THE UNIVERSITY OF CALGARY

ANTENNA DIVERSITY IN ANALOG FM CELLULAR RADIO

by

Elizabeth Barbara Olasz

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF

ELECTRICAL & COMPUTER ENGINEERING

CALGARY, ALBERTA

JULY, 1991

ISBN 0-315-71176-0

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "Antenna Diversity in Analog Cellular FM Radio", submitted by Elizabeth Barbara Olasz in partial fulfillment of the requirements for the degree of Master of Science.

Michel Fattouche

Dr. M. Fattouche, Supervisor
Dept. of Electrical and Computer Engineering

Dr. S.T. Nichols
Dept. of Electrical and Computer Engineering

Dr. A. Sesay
Dept. of Electrical and Computer Engineering

Gérard Lachapelle

Dr. G. Lachapelle
Dept. of Surveying Engineering

Date: July 18, 1991

Abstract

In a mobile analog cellular system the transmitted signal encounters a continuously changing number of propagation paths which distorts the signal so that the reception is affected at the receiver. These distortions are time varying in nature and are difficult to compensate for.

Antenna diversity is suggested as a means of improving the reception quality of cellular FM radio in fading environments. The performance of diversity in cellular FM radio is analyzed using a digitally simulated FM system which includes such components as a compandor, pre-emphasis and de-emphasis, limiter, filters as well as the modulator and limiter-discriminator type demodulator. The propagation model used is a modified Hashemi model representative of a mobile urban multipath environment.

The diversity combining methods include: selective switching using a switch and stay strategy, audio level combining employing two receivers and equal ratio combining using a co-phasing technique. These are compared to the performance of the system with no diversity. The simulation using digitized speech provides a method to compare the performance of various diversity combining schemes. Comparisons are based on output signal to noise ratios which are determined from mean square comparisons between input and output sources as well as from synchronous sampling methods for tone modulation. Results indicate that diversity using equal ratio combining provide desirable performance enhancements.

ACKNOWLEDGEMENTS

The author wishes to express appreciation to Dr. M. Fattouche for accepting the author as his student, providing the opportunity to do this work and for all the understanding given, to Dr. S.T. Nichols for his helpful suggestions and to Dr. K.E. Scott for his assistance and encouragement. The author also wishes to acknowledge and thank NovAtel for funding, equipment and for suggesting the direction of the research and the Department of Electrical and Computer Engineering at the University of Calgary for additional financial support.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter One

Introduction

Analog frequency modulation (FM) systems are widely used in current 800 MHz cellular radio systems. Although much interest has been expressed in digital cellular systems there are still many details, problems and even system structure which need to be considered. This delay in digital cellular implementation implies that analog FM systems are still of concern. Many problems and the corresponding solutions may not be restricted to the digital system but may actually be applicable to both systems.

Mobile radio systems are subject to a multipath fading environment which results in somewhat less than ideal reception. A significant problem is that there exists a number of propagation paths which distort the signal being transmitted. This set of transmission paths followed is referred to as a channel. Fading is a type of destructive interference which results in instants of poor signal to noise ratio (SNR) and rapid phase changes resulting in random FM.

The transmission path may be simulated using a channel impulse response model which represents a multipath fading environment. The linear time varying filter model incorporates the instantaneous parameter values of the fading variations. These variations are statistical descriptions of observed data (Stein 1987).

The multipath environment results from the reflection and scattering of radio waves from environmental features. The effect of the multipath propagation may result in deep signal nulls or fades which then distort the transmitted signal.

The propagation model suggested by Turin (Turin 1972) and developed by Hashemi (Hashemi 1979) is a model which describes an urban environment. It is characterized by a Rayleigh distributed envelope and exhibits flat fading (little time dispersion) relative to the 30 KHz bandwidth channel. Under this flat fading propagation condition, the signal intensity is uniform across the channel bandwidth (Rummler 1986).

Random phase changes are associated with random amplitude fading. As the receiver moves an additional noise is introduced which is referred to as random FM. This has a baseband spectrum which extends from zero to twice the Doppler shift (Jakes 1971). The random FM has a student-t distribution (Rice 1948).

One solution to the problem of the multipath propagation effects is the use of antenna diversity (Rummler 1986). Diversity refers to the situation in which information is obtained from signals transmitted over separate channels. The use of diversity is suggested since there is a lower probability that any individual signal will be struck by a deep fade coincidentally from every available source (Stein 1987). In the case of spatial diversity, redundant signals are received from spatially separated antennas. The signals received from theses antennas should have independently fading channel characteristics (Jakes 1971).

The investigation of the effectiveness of a proposed technique may be carried out by actual field trials and by simulation of the system. Simulation of a system allows system performance to be investigated in a more cost effective manner than by field trial. The simulation of the system would contain the components of the actual system and would include the compandor, pre-emphasis and de-emphasis filters, the limiter, the modulator and demodulator as well as the inclusion of a propagation model.

The multipath channel model from a modified version of Hashemi's model is presented and described in chapter 2. Various methods of spatial diversity combining schemes which have evolved over the years are presented in chapter 3. The description of the simulated system components are presented in chapter 4. These include the frequency response and characteristics of the various components. The total system response is presented in chapter 5. The performance of various diversity combining methods is investigated for sinusoidal and speech inputs. Performance parameters are also presented. The results and discussions are summarized in chapter 6.

Chapter Two

Channel Model

2.1 Introduction

In a communications system, as a signal is transmitted from its source to the destination of the receiver, the signal is often subject to a varying degree of distortion. This distortion is dependent upon the transmission medium. In a radio environment, the transmission paths followed are referred to as a channel and may be represented by a propagation model. The model should represent actual measurements of the channel. The use of a channel model is desirable since by using actual measurements the user is limited to the constraints imposed during which the measurements were obtained. Such constraints would be vehicle velocity and the sampling rate.

The channel may also be thought of representing a multiplicity of paths which are all time varying. Each path is associated with an attenuation and a time delay. At baseband, each of these arrivals has a corresponding phase due to its own transmission delay. The complex envelope is the equivalent baseband representation of the baseband channel and therefore completely characterizes the effect of the channel on the baseband signal. This can be represented mathematically as an impulse response as follows (Proakis 1983):

$$\tilde{c}(\tau,t) = \sum_{n} \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)}\, \delta[\tau - \tau_n(t)]$$

where:

$\tilde{c}(\tau,t)$ = baseband channel impulse response,

$\alpha_n(t)$ = amplitude of the n^{th} path,

$\tau_n(t)$ = propagation delay of the n^{th} path,

τ = delay factor,

t = time instant variable and

f_c = carrier frequency.

This forms the basis of the channel model from which channel characteristics and phenomenon can be described.

The impulse response of the wide band channel was obtained from a mathematical model used by Hashemi (Hashemi 1979) and suggested by Turin (Turin 1972). This model represents a multipath fading channel which describes a time varying mobile urban environment. Each profile is characterized by a set of amplitudes with a temporal resolution of 100 nsec for a duration of 7μsec and a spatial spacing of 0.03048 metres (0.1 feet). An example of the impulse response of a profile is illustrated in figure 2.1.

The modified Hashemi model (Fattouche et al 1990) results from the development of phase generation procedures. If a new reflection occurs then a new arrival angle is generated. The phase is obtained by consideration of the Doppler shift caused by the vehicle motion. The Doppler frequency is determined from this arrival angle using the following expression (Lee 1982):

$$f_{Doppler} = f_m \cos\theta$$

where:

$f_m = V/\lambda$ = maximum Doppler frequency (Hertz),

θ = arrival angle (radians),

V = vehicle velocity (metres/sec) and

λ = wavelength of an unmodulated carrier (metres).

Integration of this Doppler frequency gives rise to a differential phase for each

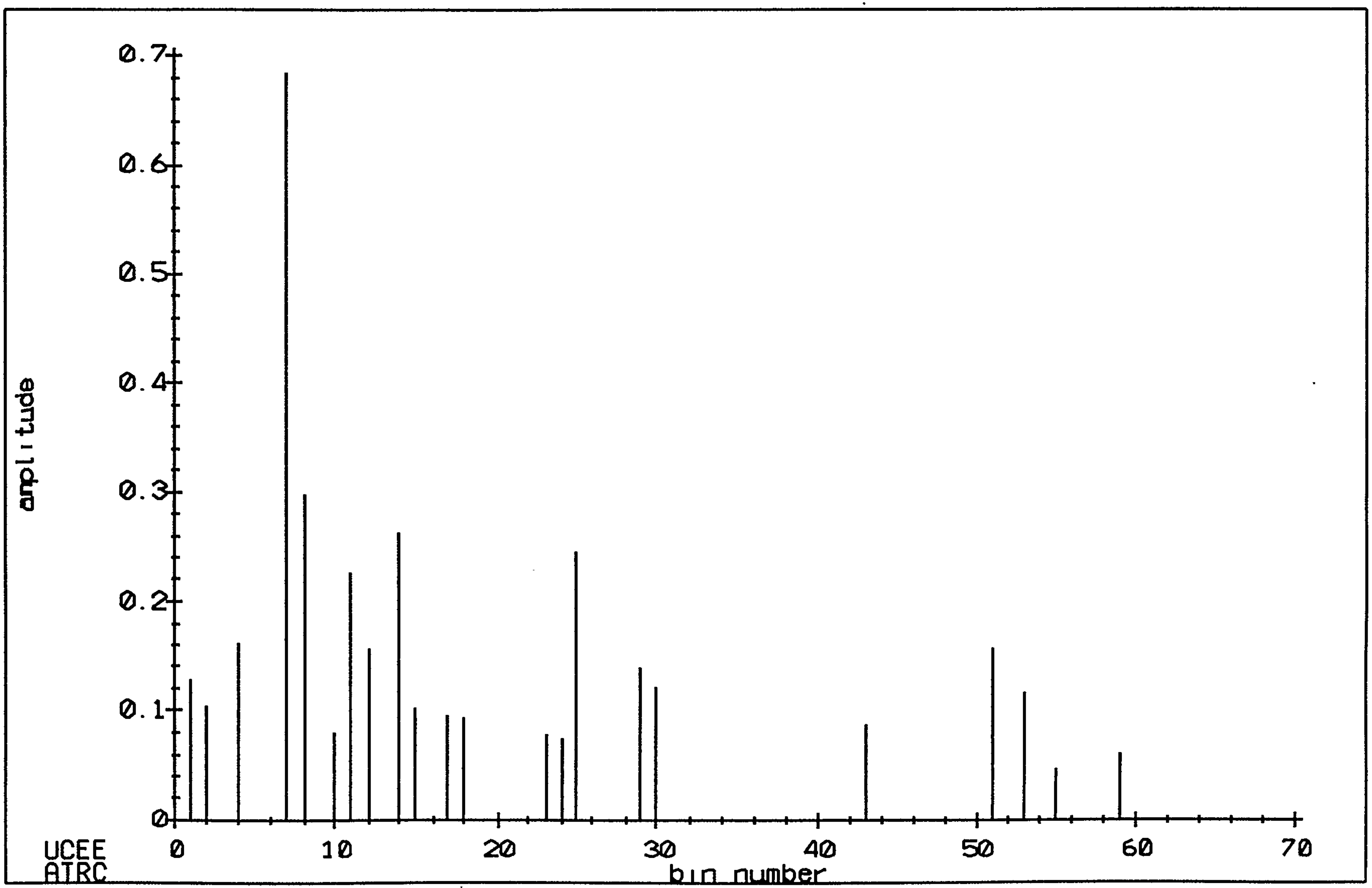


figure 2.1 - Modified Impulse Response Example

reflection between successive profiles. The complex envelope magnitude of the modified Hashemi channel model is illustrated in figure 2.2.

The modified Hashemi channel model describes a wideband channel and it may be modified so that is represents a narrowband channel corresponding to the transmission bandwidth of the signal. The narrowband channel is obtained by lowpass filtering and decimation of the wideband channel. The decimation factor is found by taking the ratio of the bandwidth of the wideband channel to the desired narrowband channel. The spatial sampling of the wideband channel was changed by interpolating between successive profiles of the channel model. The amount of interpolation is dependent upon the vehicle velocity and the sampling rate.

2.2 Channel Model Descriptions

The multipath channel exhibits certain properties which may characterize an environment. Certain aspects of these will affect the performance of a communications system (Rummler et al 1987).

As a signal is transmitted over a channel, it experiences multipath scattering which causes a time spread or delay spread to be introduced to the signal. Since the channel is time varying, this delay spread will also change. This type of channel is also referred to as being time dispersive.

Another characteristic of the channel is that of signal fading. This arises from the attenuation and delay associated with each propagation path. The phase of each path will vary with time and will cause the signal to add constructively or destructively. The received signal amplitude will thus vary according to the time varying nature of the multipath channel (Stein 1987).

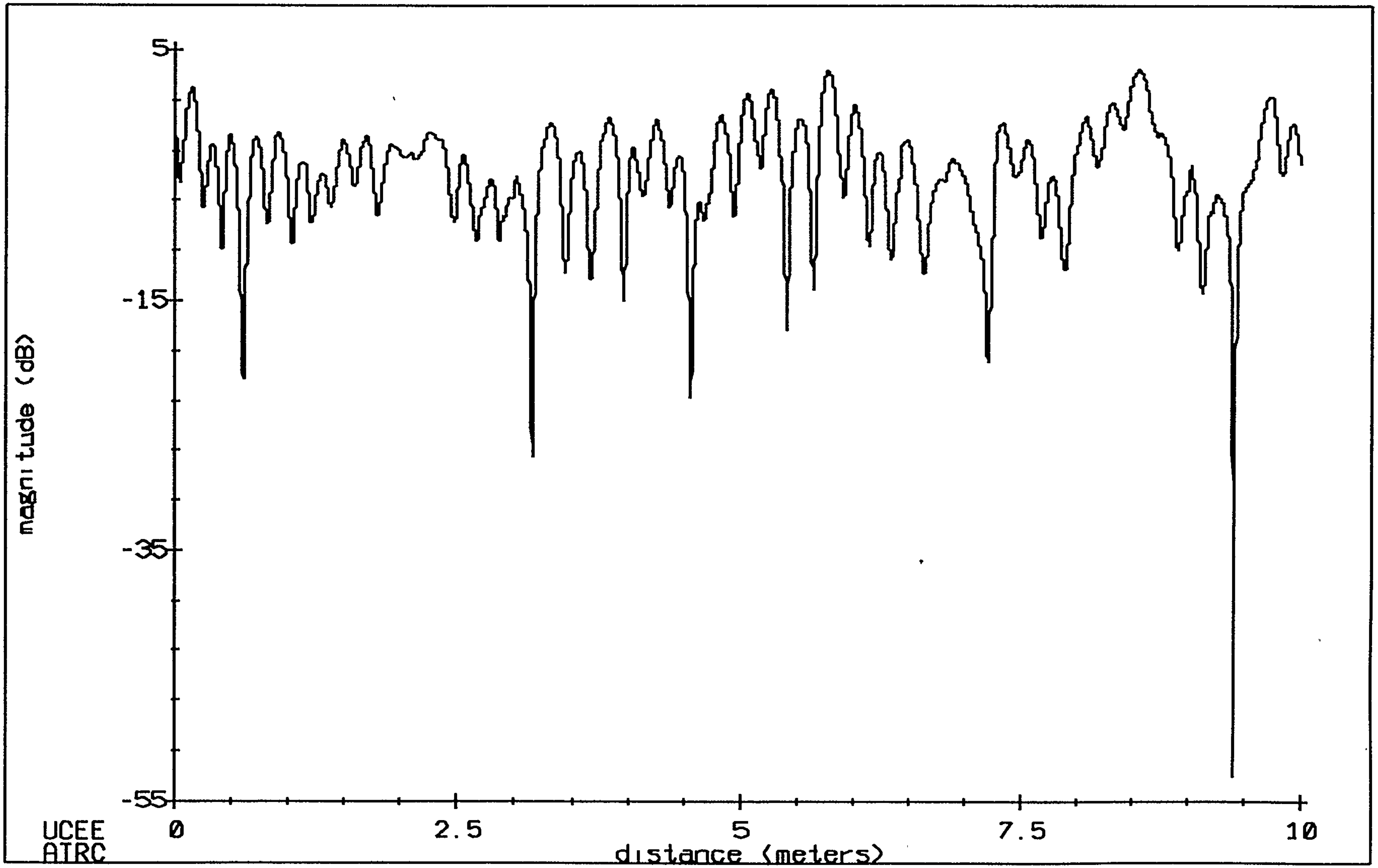


figure 2.2 - Complex Envelope Magnitude

Other terms such as long term and short term fading may be used to describe the variations in the average signal strength. Long term fading may be caused by small scale topographical changes. Short term fading would be influenced by the reflectivity of the signal scatterers and the relative motion of receivers and transmitters (Lee 1982).

