MEMORANDUM
RM-5558-NASA
MAY 1968

TRANSIONOSPHERIC PROPAGATION OF PULSED SIGNALS

W. Sollfrey

PREPARED FOR:
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

The RAND Corporation
SANTA MONICA • CALIFORNIA

May 1968 RB-5558

RM-5558-NASA, *Transionospheric Propagation of Pulsed Signals*, W. Sollfrey, RAND Memorandum, May 1968, 53 pp.

PURPOSE: To provide a procedure for computing the distortion of a pulsed signal as it is transmitted through the earth's ionosphere from a communication satellite.

RELATED TO: RAND's continuing study of communication satellite technology for NASA. The results of this study can be used in designing satellite systems which employ time-division multiplexing or other types of pulsed transmission.

THE PROBLEM: Several of the systems proposed for multiple access to communication satellites by earth stations use forms of pulse code modulation. For many, it is necessary to determine if a given signal is present at a given frequency at a specified time. However, the signal must pass through the ionosphere, where the nonlinear variation of the refractive index with frequency causes a dispersion, modifying the waveshape and spreading out the pulse in time. Thus the system may be confronted with the presence of signals in several frequency channels at a specified time. These effects limit the communication capability of the system and may determine the maximum permissible bandwidth or shortest permissible pulse length. Analyses have been made of the effects of dispersion on the transmission of pulsed signals, but so far none has been satisfactory. The exact solutions of the wave equations in terms of Bessel's functions converge too slowly to be suitable for numerical calculations. The approximate solutions using the quadratic phase distortion method break down when the deviation is comparable to the system bandwidth, which is just the condition of interest.

THE PROCEDURE: A theory of propagation in a generalized ionosphere is presented, assuming that the carrier frequency is sufficiently above the cutoff frequency so the ray paths are straight lines, that effects of the earth's curvature and magnetic fields can be neglected, and that only horizontal polarization need be considered. A new procedure is given which derives the "equivalent slab thickness" to represent the atmospheric electron density from an approximate solution of the wave equation instead of from an approximation to the equation itself. The equivalent slab parameters are determined from the integrals of the electron density and its square; however, quantities actually measured are the maximum electron density (measured by ionosondes) and the total electron content (measured by a Faraday rotation technique). After a review of the data, the experimental values are introduced into the calculation to derive a new approximation that makes use of the well-known method of stationary phase. In order to avoid the same convergence difficulties as arise in expansion about the carrier frequency, it is necessary to use a transformation of variables that is equivalent to expressing the frequency as a function of phase, instead of vice versa. Results are compared with the classical theory of Ginzburg and shown to be a nontrivial extension of the theory.

PRINCIPAL FINDINGS: The shortest acceptable pulse width for a pulsed communication system between earth and a satellite is equal to rise time and is proportional to the total electron content along the propagation path and inversely proportional to the three-halves power of the frequency. At 5 GHz carrier frequency, the typical range of shortest pulse width is 1 to 6 nanoseconds, permitting very high data rates. The overall displacement of the pulses might require a delay compensator to be built into each channel of a multichannel pulse code modulation system. Since the differential delay between channels is proportional to the electron density along the path, it may vary by a factor of 10 to 20 during the day, and a variable delay compensator may be required.

MW

MEMORANDUM
RM-5558-NASA
MAY 1968

TRANSIONOSPHERIC PROPAGATION OF PULSED SIGNALS

W. Sollfrey

This research is supported by the National Aeronautics and Space Administration under Contract No. NASr-21. RAND Memoranda are subject to critical review procedures at the research department and corporate levels. Views or conclusions expressed herein are nevertheless the primary responsibility of the author, and should not be interpreted as representing the official opinion or policy of NASA or of The RAND Corporation.

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA • 90406

Published by The RAND Corporation

PREFACE

This study, undertaken as part of RAND's continuing study of Communications Satellite Technology for the National Aeronautics and Space Administration, considers the distortion of a pulsed signal as it is transmitted through the earth's ionosphere.

The efficient use of the power available from satellites requires the use of wide-band transmission, corresponding to short pulses. These pulses may be distorted by the dispersive effects of the ionosphere, with the effects greater at the lower carrier frequencies and during periods of solar activity. This Memorandum studies the phenomena and provides a means for computing the resulting distortion. The results of the study can be used as an aid for designing satellite systems which employ time-division multiplexing or other types of pulse transmission systems.

SUMMARY

Signals transmitted between the ground and most earth satellites must pass through the ionosphere. The nonlinear variation of the refractive index with frequency causes a dispersion which can result in significant signal distortion. There have been previous theoretical investigations of the effects of dispersion on the transmission of pulsed signals, but they have suffered from various mathematical difficulties. Moreover, few experiments have been made to measure pulse distortion. Hence, the problem has been reconsidered in this Memorandum.

A theory of propagation in a generalized ionosphere is presented, and it is shown how an "equivalent slab" should be introduced. Experimental data were assembled from various sources to indicate the ranges of the relevant parameters. The integrals which arise in the theory are evaluated by a new method which does not suffer from the convergence difficulties which appeared in previous work. The solution depends on the group delay at the carrier frequency, the dispersion about the group delay, and other parameters. Explicit numerical values depend on the time of day, season, latitude, and position in the sunspot cycle. Results are computed and presented graphically for several typical ionospheric conditions with either rectangular or raised cosine pulses incident. These results extend the classical theory of Ginzburg.

The shortest acceptable pulse width is determined as a function of the parameters. It is shown to be proportional to the total electron content along the communication path and inversely proportional to the three-halves power of the frequency. At a carrier frequency of 5 GHz, the typical range of shortest pulse width is between 1 and 6 nanoseconds, permitting very high data rates.

CONTENTS

LIST OF FIGURES

I. INTRODUCTION

Several of the systems proposed for multiple access to communication satellites employ forms of pulse code modulation. These may involve time-division multiplexing or mixed frequency and time-division multiplexing (spread spectrum). Many of these systems involve determining whether a given signal is present at a given frequency at a specified time. The waveshape of the signal arriving at the receiver must be suitable for such analysis.

Signals transmitted in either direction between the earth and a communication satellite must pass through the ionosphere. The transmissions are generally at frequencies sufficiently high that the effects of ionospheric absorption and refraction are negligible. However, the transmission coefficient of the ionosphere is a function of frequency, so the medium displays dispersion. The effect will be to spread out the pulse in time and produce modifications of the waveshape. Furthermore, since the total delay of the signal in passing through the ionosphere is frequency-dependent, the system may be confronted with the presence of signals in several frequency channels at a specified time. These effects act to limit the communication capability of the system and may determine the maximum permissible bandwidth, or shortest permissible pulse length. These possibilities will be investigated in this Memorandum.

A considerable body of literature has been developed on the propagation of pulses in dispersive media. The applications have been both to the ionosphere and to waveguides operating near cutoff.

The distortion of a square wave propagating in a waveguide has been calculated by Elliott,[1] and the results of Ref. 1 have been applied to the ionosphere by Counter[2] and Dyce.[3] Reference 1 follows the procedure of expanding the phase of the transmission coefficient of the ionosphere in powers of the deviation from the carrier frequency, and then keeping only second-order terms. A parallel development is presented in Russian literature.[4,5] The method has been carried to third order by Gershman[6] and Wait.[7] The validity of Elliott's results has been questioned,[7,8] and it has been shown recently[9] that there is a mistake in integration in Ref. 1. The theoretical results of the other references cited here are correct. Furthermore, it has been demonstrated[5] that the problem can be solved exactly in terms of Lommel's functions of two variables, and several writers[10-12] have presented solutions in series of Bessel functions.

In view of this previous work, it may be asked if there is any point to additional work on the problem. The reasons for the theoretical research to be presented here may be summarized as follows. The exact solutions in series of Bessel functions converge quite slowly and are not suitable for numerical calculations. The approximate solutions using the quadratic phase distortion method are subject to the following anomaly: the expansion in powers of the deviation from the carrier frequency begins to converge slowly when the deviation is comparable to the system bandwidth, which is just the condition of interest. Furthermore, careful investigation shows

that the convergence rate depends on a parameter which is directly proportional to the carrier frequency and inversely proportional to the total electron content, and which may become on the order of unity in the gigahertz region. Hence, the problem has been reconsidered with a view to avoiding these difficulties.

A theory of propagation in a generalized ionosphere is presented in Section II. It is shown that the solution obtained satisfies a proper causality condition, which is not satisfied by the approximate solutions of Refs. 1 through 9. Section III contains a summary of the relevant experimental results. The integrals which arise in Section II are evaluated in Section IV by a method which does not suffer from the convergence difficulties indicated above. The results depend on the group delay at the carrier frequency, the dispersion about the group delay, and other parameters. Graphs are presented for several typical ionospheric conditions with either rectangular or raised cosine pulses incident. The shortest acceptable pulse width is determined as a function of the parameters.

The Appendix shows that for an ionospheric electron density profile which can be represented by an Epstein layer, the delta pulse propagation problem can be solved in closed form in Legendre functions. The analysis of this problem represents a nontrivial extension of the well-known solution in hypergeometric functions for the propagation of a sine wave through the Epstein layer.

II. PROPAGATION IN A GENERALIZED IONOSPHERE

The analysis begins with some general considerations of propagation in the ionosphere. Since the objective is communication through the ionosphere between the ground and the satellite, the carrier frequency must be above the ionospheric cutoff frequency, which means in practice above 40 MHz to take all factors into account. At such frequencies, ionospheric damping effects are clearly negligible. It will be assumed that the carrier frequency is sufficiently above the cutoff frequency so that the ray paths may be taken to be straight lines, and so that effects of the earth's curvature and magnetic field may also be neglected. Consideration of polarization effects will greatly complicate the analysis, so only horizontal polarization will be considered. With these approximations, the wave equation which describes the electromagnetic field E(z,t) becomes

$$\left[\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{e^2\mu N(z)}{m}\right] E(z,t) = 0 \qquad (1)$$

Here e and m are the electron charge and mass, μ the permeability of free space, and N(z) the electron density. MKS units are used throughout. The electron density is a slowly varying function of height, with a maximum at the F_2 layer and subsidiary maxima at lower altitudes. The characteristic lengths associated with appreciable variations of N(z) are measured in tens to hundreds of kilometers.

The field may be analyzed into its Fourier components, yielding

$$E(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e(z,\omega)e^{i\omega t}\, d\omega \qquad (2)$$

The transformed field $e(z,\omega)$ satisfies the equation

$$\left[\frac{\partial^2}{\partial z^2} + \frac{\omega^2}{c^2} - \frac{e^2\mu N(z)}{m}\right] e(z,\omega) = 0 \qquad (3)$$

with appropriate boundary conditions. Let the incident field $E_i(z,t)$ also be analyzed into Fourier components. Then the solution of Eq. (3) must satisfy the conditions

$$e(z,\omega) \rightarrow e_i(-\infty,\omega)\left[e^{-i\omega z/c} + R(\omega)e^{i\omega z/c}\right] \qquad z \rightarrow -\infty \qquad (4)$$

$$e(z,\omega) \rightarrow e_i(-\infty,\omega)T(\omega)e^{-i\omega z/c} \qquad z \rightarrow +\infty \qquad (5)$$

The experimental fact that the electron density N(z) tends to zero more rapidly than 1/z far above and far below the main part of the ionosphere has been used to simplify the boundary conditions. Equations (4) and (5) define the reflection coefficient $R(\omega)$ and the transmission coefficient $T(\omega)$. The latter is the quantity of principal interest.

There exist a few profiles N(z) for which Eq. (3) can be solved explicitly. These include the constant, linear, quadratic, exponential, and Epstein[13] layers. The solutions are all presented in Ref. 4, pp. 344-364. The application of the solution of the Epstein layer to pulse propagation is given in the Appendix to this Memorandum.

The simplest plausible approximation for the electron density N(z) is a slab of thickness H and constant density N_o. This "equivalent slab" theory has been extensively analyzed, since it is most closely akin to the waveguide transmission problem. However, the direct solution of Eq. (3) with constant N and the boundary conditions, Eqs. (4) and (5), leads to a transmission coefficient $T(\omega)$ whose amplitude is strongly dependent on frequency. This effect is caused by the spurious large gradient of electron density at the boundaries of the layer and by multiple internal reflections. A different procedure will be followed here,

which will derive an "equivalent slab" from an approximate solution of the equation, instead of from an approximation to the equation itself.

The procedure corresponds to a higher order WKB solution. Let K^2 be the coefficient of $e(z,\omega)$ in Eq. (3). At sufficiently high frequencies, ω^2/c^2 is large compared to $e^2\mu N(z)/m$, so K^2 never vanishes. Assume a solution of the form

$$e = \exp i \int^z \varphi(z')\ell z' \tag{6}$$

where the lower limit is left unspecified. Then the function $\varphi(z)$ satisfies the differential equation

$$\varphi^2 = K^2 + i\varphi' \tag{7}$$

where the prime denotes d/dz.

