Propagation Studies and Modulation Techniques for a Distributed Architecture Rural Radio-Telephone System

Simon Browne

Department of Engineering
University of Warwick

Contents

List of Figures

List of Tables

Acknowledgments

I would mainly like to thank my supervisor, Dr Steve Chandler, of Warwick University for his guidance and support throughout the period of the work. The knowledge that he would always find time to help with problems was very re-assuring, and his constant commitment greatly appreciated.

For the propagation work, thanks are due to Ken Stealey, who was also working on the system at the time, for writing the software to allow the recording, and displaying, of the received data on the Toshiba PC. Also, Prof. B. Honary, now of Lancaster University, is thanked for allowing the use of one of the Icom transceivers belonging to his group. For actual measurement work, SLNTC were extremely helpful in providing a vehicle, technician and driver for the period. Many Sierra Leonians helped in the measurement work, operating the transmitter, helping to put up aerials etc., which was very important for success of the work.

The Science and Engineering Research Council, SERC, are thanked for their financial support in the form of the studentship with which I was funded.

IF	Intermediate frequency
ISI	Inter-symbol interference
LINC	Linear transmitter using non-linear components
LO	Local oscillator
LOS	Line of sight
MSK	Minimum shift keying
NRZ	Non return to zero
PA	Power amplifier
PM	Phase modulation
PSK	Phase shift keying
PRBS	Pseudo-random binary sequence
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
RF	Radio frequency
ROM	Read only memory
SSB	Single sideband
TDD	Time-division duplex
TDMA	Time-division multiple access

IF	Intermediate frequency
ISI	Inter-symbol interference
LINC	Linear transmitter using non-linear components
LO	Local oscillator
LOS	Line of sight
MSK	Minimum shift keying
NRZ	Non return to zero
PA	Power amplifier
PM	Phase modulation
PSK	Phase shift keying
PRBS	Pseudo-random binary sequence
QAM	Quadrature amplitude modulation
QPSK	Quadrature phase shift keying
RF	Radio frequency
ROM	Read only memory
SSB	Single sideband
TDD	Time-division duplex
TDMA	Time-division multiple access

Abstract

The work described in this thesis forms part of the development of a novel digital distributed radio network. In particular, the areas of radio propagation and modulation are considered.

Field measurements of radio channel characteristics made in Sierra Leone are described. The results are presented, together with a description of the implementation of the measuring equipment. Both transmission loss and channel impulse responses were measured. Measured loss values are compared with theoretical values calculated using standard routines. The measurements were made at a frequency of 53 MHz.

The implementation of a spectrally efficient modulation scheme using a power efficient transmitter is detailed. Transmitter linearization schemes are described. Consideration is also given to filtering techniques applicable to look-up table based transmission. An overall transmitter has been produced, operating at 53 MHz, and the results are given.

Chapter 1

Introduction

1.1 Introduction

This chapter gives the background to the work presented in the thesis. After some general information on the application of the work, the radio system under development is briefly described. The modulation and propagation aspects, with which this work has been concerned, are then mentioned.

1.2 Telecommunications and the Developing World

1.2.1 The Need for Telecommunications

Whilst the telephone is generally taken for granted as being part of life in the developed world, there exists an enormous number of people throughout the developing world having no access whatsoever to this technology [12].

For the vast majority, if not all, developing countries there is some form of communications infrastructure in the larger towns and cities, however it is a very different situation in the rural areas. It can be argued that the rural areas actually provide almost all of the income for these countries from such activities as agriculture, mining and tourism, since generally the cities produce little, serving in the main as administration and service centres. The rural communities therefore deserve consideration when it comes to encouraging a country's growth, and in the distribution of its wealth. Since in many cases these communities do not actually benefit to any significant degree from the wealth generated, the standard of services available and general living conditions may be inferior to those in the cities. Certainly the cities offer some opportunities of finding material wealth and types of work not found in the rural areas, and although these opportunities may actually present themselves to only a very small minority, it is sufficient reason for many to leave the rural areas for the cities in the hope of a better life. The country's economy will rarely benefit, on the whole, from this migration however, since this extra workforce is likely to be less productive than it would be in the rural areas. The social problems associated

with this population shift are well-known, and are in evidence in many cities in the developing world. If the quality of life could be improved in the rural areas, however, people may be less inclined to leave, and instead stay and actively contribute to the wealth and well-being of the community. By creating some degree of telecommunications infrastructure, many of the hindrances to providing reasonable services may be removed and people would feel far less isolated than previously [3, 4, 10].

1.2.2 Applications of Telecommunications

Health care and agriculture are two examples of the areas in which telecommunications could play a major role. It is obvious that the ability for remote villages to have a communication link with a hospital or health centre would increase the response time to any need for medical aid, and the value of the potential to pass medical advice verbally to remote areas could be great in a large number of situations. The worth of such an infrastructure in cases of natural disaster is difficult to over-estimate, in the co- ordination of any relief effort and in the general gleaning of information on the situation. For agriculture, telecommunications can offer a great saving in time for many routines necessary in farming life in remote areas: information on markets would be instantly available, with farmers knowing where their crops could be sold, and for what price; advice on mechanical problems with machinery would be available without the need to travel to the nearest mechanic, spare parts could be ordered easily, and deals with suppliers could be decided without an actual meeting. Research in the past by such organisations as the ITU has suggested significant, and in some cases enormous, 'cost-to- benefit' ratios for telecommunication services in rural areas [4].

1.2.3 Economic Considerations

Despite the above, funding for telecommunications is likely to be inadequate for such an important service, This situation is due to a large part to the prohibitive

cost of many rural telecommunications products, the designs of which are not well matched to needs of the rural areas of developing countries. There must be, then, a great desire throughout the developing world for a low cost telecommunication system designed specifically to meet the needs of the people in the rural areas.

1.2.4 System Requirements

Having identified the need for telecommunications equipment in these areas, the first question to be asked is what type of equipment is required. This, however, is a very difficult question to answer, as it depends on large number of factors particular to the areas in where the equipment would be employed [9]. There are a number of problems faced in the installation and operation of a network in such remote areas, all of which must be carefully considered in the network design, some of the most important being the following :-

- Lack of existing infrastructure (power, roads etc.).
- Difficult terrain.
- Harsh or severe climate.
- Low density of population (and very low densities of telephones due to the low number of telephones per head of population).
- Low potential revenue.
- Poor maintenance of equipment.

From the above, it is clear that there are large differences in system requirements between the 'developed' and developing' worlds. The optimal solution to one is therefore likely to be far less optimal to the other. The general policy of applying perhaps inappropriate solutions to the above problems therefore ought to be questioned. There is a view that it would be too expensive to design equipment

specifically for the developing world, and that some of the existing lower cost equipment should be instead used. At first thought there appears to be substantial merit in this argument, however when the size of the potential market for the equipment is considered it can be seen that there are good reasons to attempt the design of a system aimed specifically at this problem.

The work initiated by Chandler [11] addresses the above problems well, with the idea of a novel radio network entirely suitable for the majority of cases in the developing world. The design is based around serving areas containing villages distributed in a fairly random manner, without the need for any existing infrastructure.

One important idea suggested was that of the spacing of telephones; it was considered that a basic service should provide sufficient telephones to limit the distance people must travel in order to reach one, to ensure that people would not be deterred from sending messages. A figure of 5 km was considered reasonable as a guideline to be used in basic system calculations. This figure would put one telephone in every 77.5 sq.km were a circular coverage used, leading to the approximation of telephone density to 0.01/sq.km. If this figure was applied to an area the size of Africa, around 300,000 telephones would be required. Figures such as these are extremely important in the design of an optimal network, and equally so in indicating the applicability of present technology to the areas being considered.

1.3 The Use of Radio

It is apparent that radio transmission is the most appropriate technique for telecommunications in by far the great majority of rural areas of developing countries [8]. The inhospitable terrain and climate create huge problems in themselves, and, with the large cost in laying any form of cable system, line transmission has many disadvantages. Radio solutions can overcome many of the problems facing line systems: installation costs are low; installation time and complexity are low; no servicing of plant is required in inaccessible locations between terminals (assuming repeaters

are not required to be placed in remote locations); additional terminals (telephones) can generally be installed and integrated into any existing system without major problems.

1.3.1 Current Systems in Use

As a means of providing communications in rural areas of the developed world, several types of system employing radio technology have been developed [8]. Multiple-access systems and satellite systems have been put into service in large areas and found to work well, although at substantial financial cost. On a less sophisticated scale, there has been use of single channel radio systems serving remote areas, with the use of repeaters where necessary. Whilst these systems may serve their purposes well, they are not, as previously stated, ideal for applications considered in the developing world. The satellite [2] and multiple-access schemes [5, 6, 7] both rely upon a reasonably high grouping of subscribers around earth stations or out-stations, respectively, to become economic. On a one-telephone-per-village basis, such subscriber densities will be reached rarely, if at all, in the developing world. Additionally, multiple-access schemes are designed to 'branch out' from existing, typically radial, telecommunication networks, and such networks are by no means common in the areas under consideration.

The fixed frequency links using repeaters are not suited to serving anything other than the smallest numbers of subscribers, being limited by the number of available channels and due to the fact that repeaters in remote locations are highly undesirable because of the problem of maintenance. The use of cellular radio technology to serve rural areas [1] is becoming increasingly popular, however this approach is not well suited to the small subscriber densities, and would require a large number of base stations and digital switches at various locations, calling for a certain degree of infrastructure and costly exchanges. Due to the low subscriber densities, each base station would effectively serve a relatively small number of stations, making

the unit station cost that much more. As an example, for a base station with a coverage radius of 30km and a telephone density of 0.01/sq.km, only approximately 30 stations would be served. The individual station unit costs, together with the share of the base station costs and of the hierarchy further up, would make the cellular approach a costly one in such an area.

Of course, where the distribution of required telephones differs significantly from the case assumed above, the relative merits of the other systems mentioned will change. If subscriber densities increase, then it may become feasible for the TDMA (time-division multiple access) or cellular type systems to be used, for example.

1.3.2 A Lower-Cost Solution

As a first step in providing a more suitable communications service, Dr Steve Chandler of Warwick University installed a basic network of solar powered, microprocessor controlled CB transceivers in the Bonthe region of Sierra Leone, West Africa [13, 17]. This system certainly met the criterion of low-cost, since the equipment in each station would total a value of around three hundred pounds, which would be far less than any other available system station costs, typically greater than 1000 pounds. Operating at 27MHz, a maximum range of approximately 50 miles was possible with the CB units over flat terrain, and with good conditions. Unfortunately, ionospheric propagation led to interference from other stations in distant countries, and, peaking in the afternoon, this interference limited range to the order of only 5 miles. The control units permitted selective calling of other stations, in addition to serving a control purpose for the provision of details of station states and other functions. This system is still in operation.

Whilst the performance of such a basic network is obviously rather limited, it serves the purpose of a local network in areas with no existing infrastructure very well indeed. Certainly, a system such as this works as a useful reference for the development of more sophisticated communication systems aimed at the same

market. The Bonthe network showed that where cost is of paramount importance, a network of CB (or VHF FM) radios, solar powered, could be a viable solution to some of the problems of rural communications in, primarily, the least developed regions of the world. The interference situation could be solved by changing to a higher-frequency system.

Unfortunately, such 'simple' radio networks are very limited in a number of ways. Range is limited to that allowed by the radio propagation, which with low power systems will be in the region of 40 - 50 km (assuming the use of VHF and relatively flat terrain), and therefore long distance communication is not generally possible, although the manual relaying of messages could be used in some instances. There would be no guaranteed privacy involved with the network, since any station could monitor the transmissions of all other stations within range, which has implications for the suitability of the system for calls of, for example, a business or very personal nature.

Analysis on the use of the network in the area has provided useful information on requirements of any service intended for this application, such as amount of time each 'telephone' is in operation, and the range of distances of calls made. Statistics like these are crucial for effective system design, and appear to be very scarce, as few small villages have been served by any telephone system in the developing world.

1.4 The Novel Network - Overview

As mentioned previously, Dr Chandler at Warwick has proposed a novel approach to providing a telecommunication service well suited to rural areas of developing countries. The system has drawn on data gathered from his pilot system in Bonthe, and uses equipment designed specifically for the task.

A basic overview of the system is as follows:- Stations (radio/telephones) are of identical form, and can act as either a terminal or a repeater at any time. An area of several hundred kilometres in diameter can be served by a network consisting of a

number of stations located in a distributed manner around the region. It is assumed that there would be a station in most reasonably sized villages, with an approximate station density, as previous given, of 0.01 per square kilometre. 'Gateway' nodes are located within the network, these being stations which are already connected to the normal public telephone network, if one exists. They enable calls to be routed into and out of the distributed radio network. The ability of stations to act as relay stations allows calls to be made over distances well beyond the normal radio horizon, and the use of digital transmission is crucial in the preservation of signal quality across such a link of relays. An illustration of the operation of the network is given in figure 1.1.

Stations would be solar powered, with battery back-up, where no mains electricity infrastructure exists. Calls are set up on a common 'calling channel', employing packet radio routing techniques, and using relay stations as necessary. A radio channel is dynamically allocated to each link, therefore several channels will be used for a call routed via a number of relays. Due to frequency reuse, however, (a single frequency channel may be used simultaneously on a number of links carrying traffic so long as the distance between these respective links is sufficient to prevent appreciable co- channel interference) this does not cause spectral problems. Once the link has been established, speech transmission, or data transmission if required, begins over the channels. Initially two approaches to speech transmission were considered. Simplex transmission, using a 'push-to-talk' and burst mode operation has the advantage that relay stations operating a store-and-forward function need only receive and transmit on single frequencies. It is, however, less user friendly than duplex systems, where the ability of communicators to speak simultaneously is far more natural. Time-division-duplex (TDD) transmission permits simultaneous transmission, however dictates that relay stations must operate in a mode of either receiving or transmitting on two different frequencies in successive time slots; this implies that each station must contain two transmitters and receivers, and hence raises the station cost and power requirements. Connection of the TDD system into any PSTN

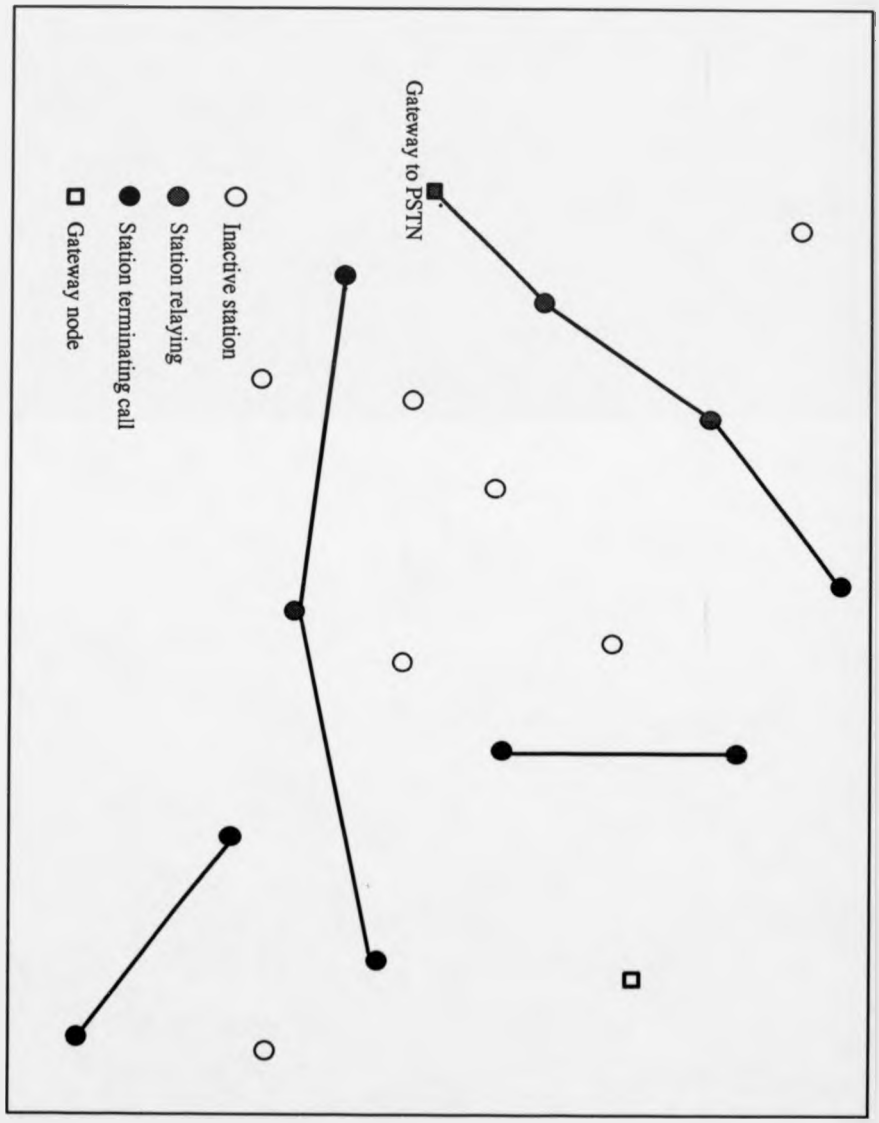


Figure 1.1: Diagram illustrating system operation

is less problematic than would be the case for the simplex system.

It was initially proposed that 32 kbps (kilobits per second) ADPCM be used to convey speech in digital form, however it is possible that advances in CODEC techniques will lower this rate to 16 kbps. Due to the burst mode operation of both systems, this rate is effectively doubled (in fact it will be slightly more than doubled to permit 'guard bands' between adjacent time slots). Typically, then, the actual transmission rate would be either 40 kbps or 80 kbps. Omni-directional antennae will be used at each station in order to give all-round coverage, the actual distances involved in hops being a function of antenna height, transmitted power and terrain type. It is anticipated that over flat terrain a coverage of up to 40 km for an individual 'hop' would be a suitable design goal for most situations, however this may have to be reduced in certain situations, as will be explained later, for reasons of congestion.

The brief overview of the system is now extended to give a more detailed description, with particular emphasis being given to the areas related to modulation and radio propagation. The Time-Division-Duplex system is assumed, which is likely to be that of main importance. Operation of a simplex network would be similar in most respects, with the exception that since duplex operation is not involved, repeater stations are only required to be able to transmit and receive on individual frequencies. A thorough, and the first detailed treatment of the system operation from a protocol point of view, is given in [16].

1.4.1 Station Description

Each station (node) consists of a dual transmitter/receiver (transceiver), and a central microprocessor acting as a small exchange. As all stations are identical, the hierarchical nature of most communication systems is avoided - an important feature. Each station will typically have one subscriber line, however consideration has been given to the provision of two lines, enabling up to about ten telephones to be

served, although this figure is obviously highly dependent upon the traffic loading of each. and there is no reason why a data terminal cannot be connected via the appropriate interface. This line, assuming the singular, can access or be accessed by any other station within the network, and hence can both initiate and terminate calls. Additionally, as each station contains dual transmitters, receivers and antennas, it has the capability to act as a relay station if required, permitting calls beyond normal radio horizon to be made.

1.4.2 Protocol Description

Call Set-Up

A common channel throughout the network is assigned as a 'calling channel'. This channel carries all call set-up data, and is therefore where stations 'negotiate' link initiation. Packet routing techniques are employed on the calling channel with access to the channel governed by a CSMA procedure. Thus, a very flexible set-up routine is possible, with a large number of potential routes available in the system for any given call. When a station attempts to call another station on the network, it sends out a 'short packet', containing information such as destination address, a synchronisation word, and an error check. It is anticipated that the short packets will be of approximately 1.6ms duration, the exact figure being determined by the CRC and synchronisation technique to be employed. Fig 1.2 shows the anticipated structure of the short packet.

Figure 1.2: Diagram illustrating short packet structure

This is received by all stations within range which are in receive mode on the calling channel, examined and re-transmitted if necessary (i.e. if the station which received the packet is not the destination station, then it will re-transmit the packet), the process continuing until the destination is reached. Call set-up then continues, involving both terminal stations and the appropriate relays, by the exchange of 'short packets' of the same form as that initiating the whole process, but containing frequency allocation information, terminal state information (engaged, free etc), and other necessary set-up data. This process takes place in an asynchronous manner. Once the set-up procedure has been completed, a link has been established, possibly over a number of channels if relays are involved, and the actual users' information can be transmitted.

An important point which arises from the possible reduction in baseband data rate due to a change of CODEC is that if the set-up procedure on the calling channel proceeds at the reduced rate then the time taken for call set-up, and therefore congestion on this frequency, will increase proportionally. The importance of this effect will obviously need appropriate consideration.

Speech Mode Operation

Once the call set-up is complete, the users' speech can begin to be sent. Circuit-switched techniques are now employed, This will be sent by burst-mode (TDD) transmission of speech digitised into a 32 kbit/s format (or less, as mentioned previously). The actual data rate of the speech information at the transmitter will be in the order of 80 kbit/s, since burst mode operation will double the initial rate, and the requirement of 'guard bands' between bursts implies a further increase in rate, so it is anticipated that 80 kbit/s will be the overall rate. Data is actually to be sent in 5 ms frames, each frame being termed a 'circuit burst', or a 'speech packet'. Within that frame will be a 0.5 ms guard band, therefore the actual duration of a circuit burst is 4.5ms. Fig 1.3 shows the structure of a circuit burst.

The 'circuit' continues in this state, with terminals and relays operating, until

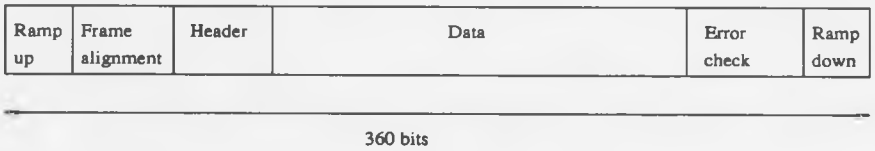


Figure 1.3: Diagram illustrating circuit burst structure

either the call reaches completion, or the circuit requires reconfiguring due to a relay being called/wanting to initiate a call or a change in the radio environment which produced an unacceptable error rate.

Null speech packets

In any speech conversation, there will obviously be large amounts of silence, as both parties will rarely be speaking continuously and simultaneously without pause. Transmitting 'silence' serves little purpose, unless a continuous signal is required for synchronisation purposes. Therefore, the time that a station is not generating useful information to be sent can be utilised for other operations. In this system, it permits stations acting in a relay capacity to switch briefly to the calling frequency and check that they are not being called to terminate a call. This is achieved by the use of a 'null speech packet', which contains data identifying itself as such, and relay stations can take the appropriate action upon its identification. Of course, the switch to the calling frequency must be of such a short duration that the receiver is once again on the appropriate conversation channel in time for reception of the following circuit burst, assuming it remains free to continue its relaying operation.

1.4.3 Radio Considerations

The operation of the system described places stringent requirements on the performance of the radio hardware [15]. This thesis concerns the investigation of appropriate radio techniques for use with this system, with particular regard to the

propagation and modulation aspects.

Spectrum

It is anticipated that around 200 channels would be required to serve a large area reliably, with relatively low densities of telephones. This relatively large number of channels will necessitate wide frequency allocation, and obviously the more the bandwidth of the individual channels can be reduced then the less the total system bandwidth required. Whilst VHF/UHF radio spectrum may not be congested in many developing countries, and therefore a large frequency allocation may be granted by regulatory bodies, this may not always be the case, and indeed the situation will be likely to change in future years. Therefore it is of major importance to keep the overall spectrum required to a minimum for a given system performance. Additionally, the bandwidth of the system places the requirement that radio hardware (antennas, power amplifiers etc) must operate satisfactorily over quite a large frequency range. Traditional analogue radio systems employing FM modulation use a channel spacing of either 12.5 or 25 kHz, however this is now only really becoming practicable in digital systems with recent advances in CODEC techniques, and work on efficient modulation techniques (high-capacity microwave systems are not considered here due to very different nature and application of such systems). Power limitations dictate the use of modulation schemes which can perform well at relatively low signal-to-noise ratios (less than 20dB), which rules out the most efficient high-level schemes. A spectral efficiency of 1 bit/sec/Hz (bps/Hz), a figure which digital cellular radio systems improve upon by only 10-20 %, would imply an overall system bandwidth of 16 MHz, assuming 200 channels and 80 kbit/sec transmission. This represents a significant percentage bandwidth (bandwidth/operating frequency) at VHF frequencies. If the spectral efficiency could be increased to 2 bps/Hz, however, the bandwidth and hence percentage bandwidth of the system would be halved. Use of a half-rate CODEC would obviously have the same effect.

Requirements of Burst-mode Operation

More detailed consideration of this issue will be given later, however, a general description of the basic problems arising due to this operation is given here. Burst-mode operation requires that the transmitter(s) should be turned on and off at a fairly high rate. This has implications for the transmitter circuitry, and for the actual spectrum of the radio signal produced.

Within the transmitter, circuits must have very short time-constants to ensure that switch on/off can be achieved in a short time, and to limit this time to several microseconds means that very careful circuit design must be followed. Standard circuits must be modified as, for example, typical decoupling circuitry would be unacceptable due to the large capacitances associated.

The frequency spectrum of a burst-mode signal may be found to be inferior to that of a signal in continuous transmission. This spectral spreading effect is due to the on/off transitions of the RF waveform, and therefore remedial measures are generally necessary if the switching is at a significant frequency. 'Ramping' up and down of the transmitted signal will be beneficial in this respect, as the harsh transients are avoided. It is insufficient, however to consider only the amplitude of the transmitted signal if ramping is used, as phase continuity is important also in minimising any spectral regeneration.

Radio Propagation

Ideally, any transmitted signal will arrive at a receiver a short time later without distortion and interference, and will be of sufficient signal strength to ensure reliable detection, hence providing perfect reception of the transmitted speech, or data, at the receiver output. Unfortunately, the radio environment is rarely so kind, and the receiver will need to recover the signal, perhaps distorted, from a certain amount of noise and interference.

In this system, propagation distances of up to 50 km are assumed, which implies

a substantial *transmission loss*, limiting the signal at the receiver to a very low level. Additionally, *smearing* of the signal in time may occur due to multipath effects, such as reflections from hills. This spreading in time can lead to *inter-symbol interference*, ISI, in digital systems, which is very likely to produce errors at the receiver. Normal point-to-point radio systems will have to operate over a given path reliably, which, if the path is poor, may necessitate large power levels and good antennas. In this system, however, there is the great advantage that a particularly useful form of diversity, *route diversity*, is inherent, and will permit a call to be made over a number of paths, some of which will obviously be much better than others. The call set-up procedure will dictate that the overall path being used will be adequate for transmission, and if, for any reason, the path should deteriorate, then reconfiguration of the route will permit the call to continue without 'dropping out', as would be the case with most other systems.

The system, being considered primarily to be a network of fixed nodes, has numerous advantages over mobile/portable networks, which are receiving much attention at present, from a propagation point of view. In the mobile situation, extremely fast fading of the transmitted signal will occur due to the motion of the vehicles, and the radio path effectively changing dramatically from one instant to another. Receivers must be able to cope with fade rates in the order of kilohertz, and with depths exceeding 30 dB. In addition, doppler shift will occur due to the motion, which will cause additional problems with the detection process. One end of a mobile link will typically have a very low, and perhaps inefficient, antenna, which will cause path losses to be large, and prevent true *line of sight* propagation for much of the time. Range will therefore be very limited. In the fixed environment, however, things are somewhat simpler, as propagation conditions will be substantially more stable, as any fading of signals will tend to be very slow, and due solely to changing atmospherics. Hence, short range links are unlikely to show significant rapid changes with time. Since stations are fixed, and permanent, antennas will be at good heights at both ends of a link, and should be very efficient. Effective range will therefore be

far greater than in the mobile environment, as path losses will be much reduced.

Since propagation data is very rare for developing countries, it was decided that it would be very useful to produce some typical figures for paths likely to be found where the system will find implementation. A period of time was spent in Sierra Leone, West Africa, characterizing a number of typical channels in different areas, and this work is detailed in Chapters 2 and 3. Sierra Leone offers a wide variety of terrain types, and a climate typical of tropical Africa, and hence was considered to be a suitable country for the field work. Measurements of transmission loss and time delay spread were made.

Power Requirements

Due to the fact that most of the stations in a network will be powered from a solar panel which charges a battery, power consumption is an important factor. Therefore, the station hardware should be power efficient. The transmitter power amplifier will be the major factor in this respect, and hence it must dissipate as little power as is practical. For this reason, it was considered worthy of some detailed research, which formed a major part of this work.

Traditional power amplifiers fall into two main categories; linear amplifiers, class A, B and AB, are used for amplifying signal of varying envelopes, and tend to be very poor with regard to power efficiency, and non-linear amplifiers, typically class C, which are for amplifying constant envelope signals and are much more efficient. Obviously, therefore, it would be desirable to use a non-linear PA, for reasons of power efficiency. However, it is the case that the most spectrally efficient modulation schemes are those which are linear, and hence a linear PA must be used to preserve the spectrum. A method of *linearising* a non-linear PA has been investigated, which, when combined with an efficient modulation scheme, can perform well in both respects.

Chapters 4, 5, 6 and 7 detail the work performed in the area of modulation, which results in the construction of the major part of the transmitter hardware.

Chapter 4 explains the choice of modulation scheme, which is $\pi/4$ -DQPSK, and describes how this scheme is particularly suitable for the system. Filtering of the signal is considered in some detail, with regard to the choice of a transmit filter which produces a good signal spectrally, with a practical implementation technique, and which will combine with a similar filter in the receiver to minimise ISI. Chapter 5 then describes the implementation of a $\pi/4$ -DQPSK modulator using EPROM look-up techniques and Nyquist filtering. Chapter 6 details the concepts behind, and implementation of the RF stages of a linearized transmitter which is both power and spectrum efficient. Chapter 7 details the work done on the distortion analysis of the modulated power amplifier stage. Intermodulation distortion (spectral spreading) of the transmit signal produces adjacent channel interference, and hence is of great importance. By characterizing the distortion characteristics of the stage, it is possible to accurately predict the resulting spectral spreading. Conclusions of the work done are presented in chapter 8.

Chapter 2

Propagation Background and Measurement Results

2.1 Radio Propagation

In any radio system, the success rate of transmission of information from one point to another will depend very heavily upon the actual propagation of the signal between the two. Certainly, it is necessary to have some knowledge of the radio environment in which any system will operate to give the designer the information required to determine such parameters as the transmitter power needed to reduce information loss to an acceptable level, antennas necessary, and receiver specifications. There are several mechanisms by which radio waves may propagate, however these are generally both frequency and distance dependent modes. Consequently, for a given frequency of radio wave there will tend to be one dominant method for a particular propagation distance. For 'line-of-sight' radio links, such as those which are under consideration in the project with which this work is associated, frequencies in the VHF, UHF and microwave range are typically employed, the exact frequency depending upon the application and the actual segment of radio spectrum allocated by the relevant regulatory body. High capacity systems must use microwave frequencies because of signal bandwidth, whereas low capacity systems will normally operate in the VHF or low UHF areas, primarily for reasons of signal bandwidth. There are advantages to operating at the lower frequencies, however, and these include better penetration of signals into valleys, less critical antennas, and lower feeder losses

2.1.1 VHF/UHF Propagation

'Ground wave' propagation is the most important mode for line-of-sight VHF/UHF links. Ground wave propagation actually divides into two main mechanisms [53]:

a) surface wave propagation is where wave energy travels along the earth's surface, however the attenuation involved increases with frequency since the conductivity of the earth decreases, with the result that propagation at VHF and above is particularly hindered. In the VHF range (30 to 300 MHz), the surface wave is generally only of secondary importance, but it can usually be neglected completely

above 300 MHz [25].

b) space wave propagation involves the wave energy propagating through the troposphere, a region extending to approximately ten miles above the earth's surface, from transmitter to receiver. This is the primary mode for line- of-sight VHF/UHF, and when a relatively unobstructed path is used with antennas at good height the surface wave signal will typically be insignificant in comparison.

Space wave propagation can result in several signals arriving at a receiver, from a single transmitter. This occurs due to a direct wave arriving together with signals, from either ground reflections or 'reflections' due to a refraction process within the troposphere. Any combination of these signals may arrive at the receiver, depending upon the local terrain and atmospheric conditions. The effects of such multipath propagation are detailed in the later section.