The channel can also be described in terms of the type and amount of distortion that it causes. When the correlation between components of the same signal which has small frequency separations is not affected by the channel, a flat fading condition results. This frequency interval determines the correlation or coherence bandwidth and is dependent upon the time delay or dispersion of the channel (Stein 1987).

Flat fading arises when the bandwidth of the signal is less than the correlation bandwidth of the channel. A signal would not be severely distorted by the channel as there is little time dispersion. As the time dispersion increases, the correlation bandwidth of the channel decreases relative to the bandwidth of the signal. This distortion leads to frequency selective fading (Proakis 1983).

2.3 Channel Model Statistical Characterizations

In a flat fading environment, the channel transfer function is constant over the bandwidth of the signal. The complex envelope may be simplified to the value of the transfer function at the dc component at baseband corresponding to the carrier frequency. Equivalently for a continuous wave the complex envelope takes the following form:

$$\tilde{c}(\tau) = \sum_{n} \alpha_n(t)\, e^{-j2\pi f_c \tau_n(t)} = r(t)\, e^{-j\theta(t)}$$

where:

$\alpha_n(t)$ = amplitude of the n^{th} path,

$\tau_n(t)$ = propagation delay of the n^{th} path,

f_c = carrier frequency,

t = time instant variable,

$r(t)$ = envelope magnitude and

$\theta(t)$ = envelope phase.

This complex envelope model is composed of an inphase and a quadrature component which are both Gaussian distributed. This is implied since the envelope represents a sum of reflections and from the central limit theorem as the number of reflections increases this sum will become a Gaussian distributed variable. This complex zero mean random variable has a magnitude which is Rayleigh distributed (Stein 1987). The probability density function of a Rayleigh distribution is given as follows (Lee 1982):

$$\rho(r) = \frac{r}{\sigma^2} \exp\left[\frac{-r^2}{2\sigma^2}\right] \quad \text{for } r \geq 0.$$

A histogram (McClave et al 1988) was constructed from the channel model samples and is plotted in figure 2.3. The Rayleigh pdf (probability density function) is also plotted. A close fit between the two functions is observed indicating that the model provides a Rayleigh distributed envelope.

In angle modulated systems, system performance is affected by phase changes. These phase changes caused by relative motion occur during transmission and are termed random frequency modulation (random FM). This random variation in the apparent frequency is due to the time varying inphase and quadrature components of the channel (Jakes 1974).

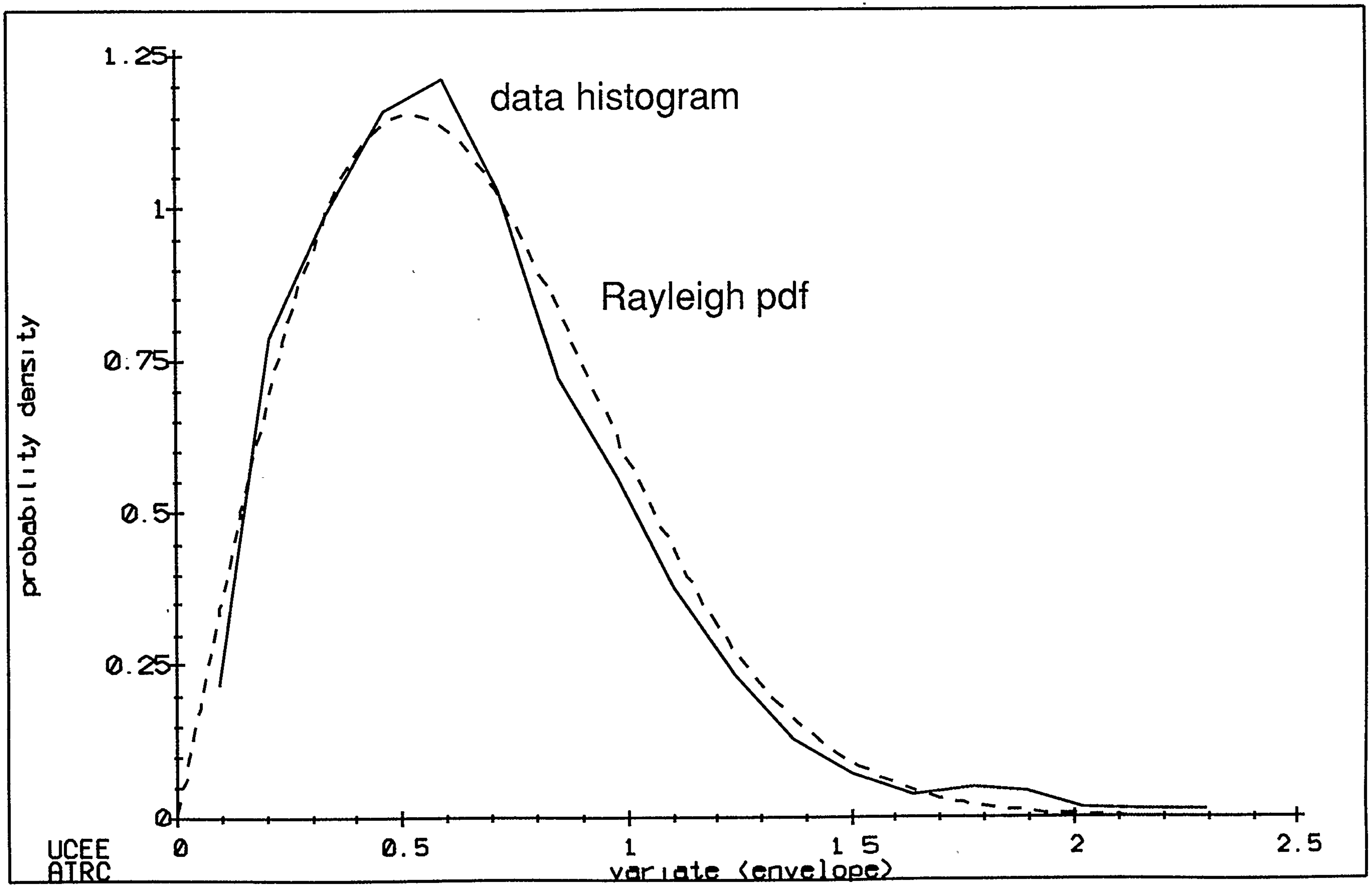


figure 2.3 - Envelope Magnitude Density Distribution

The phase variation due to random FM is determined from the phase derivative. The phase derivative with respect to time is represented by the instantaneous frequency and is related to the phase derivative with respect to distance by the velocity. The instantaneous frequency is also referred to as the instantaneous Doppler shift and is defined as follows (Eggers et al 1989):

$$w_d = \dot{\theta} = \frac{d\theta}{dt} = V \frac{d\theta}{dx}$$

where:

w_d = instantaneous frequency (radians/sec),

θ = phase (radians),

$\dot{\theta}$ = phase derivative (radians/sec),

t = time variable (sec),

V = velocity (metres/sec) and

x = spatial distance (metres).

The distribution of the instantaneous frequency has been found to follow a student-t distribution (Rice 1948). The probability density function of the phase derivative is as follows (Jakes 1974):

$$p(\dot{\theta}) = \frac{1}{w_m \sqrt{2}} \left[1 + 2\left(\frac{\dot{\theta}}{w_m} \right)^2 \right]^{-3/2} .$$

This can be shown to be a student-t distribution given by (Gibra 1974) with a variable change as follows:

$$p(t) = p(\dot{\theta}) \frac{d\dot{\theta}}{dt}$$

where $t = \dfrac{2\dot{\theta}}{w_m}$ = student-t variable,

$$\frac{d\dot{\theta}}{dt} = \frac{w_m}{2} \text{ and}$$

$$w_m = \frac{V}{\lambda} = \text{maximum Doppler frequency (radians/sec),}$$

therefore:

$$p(t) = \frac{1}{2\sqrt{2}} \left[1 + \frac{t^2}{2} \right]^{-3/2} .$$

A histogram of the phase derivative normalized by the factor $w_m/2$ is plotted in figure 2.4. The student-t probability density function is also plotted for comparison. A close similarity is observed indicating that the phase derivative of the channel is student-t distributed.

2.4 Channel Model with Diversity

The channel model is also used to create spatially separated channels. Each channel is representative of the impulse response of a particular location. In the case of spatial diversity, each antenna occupies a different location relative to the other and as the mobile receiver moves, the channel impulse responses will differ in the spatial interference pattern (different propagation paths encountered).

The spatial separation of the two channels is usually expressed in terms of a portion of the transmission wavelength. The wavelength is defined as follows:

$$\lambda = \frac{c}{f} = \frac{3 \times 10^8}{8 \times 10^8} = 0.375 \; metres/cycle.$$

With the application of spatial diversity, the intent is to have independent channels. The correlation coefficient is a measure of the independence of the channels. It is defined as follows (Lee 1982):

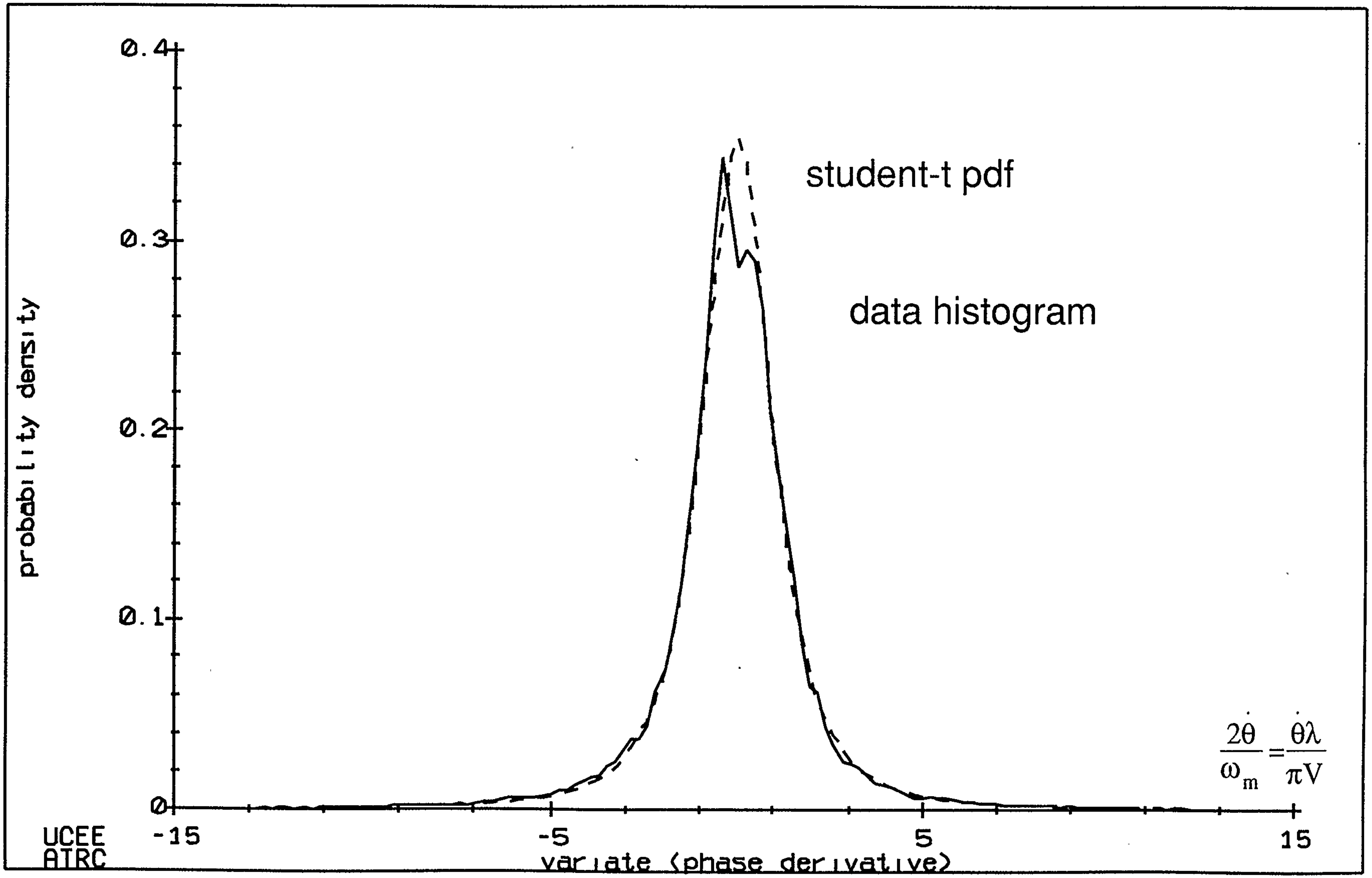


figure 2.4 - Phase Derivative Density Distribution

$$| \rho(d) |^2 = \left| \frac{E[r_1 r_2^*]}{\sigma_1 \sigma_2} \right|^2$$

where:

$r_1 = x_1 + jy_1$ = envelope of the first channel,

$r_2 = x_2 + jy_2$ = envelope of the second channel,

σ_i = standard deviation of r_i for $i = 1,2$,

d = spatial separation (metres),

x_i = the inphase component of the i^{th} channel and

y_i = the quadrature component of the i^{th} channel.

This expression can also be written as follows:

$$\rho(d) = \frac{E[x_1 x_2]^2 + E[x_1 y_2]^2}{E[x_1^2]^2}$$

assuming $E[x_1^2] = E[x_2^2] = E[y_1^2] = E[y_2^2]$.

If the angles of arrival are uniformly distributed, the correlation coefficient reduces to:

$$| \rho(d) |^2 = J_o^2(\beta d)$$

where $\beta = \frac{2\pi}{\lambda}$ = wave number (radians/metre) and

$J_o(\beta d)$ = Bessel function of order zero.

This form obtained from Lee (Lee 1982), gives a theoretical correlation coefficient. Since the envelope of the channels can be determined from the model, the correlation coefficient can be determined for the modified Hashemi model used.

The comparison of the magnitude of the theoretical and the actual correlation coefficients squared are plotted in figure 2.5. This illustrates that the channels are the most uncorrelated at a spatial spacing of 0.38 λ in the case of the modified Hashemi model.

The frequency domain counterpart of the spatial correlation is the spatial spectral density. The spectral characteristic reveals the distribution of the signal as a function of frequency. In the modified Hashemi model a Doppler frequency is associated with each reflection. This results in a spreading in the frequency domain. The spatial frequency domain representation of the spectrum is of interest since the spatial distribution is independent of the velocity. The frequency may be expressed as a function of the velocity as follows:

$$f = V \cdot \nu$$

where:

f = temporal frequency (Hertz),

V = velocity (metres/sec) and

ν = spatial frequency (cycles/metre).

The spectral density function is obtained from the Fourier transform of the autocorrelation function (Proakis 1983) and is plotted in figure 2.6.

The probability density function (pdf) of the spatial frequency describes the spatial frequency distribution of the channel as well. The pdf is determined from the following expression:

$$p_\nu(\nu) = p_\theta(\theta) \left| \frac{\partial\theta}{\partial\nu} \right|$$

where:

$p_\theta(\theta) = \dfrac{1}{2\pi}$ for uniformly distributed arrival angles and

$\left| \dfrac{\partial\theta}{\partial\nu} \right|$ = the Jacobian resulting from the transformation of variables.

