NTIA Report 87-229

Millimeter-Wave Pulse Distortion by a Single Absorption Line Simulating the Terrestrial Atmosphere

George Hufford

U.S. DEPARTMENT OF COMMERCE
C. William Verity, Secretary
Alfred C. Sikes, Assistant Secretary
for Communications and Information

October 1987

CONTENTS

LIST OF FIGURES

Millimeter-Wave Pulse Distortion by a Single Absorption Line Simulating the Terrestrial Atmosphere

George Hufford*

Molecular absorption lines in the millimeter-wave band make the atmosphere dispersive, thus causing the distortion of transmitted signal shapes. To better understand how severe this distortion is and what form it takes, it seems useful to examine the abstract problem of a single, isolated absorption line. The approximate response to such a line may be easily computed and is well described by only a few parameters. We have evaluated these parameters and rather completely characterized the distortion for eight of the more prominent lines below 600 GHz.

Key words: absorption line; atmosphere; frequency distortion; frequency response; impulse response; millimeter-wave propagation; radio signal distortion

1. INTRODUCTION

Because of the enormous amount of spectrum available, one expects that millimeter-wave radio systems will operate at very high data rates. One problem, however, is the fact that the atmosphere is a dispersive medium where frequency response depends on pressure, temperature, and humidity, and hence on the capricious weather. Even a direct, uncluttered link will suffer a kind of multiplicative noise in which the transfer characteristics are random functions of time.

To better discuss the associated distortion, it seems useful to look at the abstract problem of a single absorption line. On the one hand, this problem is amenable to some analysis, and on the other hand, it provides a fairly good imitation of parts of the actual physical spectrum. The problem is very similar to the one originally treated by Brillouin (1914) and Sommerfeld (1914) and reported in Stratton's textbook (1941; ch. 5). It is identical with a problem Vogler has discussed (1970) and it is almost the same as that used by Trizna and Weber (1982).

*The author is with the Institute for Telecommunication Sciences, National Telecommunications and Information Administration, U. S. Department of Commerce, Boulder, CO 80303 3328.

The idea is that the atmospheric channel of length x has the frequency response function

$$H(\nu) = e^{2\pi i \nu n x/c} \tag{1}$$

where the index of refraction n is both complex and a function of the frequency ν. The single absorption line enters with the description

$$n(\nu) = n_o + \frac{m(\nu_o - i\gamma_o)}{\nu_o - i\gamma_o - \nu}. \tag{2}$$

In this equation the constants n_o, m, ν_o, and γ_o are real and positive; ν_o is the resonant frequency, γ_o the line width, and m is the line strength. Note that $n(\infty) = n_o$ and $n(0) = n_o + m$.

One defect in (2) is the lack of symmetry at negative frequencies. This has the consequence that when a real-valued signal is input to the channel, the output is complex-valued. The deficiency here is minor, however, since we can insist that all signals are described in a "narrowband" representation that excludes negative frequencies. Then nonreal values introduced by the channel are merely an indication of phase distortion.

Liebe (1985) has proposed a Millimeter-Wave Propagation Model (the MPM), which employs some 78 absorption lines with frequency response functions similar to (2). The functions differ in that they are symmetrized at negative frequencies and (at least for some of the oxygen lines) the parameter m is allowed to become complex. The model also differs from (2) in that it contains continuous-spectrum components. Thus (2) is a reasonable approximation only to the extent that a particular line can be isolated so that all other components of the refractive index are sensibly constant.

In the problem that Sommerfeld and Brillouin attack, the refractive index is not only symmetrized, but also it is not n but rather n^2 that satisfies something like (2). The resulting square root that must appear in (1) introduces four different branch points in the complex ν-plane and makes analysis a bit more complicated. While there is probably good theory behind this approach, one can still claim that, as written, (2) is a good approximation provided only that the refractive index n differs very little from unity. In the case of millimeter-wave propagation through the atmosphere, this difference is, even at the most dominant lines, always quite a bit less than 10^{-3}. The use of (2) seems entirely justified.

2. ANALYSIS

The analysis we contemplate is a study of the impulse response function. This function of the time t is the inverse Fourier transform of (1)

$$h(t) = \int_{-\infty}^{\infty} H(\nu)e^{-2\pi i\nu t}d\nu \tag{3}$$

and is the output of the channel when the input is the unit impulse acting at time t = 0. We set

$$\mu_o = \nu_o - i\gamma_o \tag{4}$$

$$t_o = mx/c \tag{5}$$

and then Vogler's result (1970) is

$$h(t) = e^{-2\pi i\mu_o t_o} h_o\left(t - \frac{n_o x}{c}\right) \tag{6}$$

$$h_o(t) = u_o(t) + u_1(t)\, h_1(t) \tag{7}$$

where $u_o(t)$ is the unit impulse, $u_1(t)$ the unit step function, and

$$h_1(t) = -e^{-2\pi i\mu_o t}\, 2\pi\mu_o \sqrt{t_o/t}\; J_1(4\pi\mu_o\sqrt{t_o t}). \tag{8}$$

The first thing we note is that $h_o(t)$ vanishes for negative t and so h(t) vanishes for $t < n_o x/c$. No signal can pass through the channel in a shorter time and so the signal velocity is at most c/n_o. Then if $n_o \geq 1$, the signal velocity is less than the velocity of light and Einstein's dictum is already satisfied independent of the remaining parameters and independent of particular values of $n(\nu)$.

The impulse response in (6), (7), and (8) consists of a delayed copy of the input and an immediately following transient. There are no "precursors" such as those described by Brillouin.

The transient part described in (8) has the appearance of a ringing at the resonant frequency, which is exponentially attenuated and then modulated

by the Bessel-function term. For small t we may evaluate (8) by using a power series expansion

$$h_1(t) = -e^{-2\pi i\mu_o t}\, 4\pi^2\mu_o^2 t_o \sum_{k=0}^{\infty} \frac{1}{k!(k+1)!}\,(-4\pi^2\mu_o^2 t_o t)^k \qquad (9)$$

$$= -e^{-2\pi i\mu_o t}\, 4\pi^2\mu_o^2 t_o\, [1-2\pi^2\mu_o^2 t_o t + \ldots].$$

For large t we may use the asymptotic expansion whose first term gives

$$h_1(t) \sim -e^{-2\pi i\mu_o t} \left(\frac{4\mu_o^2 t_o}{t^3}\right)^{1/4} \cos\,(4\pi\mu_o\sqrt{t_o t} - \tfrac{3\pi}{4}) \qquad (10)$$

$$= -\,e^{-2\pi i\mu_o t} \left(\frac{\mu_o^2 t_o}{4t^3}\right)^{1/4} \quad (w + \tfrac{1}{w})$$

where

$$w = e^{4\pi i\mu_o\sqrt{t_o t}\, - i3\pi/4}. \qquad (11)$$

Since μ_o is complex, when t is really large, w begins to grow and soon the 1/w term in (10) can be ignored. When this is true we have

$$h_1(t) \sim \left(\frac{\mu_o^2 t_o}{4t^3}\right)^{1/4} e^{i\pi/4}\, e^{-2\pi i\mu_o (t-2\sqrt{t_o t})}. \qquad (12)$$

Thus the Bessel function modulation evolves into an almost pure phase modulation. The transient now appears to be an exponentially damped signal having a frequency that starts below the resonant frequency and then increases according to the formula

$$\nu(t) = \nu_o(1- \sqrt{t_o/t}). \qquad (13)$$

In order to provide an idea of magnitudes involved we have prepared Table 1, which gives eight of the more important atmospheric absorption lines and

values for several associated quantities. The basic parameters, ν_o, γ_o, and m are derived under the assumption of standard atmospheric conditions and a 50 percent relative humidity. Most of them come directly from the line definitions in Liebe's MPM. The 60-GHz "line," however, is simply a numerical fit to the point of maximum absorption and the points of half-maximum absorption. The difficulty, of course, is that this line is really a complex of some 41 lines; as we shall see, there are many purposes for which the present approximation is not adequate.

The remaining columns of Table 1 give values for several derived quantities. The quantity α, for example, is simply the specific attenuation experienced by a monochromatic signal at the resonant frequency. It is mentioned here as a reminder of the magnitude of the problem. The value t_o is that given in (5) when the distance x is 10 km. In addition, the column header indicates that this quantity is proportional to the first power of x so that values for other distances are easily obtained.

The time t_1 might be called the "time to extinction." It is the time at which the first exponential term in (8) has reached one-tenth the original value and is given by

$$t_1 = \ell n 10/2\pi\gamma_o. \qquad (14)$$

TABLE 1. Values Associated with Some Millimeter-Wave Absorption Lines.

	ν_o GHz	γ_o GHz	m $\times 10^{-8}$	α dB/km	$t_o(x^1)$ ps	$t_1(x^o)$ ps	$t_2(x^{-1})$ ps	$t_3(x^{-1})$ ns
H_2O	22	3.0	.44	.13	.147	122	1300	6.3
O_2	60	3.9	9.0	15	3.00	94	8.6	.18
O_2	119	1.7	.090	1.4	.030	216	220	97
H_2O	183	3.1	1.2	24	.40	118	6.9	2.2
H_2O	325	2.9	.44	29	.147	126	6.0	6.8
H_2O	380	3.0	2.9	250	.97	122	.66	1.0
H_2O	448	2.6	2.2	310	.73	141	.63	1.7
H_2O	557	3.2	89.	16000	30	115	.010	.028

The column headings are explained in the text. Values were computed assuming a path length of 10 km and a standard atmosphere at sea level with 50 percent relative humidity. Values for other path lengths may be computed using the indicated dependence on the distance x.

We might suppose that after a time t_1 the transient has nearly died out and that t_1 represents the amount by which a pulse would be broadened. Soon we shall correct these statements somewhat, but for now we note that t_1 is independent of the distance x and for all eight lines it varies over the limited range of about 100 to 200 ps. This means that changes in the transmitted signal will cause transients that last for only this short period of time. It means, for example, that intersymbol interference will be of little importance until the information rates begin to approach 5 Gbit/s.

The first positive zero of $J_1(y)$ occurs at about $y = 3.832$. The time t_2 is the time at which this first zero is reached, and is given by

$$t_2 = (3.832)^2/(16\pi^2\nu_o^2 t_o). \qquad (15)$$

Although the argument of the Bessel function in (8) is complex so that the transient does not have an exactly zero amplitude, at time t_2 the argument will be close to the zero and the amplitude will experience a sharp minimum. We can say that at t_2 the Bessel-function modulation asserts itself and that for larger t the asymptotic results in (10) begin to be valid.

From Table 1 we note that the two lines at 22 and 119 GHz have values of t_2 that are very large. They are either comparable to or much larger than the corresponding t_1. By the time the Bessel-function modulation becomes important the transient has died away, so the behavior here is that of a simple decaying exponential. Note, however, that at 20 km the time t_2 is half as big as indicated and that the 119-GHz transient will begin to show more complicated behavior.

For the remaining six lines in Table 1 the values of t_2 are all very short. Indeed, if we recall that at the radio frequencies we are considering, an rf cycle is about 10 ps long, we note that t_2 is even shorter than such a cycle. For any realistic pulse with a limited bandwidth the initial Bessel function behavior will be completely wiped out and the only observable results will be in accord with the asymptotic formula in (10). When the distance is only 1 km, the lines at 60, 183, and 325 GHz have values of t_2 that approach corresponding values of t_1, and in these cases the transients may show more complicated behavior.

The last column in Table 1 tries to answer the question whether the transient resembles the simpler asymptotic formula in (12). We have calculated the time at which the magnitude of w is 10 times the magnitude of 1/w,

$$t_3 = \frac{t_1^2}{16t_o}. \tag{16}$$

Most values of t_3 are even an order of magnitude larger than t_1, so there will be very little tendency to reach the simple exponential behavior. At 60, 183, 380, and 448 GHz there is some chance, especially at distances much larger than 10 km. And at 557 GHz one expects an almost pure exponential, even though a path length of more than a few kilometers is probably impossible.

3. GRAPHICAL ILLUSTRATIONS

We recall that if x(t) is the input signal to a channel, then the output is y(t) = x(t) * h(t) where h(t) is the impulse response of the channel and the star indicates simple convolution. We recall also that if

$$x(t) = p(t)e^{-2\pi i\nu_c t}, \tag{17}$$

so that the "baseband signal" p(t) modulates the "central frequency" ν_c, then

$$y(t) = q(t)e^{-2\pi i\nu_c t} \tag{18}$$

where

$$q(t) = p(t) * \hat{h}(t)$$

$$\hat{h}(t) = e^{2\pi i\nu_c t} h(t).$$

In Figure 1* we have plotted representative examples of how a short pulse is distorted after passing through a channel with a single one of the eight different lines of Table 1. Each plot shows superimposed traces of the co-phase and quadrature-phase components--i.e., of the real and imaginary parts of the baseband output signal. Vertical scales are linear in volts, but are

*This and the remaining figures may be found at the end of the present section beginning at page 12.

otherwise quite arbitrary. The plots were computed as described in the Appendix, using a discretization of the pulse p(t), the discrete Fourier transform (DFT), and the formulae in (1) and (2).

The input pulses here are Gaussian pulses modulating a central frequency equal to the resonant frequency as indicated on each plot. The pulses are extremely short--indeed, the bandwidth required to produce them is usually equal to or even greater than, the central frequency itself. While we have been careful to keep the occupied band away from zero frequency and to have somewhat more than one rf cycle in the pulse, we have still used some absurdly short pulses. Our purpose here is not to show practical examples, but to demonstrate the features described above.

As suggested, the curves for 22 and 119 GHz consist of the original pulse followed by a simple exponential transient. The curves for 60, 183, and 325 GHz show the Bessel-function transient and how the original pulse becomes cancelled out. And finally, the curves for 380, 448, and 557 GHz are almost entirely the asymptotic form of the transient, showing a cisoidal or modified cisoidal behavior; note how the oscillation frequency slows with time.

Each plot of Figure 1 has a legend showing the frequency involved, the distance assumed, and the path attenuation suffered by the pulse. For distances we have tried to stay with the 10 km of Table 1, but have sometimes deviated either to emphasize some feature of the response or to avoid some practical problems in plotting. The attenuation given is the ratio (expressed in decibels) of the energy in the input pulse to the energy in the output. One notes immediately that these attenuations are a great deal less than are implied by the values of α in Table 1. This is because the bandwidth is very large and the absorption line merely cuts a small notch out of the total spectrum.

This latter fact shows also that the choice of central frequency for these short pulses is not important. A modest change here will, because the transient rings at the resonant frequency, produce a small change in the appearance of the phase distortion but otherwise will have little effect.

