On the Diffraction of Spherical Radio Waves by a Finitely Conducting Spherical Earth

Lillie C. Walters and J. Ralph Johler

Contribution From Central Radio Propagation Laboratory, National Bureau of Standards, Boulder, Colo.

(Received August 3, 1961)

The theory for the diffraction of spherical electromagnetic waves by a finitely conducting spherical earth was developed from Maxwell's equations by Watson [1918] and the intricate computation details were later worked out by van der Pol and Bremer [1936] as the now classical series of residues. Two aspects of this computation present considerable difficulty, especially at low frequencies:

1. The calculation of the height-gain factor which takes account of an elevated transmitter and/or receiver.
2. The evaluation of the special roots, \( \tau = \tau_r \), of Riccati's differential equation,

\[
\frac{d\delta}{d\tau} - 2\delta\tau + 1 = 0,
\]

near the circle of convergence, \(|\delta\tau| = \frac{1}{2}\).

These analytic difficulties are avoided with the aid of modern analysis techniques applied to a large scale electronic computer. Hankel functions of the first and second kind of order one-third and two-thirds are calculated by numerical integral methods and then used with iteration to solve Riccati's differential equation. The amplitude and phase of the spherical radio wave diffracted in the vicinity of the earth with various altitudes above the surface of the earth, of both the transmitter and the receiver, are then calculated by a summation of the series of residues.

1. Introduction

The field of the diffracted spherical wave in the vicinity of a finitely conducting spherical earth has been discussed in detail as the ground wave by Watson [1918], van der Pol and Bremer [1937], and the particular case of a vertically polarized Hertzian dipole source, \( E_\theta \), volts/meter; i.e., the electric field directed radially with respect to the center of the earth, can be written [Johler, Kellar, and Walters, 1956],

\[
E_\theta = 2E_{\theta r} F_r,
\]

where,

\[
E_{\theta r} = \frac{10^{-7} \omega}{d} I_0 \exp \left\{ i [k d - \omega t] \right\},
\]

where, \( d \) is the distance, meters, \( f = \omega / 2\pi \) the frequency, cycles/second, \( I_0 \) is the dipole current moment, ampere-meters, \( \exp (-i\omega t) \) implies a wave varying harmonically in time, \( t \), in a medium with wave number, \( k = \frac{\omega}{c} \eta_i \), where \( c \) is the speed of light, \( c \sim 3 \times 10^8 \) m/sec, \( \eta_i \) is the index of refraction of air, \( \eta_i \sim 1 \). The attenuation function or secondary factor, \( F_r \), which takes account of the earth's curvature, finite conductivity, and dielectric constant, can be evaluated as a summation of the classical series of residues,

\[
F_r = \left[ 2\pi \alpha_2 \left( \frac{k_i a}{\alpha} \right)^{\frac{1}{2}} \sum_{a = 0}^{\infty} \frac{f_2(h_i) f_4(h_2)}{2 \tau - \frac{1}{\delta_i^2}} \right] \times \exp \left\{ \frac{i}{(k_i a)^{\frac{1}{2}} \tau \alpha_3 \frac{d}{a} + \frac{ad}{2a} + \frac{\pi}{4}} \right\},
\]

where \( a \) is the radius of the earth and \( \alpha \) is a factor which takes account of the vertical lapse of the permittivity of the earth's atmosphere.

The parameter, \( \delta = \delta_\alpha \) for vertical polarization is defined,

\[
\delta_\alpha = \frac{k_i^2}{k_1^2} \left[ \frac{k_i^2}{k_1} - 1 \right]^{\frac{1}{2}},
\]

with \( k_i = \frac{\omega}{c} \eta_i \), \( k_1 = \frac{\omega}{c} \eta_0 \), \( \sigma \) the conductivity of the medium, mhos/meter, \( \eta_0 \) the permeability of free space, henrys/meter, and \( \tau = \tau_r \), describes the special roots of Riccati's differential equation,

\[
\frac{d\delta}{d\tau} - 2\delta \tau + 1 = 0,
\]
which define the positions of the poles at which the residues for the series \( s = 0, 1, 2, 3, \ldots \) are calculated. In the case of horizontal polarization, \( E_r \) may be replaced by \( H_r \) (vertical magnetic field) providing \( \delta_e \) is replaced by \( \delta_m \), where

\[
\delta_m = \frac{k_e^2}{k_r^2} \delta_e. \tag{6}
\]

