for media with higher losses. For rich soil and poor ground the plots of $|p|$ remain above 30, and thus the contribution of the Norton wave should be negligible in the entire ρ range. For seawater, however, we may only neglect the Norton wave when ρ exceeds approximately 1000 wavelengths. Similar results for $h+z = 50\lambda_0$ are given in Fig. 19, where the numerical distance now remains above the $|p| = 30$ threshold for all three media. We may thus infer from these results that for a ground with low to moderate losses the Norton wave $\mathcal{E}^{(2)}$ may be omitted in the asymptotic field expression (49) when $h+z$ exceeds approximately 10 wavelengths. Although this conclusion is

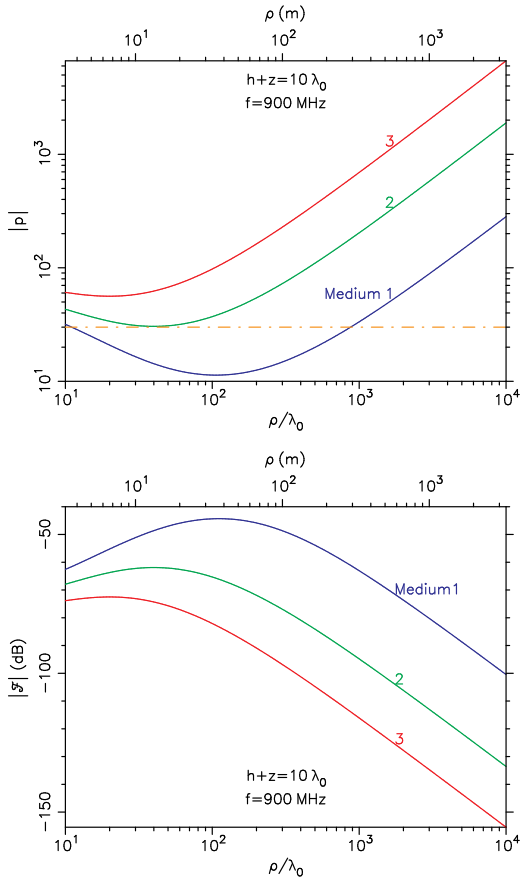

Figure 18: Plots of $|p|$ and $|\mathcal{F}|$ vs. ρ/λ_0 for the three media of Table 1 in the elevated configuration with $h+z = 10\lambda_0$ at 900 MHz. The $|p| = 30$ level is indicated by the dash-dot line.

based on numerical studies at 900 MHz, it should be applicable at higher frequencies as well, since the conductivity losses decrease with frequency, which results in an increase of the numerical distance p, and this in turn further deemphasizes the Norton wave.

Acknowledgment

The authors acknowledge with gratitude the assistance of Minyu Gu, who provided the Altair FEKO™ validation results.

Notes

[1]Norton's solution was later reformulated by Bannister [64], who also retained the usually omitted second- and third-order geometrical optics terms. In the early 1990s King [65, 66] developed a new solution for the Sommerfeld half-space problem, which however was later shown by Michalski and Jackson [67] to be equivalent to the Norton-Bannister formulation.

[2]The papers of Pauli and Ott commemorated the 70th and 75th anniversaries of Sommerfeld, respectively.

[3]Although Collin's solution was published in 2004, it enjoyed a limited distribution much earlier, in 1968.

[4]The last decade has also witnessed the introduction of a new "physics-based" solution of the Sommerfeld half-space problem [68], which has been shown to be invalid [37, 54, 69, 70], as a result of an incorrect application of the Schelkunoff integral identity [71]. Unfortunately, this criticism was

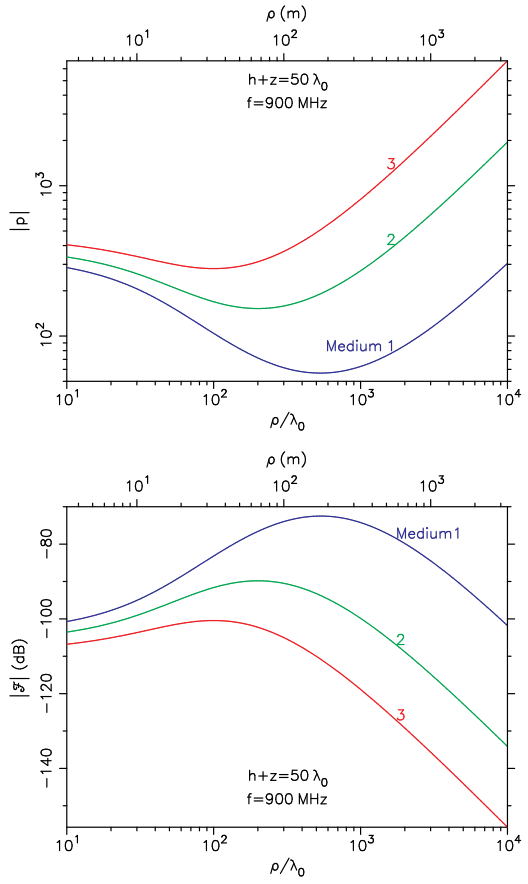

Figure 19: Plots of $|p|$ and $|\mathcal{F}|$ vs. ρ/λ_0 for the three media of Table 1 in the elevated configuration with $h+z=10\lambda_0$ at 900 MHz.

rejected by the authors [72], who continue to advocate their "Schelkunoff formulation" and claim that it allegedly provides "accurate results in certain limited regions" [73].