Several of the plots in Figure 1 show transients that continue on for a time significantly larger than the t_1 of Table 1. There are several reasons for this. First, the asymptotic result contains an exponential term that actually grows with time. If we ask when the exponential in (12) reaches a tenth its initial value, we find a time

$$t'_1 = (\sqrt{t_o} + \sqrt{t_o + t_1})^2, \qquad (19)$$

which should be a somewhat more accurate representation of the length of the transient. This value is certainly larger than t_1, but when t_o is small (as seems in Table 1 to be almost always the case), the increase is only minor. On the other hand one should note that when t_o is even one-eighth the value of t_1, then t'_1 is already twice t_1.

Probably a more important reason for extended transient times, however, is the simple fact that the plots of Figure 1 have been normalized. The large and highly oscillatory beginning of a transient is reduced to very small values by the smoothing action of convolution. We therefore scale up the function to provide a more pleasing plot, and in so doing we reveal parts of the transient's tail that were formerly unobservable. Thus the "one-tenth" used above is perhaps inappropriate.

In Figure 2 is a second set of examples of pulse distortion. Here, only the single line at 183 GHz enters, but results are shown for distances ranging from 0 to 50 km. The zero-distance plot shows the input signal. In the remaining plots one sees the output changing from a simple exponential transient through the Bessel-function behavior, and finally to the asymptotic form.

Figures 3 and 4 repeat two of the plots in Figure 2. Each cophase, quadrature-phase plot is accompanied by two additional plots of the response function. These new plots, which attempt to provide a better understanding, are first an "amplitude" plot and then a "polar" plot. The first is simply a plot of the amplitude of the response versus time. Scales are the same as before. The polar plot shows the complex plane and the trajectory formed by the real and imaginary parts of the response. It is intended to display phase distortion, but loss of the time dimension is a disagreeable drawback.

From the polar plots we might conclude that even in extreme conditions there is, if the 180° shift is tolerable, only minor phase distortion. One must, however, remember that the transient is ringing at the resonant frequency and when the central frequency is removed a few gigahertz, this extra frequency may be bothersome.

Beginning in Figure 5, we turn to an examination of how more practical pulses are distorted. The distortion suffered by a simple rectangular pulse of length 0.5 ns is shown in Figures 5 and 6. Figure 5 shows the pulse at a constant central frequency and a variety of distances while Figure 6 shows the pulse at a constant distance and a variety of central frequencies. The central frequency is now important because the bandwidth is reduced and because the signal is constant for a long enough period of time to take on the appearance of a cw signal. Note, for example, that when the central frequency approaches the resonant frequency, the principle part of the pulse drops out and only the transients at the leading and trailing edges remain. There is still enough bandwidth to make this "sideband splatter" important.

In Figures 7 and 8 we go one step further and show the distortion present in a string of bits. The string has a 2 Gbit/s rate, and in the small portion shown one can observe the five bits 10100. In Figure 7 modulation is by means of bi-phase shift keying (BPSK) in which the two kinds of bits are represented as rectangular pulses with phases 180° apart. It is interesting to note that on changing the phase the trace does not pass through zero but instead goes from one point to its negative in the complex plane with a large quadrature-phase component.

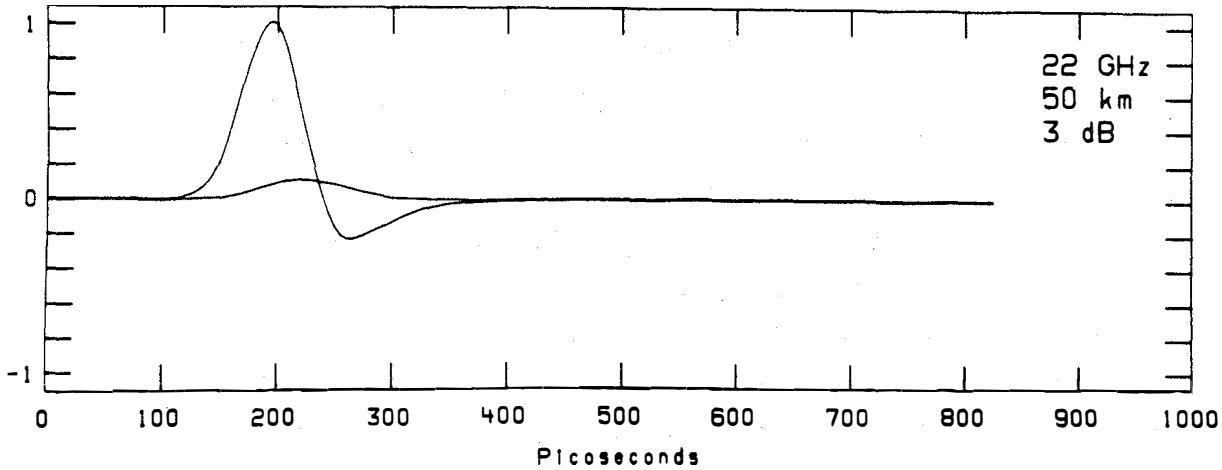
In Figure 8 the modulation used is minimum shift keying (MSK). This is a frequency shift keying in which the frequency separation is just equal to the bit rate. In the present case the two frequencies are 176 and 174 GHz. The input function now has sudden changes only in its first derivative and hence the transients are rather small. On the other hand, the fact that the attenuation at 176 GHz is greater than at 174 GHz introduces considerable amplitude distortion and may make signal demodulation difficult. Clearly, any kind of frequency modulation may have similar troubles.

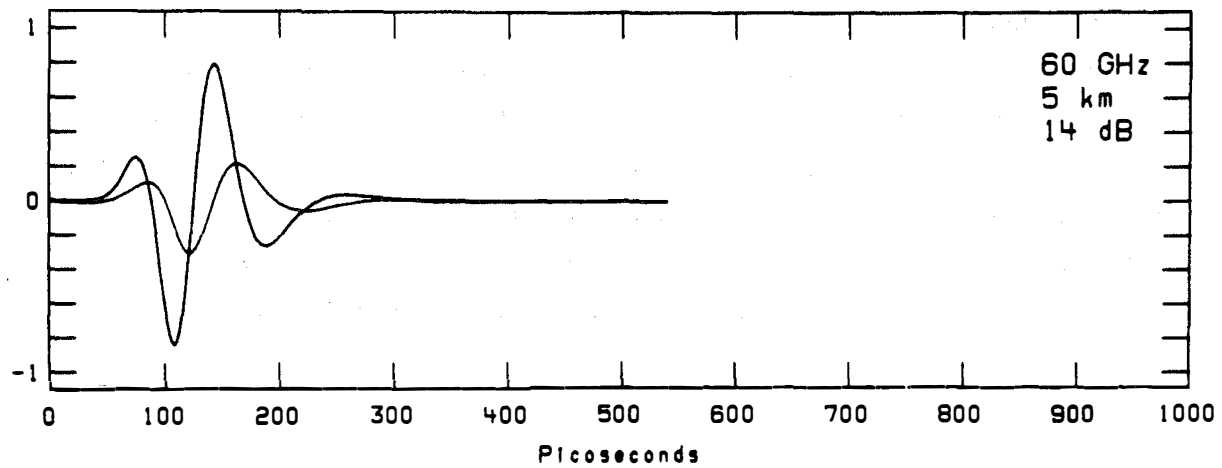
In all that we have said we have emphasized that induced transients last for only a short time. We conclude this note by pointing out situations where this might not be true.

The extinction time t_1 is immediately related to the parameter γ_o which, in turn, is almost directly proportional to the atmospheric pressure. At higher altitudes, therefore, γ_o decreases and t_1 increases. Of course, the line strength m will also decrease as will the frequency range over which the line is important.

The situation near 60 GHz is another one that gives rise to longer-lived transients. The "line" there is not approximated very well by (2). At high altitudes it breaks up into a number of sharp lines, but even at sea level its edges have steeper slopes than those derived from (2) and its top is flatter. At the longer distances, the portion of the spectrum near 60 GHz will be greatly attenuated, and whether it is greatly attenuated or enormously attenuated really makes no difference since a response will depend almost entirely on the sideband spectra. Thus a response will be independent of the maximum line absorption but will show mostly the effects of the edges. It is the steepness of the edges that is most important, and steep edges imply long transient times.

In addition to long transient times, the two isolated edges give rise to some interesting forms of pulse distortion. In Figure 9 we show a variety of responses when a Gaussian pulse is transmitted through a close approximation to the actual atmosphere. The frequency response used is that given by (1) where now n is described by the MPM, and the pulse width is long enough so that its spectrum avoids the adjacent lines at 22 and 119 GHz.





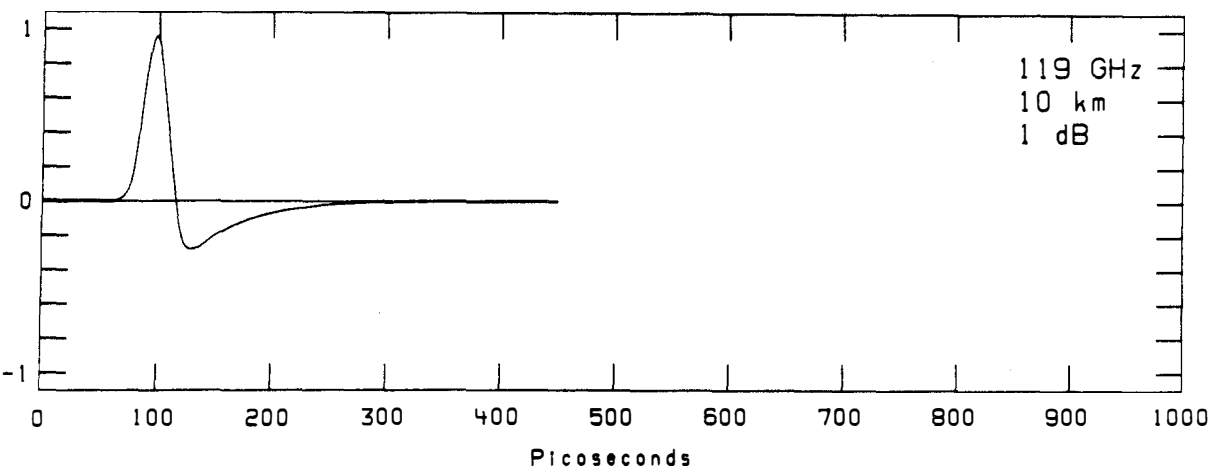


Figure 1. Examples of pulse distortion over a channel with a single absorption line. The plots show superimposed traces of the cophase (the darker line) and quadrature-phase components. (This is page 1 of 3.)

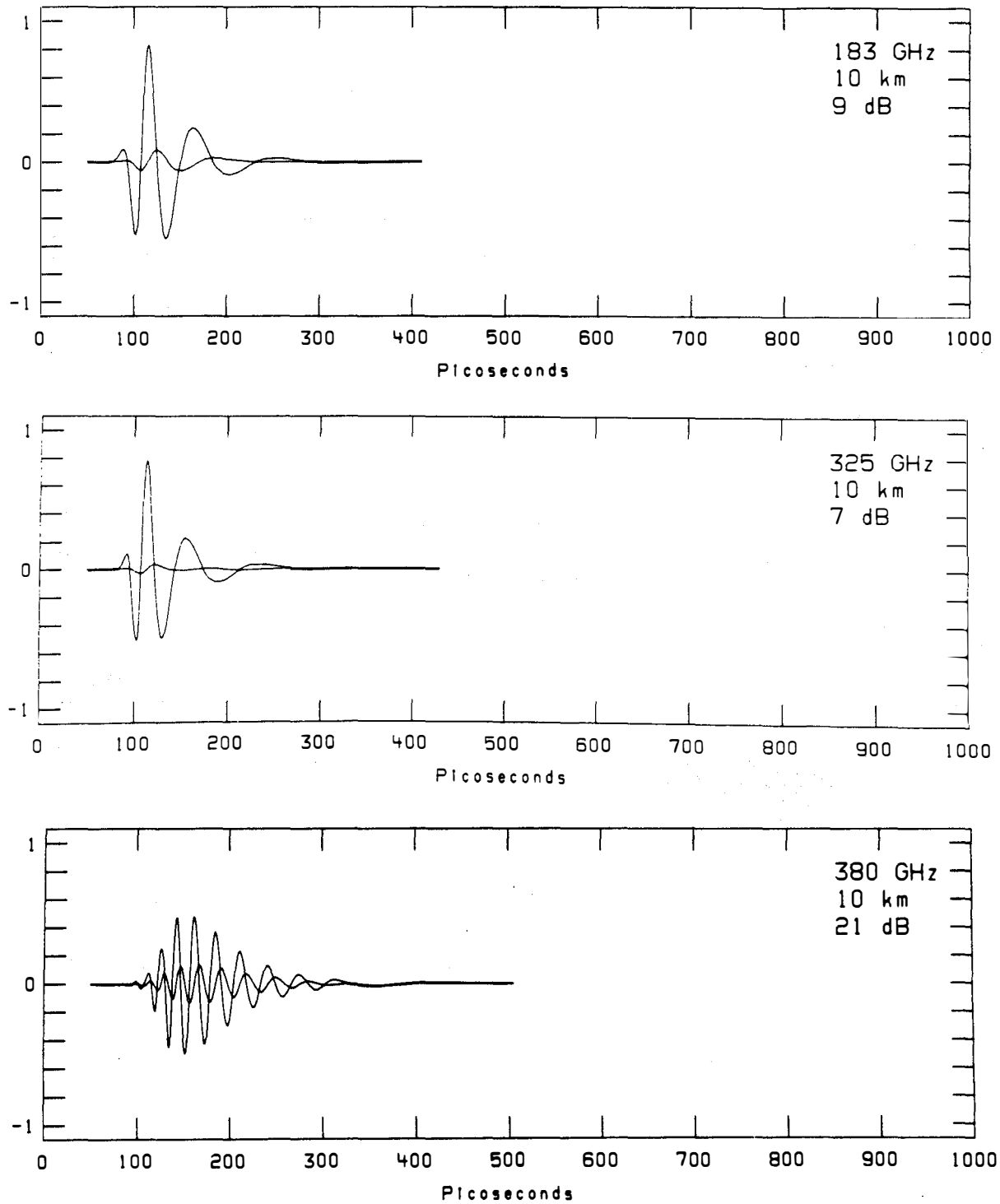
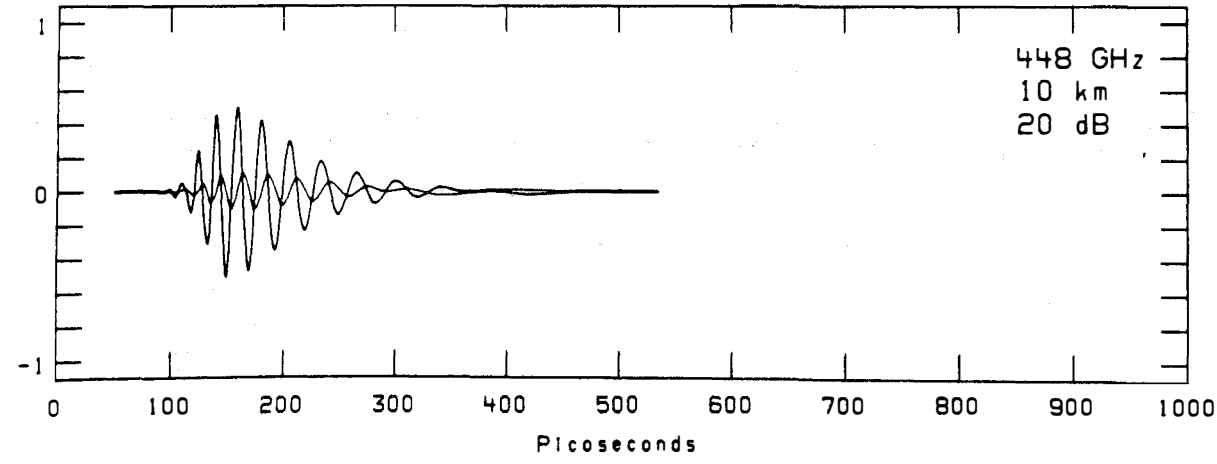


Figure 1. Continued. (Page 2 of 3).