The isolated factors, \( f_s(h_1), f_s(h_2) \), which are in each term of the series describe the effect of elevating either the transmitter, \( h_1 \), or the receiver, \( h_2 \), or both, and can be evaluated without further approximations [Johler et al., 1956] as follows:

\[
f_s(h) = \left[ \frac{(k_d a')^{2/3}}{2 a} \frac{2 \omega a^{1/3}}{a - 2 \tau_s} \right]_{1/2} \frac{H_{1/3}^{(1)}}{H_{1/3}^{(2)}} \left\{ \frac{1}{3} \left[ (k_d a')^{2/3} \frac{2 \omega a^{1/3}}{a - 2 \tau_s} \right]^{3/2} \right\}.
\]

It is often more convenient to use some sort of modified Hankel function, \( h_1(z) \) [Furry, 1945], for which values of complex argument \( z = x + iy \) have been tabulated,

\[
f_s(h) = \frac{\left( \frac{1}{2} \right)^{2/3} \frac{2 \omega a^{1/3}}{a - 2 \tau_s} }{h_1(1 - \frac{1}{2^{1/3}})} \tag{8}
\]

The roots, \( \tau_s \), are found [Johler, Walters, Lilley, 1959] by expanding the Riccati differential equation in a power series,

\[
\tau_s = \sum_{n=0}^{\infty} \frac{1}{n!} \left[ \frac{d^n \tau_s}{d \delta_e^n} \right]_{\delta_e = 0} \delta_e^n, \tag{9}
\]

where the limiting roots, \( \tau_{s,0} \) and \( \tau_{s,\infty} \) for \( \delta_e = 0 \) and \( \delta_e = \infty \) have been tabulated [Johler, Walters, Lilley, 1959].

The power series in \( \delta_e \) and \( \delta_e^{-n} \) both seem to lose precision near the circle of convergence, \( |\delta_0| = \frac{1}{2} \), for a finite number of terms. Although a considerable number of terms have been determined [Johler, Kellar, Walters, 1956], the calculation of more terms becomes very laborious. Furthermore, the tabulation of the Hankel functions [Furry, 1945] does not seem to provide an adequate range of values, especially at low frequencies and indeed is not especially suitable for application to large scale computers.

Since the Riccati differential equation can be expressed as a ratio of Hankel functions,

\[
\frac{1}{\sqrt{-2 \tau_s}} \frac{H_{1/3}^{(1)}[(2 \tau_s)^{3/2}]}{H_{1/3}^{(2)}[(2 \tau_s)^{3/2}]} \exp \left[ i \frac{\pi}{3} \right] \delta_e = 0, \tag{11}
\]

a basic technique to evaluate Hankel functions \( H_{1/3}^{(1,2)}(z) \) would permit an iterative solution of Riccati's differential equation and at the same time could be employed to evaluate the Hankel functions \( H_{1/3}^{(1)}(z) \) in the height gain factor, \( f_s(h) \). The development of such a technique is the prime task of this paper.

### 2. Evaluation of Hankel Functions by Gaussian Quadrature

Bessel functions of order \( \nu \) are solutions of the differential equation,

\[
z^2 \frac{d^2 y}{dz^2} + z \frac{dy}{dz} + (z^2 - \nu^2) y = 0 \tag{12}
\]

where \( \nu \) and \( z \) are real or complex. The Bessel functions of the third kind, known as Hankel functions, are linearly independent solutions of eq (12) and are denoted by \( H_{\nu}^{(1)}(z) \) and \( H_{\nu}^{(2)}(z) \) where \( \nu = \frac{1}{2} \) or \( -\frac{1}{2} \) for the purpose of this discussion. Hankel functions have complex values for a real argument, but are real for \( \nu + i \) \( H_{\nu}^{(1)}(iy) \) and \( \nu - i \) \( H_{\nu}^{(2)}(-iy) \) when \( y \) is positive. These functions are important in applications because they are the only Bessel functions that vanish for an infinite complex argument, \( H_{\nu}^{(1)}(z) \) if the imaginary part is positive and \( H_{\nu}^{(2)}(z) \) if it is negative [Jahnke and Emde, 1945].