Free-Space Transmission Loss

A useful starting point in the consideration of a point to point radio link is the *free-space transmission loss*; this is the loss which would occur were the path between the transmitter and receiver a straight line in a vacuum or ideal atmosphere, and unaffected by absorption or reflection resulting from any objects. Isotropic antennas are assumed, however the antenna aperture is assume to decrease directly with frequency. Quantitatively, this loss, A_o, can be determined by:

$$A_o = 20\log_{10}\frac{4\pi d}{\lambda}\text{dB} = 32.5 + 20\log_{10} f\ (\text{MHz}) + 20\log_{10} d\ (\text{km})\ \text{dB} \qquad (2.1)$$

where d is the distance between transmitter and receiver, f is the carrier frequency and λ its wavelength.

Of course, in any practical system the actual loss is likely to differ significantly from this figure, due primarily to obstructions, reflections, absorption and non-isotropic antennas. Attempts to predict exact path losses for a given radio link are

not straight-forward, however, due to the various mechanisms involved (diffraction, refraction and reflection primarily). Examination of a path profile will, however, illustrate the general propagation situation and an approximate calculation can be performed based upon equations for free-space loss, diffraction loss, ground clearance and so on.

Diffraction

General line-of-sight system design will strive to ensure that transmit and receive antennas are in the clear, and no obstructions lie between, hence providing a direct signal of suitable strength. However, in some situations this is not possible, as will often be the case where transmitters are located in villages distributed around a non-flat region, and paths may be required between them all, regardless of terrain. If the terrain between the transmitter and receiver is rough, possibly with obstructions, propagation is still possible as a result of diffraction. Diffraction is the bending of waves around objects. The bending decreases with increasing thickness of the obstruction and frequency of the radio wave. Thus, particularly in the VHF region, usable signal levels may be present at a receiver despite an obstruction in a path, such as a low hill. Additionally, it enables propagation beyond normal line-of-sight, allowing for the curvature of the earth, by diffraction around the horizon, albeit with a certain degree of signal attenuation. The two extreme cases of diffraction are those of diffraction over a smooth sphere and over a knife-edge. For a given obstruction height, the loss due to knife-edge diffraction will be considerably lower than that due to smooth sphere diffraction. The above two diffraction situations are shown in figs 2.1 and 2.2.

If the height of the obstruction is decreased until grazing incidence occurs, a loss still occurs, which in the case of perfect knife-edge diffraction is 6 dB, and for a smooth-sphere is 20 dB [46, 53]. It is not until a certain clearance is obtained that a figure approaching the free- space value is found to occur. Commonly, the amount of clearance of a radio path over obstacles is more importantly considered

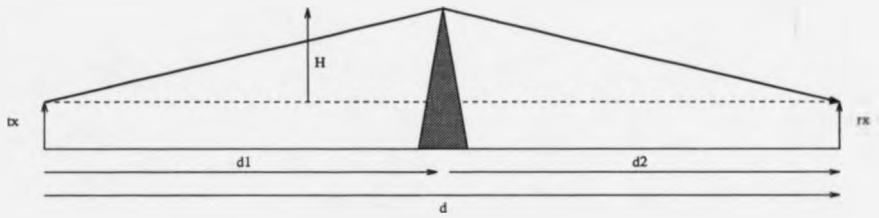


Figure 2.1: Diffraction around knife edge

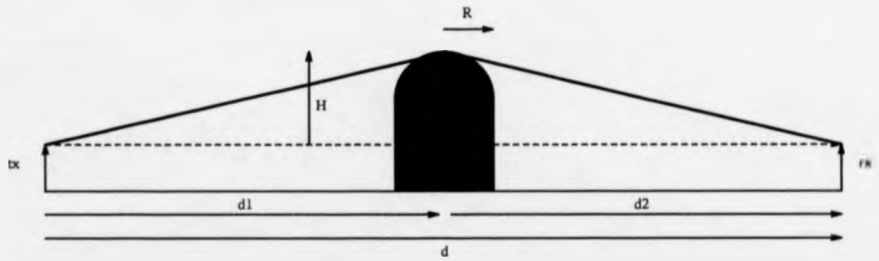


Figure 2.2: Diffraction around smooth sphere

in terms of Fresnel zones, where the nth Fresnel zone is that boundary of points from which a wave could be reflected with a path difference of n half wavelengths to the direct signal when incident at the receiver. Using this concept, it is found in practice that for minimal diffraction loss to occur over a link there should be no obstacles within about half the clearance of the first Fresnel zone [25]. On medium to long range UHF or microwave links additional clearance may be required under certain atmospheric conditions due to refraction (see next section).

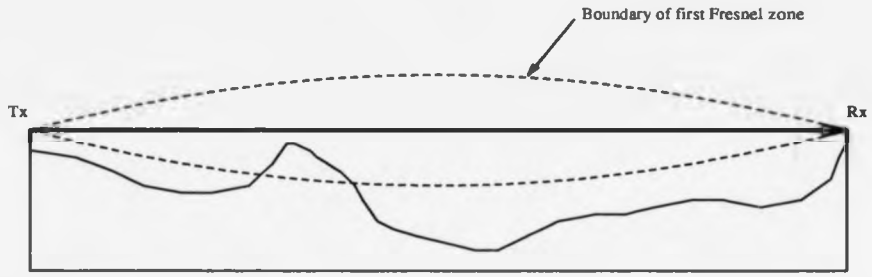


Figure 2.3: Diagram illustrating first Fresnel zone

Basic equations are given in [52] for the calculation of diffraction loss around a knife edge: as a starting point a parameter ν is defined as

$$\nu = h\sqrt{\frac{2}{\lambda}(\frac{1}{d1} + \frac{1}{d2})} \tag{2.2}$$

and then the actual loss, $L(\nu)$ is given (for $\nu \geq -.5$) by

$$L(\nu) = 6.4 + 20\log_{10}((\nu + 1)^{\frac{1}{2}} + \nu) \text{ dB} \tag{2.3}$$

where h is the height of the obstacle above the direct line between transmitter and receiver, $d1$ and $d2$ are the distances between the edge and either end of the link.

Diffraction around mountainous ridges located between a transmitter and re-

ceiver normally well beyond the mutual radio horizon can permit reliable transmission of a signal over distances much greater than those possible over flat ground, and this effect has been referred to as 'obstacle gain'. As an example, it was reported [43] that a transmission loss of only 134 dB was found on link operating over 160 miles via diffraction over an 8000 ft mountain in Alaska. Such effects are obviously of primary importance where systems must operate in mountainous regions, as the potential for both reliable transmission and interference will exist.

In hilly regions many non LOS paths will be dependent upon diffraction around more than one obstacle. In this case, the method used to calculate the total diffraction loss is not obvious, and several techniques have been proposed [44].

Of these, the two which are most widely referred to are illustrated in fig 2.4

The upper figure in the diagram shows the technique attributed to Epstein and Peterson [51]. Each obstruction in the path is treated with equal importance, and as a source/sink if the transmitter/receiver are obstructed when diffraction losses are calculated for adjacent ridges. In the diagram, the first loss occurs around the first (highest) ridge. This loss is calculated by taking the distances to the source/sink as $d1$ and $d2$, respectively, and the height of the ridge as $h1$, the difference between the peak of the ridge and what would be the line-of-sight path between transmitter and second ridge. The loss around the second peak is calculated from the values $d2$, $d3$ and $h2$. The total loss is simply the sum of these diffraction losses and the free-space loss for the path between the transmitter and receiver.

The lower diagram shows the technique which was proposed by Deygout [50]. Initially, the case of two hills is used, however this situation can be extended to include any number of hills, albeit with additional computation over the other method above. In the case of two hills, one of these is defined as the *main hill*; this is done by calculating which hill has the highest h/r figure, i.e. greatest obstruction of first Fresnel zone. The diffraction loss around the main hill is calculated using the total distances between the hill and the transmitter/receiver, and the height of the hill above the direct line between the two, as previously described for the single ridge

case. In the diagram the appropriate distance used in the calculation are $d1$, $d2+d3$ and $h1$. The loss around the other hill is then calculated by taking the distances between adjacent main hill and transmitter or receiver, assuming adjacency to one or other. Therefore, the values used here are $d2$, $d3$ and $h2$. If there are more than two obstructions, then the matter is made slightly more complex by the systematic repetition of the process used for two hills. That is, further groups of hills are split into normal/main importance. The process repeats until there are no two adjacent 'normal' hills.

This technique therefore differs from the previous one in that importance is attached to hills in a hierarchical manner. The actual results obtained are often very similar for the two, but can occasionally differ by up to around 10 dB [44] for certain types of profile. Since the diffraction loss figure for the main hill in the Deygout method is always greater than that which would be calculated in the first method, the former always predicts a greater path loss. Over the majority of paths, the small difference in the results (typicall less than 5dB) of the two techniques do not justify the extra calculation involved in Deygout's technique, and the Epstein-Peterson method is sufficient.

Obviously, the assumption that ideal knife-edge diffraction occurs is inaccurate for many cases, and, additionally, account is rarely taken of the transversal profile of the terrain at the diffracting edge, which can give rise to significant errors [52]. Accurate prediction of ground reflections is very difficult [46], and this adds further potential errors to the procedure. Therefore, the accuracy of path loss calculations is often only an approximate procedure, and, in general, an error in the order of 5 dB is to be expected, but may be considered sufficiently accurate for almost all purposes.

Refraction

Refraction, mentioned previously, also produces bending of the radio wave due to refractive index variations in the atmosphere. An important consequence of this,

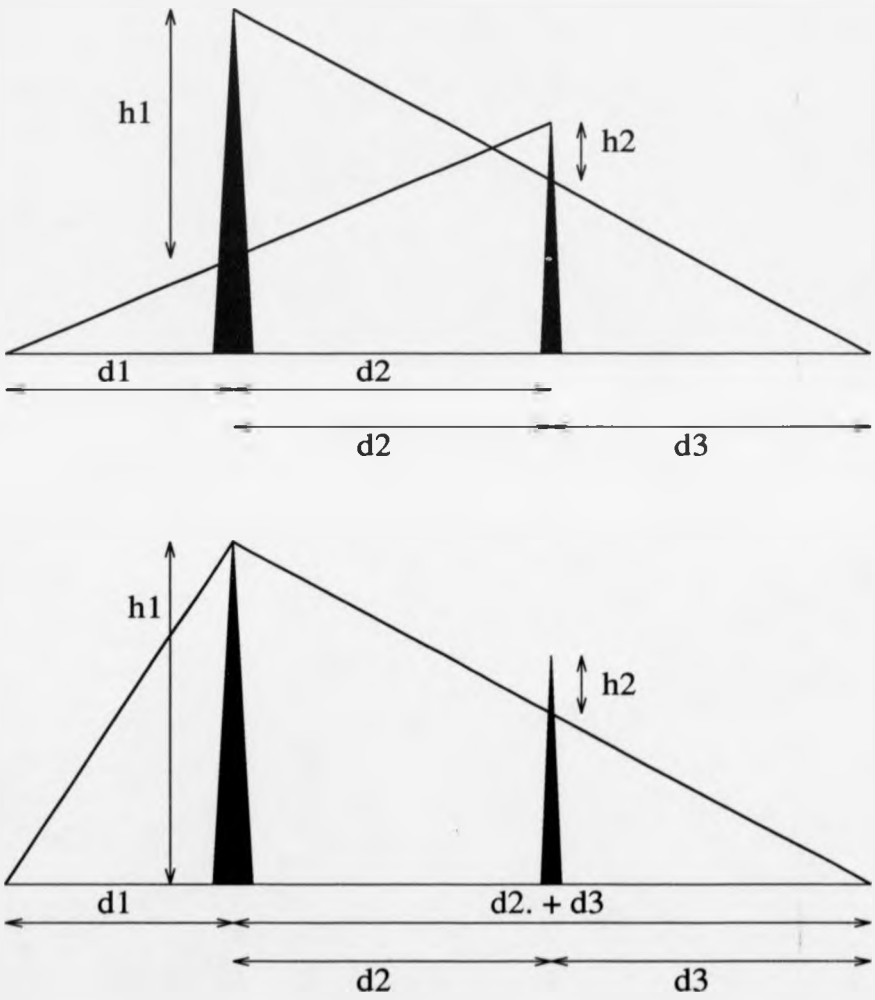


Figure 2.4: Methods of calculating loss due to multiple knife edge diffraction

under normal atmospheric conditions, is that the radio horizon is further than the physical horizon. This results form the fact that the refractive index of the atmosphere generally decreases with altitude, causing velocity of radio transmission to increase with height, and therefore a curving towards the earth of the radio signal, hence extending the apparent horizon. If the change in refractive index is linear with height, then the result is that the radio signal apparently travels in a straight line over an earth with modified radius

$$ka = \frac{1}{\frac{1}{a} + \frac{\mathrm{d}n}{\mathrm{d}h}} \tag{2.4}$$

where a is the true radius of the earth, and $\mathrm{d}n/\mathrm{d}h$ is the rate of change of refractive index with height.

A common value accepted for k is 4/3, which results from a decrease of refractive index with height of approximately $3.9 * 10^{-8}$ per metre.

However, as $\mathrm{d}n/\mathrm{d}h$ varies so the effect on the propagation of the radio signal can vary significantly. Under certain conditions the refractive index may actually increase with height (for a reasonable height), with the result that the signal will bend away from the earth. Conversely, as the rate of change increases above the normal value, so the effective bending of the signal back to earth will increase. For a factor of four increase, the earth appears flat since propagation will occur parallel to the earths surface. Above this rate, signals may be bent back down to earth to be reflected form the surface, and subsequently bent back down and reflected again, with the process continuing indefinitely while the atmospheric conditions remain similar. This effect is known as ducting, as the signal propagates in a *duct* between the earths surface and the upper level of the radio path. Ducting can cause propagation of signals over distances many times those encountered under normal conditions, with obvious possibilities for interference to other users of the frequency employed. An *elevated duct* occurs when the signal is effectively trapped within a layer of air of increasing refractive index with height, which lies between layers of

decreasing refractive index with height [53].

Fig 2.5 shows the ray profiles for several values of k. It should be noted that these are drawn above a 'flat' earth: had they been drawn above a true earth profile, then it would be noted that the apparent radio horizon is increased for k=4/3 and 2/3 cases, while further decreased for the k=-1/2 case.

Knowledge of variations of refractive index with height over a given radio path will give a good indication of the problems which may be encountered as a result of refraction process in the troposphere. In many areas of the developed world such data has been recorded over many years and therefore can aid system design considerably. In the developing world, this is not often the case, and due to the lack of radio transmission systems in operation there is still a great lack of information on typical atmospheric conditions to be encountered in the troposphere. Since in many cases the atmospheric conditions are likely to be very different to those found in the areas where data is readily available, it is quite possible that propagation will be substantially different at times.

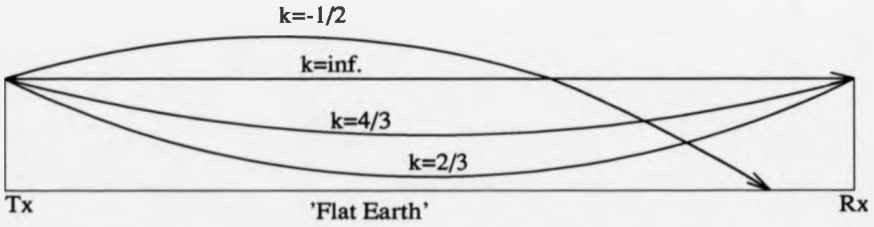


Figure 2.5: Diagram illustrating ray profiles for several values of k

2.1.2 Multipath Propagation

Propagation of a transmitted signal by more than one path to a receiver - *multipath propagation*, results in interference at the receiver. The relative phase of the incoming signals at the receiver will obviously depend upon the modes of propa-

gation from the transmitter. In addition to the phase differences resulting from different path lengths, phase shifts also occur upon reflection of a radio wave from a surface. On line-of-sight links attempts to predict the phase shifts are possible in some cases. For low grazing angles, values of +1 and -1 are often assumed for the reflection coefficients of vertically and horizontally polarized waves incident upon flat earth, respectively . This is not the case for the mobile/portable environments, particularly in urban areas where propagation generally occurs by scattering around obstacles, as the channel is assumed to produce a number of signals at the receiver with random phase.

The difference in propagation mechanisms between the fixed-link situation and the mobile/portable environment leads to different statistical models for these cases. The latter is generally modelled as a Rayleigh distribution [49], while with the former, a Rician distribution is assumed (rather than the received signal consisting of numerous independent scattered components of random phase and amplitude, it is assumed to contain one dominant fixed component and a second component which has a Rayleigh distribution) [54].

Vectorial addition of the signals will produce interference, either constructive or destructive; in the worst case two signals arriving at a receiver may cancel each other out to such a degree that signal to noise ratio is degraded below a workable level. If the signals are in phase, then constructive interference occurs and the received signal level will be enhanced. Changing the frequency of the signal used will alter the interference effects, since the relative phases of incident signals will change. Sweeping across a wide frequency band it would therefore be observed that the received signal level increases and decreases as the interference changes between destructive and constructive a number of times. This effect is known as *frequency selective fading*[26]. On high-capacity links, the effect is likely to produce a null in part of the channel, while in low-capacity links the whole channel is likely to be affected in a similar manner. In the mobile environment the channel is continually changing, with the effect that rapid fading will occur. In the static situation, however, time-variant

fading on a channel only occurs as a result of changing atmospheric conditions, ignoring the effects produced by moving reflecting objects, although it is well known that aircraft act as good reflectors, and as such could produce time-variant effects. This type of fading, as stated earlier, increases with propagation distance and with frequency, so that on long distance microwave links methods of countering such effects are commonly employed. On short range (less than 20 miles) VHF and UHF links, channels are unlikely to vary greatly with time for all but the most unusual atmospheric states. For medium range links in the same frequency range, slow fading may occur as the atmospheric conditions change.

In digital transmission, the difference in propagation time for the respective paths may cause problems with *inter- symbol interference* (ISI) and timing recovery, in addition to fading of the signal envelope. If ISI prevails, then an irreducible bit error rate (BER) may result since increasing the transmitter power will not affect the error rate. Severe ISI may render a path unusable if remedial measures are not taken, such as the use of equalization. The ISI effect will be dependent upon the symbol rate employed in transmission, low rate transmission systems obviously suffering less from ISI than high rate: a measure of ISI therefore can provide a useful guideline for the maximum channel capacity allowed.

Modelling multipath propagation

In the most simple form, the signal incident at a receiver $s_r(t)$ due to a transmitted signal $s(t)$ can be expressed as

$$s_r(t+\tau) = A(s(t) + \sum_{m=1}^{M} \Gamma_m s(t-\tau_m)) \tag{2.5}$$

where M is the number of paths in addition to the direct path, A is an attenuation factor, Γ_m is the relative attenuation of the mth path relative to the direct path, τ is the propagation time between transmitter and receiver, and τ_m is the relative delay of the mth path relative to the direct path. Alternatively, the channel frequency

response may be expressed for an M-ray model as

$$H(\omega) = H_o + \sum_{m=1}^{M} a_m e^{-j\omega\tau_m} \tag{2.6}$$

where H_o is the reference, and a_m and τ_m are the amplitude and delay of the mth path relative to the reference, respectively.

A typical situation in a hilly region is shown in fig. 2.6, where the received signal consists of the direct signal and a number of reflected signals.

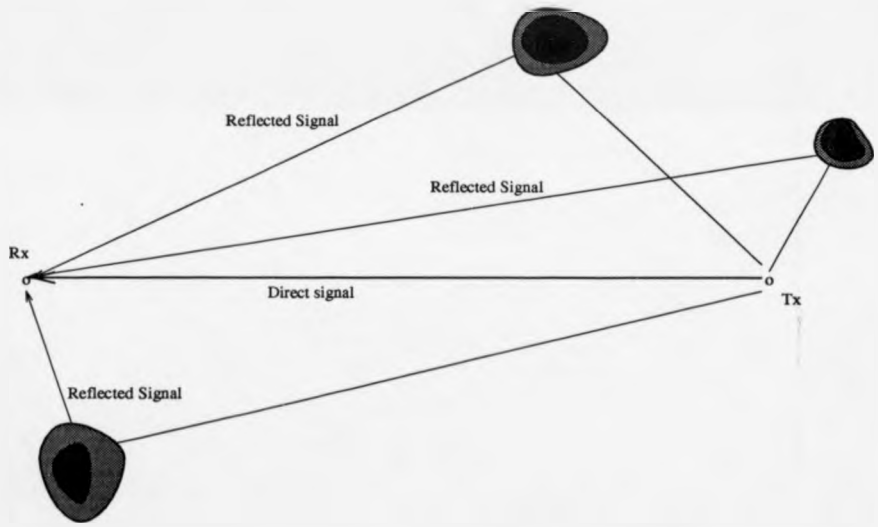


Figure 2.6: Diagram illustrating multipath propagation in a hilly region

In line-of-sight propagation with high antennas, it will often be the case that the only two signals arriving at the receiver are the direct signal and the signal resulting from a ground reflection between transmitter and receiver. In this case the actual signal level at the receiver will oscillate in a regular manner about a median level as the antenna height changes, due to the interference of the two paths changing from destructive to constructive, and vice versa. For non-line-of-sight paths, or

paths where a ground reflection is insignificant, the above effect is unlikely to be important.

2.1.3 Characterizing multipath channels

There are several methods by which a multipath channel may be characterized, the most popular being the following:

a) Transmission of a very narrow pulse. The time spreading effect is then observed directly at the receiver, as amplitude fluctuations with time. The impulse response of the channel can therefore be observed directly, assuming a pulse of an ideal impulse form. The technique was used by Schmid [22], in modelling a UHF channel over irregular terrain. The main disadvantage with this method is the need for very high power pulses, to provide sufficient signal level at the receiver.

b) Correlation system employing *pseudorandom binary sequences.* This technique also provides a means of measuring the impulse response of a channel due to the properties of the autocorrelation function of pseudorandom binary sequences, as shown in fig. 2.7, and explained in [18, 19, 27, 29, 36]. For a bipolar binary sequence of amplitude +/- a, the autocorrelation function is a series of triangular impulses of amplitude a^2, and base-width $2t_o$. Between impulses the response is flat, and of magnitude -a^2/L, where L is the sequence length. By transmitting a carrier modulated by such a sequence, and in turn multiplying the received signal with an identical sequence, clocked at a slightly different rate, the channel impulse response is effectively produced, as time smearing of the transmitted signal will cause correlations at the receiver at times other than those due to a single sequence only. By measuring the time of these 'images' relative to the main impulse, the difference in path lengths can be calculated, and measurement of the impulse magnitude gives the relative amplitude of each signal incident at the receiver. As an example, if the channel shown in figure 2.6 were measured this way, then the result would be of similar form to that shown in figure 2.8. If quadrature detection is employed and

the correlation performed on both the I and Q channels, then a complex impulse response can be determined which will contain the additional information of relative phases. In a practical measurement system, some filtering of the PRBS may well be required, and this will tend to 'round off' the triangular impulse form.

Figure 2.7: Autocorrelation of pseudo random binary sequence

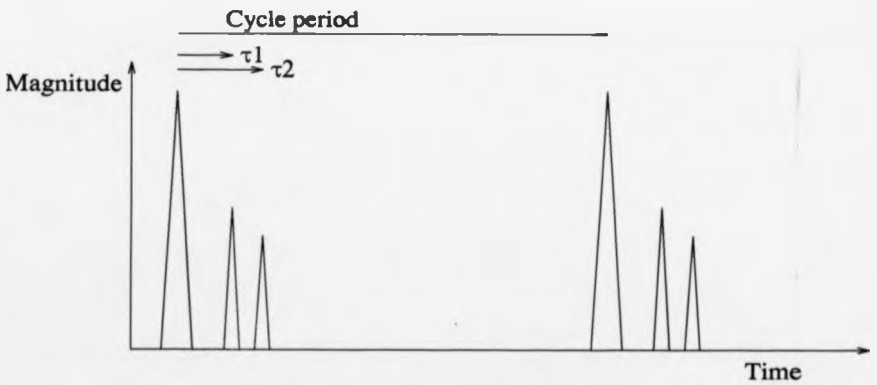


Figure 2.8: Channel measurement output for a multipath channel

Practically, this method of channel characterization has the advantage of being a reasonably simple technique to implement. On the negative side, the bandwidth of the signal is large (actually the same as that in (a)), although filtering can be employed at either baseband or RF levels to reduce this. This method was chosen

to be employed in the measurements which were to be taken as part of this research, due to its practicality. The measuring equipment is described in Chapter 3, together with a more thorough treatment of the theory of the technique.

2.1.4 Propagation Measurements in Sierra Leone

Due to the lack of information on propagation in the areas of the developing world which differ markedly from those in the developed world, it was decided that a series of measurements on typical channels would be useful as part of the development of the digital radio network. Obviously, it would be a huge task to attempt to perform such measurements in all the representative areas for which the system is aimed, however very useful results would be obtained by considering an area which has much in common with a large proportion of such areas. Since Africa is considered as the main area which would usefully be served by the network, it was decided that field studies should be carried out there. Sierra Leone was deemed the most appropriate choice for the reasons below:

- Climate typical of tropical Africa
- Wide range of terrain types
- Country in which the radio network would obviously be most useful
- Good links with relevant authorities

Measurements were made in the following areas of Sierra Leone:

a) Freetown Peninsula. This is a hilly peninsula with mountains rising to around 1000 metres. The city of Freetown, and all habitation, lie on a thin coastal strip. The transmitter was installed at the western edge of the city, at Aberdeen Point. The receiver was taken into the eastern edge of the city, with delay spread measurements being made at several locations along the way. Transmission loss measurements were not made.

b) Bonthe district. The area around Bonthe, in the south west of the country is low-lying and very flat. Much of the area is coastal swamp. Measurements were made between Bonthe and Matru, a distance of 37 km., and some transmission loss measurements were made on a boat when travelling between the two, with the transmitter located at Matru.

c) Kenema district. This region contains a range of terrain types, from very flat land to hilly areas vegetated with thick forest. Measurements were made to the north of the town of Kenema, moving at first along the side of a ridge, and then to locations behind the hills. The transmitter was installed at the main post office in Kenema, with the antenna being at a height of approximately 23 metres.

d) Kabala district. The region around Kabala is, generally, very hilly, particularly to the east and west. The transmitter was set up at the police station in the town, and measurements made along the roads to the north, east and south-east. The antenna was mounted on the police radio mast at a height of 20 metres. Several path profiles are shown for the Kabala region, in figs 2.9 and 2.10.

Where measurements were with a mobile receiver, the receive antenna height was approximately 9 metres, in all cases.

2.2 Results

The results of the propagation work are now given. Firstly, the path loss results are considered, followed by the results of the multipath characterization. Errors in the measured path losses are assumed to be a maximum of 3dB. These errors can result from distortion in the antenna radiation patterns and feeder cable losses, for example.

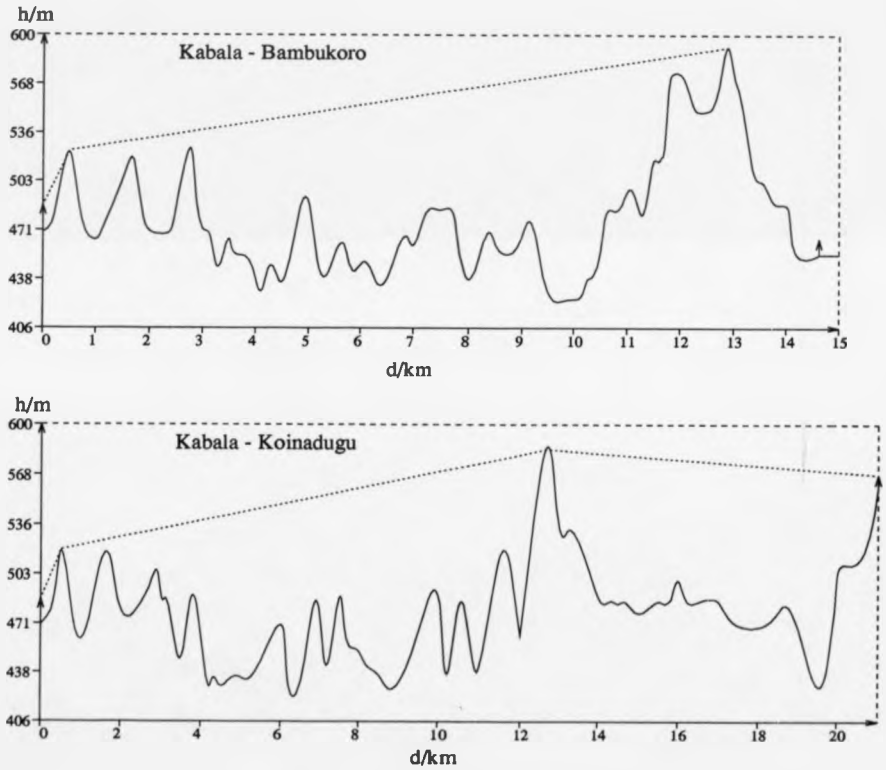


Figure 2.9: Path profiles from Kabala

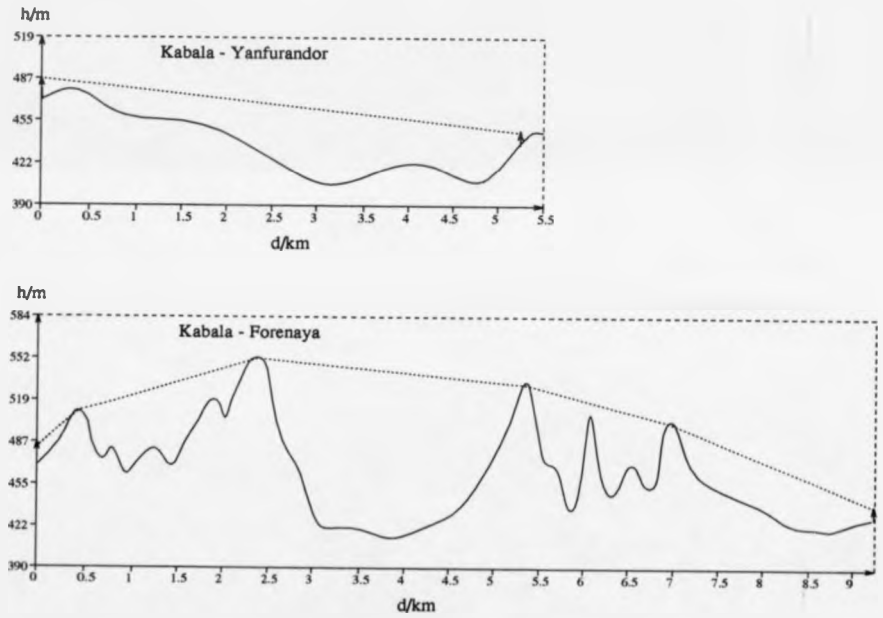


Figure 2.10: Path profiles from Kabala

2.2.1 Transmission Loss Results

An example illustrating the procedure used in the theoretical calculation of transmission loss is given, using a path profile produced for a radio link over which a measurement was made. The path was between Kabala and Bambukoro, a distance of 14.6 km. The profile shows two main obstructions - ridges located near the ends of the link. In addition there are several further ridges which approach the path sufficiently to encroach upon the first Fresnel zone, and hence add further loss to the path. These are shown in fig 2.11. The Epstein-Peterson technique is used.