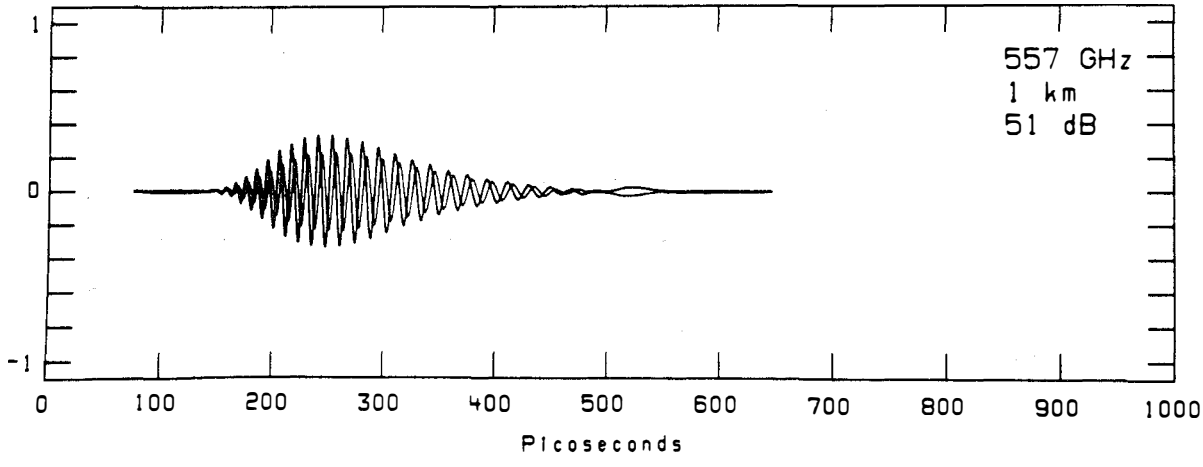


Figure 1. Continued. (Page 3 of 3)

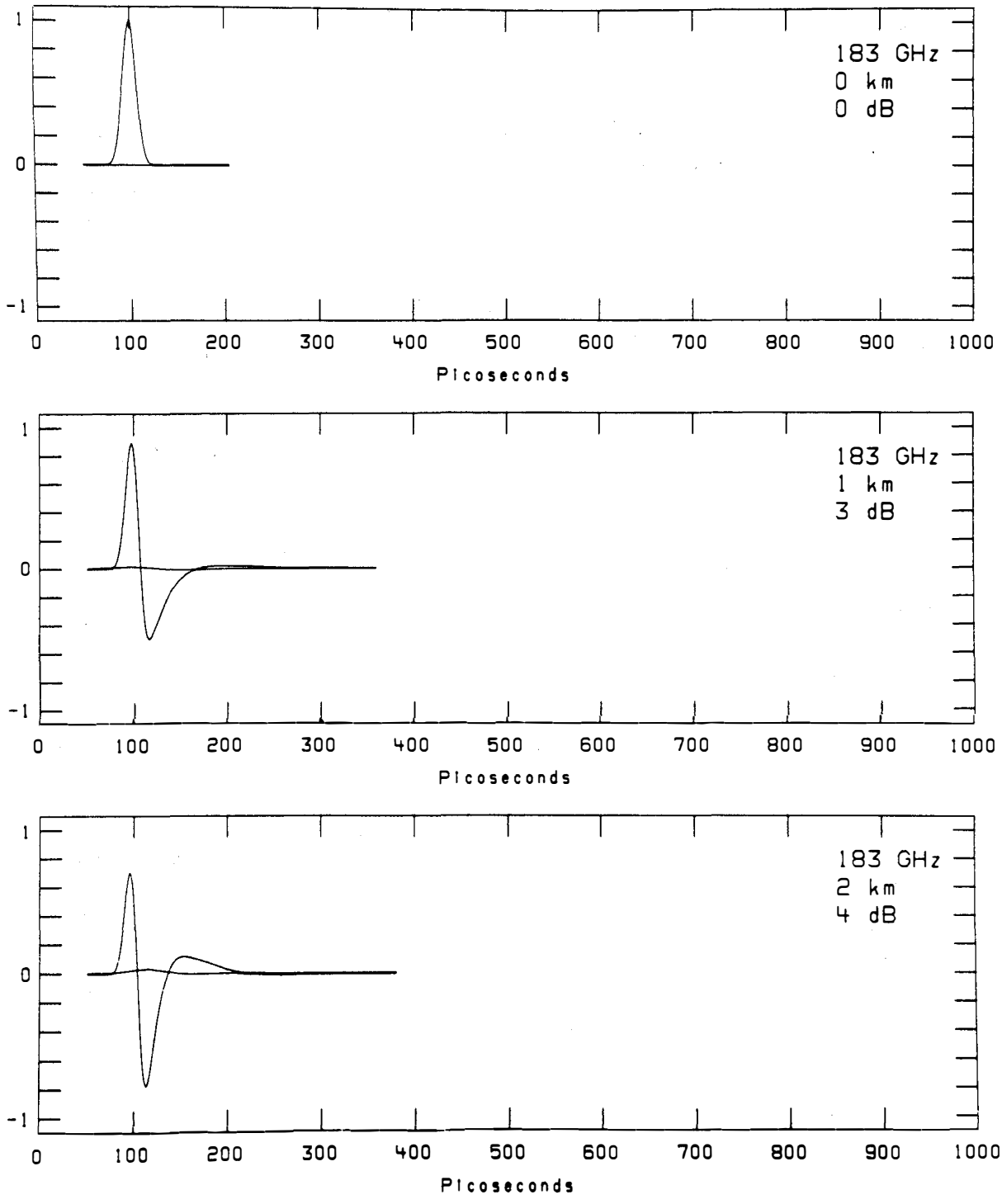


Figure 2. Pulse distortion at 183 GHz for a variety of distances. (This is page 1 of 3.)

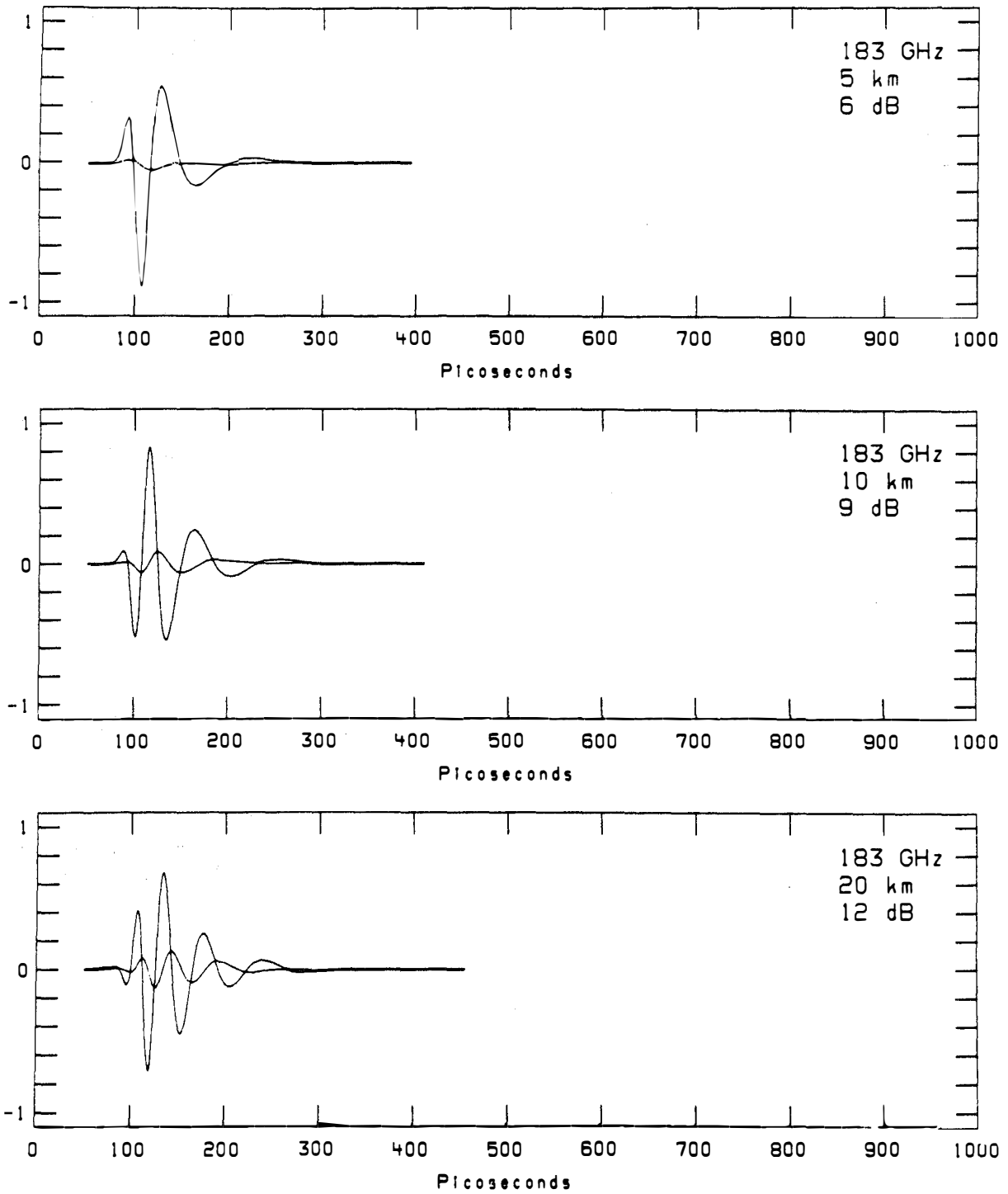
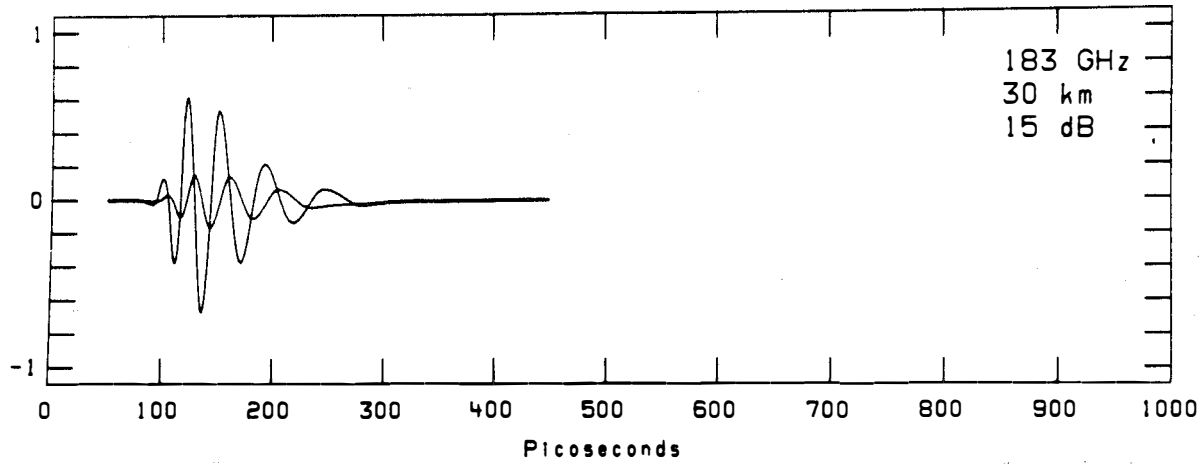
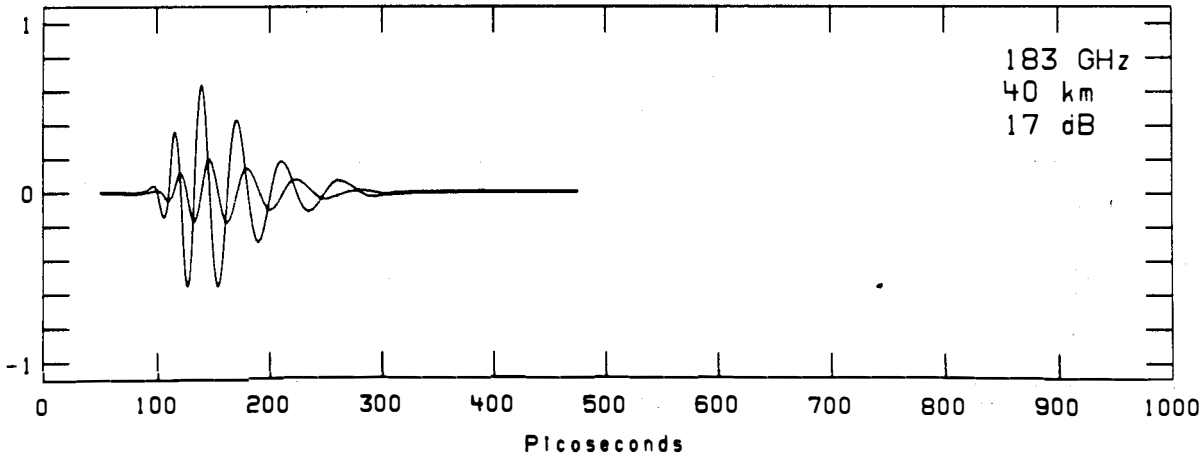


Figure 2. Continued. (Page 2 of 3.)