Integral representations for \( H_{\nu}^{(1)}(z) \) and \( H_{\nu}^{(2)}(z) \) are as follows [Jeffreys and Jeffreys, 1956]:

\[
H_{\nu}^{(1)}(z) = \frac{1}{\pi i} \int_{-\infty}^{\infty} \exp \left[ i \frac{z}{2} \left( \lambda - \frac{1}{\lambda} \right) \right] \frac{d\lambda}{\lambda^{\nu+i}}, \tag{13}
\]

\[
H_{\nu}^{(2)}(z) = \frac{1}{\pi i} \int_{-\infty}^{0} \exp \left[ i \frac{z}{2} \left( \lambda - \frac{1}{\lambda} \right) \right] \frac{d\lambda}{\lambda^{\nu+i}} \tag{14}
\]

where the expression of the limits in eq (13) means that the path is taken from \( 0 \) to \( -\infty \) by way of \( i \).

Integrals with real limits are derived by selecting a path as shown in figure 1.

**Figure 1. Contour of integration of Hankel functions, \( H_{\nu}^{(1)}(z) \).**
If $\lambda = u$, the part of the integral for $H_{-1}^{(1)}(z)$ from 0 to 1 becomes,
\[
\frac{1}{\pi i} \int_0^1 \exp \left[ \frac{z}{2} \left( u - \frac{1}{u} \right) \right] u^{-v-1} du.
\] (15)

On the semicircle from $+1$ to $-1$, if $\lambda = \exp(i\theta)$, this part is,
\[
\frac{1}{\pi i} \int_0^\pi \exp \left( iz \sin \theta \right) \exp (-v \pi i) d\theta.
\] (16)

From $-1$ to $-\infty$, if $\lambda = u^{-1} e^{i\pi}$, this part is,
\[
\frac{1}{\pi i} \int_0^1 \exp \left[ \frac{z}{2} \left( u - \frac{1}{u} \right) \right] u^{-v} \exp (-v \pi i) du.
\] (17)

and together,

\[
H_{-1}^{(1)}(z) = -\frac{1}{\pi} \int_0^\pi \exp \left( iz \sin \theta - v \pi i \right) d\theta
\] (18)

for $\text{Re}(z) > 0$.

As $H_{-1}^{(1)}(z) = [H_{-1}^{(2)}(z^*)]^* [\text{Erdelyi}, 1953]$.

\[
H_{-1}^{(2)}(z) = \frac{1}{\pi} \int_0^\pi \exp (-iz \sin \theta + v \pi i) d\theta
\]
\[
+ \frac{i}{\pi} \int_0^\pi \exp \left[ \frac{z}{2} \left( u - \frac{1}{u} \right) \right] (u^{-v-1} + u^{-v} \exp v \pi i) du
\]

for $\text{Re}(z) > 0$. Note that this method is applicable not only for $v = \frac{3}{2}$ and $\frac{7}{2}$, but also for complex $v$.

The integrals in (18) and (19) may be evaluated by Gaussian quadrature [Kopal, 1955]. In this method a finite integral is expressed as a sum:
\[
\int_a^b F(x) dx = \sum_{m=1}^M W_m F(y_m) + \epsilon(M)
\] (20)

where $\epsilon(M)$ is an error term which may, in general, be made arbitrarily small by increasing $M$ where $m = 1, 2, 3, \ldots M$.

\[
y_m = \frac{1}{2} [(b-a)x_m + (b+a)].
\] (21)

The $x_m$'s are the Gaussian abscissas and $M$ determines the number of values of $x$ to be used in the quadrature. The Gaussian weights and abscissas may be determined from the following:
\[
\int_{-1}^1 f(x) dx = \sum_{m=1}^M W_m f(x_m)
\] (22)

\[
W_m = \frac{1}{2} (b-a) H_m.
\] (23)