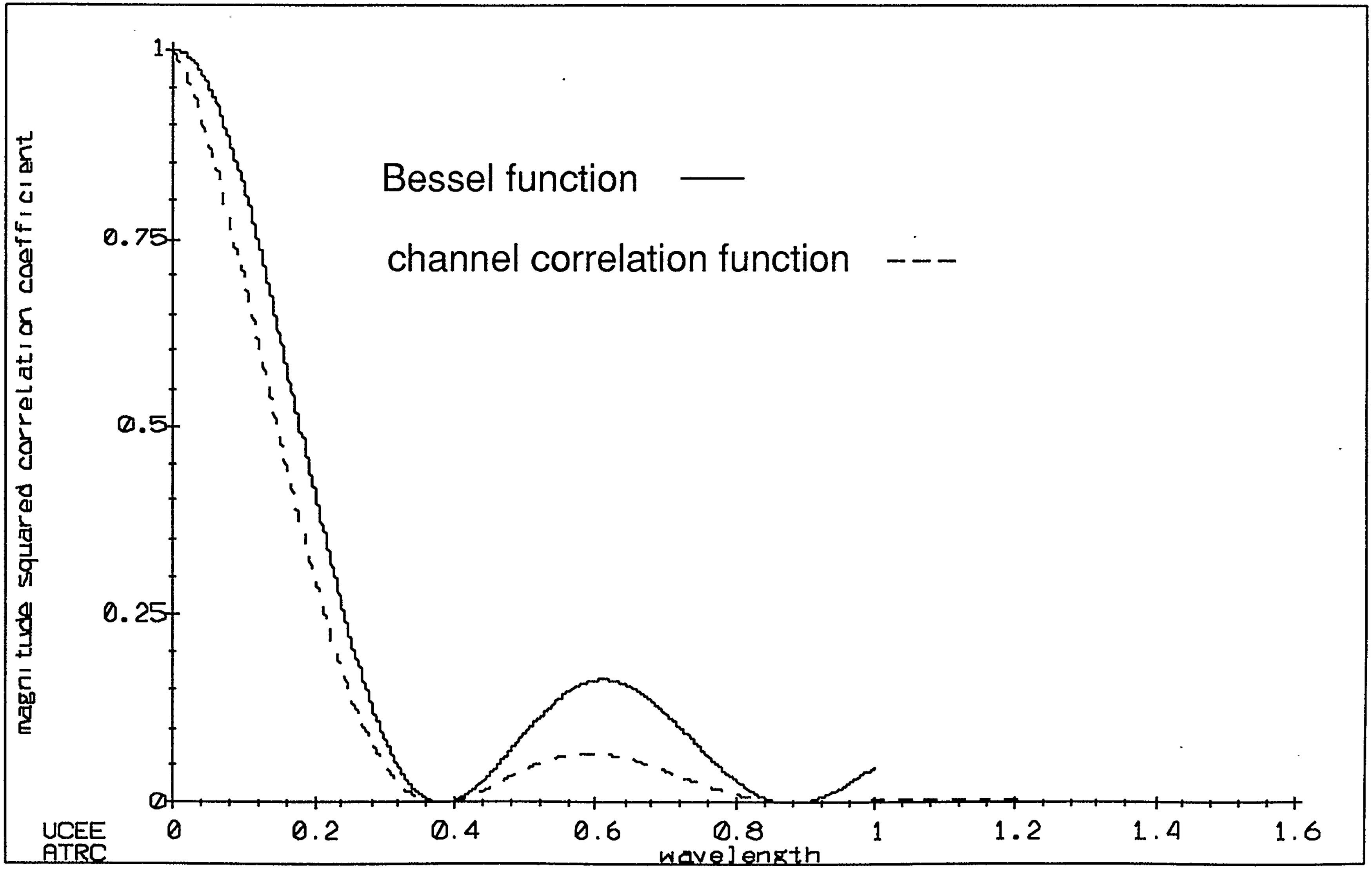


figure 2.5 - Channel Correlation Functions

The Jacobian of the variable transformation is determined from the expression for the Doppler frequency as follows:

$$f = \frac{V}{\lambda} \cos \theta$$

where:

f = Doppler frequency = $V\ \nu$ (Hertz),

V = velocity (metres/sec),

λ = wavelength (metres) and

θ = angle of arrival (radians).

Then changing to spatial frequencies ν results in:

$$\nu = \frac{\cos\theta}{\lambda}$$

then

$$\frac{\partial\theta}{\partial\nu} = \frac{-\lambda}{\sin\theta} = \frac{-\lambda}{\sqrt{1-\cos^2\theta}} = \frac{-\lambda}{\sqrt{1-(\nu\ \lambda)^2}}.$$

This is used to yield the expression for the spatial frequency pdf as follows:

$$p_\nu(\nu) = p_\theta(\theta)\left|\frac{\partial\theta}{\partial\nu}\right| = \frac{\lambda}{2\ \pi}\ \frac{1}{\sqrt{1-(\nu\ \lambda)^2}}.$$

This function is defined only for spatial frequencies with magnitudes less than $1/\lambda$.

The plot of the spectral density function for the complex envelope of the channel in figure 2.6 illustrates that the frequency content of the signal is limited to the spatial frequency band of ± 2.67 cycles/metre. This corresponds to the maximum Doppler frequency given by:

$$f = V\ \nu = \frac{V}{\lambda} = 2.67\text{cycles/metre}$$

where:

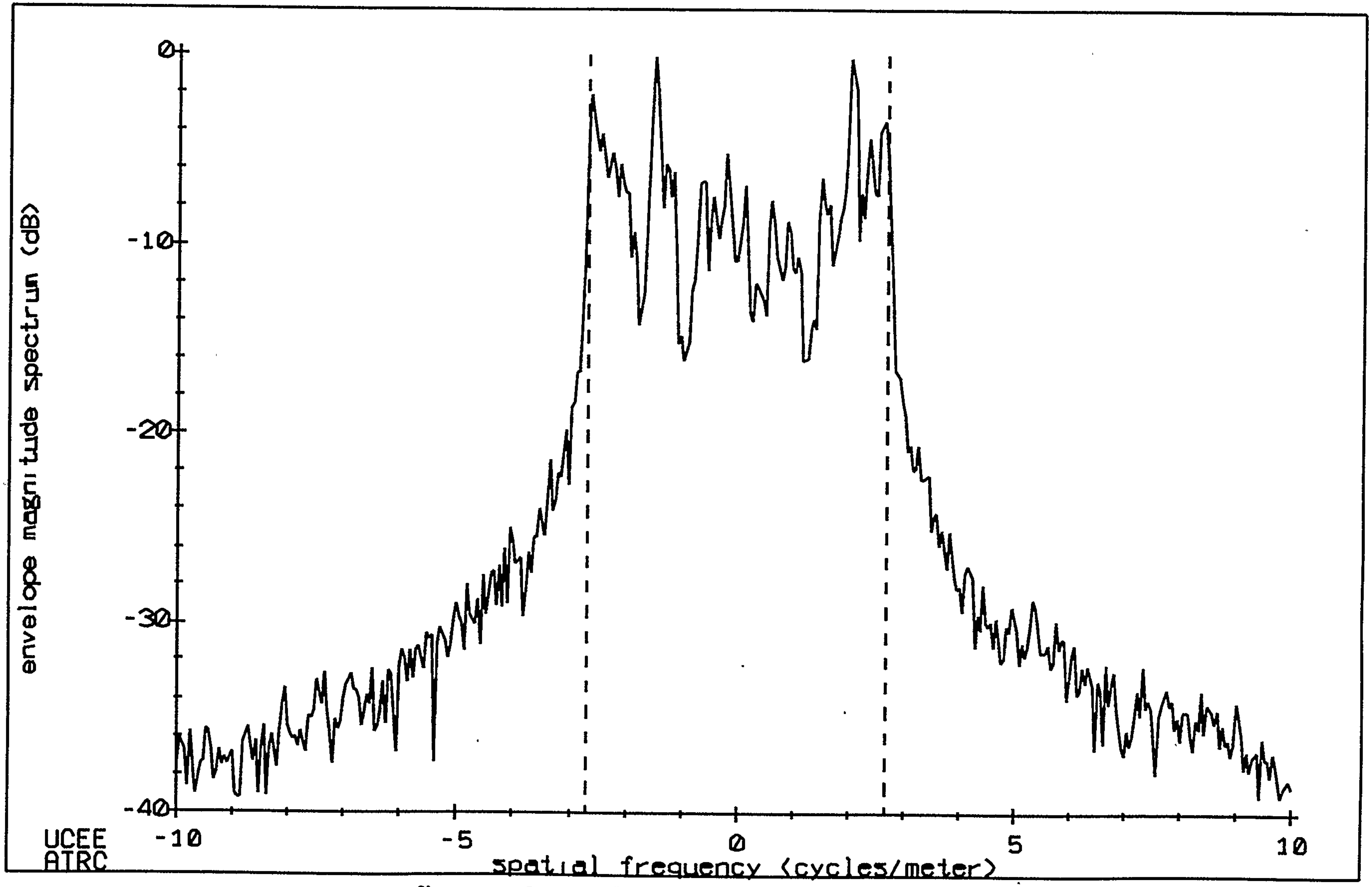


figure 2.6 - Spatial Frequency Spectrum

V = 50 km/hr and

λ = 0.375 metres/cycle.

2.5 Conclusions

The modified Hashemi model appears to display the statistical characterizations typical of a mobile radio environment at 800 MHz. The envelope displays a characteristic Rayleigh distributed envelope. In addition the phase derivative displays a student-t distribution.

With the application of spatial diversity, independent channels are desirable. From plots of spatial correlation functions, ideal spatial separations were determined to achieve minimum correlation values. These spatial separations were equivalent to those obtained from Bessel functions derived from ideal Rayleigh distributed envelopes.

The spectral density of the complex envelope of the channel illustrated that the frequency content of the channel is limited to the maximum Doppler frequency obtained by vehicular movement.

These statistics and characterizations give confidence in the validity of the modified Hashemi model and thus give confidence in its use to simulate an urban mobile radio environment.

Chapter Three

Diversity Combining Methods

3.1 Introduction

Diversity is a term used to describe a situation in which a signal is transmitted over a set of independently fading channels. The diversity may be achieved by a number of methods such as spatial separation of antennas, field component diversity from different antenna types, polarization diversity achieved by using horizontal or vertical electric fields, angle diversity using directional antennas, frequency diversity arising from the transmission of the signal over different carrier frequencies and time diversity from successive transmission of data (Lee 1982). The diversity scheme employed is intended to minimize the effects of fading and therefore to increase the reception quality of the signal.

In spatial diversity two antenna are spaced a distance apart whereby the received signals will fade independently. This implies that the channels have to be uncorrelated at the chosen antenna separation. The correlation coefficient may be used as a measure of the correlation of the two channels. Performance improvements are expected as the channel correlation decreases (Jakes 1971).

3.2 Combining Techniques

Through the use of diversity two or more signals are available from which the best signal must be selected or combined. The method or technique by which this is done is referred to as diversity combining. Each technique will have advantages and disadvantages. Diversity systems generally employ only two signals for practical reasons and since performance gains are marginal for additional signals (Jakes 1971).

3.2.1 Switched Selection

In the switched selection method of combining, a switch and stay strategy is used. This implies that only when the signal level from an antenna drops below a predetermined threshold does a switch to the stronger of the two signals occur. This combining method is perhaps the least complicated type of combining. The method selects the signal with the highest power level and discards the other signals.

This method reduces the number of switches occurring compared to a selective switching method where continuous switching occurs. Switching causes phase discontinuities which may affect system performance. The switching threshold should therefore be set so that switching does not occur very often.

3.2.2 Audio-level Combining

In order to reduce the phase discontinuities arising from switched selection, selection may be done at baseband at the system output. In audio level combining two receivers are used. Selection of the strongest signal is done at IF but both signals are passed through the receiver and the actual signal selection is implemented at the audio output at baseband. An advantage of this method would be that antenna switching does not occur before the demodulator implying less phase discontinuities or random FM.

3.2.3 Equal Ratio Combining

The equal ratio combining method arises from the idea that since two signals are available, both should be used. In this method, individual signals are cophased before being added together. This results in a coherent summing of the signals and an incoherent summing of the noise elements. An additional advantage is the reduction of random FM.

To prepare signals for coherent addition, one signal may be referenced to another or the phase introduced by the channel may be eliminated. In an actual system the phase of the channel is not known. Various techniques attempt to predict, estimate or eliminate the phase difference introduced from the channels of the two antennas. These are often complicated and the effectiveness of the technique may be compared to the limit expected when phase information is known. Since the phase of the channel in the simulation is available, the cophasing is done by subtracting the channel induced phase from each signal. This is done by a complex multiplication of the signal with the conjugate of the sum of the channel impulse response which has been normalized. This eliminates the phase introduced by the channel. Taking the complex sum of the channel impulse response is equivalent to performing a Fourier transform and finding the phase of the dc component.

3.3 Conclusions

In this chapter various types of diversity combining methods are discussed. In using a diversity scheme it is thought that performance increases will be observed when diversity signals are available which have encountered different propagation paths. The different signals would perhaps not be experiencing the same channel fades, therefore one signal or a combination of the signals will result in a better received signal and increased system performance.

Performance degradations of the various methods are often due to phase differences between received signals and also phase distortions within each received signal bandwidth. Diversity combining methods not employing cophasing techniques are affected to a greater degree by phase changes between channels. The effectiveness of these techniques will be influenced by the degree to which the channel is flat fading.

Chapter Four

FM System Simulation

4.1 Introduction

The performance of various spatial diversity combining techniques in analog FM may be investigated by simulating the communication system. Simulation is desirable since changes and modifications are more easily made before more costly actual field trials are done. It is therefore important that the simulation contain similar components to the actual system (EIA/TIA 1989) so that the simulation respond in a similar manner to the actual system.

An illustration of the simulated analog FM system component configuration is presented in figure 4.1. The system consists of the following components:

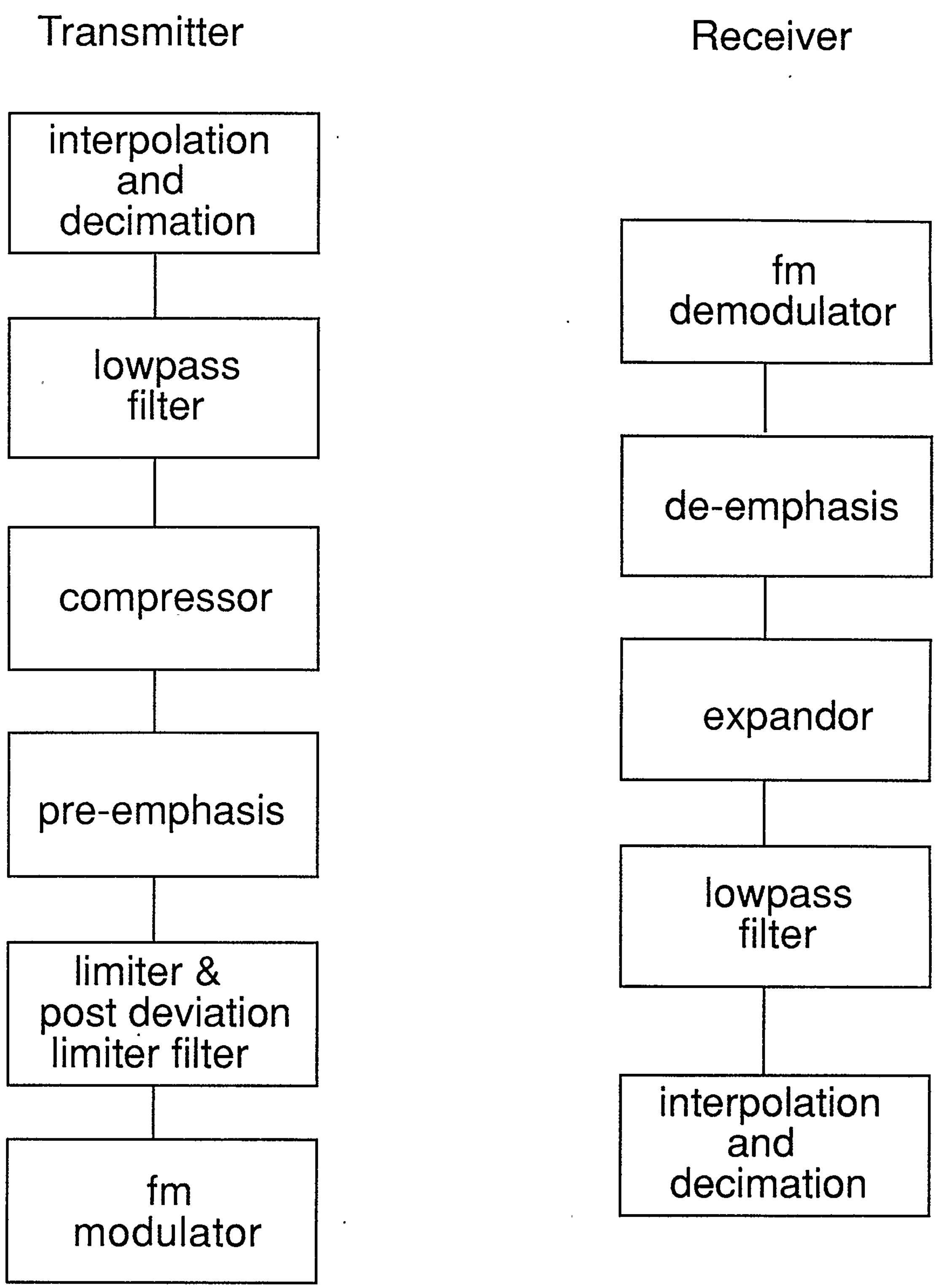
- interpolation and decimation filters,
- a compandor (compressor and expandor),
- pre-emphasis and de-emphasis,
- a limiter and a post deviation limiter filter,
- an FM modulator and demodulator/discriminator and
- lowpass filters.

In addition to the above, the channel model by Hashemi was used as discussed previously.