Under the physical conditions which exist in the ionosphere, the percentage variations of K are small at sufficiently high frequencies. The same may be assumed to hold for φ, whence the derivative term on the right side of Eq. (7) may be treated as a perturbation. Applying successive substitution to Eq. (7) to the second order produces the two solutions

$$\varphi_\pm = \pm K + \frac{iK'}{2K} \pm \left(\frac{3}{8}\ \frac{K'^2}{K^3} - \frac{K''}{4K^2}\right) \tag{8}$$

Substituting these expressions in Eq. (6) and performing some of the integrations yield the second-order WKB solution:

$$e_\pm \sim \sqrt{\frac{k}{K}} \exp \pm i\left[\int^z K\, dz' - \frac{1}{4}\ \frac{K'}{K} - \frac{1}{8}\int^z \frac{K'^2}{K^3}\, dz'\right] \tag{9}$$

where $k = \omega/c$. Solutions of the form of Eq. (9) must be fitted to the boundary conditions Eqs. (4) and (5). When the lower limit of the integrals is set equal to $-\infty$, and k is added and subtracted under the

integral sign to make the expressions converge, it follows that the approximate solution of the wave equation Eq. (3), which reduces properly to the incident wave at $-\infty$, is

$$e \sim e_i(-\infty,\omega)\sqrt{\frac{k}{K}}\left[\exp - i\left\{k\,z - \frac{1}{4}\frac{K'}{K} + \int_{-\infty}^{z} dz'\left(K - k - \frac{1}{8}\frac{K'^2}{K^3}\right)\right\}\right.$$

$$\left. + R \exp i\left\{k\,z - \frac{1}{4}\frac{K'}{K} + \int_{-\infty}^{z} dz'\left(K - k - \frac{1}{8}\frac{K'^2}{K^3}\right)\right\}\right] \qquad (10)$$

There is a corresponding solution which reduces to the transmitted wave at $+\infty$. The solutions may be matched at the point where $K' = 0$, which corresponds to the maximum electron density. The transmission coefficient may be determined from the matching conditions. Define $k_o^2 = e^2\mu/m$. The derivatives of K may be replaced by derivatives of N. After the matching procedure, the transmission coefficient may be expanded in inverse powers of frequency. Correct to inverse fifth powers (fifth order), the amplitude of the transmission coefficient is unity, and there results:

$$T(\omega) = \exp i\left[\int_{-\infty}^{\infty}\left\{k - \sqrt{k^2 - k_0^2\,N(z)}\right\} dz + \frac{1}{32}\frac{k_0^4}{k^5}\int_{-\infty}^{\infty}\left(\frac{dN}{dz}\right)^2 dz\right] \qquad (11)$$

The first integral in the exponent is the familiar WKB approximation. The second term, involving the square of the gradient of the electron density, is small when the physical problem is considered correctly. However, if the actual continuous variation of electron density is replaced by a slab, this integral becomes singular. Corresponding to this effect, the expansion of the Fresnel

transmission coefficient of a slab displays apparent fourth-order terms in the amplitude of $T(\omega)$. These are produced by the spatial singularities at the slab boundary where the electron density varies rapidly compared to the wavelength. Such configurations do not occur in the actual ionosphere. The proper definition of an equivalent slab should therefore be derived from Eq. (11) by matching the parameters in the expansion of the phase of $T(\omega)$ in inverse powers of frequency. There results for the equivalent slab

$$N_0 = \int_{-\infty}^{\infty} N^2(z)dz \Big/ \int_{-\infty}^{\infty} N(z)dz \tag{12}$$

$$H = \left[\int_{-\infty}^{\infty} N(z)dz\right]^2 \Big/ \int_{-\infty}^{\infty} N^2(z)dz \tag{13}$$

Quantities which may be actually measured are the electron density at the maximum N_m and the total electron content

$$N_t = \int_{-\infty}^{\infty} N(z)dz \tag{14}$$

The cutoff frequency of the ionosphere is given by

$$f_c(\text{Hz}) = 9[N_m(e/m^3)]^{\frac{1}{2}} \tag{15}$$

Define $\omega_c = 2\pi f_c$, $\tau = H/c$. The experimental conditions are generally such that inverse fifth powers of the ratio ω_c/ω may be neglected. Then the transmission coefficient $T(\omega)$ may be represented by the form

$$T(\omega) = \exp\left\{i\tau\left[\omega - \sqrt{\omega^2 - \omega_c^2}\right]\right\} \tag{16}$$

Previous investigations of ionospheric pulse propagation have begun with this expression, which is the transmission coefficient of a slab of electrons. The investigation to this point in this Memorandum has shown the conditions under which the transmission coefficient of the wave propagating through an arbitrary electron distribution may be replaced by the transmission coefficient of a slab.

In terms of the transmission coefficient, there results for the original time-dependent field

$$E(z,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int d\omega\, dt' E_i(-\infty,t')T(\omega)e^{i\omega(t-t'-z/c)} \qquad z \to \infty \tag{17}$$

or, if time is measured from the arrival of the first disturbance at z, $t_1 = t - z/c$

$$E(t_1) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int d\omega\, dt' E_1(-\infty,t')T(\omega)e^{i\omega(t_1-t')} \tag{18}$$

When the incident wave is a rectangular pulse, it is expressed as

$$E_i(-\infty,t') = \sin \omega_o t' \, [1(t')-1(t'-T)] \tag{19}$$

Here $\omega_o = 2\pi f_o$; f_o is the carrier frequency, T is the pulse width, and $1(t')$ denotes the Heaviside unit function, which is zero for t' negative, and 1 for t' positive. To avoid having discontinuities in E_i, assume that f_oT is an integer. Since f_oT is usually a large number, this is not a significant restriction.

With E_i in the form provided by Eq. (19), the integration over t' may be performed. The result will be given a subscript p for pulse solution, and the dependence on the carrier frequency f_o will be indicated explicitly. Thus

$$E_p(f_o,t_1) = \frac{1}{2\pi} \int_{-\infty}^{\infty} d\omega \; T(\omega) \left[e^{i\omega t_1} - e^{i\omega(t_1-T)} \right] \omega_o \Big/ \left(\omega_o^2-\omega^2\right) \tag{20}$$

The integrand has no singularity at $\omega = \omega_o$. However, it is desired to separate the two exponential terms, each of which then represents a singular integral. The integration path must be deformed into the complex plane to avoid these singularities. Furthermore, to maintain causality, the integral should vanish for $t_1 < 0$. This requires that the singularities be evaded by arcs passing below them. This leads to the representation

$$E_p(f_o,t_1) = E_s(f_o,t_1) - E_s(f_o,t_i-T) \tag{21}$$

$$E_s(f_o,t_1) = \frac{1}{2\pi} \int_{-\infty-i\epsilon}^{\infty-i\epsilon} d\omega \; T(\omega) e^{i\omega t_1} \, \omega_o \Big/ \left(\omega_o^2 - \omega^2\right) \tag{22}$$

where E_s represents the response to a carrier frequency step-function, and ϵ is a small positive number, which ensures that the singularities are avoided in a proper manner. The phase of the square root in Eq. (16) is to be interpreted as follows:

$$\begin{aligned} \arg\sqrt{\omega^2 - \omega_o^2} &= 0 \qquad \omega > \omega_c \\ &= -\frac{\pi}{2} \quad |\omega| < \omega_c \\ &= -\pi \qquad \omega < -\omega_c \end{aligned} \tag{23}$$

which ensures that the frequency components below cutoff are attenuated.

If the input waveform is a raised cosine function, it may be written in the form

$$E_i(-\infty, t) = \frac{1}{2}\left(1-\cos\frac{2\pi t}{T}\right)\sin\omega_o t[1(t)-1(t-T)] \tag{24}$$

Denoting the response by E_c, the relationship between rectangular and raised cosine pulse is obtained

$$E_c(f_o,t_1) = \frac{1}{2}E_p(f_o,t_1) - \frac{1}{4}\left\{E_p\left(f_o+\frac{1}{T},t_1\right)+E_p\left(f_o-\frac{1}{T},t_1\right)\right\} \tag{25}$$

This equation explains why the carrier frequency was explicitly indicated in the pulse response. The problem is to evaluate $E_s(f_o,t_1)$ when $T(\omega)$ has the form of Eq. (16). Although the integral can be evaluated in closed form in terms of Lommel's functions,[5] the result is not useful for computation. Approximations to the integral in Eq. (22) must be consistent with experimental results. Accordingly, a summary of relevant measurements of ionospheric properties will now be presented.

III. SUMMARY OF EXPERIMENTAL RESULTS

According to Eqs. (12) and (13), the equivalent slab parameters N_o and H are to be determined from the integrals of the electron density and its square. However, the quantities actually measured are N_m, the electron density at the maximum, and N_t, the total electron content. The first is measured by ionosondes, the second by a technique involving Faraday rotation which has been developed extensively in recent years. The electron density below the maximum may also be measured. The integral involving N^2 should therefore be eliminated in terms of measurable quantities. To do this requires specific assumptions about the shape of the ionospheric layer.

Three forms which have been used for the ionosphere are the parabolic, Epstein, and Chapman distributions. These lead respectively to the results

$$\begin{aligned} &\text{Parabolic:} && N_0 = .8\ N_m; && H = 1.25\ N_t/N_m \\ &\text{Epstein:} && N_o = .667\ N_m; && H = 1.5\ N_t/N_m \\ &\text{Chapman:} && N_0 = .658\ N_m; && H = 1.52\ N_t/N_m \end{aligned} \tag{26}$$

The parabolic and Epstein layers have equal numbers of electrons above and below the maximum, while the Chapman layer displays a top to bottom ratio of 2.15. Experimental values of the ratio display a wide range. Thus, Alouette sounder data[14] display a ratio near unity at midday, and between 2 and 3 in early morning and evening. Lunar radar observations in England[15] give a daytime ratio of

2.5 to 3.5, while at night the ratio is between 3 and 9. These results are for solar maximum to average conditions, and are reasonably corroborated for solar minimum.[16] If the profile is taken to be a Chapman layer up to one scale height above the maximum, and an exponential at higher altitudes, the top-to-bottom ratio can be used to determine the constants of the Chapman and exponential functions. At ratios of 4, 6, and 9, the coefficient .658 in Eq. (16) is reduced to .553, .512, and .482, respectively. Experimentally, ratios over 6 occur rarely. In view of these considerations, the representative values

$$N_0 = .625\ N_m, \quad H = 1.6\ N_t/N_m \tag{27}$$

will be used in calculations. The properties of the solution to the pulse propagation problem principally depend upon N_t, the total electron content, and the selection of the constant indicated in Eq. (17) is not critical.

The measured values of the major parameters N_m and N_t depend upon time of day, season, latitude, and position in the sunspot cycle, and they may display rapid fluctuations. Instead of N_m, the ionospheric cutoff frequency f_c defined in Eq. (15) will be presented. For vertical incidence, f_c may be as low as 2 MHz, and as high as 15 MHz.

Data on total electron content usually give the total content itself N_t and an equivalent slab thickness $H_s = N_t/N_m$. There have been very many measurements in recent years, and the following presentation does not pretend to be exhaustive.

Experimental values of the total electron content N_t generally lie in the range 10^{16} - 10^{18} electrons per square meter. Thus, it

proves advantageous to the presentation to introduce a parameter I by the relation

$$I = 10^{-16} N_t \ (e/m^2) \tag{28}$$

Measurements of the total electron content under solar maximum conditions are relatively few, since the experimental techniques were not sufficiently developed during the period of the last solar maximum (1958-60). Methods include measurement of the Faraday rotation of the signals from satellites, lunar radar, and ionospheric backscatter. Data from Faraday rotation of the signals from Sputnik 3 between September 1958 and December 1959 are presented by Lawrence et al.[17] Marked diurnal and seasonal variations appear in the measurements. Thus, at a local time just before sunrise, I has a mean value of 8 throughout the series of observations, but drops as low as 2.5 on some winter measurements. The largest single value of I, observed in the early afternoon during the spring of 1959, was about 160. A smooth curve drawn through the data shows midday peak values as 80 (Fall 1958), 70 (Winter 1958), 90 (Spring 1959), 35 (Summer 1959), 35 (Fall 1959), and 55 (Winter 1959). The sunspot number decreased from about 180 to about 130 during the observation period, which at least partly accounts for the drop between comparable seasons in the two years. Thus, the daytime values display a strong seasonal variation, with a maximum at the equinoxes and winter, while the nighttime values of total electron content are much more nearly constant.

The equivalent slab "thickness" $H_s = N_t/N_m$ displays much less variation. Reference 17 presents data for this quantity as an equivalent Chapman layer. Converting to slab thickness shows a range from 350 km (winter night) to 500 km (summer day and night) during the period September 1958 to September 1959.