The $x_m$'s are the roots of the Legendre polynomials defined by
\[
\frac{d^m}{dx^m} (x^2 - 1)^m = 2^m m! P_m(x)
\] (24)

\[
P_0(x) = 1
\]
\[
P_1(x) = x
\]
\[
P_2(x) = \frac{3}{2} x^2 - \frac{1}{2}
\]
\[
P_3(x) = \frac{5}{2} x^3 - \frac{3}{2} x
\]
\[
P_4(x) = \frac{35}{8} x^4 - \frac{15}{4} x^2 + \frac{3}{8}
\]

... .

Polynomials of higher degree are determined by use of the recursion formula:
\[
(m+1)P_{m+1}(x) + mP_{m-1}(x) = (2m+1)xP_m(x).
\] (25)

Upon determination of the roots, the weight coefficients, $H_m$, of the corresponding quadrature formulas are evaluated as follows:
\[
H_m = \frac{2}{(1-x_m^2)[P'_m(x_m)]^2}
\] (26)

Forty-eight weights and abscissas were used in this quadrature [Davis and Rabinowitz, 1956] approximating the integrand in this case with a polynomial of ninety-fifth degree [Kopal, 1955].

Equations (18 and 19) are valid for $\text{Re}(z) > 0$, but the values of the Hankel functions in the second and third quadrants may be found from the equations:
\[
H_{-1}^{(1)}(z) = [H_{-1}^{(2)}(z^*)]^*,
\] (27)

\[
H_{-1}^{(1)}(z e^{i\pi}) = \frac{\sin (1-m)\pi}{\sin m\pi} H_{-1}^{(1)}(z)
\]
\[
- e^{v\pi i} \frac{\sin mv\pi}{\sin v\pi} H_{-1}^{(2)}(z),
\] (28)

and
\[
H_{-1}^{(2)}(z e^{i\pi}) = \frac{\sin (1+m)\pi}{\sin m\pi} H_{-1}^{(2)}(z)
\]
\[
+ e^{v\pi i} \frac{\sin mv\pi}{\sin v\pi} H_{-1}^{(1)}(z)
\] (29)

with $m = -1$ in this instance [Watson, 1948].

3. Method of Iteration

An iterative method with the above procedure for finding Hankel functions may be used in (11) to find the roots of (5). The iterative method described by Muller [1956], although it was developed to solve polynomial equations, may be used for any function which is analytic in the neighborhood of the roots.
The functions for $\tau$ are from (11):

$$F(\tau) = -H_{1/3}^{(1)}(-2\tau)^{3/2} \frac{i \pi}{3} \delta_{\sigma} |\delta_{c}| < 1$$

and

$$F(\tau) = -\sqrt{-2\tau} H_{1/3}^{(1)}(-2\tau)^{3/2} \frac{i \pi}{3} \frac{1}{\delta_{\sigma}} |\delta_{c}| > 1.$$  

(30)

Alternate expressions for $F(\tau)$ are used so that $\delta_{c}$ does not completely dominate the function and prevent convergence.

Each iteration in the process of finding one root is obtained from the calculation of the nearer root which passes through the last three points of the function, $f(\tau)$. The quadratic equation is, in general, complex; i.e., it possesses complex roots and complex coefficients. The function $f(\tau)$ has the same ordinates as the Lagrange interpolation formula through the three points $\tau_{i-2}$, $\tau_{i-1}$, $F(\tau_{i-2})$, and $\tau_{i}$, $F(\tau_{i})$ of $F(\tau)$ is,

$$f(\tau) = A_{3} \tau^{2} + A_{1} \tau + A_{0}$$  

(32)

where $A_{3}$, $A_{1}$, and $A_{0}$ are evaluated from the requirement:

$$f(\tau_{i+2}) = F(\tau_{i+2}), f(\tau_{i+1}) = F(\tau_{i+1}) \text{ and } f(\tau_{i}) = F(\tau_{i}).$$