4.2 Interpolation and Decimation

The simulation uses as input, digitized speech data obtained from an analog to digital (A/D) converter operating at 48 kHz. Cellular FM occupies a bandwidth of 30 kHz.

Figure 4.1 FM SYSTEM COMPONENTS



An interpolation filter is therefore required so that the signal may be resampled at a rate compatible with the allocated bandwidth of the FM system. The signal is therefore interpolated by a factor of two yielding a sampling frequency of 96 kHz, then lowpass filtered and decimated by a factor of three to yield a sampled frequency of 32 kHz.

The output of the simulation also requires the use of interpolation and decimation to reconstruct the speech data to have a sampling frequency of 48 kHz in order to match the requirements of the digital to analog (D/A) converter.

Sampling theory states that a finite energy function having a band-limited Fourier transform can be reconstructed from its sampled values provided that the sampling frequency is twice that of the highest frequency component of the original signal. This reconstruction may be done by convolving the data with a sinc function (Brigham 1977).

Whenever a function is convolved with a truncated sinc pulse, oscillations occur whenever a frequency domain discontinuity is approximated. These oscillations are known as the Gibbs phenomenon and may be reduced by applying a window to the pulse. The Hanning window was used since it is known to have good sidelobe attenuation. It is given by (Press et al 1988):

$$w(n) = \frac{1}{2}\left[1 - \cos\frac{2\pi n}{N-1}\right] \qquad 0 \le n \le N-1$$

where N = filter length.

The interpolation filter is therefore given by the multiplication of the sinc pulse with the Hanning window as follows:

$$p(n) = \frac{1}{2}\left[1 - \cos\frac{2\pi n}{N-1}\right] sinc\left[n\frac{T_1}{T_2}\right] \qquad 0 \le n \le N-1$$

where:

T_1 = original sample period and

T_2 = desired sample period.

The frequency response of the resulting interpolation filter is illustrated in figure 4.2.

4.3 Compandor

A compandor consists of a compressor and an expandor. The compressor is used to compress the intensity range of the speech signals at the input to a communications channel. This is done by applying a larger gain to weak signals and a smaller gain to large signals. The expandor at the receiver or output of the communications channel performs the function of restoring the intensity range back to its original range. The purpose of this device is to keep the intensity of the speech signals above the noise floor (Lenkurt 1964).

The compandor contains a device which determines the gain to the incoming signal. This gain is determined and controlled by the following three characteristics:

1) the actual companding ratio 2:1,

2) the companding intensity range and focal point and

3) the time constants.

The compressor-expansion ratio is an expression of the degree to which the speech energy is compressed and expanded. It is determined by the ratio of the input to output power in units of decibels. A ratio of two was used for the compressor while a ratio of one half was used for the expandor.

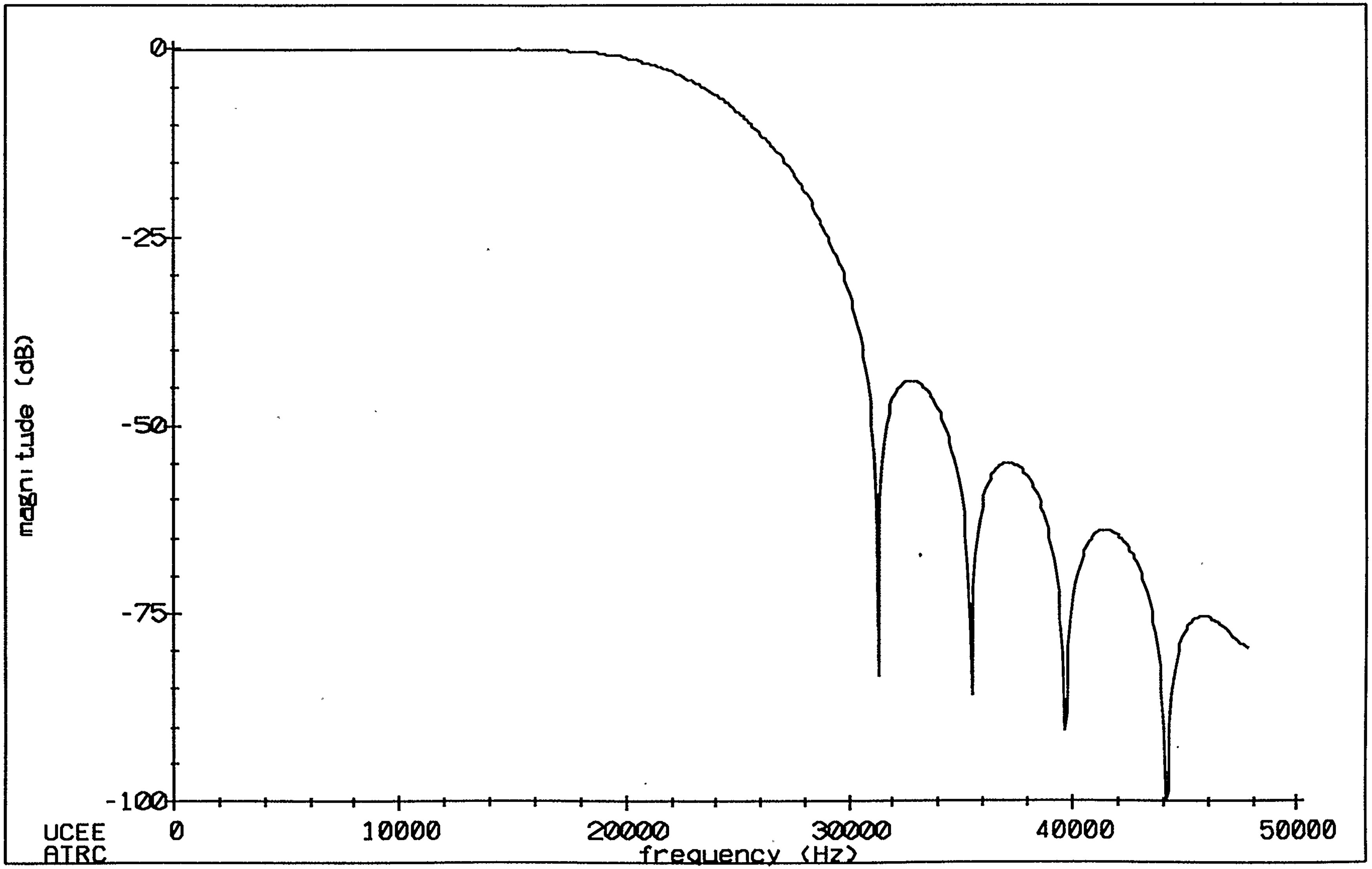


figure 4.2 - Interpolation Filter Frequency Response

The intensity range is selected so that it covers a wide energy range of speech signals. Signals occurring outside this range are limited and therefore this range should be selected such that distortion is minimized. Since the gain is adjusted according to the input signal power it follows that the range will then determine the minimum and maximum gain. The focal point refers to the energy level which results in no loss or gain (unaffected point).

From the range, focal point and ratio of the compandor, the maximum and minimum gain values can be determined. The range used varied from 3 dBmO to -55 dBmO. The logarithmic unit of dBmO applies to a power ratio where this ratio is the power dissipated in a 600 ohm resistor at a given voltage level to the power level of 1 mWatt (Lenkurt 1966). The maximum and minimum gains in dB are found from the following expression:

$$gain_{\max} = \frac{1 - ratio}{ratio} * (min\ power - focal\ point)$$

and

$$gain_{\min} = \frac{1 - ratio}{ratio} * (max\ power - focal\ point).$$

The application of the gain can be best determined by examining the circuitry of the actual system. The difference between the past value of the gain and the final steady state gain value resulting from the input signal will determine the time constant to be used. The updated gain value may be determined using the following expression:

$$g(n) = g(n-1) + \alpha\,(g_{ss} - g(n-1))$$

where:

g_{ss} = final steady state gain value corresponding to the input,

$g(n-1)$ = previous gain value,

$g(n)$ = current gain value and

$\alpha = 1 - \lambda =$ a time constant factor for memory.

This expression can be rewritten as follows:

$$g(n) = g(n-1) + (1 - \lambda)\,(g_{ss} - g(n-1))$$
$$= g(n-1) - g_{ss} - g(n-1) - \lambda\, g_{ss} + \lambda\, g(n-1),$$
$$g(n) = \lambda\, g(n-1) + (1 - \lambda)\, g_{ss}.$$

For a constant input the steady state gain will not change and enables the current gain to be expressed in terms of the initial gain value as follows:

$$g(n) = \lambda^n\, g(0) + (1 - \lambda^n)\, g_{ss}$$

or equivalently:

$$g(n) = \lambda^n\,(g(0) - g_{ss}) + g_{ss}.$$

By solving for λ^n, the time for the current gain to reach the steady state gain may be expressed as an exponential weighting factor with a time constant as follows:

$$\lambda^n = \frac{g(n) - g_{ss}}{g(0) - g_{ss}} = \exp(-n\,\Delta t/\,\tau),$$
$$\Longrightarrow \lambda = \exp(-\Delta t/\,\tau).$$

Solving for τ:

$$\tau = \frac{n\,\Delta t}{\ln\left\{\dfrac{g(0) - g_{ss}}{g(0) - g_{ss}}\right\}}$$

where:

Δt is the sample period and

τ is a time constant.

Time constants are also required in the implementation since the gain of the compandor cannot change instantaneously with a change in the input signal. If an instantaneous change occurred then the output signal would be severely distorted. The gain of the compandor varies as a function of the signal envelope rather than with it's instantaneous value.

The two time constants defined are the attack time constant and the recovery time constant. These time constants are found by solving the previous equation at the appropriate gain values to be defined:

The attack time is the interval of time in which a signal decays to 1.5 times its steady state value after an input increase. This interval or attack time is specified to be 3 msec. For an initial gain, from -16 dBmO to -4 dBmO a change of 12 dB occurs resulting in an initial gain value of 12 dB since the gain cannot change instantaneously. For the final gain, with a ratio of 2:1 a change of 6 dB occurs resulting in a final steady state gain value of 6 dB.

Solving for τ_{attack} at:

$g_{ss} = 6$ dB,

$g(n) = (1.5)(\text{steady state gain value}) = 9$ dB and

$n\Delta t = 0.003$ (3 msec)

results in: $\tau_{attack} = 0.004328$ (sec).

The recovery time is defined as the interval in which the input signal power is decreased and the signal recovers to 0.75 times its steady state value. This interval is specified to be 13.5 msec. For an initial gain, from -4 dBmO to -16 dBmO, a change of -12 dB occurs resulting in an initial gain of -12 dB since the gain cannot change instantaneously. For the final gain value, a change of -6 dB occurs resulting in a final gain value of -6 dB. The recovery time constant is solved by substitution of the appropriate variables as before.

Solving for $\tau_{recovery}$ given:

g_{ss} = -6 dB,

$g(n)$ = (0.75)(steady state gain) = -9 dB and

$n\Delta t$ = 0.0135 (13.5 msec)

results in $\tau_{recovery}$ = 0.019476 (sec).

The same time constants are used for the expandor as for the compressor. The time constant is selected based on whether the final steady state gain is greater or less than the previous gain value. Thus if the final steady state gain is greater than the previous gain then this implies an increase in the input and thus the attack time constant is used. Otherwise the recovery time constant is used.

The final steady state gain is determined from the focal point, the compandor ratio and the input power level of the signal. The relationship between the compandor input and output may be expressed in dBmO as:

input power - focal point = ratio (output power - focal point)

Since all signals are expressed in non dB formats, this may be written in a non dB format as follows:

$$10\log\frac{P_{in}}{P_{ref}} - 10\log\frac{P_{fp}}{P_{ref}} = ratio\left[10\log\frac{P_{out}}{P_{ref}} - 10\log\frac{P_{fp}}{P_{ref}}\right]$$

with $P_{ref} = 1mW$

and rearranging gives:

$$ratio \ * \ \log\frac{P_{fp}}{P_{ref}} - \log\frac{P_{fp}}{P_{ref}} = ratio \ * \ \log\frac{P_{out}}{P_{ref}} - \log\frac{P_{in}}{P_{ref}}$$

removing the logarithms and dividing both sides by:

$$\left[\frac{P_{in}}{P_{ref}}\right]^{ratio\ -\ 1}$$

results in:

$$\frac{\left[\frac{P_{fp}}{P_{ref}}\right]^{ratio\ -\ 1}}{\left[\frac{P_{in}}{P_{ref}}\right]^{ratio\ -\ 1}} = \frac{\left[\frac{P_{out}}{P_{ref}}\right]^{ratio}}{\left[\frac{P_{in}}{P_{ref}}\right]^{ratio}}$$

simplifying gives:

$$\left[\frac{P_{fp}}{P_{in}}\right]^{ratio\ -\ 1} = \left[\frac{P_{out}}{P_{in}}\right]^{ratio}$$

substituting for the power values in terms of voltages and resistances:

$$P = \frac{V^2}{R},$$

$$\left[\frac{V_{fp}^2}{V_{in}^2}\right]^{ratio\ -\ 1} = \left[\frac{V_{out}^2}{V_{in}^2}\right]^{ratio}$$

since

$$gain = \frac{V_{out}}{V_{in}}$$

and

$$\left[\frac{V_{fp}}{V_{in}}\right]^{2\,(ratio\,-\,1)} = gain^{2\,ratio}$$

solving for the gain:

$$\left[\frac{V_{fp}}{V_{in}}\right]^{\frac{ratio\,-\,1}{ratio}} = gain$$

then for the compressor the ratio = 2:

$$final\ compressor\ gain = \left[\frac{V_{fp}}{V_{in}}\right]^{1/2}$$

and with the expandor ratio = 1/2:

$$final\ expandor\ gain = \left[\frac{V_{fp}}{V_{in}}\right]^{\frac{1/2\,-\,1}{1/2}} = \left[\frac{V_{fp}}{V_{in}}\right]^{-1} = \frac{V_{in}}{V_{fp}}.$$

If these gain values are outside the range of the minimum and the maximum gain values, the steady state gain is limited accordingly.

The focal point voltage is determined from the focal point by solving for V_{fp} from the following equation:

$$focal\ point = 10\log\frac{{V_{fp}}^2}{R * P_{in}}$$

where:

$$P_{in} = 0.001\ \ \mathrm{W}\ \ \text{and}$$

$$R = 600\ ohms.$$

The gain is inversely proportional to the speech level and as such strong signals are attenuated more than weak signals. Thus where the signal level is decreasing the attack time constant is used, otherwise the recovery time constant is used for the increasing signals.

The response of the compressor to input signal level increases and decreases is illustrated in figure 4.3. The compressor response to a 12 dB input increase is an output steady state increase of 6 dB with an attack time delay of 3 msec. The response to an input decrease of 12 dB is a steady state output decrease of 6 dB with a recovery time delay of 13.5 msec. In this figure the effect of the focal point is also illustrated. The intensity of the signal is unaffected at the -3 dBmO level and signals at this level pass through the system unaffected.

The response of the expandor to input signal level increases and decreases is illustrated in figure 4.4. The expandor response to a 6 dB input increase is an output steady state increase of 12 dB with an attack time delay. The response to an input signal decrease of 6 dB is a steady state output decrease of 12 dB with a recovery time delay.

4.4 Pre-emphasis and De-emphasis

The pre-emphasis and de-emphasis filters are used to increase the output SNR by redistributing the power spectral distribution (PSD) of the baseband signal.