Measurements during 1960 and 1961 have been reported by Roger,[18] Taylor,[19] and Millman.[20] The measurements show that the electron content maximizes in the afternoon, with a strong, rather narrow peak in winter, and reaches a broad but not so high maximum in summer. Values of electron content at 3 p.m. in winter are 40,[18] and 70,[20] while in summer they are 18,[18] 25,[9] and 64.[20] The much larger values reported in Ref. 20 are associated with the low latitude (Trinidad) of the observations. The slab thickness H_s at 3 p.m. throughout the year is between 250 and 300 km,[17] compared to 250-350 km.[20] Slab thicknesses from 200 to 500 km during the winter day, 200 to 950 km during the winter night, and 350 km during the summer day were reported in Ref. 15, while Ref. 20 gives comparable daytime data and values up to 700 km at night in both summer and winter. Assembling these assorted results, representative values for the solar maximum conditions will be chosen as:

Summer Day	$I = 35$	$H_s = 350$ km	$f_c = 9$ MHz
Summer Night	$I = 10$	$H_s = 500$ km	$f_c = 4$ MHz
Winter Day	$I = 80$	$H_s = 400$ km	$f_c = 12.7$ MHz
Winter Night	$I = 8$	$H_s = 500$ km	$f_c = 3.6$ MHz

Measurements have been performed during the declining portions of the solar cycle[21-25] and under solar minimum conditions.[16,26-30] The data display very good correlation among the various experimenters. A set of representative values for solar minimum conditions at mid-latitudes will be chosen as:

Summer Day	$I = 12$	$H_s = 300$ km	$f_c = 5.7$ Mc
Summer Night	$I = 3$	$H_s = 150$ km	$f_c = 4$ Mc
Winter Day	$I = 18$	$H_s = 240$ km	$f_c = 7.7$ Mc
Winter Night	$I = 3$	$H_s = 150$ km	$f_c = 4$ Mc

At low latitudes,[26] the nighttime thickness is considerably higher, while in the Southern Hemisphere[29] the winter night values of electron content are much higher. Values of the parameters at intermediate parts of the solar cycle lie between the values at maximum and minimum conditions.

The data as presented show a diurnal range in electron content of 3-4 to 1 in summer, 6-10 to 1 in winter. The nighttime values have little seasonal effect but show a 3 to 1 variation over the solar cycle. The thickness parameter H_s has a moderate (2 to 1) diurnal variation with little seasonal variation. The daytime thickness parameter varies less with the solar cycle than does the nighttime thickness.

The theoretical expressions developed in Section II will now be simplified using these experimental results.

IV. DETAILED CONSIDERATIONS OF PULSE PROPAGATION

When the form of Eq. (16) for the transmission coefficient $T(\omega)$ is inserted in the integral of Eq. (22), there results:

$$E_s(f_o,t) \equiv \frac{1}{2\pi} \int_{-\infty-i\epsilon}^{\infty-i\epsilon} \frac{\omega_o^2 d\omega}{\omega_o^2-\omega^2} \exp\left[i\omega t + i\tau\left\{\omega - \sqrt{\omega^2 - \omega_c^2}\right\}\right] \tag{29}$$

The standard procedure for evaluating integrals of the general type, Eq. (18), is to expand $-\varphi(\omega)$, the phase of the transmission coefficient $T(\omega)$, in powers of the deviation from the carrier frequency and assume that two or three terms are sufficient. Following the notation of Ginzburg,[4] the result becomes

$$E_s(f_o,t) \sim \frac{1-i}{2}\left[\frac{1+i}{2} + F\left(\frac{\theta}{\sqrt{\pi\varphi''(\omega_o)}}\right)\right]\sin(\omega_o t - \varphi(\omega_o)) \tag{30}$$

$$\theta = t - \varphi'(\omega_o) \tag{31}$$

$$F(u) = \int_0^u e^{i\pi/2\, u^2}\, du = C(u) + iS(u) \tag{32}$$

Here $\varphi(\omega_o)$ is the phase delay at the carrier frequency, $\varphi'(\omega_o)$ is the group time delay at the carrier frequency, and $\varphi''(\omega_o)$ is the second derivative of the phase with respect to the radian frequency ω, evaluated at the carrier frequency. The amplitude of E_s is plotted on page 417 of Ref. 4 and is reproduced in Fig. 1. In terms of the variable u, the function $|E_s(u)|$ is small for $u < -4$, reaches .5 at $u = 0$, reaches a first maximum value of 1.17 at $u = 1.25$, oscillates rapidly thereafter, and remains within 5 percent of unity for $u > 4$. Reference 4 defines the establishment time of the signal as the period

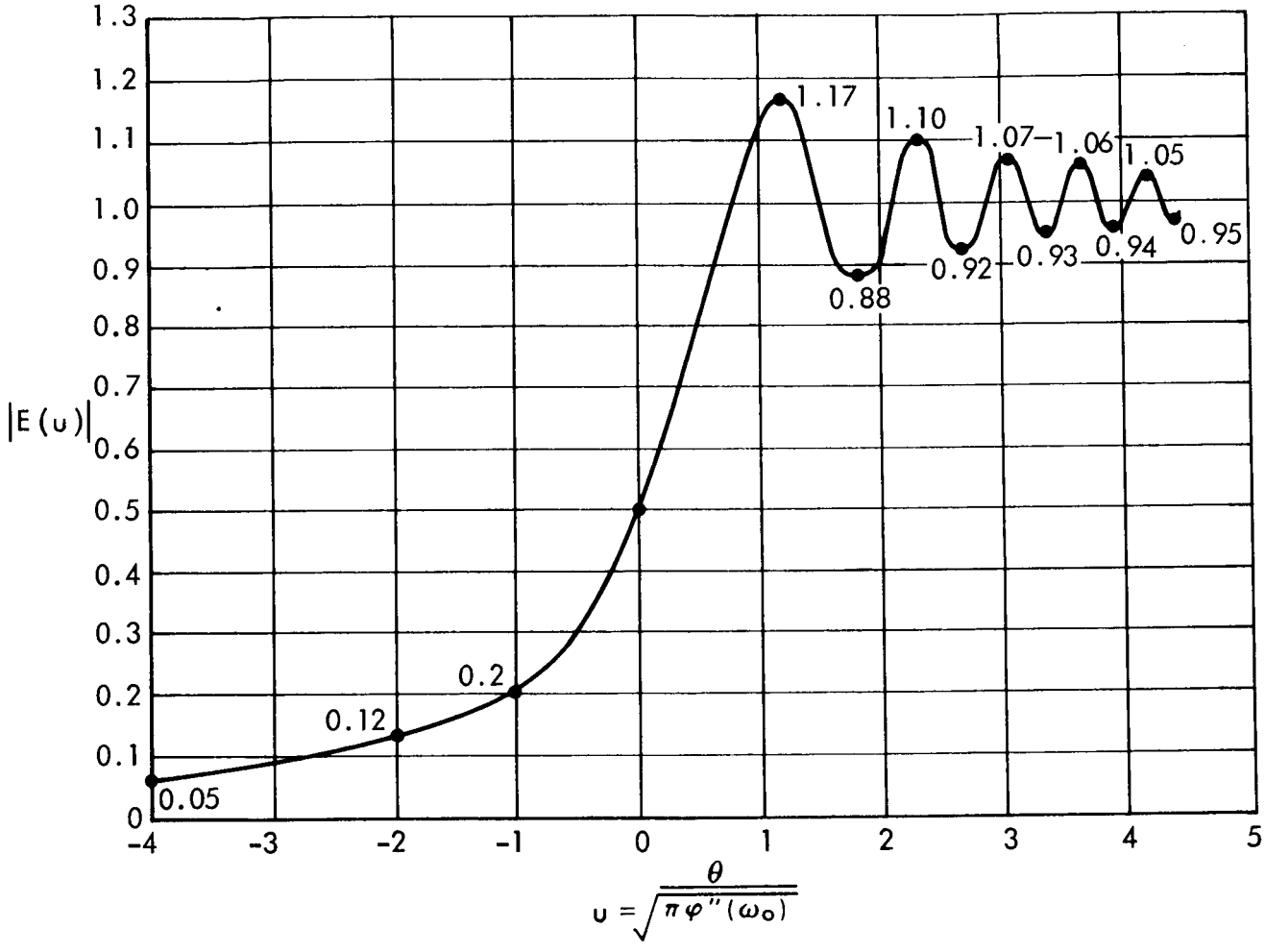

Fig. 1—Distortion of a rectangular pulse according to Ginzburg's theory

between the time the signal reaches 5 percent of the final value ($u = -4$) and the time it remains within $\pm$ 5 percent of the final value. Therefore, this establishment time is $8\sqrt{\pi\varphi_o''(\omega_o)}$. The customary definition used in this country is from the 10 percent point to the 90 percent point, which yields an establishment time equal to $3.3\sqrt{\pi\varphi''(\omega_o)}$.

According to this analysis, the rise time is determined by the parameter $\varphi''(\omega_o)$. For the phase function of Eq. (19), the situation of principal interest occurs when the carrier frequency f_o is much greater than the ionospheric cutoff frequency f_c. In this case, the rise time t_r becomes

$$t_r = 2.3\ \tau^{1/2}\ f_c\ f_o^{-3/2} \tag{33}$$

while the group time delay $\varphi'(\omega_o) = t_g$ becomes

$$t_g = \tau\ f_c^2/2\ f_o^2 \tag{34}$$

For the solution to make reasonable physical sense, the group delay must significantly exceed the rise time. While this condition was well-satisfied in the problem of reflection of pulses from the ionosphere which was the subject considered in Ref. 4, it is not necessarily satisfied in the transmission problem. The value of carrier frequency at which the group delay and the rise time become equal is $f(\text{GHz}) = .13\ I$. This frequency generally lies in the low gigahertz region. At higher frequencies, the approximations used in deriving Eq. (30) cease to be valid, since the expansion of the phase in powers of the deviation from the carrier frequency becomes slowly convergent. A different approximation is required.

The new approximation makes use of the well-known method of stationary phase. The application of this method to pulsed signals is clearly presented in Ref. 7. However, a direct application of stationary phase to the integral of Eq. (27) leads to the same convergence difficulties as arise in the expansion about the carrier frequency. It is necessary to make a transformation of variables which is equivalent to expressing the frequency as a function of phase, instead of vice versa. The transformation in question was introduced by Demisov,[5] but he did not employ stationary phase.

In the integral, Eq. (29), a new independent variable z is introduced defined by the relation

$$\omega_c z = \omega - \sqrt{\omega^2 - \omega_c^2} \tag{35}$$

As the variable ω traverses the integration path below the real axis indicated in Eq. (29), the variable z traverses a contour C which begins just left of the origin, follows the negative real axis to z = -1, then follows the unit semicircle in the upper half-plane to z = 1, returns along the positive real axis to the origin, and closes by a small semicircle in the lower half-plane. The singularities of the integrand at $\omega = \pm \omega_o$ lie outside the contour in the z-plane, but the integrand has an essential singularity at z = 0. In terms of the variable z, the integral becomes

$$E_s(f_o,t) = \frac{1}{4\pi} \int_C \frac{(z^2 - 1)dz}{z} \left[\frac{1}{z^2 + 2\frac{\omega_o}{\omega_c} z + 1} - \frac{1}{z^2 - 2\frac{\omega_o}{\omega_c} z + 1} \right] \times \exp\left(\frac{i\omega_c}{2} \left[(t+2\tau)z + \frac{t}{z} \right] \right) \tag{36}$$

The integrand has four poles, located at

$$z_1 = \frac{\omega_o - \sqrt{\omega_o^2 - \omega_c^2}}{\omega_c}, \quad z_2 = - z_1 \tag{37}$$

$$z_3 = \frac{\omega_o + \sqrt{\omega_o^2 - \omega_c^2}}{\omega_c} \qquad z_4 = - z_3 \tag{38}$$

When the algebraic part of the integrand is expanded in partial fractions, there results

$$E_s(f_o,t) = \frac{1}{4\pi} \int_C dz \left[\frac{1}{z + z_1} + \frac{1}{z + z_3} - \frac{1}{z - z_1} - \frac{1}{z - z_3}\right]$$

$$\times \exp\left(\frac{i\omega_c}{2}\left[(t+2\tau)z + \frac{t}{z}\right]\right) \tag{39}$$

The case of principal interest is $\omega_o \gg \omega_c$, so z_1 is small, and z_3 is large. Also, τ is on the order of 500-2000 μsec, while t is comparable to the group delay which is at most a few microseconds. The part of the integral, Eq. (39), which involves z_3 may easily be shown to be on the order of $(t/2\tau)^{\frac{1}{2}} z_3^{-1}$, which is less than .01 for all cases of interest.

To provide a consistent interpretation of the transformation from the ω-plane to the z-plane, the parameter z_1 should be considered as having a small negative imaginary part. The singularities of the integrand are at z_1 and $-z_1^*$, where the complex conjugate in the second singularity keeps it in the lower half-plane. The integrand now has no singularities in the upper half-plane except at the origin.