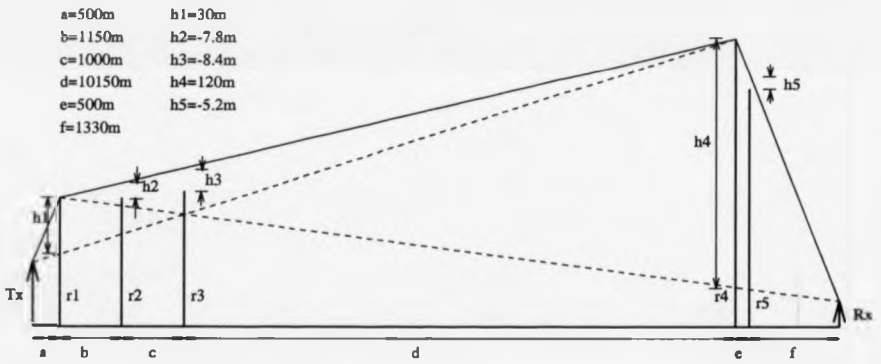


Figure 2.11: Path profile from Kabala to Bambukoro, showing main ridges

Ridge 1 (r1). $d1$=500m, $d2$=12300m, h=30m. The loss around the ridge can be calculated by

$$\begin{aligned} \nu &= h\sqrt{\frac{2}{\lambda}(\frac{1}{d1}+\frac{1}{d2})} \\ &= 30\sqrt{\frac{2}{5.66}(\frac{1}{500}+\frac{1}{12300})} \\ &= 0.814 \end{aligned} \tag{2.7}$$

and, from this, the loss, $L(\nu)$, is given by

$$
\begin{aligned}
L(\nu) &= 6.4 + 20\log_{10}((\nu+1)^{\frac{1}{2}} + \nu)\text{dB} \\
&= 6.4 + 20\log_{10}(\sqrt{1.814} + 0.814) \\
&= 13.1\text{dB}
\end{aligned}
\tag{2.8}
$$

For ridge 2, $d1$=1150m, $d2$=11150m and h=-7.8m, giving ν = -0.14, and $L(\nu)$=4.3dB. For ridge 3, $d1$=2150m, $d2$=10150m and h=-8.4m, giving $L(\nu)$ = 4.7dB. For ridge 4, $d1$=12300m, $d2$=1830m and h=120m, giving $L(\nu)$ = 17.2dB. For ridge 5, $d1$=500m, $d2$=1330m and h=5.2m, giving $L(\nu)$ = 3.9dB.

The free-space loss for a distance of 14.6 km is 90.2dB, and the sum of the above diffraction losses is 43.2dB. The estimated path loss for this link would therefore be 133.4dB. Actually, the loss was measured as 133dB. Since the technique is only approximate, such close agreement would generally not be expected. Perfect knife-edge diffraction was assumed, and no account was taken of the lateral profile of the diffracting edges, which must lead to some errors. In addition, there will be errors in the assumed path profile, due to inaccuracies in transmitter/receiver locations, and perhaps even in the maps used to draw the profiles. No account was taken of ground reflections, and these could contribute to the overall figure, but would be very difficult to predict.

The above procedure was used to compare predicted path losses with measured path losses over a number of paths in the Kenema and Kabala regions. These are shown in the table of fig. 2.12. Where necessary terrain data was not available, the measured path loss is simply given.

2.2.2 Multipath Results

This section presents some of the delay spread results obtained. Although it is possible to use the data to generate figures such as rms delay spread, it is felt

Path	Path length (km)	No. of points of diffraction	Free-space loss (dB)	Predicted diffraction loss (dB)	Total predicted loss (dB)	Measured loss (dB)	Remarks
Kabala - Forenaya	9.2	4	86	38	124	125	Close agreement
Kabala - Sengbebambu	16.3	3	91.2	30.3	121.5	136	Poor agreement
Kabala - Koinadugu	21.0	3	93.4	26.5	119.9	119	Close agreement
Kabala - Yanffurandor	5.4	1	81.6	5.1	86.7	104	Probable siting error
Kabala - Sonkonya	3.5	1	77	16	93	76	Probable siting error
Kabala - Haffia	7.5	3	84	25.7	109.7	115	Close agreement
Kabala - Dogoloya	13.8	3	89.7	33.2	122.9	116	Close agreement
Kabala - Kamasiri	16.0	4	91.0	37.7	128.7	136	Fair agreement
Kabala - Bendugukura	-	-	-	-	-	128	-
Kabala - Bendukoro	-	-	-	-	-	137	-
Kabala - Kondembaia	-	-	-	-	-	139	-
Kenema - New Kambona	29.3	2	96.3	34.1	130.4	125	Close agreement
Kenema - Yangchun/Penguma	33.8	3	97.5	45.9	143.4	133	Fair agreement
Kenema - Panderu	-	-	-	-	-	107	-
Kenema - Lago	-	-	-	-	-	126	-
Kenema - Penguma	-	-	-	-	-	139	-
Kenema - Bambawo	-	-	-	-	-	124	-
Matru - Bonthe	37.0	-	98.4	-	-	117	Link beyond l.o.s.

Figure 2.12: Table showing measured and predicted transmission losses

that such figures can depend too greatly on the measurement interpretation and the receiving system performance to be treated with any great importance. This is due to the thresholding required to differentiate between noise and actual delays: as the threshold is varied, so the rms delay spread value will vary strongly, from zero where nothing is above the threshold value to a very large value where the threshold lies below the noise floor. The delay spread profiles themselves contain all the information necessary for any reasonable path analysis and for the prediction of potential problems when practical radio signals are transmitted over the channel measured. The profiles contain only amplitude information, as the equipment used was not capable of measuring complex profiles.

In chapter 3, the equipment used is described, and consideration given to its limitations. Here, it suffices to say that in some cases the profile data recorded contained too high a noise level to provide useful information on the channel, and consequently such results are not presented. However, there were still numerous paths where data was successfully recorded, and the results obtained on a number of these are shown. Some post-processing, in the form of profile averaging, has been employed. The technique used is described in the following chapter.

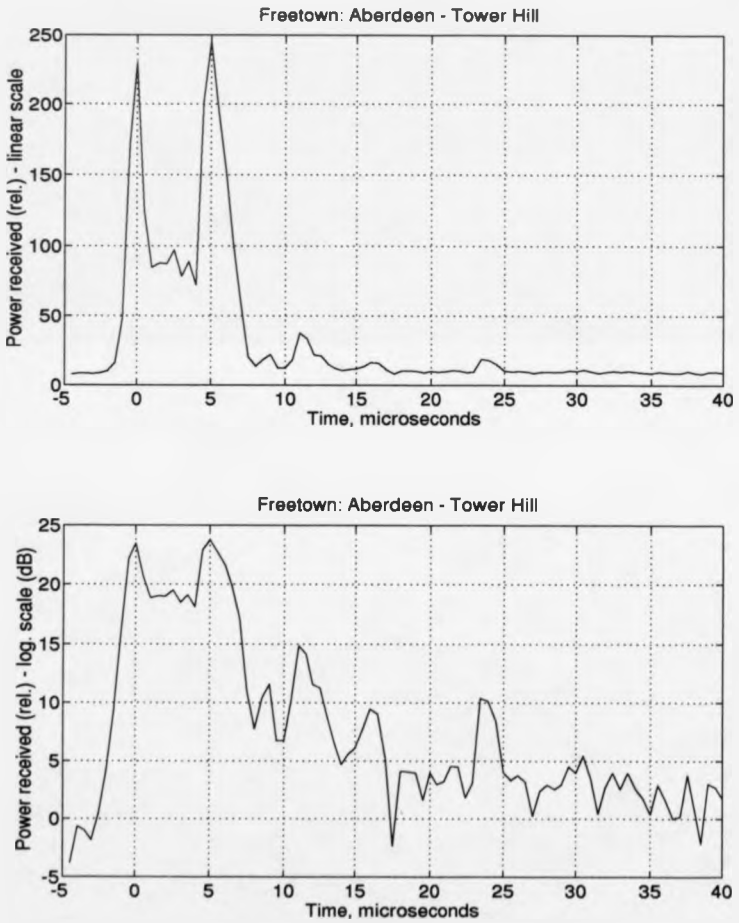


Figure 2.13: Delay profile: Aberdeen Pt. - Tower Hill, Freetown, averaged over 16 measured profiles

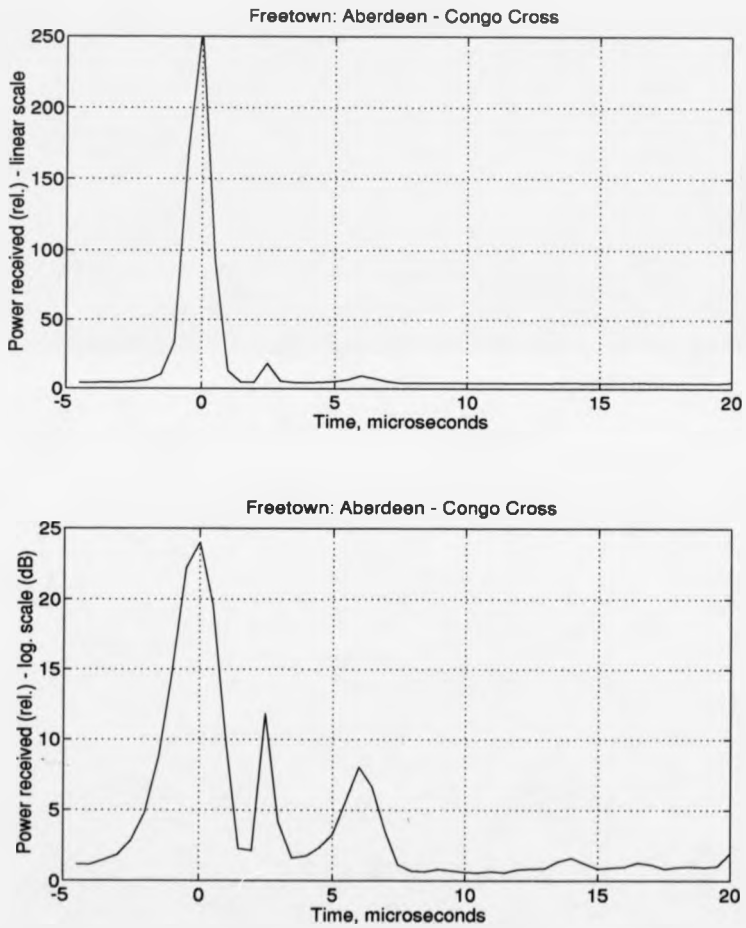


Figure 2.14: Delay profile: Aberdeen Pt. - Congo Cross, Freetown, averaged over 11 measured profiles

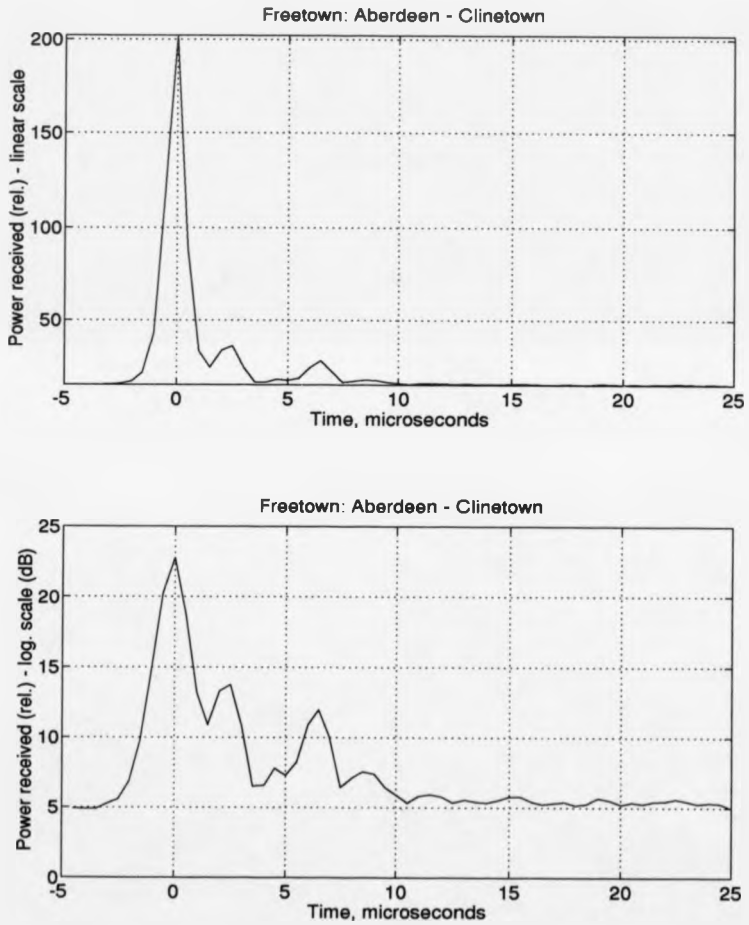


Figure 2.15: Delay profile: Aberdeen Pt. - Clinetown, Freetown, averaged over 38 measured profiles

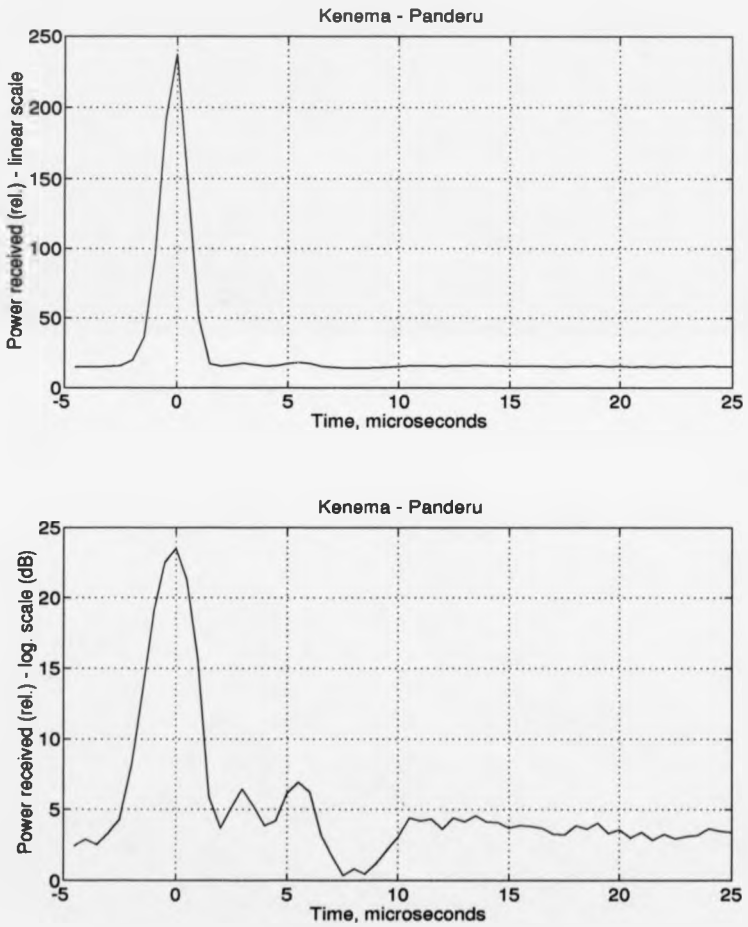


Figure 2.16: Delay profile: Kenema - Panderu, averaged over 62 measured profiles

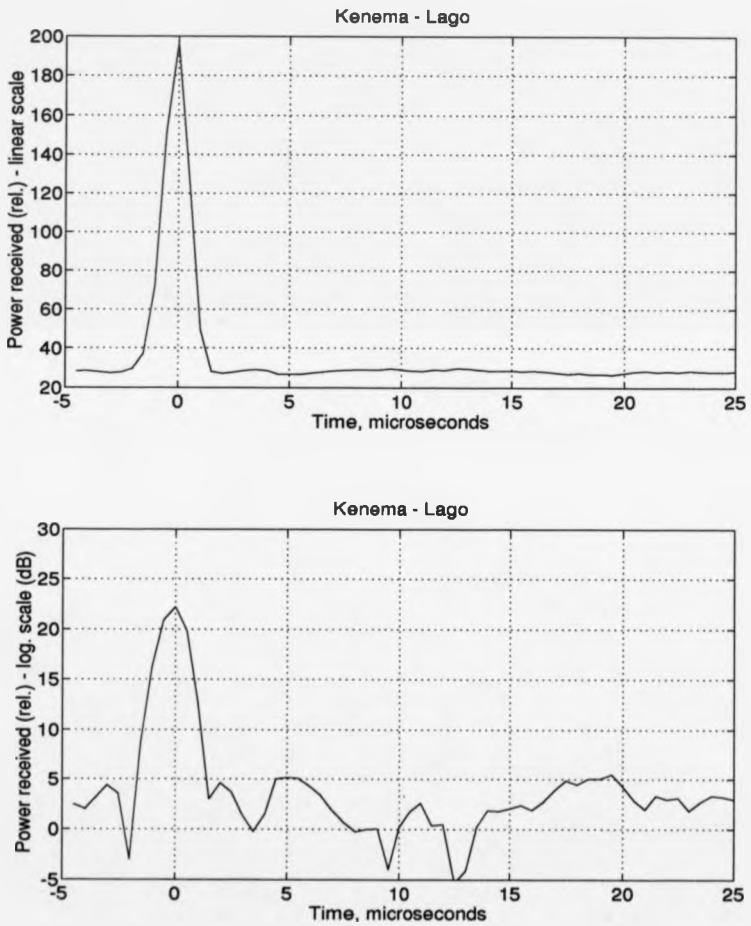


Figure 2.17: Delay profile: Kenema - Lago, averaged over 86 measured profiles

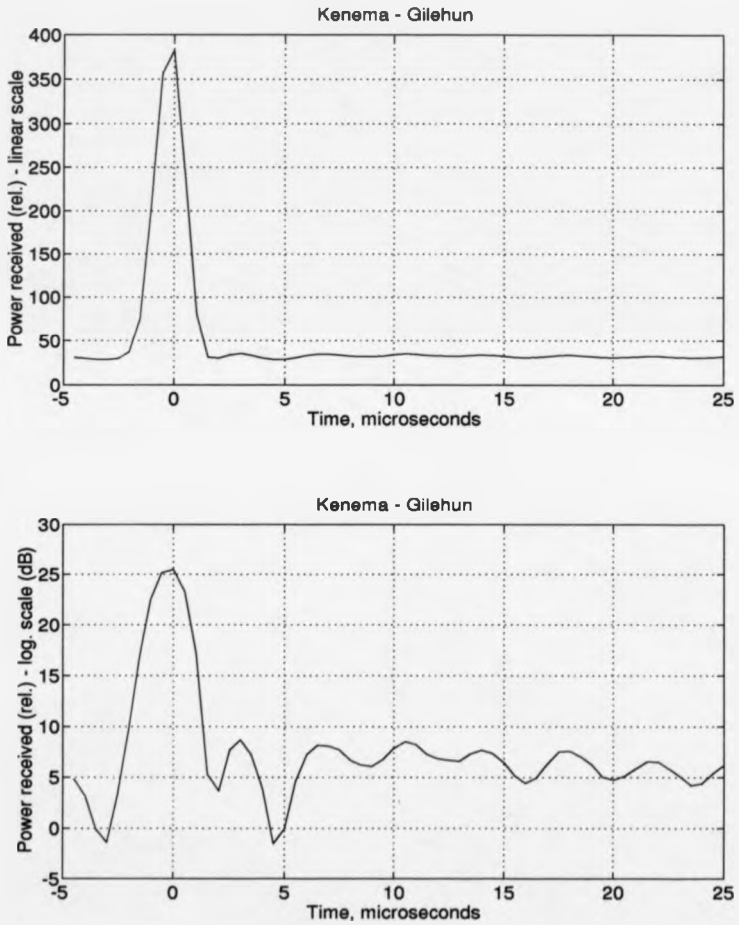


Figure 2.18: Delay profile: Kenema - Gilehun, averaged over 22 measured profiles

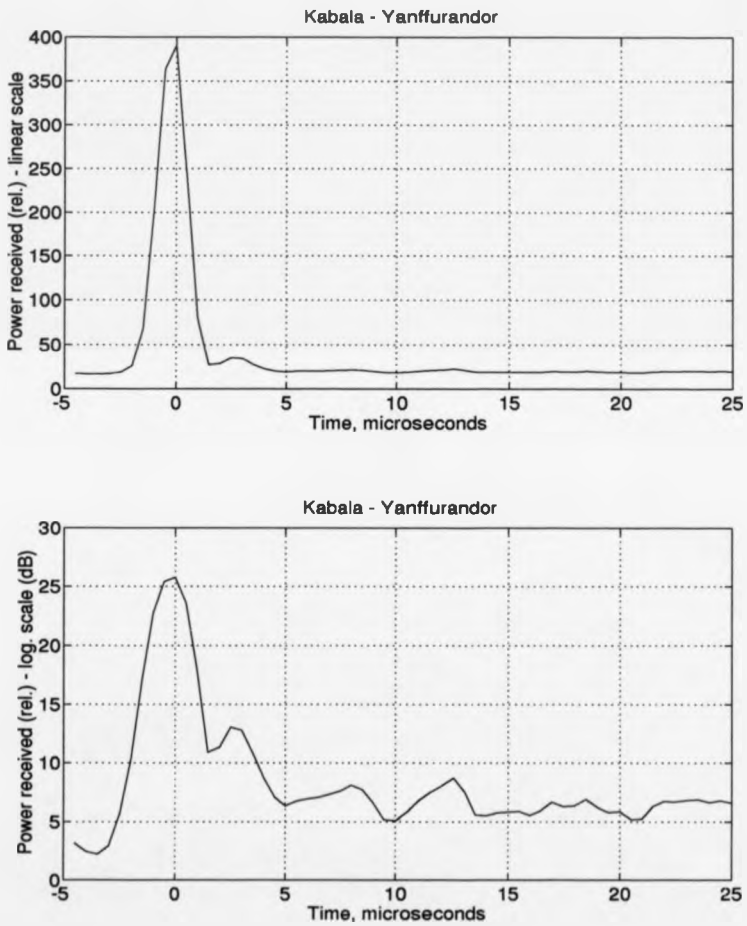


Figure 2.19: Delay profile: Kabala - Yanffurandor, averaged over 22 measured profiles

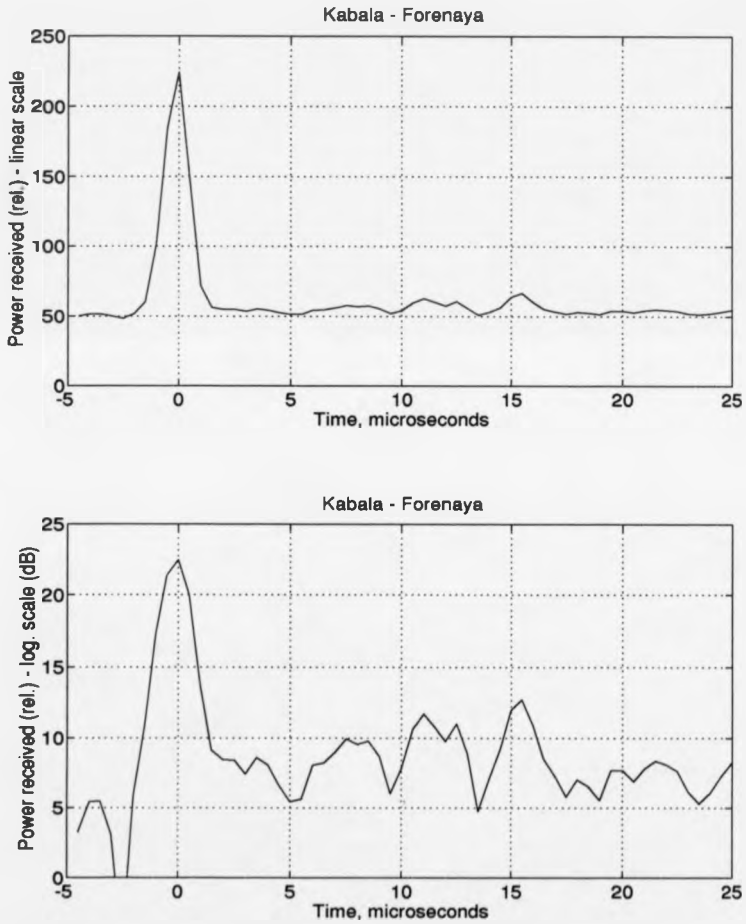


Figure 2.20: Delay profile: Kabala - Forenaya, averaged over 59 measured profiles

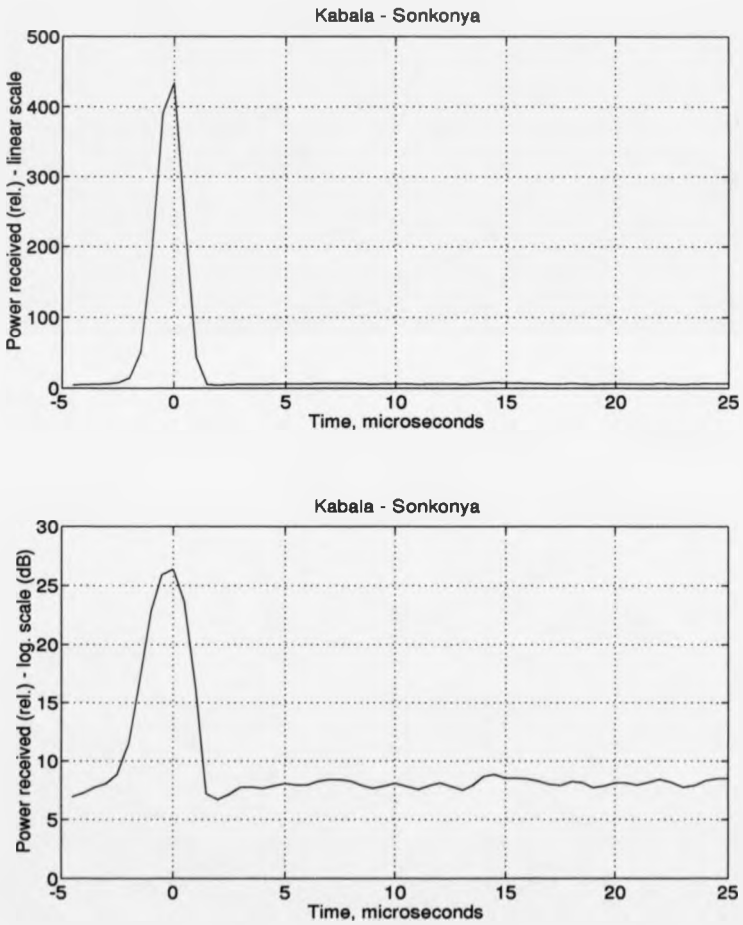


Figure 2.21: Delay profile: Kabala - Sonkonya, averaged over 3 measured profiles

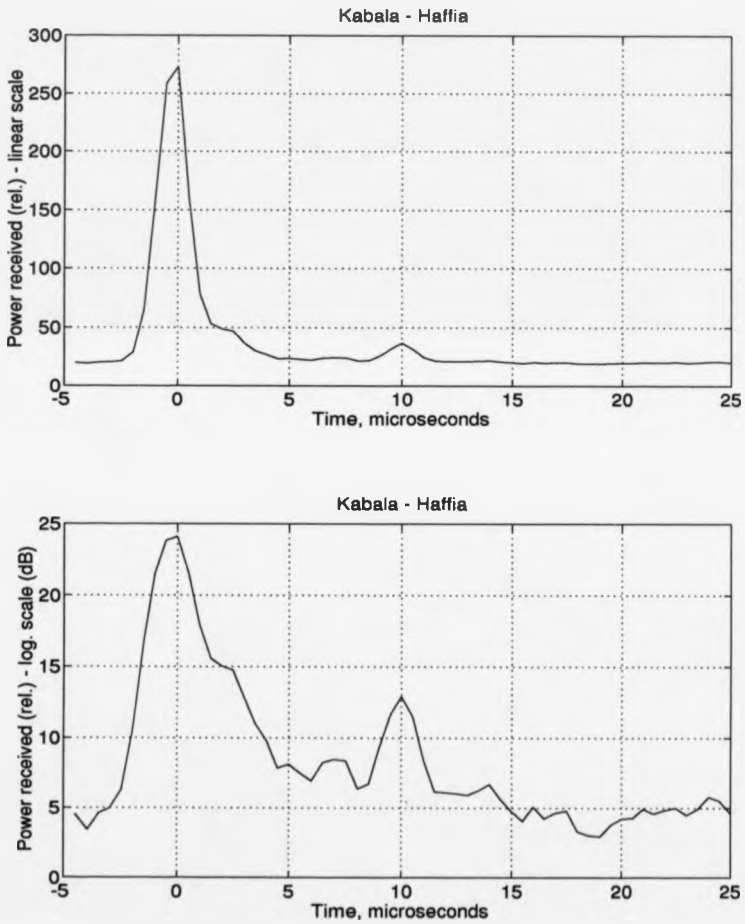


Figure 2.22: Delay profile: Kabala - Haffia, averaged over 26 measured profiles

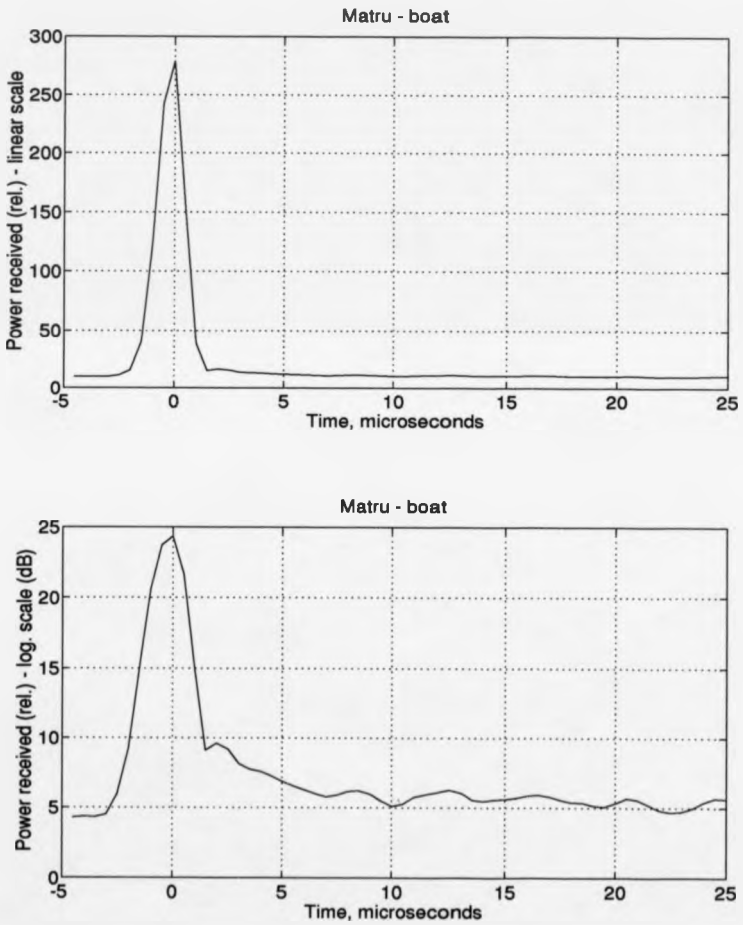


Figure 2.23: Delay profile: Matru - 'Boat', averaged over 26 measured profiles

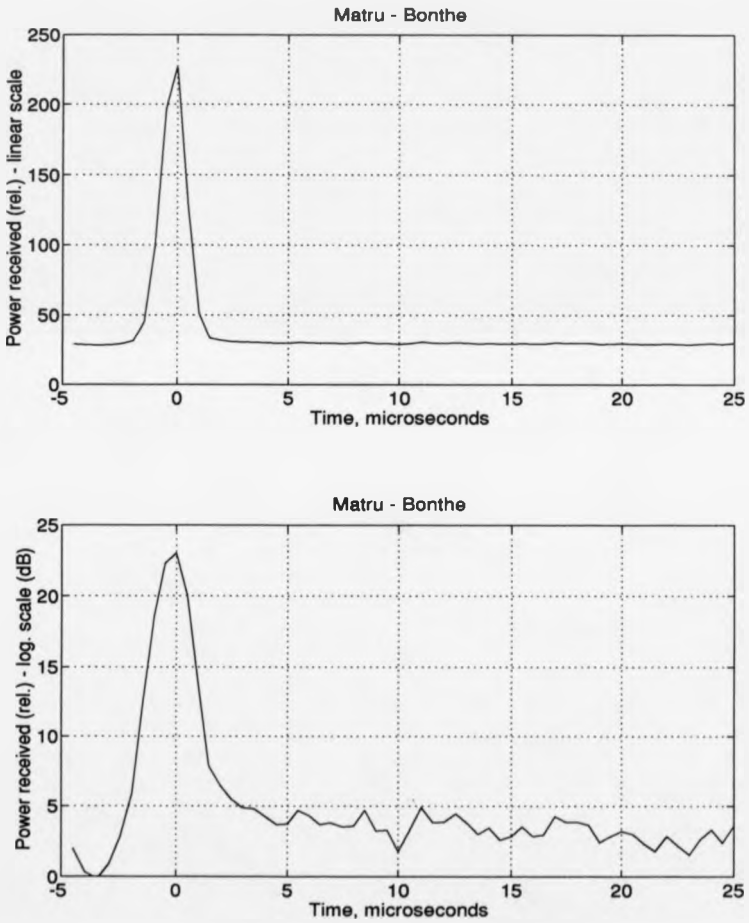


Figure 2.24: Delay profile: Matru - Bonthe, averaged over 99 measured profiles

Chapter 3

Propagation Measurement Hardware

3.1 Introduction

This chapter describes in detail the principle of the delay spread measurement technique, and the actual practical implementation of the measurement hardware.