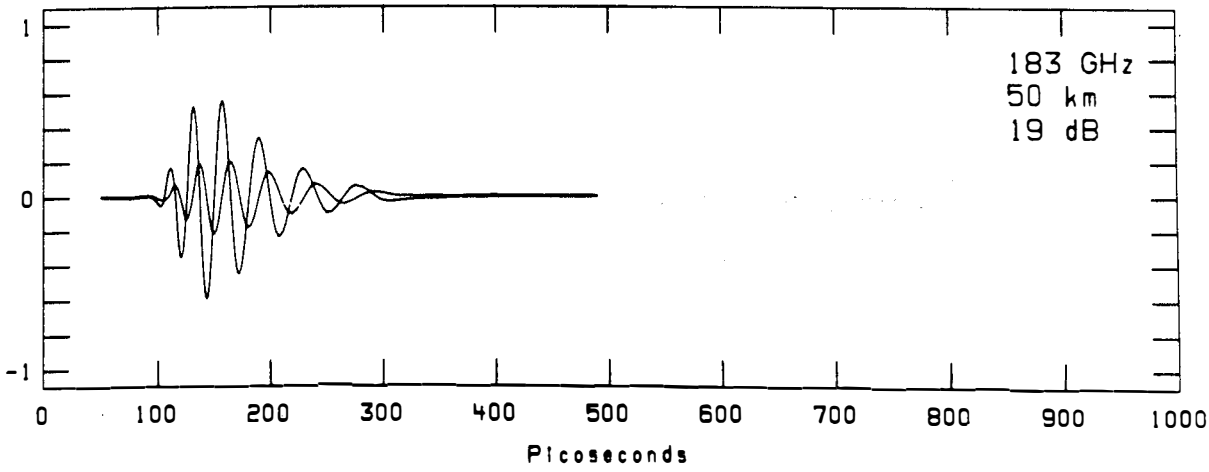


Figure 2. Continued. (Page 3 of 3.)

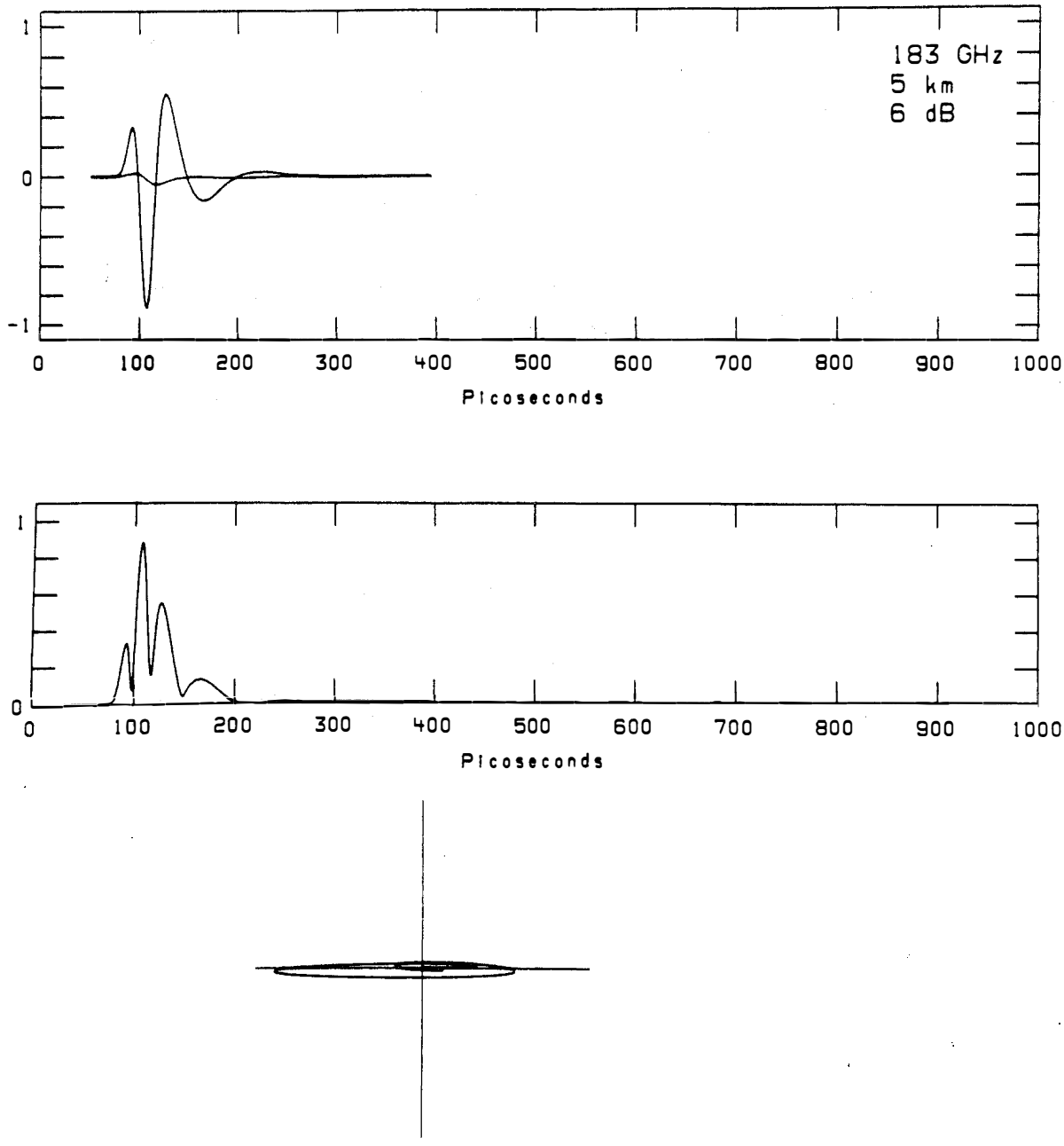


Figure 3. Pulse distortion at 183 GHz and 5 km, showing three different aspects of the same response. The two lower plots show amplitude distortion and phase distortion.

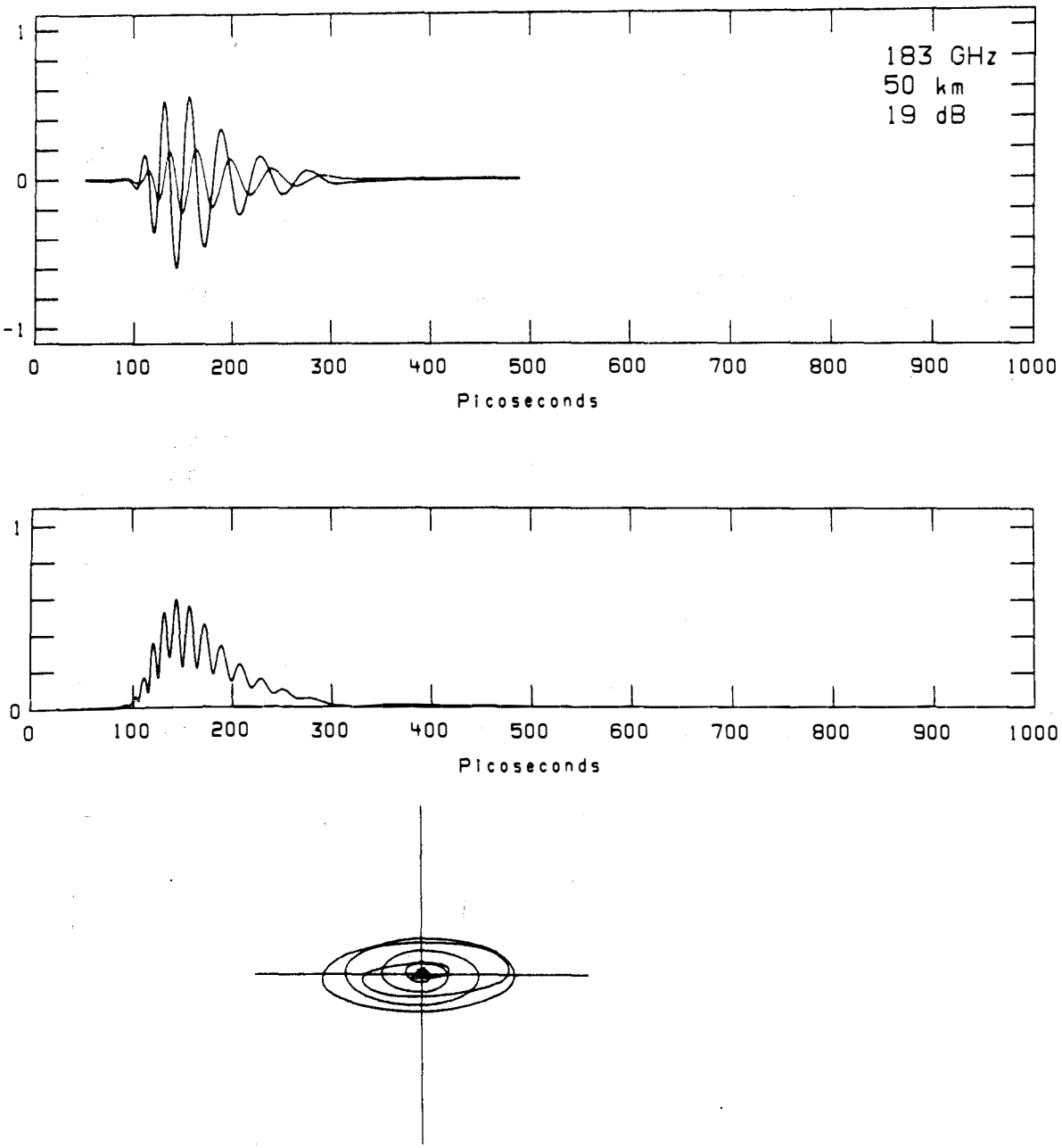


Figure 4. Pulse distortion at 183 GHz and 50 km, showing three different aspects of the same response.

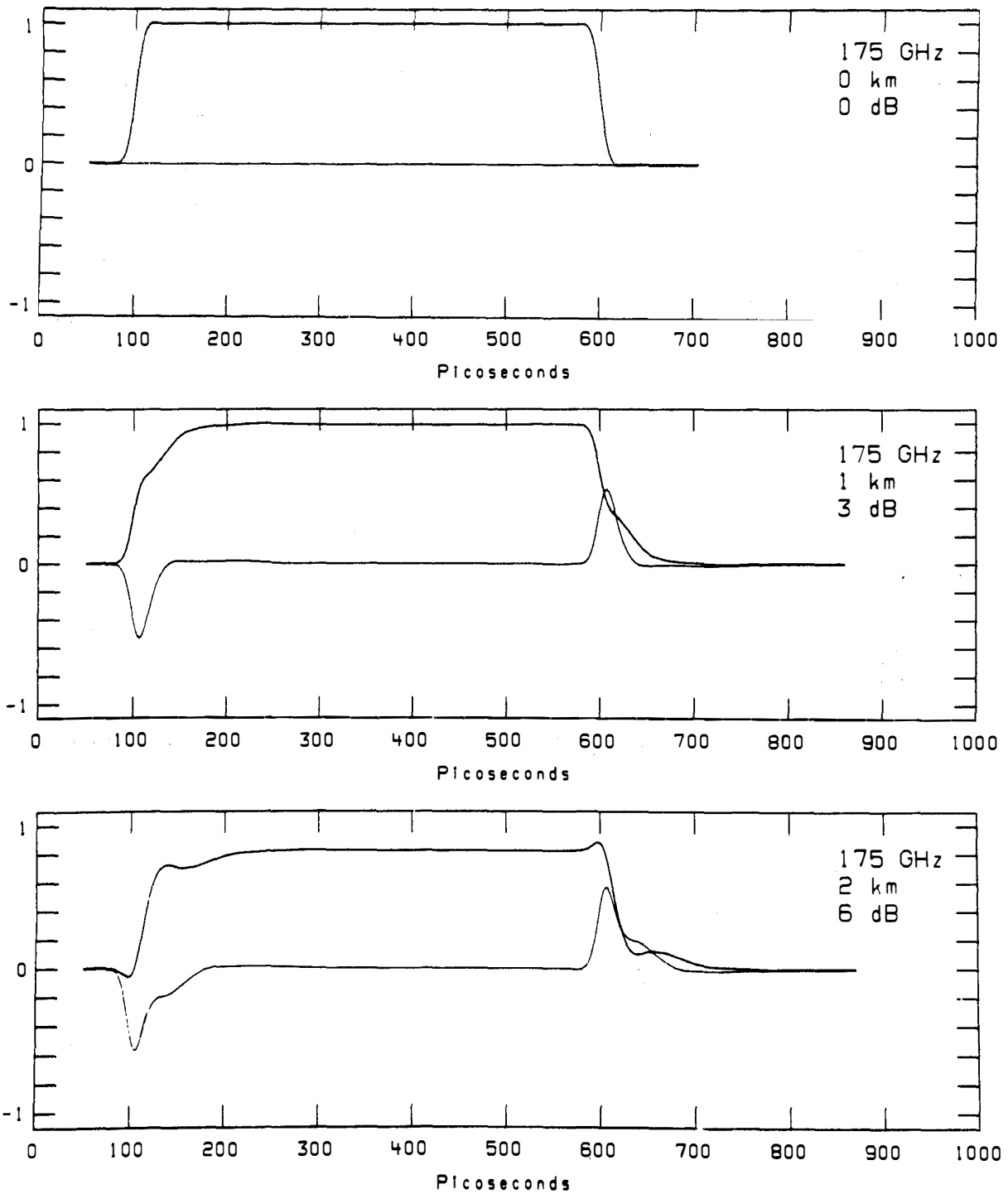


Figure 5. Distortion of a rectangular pulse. The central frequency is 175 GHz and the distance varies. (This is page 1 of 3.)