The contour C may be expanded to the entire real axis indented down at the origin, and a large semicircle in the upper half-plane. The latter makes no contribution to the integral. The contribution to the integral from the term involving $-z_1^*$ may be shown to be the complex conjugate of the contribution from the term in z_1, from which is derived

$$E_s(f_o,t) = \frac{1}{2\pi} \text{Re} \int_{-\infty}^{\infty} \frac{dz}{z - z_1} \exp\left(\frac{i\omega_c}{2}\left[(t+2\tau)z + \frac{t}{z}\right]\right) \tag{40}$$

To this point the analysis does not differ significantly from that of Ref. 5. The method of stationary phase is now to be applied to the integral in Eq. (40). The analysis is complicated by the presence of the pole of the integrand at $z = z_1$, which requires a careful treatment of the respective positions in the complex plane of the pole and the saddle points of the integrand. Let

$$g(z) = \frac{\omega_c}{2}\left[(t+2\tau)z + \frac{t}{z}\right] \tag{41}$$

The derivative of this is

$$g'(z) = \frac{\omega_c}{2}\left[t + 2\tau - \frac{t}{z^2}\right] \tag{42}$$

which vanishes at the two points

$$\pm z_o = \pm \sqrt{t/(t+2\tau)} \tag{43}$$

The second derivative of g, evaluated at $\pm z_o$, becomes

$$g''(\pm z_o) = \pm \omega_c (t+2\tau)^{3/2} t^{-1/2} = \pm \alpha \tag{44}$$

and the value of g itself at the saddle points is

$$g(\pm z_o) = \pm \omega_c \, t^{1/2}(t+2\tau)^{1/2} = \pm \beta \tag{45}$$

The stationary phase technique now replaces g(z) by its expansion to second-order terms around the two saddle points $\pm z_o$. For this approximation to be reasonable, the third-order terms must make a negligible contribution when the second-order terms are already rapidly varying. Thus, keeping third-order terms, the expansion around z_o is

$$g(z) \sim g(z_o) + \frac{1}{2} g''(z_o) \, (z-z_o)^2 + \frac{1}{6} g'''(z_0) \, (z-z_o)^3 \tag{46}$$

and there is a corresponding expansion around $-z_o$. The effective range of the second-order terms will be defined by

$$\frac{1}{2} g''(z_o) \, (z-z_o)^2 = \frac{\pi}{2} \tag{47}$$

$$\Delta z = z - z_o = \sqrt{\pi / g''(z_o)} \tag{48}$$

At this value of z, after much simplification, the third-order contribution to g(z) is

$$g_3(z) = \frac{\pi^{3/2}}{6} \quad \frac{g'''(z_o)}{g''(z_o)^{3/2}} = \frac{\pi^{3/2}}{2\beta^{1/2}} \tag{49}$$

When t is small compared to τ, β may be expressed as

$$\beta = 4.6 \times 10^3 \sqrt{It} \tag{50}$$

where t is in microseconds. For the smallest listed value of I, β is greater than 10^3 for t greater than .02 μsec. Thus, the third-order contribution to g may be neglected except in the very early time period, and this approximation becomes better for larger values of I. Under the same conditions, the saddle points at $\pm z_o$ are sufficiently separated so that their contributions may be added independently. Therefore, the integral in Eq. (40) is approximated by

$$E_s(f_o,t) \sim \frac{1}{2\pi} \operatorname{Re}\left[e^{-i\beta}\int_{-\infty}^{0} \frac{dz\, e^{-[i\alpha/2\,(z+z_o)^2]}}{z - z_1} + e^{i\beta}\int_{0}^{\infty} \frac{dz\, e^{[i\alpha/2(z-z_o)^2]}}{z - z_1}\right] \tag{51}$$

Translate the origin of the first integral to $-z_o$ and that of the second to z_o. Under the approximations already made, the limits may be extended to infinity. Thus

$$E_s(f_o,t) \sim \frac{1}{2\pi} \operatorname{Re}\left[e^{-i\beta}\int_{-\infty}^{\infty} \frac{dz\, e^{-[i\alpha/2\, z^2]}}{z - (z_o+z_1)} + e^{i\beta}\int_{-\infty}^{\infty} \frac{dz\, e^{[i\alpha/2\, z^2]}}{z - (z_1-z_o)}\right] \tag{52}$$

Again under the same approximations, the exponential begins to oscillate rapidly before the denominator has changed significantly from its value at the origin. In the first integral the denominator may be replaced by its value at the origin and removed from under the integration sign. The integral may then be evaluated, and the

resulting expression shown to be less than .01 for all interesting values of the parameters. It will therefore be neglected. However, since $z_1 - z_o$ may be small, this approximation cannot be applied to the second integral. Consequently, it remains to evaluate the expression

$$E_s(f_o,t) \sim \frac{1}{2\pi} \text{Re } e^{i\beta} \int_{-\infty}^{\infty} \frac{d\tau \, e^{[i\alpha/2 \, z^2]}}{z - (z_1 - z_o)} \tag{53}$$

This integral may be evaluated in terms of Fresnel integrals. First make the transformation $z = \sqrt{\pi/\alpha}\, u$. Define a parameter x by the relation

$$x = \sqrt{\frac{\alpha}{\pi}} (z_1 - z_o) = \sqrt{\frac{\alpha}{\pi}} \left[\frac{\omega_o - \sqrt{\omega_o^2 - \omega_c^2}}{\omega_c} - \sqrt{\frac{t}{t + 2\tau}} \right] \tag{54}$$

where x is to be given a small negative imaginary part. The real part of x is positive for $t = 0$, vanishes at $t = t_g$ (the previously defined group time delay) and is negative for $t > t_g$. The group time delay t_g now receives the interesting physical interpretation that it corresponds to that time at which the pole and saddle point of the integral, Eq. (40), are closest together. The integral now takes the form

$$E_s(f_o,t) \sim \frac{1}{2\pi} \text{Re } e^{i\beta} \int_{-\infty}^{\infty} \frac{du \, e^{[i\pi/2 \, u^2]}}{u - x} \tag{55}$$

After a lengthy sequence of transformations, Eq. (55) may be brought into the same functional form as Ginzburg's result, Eq. (30),

$$E_s(f_o,t) \sim \frac{1 - i}{2} \left[\frac{1 + i}{2} + F(-x) \right] \sin(\omega_o t - \varphi(\omega_o)) \tag{56}$$

but the argument of the Fresnel integral in Eq. (56) is not the same as that in Eq. (30). Comparing the arguments

$$\frac{\theta}{\sqrt{\pi\varphi''(\omega_o)}} \sim \sqrt{\frac{\omega_o t_g}{2\pi}} \left(\frac{t}{t_g} - 1\right) \tag{57}$$

$$-x \sim \sqrt{\frac{2\omega_o t_g}{\pi}} \left[\left(\frac{t}{t_g}\right)^{\frac{1}{4}} - \left(\frac{t}{t_g}\right)^{-\frac{1}{4}}\right] \tag{58}$$

These expressions both vanish at $t = t_g$ and have the same slope there. However, use of Eq. (58) in the argument of the Fresnel integral leads to a better treatment of the precursor wave than does Eq. (57). At $t = 0$, Eq. (56) actually goes to zero, while Eq. (30) tends to a small but finite value. Equation (56) is really not a correct reduction of Eq. (55) because the real part has not been taken correctly. However, the amplitude of the coefficient of $\sin[\omega_o t - \varphi(\omega_o)]$ in Eq. (56) equals the amplitude of the combination of sin and cos which results when Eq. (55) is reduced correctly. In deriving Eq. (56) from Eq. (55), the symmetry properties of the Fresnel integrals have been employed [$C(x) = -C(-x)$, $S(x) = -S(-x)$]. The equality of amplitudes of Eqs. (56) and (55) depends on z being real, which means the imaginary part of ω_o has been allowed to become zero.

The group time delay t_g may be written in terms of the parameter I as

$$t_g(\mu\text{sec}) = \frac{4}{3} \times 10^3 \, I/f_{\text{MHz}}^2 \tag{59}$$

The parameter I, which is proportional to the number of electrons along the path from transmitter to receiver, should be multiplied by the secant of the zenith angle of this path. For ground stations in temperate latitudes and satellites in synchronous equatorial orbit, this factor is typically between 2 and 3. Therefore, under solar minimum conditions, the diurnal, seasonal, and observation angle range of I is from 6 to 48. Under solar maximum conditions, the range is from 16 to 240.

Figure 1, the graph of Ginzburg's result, is plotted against the universal variable $(t-t_g)\sqrt{\pi\varphi''(\omega_o)}$. A graph of Eq. (56) against -x would be identical with Fig. 1. However, the variable x cannot be expressed solely in terms of the time difference $t - t_g$, so there is no convenient plot of Eq. (56) against a suitably referenced and scaled time. Therefore, a set of graphs has been drawn to cover many of the cases of interest, and these graphs of the output against $t - t_g$ are presented in Figs. 2-13. Rectangular and raised cosine pulses are treated under solar maximum and solar minimum conditions. Parameter values chosen were: (a) carrier frequency 300 MHz, pulse width 0.1 μsec; (b) 1 GHz, 0.05 μsec; (c) 5 GHz, 0.01 μsec. The calculations were performed using JOSS.*

The figures all display an oscillatory precursor and an extended tail. For the rectangular pulses, there is a high-frequency oscillation on the top of the pulse, which can be eliminated by a receiver.

---

*JOSS is the trademark and service mark of The RAND Corporation for its on-line, time-shared computer program and services using that program.