[5]The sign of the square root in (2) should be selected by the condition $\Im m\,\kappa(\xi) < 0$. In what follows, unless otherwise stated, the principal square-root branch with the positive real part is implied.

[6]It should be noted here that our usage of the term "lateral wave" is unconventional. In geophysics this term has a very specific meaning and it refers to the wave excited by a buried source, which propagates along the surface with the wavenumber of the air region. However, there has been a recent trend to also apply it to the surface waves excited by a dipole, which exhibit inverse-square range dependence, irrespective of the location of the source [3, 74, 75].

[7]It appears that in the relevant field plots in [32] the $\rho^{-1.5}$ asymptote has been arbitrarily inserted in the transitory range between the ρ^{-1} and ρ^{-2} asymptotes, where in fact the slope changes continuously. Furthermore, it is claimed that the $\rho^{-1.5}$ behavior is associated with the Norton surface wave, under the assumption that the numerical distance p is small in the intermediate ρ range. However, as we have demonstrated in Appendix F, $|p|$ is always much larger than unity in a typical elevated transmitter-receiver configuration with $(h+z) > 10\lambda_0$, and consequently the Norton wave is insignificant and cannot dominate the field behavior in this range.

[8]At the complex Brewster angle the reflection coefficient vanishes and there is no reflected wave. This Brewster zero is at the same location in the $k_\rho-$plane as the Sommerfeld pole k_p, except on Sheet II, as mentioned in Appendix A. Accordingly, both have exactly the same frequency dependence and an observation that a dip in the reflected field pattern is unaffected by even a large change of the frequency in a low-loss case should be unsurprising—and definitely does not indicate that the Zenneck wave corresponds to the zero of the reflection coefficient, rather than the pole, as claimed by in [73]. The Zenneck wave is of course absent in the asymptotic field of the dipole (see Appendix E), and thus cannot contribute to the radiation pattern. Yet, these authors state that "for a two-media problem, there is a Brewster zero and hence a Zenneck-type wave propagating over the surface where the evanescent field distribution is independent of frequency" and that "cellular wireless communication takes place through the Zenneck/Sommerfeld wave type," which they further amplify in a book section entitled "Cellular wireless propagation occurs through the Zenneck wave and not surface waves" [31]. It thus appears that, having associated the Zenneck wave with the zero of the reflection coefficient, and having observed a dip in the reflected field pattern at the Brewster's angle, these authors have incorrectly concluded that the near-ground propagation in wireless communication occurs through a Zenneck wave.

[9]Unless the correct choice of the square-root branches in k_{zn} is enforced, a numerical computation of the root of the dispersion function (63) may end in failure. For example, it has been reported [73, 76] that computer search can locate the pole k_p in the lossy case, but not in the lossless case, which is almost certainly due to the fact that the intrinsic square root function in Matlab or any computer language always returns a value with a positive real part, whereas $\Re e\, k_{z1} < 0$ should be selected at the real pole in the lossless case, according to (65). Hence, the Sommerfeld pole does not disappear in the lossless limit, but becomes real-valued and inconsequential, being squeezed between the two branch cuts in Fig. 14. The incorrect choice of the square-root branch was also made in some false "proofs" that the Sommereld pole did not exist [77, p. 431], [30, 78], and was probably responsible for placing the pole on the wrong Riemann sheet in [79, p. 99]. (Incidentally, a reprise of the Sommerfeld's proof [5] of the existence of the pole on the topmost Riemann sheet was recently given by Michalski and Mosig [54].) Furthermore, the alleged dilemma "whether the reflection coefficient has a zero or a pole" [73] does not in fact exist, since the reflection coefficient has both poles and zeros, albeit on different Riemann sheets, as already discussed above. Similarly, statements that "whether it will be a pole or a zero depends on the value of the dielectric constant" or that "zero may become a pole" for a complex dielectric constant [50] are questionable, since the value of the dielectric constant only affects the *location* of a pole (or zero).