The multipath propagation measurement technique used, described briefly in chapter 2, is based on the work of Cox [19]. In his paper, Cox describes measurement of both delay and doppler profiles of paths in a suburban environment, with obvious emphasis on a mobile radio situation, therefore. When considering a system of fixed stations, the propagation environment is significantly less harsh, and doppler measurement is almost totally redundant, for example, which reduces the hardware requirements of the measurement equipment somewhat.

3.2 Multipath Characterization using Pseudo-Random Binary Sequences

The characterization technique works around the autocorrelation properties of pseudo-random binary sequences (PRBSs), (or, $m - sequences$). As stated in chapter 2, the autocorrelation function of an unfiltered PRBS (NRZ square wave) is a series of triangular impulses.

The base width, w, of the impulses is given by

$$w = 2t_o \tag{3.1}$$

where t_o is $1/f_c$, f_c being the clock (chip) rate, and the impulses are spaced by T, where T is given by

$$T = \frac{L}{f_c} \tag{3.2}$$

and L is the sequence length.

Pseudo-random binary sequences are relatively simple to generate, using a shift-

register with feedback [55]. The length of the sequences are given by 2^{N-1}, where N is the number bits in the shift register. By *exclusive or-ing* certain (known) outputs and feeding the result back to the shift register input, the output of the shift register will be the PRBS required, in continuous repetition.

The autocorrelation, when channel characterization is concerned, may be achieved by modulating a carrier in a transmitter, correlating (mixing) the received version of the signal with a similarly modulated receiver oscillator signal, and integrating (low-pass filtering) the result. By clocking the transmit and receive sequences at slightly different rates, however, a time scaling factor is introduced. This factor, typically denoted k, is given by

$$k = \frac{f_c}{\delta f_c} \tag{3.3}$$

where δf_c is the difference between transmitter and receiver clock rates. This then produces an output of triangular impulses of base width $2kt_o$, spaced by kT. Due to the time scaling factor, the bandwidth of the signal at the correlator output is greatly reduced from that transmitted, and can be contained within the audio spectrum, for example. This permits relatively simple recording of the received data, which is particularly useful where post-processing of the data is desired.

Over a practical radio channel, multipath components may be present, and these will produce additional correlations. The correlator output will then be of the form of that shown in chapter 2.

From the delay profile produced by the correlation process it is possible to define some important parameters, which can be used to describe the radio channel. Of these, the *mean excess delay*, $\overline{\tau}$, and the *root mean square (rms) delay spread*, σ, are the most important.

Mathematically, the mean excess delay can be expressed as the first moment of the channel delay profile, and the rms delay spread as the second central moment of the profile [18, 36]:

$$\overline{\tau} = \frac{\sum_k \alpha_k^2 \tau_k}{\sum_k \alpha_k^2} \tag{3.4}$$

and

$$\sigma = \sqrt{\overline{\tau^2} - \overline{\tau}^2} \tag{3.5}$$

where

$$\overline{\tau^2} = \frac{\sum_k \alpha_k^2 \tau_k^2}{\sum_k \alpha_k^2} \tag{3.6}$$

In the above, k is the number of delayed impulses after the first detected signal, τ_k is the delay of the k -th impulse and α_k its magnitude.

The above equations permit quantitative descriptions of a given channel. One problem, however, in the use of the above, is the fact that a threshold must be applied to the received delay profile to distinguish between actual delay components and the receiver noise. A low threshold would give rise to higher values of rms delay spread than a high threshold, as in the latter case some weak signals will be neglected. Since there is no standardization around a given threshold value/noise floor in published work, rms delay spread figures can vary with instrumentation and interpretation. In [36] it is estimated that rms delay spreads can differ by almost an order of magnitude for different practical threshold levels, and this must be taken into account when comparing published results.

3.3 Measurement Equipment Implementation

Much of the equipment used in the propagation measurement work was designed and built as part of the work. This is now described.

3.3.1 Transmitter Hardware

The transmitter which was used is shown in block diagram form in fig. 3.1. A PRBS generator, clocked at 1 MHz, produces the required sequence for the measurements. Its output is buffered by a complementary emitter-follower pair to produce a low-impedance source of appropriate magnitude for the 50Ω low-pass filter, FL1, and subsequent mixer, M1. At the mixer, the output of the 53 MHz crystal oscillator is modulated by the filtered PRBS. The output is then amplified by several stages. Q2 and Q3 are broadband amplifiers of 50Ω input/output impedance, giving around 25 dB gain. Q4 further amplifies the signal to drive the PA, Q5. The output level of the RF from Q5 is up to 15 Watts, variable by tuning the driver input matching network. The output of the PA then feeds a half-wave dipole.

3.3.2 Receiver Hardware

The down-converter and correlator sections of the receiver are shown in fig. 3.2. The operation of this unit is as follows. A half-wave dipole antenna is used as the receive antenna and feeds the receiver via a length of co-axial cable and a calibrated attenuator. The signal is first filtered in the down-converter by an L/C bandpass filter, and then amplified by Q1, a dual-gate MOSFET. The output of Q1 is mixed with a 15 MHz local oscillator, the output of which is amplified by Q2 and filtered with by FL2, a TV IF SAW filter to produce the required 1st IF signal at 38 MHz. A Plessey amplifier chip, IC1, then further amplifies the 38 MHz signal and its output fed to mixer M2 which is driven by the 15 MHz LO signal bi-phase modulated by the receiver PRBS in M2. The output of this mixer is then fed to the receiver of the IC-735, tuned to the output at 23 MHz, for detection.

To prevent distortion of the impulses at the receiver due to the AGC operation of the IC-735, it was necessary to adjust the RF gain control on the receiver to the point where AGC operation ceased. In this way it was possible to produce an almost constant output signal magnitude for the following processing stages.

The signal at the output of the detector is fed to an 8-bit A/D converter. After conversion, the data is read in and stored by a Toshiba T1000 lap-top PC.

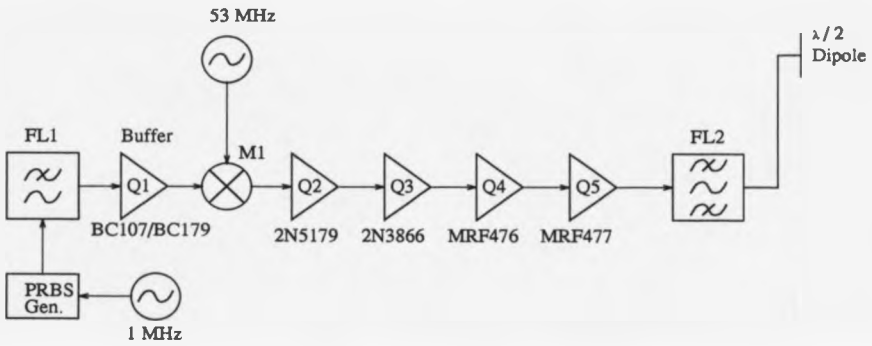


Figure 3.1: Transmitter for multipath measurements

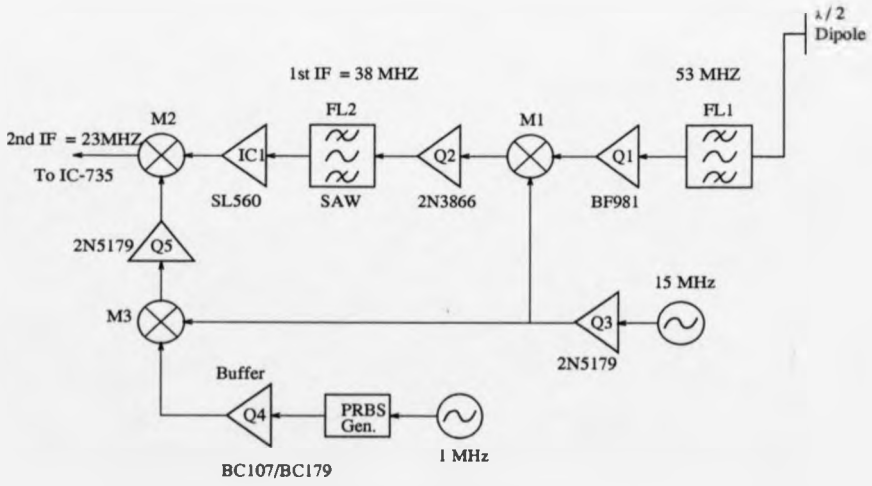


Figure 3.2: Receiver for multipath measurements

3.4 Transmission-loss measurements

In the field work undertaken, both transmission loss and multipath characterization measurements were made. The method used to measure the path loss was simple: the transmitter would be switched to transmit a constant carrier, and the calibrated attenuator ahead of the receiver would be adjusted until a given signal-strength meter reading was observed on the Icom receiver. By later calibration of the receiver (see next section) the signal level producing this reading was known, and the attenuation needed to reduce the input signal to the required level was recorded at each location, thereby permitting the actual signal level at the receiver input to be calculated.

3.5 Calibration of the receiving equipment

Due to the approach taken in the receiver implementation, calibration of the equipment was essential. This was necessary to permit the calculation of the path loss, and also to characterize the distortion introduced in the impulse response measurements by the non-linear detector in the receiver.

3.5.1 Receiver detector calibration

The output of the receiver is taken from the AM envelope detector, which is very non-linear in operation. As a means of characterizing the detector, it's amplitude response was calibrated against that of the receiver product detector, which is assumed linear for signals over the voltage range of interest. A plot of AM detector output vs. product detector output is shown in fig 3.3. A smooth-curve approximation to this was produced using *Mathematica*, with the resulting polynomial (containing an offset and a scaling factor to take into account the A/D process) being;

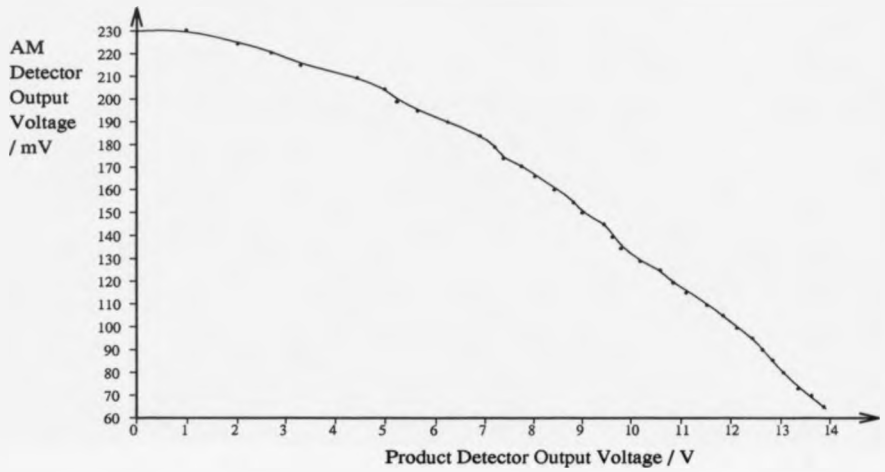


Figure 3.3: Graph showing AM detector vs. product detector responses

$$y = 7.26.10^{-7}x - 0.0202x^2 - 2.29.10^{-6}x^3 + 5.02.10^{-9}x^4 + 5.87.10^{-12}x^5 \qquad (3.7)$$

where y represents the AM detector output and x the product detector output.

Using the above, it was possible to calibrate the received and stored data, and so produce the equivalent data which would have been produced by a linear detector. Since this would result in a measure of received voltage, the received power level (relative) is found by the square of these results,

3.5.2 System transmission loss calibration

In calculating the transmission loss over a channel, it is necessary to calibrate the losses and gains of the measuring equipment. To give a measure of received signal level, attenuation is switched into the receiving system, between the down-converter and main receiver, until the signal strength meter in the main receiver showed a given

level. The level chosen was a reading of $S1$, which corresponds to an input signal of -97dBm, at 23 MHz (as measured in the lab.). The down-converter board has a gain of 27dB (this was changed to 11dB for later measurements due to a modification to the circuit configuration), and therefore the level required for the given meter reading is -(97 + 27) = -124dBm (-108dBm in the latter case) at 53 MHz. With knowledge of the transmitter output power, the remaining system parameters needed are feeder losses and antenna gains. Half-wave dipoles were used, and these antennas have a theoretical gain marginally greater than 2dB. The co-axial cable used (URM-43) has a loss of approximately 1dB per 10m at 53 MHz, and therefore a loss of 1dB would be expected in the receiving system. Overall, therefore, a signal level of -125dBm (-109dBm) at the antenna would give the meter reading considered. Signal levels at the antenna of a level greater than -125dBm could therefore be measured by the amount of attenuation required to give a reading of $S1$, by adding the attenuation figure to -125dBm (for example, 30dB attenuation would be used for a signal of magnitude -95dBm).

3.5.3 Post-processing of delay profile data

It had been anticipated that a certain amount of benefit would be gained by post-processing the received data. This, using an averaging process, could reduce the background noise level and hence improve the effective dynamic range of the system. Unfortunately, the software used in the recording of the data caused regular interruptions in the recording process while data was written to disk, and this had a significant effect upon the spacing of the individual profiles. In the cases where the profile peaks could not be identified, therefore, it was not possible to produce any meaningful results. However, where the peaks could be identified, it was possible to perform an averaging process, and the procedure used is now described.

In order to average the data profiles, it was decided to use a 'comb-filter' approach. Here, the processed data file is examined, and if a profile peak it detected

a unit impulse is assigned to the comb. At all other sample points a value of zero is assigned. The result, a comb with teeth spaced exactly as the profile peaks can then be correlated with the data recorded to effectively average over a number of profiles. This is illustrated in fig. 3.4. The top figure shows the recorded data for a path. Below that, the comb generated is shown, and it can be seen how the 'teeth' are exactly coincident with the peaks in the delay profile. The result of the convolution of these two data sets is shown at the bottom.

The profile in the centre of this result is that which is taken. Demonstrating the effect of the averaging, fig 3.5 shows a single profile together with the result of the averaging process. Generally, it is useful to plot this result on a log scale, since delayed terms are often at a low level relative to the main pulse. For accurate representation, the noise floor should act as the reference, however it can be difficult to determine the exact level of this. Consequently it should be accepted that there will be an error to some degree in the log scale, and this is often not commented upon in published work. The results shown in chapter 2 attempt to minimise this error by referencing the data correctly to the noise floor, but it is anticipated that the inaccuracy in this process is such that the values plotted may, particularly when close to the noise floor, be in the region of 2-3 dB in error. The same does not apply to the data plotted on a linear scale, where such referencing is not so important, of course.

Fig 3.5 shows the single and averaged profiles plotted on logarithmic scales, and here the improvement due to processing can be more clearly seen.

3.5.4 Hardware Limitations

The hardware employed in the propagation work was found to have limitations, particularly with regard to the delay spread measurements.

The time resolution of the recorded profiles relatively poor - the clock rate of 1 MHz producing a resolution (in terms of base width of correlation pulse) of 2

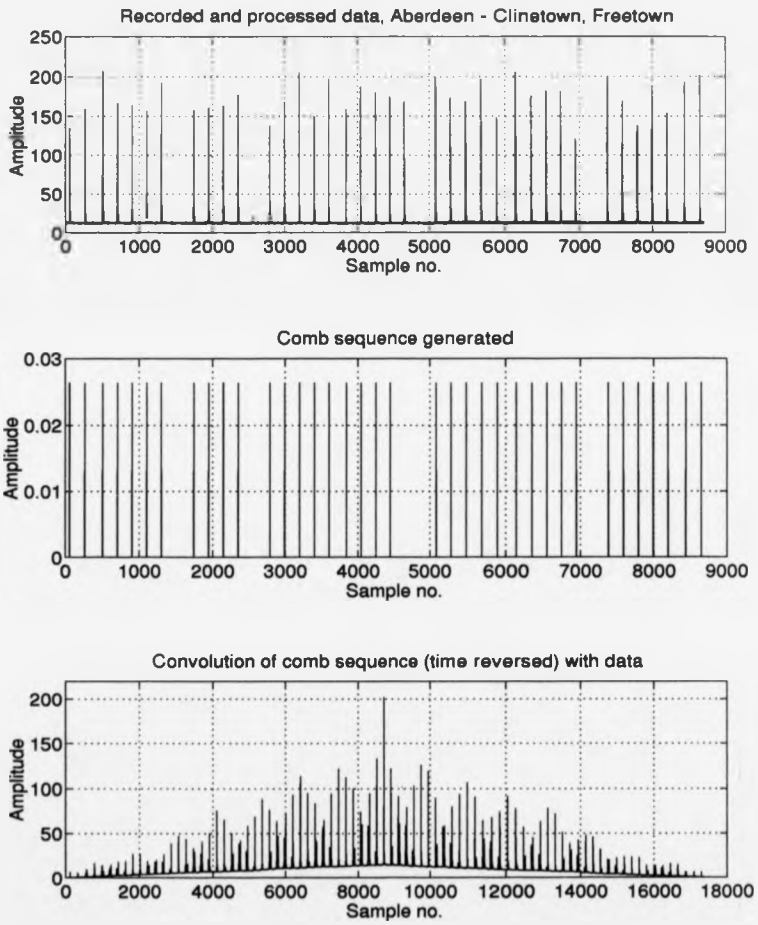


Figure 3.4: Illustration of technique used in averaging delay profiles

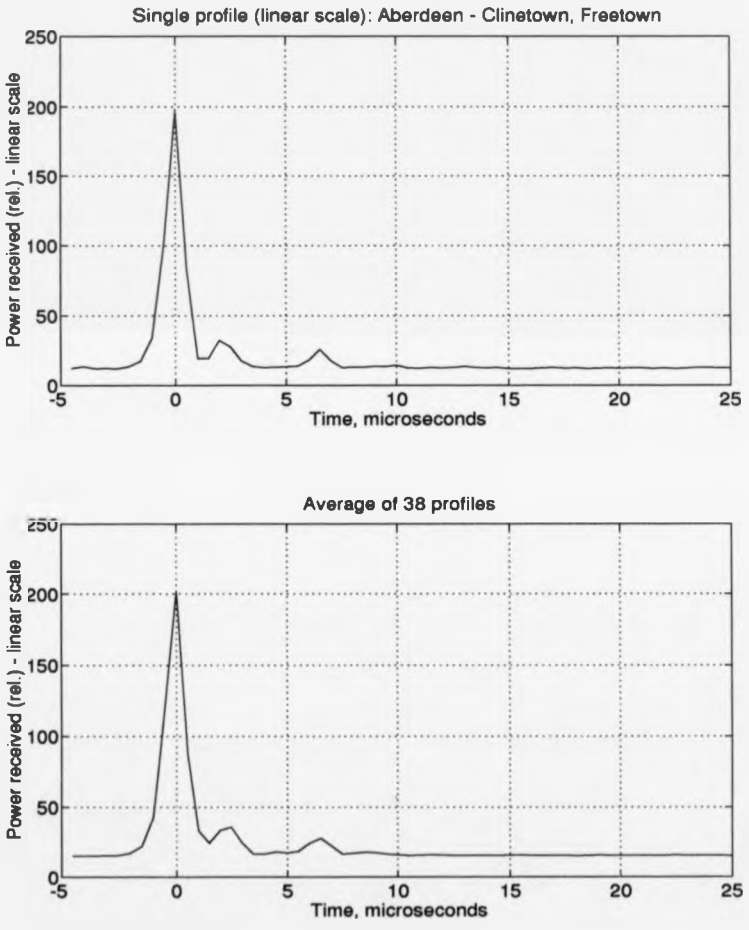


Figure 3.5: Comparison of single profile with averaged profiles

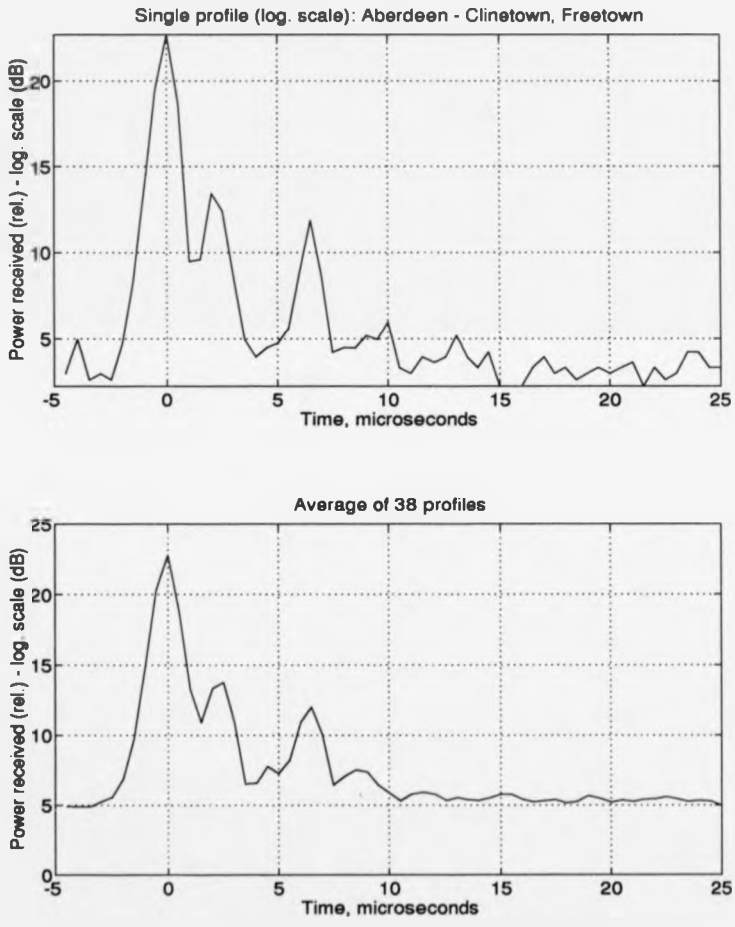


Figure 3.6: Comparison of single profile with averaged profiles, plotted on log. scales

μs. This could be improved by raising the clock rate, but the RF bandwidth of the transmitted signal would increase in proportion, and licencing considerations prohibited such a further increase. The sample rate used in the recording was quite low (250 Hz), which resulted in relatively low resolution in the recorded data.

On a number of paths the signal-noise-ratio was insufficient to provide a useful delay spread profile. This was compounded by the problem with the recording software (interruptions in the recorded data), as the benefits achievable through averaging were limited. One problem causing the above was the increase in receiver noise floor when the receive PRBS generator was operating. The PRBS generator is effectively a wide-band noise source, producing, in this case, a series of tones spaced by $(1 * 10^6/127)$ Hz, i.e. 7.87 kHz, and extending right through the HF range. Although the generator output was filtered, it did not prevent the raising of the noise-floor. This could well have been due to leakage into the IFs of the receiving system, in particular the final IFs in the ICOM transceiver, at 455 kHz and 9 MHz. The actual receiver noise-figure was not especially good (although it wasn't measured) since the receive antenna was followed by a relatively long cable run before the first stage of amplification. The loss incurred in the cable adds directly to the noise figure of the receiver itself, and can hence be a major factor in this system parameter. The pre-amplifier in the down-converter unit was designed to be low-noise, although again the actual noise figure was not measured. This amplifier was preceded by a high-pass filter to ensure that received energy in the HF region would be strongly attenuated, and this filter would have a small loss, in the region of 1 dB, which again adds considerably to the overall noise figure. It is considered that the overall noise figure for this system is probably in the region of 5-6 dB, which could obviously be greatly improved by the use of a low-noise pre-amplifier located directly at the antenna terminals (which could easily lower the noise-figure to below 2 dB).

Some way into the work, very large increases in received noise level were found to be increasingly occurring, and this was traced to oscillations in the 38 MHz IF

amplifier stage in the down-converter unit. The amplifier used a Plessey SL560 IC, and after unsuccessful attempts to cure the oscillation it was decided to modify the circuit by replacing this with a broadband (and stable) bipolar-type amplifier. The gain of the modified converter was reduced by 16dB in the process, and this meant that it was difficult to produce useful delay spread profiles on some of the final measurements as insufficient signal was presented to the 23 MHz receiver.

3.5.5 A Note on Later Work in Tanzania

As part of the further project development, and the PhD of P.J.Chitamu, of the Universities of Southampton and Dar-es-Salaam, a period of July-September 1993 was spent in Tanzania undertaking further, and more comprehensive, propagation measurements. This work drew quite heavily on the experience gained during the work in Sierra Leone, and this being the case many of the problems encountered were overcome, mainly during the design stage. The measurements were taken at a frequency of 300 MHz, which had been adopted as a frequency to be used in the field trials of the initial system equipment. Changes made for this work included:

- Use of low-noise mast-head pre-amplifier to give overall receiver noise figure of around 2dB.
- Use of VHF/UHF transceivers as tunable IF and LO source in receiver and transmitter respectively, at 430.2 MHz.
- Recording of audio output (i.e. correlation output modulating audio carrier), giving good linearity.
- Commercial A/D hardware and software interface used, allowing un-interrupted recording of delay profile data at a sample rate of 5 kHz.
- 511-bit PRBS used, clocked at 5 MHz, giving resolution of 0.4 microseconds.

Chapter 4

Modulation Background and Implementation Considerations

Inevitably, the delay spread results obtained were, on the whole, more satisfactory. The receiver noise floor was substantially lower, permitting useful results to be obtained on a larger proportion of the paths considered. Since there is an audio carrier present in the received profile data, it will be possible to generate a complex channel impulse response, hence giving a full channel characterization. The results of this work will be presented in full in the thesis of Chitamu in 1995.

4.1 Introduction

This chapter describes the choice of modulation scheme for the system, starting from some basic modulation background and proceeding through the important considerations of power efficiency, spectrum efficiency, effects of filtering and implementation practicability.

4.2 Modulation

It is necessary to generate a signal at the required radio frequency (VHF/UHF) which contains the information to be transmitted. In this application, digital modulation is necessary, since the information to be transmitted is binary data.

The subject of modulation is covered well in standard texts, such as [80, 81, 82], and so only a very brief background to the basic theory is appropriate here.

4.2.1 Digital Modulation Techniques

Modulation is the process of varying one or more characteristics of the RF carrier in order to convey the information being transmitted. If the modulating signal is of a digital nature, using single bits, or groups of bits, then digital modulation results. The three basic properties which can be varied are the amplitude, the phase and the frequency, leading to the terms amplitude modulation (AM), phase modulation (PM), and frequency modulation (FM). With a digital modulation source, the above schemes are termed ASK (amplitude shift keying), PSK (phase shift keying) and FSK (frequency shift keying).

Amplitude Shift Keying

ASK involves varying the amplitude of the carrier in response to the modulating signal. Pure ASK schemes, however, tend not perform well in the presence of noise and fading in comparison with PSK and FSK. They are rarely used, therefore, in

modern communication systems. When combined with PSK, however, they can produce spectrally efficient modulation schemes. One major disadvantage of ASK is the requirement for a linear transmitter, as this has tended to imply poor power efficiency: this fact led to the development of high-level modulation, where the amplitude information is used to modulate the (non-linear) amplifier power supply voltage, and thereby permit good power efficiency.

Phase Shift Keying

By varying the phase of the carrier, PSK is produced. Phase modulation has the advantage that the signal produced is of constant envelope, and can therefore be amplified using non-linear, and efficient, power amplifiers. A major disadvantage, however, is the need at the receiver for a replica of the original carrier, of correct phase, to facilitate the demodulation process. Differential techniques may be employed, however, and here the carrier received during the previous bit/symbol period may be delayed and used, with the obvious penalty of it degrading the required signal-to-noise ratio required for a given error rate.

Frequency Shift Keying

Frequency modulation of a carrier has the advantages that a constant envelope signal is produced, as with PSK, and that the detection process is relatively simple at the receiver, using, for example, a limiter-discriminator technique. FSK does require a higher signal-to-noise ratio than comparable PSK schemes, however, particularly when coherent detection can be employed. Additionally, PSK schemes tend to be more spectrally efficient than FSK.

Multi-level keying

All of the above modulation methods can use straight binary data as the modulating signal, or data that has been grouped together into 'symbols' of required length.

Thus, *multi-level* or *M-ary* schemes can be produced, offering greater spectral efficiency than the simple binary schemes. If L is the number of bits in a symbol, then M, the modulation scheme level, is given by $M = 2^L$. The theoretical bandwidth efficiency of M-ary schemes (assuming ISI-free data shaping) is L $b/s/Hz$. However, the higher the level of scheme used then the greater the signal/noise ratio required for a given error rate. High level schemes can therefore be thought of as less power efficient than lower level schemes. For example, to increase the spectral efficiency by 1 $b/s/Hz$ typically requires an increase of at least 3 dB in required E_b/N_o. Due to the steepness of error rate versus E_b/N_o curves [83], changing the modulation scheme level will have a great affect on the error rate encountered. On a channel where QPSK has an error rate of 10^{-4}, 8-PSK would have an error rate of around 10^{-2}, which means the difference between a usable and a non-usable channel (a system *outage* is often defined as an error rate of greater than 10^{-3}). The higher the level of the modulation scheme, the greater is the susceptibility to errors resulting from such effects as multipath propagation, non-linearities and ISI, and since these errors are effectively irreducible (in the sense that raising the transmit power will have no effect), low level schemes can offer distinct advantages in such situations.

4.2.2 Power and Spectrum Efficiency

From the above, there will obviously be a compromise to be made between power efficiency and spectrum efficiency, and the particular application of the system being designed will tend to dictate how that compromise should be made. Two examples of this trade-off are terrestrial microwave links, and satellite links. The terrestrial microwave links employ such schemes as 64-QAM, permitting highly spectrum efficient communication, but with an increased energy-per-bit (or, transmit power) requirement of around 10 dB over 4-QAM (QPSK). In satellite communications the huge path losses on the up- and down-links, together with the limited power available at the satellite terminal, have tended to limit the modulation used to QPSK-type

schemes, as the signal/noise level required for higher level schemes was not available.

In [113] a detailed treatment of this subject, with particular regard to satellite links, is presented. Feher [113] defines spectrum efficient modulation schemes as being those schemes offering efficiencies of 2 $b/s/Hz$ or greater. Similarly, power-efficient schemes are defined as those which can offer a P_e of 10^{-8} for an E_b/N_o of less than 14 dB. Here, P_e is the error probability and E_b/N_o is the *average received bit energy-to-noise density ratio.* The ratio E_b/N_o is the measure most often used in system calculations for error rate performance, and is related to the *average carrier-to-average noise power ratio*, C/N, by

$$\frac{E_b}{N_o} = \frac{C}{N}.\frac{B_w}{f_b} \tag{4.1}$$

where B_w is the receiver noise bandwidth and f_b is the transmit bit rate.

4.2.3 Efficiency Requirements of the Rural Radio System

There are similarities regarding choice of modulation scheme for the digital rural radio network and the satellite communications systems mentioned above. Both systems are strictly limited by the power available for transmission, as solar power is the source both in the satellite transponders and in the rural network transceivers. Both systems have tight requirements on the signal spectra since many channels need to be accommodated within an allocated bandwidth. It is anticipated that around 200 channels will be required for the rural digital radio system. If the baseband data rate is 80 kb/s, then a spectral efficiency of $1b/s/Hz$ would imply a total frequency allocation of 16 MHz. With an efficiency of $1.5b/s/Hz$, the corresponding system bandwidth would be 10.7 MHz. Such a reduction would obviously be of great benefit. The dependence of the system upon solar power, however, means that the transmit power available is strictly limited, and consequently the complexity of the modulation scheme is accordingly limited. This factor is compounded by the necessary use of *omni-directional* antennae, which will consequently provide minimal

gain over the radio paths.

One important factor in the operation of this system is that of frequency re-use, which will be dependent to a large degree on the chosen modulation scheme's susceptibility to co-channel interference. If it is assumed that all transmitters are operating at the same output power, then errors occurring due to co-channel interference will be insensitive to changes in that general transmit power since both the wanted and the interfering signal will alter together. Since required signal-to-interference ratios for a given error rate increase as the level of modulation scheme increases, it follows that frequency re-use will become less efficient, which in this system would imply a decrease in overall system efficiency, leading, for example, to higher rates of unsuccessful call attempts, and greater call set-up times.