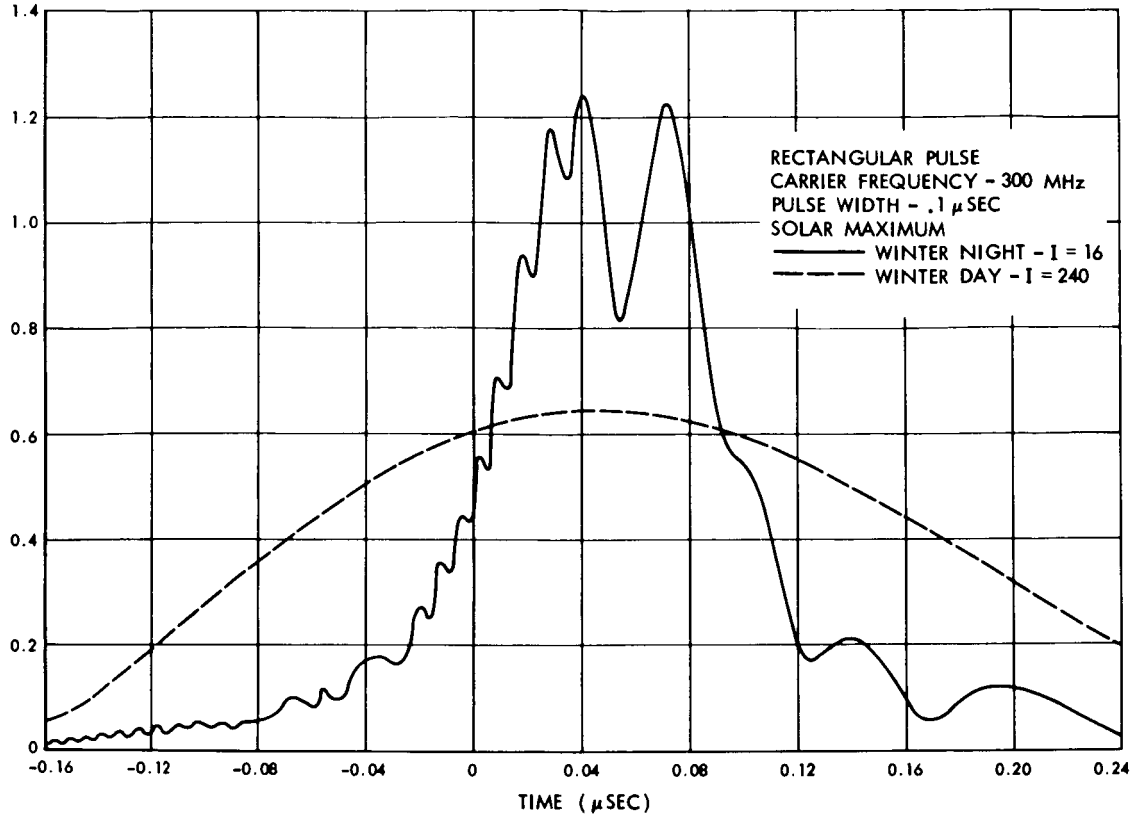

Fig. 2—Distortion of a 300 MHz, 0.1 μsec rectangular pulse at solar maximum

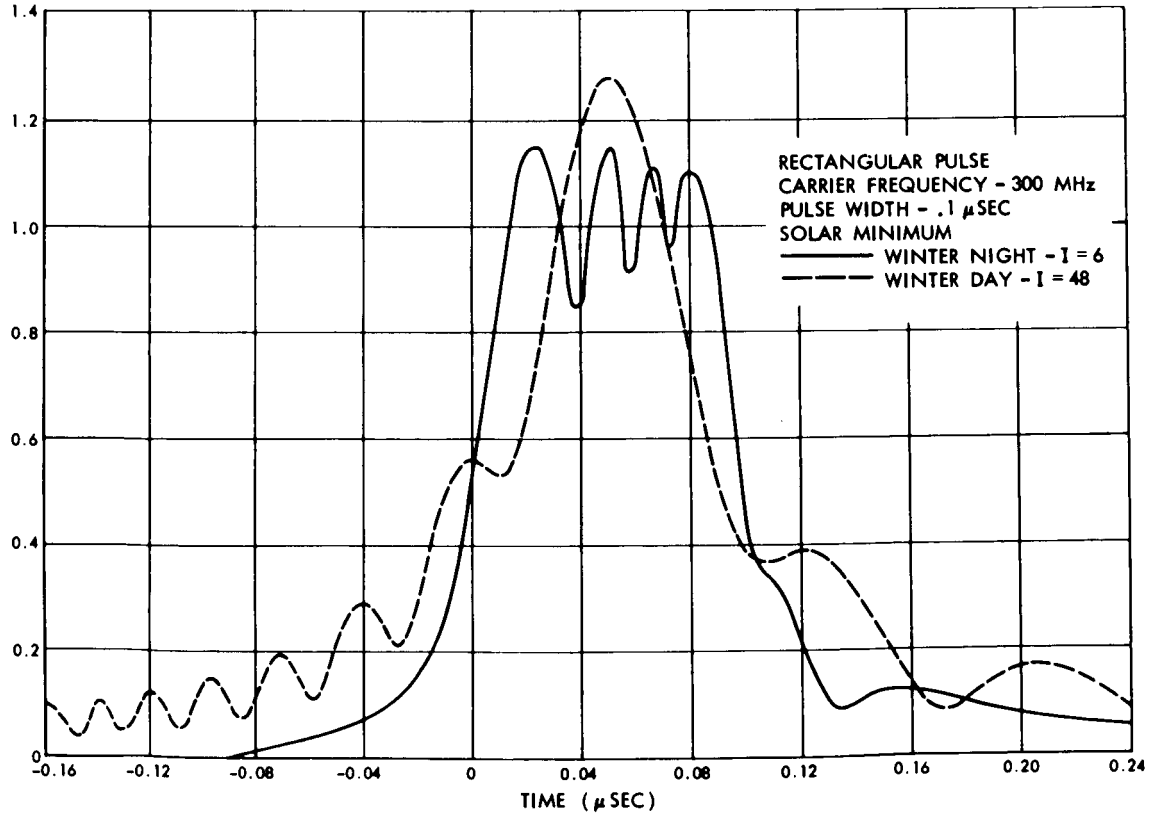

Fig. 3—Distortion of a 300 MHz, 0.1 μsec rectangular pulse at solar minimum

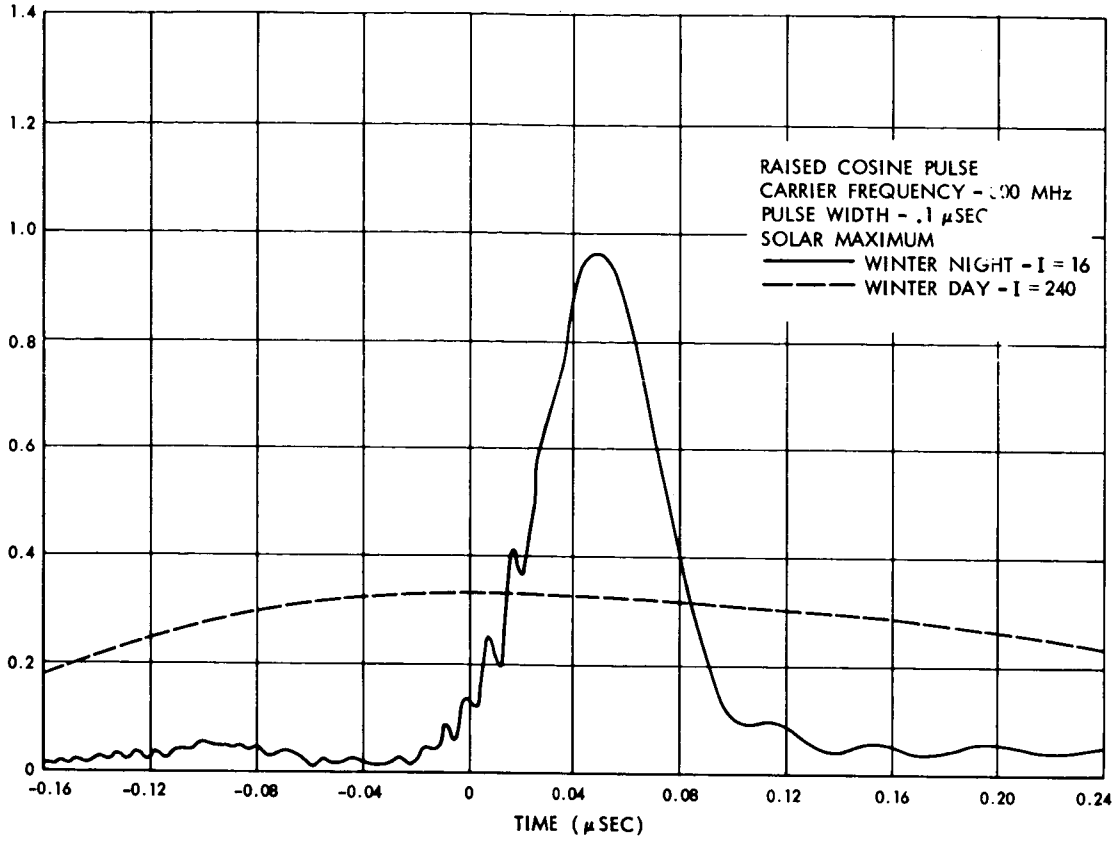

Fig. 4—Distortion of a 300 MHz, 0.1 μsec raised cosine pulse at solar maximum

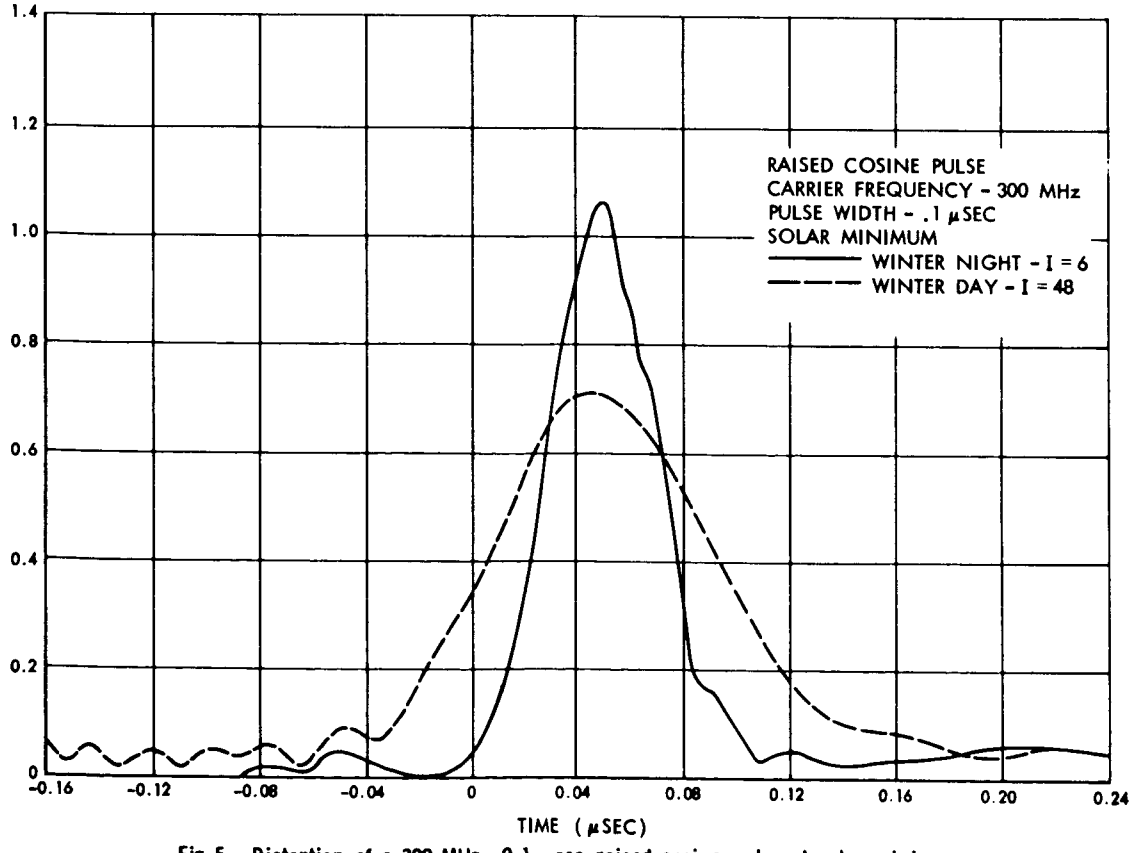

Fig. 5—Distortion of a 300 MHz, 0.1 μsec raised cosine pulse at solar minimum

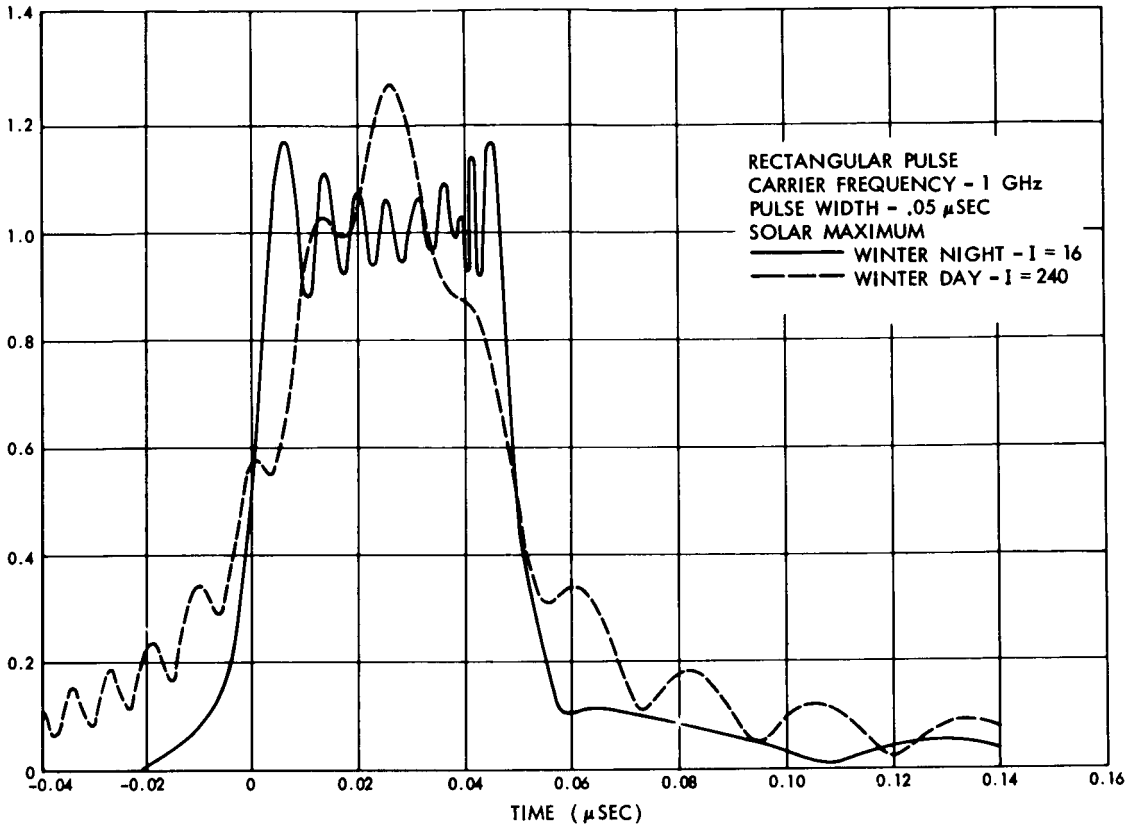

Fig. 6—Distortion of a 1 GHz, 0.05 μsec rectangular pulse at solar maximum

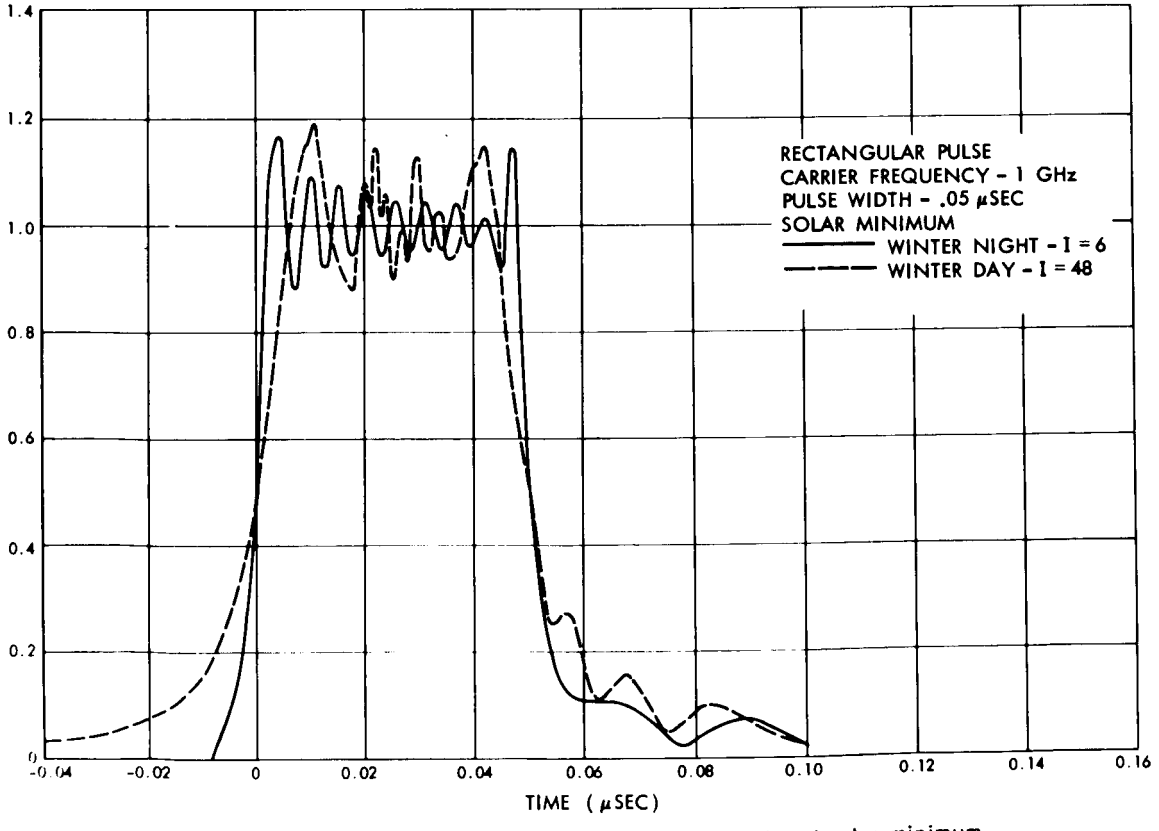

Fig. 7—Distortion of a 1 GHz, 0.05 μsec rectangular pulse at solar minimum

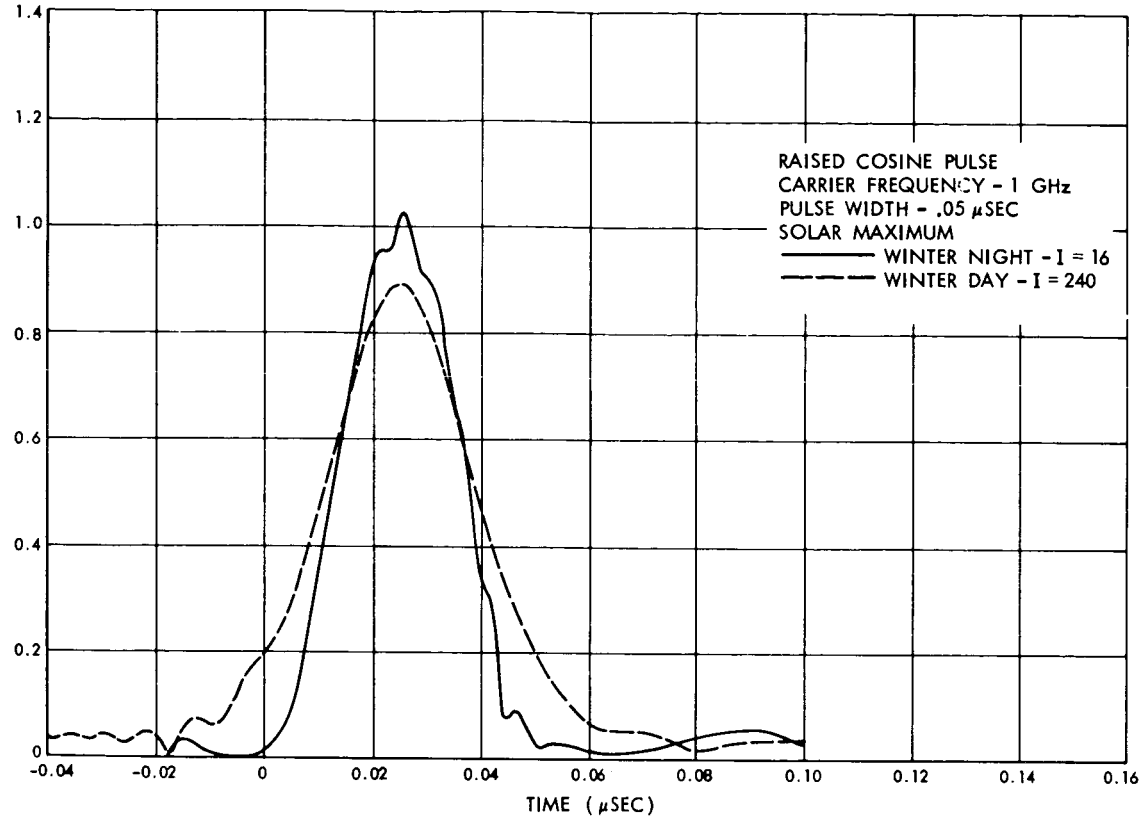

Fig. 8—Distortion of a 1 GHz, 0.05 μsec raised cosine pulse at solar maximum

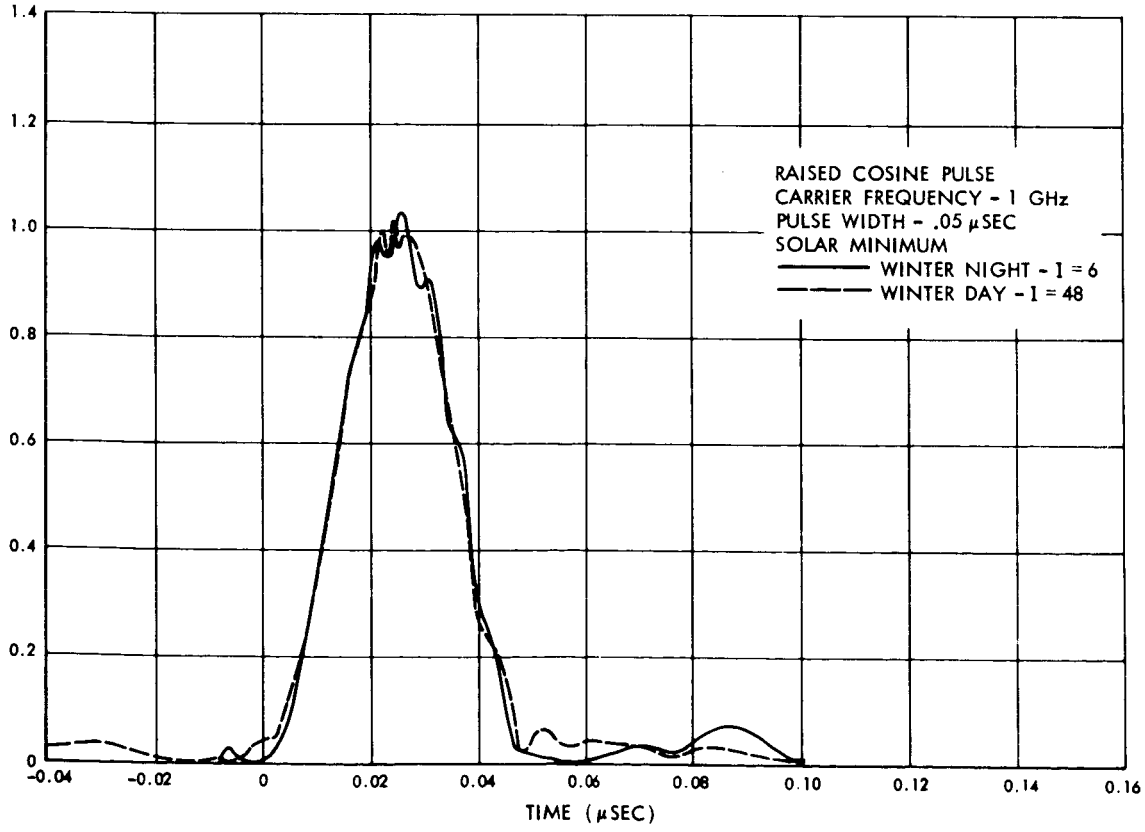

Fig. 9—Distortion of a 1 GHz, 0.05 μsec raised cosine pulse at solar minimum

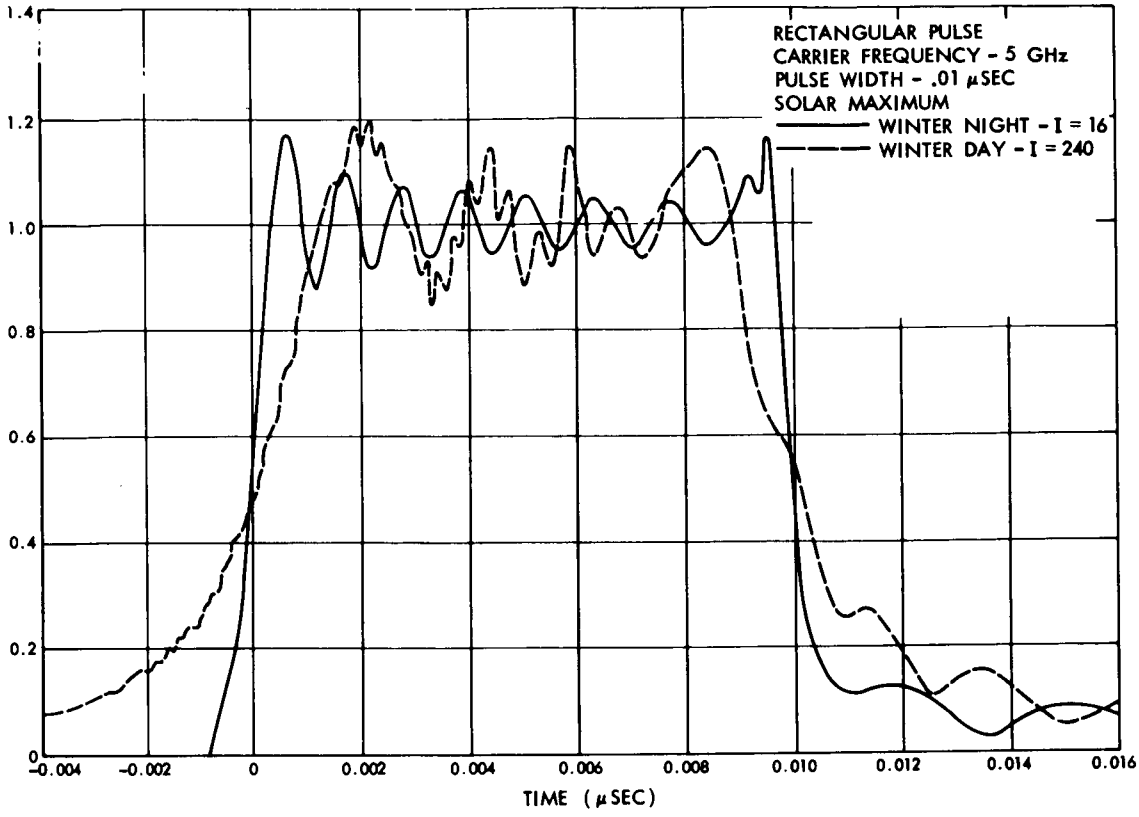

Fig. 10—Distortion of a 5 GHz, 0.01 μsec rectangular pulse at solar maximum

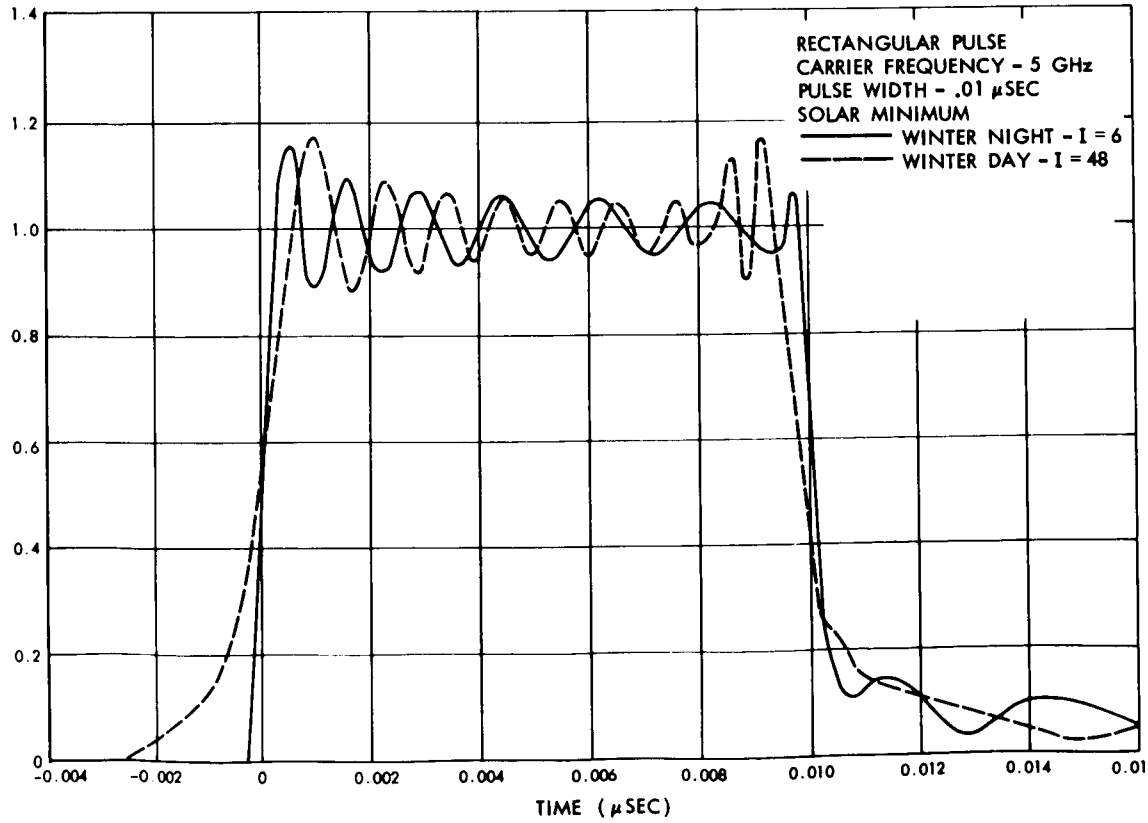

Fig. 11—Distortion of a 5 GHz, 0.01 μsec rectangular pulse at solar minimum

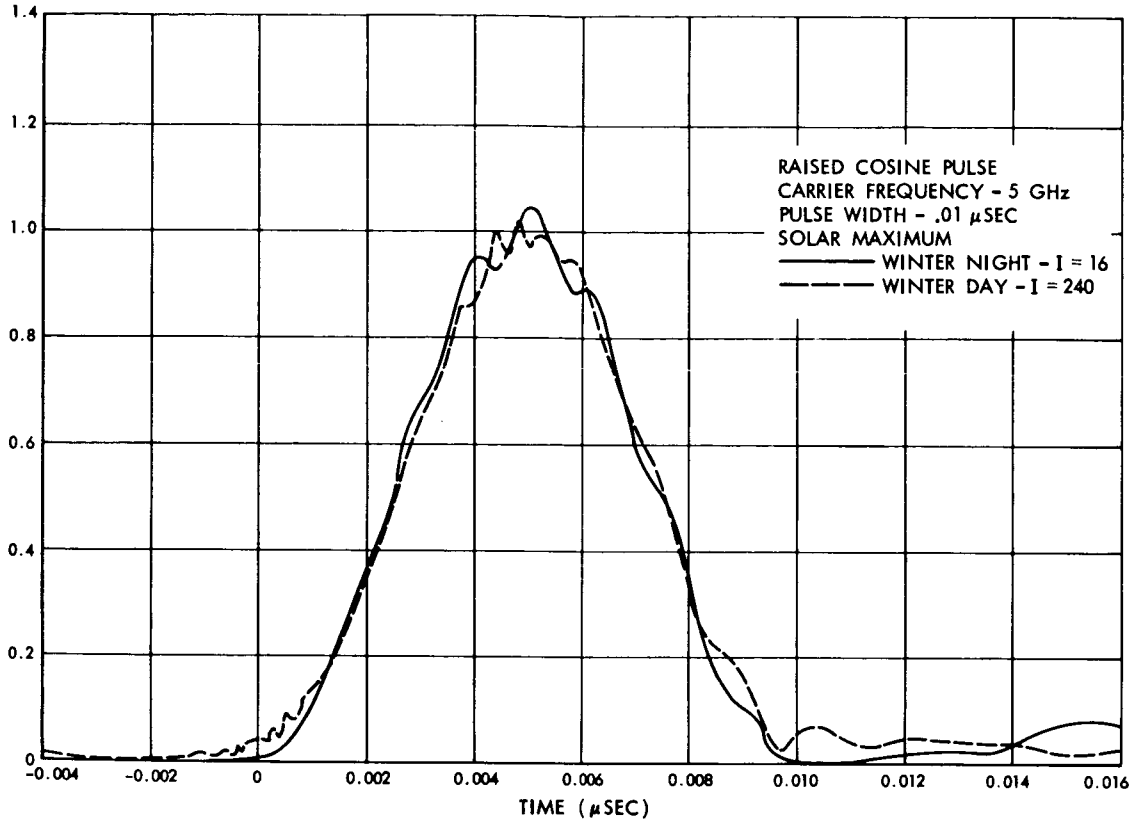

Fig. 12—Distortion of a 5 GHz, 0.01 μsec raised cosine pulse at solar maximum

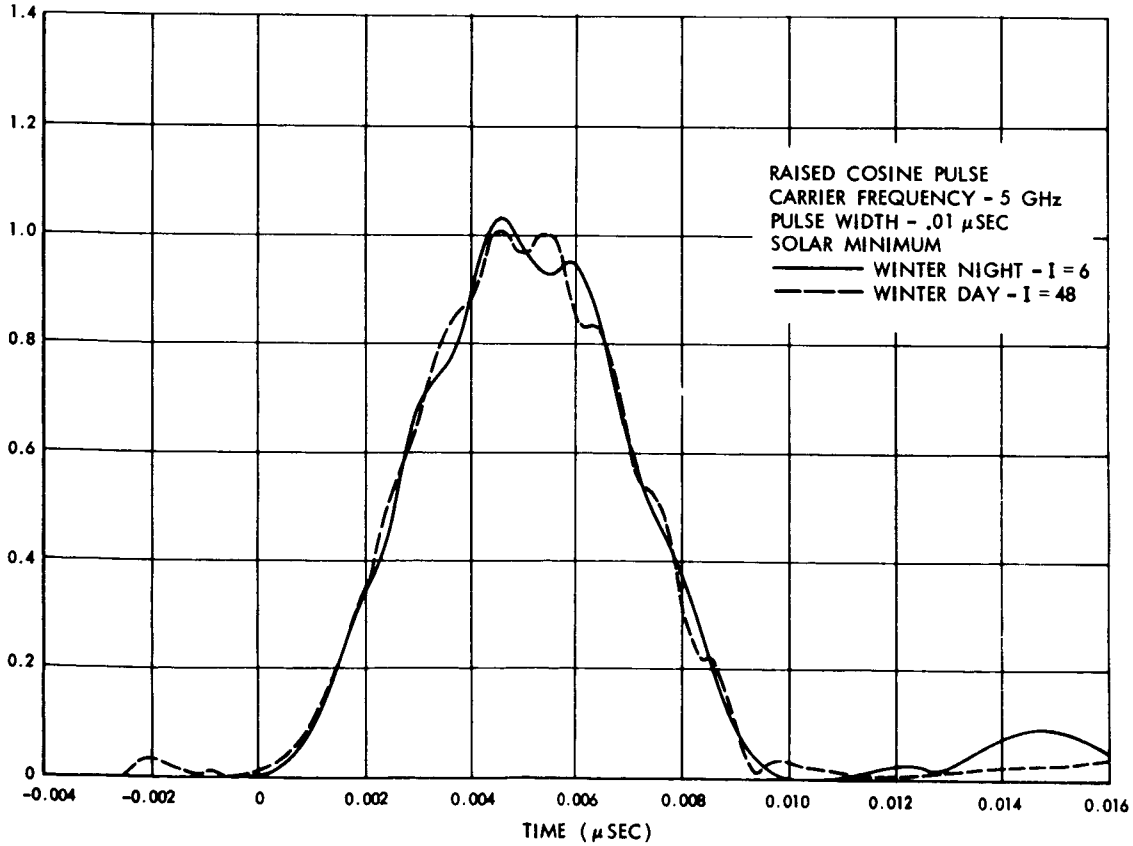

Fig. 13—Distortion of a 5 GHz, 0.01 μsec raised cosine pulse at solar minimum

However, the long tails can cause a system to regard a pulse as present when it is actually absent.

The figures show that the 300 MHz, 0.1 μsec pulse is very severely stretched under winter day conditions. The 1 GHz, 0.05 μsec pulse is borderline, and the 5 GHz, 0.01 μsec pulse seems acceptable. In terms of Ginzburg's parameter $T/\sqrt{\pi\varphi''(\omega_o)} = a$, values of a above 8 provide good pulse transmission, values between 3 and 8 are borderline, and values below 3 are bad. (The highly elongated pulses of Figs. 2 and 4 have a = 1.0.)