[10]Since the topic of the Zenneck wave—including its very existence and physical reality—has been fraught with confusion, it is appropriate to unequivocally state here that the Sommerfeld pole *does* exist on the topmost sheet of the Riemann surface with hyperbolic branch cuts, even in the lossless case, and that Sommerfeld correctly included the Zenneck wave in his field representation, which served as the point of departure for the subsequent (unfortunately, flawed) asymptotic analysis [5]. However, the presence or not of the Zenneck wave in the field complex depends in the particular field representation, and there is no obligation from the mathematical point of view to choose one formulation over another. For example, there is certainly no Zenneck wave in the dipole field expressed as a Sommerfeld integral along the real axis in the $k_\rho-$plane, and no Zenneck wave appears when this path is deformed around vertical lines emanating downwards from the branch points, rather than around the hyperbolic branch cuts [38]. Furthermore, there is no Zenneck wave in the asymptotic field expression derived by the modified saddle-point method, as explained in Appendix E. This apparent paradox is easily explained: the appearance of a pole wave in the wave complex does not a priori mean that this wave will dominate the behavior of the *total* field of the dipole, and thus be physically meaningful. Although this *does* occur in the case of surface plasmon polariton, which is also associated with the Sommerfeld pole, except positioned differently with respect to the wavenumber of the upper half-space, the same does not hold for the Zenneck wave, which is asymptotically cancelled in the total field by the contribution from a branch cut integral. Theoretically, the Zenneck wave may be enhanced by an antenna or antenna array with a suitably tailored taper matching the transverse field distribution of the Sommerfeld pole wave [80–83], [52, Ch. 7], but the required aperture is too large to be realizable in practice, hence any attempts to experimentally detect the Zenneck wave over a smooth ground are doomed to fail. For this reason some authors call the Zenneck wave non-physical, which—considering that a great deal of confusion has already arisen regarding this wave—should probably be avoided, as this adjective can be misconstrued to imply that the Zenneck wave is contrary to some laws or fundamental postulates of physics. Indeed, there have been attempts to

discredit the Zenneck wave on the physical basis, by claims that it violates the radiation condition (which is irrelevant, since the radiation condition holds for the total Maxwellian field and not necessarily for its constituents) or that its group velocity exceeds the speed of light (which may well be true, however its energy velocity always remains below this physical limit [84]). There have been strenuous objections in the literature [50, 73] to calling the "fast" Zenneck wave a *surface* wave, reserving this term exclusively for the "slow" waves, such as a "trapped" mode guided by a dielectric slab. But, "What's in a name?" Note that the Zenneck wave propagates along the surface and decays exponentially in the transverse direction at a rate significantly greater than that in the radial direction, which are indisputably the necessary attributes of a surface wave, and it seems unnecessary to additionally impose the condition that the wave must be slow—a limitation which does not derive from any boundary conditions [85]. Furthermore, both the Zenneck wave and a trapped surface mode may be obtained as source-free solutions of the Maxwell equations at the pole of the reflection coefficient looking into the lower half-space, and thus both are pole waves and their fast or slow attribute is solely determined by the pole location with respect to the wavenumber of the external medium. Another argument used recently against calling the Zenneck wave a surface wave is that, whereas the latter is generated by the pole of the reflection coefficient Γ, the former supposedly arises from the zero of Γ [73]. However, this is a false dichotomy, since the Zenneck pole wave may *also* be synthesized by illuminating a lossy half-space with a plane wave incident at the (complex) Brewster angle [86], [87, Sec. 2.5], [88]. We should note here that the Sommerfeld pole, which gives rise to the Zenneck wave, is weakly dependent on the frequency in the case of a low-loss lower half-space medium, whereas the poles associated with trapped surface modes exhibit a strong variation as the frequency is changed. This explains why an increase in frequency has little effect on the transverse field distribution of the Zenneck wave—which remains loosely bound to the surface—but it greatly affects the transverse field of a trapped mode, which becomes more surface-bound as the electrical thickness of the guiding layer increases. These differences notwithstanding, whether the Zenneck wave is classified as a "true" surface wave or not is in our opinion just a matter of semantics.

[11] Wait [89] has suggested that the trapped wave component, which has the inverse-square-root range dependence, be called the Barlow surface wave, to distinguish it from the Norton surface wave, where the trapped wave is absent, and which has inverse-square range dependence, associated with the lateral wave.