The result of the above constraints suggests the use of a modulation scheme in the QPSK family. Higher level schemes are ruled out because of power efficiency and frequency re-use issues, and lower level schemes because of spectral efficiency. QPSK is therefore a good compromise, and is indeed widely used in situations where these constraints prevail.

4.2.4 QPSK, OQPSK and MSK

Having stated that the QPSK family of modulation schemes represent a suitable compromise between power and spectral efficiencies, a more detailed description of these schemes is now appropriate.

Phase modulated signals can be described (as can any modulated signal) by the general equation

$$s(t) = I(t)\cos\omega_c t + Q(t)\sin\omega_c t \tag{4.2}$$

where I and Q are the amplitudes of the modulating signals on the orthogonal carriers. In QPSK, the incoming data stream is split into two streams, forming the I and Q baseband channels, each with a bit rate one half of that of the source. The quadrature carriers are then modulated by this data, and the outputs summed to

give the QPSK signal, as shown in figure 4.1. Four possible phases of the carrier therefore result, and these can be drawn on a *signal space diagram*, which plots $I(t)$ on the x-axis, and $Q(t)$ on the y-axis. The signal space diagram is the standard way to illustrate the phase/amplitude properties of digital modulation schemes. A signal space diagram of QPSK is shown in figure 4.1. The carrier may shift from one to any other phase location upon the transition of the data on the I and Q channels, as defined by equation 4.2. These phase changes are also shown in figure 4.1. Phase shifts of $0,\pm\pi/2$ and $\pm\pi$ are possible with QPSK. In the time domain, this results in phase reversals in the carrier waveform at the bit transitions.

A modified version of QPSK, called *OQPSK* (offset QPSK) was introduced as a means of improving the transmission performance of QPSK in practical systems employing filtering and hard-limiting (see next section). This technique avoided the transitions of π by staggering the I and Q data streams by T_b, where T_b is the data source bit period. The result of this staggering of the channels is that only the data on one channel can change at any instant, hence allowing phase shifts of 0 or $\pm\pi/2$ only.

A modification to OQPSK was later introduced, known as *MSK*. It was realized that the instantaneous transitions in the carrier waveform contained high frequency energy, and thus the sidelobes of the QPSK and OQPSK spectra were rather large. MSK, by using half-sinusoid weighting on the I and Q channels, smoothed these transitions whilst retaining a constant envelope signal. The result of this, in the frequency domain, is a much faster roll-off of sidelobes away from the carrier frequency, but at the expense of a larger main lobe.

4.2.5 The need for filtering

Any modulation scheme that takes square pulses as the modulating signal will be spectrally inefficient, in that sidelobe levels will be significant (see previous section for QPSK-type modulation). This causes considerable interference to adjacent

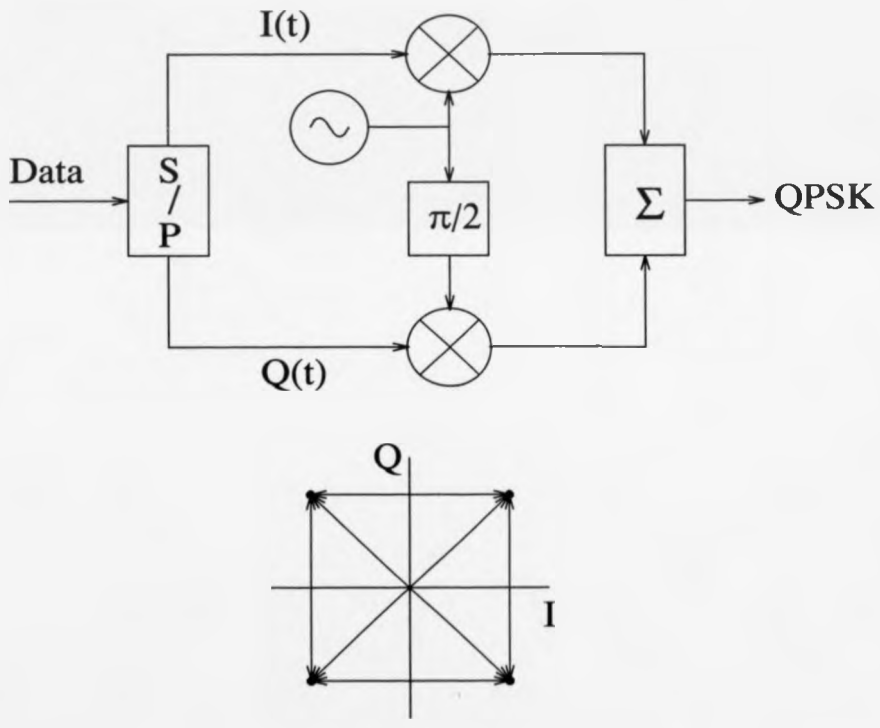


Figure 4.1: QPSK modulator implementation and signal constellation diagram with phase transitions

channels in multi-channel radio systems, and is therefore unacceptable. Filtering of the data is therefore necessary to reduce the sidelobe levels. A consequence of this filtering, however, is that envelope fluctuations may be introduced into the rf signal to be transmitted, and this is particularly noticeable in PSK systems. If a non-linear amplifier is then used to amplify this signal, sidelobe regeneration will occur since the amplifier will tend to produce a constant envelope signal at its output if operated in saturation (i.e. at maximum power efficiency), thereby reversing much of the effect of the previous filtering. The data filtering can, of course, be performed prior to phase or frequency modulation, which leads to CPM, *continuous phase modulation*. CPM divides into two subsets - *partial response* and *full response* [61, 62]. Partial response CPM results in spreading of the pulse over adjacent symbol intervals, and hence ISI, while full response CPM restricts pulse energy to the current interval only. Unfortunately, schemes of the former type tend to be spectrally inefficient, and those of the latter type require complex receivers in order to perform well. Consequently, performing filtering after modulation is in many cases preferable.

The fact that non-linear amplification, a form of *limiting*, causes spectral regeneration is apparent when consideration is given to the limiting process. In standard QPSK, data is split into in-phase (I) and quadrature-phase (Q) channels. This data may then be filtered identically and both channels will have similar frequency spectra. When used to modulate quadrature carriers which are then summed, the resultant rf spectrum is merely the two-sided version of the sum of the baseband spectra, symmetrical about the carrier frequency. The envelope of the rf waveform is given by

$$E(t) = \sqrt{I(t)^2 + Q(t)^2} \tag{4.3}$$

If the I and Q data is smoothed by filtering, then E(t) will not be constant. With I and Q both unfiltered NRZ (that is, $\pm$ 1) the envelope would always be

$\sqrt{2}$. The effect of limiting is to produce a constant envelope signal with ideally zero phase shift between input and output. As shown in [66], this effect produces *crosstalk* between the I and Q channels at the limiter output. In QPSK where both channels have common transition times, the crosstalk can lead to squaring of the filtered pulses (i.e. reverses the filtering process) where both channels change simultaneously, and glitches where only one channel has a change of data. These glitches do not occur at the sampling instants, however, and so the effect is not so critical. With OQPSK, however, both channels do not change simultaneously, and so the effect is limited to producing glitches on one channel when the other is in transition. Due to the staggering process in OQPSK the glitches now occur at the sampling instants on each channel, and therefore the detection process is affected. The fact that re-squaring of the data occurs in QPSK, but does not in OQPSK means that greater spectral re-spreading occurs in QPSK.

It was stated previously that MSK, by smoothing the data on the I and Q channels, has a constant envelope. It would therefore be expected that limiting would have little effect upon the spectrum. However, in practical systems the sidelobe levels in MSK are too large, and therefore further smoothing of the I and Q data is required. Such further filtering will usually produce envelope fluctuations. In this respect, MSK is practically found to perform similarly to OQPSK in terms of sidelobe regeneration, however since the main lobe of MSK is wider than that of OQPSK, the spectrum of filtered and non-linearly amplified OQPSK remains superior to that of filtered and non-linearly amplified MSK.

For the above reasons, OQPSK is often used in satellite links where non-linear amplification is employed in the transmitter and coherent detection is possible.

4.2.6 Variations on QPSK and MSK : $\pi/4$-QPSK and GMSK

The digital cellular revolution, in particular, provided the need for power and spectrum efficient modulation schemes to be developed further. Frequency spectrum

is at a premium due to the congestion present in the VHF/UHF bands, and the widespread predicted use of portable transceivers dictated that the power consumption of the radio units should be minimized. From this requirement, two modulation schemes in particular have been pushed to the forefront of development, although the ideas are by no means recent. These are a) GMSK and b) $\pi/4$-QPSK.

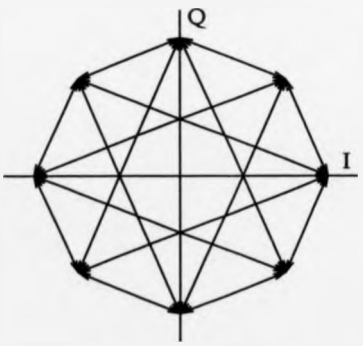
GMSK

GMSK, or Gaussian-filtered MSK, was first introduced in [56]. Here, the baseband data is filtered by a Gaussian filter before phase modulating the carrier. The Gaussian form of the pulses has the effect that the signal remains of constant amplitude with much reduced sidelobes in the frequency domain, but ISI is introduced, which is the main short-coming of the scheme. Coherent and differential detection of GMSK are possible, however for efficient demodulation the use of a Viterbi decoder is important, due to the inherent ISI. The receiver complexity is therefore greater than that for $\pi/4$-DQPSK for similar performance. GMSK does have the advantage, however, that the constant envelope permits non-linear amplification without spectral degradation. The scheme is, however, susceptible to spectral spreading effects when non-linear amplification is performed on a signal containing phase errors which have resulted in amplitude fluctuation; such errors arise from, for example, d.c. offsets and unequal I/Q levels. Such spreading effects can be quite significant, even for low values of these errors [131].

$\pi/4$-QPSK

The original idea for this scheme was presented in [71], as a variation of DQPSK. Rather than the usual phase shifts of QPSK, $\pi/4$-QPSK involved the modification that phase shifts of $\pm\pi/4$ and $\pm 3\pi/4$ are those permitted. In this way, the phase trajectories of the signal would not pass through the origin, and hence the signal envelope would never decrease to zero. The spectral spreading caused by limiting

would therefore be reduced from that found with QPSK, while if a linear transmitter was to be employed, the scheme would place far less stringent requirements on the operation of the PA, due to the reduced envelope fluctuation. The signal space diagram for $\pi/4$-DQPSK, and the symbol to phase shift mapping is shown in figure 4.2. When serious consideration began to be given to linear modulation schemes due to their spectral efficiency, probably around the time of the publication of [64], $\pi/4$-DQPSK became popular. Amongst other applications, it has been adopted for ADC, the North American second generation cellular system, JDC, the Japanese counterpart [105], and TETRA, a European trunked radio system. Non-coherent, coherent and differential detection of $\pi/4$-DQPSK are all possible, and therefore quite simple receiver structures can result. [84, 85]



Information Symbol	Phase Shift
1 1	$\pi/4$
0 1	$3\pi/4$
0 0	$-3\pi/4$
1 0	$-\pi/4$

Figure 4.2: $\pi/4$-DQPSK signal space diagram and phase mapping

4.2.7 Constraints Placed Upon the Modulation Scheme by the System

The rural radio system, by its nature, places certain constraints on the transmission performance. Some of these are:

- good power efficiency

- good spectral efficiency
- resistance to propagation impairments (multipath)
- operation in burst-mode
- non-coherent and low-complexity receiver structure
- resistance to co-channel interference

The power and spectral efficiency requirements have been mentioned previously, therefore attention is concentrated on the remaining points above.

Resistance to Propagation Impairments

Chapter 2 described the radio environment in which the system is likely to operate. It was shown that in hilly terrain multipath effects may become significant. This spreading of the transmitted signal in time can cause fading and ISI, thereby introducing a potentially large error rate. In this system, omni-directional antennae will be employed, offering no rejection of the multipath components. Additionally, the large antenna heights may mean that the number of potential 'reflectors' of the transmit signal is increased over the situation where antennae would be relatively low, and hence the delay spread could be increased.

The modulation scheme used must not, therefore, be particularly susceptible to performance degradation under such circumstances. Consideration is given to the effect of multipath on several modulation schemes in [21], and the results are now summarized.

The work concentrated on the portable radio environment in particular, and therefore path models of Rayleigh statistics were used. The three main sources of system errors were given as

- Frequency selective fading of signal due to constructive/destructive interference effects

- Inter-symbol interference
- Performance degradation of timing recovery hardware

The general conclusions of Chuang were that for low values of delay spread normalized to the transmission bit rate the main error mechanism was that of signal fading. For high values, timing problems were thought to be the likely source of most errors. The channel profile appeared not to be significant for the range of delay spreads simulated. Irreducible bit error rates resulting from multipath were believed to be only possible when the sum of the magnitudes of the multipath components exceeded the magnitude of the direct signal, (as it is in this case that increasing the transmitted power (and hence received signal-noise ratio) has no effect. The above held for all modulation schemes considered. In comparing the effects on several digital modulation schemes, the following was found. Cross-rail interference can be a serious problem, and hence BPSK offers a certain resistance not found in quadrature (and higher level) schemes such as QPSK, GMSK etc. Also, OQPSK and MSK, with their staggering of quadrature data, offer less resistance than standard QPSK since the sampling instant on one channel is removed by only half a symbol period from the symbol transition on the other channel, which implies that ISI would be more significant at sampling instants due to cross-rail effects. The effects of the filtering employed in QPSK and GMSK were also considered. It was shown that with root raised cosine filtered QPSK, the higher the roll-off factor, the better the performance in multipath. This was largely attributed to ISI effects due to the different pulse shapes. In this respect, the error rate with a roll-off of 1 was generally in the region of .1 that of the error rate with roll-off of 0. For GMSK, while there was a slight effect of filter bandwidth on error rate, it was not so pronounced as with QPSK. Comparing the two schemes, it was found that QPSK with a roll-off of 0.5 would in practice have a superior error rate to GMSK by a factor of between 2 and 4. QPSK with a roll-off of 1, however, was in the region of 10 times more resistant than GMSK. These irreducible error rates were calculated

using the assumption that the additive noise was zero, hence the signal-noise ratio was infinite.

Operation in Burst-Mode

The fact that TDD operation will be used places important requirements on the transmitter implementation. Obviously, for reasons of power consumption, the transmitter should only be on when required for transmission, and therefore rapid switching is essential. However, the effect which is of major concern in this work is the spectral spreading of the rf signal due to its burst nature. Instantaneous on/off transitions of this signal would produce severe spectral spreading, with the effect determined theoretically by burst duration. This spectral spreading would cause interference to neighbouring channels, and is therefore of great concern.

It is necessary, therefore, to employ some form of *shaping* at the on/off transitions to reduce or prevent the effect. This does, however, dictate that the power amplifier used must be able to reproduce the shaping waveform, and this weighs heavily against the use of a standard saturating non-linear amplifier with no power control, and, perhaps more importantly, suggests that the benefits due to transmitter simplicity of using a constant-envelope modulation scheme, such as GMSK, can be greatly reduced.

The above points to the use of an appropriate linear modulation scheme for this application, provided that power efficient transmitter hardware will be employed.

Receiver Structure

Although work on the receiver does not form part of this work, it is obviously a crucial factor in the choice of modulation scheme, and therefore needs consideration.

While coherent detection schemes do offer the significant advantage that the required E_b/N_0 is less than that required for non-coherent schemes, for a given error rate, they do require effective carrier recovery at the receiver. It is the need

for this carrier recovery which is the major drawback of coherent detection. The nature of this radio system, with the need for very rapid detection of short bursts of data, implies that non-coherent detection may be necessary, since the acquisition time associated with coherent detection may be too long. Whilst it is possible that synchronization techniques may be developed to permit the coherent route to be taken, initially it is assumed that non-coherent detection will be employed, and hence the modulation scheme employed should be capable of this type of detection.

The use of a low-complexity receiver, such as a limiter-discriminator integrate-and-dump type, can offer reasonable performance at low-cost [14]. A receiver structure of this type has been considered a suitable choice for this system, as has a relatively simple differential detection type structure. With the advances in DSP technology, the use of DSP hardware in the implementation of a significant part of the transceiver hardware is a distinct possibility. Whilst of ever-decreasing cost, the complexity of any implementation may be increased, and these issues will need addressing when decisions on the selected implementations are to be made. Initially, though, the choice of any transmission standard should not restrict the actual technology used in implementing the receiver.

4.3 Use of $\pi/4$-DQPSK

Taking all of the above into consideration, $\pi/4$-DQPSK was chosen to be the most appropriate modulation scheme for the system. To achieve good power efficiency, the use of a linearized transmitter is necessary, with the power amplifier operating in a non-linear mode.

4.4 Filtering

Filtering is a necessary aspect of digital radio systems seeking spectral efficiency. The transmission of unfiltered NRZ baseband data produces large sidelobe levels,

which cause severe adjacent channel interference. Smoothing of the baseband data is therefore required to minimise this interference.

The most popular type of filter employed in data transmission is the *raised-cosine*. This family of filters theoretically can provide significant attenuation of sidelobes without intersymbol interference [80, 81, 82]. Raised cosine filters are generally specified by their *roll-off factor*, α, or *excess bandwidth*. For example, filtering a data stream of bit rate r kb/s with a filter of roll-off factor α will produce a filtered stream containing frequency components up to $r(1+\alpha)$ kHz, or with a bandwidth of $100(1+\alpha)$ percent of r. The use of a raised cosine filter does lead to a long duration time domain response (theoretically infinite), resulting in problems with practical implementations, as truncation will obviously be necessary in any FIR system, and, at least on the leading edge, in an IIR implementation.

The filter with $\alpha = 0$, is the ideal *brick wall filter*, offering a flat response within the passband and allowing no energy outside the passband to pass through. The lower the roll-off factor, however, the more the energy of a pulse will be spread out in time. This causes the degradation of the output spectrum due to truncation of filter responses with high roll-off to be much less harsh than that for low roll-off factors. In addition, the lower the roll-off factor the higher the resulting *peak-to-mean* power ratio of the filtered signal. This issue is particularly important when considering transmitter implementation, as it determines the range of amplitude swing over which the transmitter must efficiently operate, and also for co-channel interference purposes, where the higher the peak amplitude of the interfering signal the greater the probability of received errors due to threshold crossing. A compromise is therefore required between the time and frequency responses, in order to minimise the detrimental effects of the above.

In a practical radio system, filtering will be divided between the transmitter and the receiver. Ideally, for an overall raised cosine response, the transmitter and receiver will both have *square root raised cosine filters* (often shortened to root raised cosine), thus combining to give the required overall response while meeting

the matched-filter requirement.

The frequency response of the root raised cosine filter is

$$R(\omega) = \begin{cases} \sqrt{T} & 0 \leq |\omega| \leq \pi(1-\alpha)/T \\ \sqrt{T}\cos[T(\omega - \pi(1-\alpha)/T)/4\alpha] & \pi(1-\alpha)/T \leq |\omega| \leq \pi(1+\alpha)/T \\ 0 & |\omega| > \pi(1+\alpha)/T. \end{cases} \tag{4.4}$$

with the corresponding time domain response being

$$r(t) = \sqrt{T}\left(\frac{sin[\pi t(1-\alpha)/T] + 4\alpha t/T\cos[\pi t(1+\alpha)/T]}{\pi\frac{t}{T}[1-(4\alpha\frac{t}{T})^2]}\right. \tag{4.5}$$

Truncation in the time domain is effectively equal to *windowing* the impulse response of the filter with a rectangular window. Since multiplication in the time domain equates to convolution in the frequency domain, we have

$$s(t) = r(t)f(t) \tag{4.6}$$

and

$$S(\omega) = R(\omega) \otimes F(\omega) \tag{4.7}$$

where $s(t)$ is the truncated impulse response and $f(t)$ the windowing function, and their frequency domain equivalents given by $S(\omega)$, and $F(\omega)$, with $\otimes$ representing convolution.

Simulation work was carried out to investigate the performance of root-raised cosine filters with different roll-off factors and truncation lengths. Plots of the impulse responses of root raised cosine filters of various roll-offs and resulting constellation diagrams for a $\pi/4$-DQPSK system are shown in figure 4.3 and figure 4.4, respectively. In addition, the resulting constellation diagrams which would be found at the output of a second identical filter are shown in fig 4.5; these are the ISI-free

waveforms that are ideally produced at the output of the receiver filter.

The effect of varying the filter truncation length is shown in figs 4.6 - 4.10. The truncation length, L, is defined as the number of bit (symbol) periods over which the response extends. For example, in the case of L=7, the data in a given bit (symbol) period will affect the three periods before and the three following.

These figures demonstrate well the effect of filter truncation on the output spectra, and highlight the change of effect with roll-off factor. It can be seen that the lower the roll-off factor the higher is the degradation in output spectrum resulting from finite truncation of the filter response. Longer truncation lengths are therefore required for a given value of adjacent-channel interference. A comparison of the spectra also show how, as expected, the width of the main lobe increases as the roll-off factor is increased. This has obvious implications for channel spacing in radio systems. In terms of the signal constellation, which is an important consideration in practical implementations of transmitters, it can be seen that the constellation is much cleaner for the larger roll-off factors. Since the distance of any constellation point from the origin represents the relative amplitude of the output signal, and as it is striven to avoid amplitudes tending towards zero, there is a strong incentive to choose a roll-off factor which avoids transitions passing very close to the origin. There is obviously, therefore, a large trade-off in choice of roll-off factor for raised cosine filtering. Too low a roll-off factor produces very high modulation depths and results in large side-lobe levels due to truncation effects, but does produce a narrow main lobe. Too high a roll-off factor, while performing well in the respects of sidelobe levels and envelope fluctuation, does result in poor spectral efficiency due to main lobe width. With ideal filtering and $(\pi/4)$ QPSK, spectral efficiencies of 2 $b/s/Hz$ occur with a roll-off factor of 0, and 1 $b/s/Hz$ with $\alpha = 1$.

Although it isn't possible to tell from the constellation diagrams due to the absence of any time reference, the output waveforms after full raised filtering are free of ISI at the sampling instants, as would be required for efficient detection the received waveform.

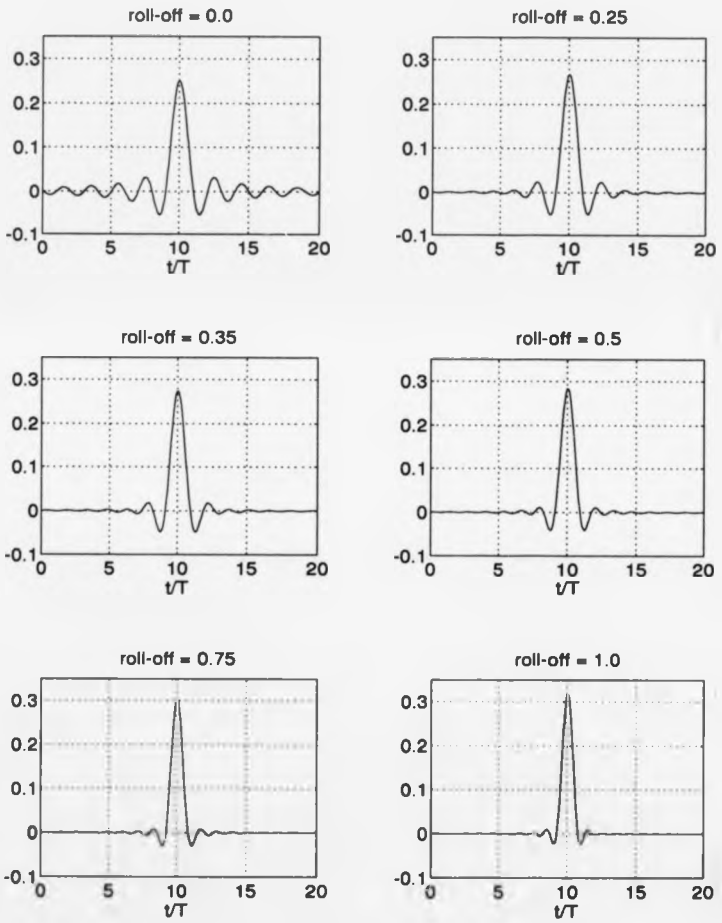


Figure 4.3: Impulse responses of root raised cosine filters

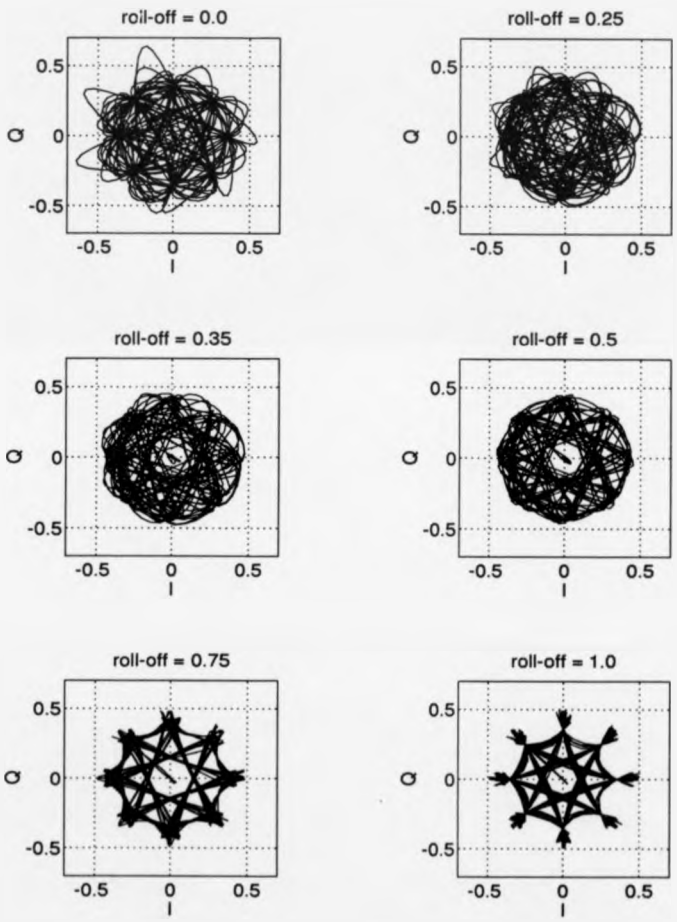


Figure 4.4: Constellation diagrams of $\pi/4$-DQPSK signals with root raised cosine filters

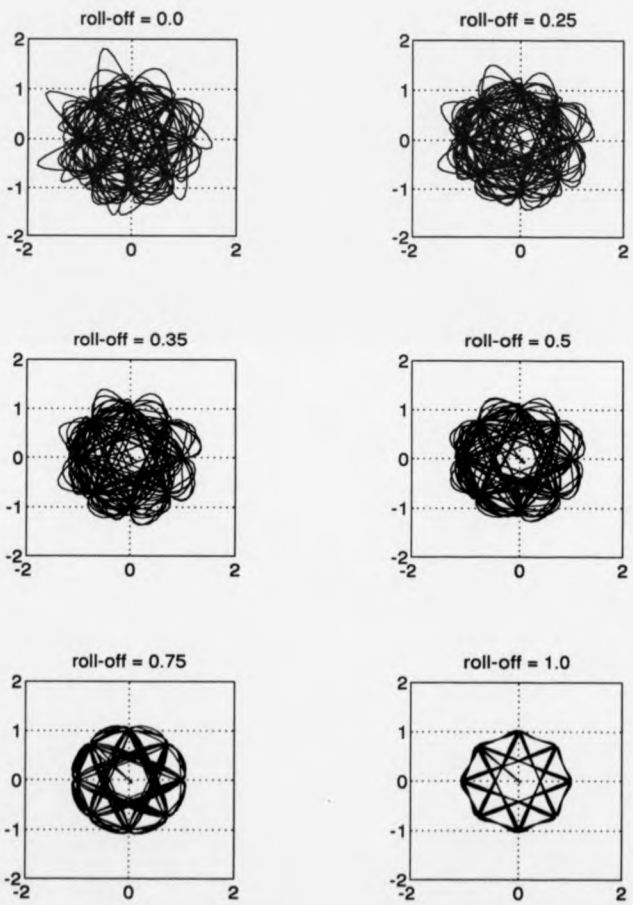


Figure 4.5: Constellation diagrams of $\pi/4$-DQPSK signals with cascaded root raised cosine filters

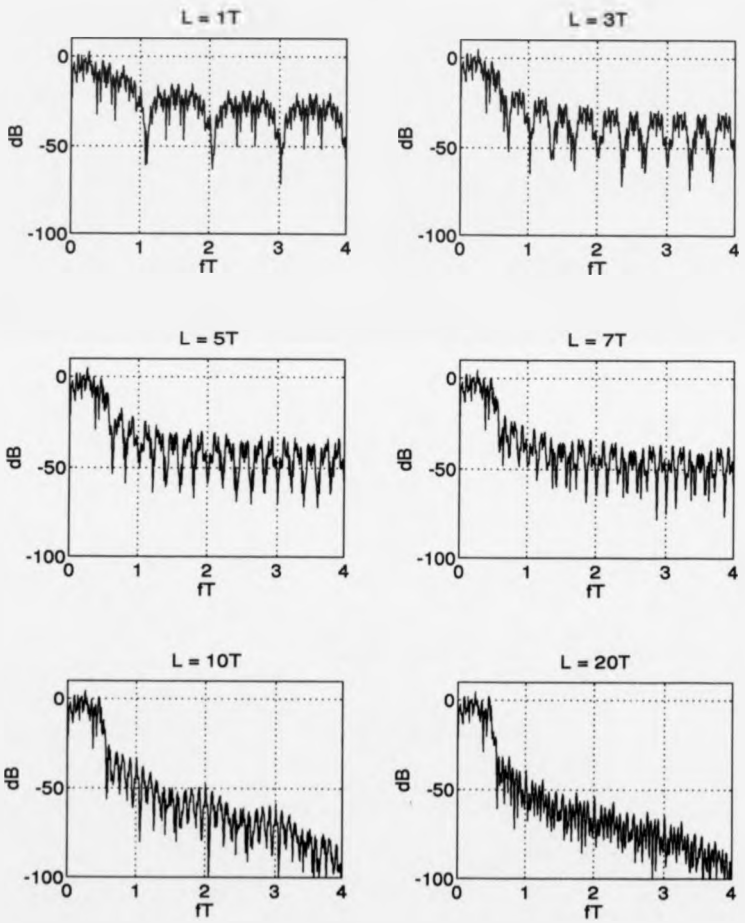


Figure 4.6: Baseband spectra with root raised cosine filtering of roll-off 0.0 and different truncation lengths

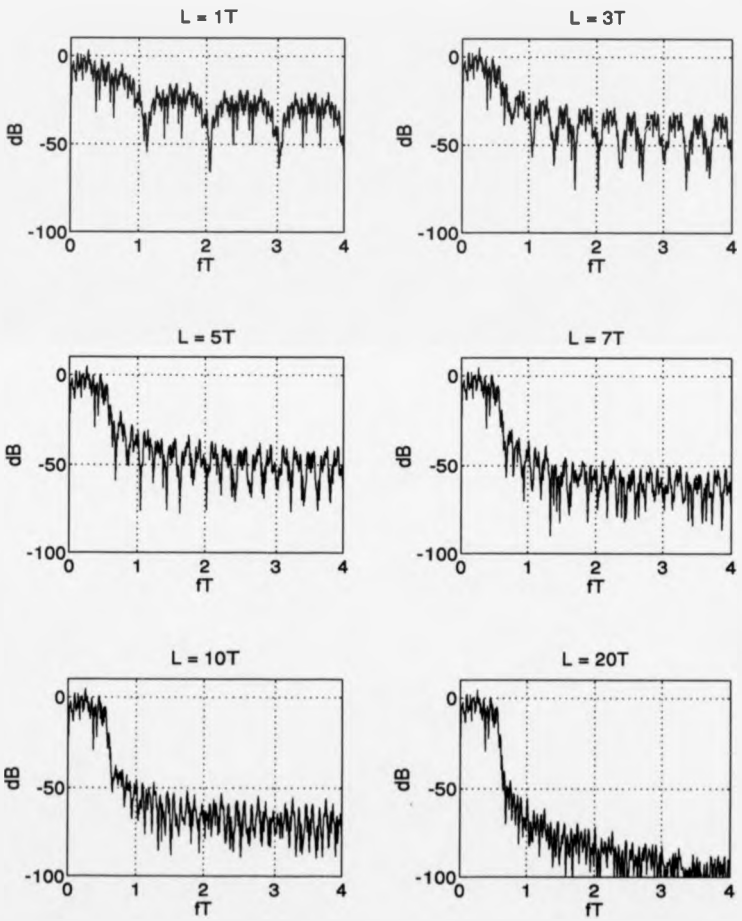


Figure 4.7: Baseband spectra with root raised cosine filtering of roll-off 0.25 and different truncation lengths