A careful comparison between the theory of Ginzburg and the results depicted in Figs. 2-13 shows that the minimum acceptable pulse width is given by about the same value in each theory. The fine structure on top of the pulse differs, but this is seldom important. The use of Eq. (56) to treat the precursor permits extension of the previous theory (Eq. (30)) to circumstances under which Eq. (30) is not valid.

It may be concluded that the shortest acceptable pulse width which may be employed in a pulsed communication system between earth and a satellite is given by equating pulse width and rise time, and is approximately

$$T(\text{nsec}) = 4\, I^{1/2}\, [f(\text{GHz})]^{-3/2} \tag{60}$$

where the quantity I is the number of electrons per square meter along the path in units of 10^{16}. In practice I is between 6 and 240, so $I^{1/2}$ is between 2.5 and 16. Thus, at 1 GHz, the shortest possible

pulse width ranges from 10 nsec (solar minimum, night) to about 60 nsec (solar maximum, winter day). At 5 GHz these become between 1 and 6 nsec, permitting very high data rates.

Another important effect is the overall displacement of the pulses. If I = 240, the group delay at 1000 MHz is .32 μsec, while at 1080 MHz it is .274 μsec. Thus, if a raised cosine pulse of width .05 μsec were transmitted at 1000 MHz, immediately followed by a .05 μsec pulse at 1080 MHz, the two pulses would arrive at the receiver almost simultaneously. Each distorted pulse can put appreciable energy into the receiver designed to accept the other. This condition might require a delay compensator to be built into each channel of a multichannel pulse code modulation system. Since the differential delay between channels is proportional to I, and hence may vary by a factor of 10-20 during the day, a variable delay compensator may be necessary.

Appendix

THE EPSTEIN LAYER

The symmetrical Epstein layer[13] represents what is most likely the simplest form of electron density profile which contains no discontinuities, tends to zero both above and below, and can be solved in terms of known functions. The functional form for the electron density is

$$N(z) = 4N_o \, e^{z/H} \Big/ (1+e^{z/H})^2 \tag{A-1}$$

Upon writing $e(z,\omega) = F(z,\omega) \exp - ikz$, the wave equation (3) takes the form