Pre-emphasis applied at the transmitter is used to amplify or boost the high frequency components of the message signal. Signal quality is associated with the frequency content of the signal and as higher frequency signal components have less energy, they are more easily affected by noise. The transfer function for the pre-emphasis filter is given by (Lathi 1983):

$$H_p(w) = \kappa \frac{jw + w_1}{jw + w_2}$$

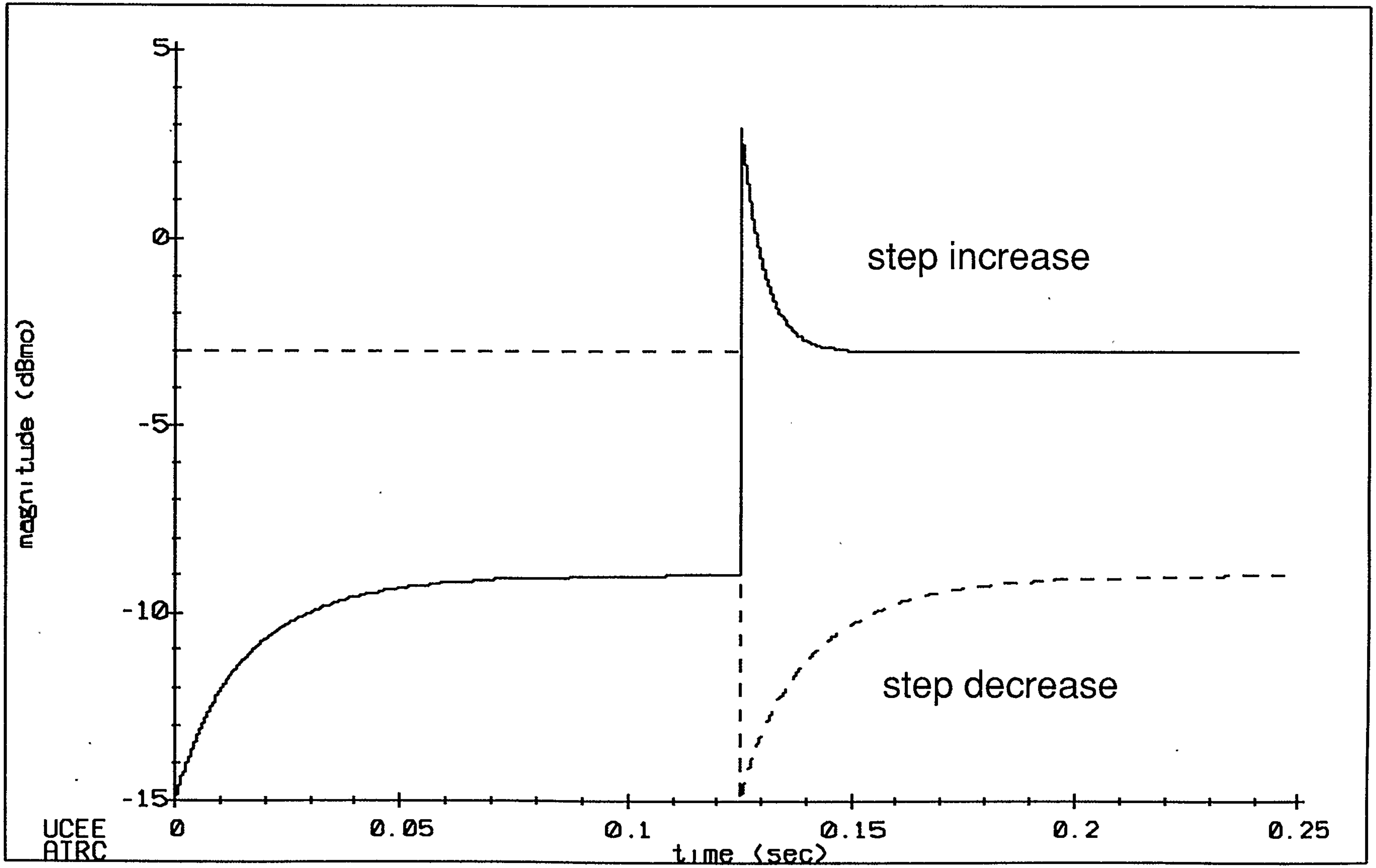


figure 4.3 - Compressor Response

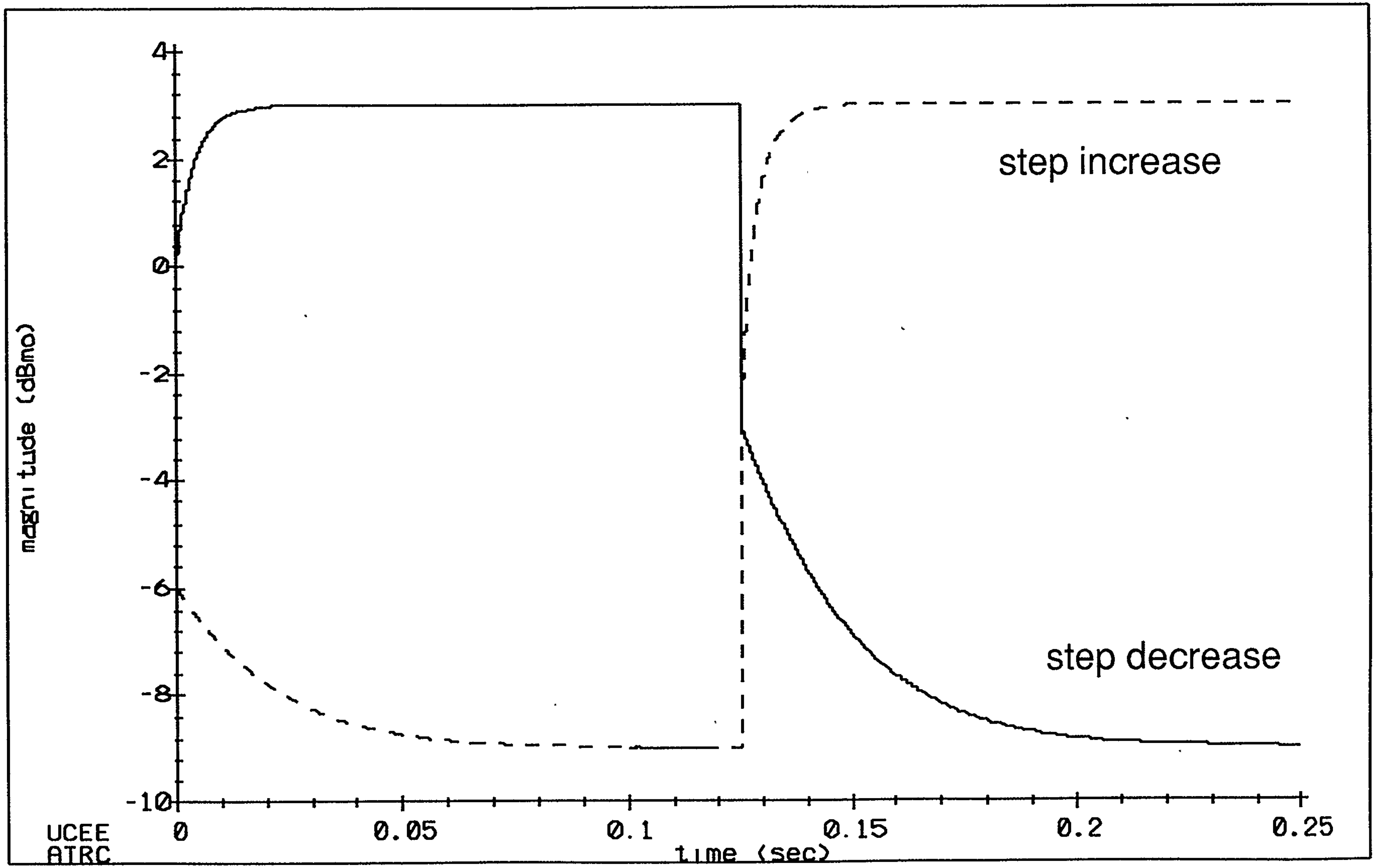


figure 4.4 - Expandor Response

with $\kappa = \dfrac{w_2}{w_1}$.

Substitution with $s = jw$ gives the analog transfer function as follows:

$$H_p(s) = \frac{1 + s/w_1}{1 + s/w_2} = \frac{1 + \tau_1 s}{1 + \tau_2 s}$$

where:

$$\tau_i = \frac{1}{w_i} = \frac{1}{2\pi f_i},$$

$f_1 = 300$ Hz and $f_2 = 3000$ Hz.

The analog filter can be made digital by applying the bilinear transform (Poularikas et al 1991):

$$s = \frac{2}{T} \frac{(1 - z^{-1})}{(1 + z^{-1})}$$

with T = sampling period,

then

$$H_p(z) = \frac{T(1 + z^{-1}) + \tau_1 2(1 - z^{-1})}{T(1 + z^{-1}) + \tau_2 2(1 - z^{-1})} = \frac{Y(z)}{X(z)}$$

and solving for $Y(z)$ gives:

$$Y(z) = \frac{X(z)\,(T + 2\tau^1) + z^{-1} X(z)\,(T - 2\tau_1) - z^{-1} Y(z)\,(T - 2\tau_2)}{(T + 2\tau_2)}.$$

The frequency response of the pre-emphasis filter is illustrated in figure 4.5. The high frequency signal components are amplified from unity gain at the lower frequencies up to a maximum of 20 dB at the higher frequencies.

The inverse operation at the receiver is accomplished using a de-emphasis filter. The

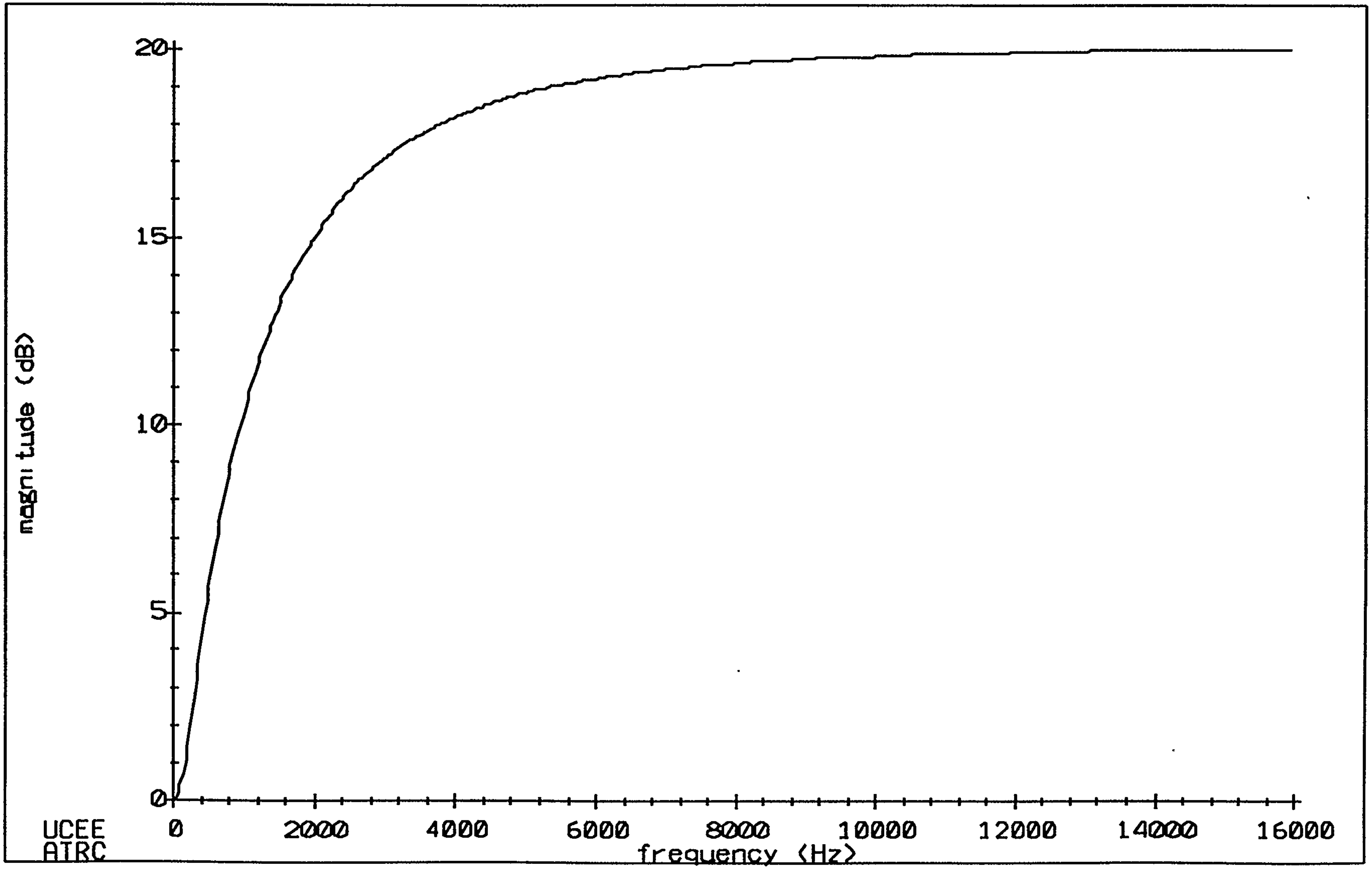


figure 4.5 - Pre-emphasis Filter Frequency Response

high frequency signal components resulting from the channel noise are attenuated by as much as 20 dB by the de-emphasis filter thus increasing the output signal to noise ratio. The transfer function of the de-emphasis filter is the inverse of the pre-emphasis filter and is given as:

$$H_d(s) = \frac{1}{H_p(s)} = \frac{1 + \tau_2 s}{1 + \tau_1 s}$$

$$\text{where} \quad \tau_i = \frac{1}{w_i} = \frac{1}{2\pi f_i}.$$

Applying the bilinear transform as before gives:

$$H_d(z) = \frac{T(1 + z^{-1}) + \tau_2 2(1 - z^{-1})}{T(1 + z^{-1}) + \tau_1 2(1 - z^{-1})} = \frac{Y(z)}{X(z)}$$

and solving for $Y(z)$ results in:

$$Y(z) = \frac{X(z)\,(T + 2\tau_2) + z^{-1}\,X(z)(T - 2\tau_2) - z^{-1}\,Y(z)(T - 2\tau_1)}{(T + 2\tau_1)}.$$

In figure 4.6 the frequency response of the de-emphasis filter is plotted. The high frequency components are seen to be attenuated by as much as 20 dB. The high frequency components have been affected by the channel and noise and therefore do no contribute to the signal quality. Since the de-emphasis filter is the inverse of the pre-emphasis filter the overall gain from both filters is unity.

4.5 FM Modulation and Demodulation

4.5.1 Modulation

Current mobile cellular radio systems employ frequency modulation which is a form of angle modulation in which the instantaneous frequency is varied linearly with the

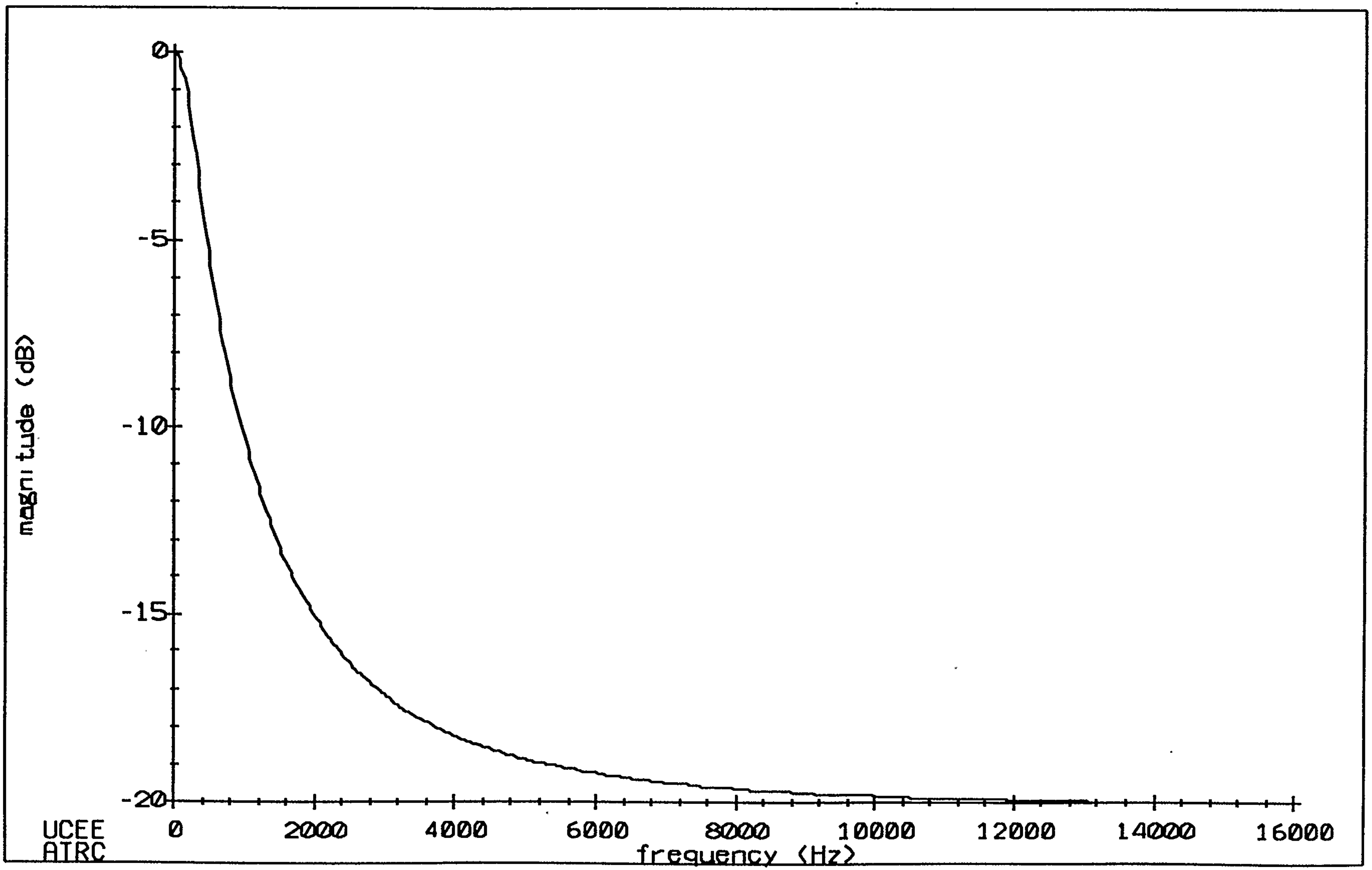


figure 4.6 - De-emphasis Filter Frequency Response

message signal. This is in contrast to phase modulation where the phase is varied linearly with the message signal. The frequency modulated waveform my be described simply as follows (Lathi 1983):

$$s(t) = A_c \cos \theta(t)$$

where:

A_c = the carrier amplitude and

$\theta(t)$ = the phase of the modulated signal.