Upon solving for $A_{3}$, $A_{1}$, $A_{0}$ and using the quantities $\beta = \tau - \tau_{0}$, $\beta_{i} = \tau_{i} - \tau_{i-1}$, $\beta_{i-1} = \tau_{i-1} - \tau_{i-2}$, $\lambda = \frac{\beta}{\beta_{i}} = \frac{\beta_{i}}{\beta_{i-1}}$, and $\delta_{i} = 1 + \lambda_{i}$ the Lagrange interpolation formula becomes a quadratic in $\lambda$:

$$f(\tau) = \lambda^{3} \delta_{i}^{3} \left[ F(\tau_{i-2}) \lambda_{i}^{2} - F(\tau_{i-1}) \lambda_{i} \delta_{i} ight] + F(\tau_{i}) \lambda_{i} + \lambda \delta_{i} \left[ F(\tau_{i-2}) \lambda_{i}^{2} - F(\tau_{i-1}) \delta_{i} \right] + F(\tau_{i}) \left( \lambda_{i} + \delta_{i} \right) \right] + F(\tau_{i}).$$

(33)

$\tau_{i+1}$ is found for the condition, $f(\tau) = 0$, solving for $\lambda$ and employing the relationship,

$$\lambda = \lambda_{i+1} = \frac{\tau_{i+1} - \tau}{\tau_{i} - \tau_{i-1}}.$$  

(34)

Rationalizing the numerator of the standard quadratic formula, and solving for $\lambda_{i+1}$,

$$\lambda = \frac{\sqrt{g_{i}^{2} - 4 F(\tau_{i}) \delta_{i}}}{2 F(\tau_{i}) \delta_{i}} \frac{F(\tau_{i-2}) \lambda_{i} - F(\tau_{i-1}) \delta_{i} + F(\tau_{i})}{g_{i} = F(\tau_{i-2}) \lambda_{i}^{2} - F(\tau_{i-1}) \delta_{i} + F(\tau_{i}) (\lambda_{i} + \delta_{i})}.$$  

(35)

where,

$$g_{i} = F(\tau_{i-2}) \lambda_{i}^{2} - F(\tau_{i-1}) \delta_{i}^{2} + F(\tau_{i})(\lambda_{i} + \delta_{i}).$$

(36)

The sign choice in the denominator of $\lambda_{i+1}$ (35) is resolved by selecting the value for the larger denominator. This choice of $\lambda_{i+1}$ gives $\tau_{i+1}$ the root of the Lagrange interpolation formula (32) which is nearer $\tau_{i}$, $\tau_{0}$, $\tau_{1}$, and $\tau_{2}$, the initial value of $\tau$ are estimated in this instance from the limiting roots $\tau_{s,0}(\delta_{\sigma} = 0)$ and $\tau_{s,\infty}(\delta_{\sigma} = \infty)$ for any desired $\delta$ [Johler et al., 1956], with the corresponding $F(\tau)$ evaluated from (30 or 31).

The actual values of $\tau_{s,0}$ and $\tau_{s,\infty}$ are used for $s < 4$ while approximations for other $s$ are determined from the following [Miller, 1946]:

$$\tau_{s,0} \sim \frac{\sqrt{g_{i}^{2} - 4 F(\tau_{i}) \delta_{i}}}{2} \left[ 1 + \frac{5}{48 y^{2} - 36 y} \right]$$

$$\tau_{s,\infty} \sim \frac{\sqrt{g_{i}^{2} - 4 F(\tau_{i}) \delta_{i}}}{2} \left[ 1 - \frac{7}{288 y^{2} + 35} \right]$$

(37)

(38)

where,

$$y_{1} = \frac{3\pi}{8} (4s + 3)$$

(39)

$$y_{2} = \frac{3\pi}{8} (4s + 1).$$

(40)

The iterative process is terminated with $\tau_{i}$, the desired root when,

$$\left| \frac{\tau_{i+1} - \tau_{i}}{\tau_{i}} \right| \leq \epsilon$$

(41)

where $\epsilon$ is a predetermined number.