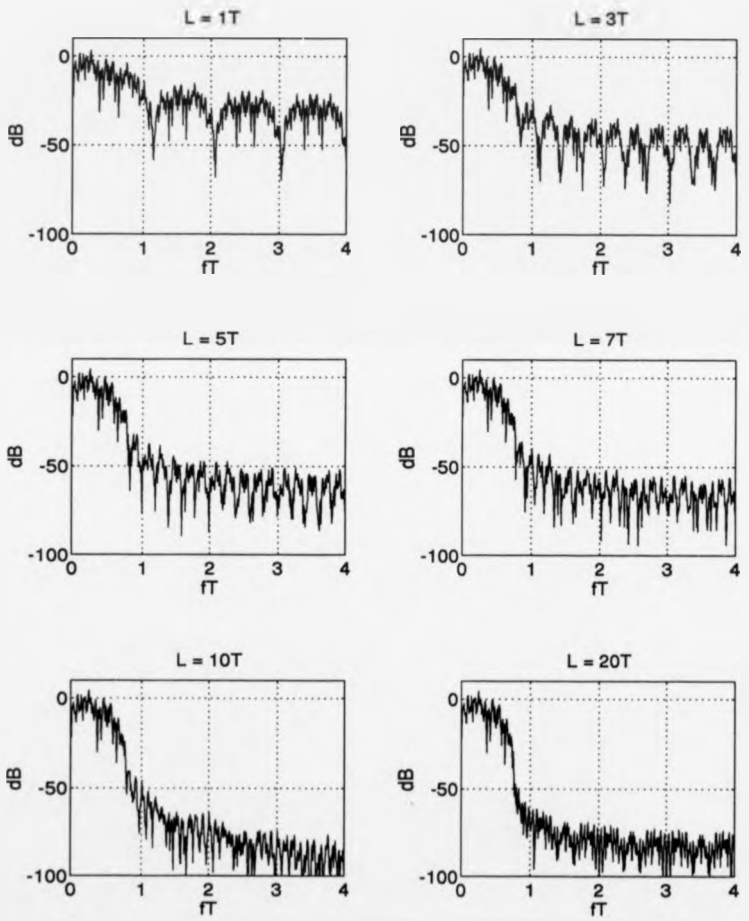


Figure 4.8: Baseband spectra with root raised cosine filtering of roll-off 0.5 and different truncation lengths

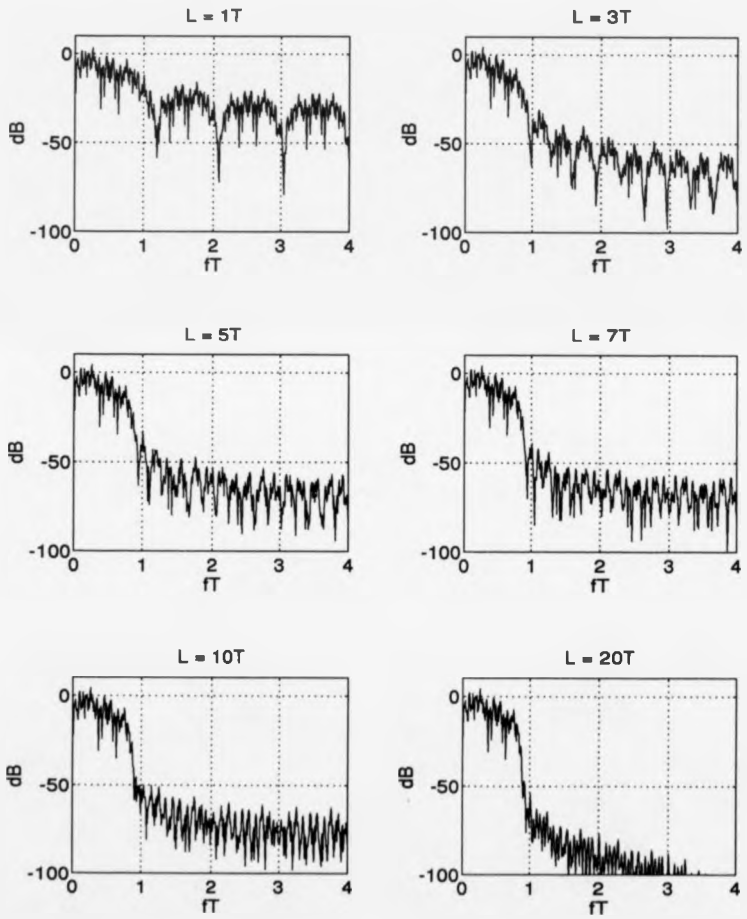


Figure 4.9: Baseband spectra with root raised cosine filtering of roll-off 0.75 and different truncation lengths

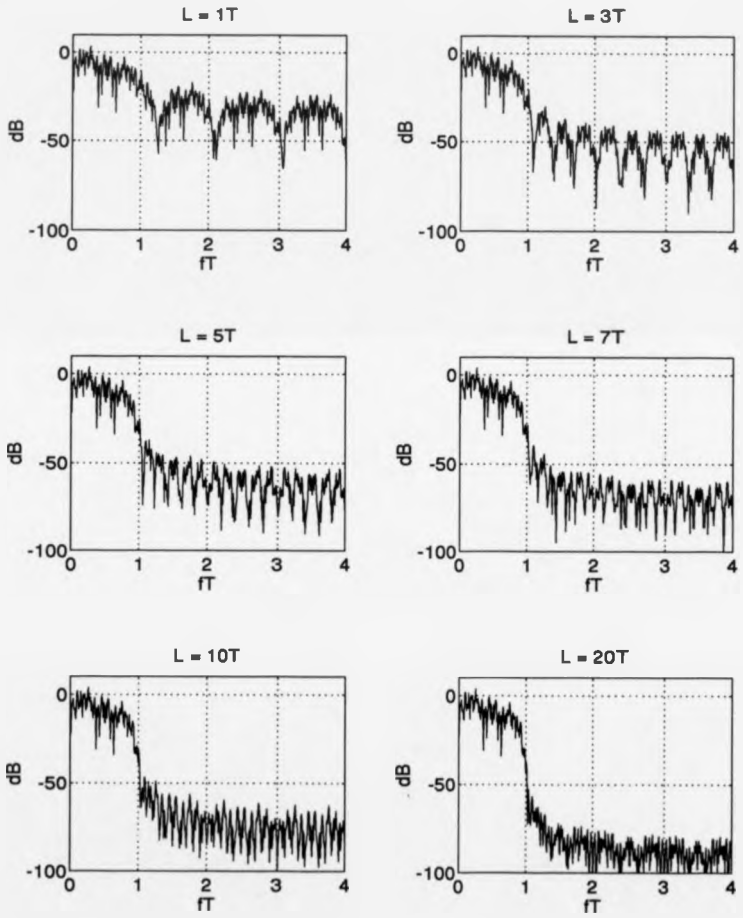


Figure 4.10: Baseband spectra with root raised cosine filtering of roll-off 1.0 and different truncation lengths

A series of filters has however been designed [107] with finite truncation of the impulse response as a major consideration. Filters with truncation lengths of 4 and 5 bits are described, which give far-out spectra superior to the case for truncated raised cosine filters. These filters are suitable for digital implementation techniques, such as the EPROM look-up system used in this work. Of the filters described in [107], the two showing the most promising spectral characteristics, labeled (a) and (c) in the paper, are defined by their impulse responses:

$$g(t) = 1 - 1.94\cos\omega t + 1.41\cos\omega t - 0.47\cos\omega t, 0 < t < 4T \tag{4.8}$$

$$g(t) = 1 - 2\cos\omega t + 1.823\cos\omega t - 0.823\cos\omega t, 0 < t < 5T \tag{4.9}$$

The filters are referred to as filter 1 and filter 2, respectively from now on. The above impulse responses are shown in fig 4.11. The $\pi/4$-DQPSK baseband spectra, after filtering with these filters, are shown in fig 4.12. The signal constellation diagrams which are theoretically found at the output of single filters (transmitter output), and cascaded filters (receiver output) are shown in fig 4.13.

It can be seen that for similar truncation lengths, the Kingsbury filters perform much better than the root raised cosine in respect to the sidelobe (ACI) levels. As with the root raised cosine filters, signals are free from ISI at the sampling instants after passing through two identical cascaded filters.

4.5 EPROM Look-Up Implementation of $\pi/4$-DQPSK filtering

The use of digital techniques permits very accurate filtering to be achieved without the practical limitations of analogue filtering (component tolerances etc). Using look-up techniques, digital modulation waveforms may be generated at baseband,

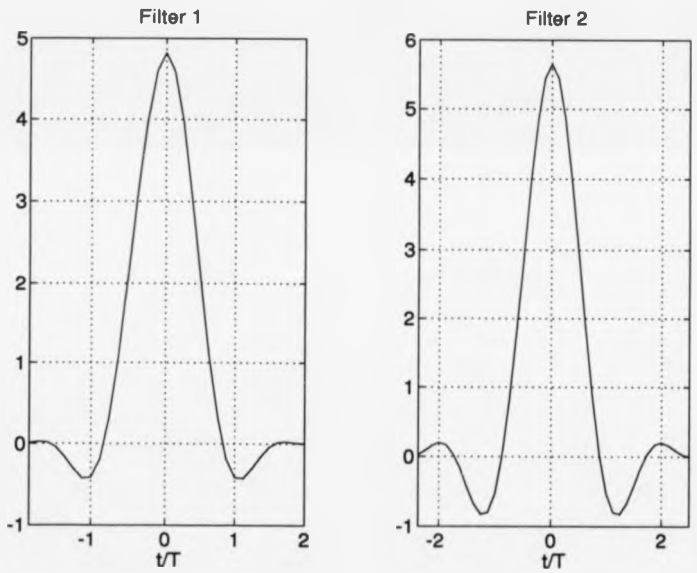


Figure 4.11: Impulse responses of Kingsbury filters

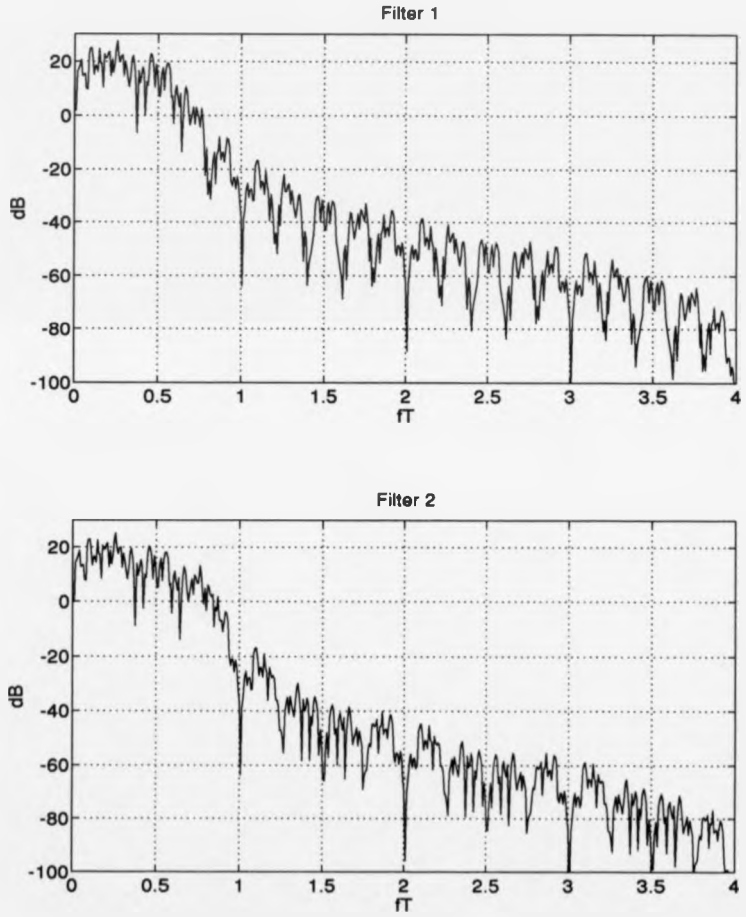


Figure 4.12: Baseband spectra for $\pi/4$-DQPSK filtered by Kingsbury filters

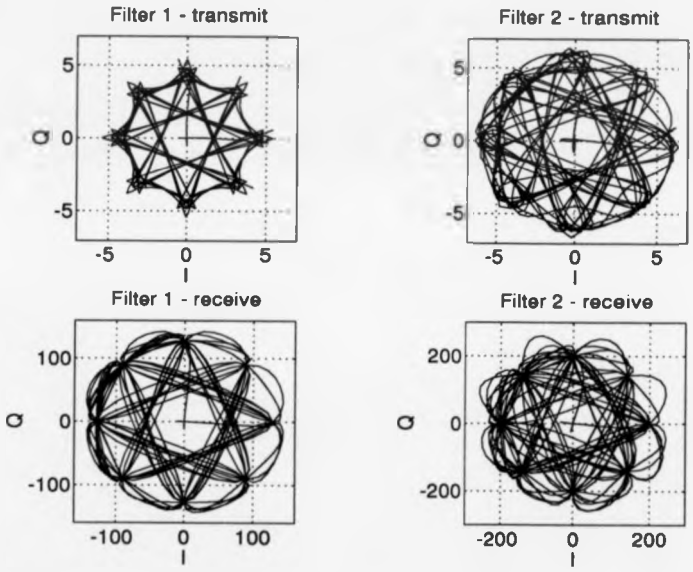


Figure 4.13: Signal constellation diagrams for $\pi/4$-DQPSK filtered with Kingsbury filter: upper traces - transmitter output, lower traces - output from two identical filters cascaded

or at an intermediate frequency, however hardware limitations can affect the latter, particularly when high data rates are being used.

Whichever type is to be used, the filtering process requires a large amount of memory. For $\pi/4$-DQPSK, in order to be able to calculate output waveforms for a given data sequence the following procedure is used: calculate initial phase location, the data sequence defines a certain sequence of phase shifts given by the $\pi/4$-DQPSK mapping process. However, when the baseband data is filtered, each data symbol will produce a response in neighbouring symbol periods, and hence an observation window is required, the length of which is defined by the filter truncation length. For example, if the truncation length, L, is five symbol periods, then to calculate the output waveform(s) for each input symbol, it is necessary to consider the two preceding symbols and the two proceeding symbols, together with the initial phase location. Since two bits form one symbol, this leads to an observation window of ten bits, together with a three-bit word defining the initial phase. In all, thirteen bits are needed in order to define the actual state of the $\pi/4$-DQPSK signal. To actually generate practical waveforms, a certain number of samples must be produced per symbol period: if M samples/symbol are needed, then an n-bit *interpolating* counter must be used, where n is $\log_2 M$, assuming M to be a power of two.

In the case of an envelop elimination and restoration (EER) system, it would be ideal to generate two waveforms by look-up techniques: the envelope of the rf waveform and its phase, in a limited form. The phase would be generated at an intermediate frequency and up-converted due to the high carrier frequencies used. Such an arrangement would be as shown in figure 4.14. It is found that the envelope waveform contains significant frequency components up to approximately $3f_s$, and therefore at least eight samples would be required per symbol, which could be achieved with three further address lines, totalling 16.

With the phase waveform, however, the situation is less practical. The (one-sided) 60 dB bandwidth of the limited signal is approximately $5f_s$. Ten samples would be required per symbol to reproduce this exactly. It also means that f_{IF}, the

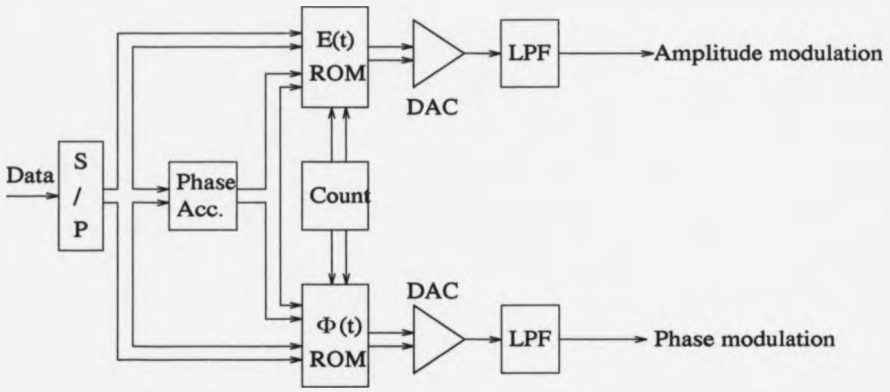


Figure 4.14: Implementation of $\pi/4$-DQPSK phase and amplitude generation using EPROM look-up

intermediate frequency at which the waveform is generated must be greater than $5f_s$ Hz, to ensure that the modulated signal components are all positive in frequency. However, even satisfying this, the IF needed must be greater for the following reason. Up-conversion to RF, or, more likely, another IF produces a two-sided spectrum, with components centred around $f_{IF2} \pm f_{IF1}$. If $f_{IF1} = 5f_s$, then the upper and lower sidebands will just meet. It is impractical, however, to remove one of these sidebands by filtering in this case, particularly with a filter of linear phase characteristics. The sidebands need to be separated by a certain amount in frequency before practical filtering becomes possible, and this raises the necessary frequency of the first IF. If, say, f_{IF1} is doubled to satisfy this, then maximum frequency components of $15f_s$ will be present in the first IF. Sampling is therefore necessary at a rate of at least $30f_s$ z. In actual fact, even this is unlikely to be sufficient, due to the fact that the modulation will not only appear around the desired IF, but also around all multiples of the sampling frequency, and it will be found that the lower sideband of the signal centred on the sampling frequency will near the upper sideband of the desired signal unless the sampling frequency is raised yet further. Therefore,

ultimately, a sample rate of $35 - 40f_s$ could be considered appropriate, to prevent implementation difficulties and performance degradation with analogue filtering.

Such an oversampling rate would require an extra 6 address lines, which added to the 13 already needed implies the use of a 4 Mb EPROM - currently a costly device. For ramping up/down considerations one extra address line is required for the approach taken, and this would lead to an 8 Mb EPROM, or two 4 Mb EPROMs, which either way, is not a practical proposition at present for cost reasons.

The above shows that using EPROM look-up, it is not yet practical to generate the (amplitude limited) phase waveform at an IF. Therefore, generating the baseband I and Q components of the $\pi/4$-DQPSK signal is the most suited to the technique. This is easily done using a 16-bit look-up system, which use two 512 kb EPROMs. Due to the burst-mode operation of the system an extra address line is required, as explained later, for ramping up/down purposes, and so 1 Mb EPROMS are needed.

4.5.1 Ramping Up/Down

In transmitting a burst-mode signal, the transmitter is effectively being turned on and off at high speed. If this switching is not done smoothly, then relatively high levels of adjacent channel interference may result. In burst-mode radio systems the switching rate is often very high, and hence this matter can be very important.

The problem is identical in nature to that arising from finite truncation of filter impulse responses.

If $s(t)$ is the time-domain representation of the signal under consideration then, by switching it on/off, multiplication by a (rectangular) window ($w(t)$) is effectively being performed. In the frequency domain, the resulting response is given by the convolution of $S(\omega)$ and $W(\omega)$, where these are the respective frequency domain representations.

The power spectral density of the resultant signal, $T(\omega)$, is given by

$$T(\omega) = \mid S(\omega) \otimes W(\omega) \mid^2 \tag{4.10}$$

One implication of the above is that as the burst duration shortens, so the spectral spreading will increase, as would be anticipated. In simple terms this is easily understood by considering that there are more on/off transients per unit time, and as it is these transients that contain the high-frequency energy, the relative magnitude of the high-frequency components (i.e. those which create adjacent channel interference) will increase accordingly. To be more accurate, the spectral spreading occurs only at the burst edges, when the transients occur. In this way, adjacent channel interference will tend to be of a burst nature, since during transmission of the burst data itself there will be no spreading effect. The *average* error rate due to burst mode operation will therefore depend on the proportion of time spent at burst edges, and consequently on the burst rate and duty cycle.

The method of implementation of the transmitter in this work is such that problems due to the burst mode operation can be reduced to a very low level. This follows from the fact that the EPROM look-up method strictly limits the filter truncation length, and as such there is very little energy in the transmit signal which lies outside of the actual baseband data sequence period. To transmit the signal in its entirety, hence eliminating the extra 'windowing' function, therefore involves only several symbol periods. In the case of a filter with a truncation length of 5 symbols, it is necessary to transmit filter pre- and post-cursors for a duration of only two symbol periods at the start and end of the burst, respectively. If this is accomplished without non-linearity, then there will be no spectral spreading resulting from the burst nature of the system. This is a very important result, and one which receives little comment in published work concerning TDMA-type radio systems.

It is interesting, however, to consider the situation where the filter 'tails' are not transmitted, and to observe the effect of spectral spreading. Simulation work using MATLAB was performed to demonstrate the spreading effect at the a burst rate

and duty cycle appropriate to the system. In fig 4.15 the result of this simulation is shown, and it should be noted that the lower spectrum, which shows the spreading effect, is an average spectrum over several bursts. As stated above, at the burst edges the transmit spectrum will exhibit far more spreading than this, and, of course, in the centre of the bursts the spectrum will not show any spreading.

Practical method of achieving ramping

When it comes to transmitting the signal energy at the beginning and end of each data burst, the question of practicality arises. To solve the problem, it was decided that the EPROM look-up circuitry should be able to operate in two modes - normal and ramp. The normal mode has already been described, and is not changed. The operation in ramp mode is now explained.

It became apparent that there would need to be an extra control line to the EPROM circuitry to set the mode of operation. Therefore, it was thought that if in normal mode the area of EPROM containing the burst data waveforms would be accessed and if in ramp mode the area containing ramp waveforms would be accessed. It would be a relatively simple task to produce a circuit which would generate either pre-cursors or post-cursors, but it was found that these techniques could not work for both. In the first case, data is being read in at one end of the shift registers, and in the other case it is being read out at the other end. Problems arise since EPROM address lines are restricted to binary levels, and therefore even when all the data to be transmitted has been read in, with continued operation the registers would begin to fill with logic '0's, but symbol '(0,0)' actually represents a phase shift, and so the system would continue as normal if an alternative technique wasn't found. This inability to have 'null' information was a large drawback.

After much consideration, it was decided that an approach that would be suitable, but as a compromise, was to define the symbol (0,0) as null information when in ramp mode; in this way, the data for EPROM storage would be similar, but symbol (0,0) would produce an impulse of zero magnitude.

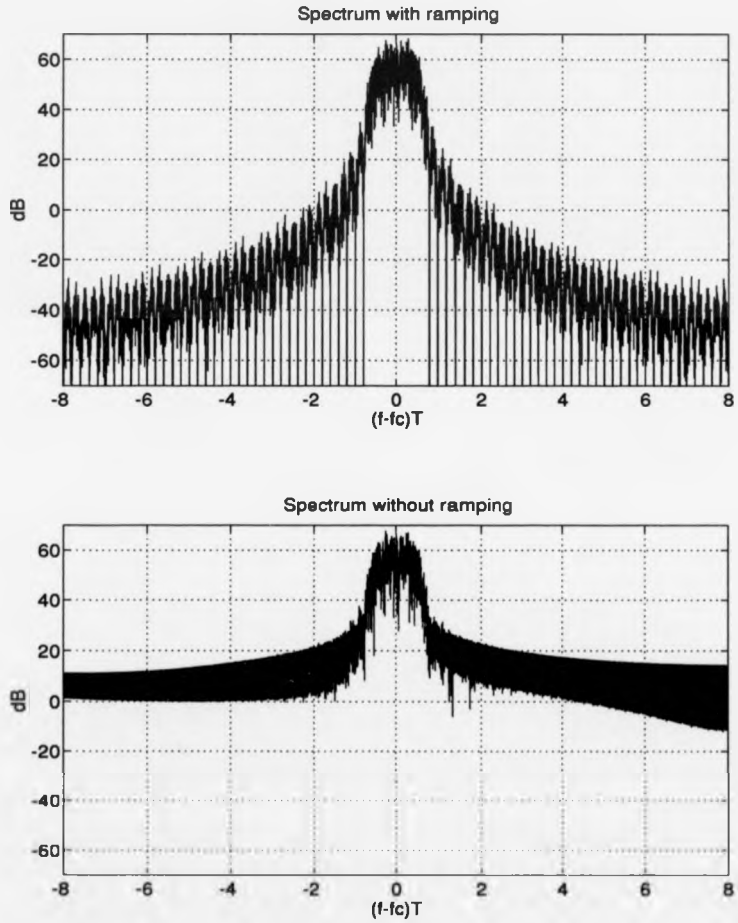


Figure 4.15: Spectra for $\pi/4$-QPSK transmitted with and without filter 'tails'

Chapter 5

Digital Modulator Implementation

5.1 Digital Modulator Implementation

Modulated (and unmodulated) carrier signals can be defined by the equation below:

$$s(t) = x(t)\cos\omega_c t + y(t)\sin\omega_c t \tag{5.1}$$

The terms $x(t)$ and $y(t)$ define the instantaneous point in signal space of the signal, referenced to the carrier. Modulated RF (or IF) signals can therefore be generated by a quadrature multiplication arrangement, with a summation of the outputs of the two multipliers. With unfiltered QPSK, $x(t)$ and $y(t)$ are ± 1, and the output $s(t)$ can take one of four positions in signal space, referenced to the carrier, corresponding to the possible combinations of $x(t)$ and $y(t)$. With DQPSK, $x(t)$ and $y(t)$ result from differential encoding of the baseband data stream, but as far as the modulation process is concerned, the situation is identical. The filtering of the data to be transmitted will generally spread the pulse duration in time, causing interference between adjacent pulses (ISI). If $h(t)$ is the pulse shape that is employed, which is defined as being symmetrical around the time axis (i.e. the centre of the pulse is at t=0) then the filtered baseband data can be described, at time t in the m-th symbol period by

$$x(mT + t) = \sum_{i=-\infty}^{\infty} a_{m+i} h(t - T/2 + iT) \tag{5.2}$$

and

$$y(mT + t) = \sum_{i=-\infty}^{\infty} b_{m+i} h(t - T/2 + iT) \tag{5.3}$$

where a_i and b_i represent the unfiltered baseband data, i is the symbol identifier, and T the symbol duration.

The generation of $x(t)$ and $y(t)$ above may be achieved digitally, with benefits over the corresponding analogue approach. In doing this, the above waveforms are produced in a sampled, rather than continuous, form. Since the waveforms are

bandlimited, they can be reproduced exactly from a succession of samples as long as the number of samples per symbol is sufficient (i.e. sampling rate to be a minimum of twice the highest frequency component of the waveform).

The equations defining the sample values with time are as above, with t replaced by $n\tau$, where n is the sample number within the current symbol period, and τ the sample spacing.

5.1.1 Digital vs analogue approaches

The use of digital generation of the filtered data does have important practical advantages over corresponding analogue filtering, which makes it generally far more suitable in the implementation of data transmission equipment.

Analogue filtering suffers from accuracy and reproducibility problems. Finite component tolerances place a fundamental limit on the results achievable. In addition, it is very difficult to realise filters which have excellent phase and amplitude responses, as is required in the distortion-free transmission of data.

By using digital techniques, however, many of theses problems are overcome. Circuits are extremely reproducible, and it is not difficult to implement high performance filters. The associated cost is now low, and decreasing all the time. Accuracy is limited in part by the quantization of the waveform samples, however the distortion resulting from this can be made to very small indeed. Additionally the linearity of the D/A conversion will in practice not be perfect, but again the resulting errors will be small. While the digital technique does lead to the generation of unwanted outputs (see later), the majority of these are sufficiently removed from the wanted output(s) in frequency to be easily removed by simple analogue filters.

One major advantage of using a digital technique is that it can be simple to change from one filter type to another, or one modulation scheme to another. This can be achieved by simply switching to a different section of memory, or by reprogramming, for example. Such versatility is not possible with analogue hardware.

There are two principal methods of performing this digital generation - the DSP-based method (that employing a Digital Signal Processor microprocessor) and the table look-up method. The DSP approach calculates the instantaneous values for filtered data by a succession of multiply-and-add routines, having a sampled version of the pulse shape stored in ROM. Using this technique and contempory DSP chips, practical filtering of data at rates of at least several megahertz is possible [104]. To give a relevant example of the process, with a baseband data rate of 80 kb/s, equating to a 40 ksymbol/s rate in QPSK, 8 samples used per symbol, and a pulse length of 5 symbol periods, a multiply-and-add must be performed every 625 ns. In reality the rate would be higher than that since a larger pulse truncation length would be used to gain the advantages of DSP. A disadvantage of DSP is that they can consume significant power when operating at high speeds.

The table look-up technique can offer simplicity at the expense of possibly large memory requirements. By pre-calculating the responses $x(t)$ and $y(t)$ due to all possible baseband data streams of a given length, the required samples can be read out of memory when that memory is addressed by any arbitrary data sequence of the same length. The length of the addressing data sequence is set by the pulse truncation length employed. This technique is constrained by the size of the memory (or, number of address lines), and the access time of the memory.

These constraints can degrade the filtering performance achievable, due, for example, to limited pulse truncation lengths. For this work, a truncation length of 5 symbols would be the maximum allowed practically, which places restrictions on the filters usable (see Chapter 4 for details). In particular, the use of root raised cosine filters with very low roll-off factors is not suitable due to the large impulse response lengths; the truncation of such responses leads to significant spectral spreading.

5.2 An EPROM look-up source of $\pi/4$-DQPSK I and Q data

For this work, it was decided that the look-up approach was suitable as the source of the I and Q data for a $\pi/4$-DQPSK signal. It offers a relatively simple solution to the problem of precise filtering of the baseband data, with low power requirements and relatively low cost. Despite the relative simplicity, the performance of such a system can be very good. It is anticipated, though, that DSP may be used at a later stage in the development of the system. A conceptual diagram of a $\pi/4-$DQPSK quadrature baseband components generator is shown in fig 5.1.

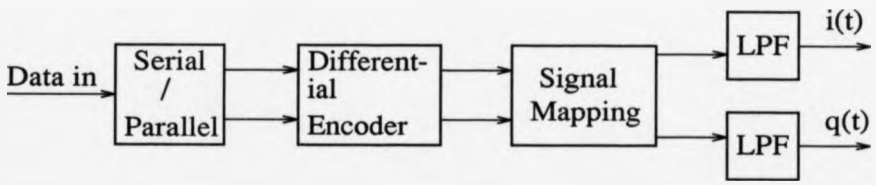


Figure 5.1: $\pi/4$-DQPSK baseband generator

The source generator feeds a serial-to-parallel converter, which produces parallel streams, denoted u and v. Differential encoding is next performed, and a signal mapping function converts to the required x and y data. This data is then filtered, ideally with a Nyquist filter, to prevent ISI, and the filtered outputs drive a quadrature modulator. [67, 70]

In more detail, the process of generating the filtered data, $x(t)$ and $y(t)$, is now described.

The binary source data (NRZ) feeds the serial-to-parallel converter. The data is grouped into 2-bit words (symbols), and the four possible symbols translate to phase-shifts as in table 5.1.

Since the scheme is differential, these phase-shifts represent the change in phase

$\pi/4$-DQPSK Mapping Scheme	
Information Symbol	Phase Shift
1 1	$\pi/4$
-1 1	$3\pi/4$
-1 -1	$-3\pi/4$
1 -1	$-\pi/4$

Table 5.1: $\pi/4$-DQPSK phase mapping

of the carrier over the appropriate symbol period. Representing the phase-shifts in symbol k as θ_k, the following mapping can be applied to generate the coefficients u and v.

$$u_k = u_{k-1}\cos\theta_k - v_{k-1}\sin\theta_k \tag{5.4}$$

and

$$v_k = u_{k-1}\sin\theta_k - v_{k-1}\cos\theta_k \tag{5.5}$$

The terms u and v can take levels of ± 1, $\pm\frac{1}{\sqrt{2}}$, and 0. The resulting eye diagrams for the unfiltered quadrature data is shown in fig 5.2. From the diagram, it can be seen that the eye alternates from 2-level to 3-level from symbol to symbol. Additionally, while the I data is in a 3 level state, the Q data will be in a 2 level state, and vice versa.