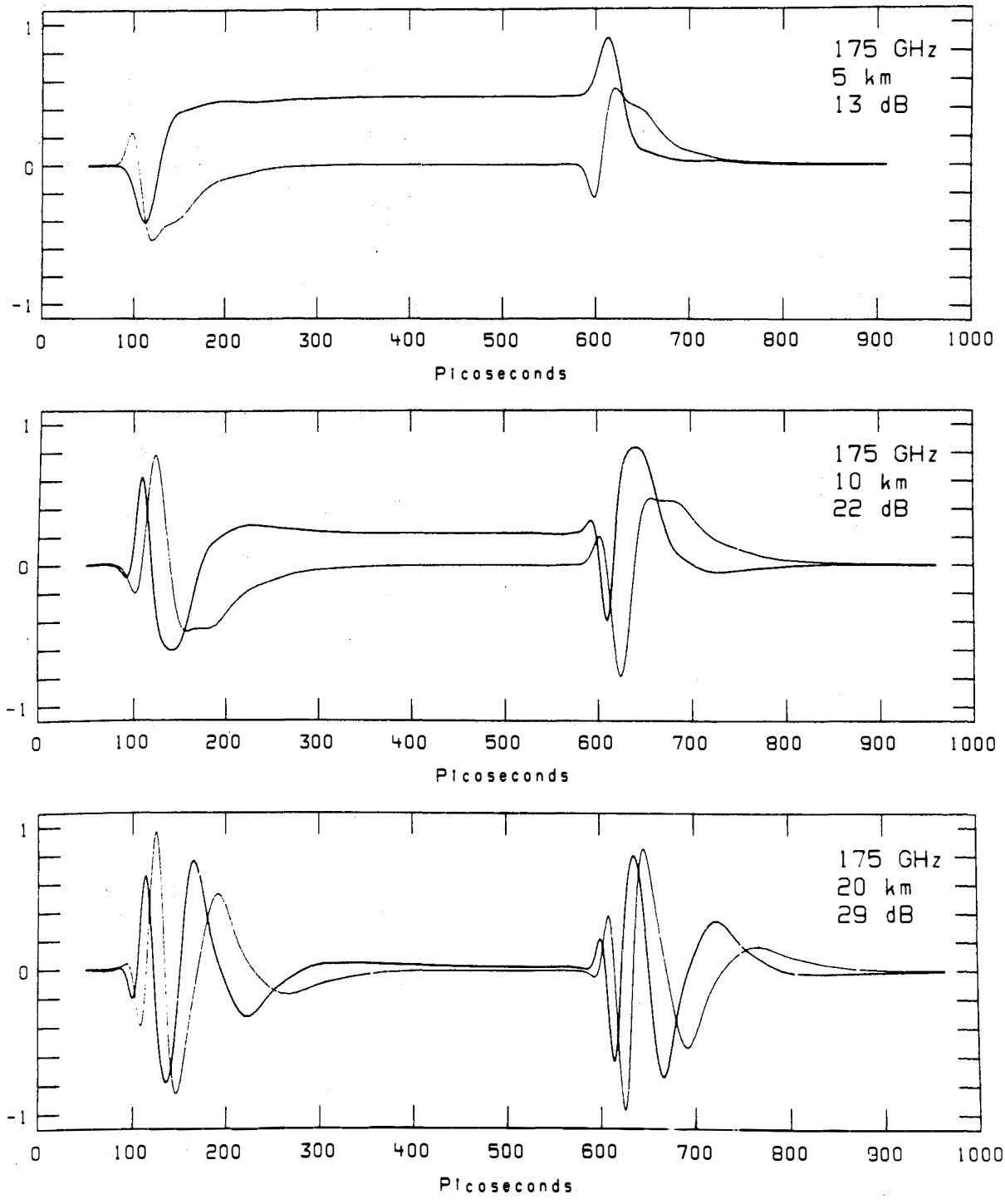


Figure 5. Continued. (Page 2 of 3.)

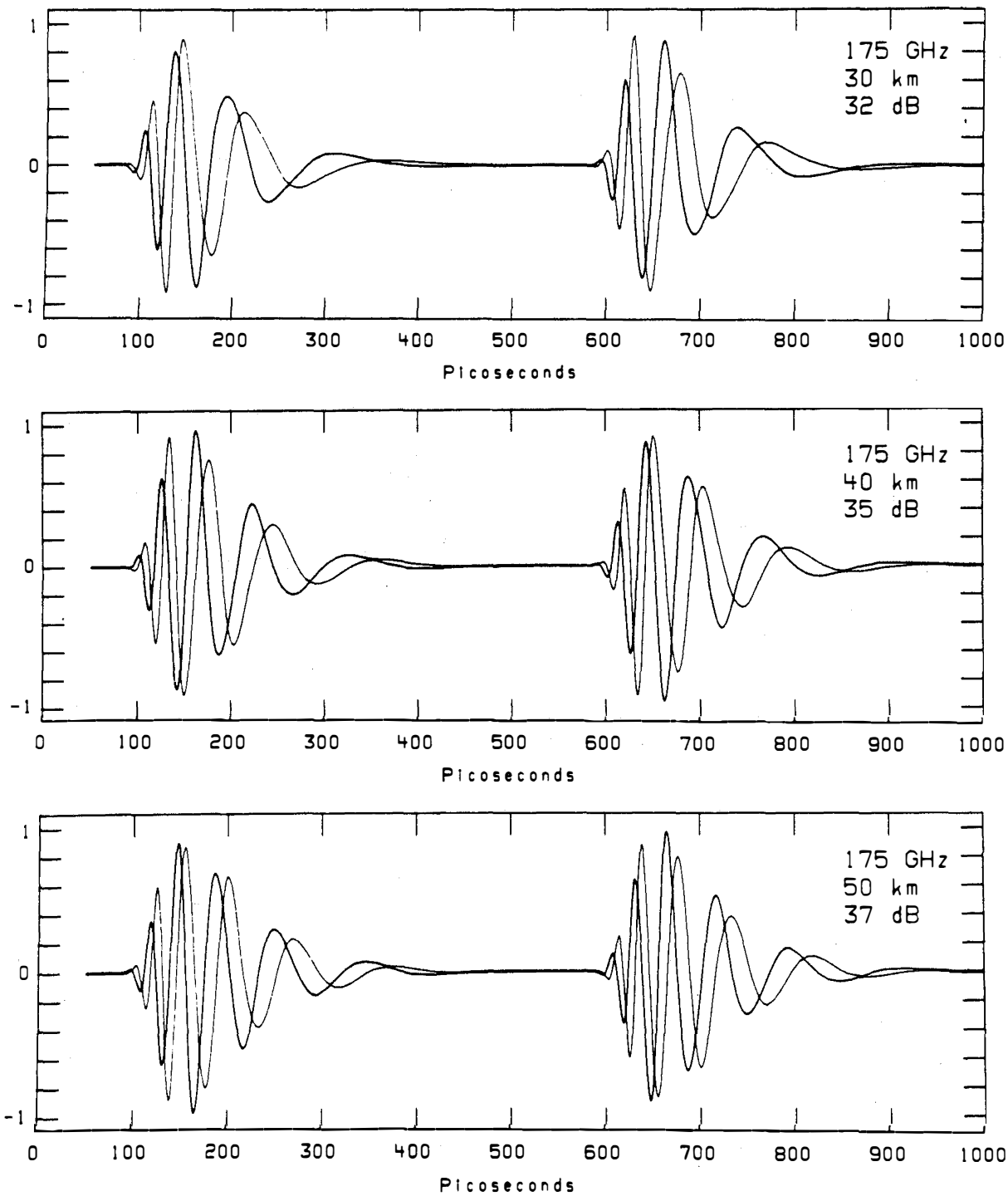


Figure 5. Continued. (Page 3 of 3.)

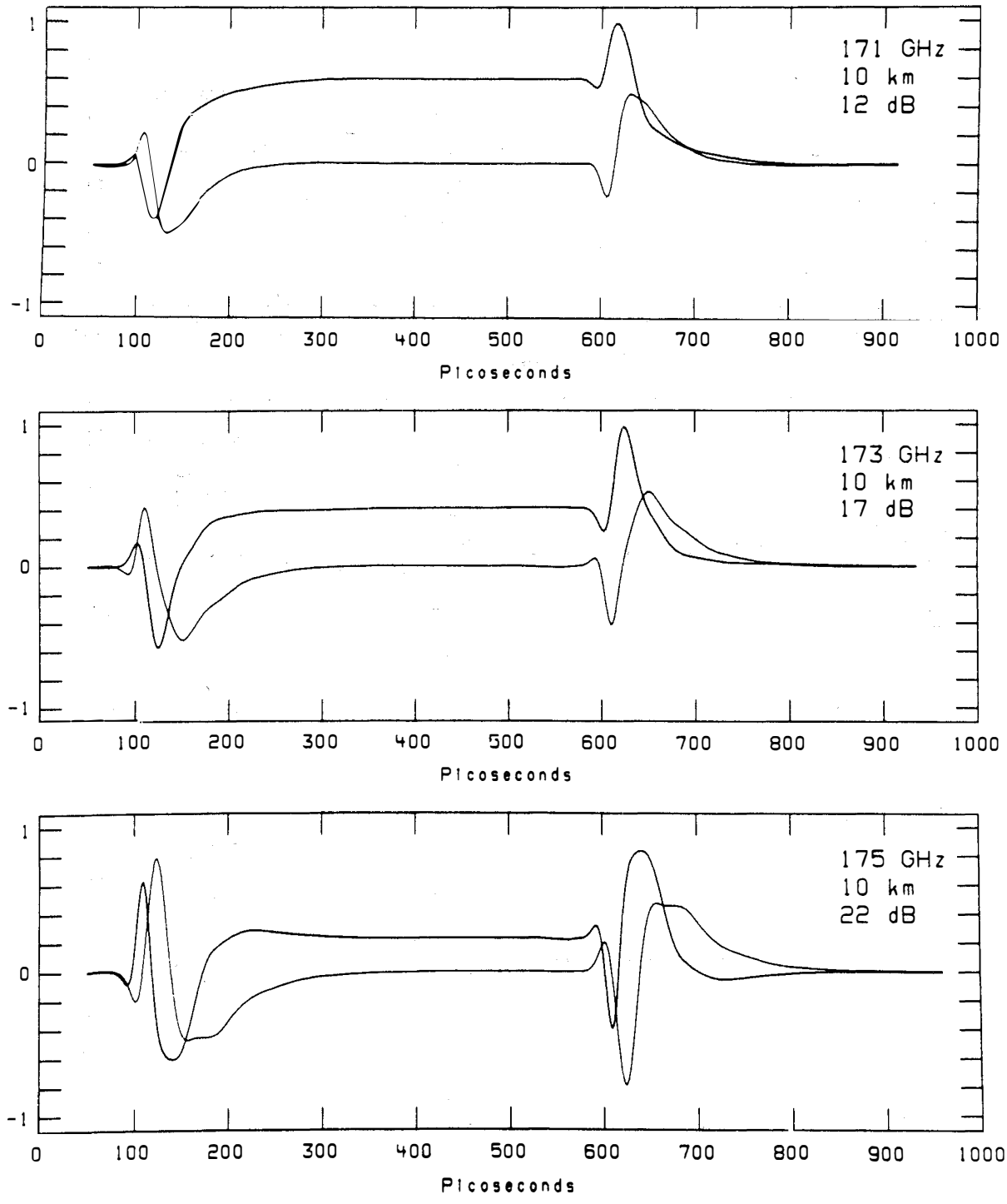


Figure 6. Distortion of a rectangular pulse. The distance is 10 km and the central frequency varies. (This is page 1 of 4.)

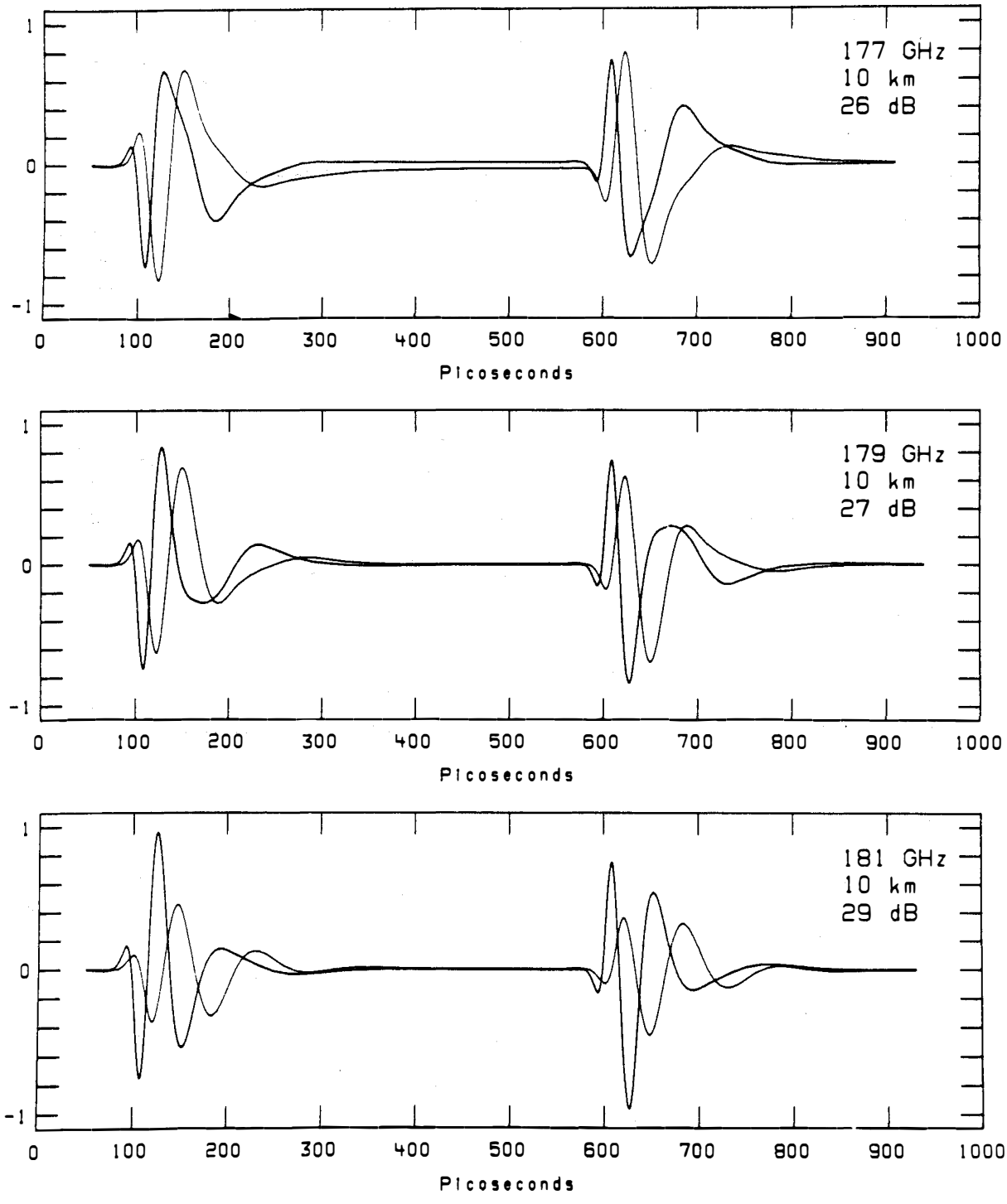


Figure 6. Continued. (Page 2 of 4.)