[12] In his seminal paper, Sommerfeld [5] considered the case of surface-to-surface propagation and obtained an asymptotic solution, which unfortunately had an elusive but consequential error [24]. Although Sommerfeld later derived the correct expression using a different method [8, 58], a belated discovery by Norton [90] of the discrepancy between his two solutions inexplicably resulted in prolonged and recurring controversies in the literature, most of which have already been adequately addressed by Baños [1, Sec. 4.10], Wait [91], Collin [24], and Michalski and Mosig [54]. Incidentally, Collin [24] appears to have unwittingly created another controversy when he stated in his paper that the Sommerfeld sign error was "a myth." By this seemingly alarming announcement Collin meant that in his asymptotic solution Sommerfeld did not in fact err in the choice of the sign of the complex square root of the numerical distance parameter—as was the prevailing belief since the paper by Niessen [92], but he did not imply that there was no error. Yet, in the wake of Collin's paper, statements have appeared in the literature that "there was no error in Sommerfeld's work" [50] or that "Sommerfeld had no error in sign" [73]. In fact, there of course was an error in Sommerfeld's original paper, manifesting itself as a wrong sign in his asymptotic surface field expression: if this sign were to be changed to the opposite, a correct result would be obtained. Therefore, although Collin has shown that such a "fix" could not be justified on mathematical grounds, one would not be incorrect in saying that there *was* a sign error in Sommerfeld's work—and, paradoxically, in more than one sense. We are referring here to the seemingly never mentioned misprint in one of the final formulas in the Sommerfeld's 1909 paper, which mistakenly omitted the incorrect sign, thus making the formula appear correct! This is possibly why Sommerfeld in his 1926 paper could state that the (correct) surface field formula he obtained there agreed not only with Weyl's result, but also with his own expression derived previously. Although in his "myth paper" Collin debunked the prevailing notion as to the source of the error in Sommerfeld's asymptotic solution, he was not able to pinpoint the true source—nor was anyone else. (This is not meant as a criticism, since finding a *single* source of the error may well be an impossible task, in view of the numerous approximations made by Sommerfeld in his derivation, usually invoking the high-contrast assumption.) Consequently, although Collin's discovery is an important contribution to the history of the problem, the fact remains that the Sommerfeld's original asymptotic formula is incorrect and, furthermore, the correct asymptotic surface field expression has been available since at least the second paper of Sommerfeld—and has definitely been well known since the work of Norton.

[13] This report is no longer available in the US, but we have located a copy at the National Diet Library, Kyoto, Japan. It is interesting that the $\sin^2\theta_2$ denominator factor also appears in the numerical distance when the Norton surface wave is derived using an integral equation approach [61, 93], but was not included by Bannister [64], nor does it arise in the modified saddle-point method formula (34).

[14] The surface plasmon polariton and the Zenneck wave are thus closely related, as they represent two aspects of the same electromagnetic boundary-value problem [94]. Both waves are associated with the Sommerfeld pole on the top sheet of the Riemann surface and are classified as belonging to one or the other surface wave species depending on the pole location with respect to the wavenumber of the upper half-space—which in turn depends on the sign of $\Re e\,\epsilon$. Furthermore, it is possible for a "slow" surface plasmon to evolve into a "fast" Zenneck wave, and vice versa, with a change in frequency [42, 95, 96]. Yet, some authors [97] have mistakenly claimed that the Sommerfeld pole "does not exist" and thus surface plasmons have "no relation, whatsoever, with Sommerfeld surface waves!"

[15] The symbol Ω_N in (84), where the subscript N stands for Norton, is not related to the symbols Ω_1 and Ω_2 used elsewhere in this paper.

References

[1] A. Baños, *Dipole Radiation in the Presence of a Conducting Half-Space*. New York: Pergamon Press, 1966.

[2] L. M. Brekhovskikh, *Waves in Layered Media*, 2nd ed. New York: Academic Press, 1980.

[3] R. W. P. King, M. Owens, and T. T. Wu, *Lateral Electromagnetic Waves. Theory and Applications to Communications, Geophysical Exploration, and Remote Sensing*. New York: Springer-Verlag, 1992.

[4] T. S. M. Maclean and Z. Wu, *Radiowave Propagation Over Ground*. London: Chapman & Hall, 1993.

[5] A. Sommerfeld, "Über die Ausbreitung der Wellen in der drahtlosen Telegraphie," *Ann. Physik*, vol. 28, no. 4, pp. 665–736, 1909.

[6] J. Zenneck, "Über die Fortpflanzung ebener elektromagnetischer Wellen längs einer ebener Leiterfläche und ihre Beziehung zur drahtlosen Telegraphie," *Ann. Physik*, vol. 328, no. 10, pp. 846–866, 1907.

[7] H. Weyl, "Ausbreitung elektromagnetischer Wellen über einem ebenem Leiter," *Ann. Physik*, vol. 365, no. 21, pp. 481–500, 1919.

[8] A. Sommerfeld, "Über die Ausbreitung der Wellen in der drahtlosen Telegraphie," *Ann. Physik*, vol. 81, pp. 1135–1153, 1926.

[9] B. van der Pol and K. F. Niessen, "Über die Raumwellen von einem vertikalen Dipolsender auf ebener Erde," *Ann. Physik, 5. Folge*, vol. 402, no. 4, pp. 485–510, 1931.

[10] B. Kockel, "Die Sommerfeldsche Bodenwelle," *Ann. Physik, 7. Folge*, vol. 1, pp. 146–156, 1958.

[11] K. A. Norton, "The propagation of radio waves over the surface of the earth and in the upper atmosphere–Part II: The propagation from vertical, horizontal, and loop antennas over a plane earth of finite conductivity," *Proc. I.R.E.*, vol. 25, no. 9, pp. 1203–1237, 1937.