Upon filtering with a Nyquist filter, to preserve the ISI-free condition, the eye diagram takes on the same values at the centre of each symbol period (sampling instant), but the transitions between these points in time are greatly smoothed, reducing the bandwidth of the resultant signal. The filtering will be divided between transmitter and receiver, ideally, and hence ISI will actually be noted at the transmitter output, as it will only be after filtering in the receiver that the ISI-free signal is produced (ignoring multipath effects). The block diagram of the practical transmitter is shown in fig. 5.3.

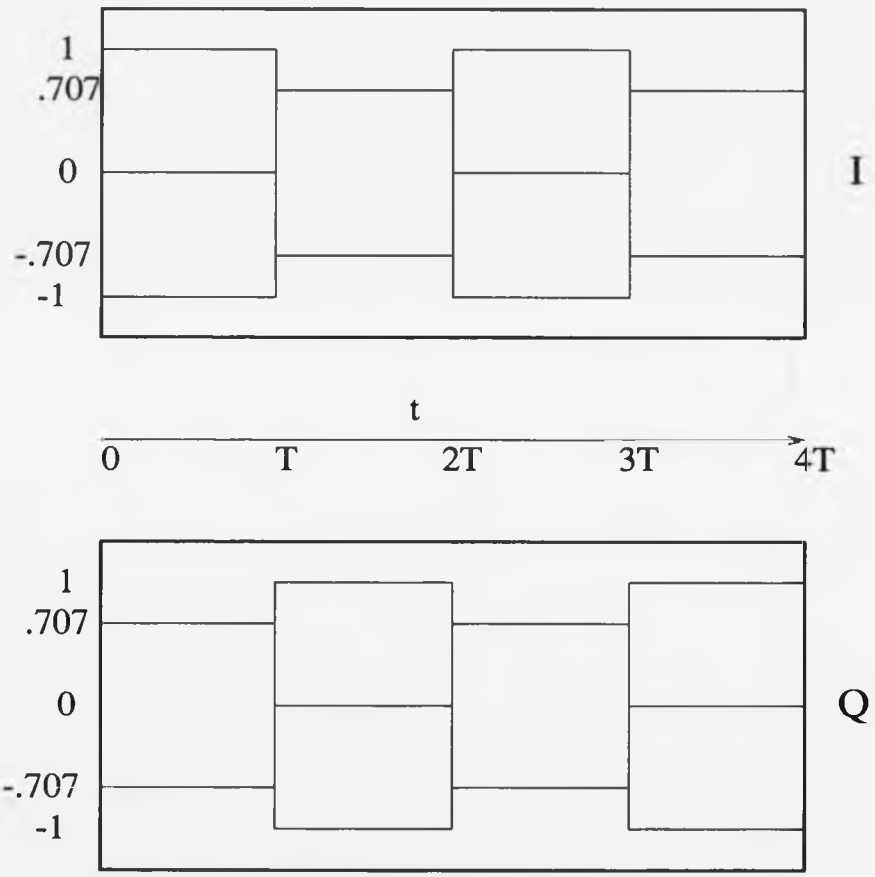


Figure 5.2: $\pi/4$-DQPSK I and Q eye diagrams for rectangular baseband data

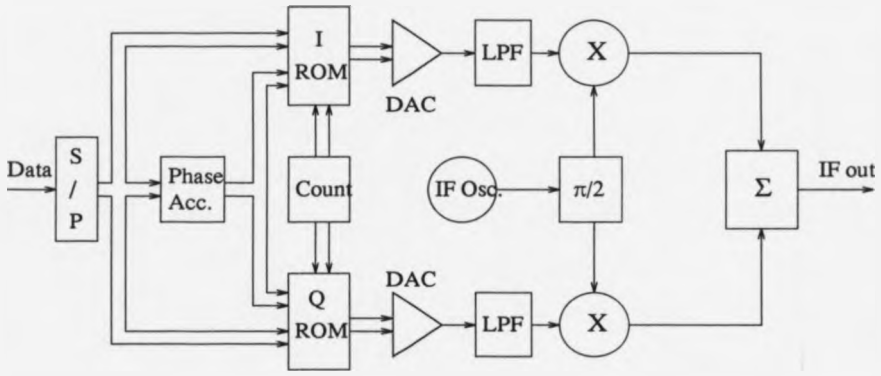


Figure 5.3: $\pi/4$-DQPSK modulator implementation using EPROM look-up

The data to be transmitted is read serially into shift registers (serial-to-parallel conversion). Since filter truncation to 5 symbols is used, ten bits are required as address lines for the I and Q EPROMs. As the modulation scheme is differential in nature, it is necessary to have the starting phase for each symbol block. This phase is updated every symbol period, and this is accomplished by some basic logic circuitry, described later. The 8 possible phase locations in the unfiltered $\pi/4$-DQPSK scheme require 3 bits for identification. Finally, since a number of samples are required per symbol period, an interpolation counter is used. This, for an 8 sample per symbol rate, requires 3 bits. In total, therefore, unique samples are addressed by a 16-bit word. For the ramping considerations 17 bits are used in practice, and this is explained in the later section on ramping.

The outputs from the EPROMs are latched to prevent transients when the output lines change state, and the 8 bit outputs are then converted into analogue form by DACs, before being filtered and feeding the respective quadrature modulators. The outputs of the modulators are then summed, so generating the $\pi/4$-DQPSK signal at IF or RF, as desired.

5.2.1 Distortion effects in the look-up technique

Apart from non-ideal filtering due to time-truncation, there are other distortion effects caused by the digital nature of the technique. The two effects which are most important are quantization noise, and noise resulting from the sampling process [102]. The outputs resulting from the sampling occur as modulation around multiples of the sampling frequency, and are therefore quite easily removed by filtering at the output of the DAC. Quantization of the digitally generated signals has to be accepted, since each sample will be represented by a finite number of bits. A general rule [102] predicts non-harmonic spurious components at a level of $6n$ dB below a carrier generated by digital synthesis, where n is the number of bits used. Hence, with an 8-bit representation, spurii would be approximately 48 dB below the wanted output. A diagram of a typical output spectrum from a digital synthesizer is shown in fig 5.4. This illustrates the distortion effects due to the sampling and quantization processes. The effects of filter truncation, which are often considered together with these effects [106], were shown in Chapter 4.

5.2.2 Differential phase encoder implementation

In unfiltered $\pi/4$-DQPSK, there are eight possible phase states. These are shown in fig 5.5, each represented by a 3-bit word. The figure also shows the relationship between input symbol and phase shift; each phase shift can be described in terms of an incremental 3-bit word, which, when added to an instantaneous phase state, will give the phase state in the following symbol period. A hardware implementation is shown which performs the required function. Here, exclusive-OR gates convert the symbol (a_k, b_k) into the phase increment word, (c_k, d_k, e_k). In symbol period k, (c_k, d_k, e_k) is modulo-8 added with the 3-bit phase state in symbol $k-1$ to give phase state k. A latch arrangement ensures that the phase state of the previous symbol is locked into the adder input until the next symbol period, when the information is updated.

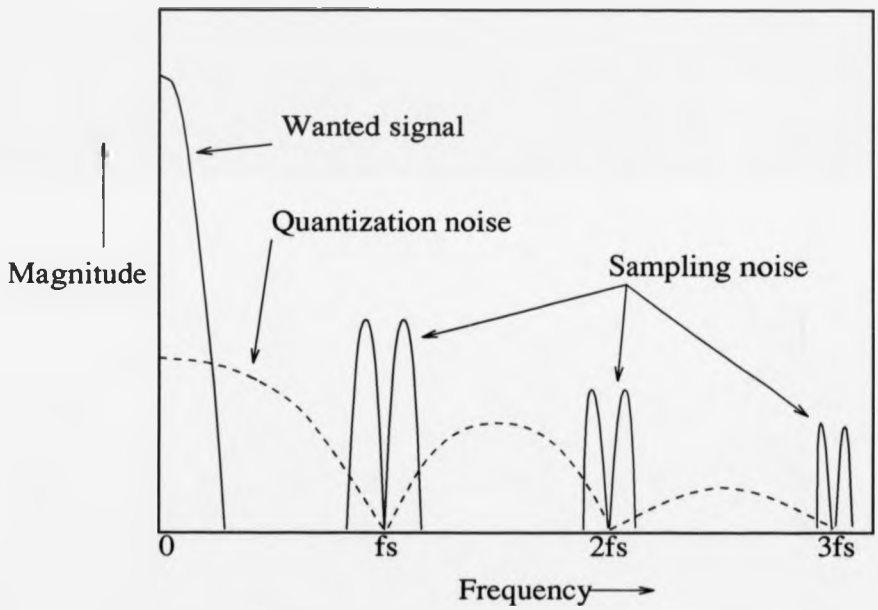
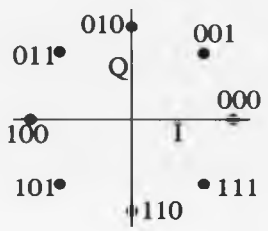


Figure 5.4: Diagram showing typical output spectrum from digital synthesis circuitry

Symbol	$\Delta\phi$	Increment
1 1	$\pi/4$	001
0 1	$3\pi/4$	011
0 0	$-3\pi/4$	101
1 0	$-\pi/4$	111
a b		c d e



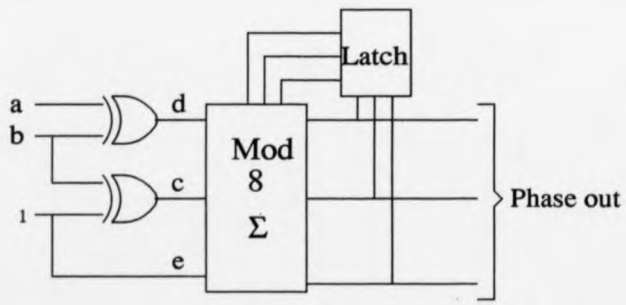


Figure 5.5: $\pi/4$-DQPSK differential phase encoder

An example of the operation follows. If the $\pi/4$-DQPSK signal is at phase location (0 1 1) and symbol (1 1) is input, a code increment (0 0 1) is produced by the encoder. 3-bit addition of (0 1 1) with (0 0 1) results in (1 0 0), and it can be seen from the diagram of phase locations that the new phase state is indeed that state which is reached by a phase rotation of $\pi/4$, as required by the mapping code. From the phase state word, the quadrature coefficients, u and v, can be produced by look-up techniques.

5.2.3 Generation of filtered coefficients

Practically, filtering to a truncation length of 5 symbol periods is used, for reasons given in Chapter 4. This means that to generate the quadrature coefficients in symbol period k, it is necessary to observe symbols $k-2, k-1, k, k+1$ and $k+2$, in addition to having knowledge of the initial phase state. The actual process used to generate the coefficients for later storage in ROM is as follows:

- Read initial phase state
- Read input data (5 symbols = 10 bits in serial form)
- Generate differential phase shifts from input symbols
- Generate coefficients u, v (5-level) from phase information by $\pi/4$-DQPSK mapping process
- Generate filtered u, v by convolving filter impulse response with series of impulses u, v. Impulses spaced by 7 '0's to give 8 samples per symbol.
- Take appropriate 8 samples from filtered u, v streams.

When generated in this way, each I or Q sample is represented by a 16-bit word. The structure of this word is shown in fig 5.6.

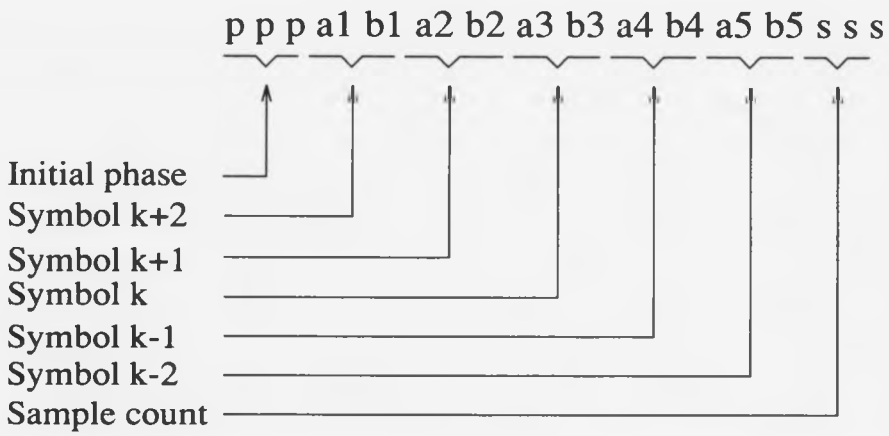


Figure 5.6: Address word structure

5.2.4 Generation of ramping coefficients

As stated in chapter 4, it was decided that ramping of the data streams would be achieved by assigning the symbol (0,0) to be a null symbol. An extra address line (and hence twice the EPROM memory) is required, to divide the EPROM operation into 'normal' and 'ramp'. With (0,0) representing zero pulse energy in ramp mode, baseband coefficients are calculated and stored in the same way as with normal operation.

Testing ramp operation

In order to test the ramping operation, it was necessary to make some changes to the test board. Some timing circuitry was added to control the burst rate, and set the transmitter ramp state high or low. When a burst begins, a clock-derived line holds the state in ramp mode for 5 symbol periods, to allow ramp-up. While in this state, the (0,0) symbol is forbidden, as a transient would be caused when the system goes into normal operation due to the symbol suddenly representing energy

in that mode. To prevent this from occurring (for testing purposes only), the I-line is tied high while in ramp mode. Ramp-down mode was not tested thoroughly. Since the same process will occur as that during ramp-up, only in reverse, spectral spreading effects will be identical. The hardware implementation requires a control line indicating that ramp down should begin, and it was felt that the extra work in realising this for testing purposes was not warranted.

5.2.5 Quadrature modulator

Having generated the I and Q filtered baseband components, it is next necessary to feed them to a quadrature modulator to produce the $\pi/4$-DQPSK signal at an intermediate frequency. A diagram showing the implementation of the quadrature modulator is fig 5.7.

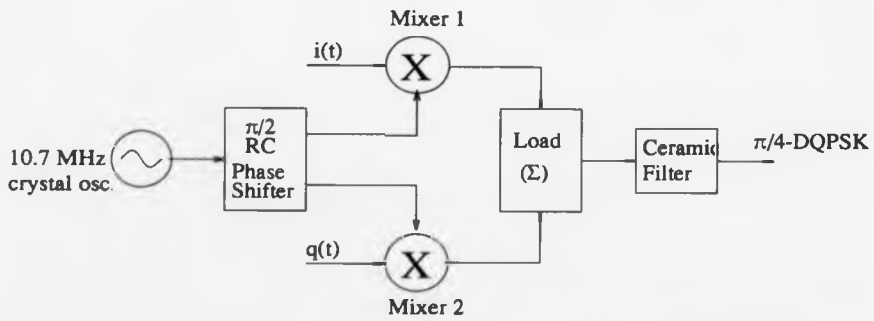


Figure 5.7: $\pi/4$-DQPSK quadrature modulator implementation

Here, $x(t)$ and $y(t)$, the filtered I and Q components feed two separate mixers, of the double balanced bipolar transistor type, together with the 10.7 MHz local oscillator, of required phase. The $\pi/4$ phase shifter, on the output of the 10.7 MHz crystal oscillator, using simple R/C circuitry, generates the two local oscillator signals (sine and cosine). The outputs of the mixers are summed in a common load (low impedance in comparison with mixer output impedance) to give the modulated

IF signal. A ceramic IF filter removes unwanted mixing products, including spurii on the digitally synthesized baseband signals. The bandwidth of this filter is set to be greater than the bandwidth of the generating signal since it is assumed that the group delay effects of the filter will be significant at in the vicinity of the cut-off frequencies, thus these points should be far enough removed from the signal frequencies to prevent phase distortion effects. With high-quality (and expensive) IF filters which can offer a satisfactory group delay response this factor would be less of an issue.

Throughout this work, ideal quadrature modulation is assumed. It should be noted, however, that errors in the quadrature local oscillator signal phases will result in distortion effects, including *crosstalk* between I and Q channels. The results of such errors have been considered in [113], and the results demonstrate the importance of minimising such phase errors. The severity of the distortion produced by these errors is dependent upon the type of power amplifier used. A linear PA will tend not to produce further signal degradation, however a non-linear PA can introduce significant distortion. This was found to be the case in [131] where phase errors in a GSM system employing quadrature modulation were shown to be the source of significant spectral spreading when the signal was non-linearly amplified. As the amplification scheme used in this work is linear, then such spectral spreading would be minimized.

The simple RC phase shift network used here was found to work extremely well, and it is anticipated that there ought to be little problem reproducing the circuitry. As with all analogue circuitry the result will depend on component tolerances and circuit construction effects, but with care in the choice of components and the circuit lay-out there should be little problem.

5.2.6 Practical results

The look-up and quadrature modulator circuits all worked well. Fig 5.8 shows a sequence of filtered data on the output of the I-channel. The spectrum of the data stream is shown in fig 5.9. In addition, a spectrum of the $\pi/4$-DQPSK produced by the quadrature modulation process is shown in fig 5.10.

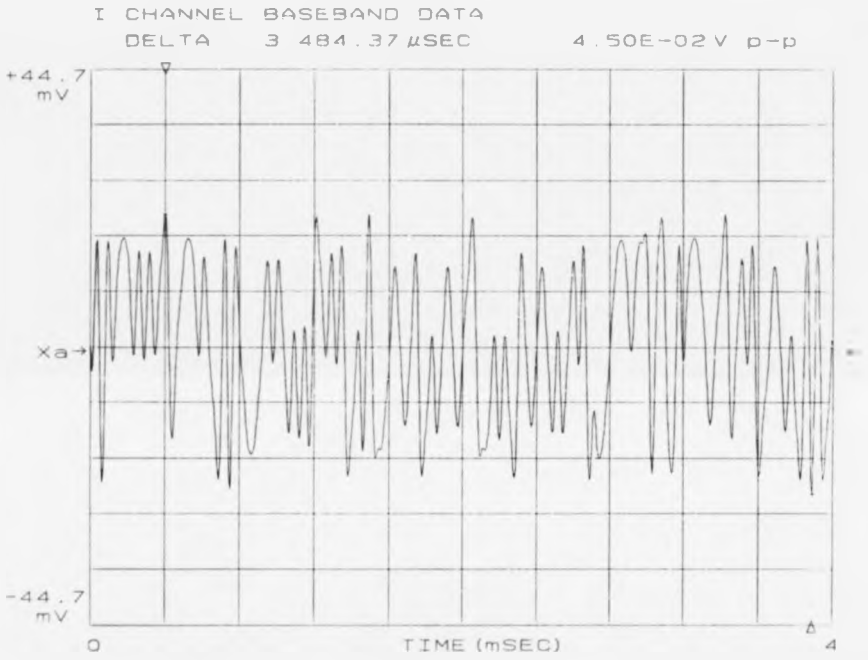


Figure 5.8: I-channel data generated by look-up technique

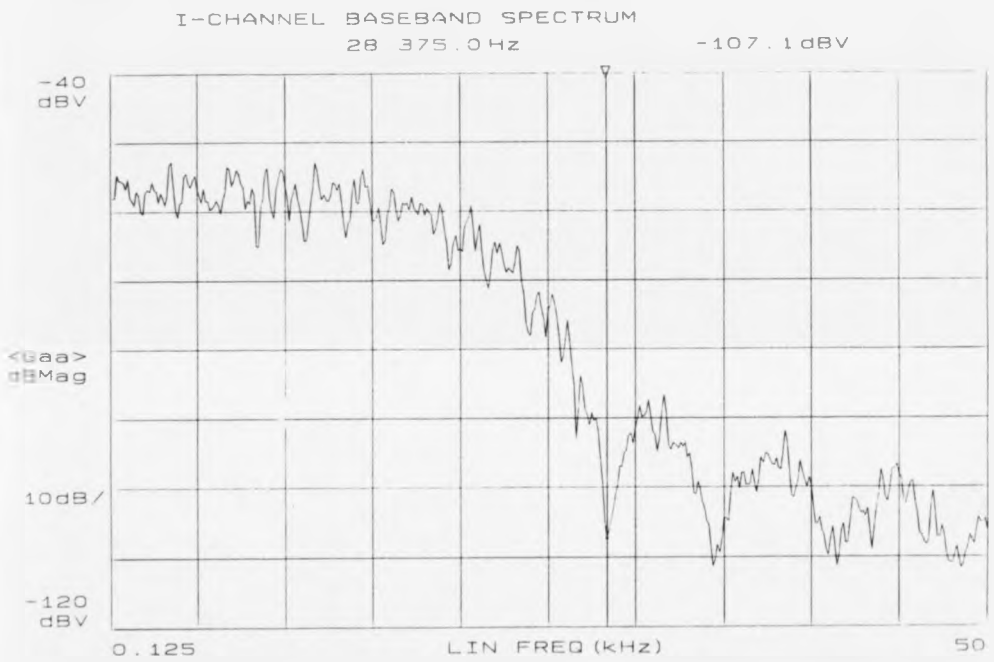


Figure 5.9: Spectrum of I-channel data generated by look-up technique

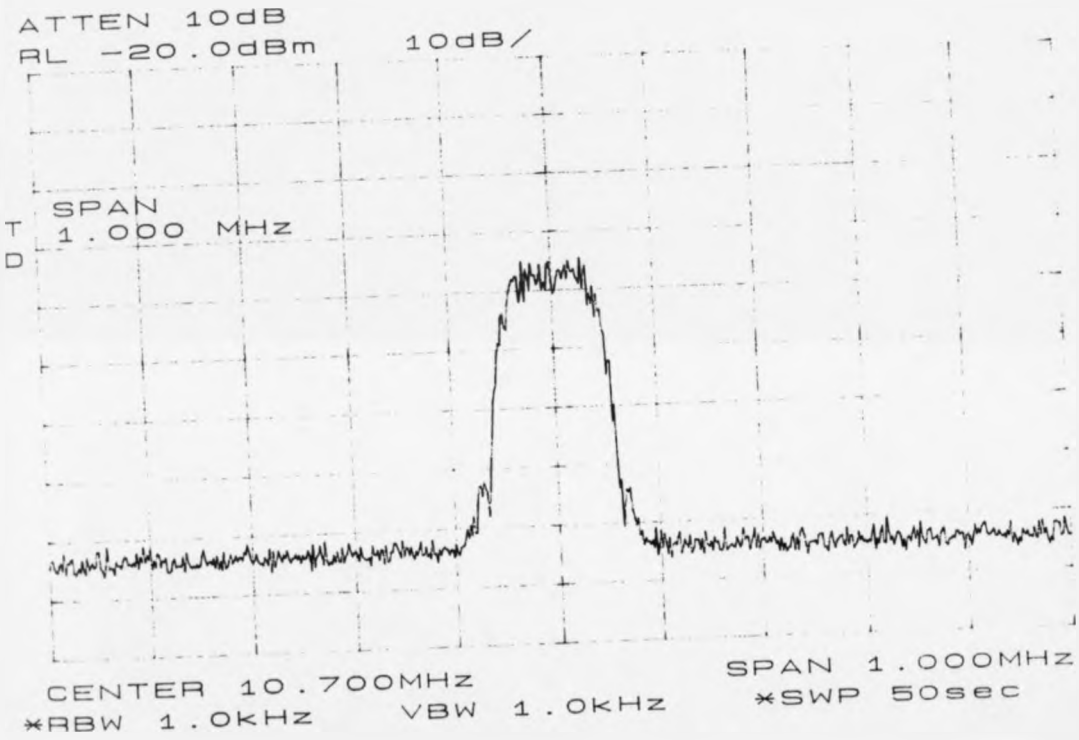


Figure 5.10: Spectrum of $\pi/4$-DQPSK generated by quadrature modulation

Chapter 6

Envelope Elimination and Restoration Transmitter Implementation

6.1 Introduction

This chapter describes the RF hardware which takes in the $\pi/4$-DQPSK signal at 10.7 MHz, converts this to the required carrier frequency (53 MHz) and amplifies to a practical power level for transmission. Methods of achieving the final amplification at high efficiency are discussed.

For this application, if high efficiency amplification of the linear signal is not achievable, then the use of a linear modulation scheme would not be appropriate, as its disadvantages would outweigh those which result from the use a non-linear scheme, e.g. GMSK.

6.1.1 Modes of nonlinear power amplifier operation

High efficiency power amplifiers (PAs) generally operate in a mode of type class C, class C-E, class E or class F [132]. The basic difference between these is that of the output network configuration.

Efficiency considerations

Apart from losses in components in power amplifier output networks, which are usually small anyway, the main concern in amplifier efficiency is dissipation within the active device. Dissipation occurs when both the voltage across the device and the current it conducts are non-zero, and it is simply defined by the product VI. Consequently to achieve high efficiency it is desirable to ensure that when the voltage across the device is high the current flowing is low, and vice versa. This is the principle behind all high-efficiency circuit configurations, and output networks are used to force this situation, in conjunction with appropriate drive conditions.

Class C operation

In 'pure' class C operation, the output network consists of a parallel resonant circuit, tuned to the fundamental of the carrier. This results in an almost sinusoidal drain (collector) voltage waveform, as the parallel resonant circuit appears as a low impedance to harmonics. At resonance, the transistor conducts for a short time, and this conduction occurs when the voltage waveform is at minimum, hence the dissipation in the device is low and overall efficiency high. Saturation if it does occur, can only be for a very short duty cycle, and this implies very large values of current for reasonably high output powers to be produced, due to the low voltage levels used in solid state circuits. In addition, this mode of operation is rarely employed in its pure form in solid-state circuits at high frequencies due both to device characteristics (non-linear capacitances in particular), and the impractical component values of the output tuned circuit at the low impedance levels used [132]. Instead, one of the following modes below is usually preferred.

Class C-E operation

Class C-E, or *mixed mode* operation, is that which is generally used in the amplification of constant-envelope VHF/UHF signals [93, 132]. A series resonant output network is used, and this permits harmonic voltages to be present on the drain. The amplifier is driven hard enough to ensure that saturation occurs, and this together with the effect of the output network permits the drain voltage to be low and relatively flat for a substantial fraction of the cycle. High efficiencies can result, and figures of 70-80 % are common.

Class E operation

When very high efficiency is sought, the class E mode of operation can be used. In this mode, the output network is designed to ensure that at the moment the PA device begins to conduct the drain voltage had dropped to zero and is flat, and

remains at that level until conduction ceases. Ideally this can give rise to an efficiency of 100 %, however circuit imperfections such as finite device saturation resistance reduce this slightly. Analysis of this amplifier configuration is well documented, under both optimal and non-optimal load conditions [87, 88, 89, 91, 96, 97, 98, 99, 125, 128].

Class F operation

Class F operation uses parallel resonant circuits in the load to prevent harmonic currents (typically 3rd and 5th harmonics) from flowing into the load network [132]. This produces a 'squaring' of the drain waveform, and if the current pulses occur when the waveform is at its minimum, then the efficiency can be high. The effect can be produced by having a parallel resonant circuit tuned to the carrier frequency situated at the end of a quarter wave transmission line, the other end of which connects to the drain. Impedance transformation by the line produces, at the drain, short circuits at even harmonics, and open circuits at odd harmonics, thereby permitting a square wave voltage.

6.2 Power amplifier linearization techniques

Filtered $\pi/4$-DQPSK contains substantial envelope fluctuation, and therefore requires linear reproduction of this envelope if spectral spreading is to be avoided. Standard linear amplifiers (class A/B) are generally operated in a power inefficient mode, particularly when amplifying signals which spend a significant proportion of time substantially below their peak level. Such amplifiers can be driven into saturation, becoming class E, or other, and efficiency will rise significantly, at the expense of linearity, of course. A power efficient, and therefore non-linear, mode of operation is desirable for this system, due to power supply constraints, however the envelope fluctuation is needed for spectral reasons.

Several techniques have been developed as means of linearizing non-linear (and

power efficient) power amplifiers, in order to satisfy the demands of power and spectrum efficiency. These are now considered.

6.2.1 Linear amplification using non-linear components (LINC)

Originally described by Cox [112], LINC operates as follows. A bandpass input signal (containing amplitude and/or phase modulation) is split into two constant amplitude signals, which have the phase relationship such that when summed they re-form the original signal. Mathematically, this can be described by

$$S(t) = E(t)\cos\left(\omega_c t + \phi(t)\right) \tag{6.1}$$

$$= S_1(t) + S_2(t) \tag{6.2}$$

$$\tag{6.3}$$

where

$$S_1(t) = \frac{E_{max}}{2}\cos[\omega_c t + \phi(t) + \alpha(t)] \tag{6.4}$$

$$S_2(t) = \frac{E_{max}}{2}\cos[\omega_c t + \phi(t) - \alpha(t)] \tag{6.5}$$

$$\alpha(t) = \cos^{-1}[E(t)/E_{max}] \tag{6.6}$$

$$\tag{6.7}$$

with $S(t)$ being the input signal and $S_1(t)$ and $S_2(t)$ the two resulting constant envelope signals.

These signals, being of constant envelope, can be amplified by separate non-linear, and power efficient, amplifiers, without significant distortion. When amplified to a suitable level in this way, they are then combined, and the result is ideally a high-power replica of the input signal. A block diagram of the transmitter is shown

in fig 6.1. As far as power efficiency is concerned, LINC suffers from the obvious penalty of requiring a hybrid output combiner, which produce a loss of 3 dB, hence halving the transmitter power efficiency.

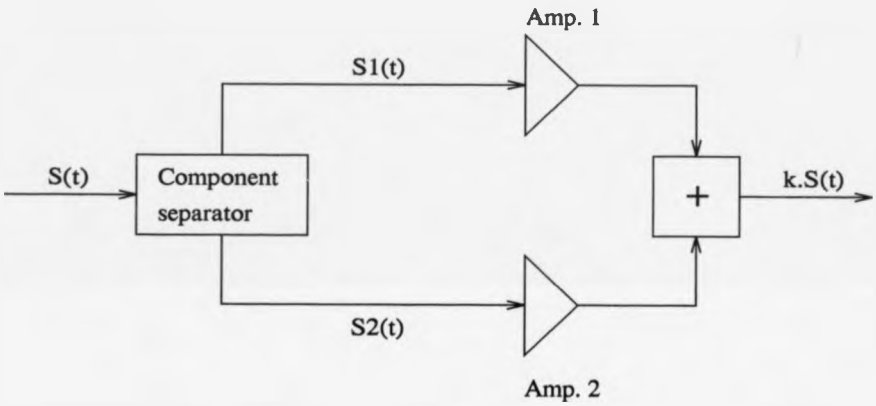


Figure 6.1: Diagram illustrating LINC transmitter implementation

Practically, ideal operation of such a transmitter is very difficult to obtain. Generation of the two constant envelope signals from the given input signal using analogue techniques was a major problem, but recent advances in DSP techniques are overcoming this [117]. The problem of maintaining phase and amplitude match of the RF stages is problematic, particularly when the carrier frequency changes, and small phase errors can give rise to significant spectral spreading [101]. A method of correcting such phase errors by means of feedback has been shown [110], however the system complexity is large for the power efficiency obtained (21 % with OQPSK, α =1).

6.2.2 Cartesian feedback

The cartesian feedback scheme uses quadrature demodulation of the transmitter output, comparison of the results with the baseband quadrature components and

adjustment of the latter as necessary to ensure that the output signal is a reliable reproduction of the input, as shown in fig 6.2.

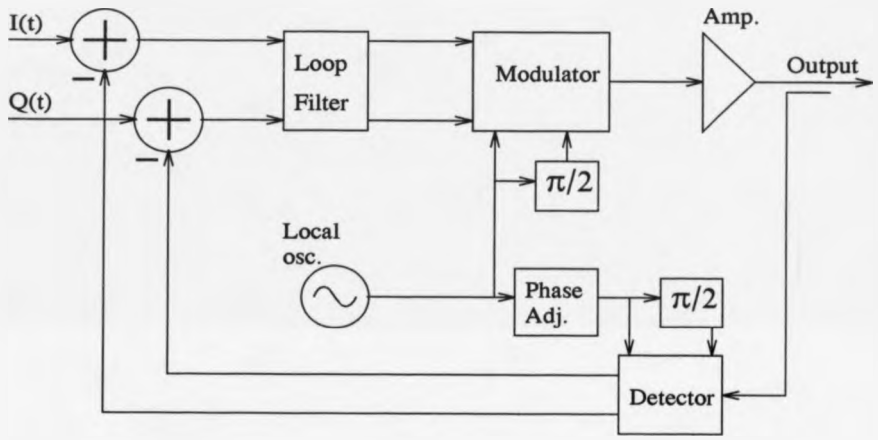


Figure 6.2: Diagram illustrating cartesian feedback implementation

This scheme permits the PA to be operated in a non-linear mode, however it is necessary to avoid saturation at all times except when the output signal is at its maximum value, as attempts to raise the output amplitude by increasing the input amplitude would fail if the device were in saturation. As the highest power efficiencies in non-linear amplifiers occur in a state of saturation, this scheme suffers from the disadvantage of having limited power efficiency (typically less than 40 %), although of course the resulting efficiency may still be much higher than that of traditional linear amplifiers.