The instantaneous frequency can be represented as follows:

$$f_i(t) = f_c + \frac{k_f}{2\pi} m(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt}$$

where:

$f_i(t)$ = the instantaneous frequency (Hertz),

f_c = the unmodulated carrier frequency (Hertz),

k_f = the frequency sensitivity of the modulator and

$m(t)$ = the message signal.

The frequency modulated waveform is a non linear function of the message signal $m(t)$ and results in the following form:

$$s(t) = A_c \cos [\, 2\pi f_c t + 2\pi k_f \int_0^t m(\tau)\, d\tau + \phi_o \,]$$

where ϕ_o = arbitrary initial phase (radians).

The frequency modulated waveform results in a signal centered around the carrier frequency. Since the signal is narrowband (30 kHz) the signal can be represented at baseband using its complex envelope equivalent. This form may be expressed as follows:

$$\tilde{s}(t) = A_c \exp [\, j2\pi k_f \int_0^t m(\tau)\, d\tau + \phi_o \,].$$

For simulation and analysis purposes the use of the complex envelope enables the use of a lower sampling rate (32 kHz). Otherwise at radio frequencies of 800 MHz, Nyquist sampling theory states that the sampling frequency would have to be approximately twice the carrier frequency thus requiring a much higher sampling rate.

4.5.2 Limiter

In FM systems, the nonlinear modulation results in the spectrum of a constant envelope FM signal not being bandlimited. However, the FM wave spectrum is found to contain a limited number of side frequencies of significant intensity. It therefore is possible to define and specify an effective transmission bandwidth. This is accomplished by specification of the maximum instantaneous frequency deviation factor (Δf) and represents the maximum departure of the instantaneous frequency of the FM wave from the carrier frequency. This may be expressed as follows:

$$f_i = f_c + \Delta f = f_c + \frac{k_f}{2\pi} m(t)_{\max}$$

$$\text{where } \Delta f = \frac{k_f}{2\pi} m(t)_{\max} .$$

The maximum instantaneous frequency deviation is therefore proportional to the amplitude of the modulating wave and is independent of the modulation frequency.

Current cellular radio is assigned a transmission bandwidth of 30 kHz. The bandwidth may be defined using Carson's rule with:

$$B_{FM} = 2\,(\Delta f + B)$$

where B = the bandwidth of the message.

The deviation limiter partially functions to bandlimit the FM wave by specification of the maximum deviation factor Δf_i and by limiting the signal so that the maximum frequency deviation specified is not exceeded. Current cellular systems limit the instantaneous frequency deviation to ± 12 kHz. The deviation limiter is followed by a lowpass filter with attenuation characteristics of 40 dB rolloff per decade. This lowpass filter also removes unwanted transients from the transmitted signal.

4.5.3 Demodulation

In FM systems, information resides in the instantaneous frequency. The detection or recovery of this information is the demodulation process.

The transmitted FM signal has a continuous phase and a constant envelope. The effects of the channel and of noise may cause the envelope of the signal to vary, however, removal of this effect facilitates the performance of the demodulation scheme. In the simulation this was accomplished by using an automatic gain control (AGC) device to insure a constant signal envelope.

The detection of the message is accomplished by finding the difference between the previous signal and the current signal (a differentiation process). The quadrature component of this difference divided by the current signal yields the demodulated signal. This scheme describes the limiter-descriminator demodulation process. It can be shown to yield the message signal as follows:

$$\tilde{s}(t) = A \exp\left[\, j2\pi k_f \int_0^t m(\tau)\, d\tau \,\right]$$

where:

$\tilde{s}(t)$ = complex envelope of the FM waveform and

A = a magnitude of the FM waveform.

If this expression is differentiated it yields the following:

$$\frac{d\,\tilde{s}(t)}{dt} = j2\pi k_f\ m(t)\ A\ \exp\left[\ j2\pi k_f \int_0^t m(\tau)\ d\tau\ \right].$$

The division of these two expressions yields the following:

$$\frac{\frac{d\,\tilde{s}(t)}{dt}}{\tilde{s}(t)} = j2\pi k_f\ m(t).$$

Further division by a quadrature factor of $j2\pi k_f$ would then yield the message component $m(t)$.

$$i.e.\ \ m(t) = \mathrm{Im}\left[\frac{1}{2\pi k_f}\ \frac{\frac{d\tilde{s}(t)}{dt}}{\tilde{s}(t)}\right]$$

4.5.4 FM Spectra

The spectrum of a tone modulated signal may be derived theoretically and as such may be compared to the spectrum obtained at the output of the FM modulator.

The theoretically derived spectrum is obtained from the complex envelope of the FM modulated sinusoidal wave which is written as (Haykin 1989):

$$s(t) = A_c \cos\left[\ 2\pi f_c t + \beta \sin\left(\ 2\pi f_m t\right)\ \right]$$

$$\text{where } \beta = \frac{\Delta f}{f_m} = \text{modulation index.}$$

The complex envelope may then be expressed as:

$$\tilde{s}(t) = A_c \exp\left[\ j\beta\ sin(2\pi \sin\left(2\pi f_m t\right)\ \right].$$

Since the complex envelope is a periodic function it may be expressed as a complex Fourier series. The spectrum is obtained by determining the complex Fourier coefficients.

The complex Fourier coefficients are determined as follows:

$$c_n = \frac{1}{T} \int_{-T/2}^{T/2} \tilde{s}(t)\, e^{-j2\pi \frac{nt}{T}}\, dt$$

$$= f_m A_c \int_{-1/2f_m}^{1/2f_m} \exp\left[\, j\, \beta \sin(2\pi f_m t) - j\, 2\pi n\, f_m t \,\right] dt.$$

This expression may be simplified by a substitution of variables such that:

$$x = 2\,\pi\, f_m t$$

then

$$c_n = \frac{A_c}{2\pi} \int_{-\pi}^{\pi} \exp\left[\, j(\, \beta \sin x - nx\,)\,\right] dx.$$

Evaluating this integral yields the nth coefficient as the nth order Bessel function of the first kind with β as the argument and results in the following form:

$$c_n = A_c\, J_n(\beta).$$

The spectrum is obtained by summing the Fourier coefficients as follows:

$$\tilde{S}(f) = A_c \sum_{n=-\infty}^{\infty} J_n(\beta)\, \delta(f - nf_m).$$

The envelope of this theoretically derived spectrum is plotted along with the spectrum obtained at the modulator output in figure 4.7. Only an outline of the shape of the theoretical spectrum is plotted in the figure to illustrate the difference between the two spectra. Both spectra exhibit impulses of varying amplitudes spaced at a frequency of

1 kHz apart. This separation corresponds to the frequency of the input tone. This tone response illustrates harmonic distortion caused by the nonlinear nature of the modulator. This plot also indicates that the FM modulator is functioning as expected.

4.6 Lowpass Filtering

Lowpass filtering is used in many parts of the simulation whenever it is desired to limit the spectrum to some specified bandwidth. The speech is lowpass filtered to have a bandwidth of 4 kHz at the output of the system since telephones are limited to this bandwidth. Lowpass filtering is also used as a discriminator filter at the demodulator in order to ensure that the bandwidth of the noise does not exceed that of the message signal bandwidth. In addition, the lowpass filters function to reduce transients occurring from sharp discontinuities such as at the output of the limiter.

Nyquist pulses such as the raised cosine pulse can be used as a lowpass filter. This type of pulse is desirable since the transfer function of the pulse looks like a lowpass filter, the characteristics are well defined and the phase is linear. The impulse response is as follows (Benedetto et al 1987):

$$q(t) = \frac{sinc(t) * cos(\alpha \pi \frac{t}{T})}{1 - 4\alpha^2 \frac{t^2}{T^2}}$$

where:

α = the rolloff parameter and

T = half width of the pulse's main lobe (zero crossing point).

As t approaches $t_0 = 1/2\alpha$ both the numerator and the denominator tend towards zero. The value of the function at this point can be determined by using l'Hospital's rule.

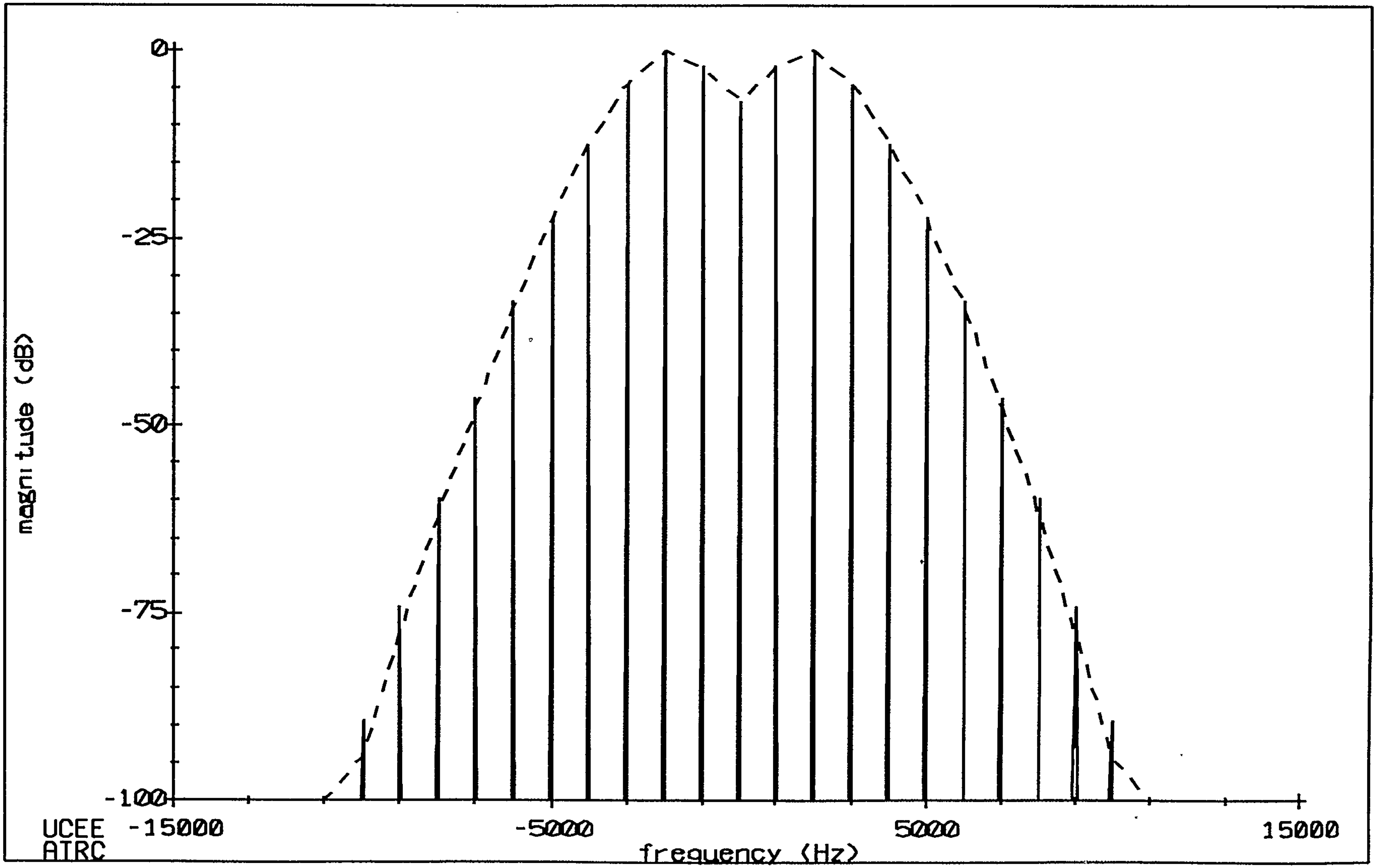


figure 4.7 - Normalized Spectrum of an FM Modulated Tone

$$ie. \quad \lim_{t \to t_0} q(t) = \frac{\left.\frac{d}{dt} n(t)\right|_{t = t_0}}{\left.\frac{d}{dt} d(t)\right|_{t = t_0}} = \frac{\left.\frac{d}{dt} cos\left(\pi\alpha\frac{t}{T}\right)\right|_{t = t_0}}{\left.\frac{d}{dt}\left[1 - 4\alpha^2\frac{t^2}{T_2}\right]\right|_{t = t_0}}$$

$$= \left.\frac{-\pi\ \alpha\ sin(\pi\alpha t)}{-8\alpha^2 t}\right|_{t = t_0} = \frac{\pi}{4}.$$

Since this pulse is symmetric in the time domain about the origin, it has zero phase and when it is shifted to become causal and realizable it then exhibits a linear phase characteristic.

In the frequency domain the raised cosine pulse consists of a flat amplitude portion between

$$f = -\frac{(1-\alpha)}{2T} \quad \text{and} \quad f = \frac{(1-\alpha)}{2T}$$

and a rolloff portion which has a sinusoidal form between

$$f = \pm\frac{1-\alpha}{2T} \quad \text{and} \quad f = \pm\frac{1+\alpha}{2T}.$$

The desired lowpass filtering is achieved by selecting the rolloff (α) and the bandwidth (pulsewidth) parameters appropriately. The rolloff parameter α specifies the rolloff of the function. It also indicates the amount of bandwidth used by the raised cosine pulse in excess of the Nyquist bandwidth corresponding to an ideal lowpass filter. This ideal solution would be a time domain sinc pulse or equivalently a frequency domain rectangular pulse.

$$ie. \quad \alpha = \frac{excess\ Nyquist\ bandwidth}{Nyquist\ bandwidth}$$

The pulsewidth or bandwidth is related to the frequency at which cutoff is desired. Since this pulse is not an ideal solution, the cutoff may be specified at the zero crossing point of the time domain pulse. It may also be equivalently specified at the 3dB power point of the frequency domain pulse which would correspond to the ideal solution cutoff:

$$ie.\ BW = bandwidth\ of\ the\ pulse = \frac{1}{2T}$$

Having determined the main lobe pulse width (T) the implementation of this filtering operation requires the function to be sampled digitally. This is done by the following substitution:

$$t = i\Delta t = \frac{i}{f_s}$$

where:

t = is the time variable,

i = integer index variable,

$\Delta t = \frac{1}{f_s}$ = data sampling period and

f_s = data sampling frequency.

Truncation is done at about ten cycles (T) on either side of the maximum value of the pulse, since with a rolloff of 0.25 the magnitude of the pulse is attenuated significantly at that point. This corresponds to a symmetric filter of length given by:

$$length = \left\lceil \frac{10}{2\,BW\ \Delta t} \right\rceil - 1.$$

The frequency response of the lowpass filter is illustrated in figure 4.8 where only positive frequencies are plotted. The flat amplitude portion continues to approximately 3 kHz. At a frequency of 4 kHz the pulse has been attenuated by approximately 40 dB.