[12] B. van der Pol, "Theory of the reflection of the light from a point source by a finitely conducting flat mirror, with an application to radiotelegraphy," *Physica*, vol. 2, pp. 843–853, 1935.

[13] W. H. Wise, "Asymptotic dipole radiation formulas," *Bell Syst. Techn. J.*, vol. 8, no. 4, pp. 662–671, 1929.

[14] H. Ott, "Die Sattelpunktsmethode in der Umgebung eines Pols Mit Anwendungen auf die Wellenoptik und Akustik," *Ann. Physik, 5th Series*, vol. 43, no. 6/7, pp. 393–403, 1943.

[15] W. Pauli, "On asymptotic series for functions in the theory of diffraction of light," *Phys. Rev.*, vol. 54, no. 2, pp. 924–931, 1938.

[16] P. C. Clemmow, "Some extensions of the method of integration by steepest descents," *Quart. J. Mech. Appl. Math.*, vol. 3, Pt. 2, pp. 241–256, 1950.

[17] ——, "A method for the exact solution of a class of two-dimensional diffraction problems," *Proc. Roy. Soc. London, Ser. A*, vol. 205, no. 1081, pp. 286–308, 1951.

[18] ——, *The Plane Wave Spectrum Representation of Electromagnetic Fields*. New York: Pergamon Press, 1966.

[19] J. R. Wait, "Theory of electromagnetic surface waves over geological conductors," *Geofisica Pura e Applicata*, vol. 28, pp. 47–56, 1954.

[20] ——, "Excitation of surface waves on conducting, stratified, dielectric-clad, and corrugated surfaces," *J. Res. Natl. Bur. Stand.*, vol. 59, no. 6, pp. 365–377, 1957.

[21] B. L. Van der Waerden, "On the method of saddle points," *Appl. Sci. Res.*, vol. B2, pp. 33–45, 1951.

[22] L. B. Felsen and N. Marcuvitz, *Radiation and Scattering of Waves*. Englewood Cliffs, N.J.: Prentice Hall, 1973.

[23] G. D. Bernard and A. Ishimaru, "On complex waves," *Proc. IEE*, vol. 114, no. 1, pp. 43–49, 1967.

[24] R. E. Collin, "Hertzian dipole radiating over a lossy earth or sea: Some early and late 20th-century controversies," *IEEE Antennas Propag. Magaz.*, vol. 46, no. 2, pp. 64–79, 2004.

[25] K. A. Michalski and H. Lin, "On the Sommerfeld half-space problem: Appraisal of approximate solutions with extensions to plasmonics," *J. Electromagn. Waves Appl.*, vol. 32, no. 4, pp. 483–503, 2018.

[26] ——, "On the far-zone electromagnetic field of a vertical Hertzian dipole over an imperfectly conducting half-space with extensions to plasmonics," *Radio Sci.*, vol. 52, pp. 798–810, 2017.

[27] F. T. Dagefu and K. Sarabandi, "Analysis and modeling of near-groun wave propagtion in the presence of building walls," *IEEE Trans. Antennas Propagat.*, vol. 59, no. 6, pp. 2368–2378, 2011.

[28] M. Bellec, S. Avrillon, P. Y. Jezequel, S. Palud, F. Colombel, and P. Pouliguen, "Measurements of vertically polarized electromagnetic waves over a calm sea in HF band. Comparison to planar earth theories," *IEEE Trans. Antennas Propag.*, vol. 62, no. 7, pp. 3823–3828, 2014.

[29] S. F. Mahmoud and Y. M. M. Antar, "High frequency ground wave propagation," *IEEE Trans. Antennas Propag.*, vol. 62, no. 11, pp. 5841–5846, 2014.

[30] T. K. Sarkar, W. M. Dyab, M. N. Abdallah, M. Salazar-Palma, M. V. S. N. Prasad, S. W. Ting, and S. Barbin, "Electromagnetic macro modeling of propagation in mobile wireless communication: theory and experiment," *IEEE Antennas Propag. Magaz.*, vol. 54, no. 6, pp. 17–43, 2012.

[31] T. K. Sarkar, H. Chen, M. Salazar-Palma, and M.-D. Zhu, "Physics-based modeling of experimental data encountered in cellular wireless communication," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 6673–6682, 2018.

[32] T. K. Sarkar, M. Salazar-Palma, and M. N. Abdallah, *The Physics and Mathematics of Electromagnetic Wave Propagation in Cellular Wireless Communication*. Hoboken, NJ: Wiley, 2018.

[33] T. K. Sarkar, H. Chen, M. Salazar-Palma, and M. Zhu, "Lessons learned using physics-based macromodel for analysis of radio wave propagation in wireless transmission," *IEEE Trans. Antennas Propagat.*, vol. 67, no. 4, pp. 2150–2157, 2019.