With procedures developed for finding Hankel functions and $\tau$, $E_{r}$ may be calculated from (1) and (3).

4. Discussion of Results

In figures 2 and 4, describing effects of land and sea water, respectively, the transmitter on the surface, the receiver at varying heights above the surface in wavelengths, $\lambda = \frac{c}{f}$, and for any distances along the surface of the earth, the ratio of the field aloft relative to that at the surface ($h = 0$), $\frac{E(\omega, d, h)}{E(\omega, d)}$, approaches a value of one as the height of the receiver approaches zero. For increasing heights of the receiver, this ratio first exhibits a minimum less than one, although slight in some instances, and then increases in an exponential manner. The analytical behavior of the height-gain function has been discussed in considerable detail by Wait [1956].

Characteristics of the amplitude ratio described above, however, with some differences, are present if both transmitter and receiver are elevated (fig. 6). For a transmitter height of 1 wavelength and varying receiver height, for any distance along the surface of the earth, the amplitude ratio approaches a value less than 1 as the receiver height approaches zero. For a transmitter of height 10 wavelengths and varying receiver height, the amplitude ratio approaches values greater than one which varies with the distance along the surface as the receiver height approaches zero. The amplitude ratios for these elevated transmitters increase very rapidly as expected because there are two height gain factors increasing in an exponential manner.

For the phase relationships (figs. 3, 5, and 7) corresponding to the amplitude ratios (figs. 2, 4, and 6).
Figure 2. Amplitude of the vertical electric field, $E_r$, relative to surface value at various distances over land, with either transmitter or receiver elevated.

In this figure as well as those which follow $h_1$ and $h_2$ can be interchanged.

Figure 3. Phase of the vertical electric field, $E_r$, relative to surface value at various distances over land with either transmitter or receiver elevated.

respectively, the difference of the phase aloft and that at the surface ($h=0$), $t(\omega,d,h) - t(\omega,d) < 0$ with a very slight minimum at low increasing heights. To illustrate details of the phase in this region, a logarithmic scale was used with $t(\omega,d,h) - t(\omega,d) < 0$ graphed as lead with dashed lines, and $t(\omega,d,h) - t(\omega,d) > 0$ designated as lag with solid lines (figs.
3, 5). The phase lag may be ambiguous by 2π radians in any case, which in microseconds is 10, 1, or 0.1 for 100, 1,000, and 10,000 kc/s, respectively. In figure 7, the lower solid lines represent lead for the transmitter at 1 wavelength, while the upper solid lines represent lag for the same case. For 10-wavelength height of transmitter, the dotted lines represent lead.

5. Conclusions

All of the amplitude ratio curves have a minimum, even though slight in some cases, and then rise exponentially. For transmitter and receiver both elevated this ratio is very pronounced. The phase curves also exhibit a minimum, sometimes very slight, with \( t(\omega, d, h) - t(\omega, d) \) changing from a negative to a positive quantity or from a phase lead to a phase lag relative to the surface.

In applying the methods used in this paper, it may be concluded that the iterative method described is applicable for roots of Riccati’s differential eq(5) to any desired accuracy consistent with the capacity of the electronic computer used. Also, the method for evaluating Hankel functions, as it is applicable for real or complex order, may be used for the many other applications of these functions.

In particular, the height gain functions and the ground wave field can be evaluated at most any distance or altitude of transmitter and/or receiver with the aid of the techniques presented, the ultimate limitation being the computer capacity and speed, and the failure of the approximation in (3).

6. References

Johler, J. R., W. J. Kellar, and L. C. Walters, Phase of the low radiofrequency ground wave, NBS Circ. 573 (March 14, 1956).
van der Pol, B., and H. Bremmer, The diffraction of electromagnetic waves from an electrical point source round a finitely conducting sphere, with applications to radio-telegraphy and the theory of the rainbow, Naturwissenschaften, Circ. 52, 823 (Nov. 1934); Supp. J. Science 52, 817 (June 1938).
Wait, J. R., Radiation from a vertical antenna over a curved stratified ground, J. Research NBS 56, 237–244 (April 1956) RP2671.

(Paper 66D 1—177)