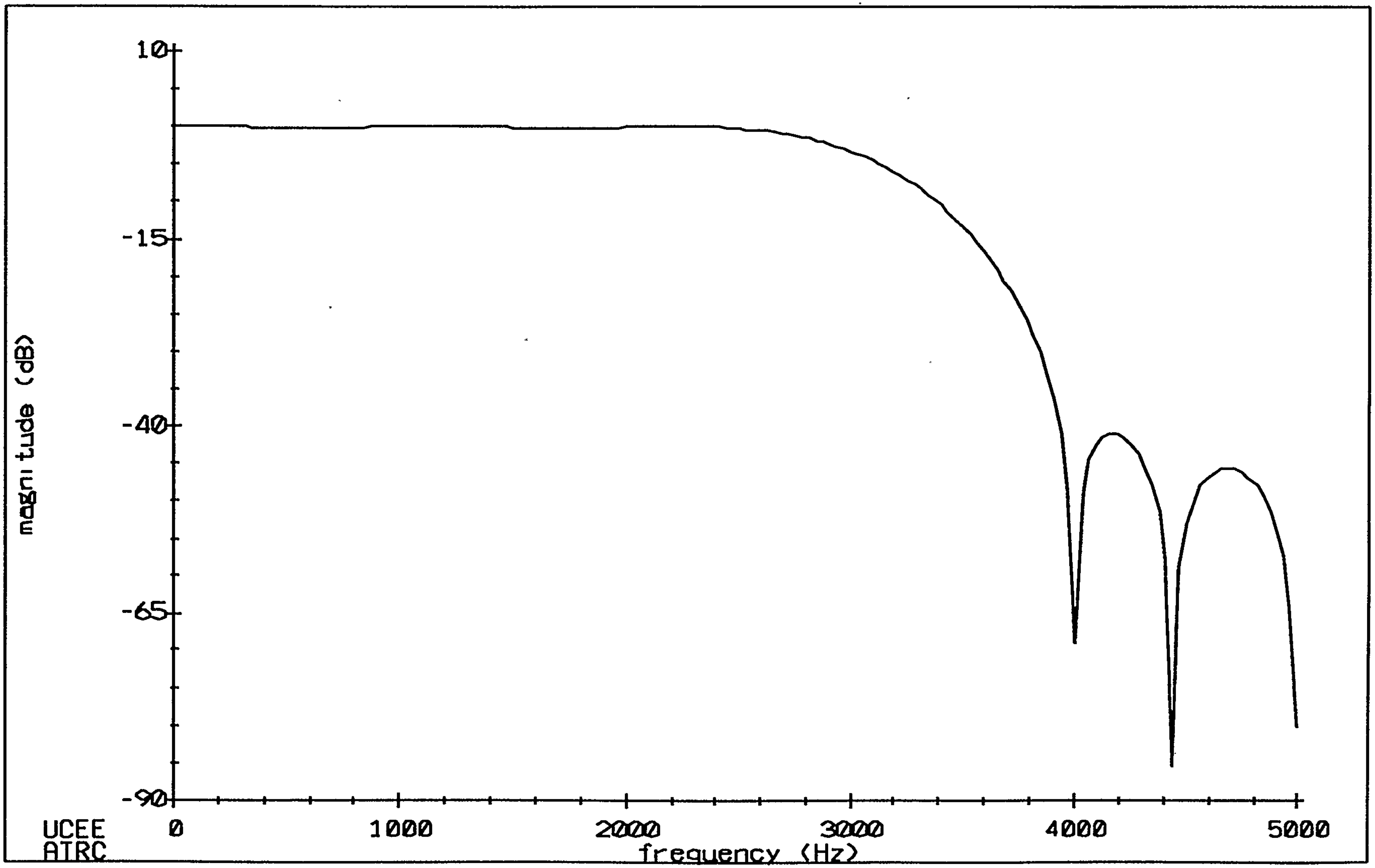


figure 4.8 - Lowpass Filter Frequency Response

4.7 Conclusions

Each component of the simulated FM system has been investigated so as to ensure that the components appear to function as their actual analog counterparts or as expected. When several components are placed together overall system response may differ somewhat from that of the analog system. With proper selection of the various component parameters this difference can be minimized. Analysis of system response and system performance to sinusoidal inputs would lead to confidence in the use of the simulated system.

Chapter Five

System Performance Results

5.1 Introduction

In order to have confidence in the performance of the system it is necessary to investigate the system response. This begins by the calibration of the system and by observing the response to a tone input. A tone is a sinusoidal wave of a single frequency.

The performance quality may be expressed in terms of output SNR values. In tone modulated systems the relationship between input and output SNR values is available in empirical form. This expected value is compared to the actual output SNR values obtained from the simulated system using synchronous sampling. For speech sources the relationship between input and output SNR values is not easily determined and therefore other means of measurement are necessary. The method used is based on correlation between the input and output to and from the system. This method is shown to achieve comparable results to that of the synchronous sampling method.

Finally the results of the simulation employing various types of diversity combining is presented.

5.2 System Calibration

The performance of the FM system is dependent upon the choice of the parameters of the various system components which function in an interrelated manner. The parameter selection or calibration of the system is generally accomplished by passing a 1000 Hz tone through the system. The parameters which are affected are the scaling factor of the input signal or data, the focal point of the compandor, the maximum gain of the limiter before clipping occurs and a rescale factor.

The scaling factor applied is dependent upon the data type and the maximum signal value allowed to the compressor. The data type may be in short integer format for binary sources and is converted to a floating point format.

The focal point of the compandor sets the distance from the maximum gain of the compandor. Selection of this parameter affects the relative gains applied to the signal. The relative gains are dependent upon the intensity range of the compandor.

The limiter functions to limit the amplitude of the signal to a maximum magnitude of unity since the signal at the limiter input has been gained from the compressor and the pre-emphasis. Selection of the limiter's maximum level must consider that signal above this value results in distortion of the signal waveform while setting the level at a high level results in attenuation of average signal levels.

The system is required to occupy a bandwidth limited to 30 kHz. In order to achieve this requirement experience has shown that the instantaneous frequency deviation must be limited to ± 12 kHz. The system should be calibrated so that values should not exceed this limit however when occasional high signal level or transients occur the limiter functions to clip these large signal variations. The amplitude of the signal will also be affected by the pre-emphasis, the compressor and any scaling of the signal level. The pre-emphasis will apply a gain from 0 to 20 dB depending upon the frequency of the signal. The transmitter is calibrated to result in a nominal ± 2.9 kHz peak frequency deviation from a 1 kHz tone input.

The rescale factor is applied at the FM demodulator output of the receiver. This is required to restore the signal to its original value. Calibration is done to achieve a nominal value of a 1 kHz tone from a ± 2.9 kHz peak frequency deviation.

5.3 System Performance Measure

5.3.1 Empirical Output SNR

System performance is often evaluated in terms of the relationship between the output signal-to-noise ratio (SNR) and the input SNR (Lathi 1983). The output SNR is defined at the output of the receiver as follows:

$$\textit{output SNR} \;=\; \frac{\text{average power of message signal}}{\text{average power of noise}}.$$

The determination of the output SNR is somewhat difficult since angle modulation is a non-linear type of modulation and as such the signal and noise form intermodulation components. This implies that a linear relationship between the input SNR and the output SNR does not exit because signal and noise distortions occur.

5.3.2 Synchronous Sampling

The output SNR for tones may be determined by synchronously sampling at the output of the system. Synchronous sampling is used in order to differentiate the signal and noise at the system output and is possible since the system is sampled such that there is an integral number of samples per cycle of the tone frequency. The periodogram of the data is obtained as a spectrum estimate which resolves the output into its frequency components so that the signal (tone frequency) can be distinguished from the noise components. Thus the output SNR can be obtained for tone modulation. Energy spillage or smearing is avoided by having an integral number of cycles in the periodogram.

5.3.3 Correlation Coefficients

The output SNR may also be determined by comparing the output signal to the input signal. This comparison is made by determining the maximum correlation coefficient of the data set and is therefore not restricted to tone modulation but may also be used for data/speech sources.

If the output signal of the FM system is compared to the input signal to the system, the difference between the two will be a measure of the distortion caused by this system. If some of this distortion can be accounted for by an ideal system with a delay and scaling factor, done because these are not considered to be true distortion, then the remaining distortion can be considered to represent the noise. This ideal system is obtained by a least squares solution where the difference between the actual system output and the ideal system output is represented by a normalized mean square error. The least squares approximation of the ideal system to the actual system results in a minimization of the mean square error.

In this method the output SNR is determined by considering the output signal and noise powers defined as follows:

$$output\ SNR = \frac{average\ signal\ power}{average\ noise\ power} = \frac{E\ [\ |\ y(t)\ |^2\ ]}{\min E\ [\ |\ e(t)\ |^2\ ]} = \frac{\sigma_y^2}{mmse}$$

where:

$y(t)$ = output signal,

σ_y^2 = variance of the signal and

$mmse$ = minimum mean square error.

The minimum mean square error is obtained by minimizing the mean square error as follows:

$$mean\ square\ error = E\ [\ |\ y(t) - \lambda\, x(t-k)\ |^2\ ]$$

where:

$y(t)$ = output signal,

$x(t)$ = input signal,

λ = scaling factor and

k = sample delay of signal.

The minimum mean square error is obtained by minimizing the mean square error as follows:

$$\textit{mean square error} = E\,[\,|\,y(t) - \lambda\, x(t-k)\,|^{2}\,]$$

$$= E\,[\,(y(t) - \lambda\, x(t-k)\,(y(t) - \lambda\, x(t-k))^{*}\,]$$

$$= E\,[\,y(t)\,y(t)^{*}\,] + |\lambda|^{2}\, E\,[\,x(t-k)\,x(t-k)^{*}\,]$$

$$- \lambda^{*}\, E\,[\,y(t)\,x(t-k)^{*}\,] - \lambda\, E\,[\,x(t-k)\,y(t)^{*}\,]$$

since $E\,[\,y(t)\,x(t-k)^{*}\,] = R_{yx}(-k) = R_{xy}(k)$

and $E\,[\,(y(t)\,x(t-k)^{*})^{*}\,] = R_{yx}^{*}(-k) = R_{xy}^{*}(k)$

$$= \sigma_y^2 + \sigma_x^2 \left\{ |\lambda|^2 - \frac{1}{\sigma_x^2}\left[\lambda^{*}\, R_{xy}(k) + \lambda\, R_{xy}^{*}(k) \right] \right\}$$

by completing the square:

$$= \sigma_y^2 + \sigma_x^2 \left\{ \lambda - \frac{R_{xy}(k)}{\sigma_x^2} \right\} \left\{ \lambda^{*} - \frac{R_{xy}^{*}(k)}{\sigma_x^2} \right\} - \frac{R_{xy}(k)\, R_{xy}^{*}(k)}{\sigma_x^2}$$

$$= \sigma_y^2 + \sigma_x^2 \left| \lambda - \frac{R_{xy}(k)}{\sigma_x^2} \right|^2 - \frac{\left| R_{xy}(k) \right|^2}{\sigma_x^2}$$

this quantity may be minimized by setting λ such that:

$$\lambda = \frac{R_{xy}(k)}{\sigma_x^2}$$

then $\quad mmse = \sigma_y^2 \left[1 - \dfrac{\left| R_{xy}(k) \right|^2}{\sigma_y^2\, \sigma_x^2} \right]$

The correlation coefficient may be defined as:

$$\rho^2 = \frac{\left| R_{xy}(k) \right|^2}{\sigma_x^2 \, \sigma_y^2}$$

then $\quad mmse = \sigma_y^2 \, [\, 1 - \rho^2 \,].$

The output SNR is then defined as follows:

$$output\ SNR = \frac{average\ signal\ power}{average\ noise\ power} = \frac{\sigma_y^2}{mmse} = \frac{1}{1 - \rho^2}.$$

From this derivation it is shown that the output SNR may also be determined by comparing the output signal to that of the input signal. This comparison is made by determining the maximum correlation coefficient of the data set and then calculating the resulting output SNR performance measure.

5.3.4 Comparisons

The synchronous sampling method and the correlation coefficient method were both used to determine the output SNR values for tone modulation. A 1 kHz tone (with β = 2.9) is put through the modulator and the demodulator of the FM system in order to compare the two methods. This comparison is illustrated in figure 5.1.

The two methods resulted in comparable output SNR values except at low input SNR values. Here the synchronous sampling method resulted in a negative output SNR. The correlation coefficient method will never result in a negative output SNR since the minimum value is zero obtained when there is no correlation between the input and output sources. From this plot confidence can be obtained in using the correlation coefficient method for speech sources.

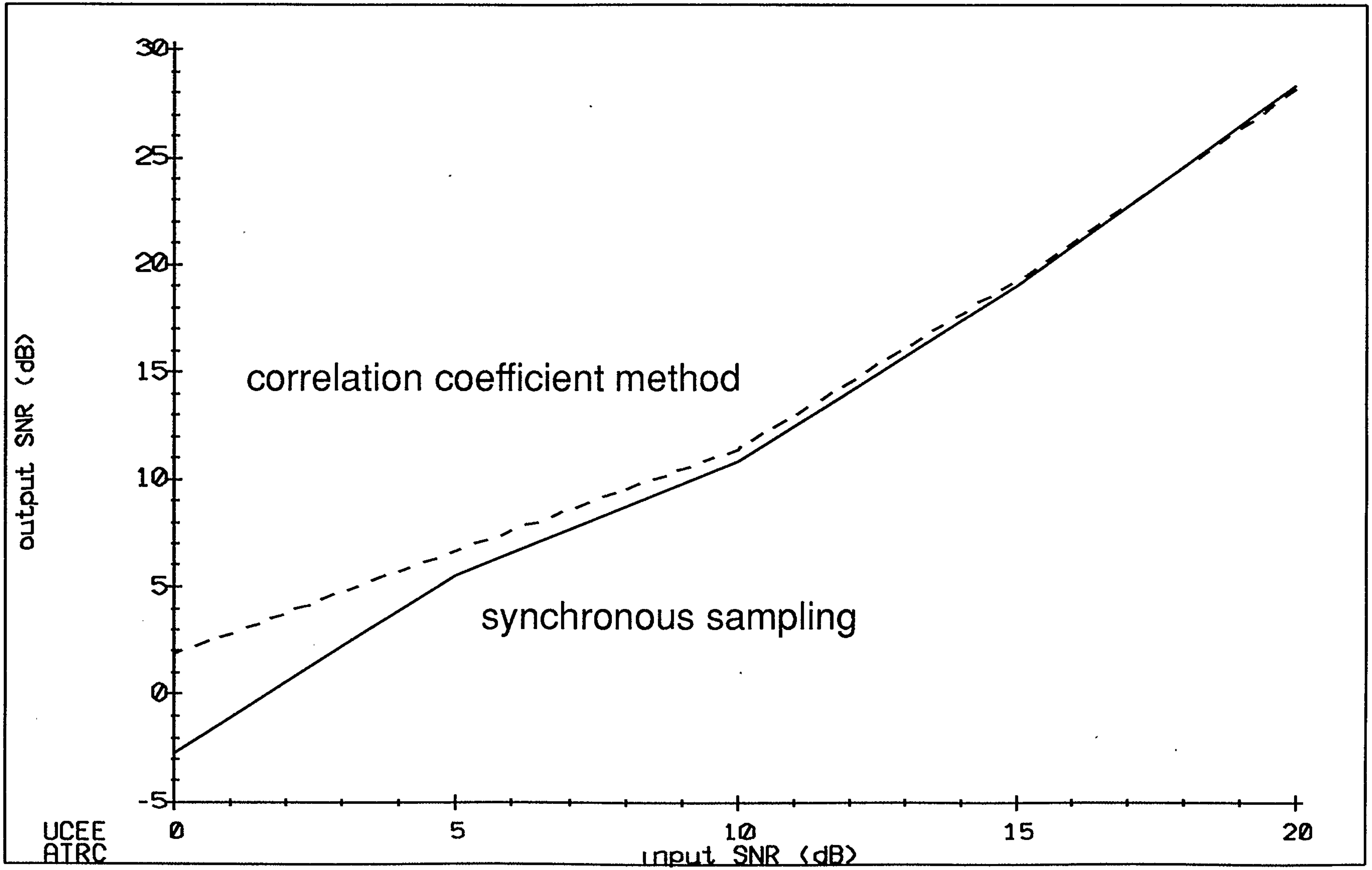


figure 5.1 - Tone Synchronous Sampling vs Correlation Coefficient Method

5.4 Diversity Combining Performance

5.4.1 Objective Speech Comparisons

The performance of various antenna diversity combining methods for an FM system is obtained through system simulation. Speech is digitized using an analog to digital (A/D) converter. This digitized speech is put through the simulated FM system which produced an output speech file. This output speech file could be converted to actual audio speech by using a digital to analog (D/A) converter. The output speech file is cross-correlated with the input speech file to determine a correlation coefficient from which the performance measure of output SNR is calculated. The procedure is repeated for various input SNR's for each system configuration studied.

The system configurations investigated included switched selection, audio level combining, equal ratio combining and flat fading types of diversity combing as well as comparisons to no diversity with a channel and with the ideal channel. These results are illustrated in figure 5.2.

Comparisons of the system with no diversity are made between the ideal channel and the system using Hashemi's channel model. The simulations indicate that performance of the system with an ideal channel is better than that with the Hashemi channel. The channel causes a distortion of the signal which may be quite severe when a deep fade is encountered. This distortion results in poorer system performance.