[34] D. Penkin, G. Janssen, and A. Yarovoy, "Impact of a half-space interface on the wireless link between tiny sensor nodes," *Radio Sci.*, vol. 49, no. 9, pp. 798–811, 2014.

[35] N. Chahat, G. Valerio, M. Zhadobov, and R. Sauleau, "On-body propagation at 60 GHz," *IEEE Trans. Antennas Propagat.*, vol. 61, no. 4, pp. 1876–1888, 2013.

[36] M. Grimm and D. Manteuffel, "Norton surface waves in the scope of body area networks," *IEEE Trans. Antennas Propagat.*, vol. 62, no. 5, pp. 2616–2623, 2014.

[37] K. A. Michalski and J. R. Mosig, "The Sommerfeld half-space problem revisited: From radio frequencies and Zenneck waves to visible light and Fano modes (Invited)," *J. Electromagn. Waves Appl.*, vol. 30, no. 1, pp. 1–42, 2016, Erratum, ibid., vol. 31, no. 2, p. 241.

[38] K. A. Michalski and D. R. Nevels, "On the groundwave excited by a vertical Hertzian dipole over a planar conductor: Second-order asymptotic expansion with applications to plasmonics," *IEEE Trans. Microwave Theory*

Tech., vol. 65, no. 4, pp. 1133–1140, 2017, Erratum, ibid., vol. 67, no. 10, p. 4287.

[39] K. A. Michalski and H. Lin, "On the far-zone electromagnetic field of a horizontal electric dipole over an imperfectly conducting half-space with extensions to plasmonics," *Radio Sci.*, vol. 53, pp. 62–82, 2018.

[40] M. H. Bezerra Cardoso, S. Mostarshedi, G. Boudoin, and J. Laheurte, "Analytical expressions for critical distances for near-ground propagation," *IEEE Trans. Antennas Propag.*, vol. 66, no. 5, pp. 2482–2493, 2018.

[41] K. A. Michalski and J. R. Mosig, "On the complete radiation pattern of a vertical Hertzian dipole above a low-loss ground (Invited Paper)," *IEEE J. Microwaves*, vol. 1, no. 3, 2021.

[42] K. A. Michalski, "On the low-order partial-fraction fitting of dielectric functions at optical wavelengths," *IEEE. Trans. Antennas Propag.*, vol. 61, no. 12, pp. 6128–6135, 2013.

[43] A. Hessel, "General characteristics of traveling-wave antennas," in *Antenna Theory-Part 2*, R. E. Collin and F. J. Zucker, Eds. New York: McGraw-Hill, 1969, ch. 19, pp. 151–258.

[44] K. A. Michalski, "On the efficient evaluation of integrals arising in the Sommerfeld halfspace problem," in *Moment Methods in Antennas and Scatterers*, R. C. Hansen, Ed. Boston: Artech House, 1990, pp. 325–331, reprinted from *IEE Proc., Pt. H: Microwaves Optics and Antennas*, vol. 132, no. 5, pp. 213–218, 1985.

[45] J. R. Wait, "Characteristics of antennas over lossy earth," in *Antenna Theory-Part 2*, R. E. Collin and F. J. Zucker, Eds. New York: McGraw-Hill, 1969, ch. 23, pp. 386–437.

[46] J. D. Kraus and R. J. Marhefka, *Antennas for All Applications*, 3rd ed. New York: McGraw-Hill, 2002.

[47] K. A. Michalski, "Extrapolation methods for Sommerfeld integral tails (Invited)," *IEEE Trans. Antennas Propag.*, vol. 46, no. 10, pp. 1405–1418, 1998.

[48] K. A. Michalski and J. R. Mosig, "Efficient computation of Sommerfeld integral tails—methods and algorithms (Invited)," *J. Electromagn. Waves Appl.*, vol. 30, no. 3, pp. 281–317, 2016.

[49] A. Z. Elsherbeni, P. Nayeri, and C. J. Reddy, *Antenna Analysis and Design Using* FEKO *Electromagnetic Simulation Software*. Raleigh, NC: SciTech Publishing, 2014.

[50] T. K. Sarkar and M. Salazar-Palma, "The demise of the hundred year old mythology of the famous Sommerfeld sign error along with the realization of the Zenneck wave and its relationship with surface, lateral and leaky waves," *FERMAT*, vol. 31, pp. 1–44, Article 1, Jul.-Aug. 2019.

[51] K. A. Michalski and J. R. Mosig, "On the deficiency of the first-order Ott-Clemmow saddle-point method as applied to the Sommerfeld half-space problem," *IEEE Trans. Antennas Propagat.*, vol. 70, 2022, under review.

[52] J. R. Wait, *Wave Propagation Theory*. New York: Pergamon, 1981.