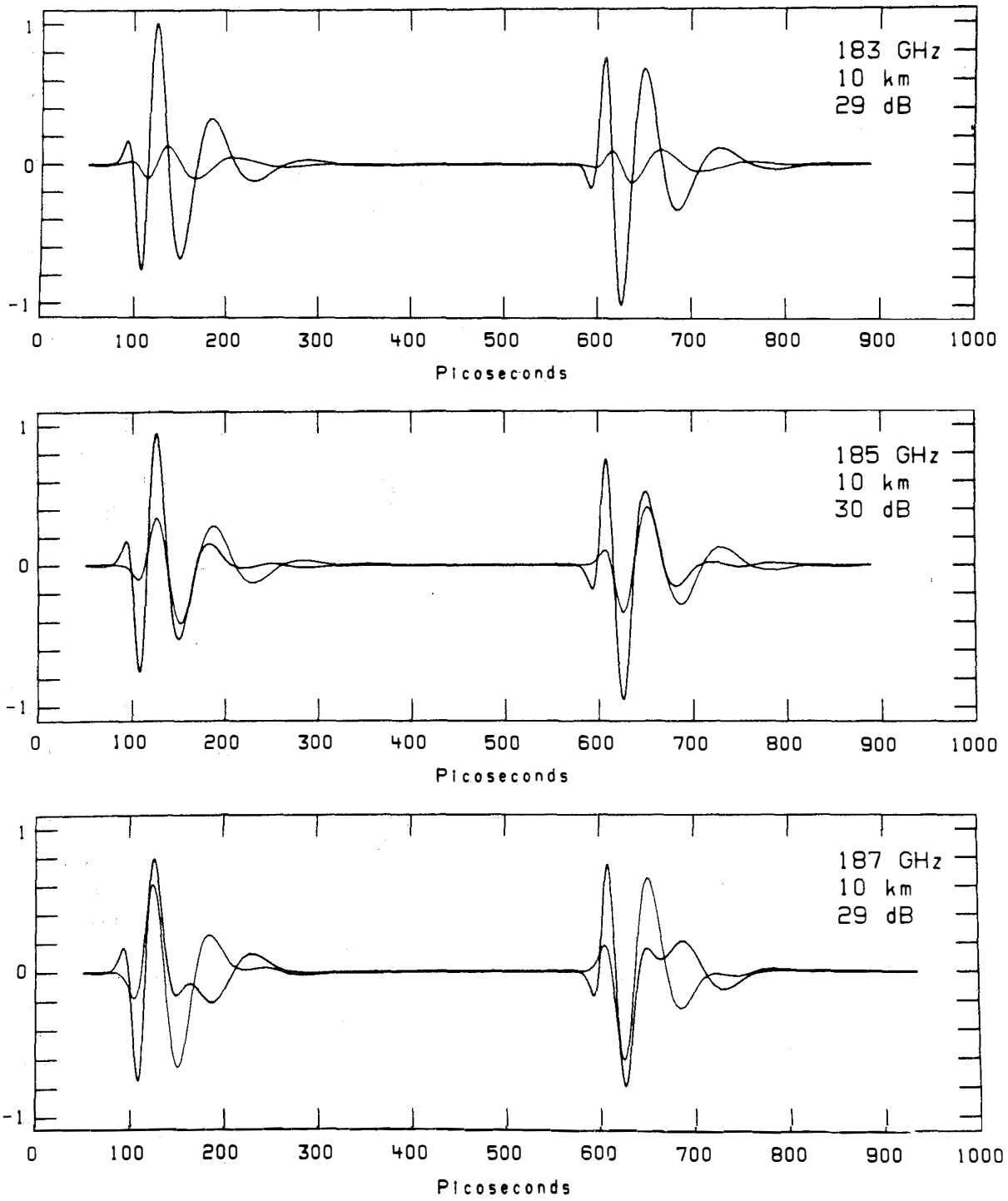


Figure 6. Continued. (Page 3 of 4.)

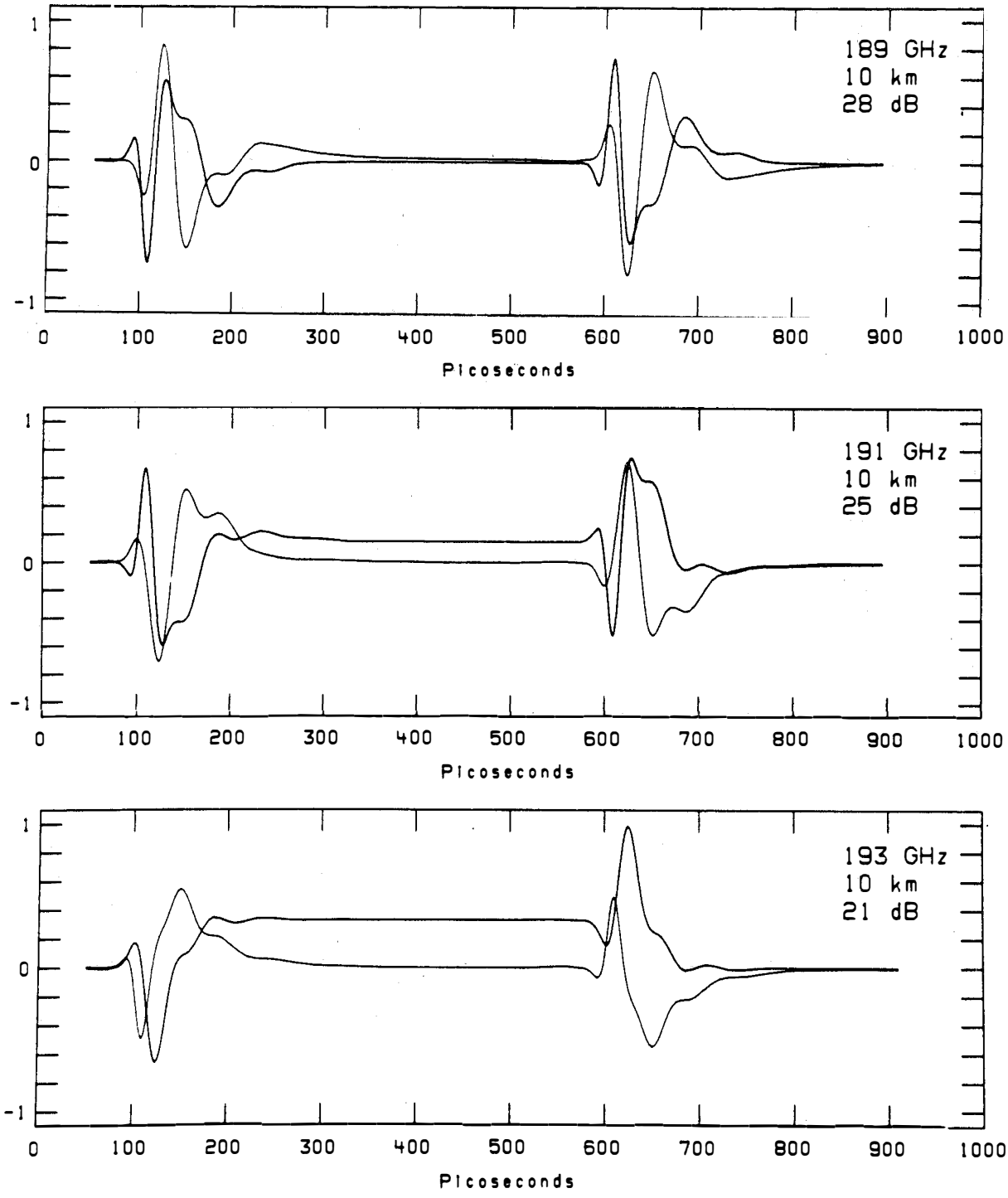


Figure 6. Continued. (Page 4 of 4.)

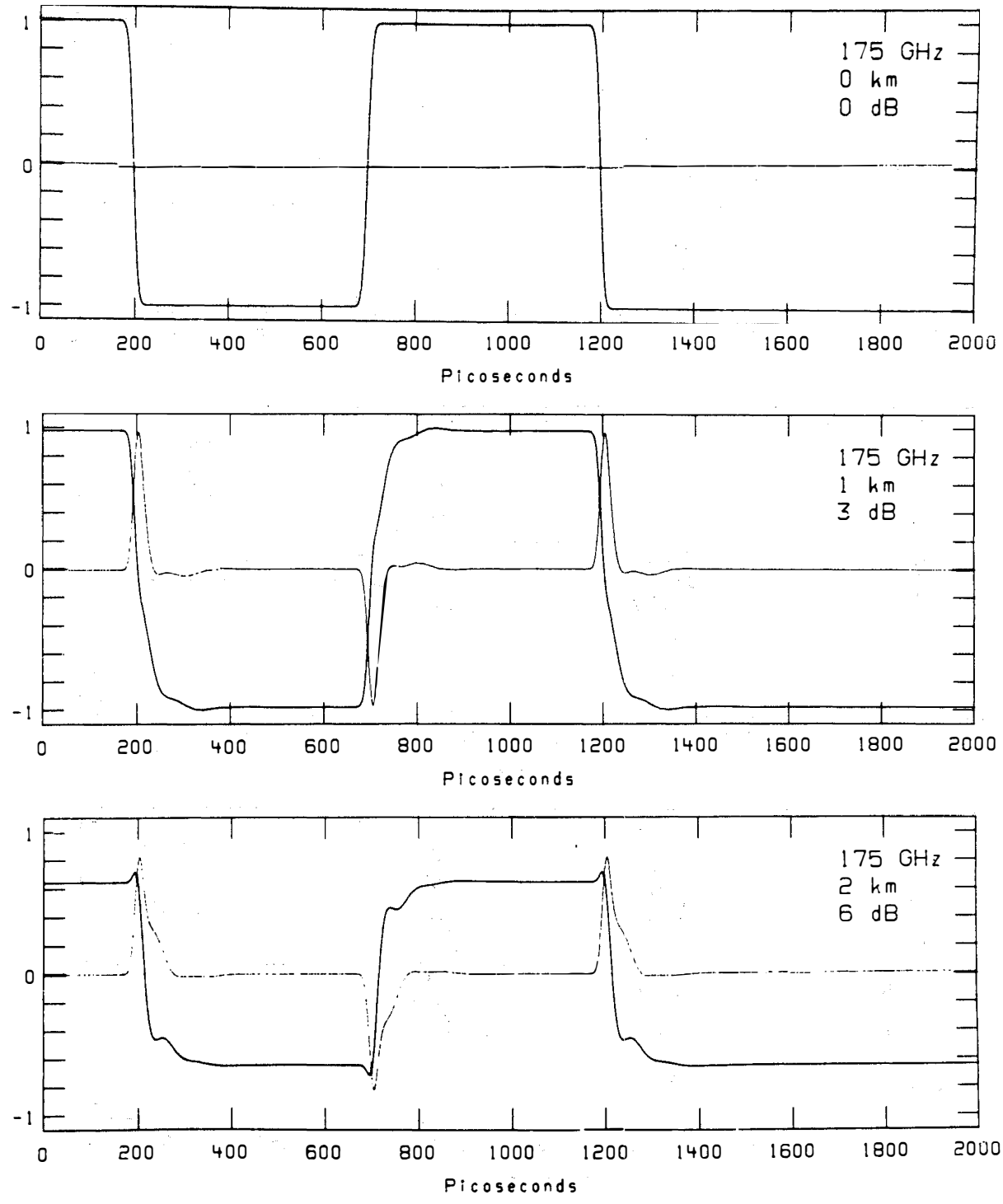


Figure 7. Distortion of a BPSK signal. (This is page 1 of 2.)

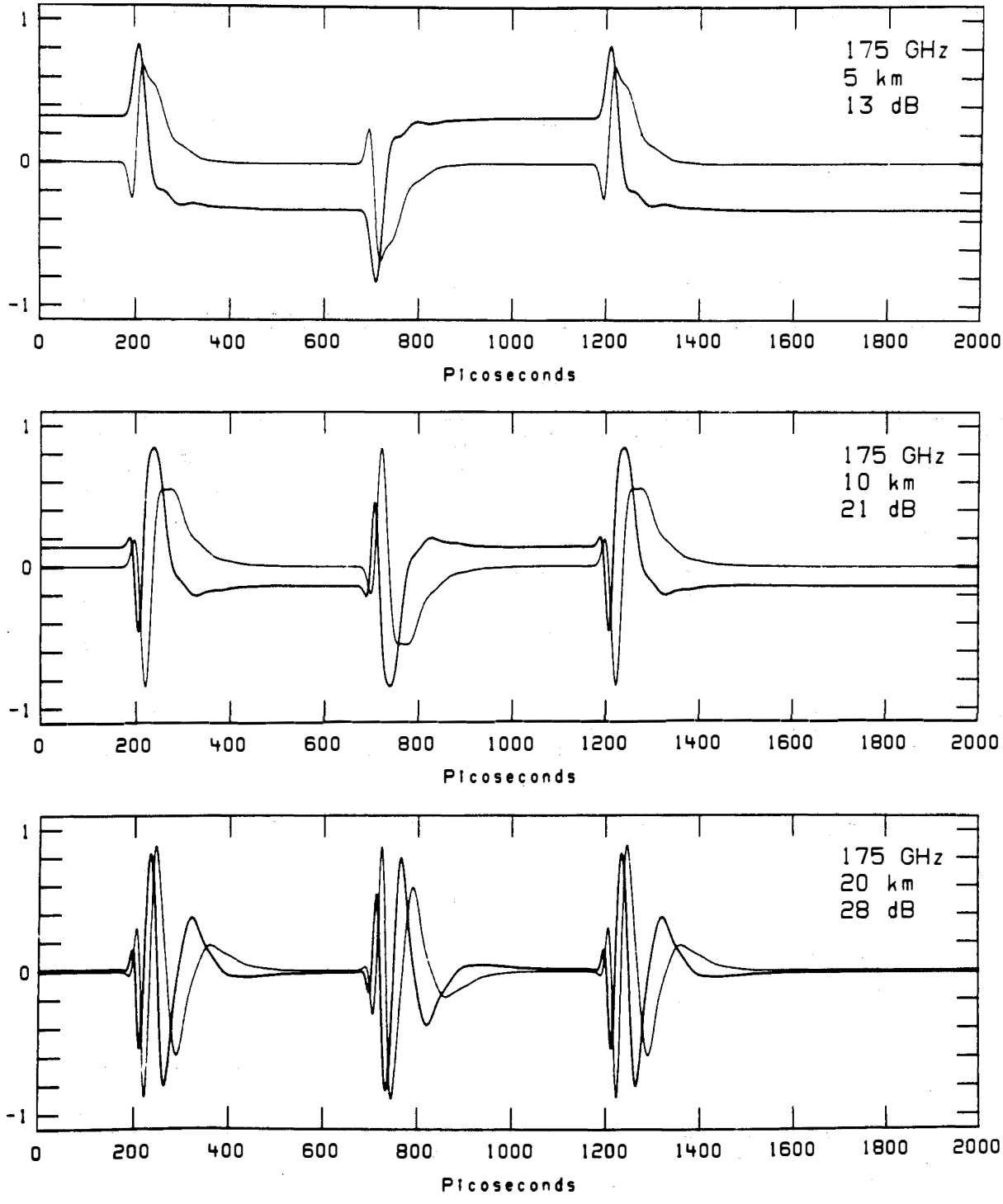


Figure 7. Continued. (Page 2 of 2.)