The switched selection and audio level combining methods investigated both involve a switch and stay strategy for signal selection. This implies that switches from one antenna to another occur only when the signal level from the current antenna drops below a predetermined threshold level. A comparison of various threshold levels using the switched selection combining method are presented in figure 5.3. This figure

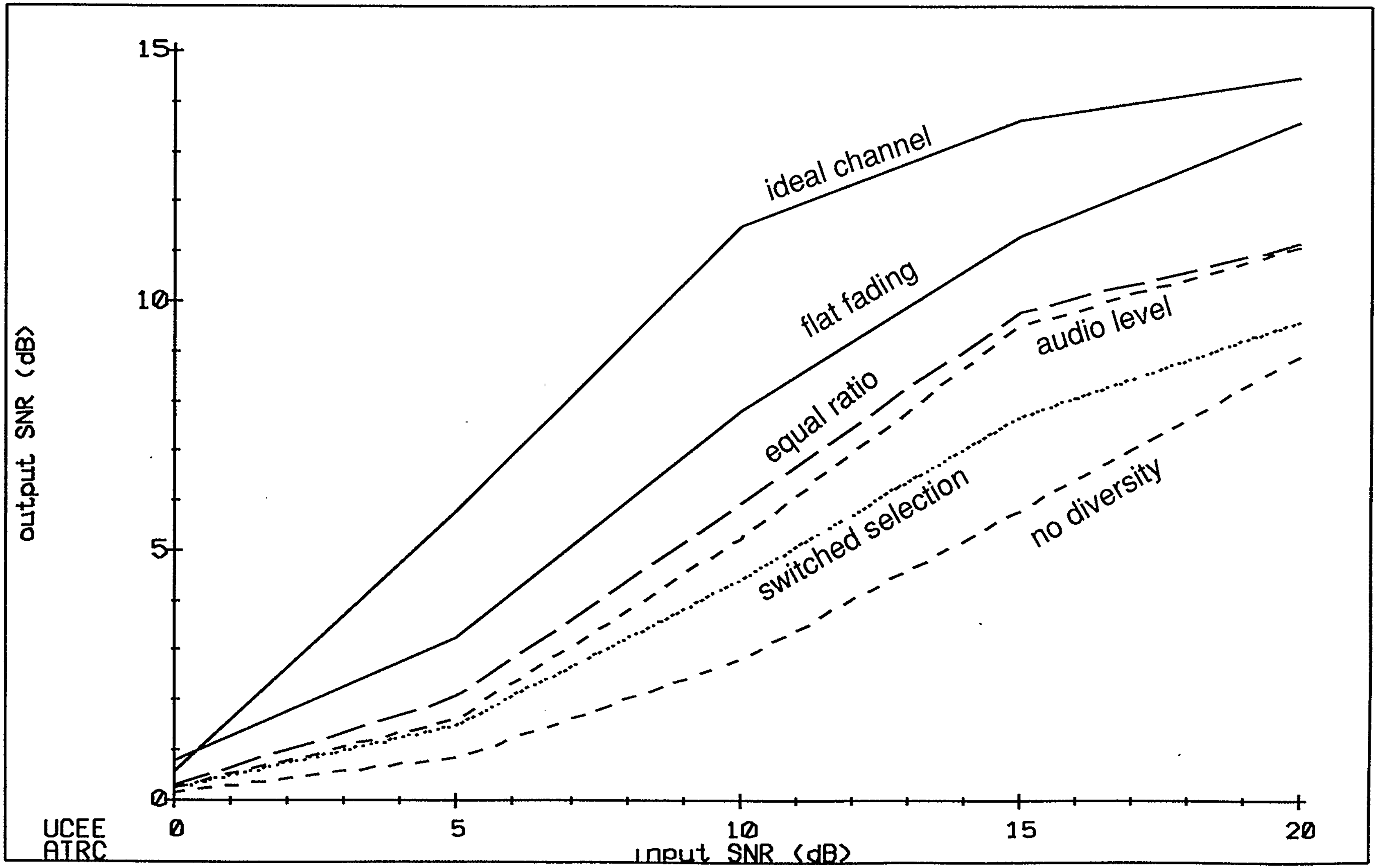


figure 5.2 - Diversity Performance Comparisons for Speech

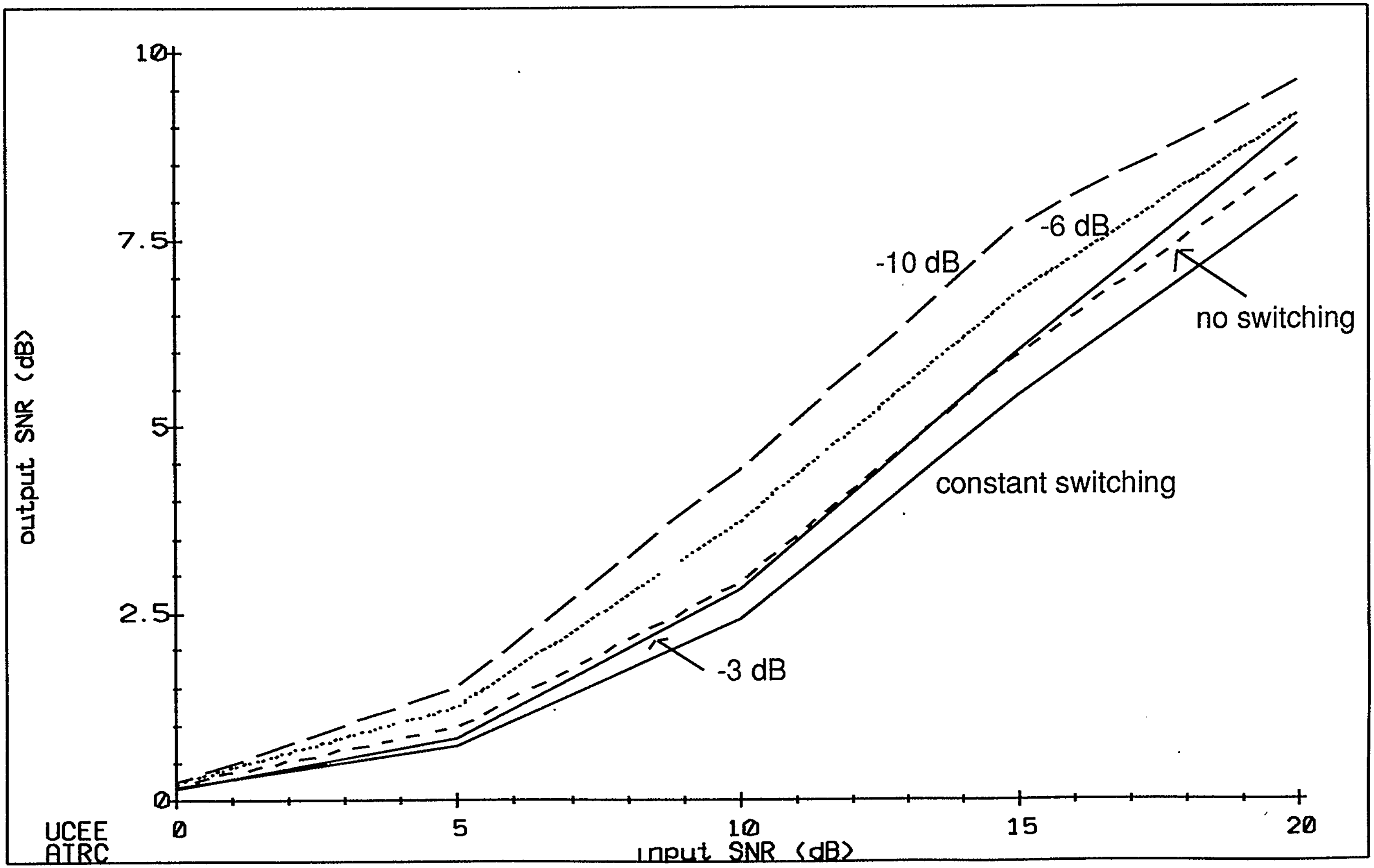


figure 5.3 - Switched Selection Threshold Comparisons

illustrates that performance is improved when the threshold level is set so that switching does not occur very often. Switching causes phase discontinuities which appear to affect system performance. Figure 5.3 also indicates that poor selection of the threshold level may result in a performance which is inferior to not having a diversity scheme. Simulation of the switched selection method of combining indicates that performance improvement is obtained over the simulation with no diversity when a threshold level of 0.1 corresponding to -10 dB is selected. This performance improvement is perhaps limited by phase discontinuities which occur during antenna switching.

A system employing two receivers as in audio level combining is suggested which would be a method of combining the two signals coherently. This is simulated and a performance improvement is obtained. Unfortunately this would be an unrealistic method of diversity combining.

Equal ratio combining is another method in which the diversity signals are combined coherently. Performance increases are expected since both antennas contribute to the final signal. In this method the signals from the two antennas are co-phased before being added together. This results in a coherent summing of the signals as well as an incoherent summing of the noise elements. An additional advantage is the reduction of random FM. Performance results indicate a significant improvement over other methods of diversity combining discussed previously.

The coherent addition of the two antennas is possible for a flat fading channel. By taking the complex sum of the channel impulse response and eliminating the phase, the resulting value may be used as a non time dispersive channel. This type of environment would illustrate the maximum performance that would be achieved with diversity. The FM system is simulated for this case and the performance achieved is somewhat better than the performance of the equal ratio combining technique.

5.4.2 Subjective Speech Comparisons

Typically the analysis of speech involves subjective evaluation by means of simple listening. This forms a part of the comparison because eventually any diversity implementation in a mobile cellular radio system will involve user satisfaction of perceived quality improvement.

The various speech samples analyzed include the diversity combining methods of switched selection, audio level combining and equal ratio combining. These are compared to the original speech source file which has been passed through the simulated FM system subject to an ideal channel as well as the simulated modified Hashemi channel involving no diversity.

From a number of surveyed participants, a number of observations were made and are presented. Originally it is difficult to detect differences between the various speech samples. It is therefore necessary to replay the speech samples in order to train the ear to the subtle differences. As the speech samples are reviewed, differences become more obvious. At higher input SNR values, the differences are not as obvious however at lower SNR values these differences are more noticeable.

The surveyed participants indicate that at high SNR values the speech sample which is passed through the system with an ideal channel sounds similar to the original source. As the noise is increased, the speech sample becomes more noisy. Significant further deterioration is observed when the non-ideal channel is included. As the various diversity combining methods are presented improvement is observed. The switched selection combining method does not indicate significant perceptible improvement. The audio level combining method results in a more significant improvement while with the equal ratio combining method somewhat greater improvement is observed. The surveyed participants generally appear satisfied with the performance of the audio level combining method with the greatest satisfaction occurring with the equal ratio combining speech sample.

Other observations generally include clicking sounds however not all participants notice this. The clarity and continuity appear to be the greatest factors in the evaluation.

These subjective tests are most significant when a comparison is made to a sample where distortion is very significant to that where clarity is increased and yields a comprehension increase.

5.5 Conclusions

The performance of the simulated analog FM system is affected by such factors as system calibration, component parameter selection and diversity implementation. The performance of the system may be evaluated by determination of the output SNR. Methods to determine the output SNR include a synchronous sampling method and a correlation coefficient method which yield comparable values for tone modulation. The correlation coefficient method is used to evaluate system performance for speech sources. Comparisons between various speech samples, obtained from using various diversity combining methods reveal that performance improvement may be obtained with diversity. These results compare favorably to subjective listening comparisons. Equal ratio combining using ideal co-phasing is found to offer the best performance improvement.

Chapter Six
Conclusions

6.1 Introduction

Diversity describes a situation in which the signal is received from independent sources such as spatially separated antennas in which case is termed spatial diversity.

The used of spatial diversity techniques in FM cellular radio is suggested as a method to improve the reception quality of signals transmitted over a flat fading channel. This possibility is most conveniently investigated by system simulation. The propagation path forms a part of this simulation.

6.2 The Propagation Path

The propagation path or channel which affects the signal quality is represented by a modified Hashemi model. This model is representative of a mobile urban environment with the associated multipath fading. The modified Hashemi model is obtained by consideration of the Doppler shift associated with each reflection caused by the vehicle motion. The spatial spectral density of the channel envelope also describes the frequency distribution. The spatial frequency components are limited to the maximum Doppler frequency resulting from vehicular movement.

Comparisons between the probability density functions of various distributions and of the channel are used to determine the distribution type. The magnitude of the complex envelop of the channel is Rayleigh distributed. The phase derivative associated with the random FM is also found to be student-t distributed. These statistical characteristics are representative of mobile channels.

Spatial diversity techniques employ two channels which are representative of the location of the two antennas and are delayed versions of the other. From the correlation of the envelope of the channel optimum spatial separations of approximately 0.4 λ are suggested. At this separation the correlation is at a minimum value and indicates the greatest independence. This optimum separation is also suggested from the theoretical correlation coefficients derived from a Rayleigh distributed envelope.

6.3 System Simulation & Performance

Simulation of an analog FM cellular system allows performance evaluations of current system configurations as well as diversity combining techniques to be evaluated.

The components of the system include lowpass filters, compressor and expandor, pre-emphasis and de-emphasis filters as well as the modulation and demodulation components.

The performance of the system has been evaluated by determination of the output SNR. For tone modulation, the synchronous sampling method is comparable to the correlation coefficient method in determining the output SNR. The correlation coefficient method was used for speech sources.

Included in the investigation of spatial diversity schemes are the combining techniques of switched selection, audio-level combining and equal-ratio combining. Results indicate that equal-ratio combining appears to yield the best performance improvements. The performance of this technique will be dependent upon the type of co-phasing technique employed to achieve coherent signal addition from the two antennas.

6.4 Future Work

Future work would be to investigate and develope a technique for co-phasing. System simulation would be used to evaluate the technique before actual hardware implementation required for field trials.

The simulation could also be used to investigate data transmission techniques such as quadrature amplitude modulation (QAM) on FM. Spatial diversity could also possibly play a role in these schemes.

Spatial diversity is also applicable to digital systems and this should also be investigated. This would involve the development of a digital system simulation involving QAM or differential phase shift keying types of modulation.

Bibliography

Benedetto, S., Biglieri, E., Castellani, V., *Digital Transmission Theory,* Prentice-Hall, Englewood Cliffs, N.J., 1987.

Brigham, E.O., *The Fast Fourier Transform,* Prentice-Hall, Englewood Cliffs, N.J., 1974.

Eggers, P.C.F. and Andersen, J.B., "Measurement of complex envelopes of mobile scenarios at 450 MHz", IEEE Trans. Vehicular Tech. Vol 38, No. 2, May 1989.

EIA/TIA-553, "Mobile Station - Land Station compatibility specification", Electronic Industries Association, September 1989.

Fattouche, M., Petherick, L., and Fapojuwo, A., "Diversity for mobile radio communications," Proceedings of the 15th Biennial Symposium on Communications, pp. 196-199, Kingston Ontario, Canada, June 1990.

Gibra, I.N., *Probability and Statistical Inference for Scientists and Engineers,* Prentice-Hall, Englewood Cliffs, N.J., 1973.

Hashemi, H., "Simulation of the urban radio propagation channel", IEEE Trans. Vehicular Tech. Vol VT-28, August 1979.

Haykin, S.S., *An Introduction to Analog and Digital Communications,* John Wiley and Sons, New York,N.Y., 1989.

Jakes, W.C.Jr., "A comparison of specific space diversity techniques for reduction of fast fading in UHF mobile radio systems", IEEE Trans. Vehicular Tech. Vol VT-20, No. 4, November 1971.

Jakes, W.C., *Microwave Mobile Communications,* Wiley, N.Y., 1974.

Lathi B.P., *Modern Digital and Analog Communication Systems,* Holt, Rinehart and Winston, N.Y. 1983.

Lee. W.C.Y., *Mobile Communications Engineering,* McGraw-Hill, N.Y., 1982.

McClave J.T., Dietrich F.H., *Statistics,* Dellen Publishing Company, San Francisco, 1988.

Poularikas W.H., Seely S., *Signals and Systems,* PWS-KENT Publishing Company, Boston, 1991.

Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., *Numerical Recipies in C,* Cambridge University Press, Cambridge, 1988.

Proakis, J.G., *Digital Communications,* McGraw-Hill, New York, 1983.

Rummler, W.D., Coutts R.P., and Linger, M., "Multipath fading channel models for microwave digital radio", IEEE Commun. Mag. pp 30-42, Nov. 1986.

Stein, S., "Fading Channel Issues in System Engineering", IEEE Journal on Selected Areas in Communications, Vol. SAC-5, No. 2, Feb. 1987.

Turin, G.L., "A Statistical Model of Urban Multipath Propagation", IEEE Trans Vehicular Tech., Vol. VT-21, Feb. 1972.

"The Theory and Use of Compandors in Voice Transmission Systems", The Lenkurt Demodulator, Oct. 1964.

"dB and other Logarithmic Units", The Lenkurt Demodulator, April 1966.