[53] A. Ishimaru, *Electromagnetic Wave Propagation, Radiation, and Scattering. From Fundamentals to Applications*, 2nd ed. Hoboken, NJ: Wiley, 2017.

[54] K. A. Michalski and J. R. Mosig, "The Sommerfeld halfspace problem redux: Alternative field representations, role of Zenneck and surface plasmon waves," *IEEE Trans. Antennas Propag.*, vol. 63, no. 12, pp. 5777–5790, 2015.

[55] L. E. Vogler, "A note on the attenuation function for propagation over flat layered ground," *IEEE Trans. Antennas Propag.*, vol. 12, no. 2, pp. 240–242, 1964.

[56] G. P. M. Poppe and C. M. J. Wijers, "More efficient computation of the complex error function," *ACM Trans. Math. Software*, vol. 16, no. 1, pp. 38–46, 1990.

[57] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, *NIST Handbook of Mathematical Functions*. Cambridge, UK: Cambridge University Press, 2010.

[58] A. Sommerfeld, "Drahtlose Telegraphie," in *Die Differential- and Integralgleichungen der Mechanik und Physik, Part II*, 2nd ed., P. Frank and R. von Mises, Eds. New York: Rosenberg, 1943, ch. 23, pp. 918–977.

[59] ——, *Partial Differential Equations in Physics*, ser. Lectures on Theoretical Physics, Vol. VI. New York: Academic Press, 1949.

[60] K. A. Norton, "The calculation of the ground-wave field intensity over a finitely conducting spherical earth," *Proc. IRE*, vol. 29, no. 12, pp. 623–639, 1941.

[61] R. J. King, "Electromagnetic wave propagation over a constant impedance plane," *Radio Sci.*, vol. 4, no. 3, pp. 225–268, 1969.

[62] K. A. Norton, "The polarization of downcoming ionospheric radio waves," Natl. Bur. Standards, Boulder, CO, FCC Rep. 60047, May 1942.

[63] H. Ott, "Gibt es im Feld eines Senders eine Zenneckwelle?" *Archiv elektr. Übertrag.*, vol. 5, pp. 15–24, 1951, with editorial comment from the publisher, K. W. Wagner.

[64] P. R. Bannister, "New formulas that extend Norton's farfield elementary dipole equations to the quasi-nearfield range," Naval Underwater Systems Center, New London, CT, Tech. Rep. 6883, Jan. 1984.

[65] R. W. P. King, "Electromagnetic field of a vertical dipole over an imperfectly conducting half-space," *Radio Sci.*, vol. 25, no. 2, pp. 149–160, 1990.

[66] R. W. P. King and S. S. Sandler, "The electromagnetic field of a vertical electric dipole over the earth or sea," *IEEE Trans. Antennas Propagat.*, vol. 42, no. 3, pp. 382–389, 1994.

[67] K. A. Michalski and D. R. Jackson, "Equivalence of the King and Norton-Bannister theories of dipole radiation over ground with extensions to plasmonics," *IEEE Trans. Antennas Propag.*, vol. 64, no. 12, pp. 5261–5261, 2016.

[68] W. M. Dyab, T. K. Sarkar, and M. Salazar-Palma, "A physics-based Green's function for analysis of vertical electric dipole radiation over an imperfect ground plane," *IEEE Trans. Antennas Propag.*, vol. 61, no. 8, pp. 4148–4157, 2013.

[69] R. Janaswamy, "Comments on 'A physics-based Green's function for analysis of vertical electric dipole radiation over an imperfect ground plane'," *IEEE Trans. Antennas Propag.*, vol. 62, pp. 4907–4910, Sep. 2014.

[70] ——, "Consistency requirements for integral representations of Grenn's functions—Part II: An erronous representation," *IEEE Trans. Antennas Propag.*, vol. 66, no. 8, pp. 4069–4076, 2018.

[71] S. A. Schelkunoff, "Modified Sommerfeld's integral and its applications," *Proc. IRE.*, vol. 24, no. 10, pp. 1388–1398, 1936.

[72] W. M. Dyab, T. K. Sarkar, and M. Salazar-Palma, "Reply to comments on 'A physics-based Green's function for analysis of vertical electric dipole radiation over an imperfect ground plane'," *IEEE Trans. Antennas Propag.*, vol. 62, no. 9, pp. 4910–4913, 2014.

[73] T. K. Sarkar, M. N. Abdallah, M. Salazar-Palma, and W. M. Dyab, "Surface plasmons/polaritons, surface waves, and Zenneck waves: Clarification of the terms and a description of the concepts and their evolution," *IEEE Antennas Propag. Magaz.*, vol. 59, no. 3, pp. 77–93, 2017.

[74] G. Gilbert, E. Raiten, and M. Visser, "Lateral wave contribution to the low-altitude radar propagation factor," *Radio Sci.*, vol. 29, no. 2, pp. 483–494, 1994.