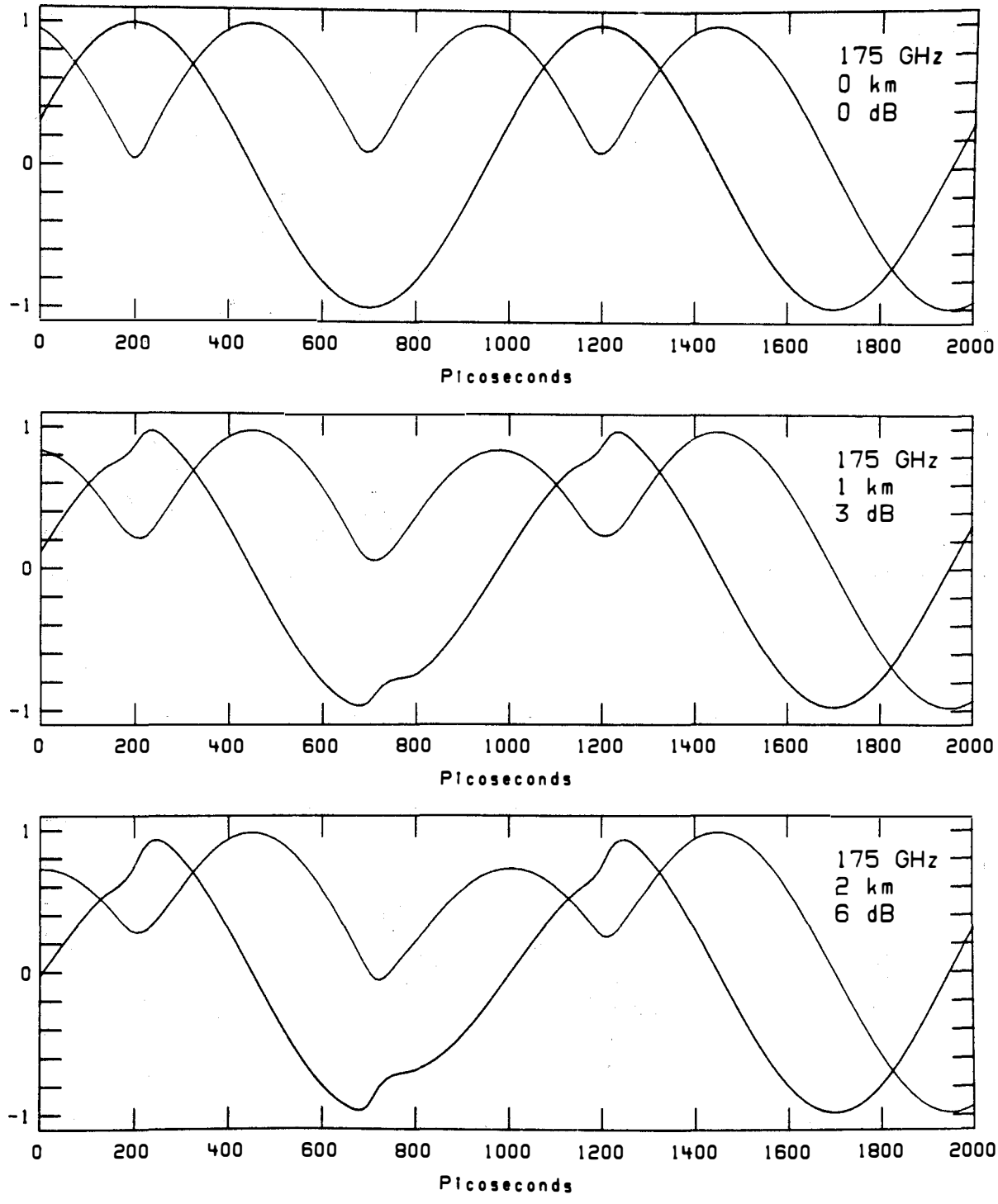


Figure 8. Distortion of an MSK signal. (This is page 1 of 2.)

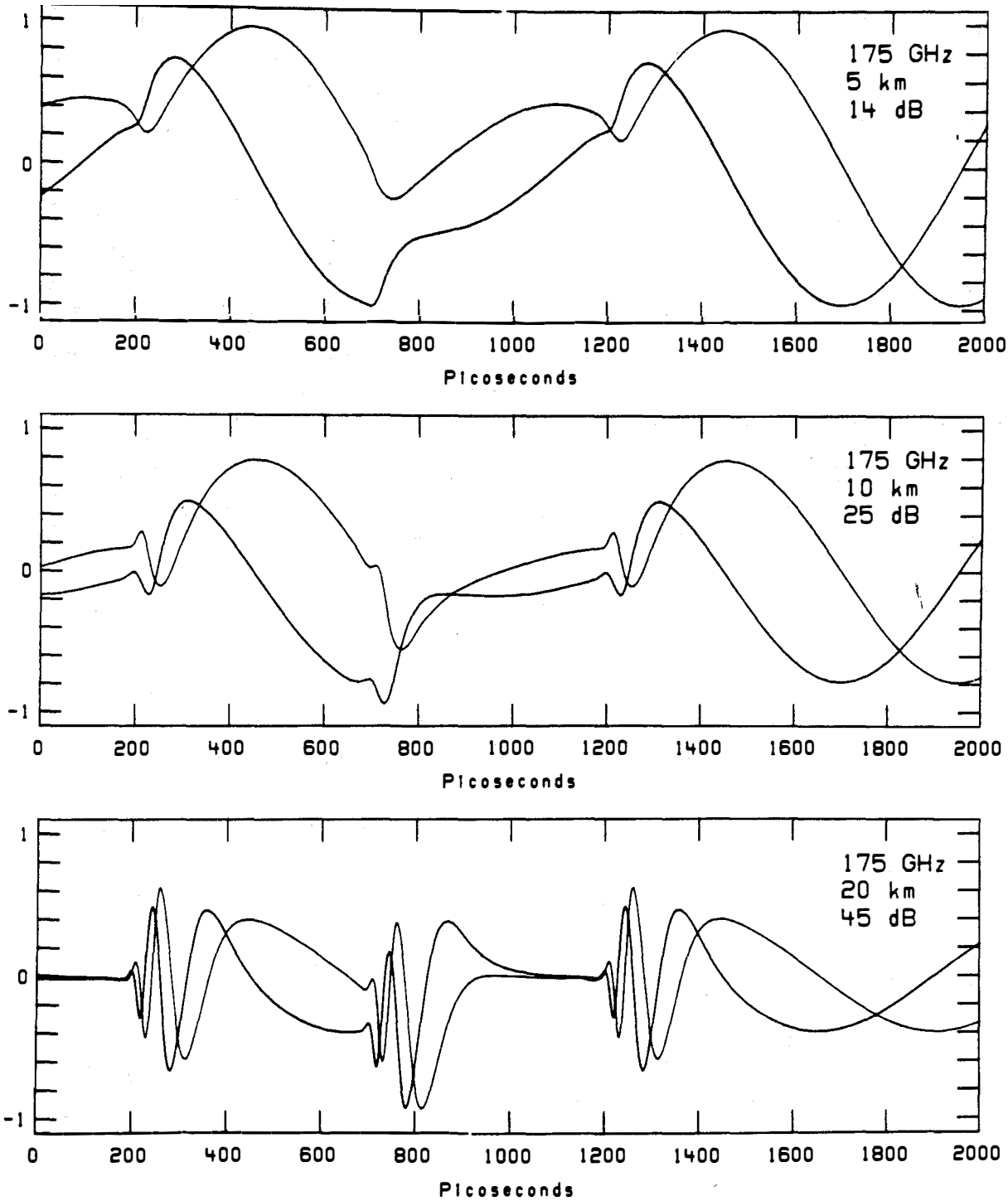


Figure 8. Continued. (Page 2 of 2.)

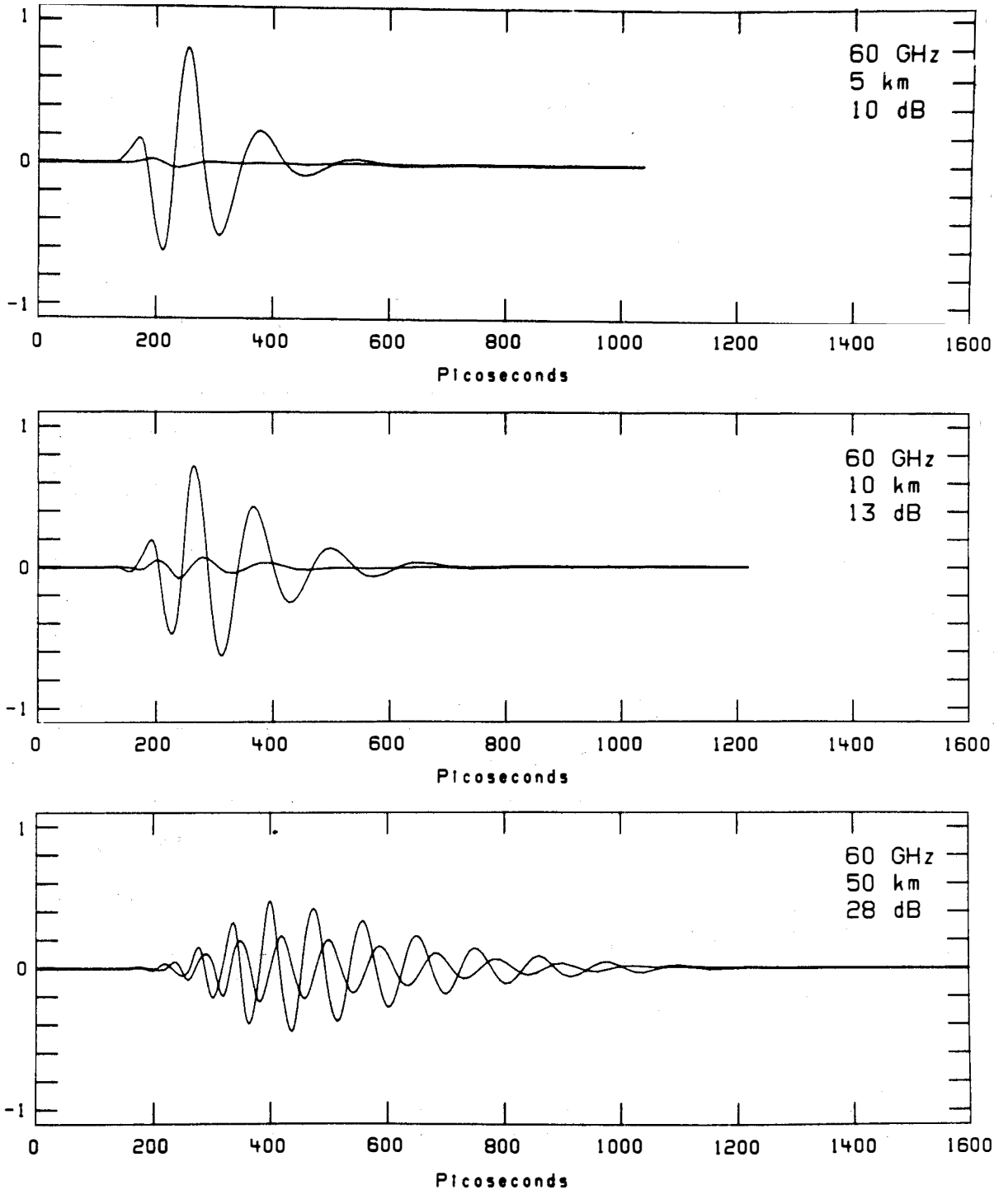


Figure 9. Distortion of a Gaussian pulse near 60 GHz when the atmosphere is represented by the MPM. (This is page 1 of 2.)