$$\left[\frac{d^2}{dz^2} - 2ik\frac{d}{dz} - \frac{4k_o^2 N_o e^{z/H}}{(1+e^{z/H})^2}\right] F(z,\omega) = 0 \tag{A-2}$$

Introduce a new independent variable x by the relation

$$x = \frac{1}{1 + e^{z/H}} \tag{A-3}$$

and new parameters $K = kH$, $K_o = k_o H$. Then Eq. (A-2) becomes, after much simplification

$$\left[x(1-x)\frac{d^2}{dx^2} + (1+2iK-2x)\frac{d}{dx} - 4K_o^2 N_o\right] F(x,\omega) = 0 \tag{A-4}$$

The boundary conditions become

$$F \to e_i\left[1+R(1-x)^{2iK}\right] \qquad x \to 1 \tag{A-5}$$

$$F \to e_i T \qquad x \to 0 \tag{A-6}$$

The differential equation (A-4) is in the standard form for the hypergeometric function.* Introduce a parameter γ by

$$\gamma = \sqrt{4K_o^2 N_o - \frac{1}{4}} \tag{A-7}$$

For all cases of physical interest γ is real. Then the general solution of Eq. (A-4) may be written in the form

$$F = e_i \left[\begin{array}{l} A\,{}_2F_1\left(\frac{1}{2} + i\gamma,\ \frac{1}{2} - i\gamma,\ 1 + 2iK, x\right) \\ + B\, x^{-2iK} (1-x)^{2iK}\, {}_2F_1\left(\frac{1}{2} + i\gamma,\ \frac{1}{2} - i\gamma,\ 1 - 2iK,\ x\right) \end{array} \right] \tag{A-8}$$

Here A and B are constants. The second term in Eq. (A-8) represents an incoming wave at $z = \infty (x = 0)$; thus the constants are determined from Eq. (A-6) as $A = T$, $B = 0$. Applying the connection formulas for the hypergeometric function** yields

$$F = e_i T \left[\begin{array}{l} \dfrac{\Gamma(2iK)\Gamma(1+2iK)\,{}_2F_1\left(\frac{1}{2} + i\gamma,\ \frac{1}{2} - i\gamma,\ 1 - 2iK,\ 1 - x\right)}{\Gamma\left(\frac{1}{2} + 2iK + i\gamma\right)\Gamma\left(\frac{1}{2} + 2iK - i\gamma\right)} \\ + \dfrac{\Gamma(-2iK)\Gamma(1+2iK) x^{-2iK} (1-x)^{2iK}\, {}_2F_1\left(\frac{1}{2} + i\gamma,\ \frac{1}{2} - i\gamma, 1 + 2iK, 1 - x\right)}{\Gamma\left(\frac{1}{2} + i\gamma\right)\Gamma\left(\frac{1}{2} - i\gamma\right)} \end{array} \right] \tag{A-9}$$

Applying the boundary condition Eq. (A-5) expresses the transmission coefficient T as

$$T = \frac{\Gamma\left(\frac{1}{2} + 2iK + i\gamma\right)\Gamma\left(\frac{1}{2} + 2iK - i\gamma\right)}{\Gamma(2iK)\Gamma(1+2iK)} \tag{A-10}$$

The reflection coefficient R may also be determined but it will not be needed.

---

*Reference 31, p. 562.