[75] J. R. Wait, "Lateral waves and the pionieering research of the late Kenneth Norton (1907-1982)," *IEEE Antennas Propagat. Magaz.*, vol. 37, no. 6, pp. 104–105, 1995.

[76] M. N. Abdallah, D. Salama, T. K. Sarkar, and M. Salazar-Palma, "An expose of zenneck waves and surface plasmon polaritons," in *Digest IEEE MTT-S Int. Symp.*, Honolulu, HI, Jun. 2017, pp. 1–4.

[77] S. A. Schelkunoff, *Electromagnetic Waves*. New York: Van Nostrand, 1943.

[78] T. K. Sarkar, W. M. Dyab, M. N. Abdallah, M. Salazar-Palma, M. V. S. N. Prasad, S. Barbin, and S. W. Ting, "Physics of propagation in a cellular wireless communiction environment," *URSI Radio Sci. Bull.*, no. 343, pp. 5–21, Dec. 2012.

[79] W. C. Chew, *Waves and Fields in Inhomogeneous Media*. New York: IEEE Press, 1995.

[80] G. Goubeau, "Über die Zennecksche Bodenwelle," *Z. angew. Phys.*, vol. 3, no. 3/4, pp. 103–107, 1951.

[81] J. R. Wait and D. A. Hill, "Excitation of HF surface wave by vertical and horizontal antennas," *Radio Sci.*, vol. 14, no. 5, pp. 767–780, 1979.

[82] L. Petrillo, F. Jangal, M. Darces, J. Montmagnon, and M. Hélier, "Towards a better excitation of the surface wave," *Progr. Electromagn. Res. M*, vol. 13, pp. 17–28, 2010.

[83] F. Mesa and D. R. Jackson, "Excitation of the Zenneck wave by a tapered line source above the earth or ocean," *IEEE Trans. Antennas Propagat.*, vol. 68, no. 6, pp. 4848–4859, 2020.

[84] V. V. Shevchenko, "Field structure and velocities of surface waves on plane interfaces of conducting media with a relatively high conductivity, Zenneck wave," *J. Comm. Tech. Electron.*, vol. 60, no. 4, pp. 335–340, 2015.

[85] J. R. Wait, "Launching a surface wave over the earth," *Electron. Lett.*, vol. 3, no. 9, pp. 396–397, 1967, with a rebuttal by H. M. Barlow.

[86] H. M. Barlow and A. L. Cullen, "Surface waves," *Proc. IEE, Pt. III*, vol. 100, no. 68, pp. 329–341, 1953.

[87] G. Tyras, *Radiation and Propagation of Electromagnetic Waves*. New York: Academic Press, 1969.

[88] K. A. Michalski and J. R. Mosig, "Zenneck wave or surface plasmon? Sommerfeld pole or Brewster zero?" *IEEE Antennas Propagat. Magaz.*, vol. 64, 2021, in print.

[89] J. R. Wait, "A note on surface waves and ground waves," *IEEE Trans. Antennas Propag.*, vol. 13, no. 6, pp. 996–997, 1965.

[90] K. A. Norton, "Propagation of radio waves over a plane earth," *Nature*, vol. 135, no. 3423, pp. 954–955, 1935.

[91] J. R. Wait, "The ancient and modern history of EM ground-wave propagation," *IEEE Antennas Propag. Magaz.*, vol. 40, no. 5, pp. 7–24, 1998.

[92] K. F. Niessen, "Zur Entscheidung zwischen den beiden Sommerfeldschen Formeln für die Fortpflantzung von drahtlosen Telegraphie," *Ann. Physik, 5. Folge*, vol. 29, pp. 585–596, 1937.

[93] H. Green, "Derivation of the Norton surface wave using the compensation theorem," *IEEE Antennas Propag. Magaz.*, vol. 49, no. 6, pp. 47–57, 2007.

[94] A. Sihvola, J. Qi, and I. V. Lindell, "Bridging the gap between plasmonics and Zenneck wave," *IEEE Antennas Propag. Magaz.*, vol. 52, no. 1, pp. 124–136, 2010.

[95] M. Sarrazin and J.-P. Vigneron, "Light transmission assisted by Brewster-Zennek modes in chromium films carrying a subwavelength hole array," *Phys Rev. B*, vol. 71, pp. 075 404–1–5, 2005.

[96] K. A. Michalski and J. R. Mosig, "On the surface fields excited by a Hertzian dipole over a layered halfspace: From radio to optical wavelengths," *IEEE Trans. Antennas Propag.*, vol. 63, no. 12, pp. 5741–5752, 2015.

[97] W. M. G. Dyab, M. N. Abdallah, T. K. Sarkar, and M. Salazar-Palma, "On the relation between surface plasmons and Sommerfeld's surface electromagnetic waves," in *Digest IEEE MTT-S Int. Symp.*, Seattle, WA, Jun. 2013, pp. 1–4.