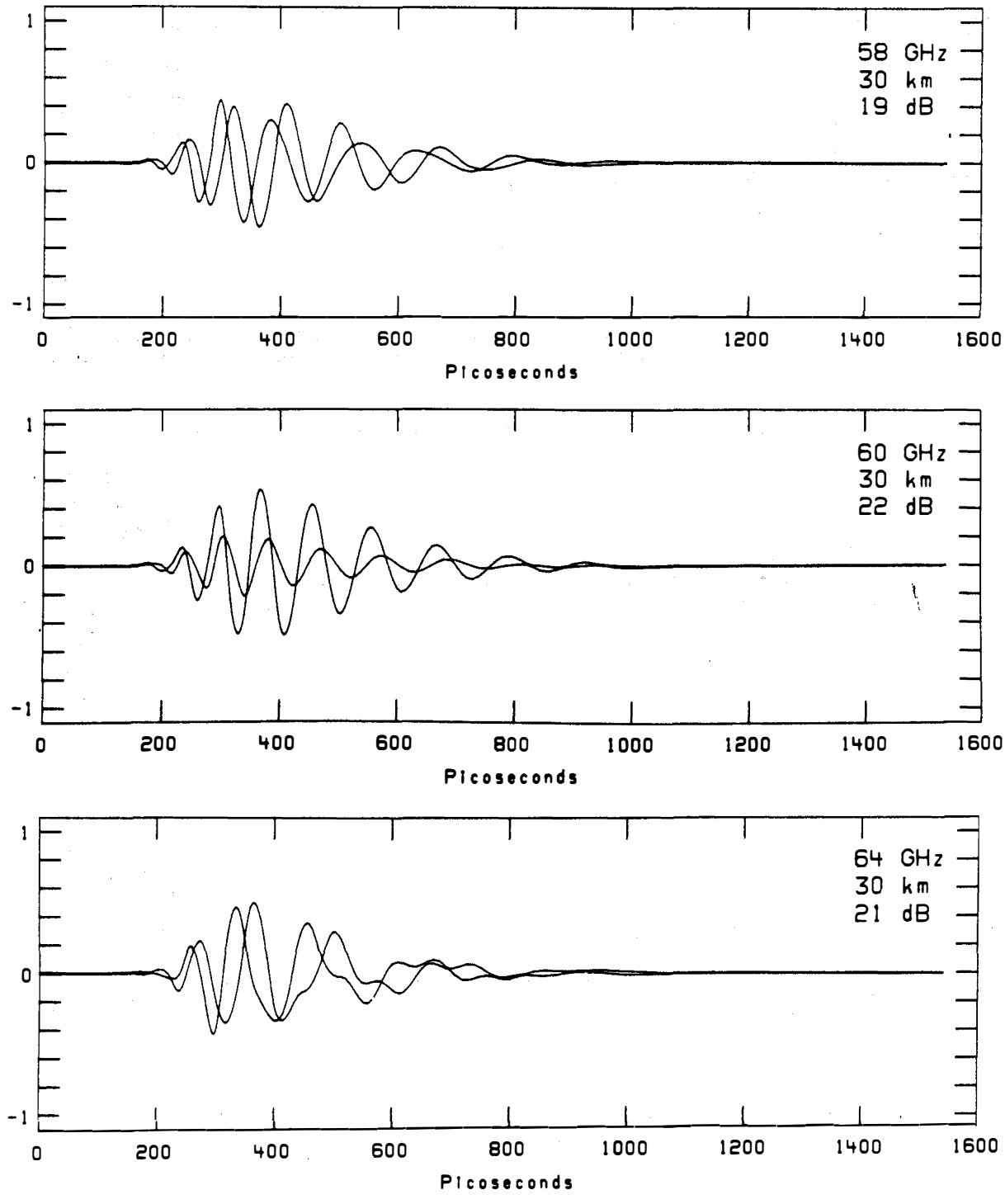


Figure 9. Continued. (Page 2 of 2.)

4. REFERENCES

Brillouin, L. (1914), Uber die Fortpflanzung des Lichtes in dispergierenden Medien, Ann. Physik 44, pp. 203-240.

Liebe, H. J. (1985), An updated model for millimeter wave propagation in moist air, Radio Sci. 20, pp. 1069-1089.

Sommerfeld, A. (1914), Uber die Fortpflanzung des Lichtes in dispergierenden Medien, Ann. Physik 44, pp. 177-202.

Stratton, J. A. (1941), Electromagnetic Theory (McGraw-Hill, New York, NY).

Trizna, D. B., and T. A. Weber (1982), Brillouin revisited: Signal velocity definition for pulse propagation in a medium with resonant anomalous dispersion, Radio Sci. 17, pp. 1169-1180.

Vogler, L. E. (1970), Pulse distortion in resonant and nonresonant gases, Radio Sci. 5, pp. 1301-1305.

APPENDIX: NUMERICAL ANALYSIS OF PULSE RESPONSES

The figures in this report have been computed and drawn by a digital computer using numerical techniques that are sometimes not entirely standard. It seems useful to document here some of these techniques.

The basic problem, of course, is to compute the convolution of a given pulse or input signal $p(t)$ with the impulse response $h(t)$ as expressed, for example, in (6), (7), and (8). This is most easily done using the services of the Fourier transform wherein the transform $P(\nu)$ of the input is multiplied by the spectral response $H(\nu)$ of the channel, and then the inverse Fourier transform applied. The approach is inviting not only because the spectral response is usually easier to compute, but also because of the opportunity to use the Fast Fourier Transform (the FFT).

The FFT is an elegant algorithm that describes how to quickly compute the Discrete Fourier Transform (DFT). To define the latter we suppose $x_0, x_1, \ldots, x_{n-1}$ is a sequence of n complex numbers. The DFT of order n is then a second sequence $X_0, X_1, \ldots, X_{n-1}$ given by

$$X_j = \sum_{k=0}^{n-1} e^{2\pi ijk/n} x_k \qquad (A.1)$$

The inverse to this transform is found to be

$$x_k = \frac{1}{n} \sum_{j=0}^{n-1} e^{-2\pi ijk/n} X_j \qquad (A.2)$$

Note that in both equations the right-hand sides have meaning whatever the values of j or k, respectively. If we suppose they take on all integer values then we obtain the "periodic extensions" of the left-hand sides, and indeed the DFT is often used as a transformation of infinitely long periodic sequences.

Our aim, then, is to recast the required Fourier transform so as to make use of the DFT. If $x(t)$ is to be a function that modulates the central frequency ν_c, we would write

$$X(\nu) = \int_{-\infty}^{\infty} x(t) e^{2\pi i(\nu-\nu_c)t} dt$$
$$\approx \sum_{0}^{n-1} x(j\delta t) e^{2\pi i(\nu-\nu_c)j\delta t} \delta t. \qquad (A.3)$$

The summation here is essentially the trapezoidal rule and assumes that $x(t)$ is sufficiently well approximated by samples taken at intervals δt and that it vanishes or nearly vanishes when $t < 0$ or $t \geq T = n\delta t$.

If we sample $X(\nu)$ at intervals $\delta\nu$ then

$$X(\nu_c + k\delta\nu) \approx X_k = \delta t \sum_{0}^{n-1} x_j e^{2\pi ijk\delta t\delta\nu} \qquad (A.4)$$

and so we have the DFT provided $\delta t\delta\nu = 1/n$. The inverse would be

$$x_j = \delta\nu \sum_{0}^{n-1} X_k e^{-2\pi ijk/n} \qquad (A.5)$$

which is very reasonable and essentially the trapezoidal rule applied to the inverse Fourier transform. We have said that ν_c is the "center" frequency, and for this to make sense we should think of k in both (A.4) and (A.5) as varying from $-n/2$ to $+n/2$. Thus the periodic nature of X_k enters immediately in an important way. In the usual enumeration where $k=0,\ldots, n-1$, one must imagine that the second half corresponds to frequencies less than ν_c and that they are better displayed when moved to the other side of $k=0$.

For the figures of this report we like to think of the pictured signal shapes as limited in both time and frequency. In time they differ sensibly from zero only in some (usually small) portion of the interval from 0 to T. And they are band limited so that their spectra differ sensibly from zero only in some portion of the interval $|\nu-\nu_c| < B/2$ where $B = n\delta\nu$ is the total bandwidth of frequencies available to the computations. This means that the x_j vanish near $j = 0$ and $j = n-1$ and that the X_k ($k=0,\ldots, n-1$) vanish near $k = n/2$. Thus (A.4) and (A.5) look a little more like the trapezoidal rule and, even better, like the trapezoidal rule with end correction. The latter is a quadrature rule (see, e.g., Lanczos, 1956, ch. 6) that uses additional information concerning the derivatives of the integrand at the two endpoints and so becomes a rule of much higher order. Various kinds of error estimates may be found in, for example, the survey article by Henrici (1979).

Note that $\delta t = 1/B$ and $\delta\nu = 1/T$. From the Nyquist criterion it follows that if the above restrictions are adhered to, then without any further qualifications both the signal and its spectrum are automatically sampled often enough to assure fidelity.

The Gaussian pulses used in Figures 1-4 were designed to approximate an impulse and hence to be as short as possible while still satisfying some restrictions. The pulse and its spectrum are given by

$$p(t) = e^{-(t/\tau)^2} \qquad (A.6)$$

$$P(\nu_c + \nu) = \sqrt{\pi}\,\tau\, e^{-(\pi\nu\tau)^2}$$

where the parameter τ defines the pulse width so that, for example, 2τ is the time between 1/e levels.

In choosing a value for τ there are some three considerations. First, we should require $B < 2\nu_c$ so that there are no negative frequencies in our calculation and our claim to be examining "narrowband" signals has merit. Second, the actual bandwidth of the pulse should be contained well within the computational bandwidth so that the quantity $\rho = P(\nu_c \pm B/2)/P(\nu_c)$ is very small. And third, since $\nu_c\tau$ is the number of rf cycles in the time interval τ it should probably be somewhat larger than 1.

Let $\tau = \alpha\delta t$ and $B = \beta\nu_c$. Then the above considerations become (i) $\beta < 2$, (ii) $\rho = \exp[-(\pi\alpha/2)^2]$ is much less than 1, and (iii) $\nu_c\tau = \alpha/\beta > 1$. With $\alpha = 1.5$ we find $\rho = 3.9\times10^{-3}$ and we feel that this is about the smallest value of α that one can use. We generally set β equal to approximately 1. Thus for the 22-GHz line in Figure 1 we have set B = 40 GHz and 2τ = 75 ps, while for frequencies above 300 GHz we have B = 200 GHz and 2τ = 15 ps. For the 183-GHz line used in all four figures we used B = 200 GHz and 2τ = 20 ps.

One further detail we have used assures that the spectrum vanishes at the edges of the total band of computations. We simply evaluate $P(\nu_c \pm B/2)$ and subtract it as a constant from the whole spectrum. When interpreted as a sequence P_k of sampled values this process will assure $P_{n/2} = 0$. Equivalently, we may subtract a similarly small amount from the peak value of the sampled time sequence p_j.

In addition, these same ideas were used in preparing the more complicated signals in Figures 5 through 8. The Gaussian pulse is, of course, equivalent to the response obtained by passing an impulse through a filter whose frequency response is the function $P(\nu)$ in (A.6) (or perhaps modified according to the previous paragraph). The sharp corners and jump discontinuities of our signals have all been smoothed by the application of just such a filter. In this way, our signals are all band limited.

In all these considerations we have not mentioned n and how large it must be. With the other parameters determined, to choose n it is now only necessary to require that T be large enough. This means it should be large enough to include the channel response, since it is that we will eventually want to plot. The value of n is thus determined not so much by the characteristics of the input pulse as by those of the channel.

Equivalent to requiring that T be large enough is requiring $\delta\nu$ to be small enough. It must be small enough to follow the variations in the frequency response of the channel. In particular, it must be somewhat less than the line width γ_o, but also it must be small enough that the changes in the exponent $\nu nx/c$ in (1) are only a fraction of a cycle. We have set n = 512 for the lower frequencies and n = 1024 for the higher. For the 557-GHz line we were forced to use n = 2048.

We have already remarked in this report that the vertical scale of the plots is an artificial one, and that the pictured responses have been normalized. For the Gaussian pulses the normalization has been such that the total energy in the resulting response equals the total energy of the input pulse--i.e., a Gaussian pulse of maximum amplitude 1. For the more extended signals the normalization has simply been so that the maximum amplitude of the response is 1.

This normalization process resolves only one of the ambiguities involved in the representation of response functions. A second ambiguity has to do with where to place the response on the horizontal axis. The origin of the time axis must be considered artificial and should be chosen only to properly display the response function. We recall that a shift of the response by a time delay s multiplies the spectrum by the factor $\exp 2\pi i\nu s$.

A third ambiguity has to do with the phase of the response. There is no good definition for the absolute phase of a narrowband signal, so we are at liberty to treat it as artificial. In the end we can multiply the complex response function by any phase-changing constant factor.

With the refractive index as in (2), the exponent in (1) can be rewritten as

$$\nu nx/c = \nu n_o x/c - \mu_o t_o + \frac{\mu_o^2 t_o}{\mu_o - \nu}. \qquad (A.7)$$

In the actual computations, therefore, we have ignored the first two terms here, thus displacing the response along the time axis and introducing a constant phase shift. It turns out that if we do nothing else the response has the same apparent onset time as the input pulse and its phase seems pleasing enough. This is evidently because the last term in (A.7) tends to zero as ν tends away from ν_o and it does this rapidly enough to have approached its final value well within the boundaries of our very large band.

This is the treatment, or rather the lack of treatment, that produced the Gaussian pulse responses. For the more extended signals we found it useful to reorient the phase somewhat. In the case of the rectangular pulses, for example, we shifted phases so that when the signal reaches its steady value it is real and positive.

REFERENCES

Henrici, P. (1979), Fast Fourier methods in computational complex analysis, SIAM Rev. 21, pp. 481-527.

Lanczos, C. (1956), Applied Analysis (Prentice Hall, Englewood Cliffs, NJ).

FORM **NTIA-29**
(4-80)

U.S. DEPARTMENT OF COMMERCE
NAT'L. TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION

BIBLIOGRAPHIC DATA SHEET

	1. PUBLICATION NO.	2. Gov't Accession No.	3. Recipient's Accession No.
	NTIA Report 87-229		

4. TITLE AND SUBTITLE

Millimeter-Wave Pulse Distortion by a Single Absorption Line Simulating the Terrestrial Atmopshere

5. Publication Date
October 1987

6. Performing Organization Code
NTIA/ITS

7. AUTHOR(S)
George Hufford

9. Project/Task/Work Unit No.
910 8108

8. PERFORMING ORGANIZATION NAME AND ADDRESS
National Telecommunications and Information Administration
Institute for Telecommunication Sciences
325 Broadway
Boulder, CO 80303 3328

10. Contract/Grant No.

11. Sponsoring Organization Name and Address
National Telecommunications and Information Administration
Herbert C. Hoover Building
14th & Constitution Avenue, NW
Washington, D.C. 20230

12. Type of Report and Period Covered

13.

14. SUPPLEMENTARY NOTES

15. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

Molecular absorption lines in the millimeter-wave band make the atmosphere dispersive, thus causing the distortion of transmitted signal shapes. To better understand how severe this distoriton is and what form it takes, it seems useful to examine the abstract problem of a single, isolated absorption line. The approximate response to such a line may be easily computed and is well described by only a few parameters. We have evaluated these parameters and rather completely characterized the distortion for eight of the more prominent lines below 600 GHz.

16. Key Words *(Alphabetical order, separated by semicolons)*

absorption line; atmosphere; frequency distortion; frequency response; impulse response; millimeter-wave propagation; radio signal distortion

17. AVAILABILITY STATEMENT	18. Security Class. *(This report)*	20. Number of pages
☒ UNLIMITED.	Unclassified	44
☐ FOR OFFICIAL DISTRIBUTION.	19. Security Class. *(This page)*	21. Price:
	Unclassified	

☆ U.S. Government Printing Office: 1980—678-495/529