**Reference 31, p. 563.

It will first be shown that T reduces to the form of Eq. (11) for k sufficiently large. For this, use the asymptotic formula*

$$\log \Gamma(z) \rightarrow (z - \tfrac{1}{2}) \log z - z + \frac{1}{2} \log 2\pi + \frac{1}{12z} - \frac{1}{360z^3} + \frac{1}{1260z^5} \qquad \text{(A-11)}$$

which is correct to inverse sixth powers of z. For k sufficiently large, K = kH is large compared to either 1 or γ. The expansion (A-11) may be applied to each gamma-function factor in Eq. (A-10), and then the result may be further expanded in inverse powers of γ/K and $1/K$. All terms involving positive powers of K drop out. The result of the laborious simplifications is

$$\log T \rightarrow \frac{2i\, k_o^2 N_o H}{k} + \frac{1}{3} \frac{i\, k_o^4 N_o^2 H}{k^3} + \frac{2}{15} \frac{i\, k_o^6 N_o^3 H}{k^5} + \frac{i\, k_o^4 N_o^2}{60Hk^5} \qquad \text{(A-12)}$$

The terms contributing negative even powers of k all cancel. The first three terms of this expression may be identified as the first three terms in the expansion of the first integral in the exponent of Eq. (11), while the fourth term corresponds to the second integral in Eq. (11) when Eq. (A-1) is substituted for N. Thus, the asymptotic expansions of the approximate solution of Eq. (10) and the exact solution of Eq. (A-10) agree to fifth-order terms. Quantities neglected in Eq. (A-11) and ensuing expansions are less than .001 for z = 1, which corresponds to frequencies far below the region of interest.

The amplitude of the transmission coefficient T may be expressed exactly by the relation

---

*Reference 31, p. 257.

$$|T(\omega)| = \frac{\sinh 2\pi K}{\left[\frac{1}{2}(\cosh 4\pi K + \cosh 2\pi\gamma)\right]^{1/2}} \tag{A-13}$$

At high frequencies, this is approximately

$$|T(\omega)| \rightarrow 1 - \cosh 2\pi\gamma\, e^{-4\pi K} \tag{A-14}$$

and thus the amplitude of the transmission coefficient tends exponentially to unity. This explains why the real terms in the asymptotic expansion drop out. It may be expected that this result is generally true for electron density profiles which are continuous, have continuous derivatives, and tend exponentially to zero far above and far below the maximum.

While the incident carrier pulse problem cannot be solved exactly with the electron distribution Eq. (A-1) it is possible to solve for the δ-function response. This is the Fourier transform of $T(\omega)$, or

$$E_\delta(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\, e^{i\omega t}\, T(\omega) \tag{A-15}$$

and represents the solution when $E_i(-\infty,t) = \delta(t)$. The integral becomes

$$E_\delta(t) = \frac{1}{2\pi}\frac{c}{H}\int_{-\infty}^{\infty} dK\, e^{iKct/H}\, \frac{\Gamma\left(\frac{1}{2}+i\gamma+2iK\right)\Gamma\left(\frac{1}{2}-i\gamma+2iK\right)}{\Gamma(2iK)\Gamma(1+2iK)} \tag{A-16}$$

$$= \delta(t) + \frac{1}{2\pi}\frac{c}{H}\int_{-\infty}^{\infty} dK\, e^{iKct/H}\left[\frac{\Gamma\left(\frac{1}{2}+i\gamma+2iK\right)\Gamma\left(\frac{1}{2}-i\gamma+2iK\right)}{\Gamma(2iK)\Gamma(1+2iK)} - 1\right] \tag{A-17}$$

The integral in Eq. (A-17) is convergent. The integrand has poles at $1/2 \pm i\gamma + 2iK = -n$, where n is a positive integer, or equivalently

$$K_n = \mp \frac{1}{2}\gamma + i\left(\frac{1}{2}n + \frac{1}{4}\right) \tag{A-18}$$

These all lie in the upper half-plane. Evaluating the residues at these poles yields

$$E_\delta(t) = \delta(t) + \frac{c}{2H}\sum_0^\infty \frac{(-1)^n}{n!}\left[\frac{\Gamma(-n-2i\gamma)e^{-[ct/2H(n+1/2+i\gamma)]}}{\Gamma\left(-n-\frac{1}{2}-i\gamma\right)\Gamma\left(-n+\frac{1}{2}-i\gamma\right)} + \frac{\Gamma(-n+2i\gamma)e^{-[ct/2H(n+1/2-i\gamma)]}}{\Gamma\left(-n-\frac{1}{2}+i\gamma\right)\Gamma\left(-n+\frac{1}{2}+i\gamma\right)}\right] \tag{A-19}$$

Use of the reflection formula*

$$\Gamma(z)\Gamma(1-z) = \pi \csc \pi z \tag{A-20}$$

enables the series to be written as

$$E_\delta(t) = \delta(t) + \frac{c}{2H}\,\mathrm{Re}\,\frac{\cosh \pi\gamma}{\pi i \sinh \pi\gamma}\sum_0^\infty \frac{\Gamma\left(\frac{1}{2}+i\gamma+n\right)\Gamma\left(\frac{3}{2}+i\gamma+n\right)}{\Gamma(1+2i\gamma+n)n!} \times e^{-[ct/2H(n+1/2+i\gamma)]} \tag{A-21}$$

This is a hypergeometric function, yielding

$$E_\delta(t) = \delta(t) + \frac{c}{2\pi}\,\mathrm{Re}\,\frac{\cosh \pi\gamma}{\pi i \sinh \pi\gamma}\,\frac{\Gamma\left(\frac{1}{2}+i\gamma\right)\Gamma\left(\frac{3}{2}+i\gamma\right)}{\Gamma(1+2i\gamma)} \times e^{-[ct/2H\,(1/2+i\gamma)]}\,{}_2F_1\left(\frac{1}{2}+i\gamma, \frac{3}{2}+i\gamma, 1+2i\gamma, e^{-[ct/2H]}\right) \tag{A-22}$$

This hypergeometric function may be identified as a Legendre function. After many horrendous details, there results

*Reference 31, p. 256.

$$E_\delta(t) = \delta(t) - \frac{d}{dt} P_{-\frac{1}{2}+i\gamma}\left(2e^{[ct/2H]}-1\right) \tag{A-23}$$

It follows that for a general input function $E_i(t)$, the output is given by

$$E(t) = \int_{-\infty}^{t} dt' \frac{d}{dt'} E_i(t-t') P_{-\frac{1}{2}+i\gamma}\left(2e^{[ct'/2H]}-1\right) \tag{A-24}$$

In the case of principal interest, the argument of the Legendre function is close to one, while its order is large. This is the transitional region in which the Legendre function is represented by an expansion in Bessel functions. The resulting expansion is very similar to the corresponding representation for a slab,[11] and there seems little point to carrying the analysis further.

REFERENCES

1. Elliott, R. S., "Pulse Waveform Degradation due to Dispersion in Waveguide," Trans. IRE, Vol. MTT-5, No. 4, October 1957, pp. 254-257.

2. Counter, V. A., and E. P. Riedel, Calculations of Ground-Space Propagation Effects, Lockheed Aircraft Corporation, Missile Systems Division, Sunnyvale, California, Report No. LMSD-2461, May 1958.

3. Dyce, R. B., "Effects of the Ionosphere on Radio Waves at Frequencies Above 50 Mc," Section II of Final Report, Upper Atmosphere Clutter Research, Part XIII, Stanford Research Institute, Menlo Park, California, SRI Project 2225, April 1960.

4. Ginzburg, V. L., Propagation of Electromagnetic Waves in Plasma, Gordon and Breach Science Publishers, Inc., New York, 1961, pp. 411-423.

5. Demisov, N. G., "Propagation of Electromagnetic Signals in an Ionized Gas," Zhur. Eksp. i Teoret. Fiz., Vol. 21, 1951, pp. 1354-1363.

6. Gershman, B. N., "On the Spreading of Electromagnetic Signals in an Ionized Gas," Zhur. Tekh. Fiz., Vol. 22, 1952, pp. 101-104.

7. Wait, J. R., "Propagation of Pulses in Dispersive Media," Radio Science, Vol. 69D, No. 11, November 1965, pp. 1387-1401.

8. Knop, C. M., and G. I. Cohn, "Pulse Waveform Degradation due to Dispersion in Waveguide," IEEE Trans. Microwave Theory and Techniques, Vol. MTT-11, September 1963, pp. 254-257.

9. Scheibe, P., and J. Huntsberger, Ionospheric Effects, Electromagnetic Systems Laboratories, Palo Alto, California, Report No. ESL-IR11, July 1966, p. 18.

10. Rubinowicz, A., "Über die Fortpflanzung Unstetiger Elektromagnetischer Signale in Wellenleitern," Acta Phys. Polon., Vol. 13, 1954, pp. 115-133.

11. Poincelot, P., "Propagation d'un Signal le long d'un Guide d'Ondes," Annales des Telecommunications, Vol. 9, No. 11, November 1954, pp. 315-317.

12. Knop, C. M., "Pulsed Electromagnetic Wave Propagation in Dispersive Media," IEEE Trans. Antennas and Propagation, Vol. AP-12, No. 4, July 1964, pp. 494-496.

13. Epstein, P., "Reflection of Waves in an Inhomogeneous Absorbing Medium," Proc. Natl. Acad. Sci., Vol. 16, 1930, p. 627.

14. King, J. W., P. A. Smith, D. Eccles, and H. Helm, The Structure of the Upper Ionosphere as Observed by the Topside Sounder Alouette, D.S.I.R. Radio Research Station, Document No. RRS/IM 94, July 1963.

15. Evans, J. V., and G. N. Taylor, "The Electron Content of the Ionosphere in Winter," Proc. Royal Soc. London, Series A, Vol. 263, 1963, pp. 189-211.

16. Tyagi, T. R., and Y. V. Somayajulu, "Some Results of Electron Content Measurements at Delhi from Faraday Fading of S-66 Transmissions," Radio Science, Vol. 1, No. 10, October 1966, pp. 1125-1130.

17. Lawrence, R. S., D. J. Posakony, O. K. Garriott, and S. C. Hall, "The Total Electron Content of the Ionosphere at Middle Latitudes near the Peak of the Solar Cycle," J. Geophys. Res., Vol. 68, No. 7, April 1963, pp. 1889-1898.

18. Roger, R. S., "Measurements of the Equivalent Slab Thickness of the Daytime Ionosphere," J. Atmos. and Terr. Phys., Vol. 26, 1964, pp. 475-497.

19. Taylor, G. N., "The Electron Content of the Ionosphere at Middle Latitudes in Summer," Proc. Royal Soc. London, Series A, Vol. 279, 1964, pp. 497-509.

20. Millman, G. H., "The Variation of Electron Content in the Ionosphere at a Low Latitude," in J. Frihagen (ed.), Electron Density Profiles in Ionosphere and Exosphere, North Holland Publishing Company, Amsterdam, 1966, pp. 555-566.

21. Garriott, O. K., F. L. Smith, and P. C. Yuen, "Observations of Ionospheric Electron Content Using a Stationary Satellite," Planetary and Space Science, Vol. 13, 1965, pp. 829-838.

22. Yeh, K. C., and B. S. Flaherty, "Ionospheric Electron Content at Temperate Latitudes During the Declining Phase of the Sunspot Cycle," J. Geophys. Res., Vol. 71, 1966, pp. 4557-4570.

23. Bhonsle, R. V., A. V. deRosa, and O. K. Garriott, "Measurement of the Total Electron Content and the Equivalent Slab Thickness of the Midlatitude Ionosphere," Radio Science, Vol. 69D, No. 7, July 1965, pp. 929-937.

24. Checcacci, P. F., and M. T. deGiorgio, "Diurnal Variation of the Total Ionospheric Electron Content over Central Italy," J. Atmos. and Terr. Phys., Vol. 28, No. 1, January 1966, pp. 113-116.

25. Hibberd, F. H., and W. J. Ross, "Total Electron Content of the Ionosphere in Middle Latitudes," J. Geophys. Res., Vol. 71, No. 9, May 1966, pp. 2243-2253.

26. Yuen, P. C., and T. H. Roelofs, "Diurnal Variation of the Ionospheric Total Electron Content," J. Geophys. Res., Vol. 71, No. 3, February 1966, pp. 849-854.

27. Klobuchar, J. A., and H. E. Whitney, "Middle-Latitude Ionospheric Total Electron Content," Radio Science, Vol. 1, No. 10, October 1966, pp. 1149-1153.

28. Checcacci, P. F., "Ionospheric Measurements by Means of the Early Bird Geostationary Satellite," Radio Science, Vol. 1, No. 10, October 1966, pp. 1154-1158.

29. Titheridge, J. E., "Continuous Records of the Total Electron Content of the Ionosphere," J. Atmos. and Terr. Phys., Vol. 28, 1966, pp. 1135-1150.

30. Rao, N. N., "Ionospheric Electron Content and Irregularities Deduced from BE-C Satellite Transmissions," J. Geophys. Res., Vol. 72, No. 11, June 1967, pp. 2929-2942.

31. Abramowitz, M., and I. A. Stegun, Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series, AMS-55, June 1964.