In addition to the efficiency considerations, successful operation of the scheme is critically dependent upon precise adjustment of the feedback loop [115], and errors in this can easily lead to system instabilities. Operation over wide bandwidths is not possible, and it is generally accepted that the scheme is suitable only for amplification of a single channel, or a closely spaced group of narrowband channels

[116]. Any errors in the quadrature demodulation process, including the introduction of noise and distortion products, transfer directly to the transmitter output, which is a further drawback with the scheme. There is obviously a large increase in system complexity with this scheme, since the feedback network is essentially an entire receiver, which, together with the other drawbacks, makes it unattractive for use in this system.

6.2.3 Pre-distortion

Transmitter linearization using pre-distortion attempts to compensate for transmitter non-linearity by pre-calculating the input signal required to produce the given output signal, and using this signal as the transmitter input. This can be achieved using analogue or digital means, however the analogue approach is limited to compensating for only low (usually 3rd) order distortion products. Digital approaches usually employ look-up techniques. Such schemes are obviously sensitive to temporal changes in transmitter operating conditions, such as temperature changes, or changes in operating frequency.

Consequently, adaptive schemes are being increasingly considered. These require relatively large amounts of memory storage, and can have relatively long adaption times [103], however, and so require efficient DSP-type implementation, with its associated additional complexity. Also, extra hardware is needed to monitor the output signal, or distortion, and this must be highly linear itself. Digital pre-distortion suffers from bandwidth limitations due to the finite processing speed used; for high order distortion products to be compensated, high IF bandwidths are needed, and the original digital signal must be over-sampled accordingly.

As with cartesian feedback, the amplifier must be operated in a mode sufficiently backed-off from saturation that saturation is only reached (if at all) at maximum power output, so, again, average power efficiency is limited, and decreases with increasing peak-mean ratio of the input signal.

6.2.4 Envelope elimination and restoration

Envelope elimination and restoration, EER, as a means of transmitter linearization, was introduced by Kahn in the 1950s. It was used in SSB transmission using a class-C PA operating at high efficiency. The concept of EER is shown in figure 6.3. The signal to be transmitted is first split into two streams - the amplitude stream and the phase stream. In the phase stream, the signal is first limited, amplified and then used to drive the PA. This signal contains the phase of the original signal, which ideally is not affected by the limiting process. In the amplitude stream, envelope detection is used and the resulting signal is amplified and, via a modulator, controls the supply voltage to the PA. Assuming a linear supply voltage to output voltage characteristic for the PA, which is reasonable when the amplifier is operating in a saturating (and hence power efficient) mode, the original envelope is restored to the signal in the process.

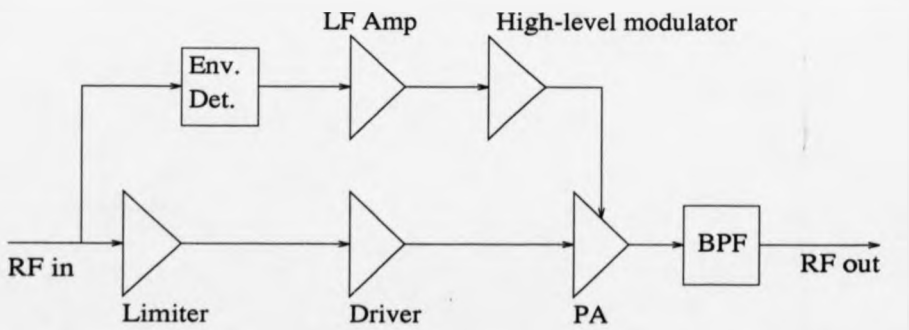


Figure 6.3: Diagram illustrating envelope elimination and restoration implementation

By the use of a high-efficiency modulator stage, the overall power efficiency can, in practice, be greater than 50 % [100], which makes the scheme very attractive for this system, and others which have power supply limitations.

The operation of the PA in this scheme can be contrasted to that occurring in the

cartesian feedback technique: in EER, a saturated mode of operation is a necessary condition for linearization, whereas it is virtually prohibited in the latter.

From the above, it is clear that EER is an appropriate choice of linearization scheme for this project. It offers a relatively simple solution to the problem of linearization, certainly far simpler than the other schemes described. The extra system complexity resulting is not large, being basically an envelope detector, LF amplifier and modulator, which has both size and cost advantages over the more complex schemes.

6.2.5 PA high-level modulation techniques

The EER concept relies upon effective high-level modulation of the PA in order to restore the original signal envelope. So far, it has been assumed that the high-level modulation of the PA will occur by varying the main DC supply line voltage, however it is possible to achieve a similar effect in FET amplifiers by gate bias variation (although this will not maintain the saturated mode of operation). The two techniques are now considered in more detail

Gate modulation

With gate modulation, the gate bias is dynamically controlled by the envelope of the desired output signal. An increase in bias turns the device harder-on, and produces an increase in output level. Likewise, decreasing the input bias will reduce the output amplitude. If the amplifier operates in class-C, as is usually the case, then the gate bias will typically vary between a positive value below the threshold to a negative amount, with actual values being dependent upon the input drive level. Unfortunately, the gate/drain voltage transfer characteristic is very non-linear, and so the output signal is actually a substantially distorted version of the input signal. Although this amplitude distortion can, to a large part, be corrected by feedback, there is appreciable AM-PM distortion caused by the effective variation in input

signal amplitude, which results in the variation of amplifier duty cycle and saturation depth, for example. In particular, the effective device input capacitance is the source of most of the distortion; this capacitance is formed by the gate-source capacitance in parallel with the input capacitance due to the Miller effect acting upon the drain-gate capacitance. The latter capacitance is proportional to the amplifier voltage gain, and since this gain changes with input signal voltage so will the capacitance. This change will produce a phase shift at the input, since the input capacitance charges and discharges through a finite source resistance.

It is apparent, therefore, that straight-forward gate modulation does not perform well in the generation at high efficiency of a non-constant envelope signal.

However, as mentioned, the use of a simple feedback system can improve the amplitude characteristic substantially. This is done by comparing the envelope of the output signal with the required envelope, amplifying the difference and applying this to the gate [94]. By this process, the AM-AM distortion is, to a large degree, removed. It does nothing to reduce the magnitude of the AM-PM distortion, of course. It would be possible, though, to improve the phase distortion effects by means of a feedback loop controlling a phase-shifter, and this wouldn't add too much to the system complexity [95].

Since input bias is used to vary the output amplitude, the amplifier can not operate in saturation, hence overall efficiency is not as high as that achievable with drain modulation, below.

Drain modulation of a non-linear PA

When a PA is operating in a mode for which part of the cycle is spent in saturation, the output signal amplitude will be an almost linear function of the DC supply voltage [86, 88, 90, 100, 132]. Amplitude modulation can therefore be applied to a non-linearly amplified signal by varying the supply voltage as required. Unfortunately, the process does begin to deviate from the ideal when the voltage applied is either very low, and feedthrough effects become significant, or when it is very high,

and the operation becomes non-saturated.

The feedthrough effect results from the drain-gate capacitance of the FET. At low voltage supply levels an increasing part of output waveform is due to the gate-drain feedthrough by the capacitive coupling, and the voltage waveform due to this is almost in phase quadrature with the desired output, as the drain-gate capcitance's reactance is likely to be much greater than the amplifier load resistance. Rather than drop linearly towards zero, therefore, when the supply voltage does, the output voltage decreases non-linearly towards a small level, with the relative output phase changing until it reaches $\pi/2$ with zero volts applied. This non-linear effect produces spectral spreading at high modulation depths, and therefore operation in this region should ideally be avoided, if possible. $\pi/4$-DQPSK has a distinct advantage over standard QPSK in this respect, since zero crossings of the signal envelope do not occur, although consideration should be given to whether the filtering scheme used will create high envelope fluctuations, as this can occur with low roll-off factor (root) raised cosine filtering.

The non-linearity which results from the supply voltage going too high and taking the operation out of saturation can obviously cause distortion, and hence spectral spreading also. To prevent this from happening, sufficient drive should be applied to ensure that, for all voltage amplitude levels of interest, saturation does occur for a small part of each cycle.

Of course, in satisfying the above, the amplifier is likely to be well over-driven at low supply voltage levels, and this will tend to make the feedthrough effect more pronounced. Therefore, it can be expected that signals with high peak-mean envelope levels will experience more spectral spreading in such a system than signals with lower peak-mean levels.

The largest disadvantage with drain modulation in comparison to gate modulation is the fact that it is the supply to the drain (i.e. that part of the circuit through which the highest current flows) which is modulated, and hence the overall efficiency can be greatly reduced by any series losses/inefficiencies in the line, whereas varying

the gate DC bias is a much simpler task, since the gate is effectively a very high impedance at the low frequencies concerned, and so negligible current is involved. As the modulator and PA are in series for the drain current, the overall PA efficiency of the drain modulated PA is given by

$$\eta_t = \eta_{pa} \cdot \eta_{mod} \tag{6.8}$$

where η_t is the total efficiency, η_{pa} is the PA efficiency and η_{mod} that of modulator stage.

Although the modulator that has been used in this work has been an emitter-follower circuit, which is obviously relatively inefficient, it will be possible to use a much higher efficiency circuit. Usually, a switching-type circuit is used, with the switching frequency being five, or more, times the maximum frequency component of the amplitude signal. However, with the high data rate being used in this system, this would require switching in the region of 2 MHz, or higher. Problems with keeping RF generated in such a modulator out of the PA output would be found, and also it would be very difficult to maintain a good phase characteristic across the frequency range of the amplitude signal, which would result in a spectral spreading effect. One solution to this problem would be to use a 'split' modulator, comprising an emitter-follower to supply the (relatively small) high frequency components, and a switching amplifier to supply the low frequency (and much larger) current. Use of a feedback network would permit the follower to operate around the average current point, with the second amplifier providing that average current. It should be possible to achieve efficiencies in the region of 80-90 % with this type of scheme, and hence it is very attractive for high level modulation purposes.

As a result of the above, drain modulation is more suitable for this work, as the modulation process is much more linear, resulting in a substantially cleaner output spectrum, and the actual operating efficiency of the amplifier can be significantly higher, especially if a high efficiency modulator is used.

6.3 A practical 53 MHz PA

It was decided to use a power MOSFET device for the PA, as these can offer benefits over bipolar devices, such as:

- Higher thermal stability.
- Ease of biasing.
- Smaller parameter changes with operating conditions.

The Motorola MRF136 appeared suitable, offering 15W output power off a 28V supply and operation throughout the VHF range. As simulation of the circuits produced was required to confirm, and predict, results of practical operation, an accurate model of the device was obtained, for use with the *PSPICE* circuit simulation package. The model is capable of effectively simulating such effects as dynamic drain-gate and drain-source capacitances, which play a major role in the distortion mechanisms occurring in this application. Using PSPICE, it was then possible to simulate the AM-AM and AM-PM distortions resulting from changes in supply voltage, for any given circuit configuration. In addition, effects such as the variation of drive level and gate bias could readily be simulated, to build up a general picture of amplifier performance.

6.3.1 Up-converter and 53 MHz amplifiers

Before practical work could commence on the PA, it was necessary to be able to provide the required $\pi/4$-DQPSK drive signal at 53 MHz. The look-up and quadrature modulator implementation for generation of $\pi/4$-DQPSK at 10.7 MHz was described in chapter 5, and the hardware built to up-convert and amplify the RF signal is detailed now.

The 10.7 MHz IF signal is translated up to a frequency of 53 MHz by mixing with a second local oscillator signal, of frequency 63.7 MHz. A similar type of mixer

was used as in the quadrature modulator (chapter 5). The output of the mixer is then filtered by a 3-pole L/C circuit with a 3 dB bandwidth of 5 MHz, centred at 53 MHz. Rejection of the 74.4 MHz mixer output is greater than 40 dB. A buffer amplifier (50 Ω input/output impedance) completes the converter board.

The output of the converter then feeds a limiter/amplifier board. Here, the input signal is amplified with another broadband amplifier, the output of which drives a limiter circuit, using a Plessey SL532 low phase shift limiter chip, and also feeds an envelope detector circuit on the modulator board. The signal at this point is of around 3 volts peak-peak. This level is necessary in order to be able to perform efficient diode detection, but is too great for the limiter, and hence some resistive padding is used on the limiter input. The limiter is followed by an emitter-follower buffer, as it requires a high impedance load. This stage then drives a broadband amplifier, which produces approximately 7-8 volts of limited signal, for the main driver stage on the PA board.

6.3.2 PA circuitry

The PA circuit is shown in fig 6.4. The diagram shows the circuit configuration used in PSPICE simulation, the difference between that and the actual PA constructed being that in the simulation the driver stage is replaced be a low-impedance (RS) voltage source (vs). Gate bias is applied via R2. C1 and C2 are DC blocks, and R1, of low value, aids stability by damping the gate circuit. The RF choke in the drain circuit, RFC, provides the DC current to the amplifier, and its value is of sufficient value (1 μ H) to prevent RF currents from flowing, but to allow the low frequency fluctuations required for the amplitude modulation. L1, C3 and C4 form the resonant output network, with a Q of 4, transforming the load resistance of 50 Ω down to 28 Ω, which theoretically will give 10W output power off a 24 volt supply.

The circuit configuration was altered slightly from time to time, however the performance did not alter significantly from that found with this configuration.

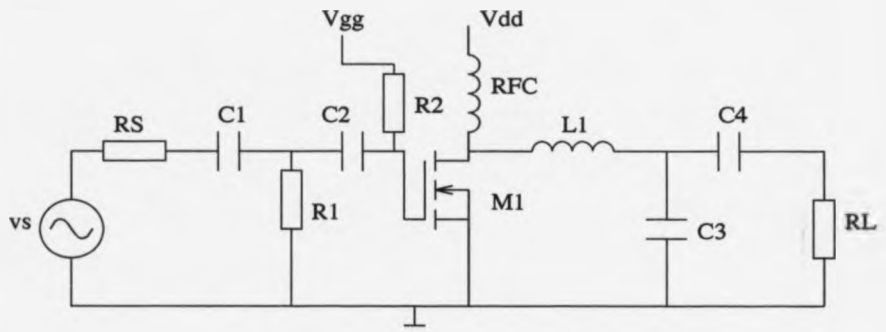


Figure 6.4: Power amplifier circuit

Practically, the driver stage was a high-frequency 1W bipolar transistor, operating in class C, although class B was later used to raise the drive level. An L/C network was used as an output/matching network.

6.3.3 Simulation results

The PA circuit described above was simulated using PSPICE. Some results found, which were typical of those generally obtained from the simulation work are shown in figs 6.6, 6.7 and 6.8. In this simulation, a sinusoidal driving source was used, with 20 volts peak-peak open circuit voltage and a source resistance of 20 Ω. Fig 6.6 shows the phase of the output signal (relative) for three different gate bias levels; it can be seen how the phase drops as the supply voltage is reduced, with the rate of change of phase increasing rapidly for low supply levels. Fig 6.7 shows the output signal amplitude for the same bias settings. The increase in output level with increasing gate bias is apparent, suggesting that saturation of the amplifier is not being achieved at the lower bias levels. It can also be seen that at high supply levels the gradients of the curves begin to fall, suggesting again that the amplifier is coming out of saturation. Fig 6.8 gives PA efficiencies under the same conditions. As would be expected from the previous two figures, efficiency is found to be higher

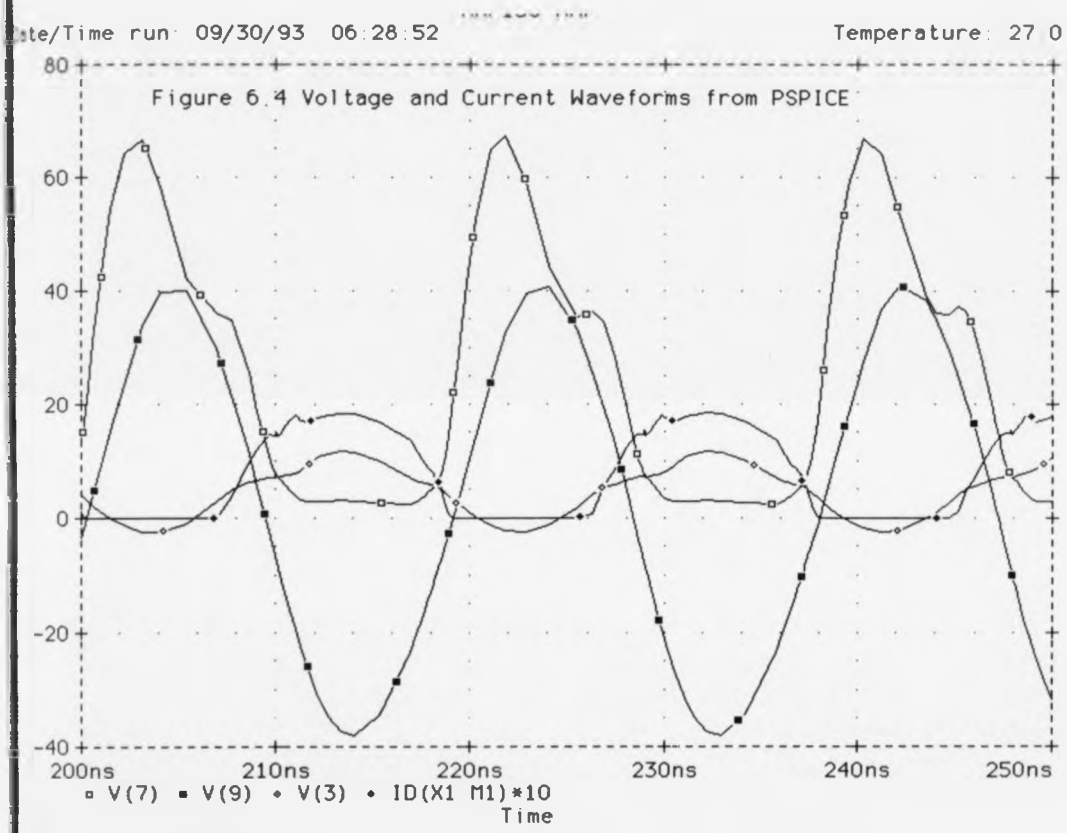
as the gate bias increases, since as the bias increases the amplifier is in more of a state of saturation, which is where the highest efficiencies will occur.

A plot of several of the waveforms occurring in the saturating PA is given - fig 6.5. The plot shows four waveforms: V(7) is the drain voltage; V(9) is the output voltage across the load; V(3) is the gate voltage, and ID(X1.M1) is the current flowing through the FET. It can be seen that the drain voltage is low and flat (i.e. the device is saturated) for a considerable period of time, and that this coincides with the current peak, hence efficiency should be good. The kinks in the gate waveform represent the times when the FET input capacitance is being charged or discharged. At theses times, the capacitance is at its largest and hence the slight flattening of the waveform. A slight glitch can also be seen on the drain current waveform, and this is due to the FET output capacitance discharging through the FET when the drain voltage drops towards zero.

Effect of load variation

In this work, the effect of variation of load impedance on PA operation is of major importance, as the PA used will be having to operate over a bandwidth large enough that the load impedance will change across the band. To observe the possible effects by simulation, the series inductance in the output network was varied either side of its calculated value. Characterization was done at two supply voltages - 3V and 24V, to demonstrate the phase shift across the range, and the effect on the output signal amplitude at 25V was also recorded. A gate bias of 5V was used, and a 20V peak-to-peak 20 Ω source. The results are shown in fig 6.9.

Practically, a series of measurements of phase shift over the supply range 5 - 25V were made, for different settings of output power resulting from output network tuning (that is, the output power produced with the supply voltage at 25V). For this practical case, the PA output network was as shown in fig 6.10. C2 it tuned for required loading, as it transforms the impedance of RL down to the required value, and the series network C1/L1 then resonates the output circuit. To produce



- V(7): Drain voltage
- V(9): Load voltage
- V(3): Gate voltage
- ID(X1.M1): FET drain current

Figure 6.5: Waveforms in saturating PA

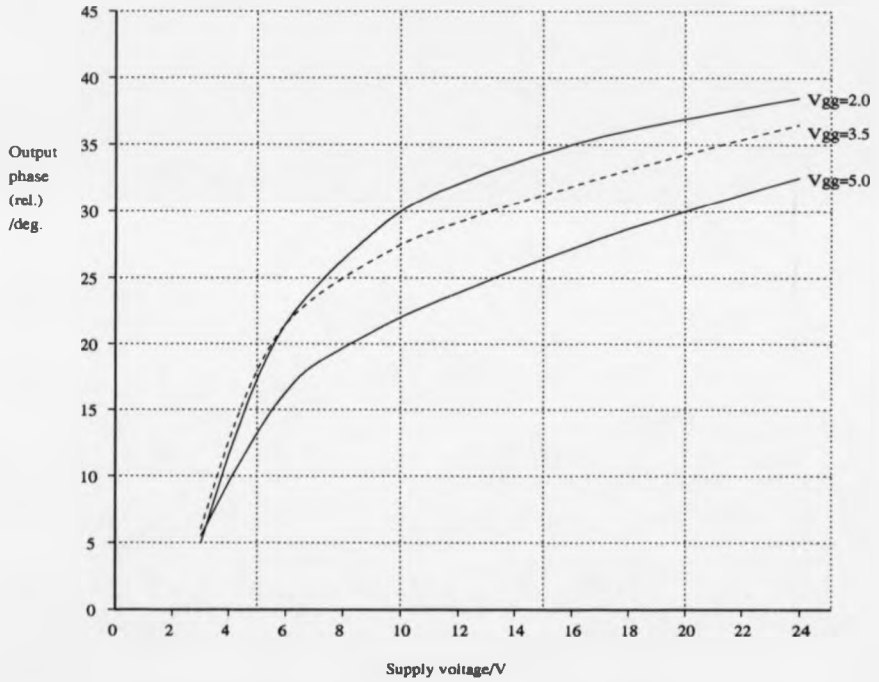


Figure 6.6: Phase vs. supply voltage for simulated circuit

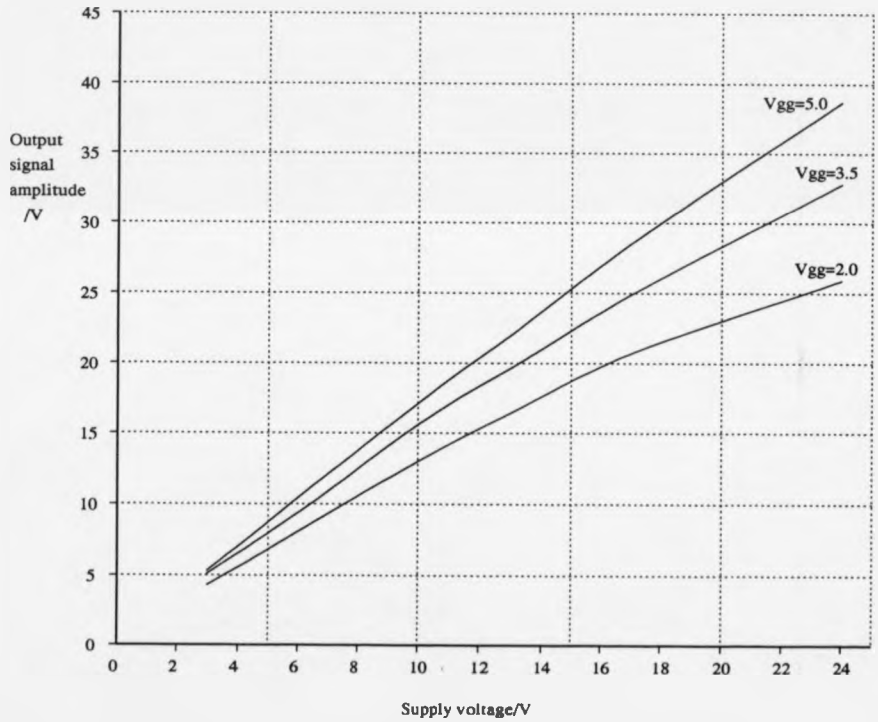


Figure 6.7: Output signal amplitude vs. supply voltage for simulated circuit

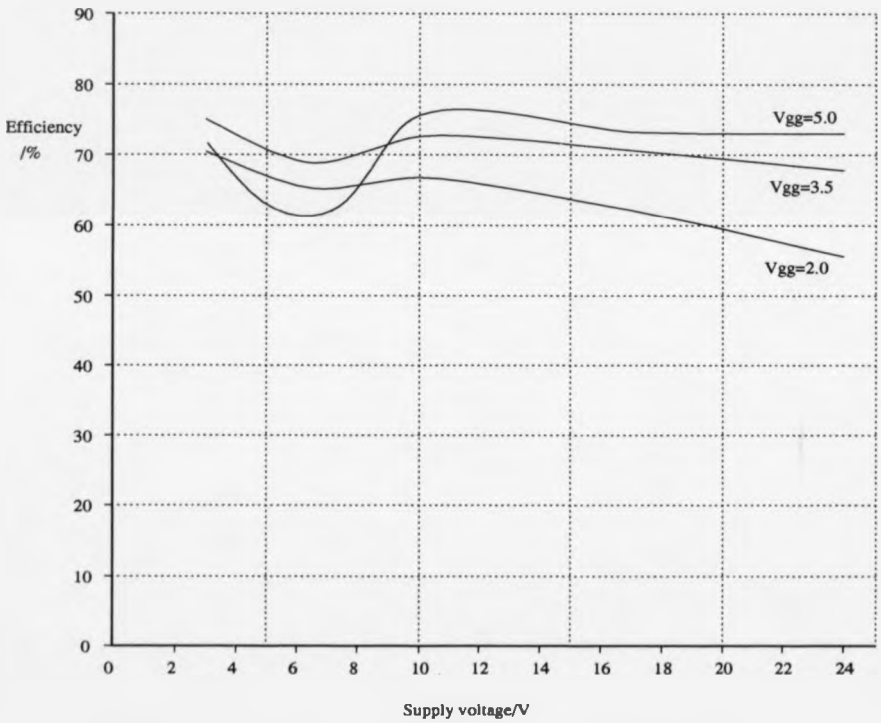


Figure 6.8: Efficiency vs. supply voltage for simulated circuit

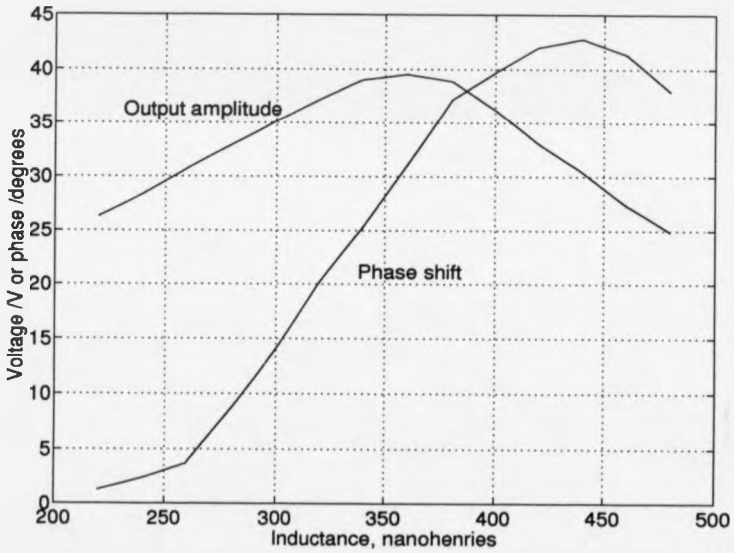


Figure 6.9: Amplifier performance vs. series inductance for simulated circuit

the output power variation for the measurements, C1 was varied about the position of resonance. The results are shown in fig 6.11 for values of power output of 1W on one side of resonance, through resonance, where the output is 7.5W, and to 1W on the other side of resonance. Hence, there are two readings for each power level, which are plotted on either side of the resonance point. It can be seen that the curve crosses the axis to one side of resonance; at this point there is zero phase shift as the PA voltage varies between a high and low level, and therefore zero AM-PM distortion.

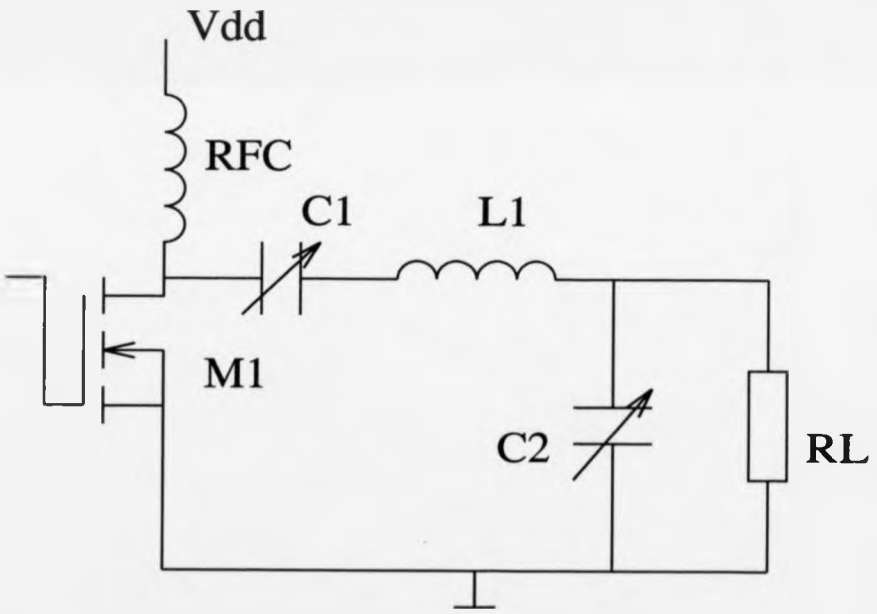


Figure 6.10: Practical PA output network

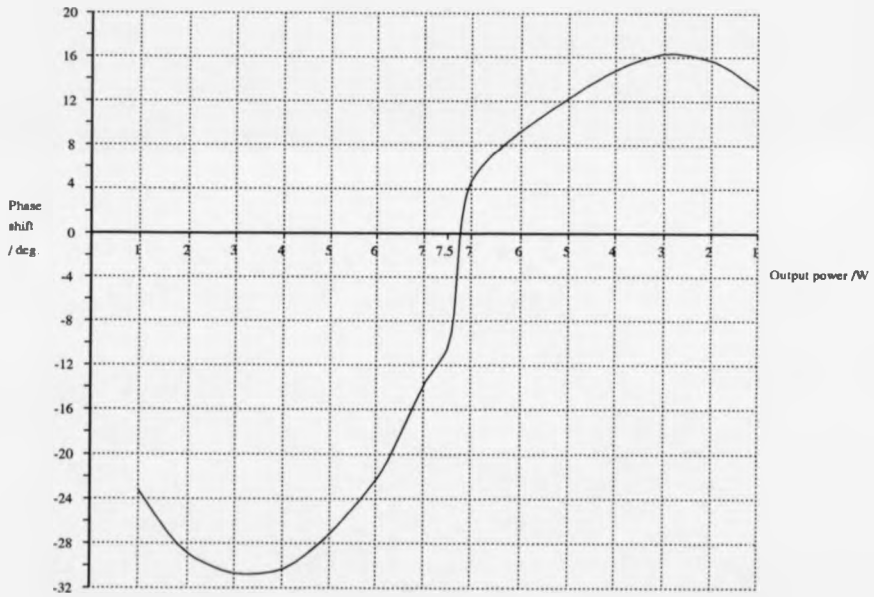


Figure 6.11: Phase shift vs. output power for practical circuit

6.3.4 Modulator circuitry

The modulator used initially is shown in fig 6.12. The input signal, containing amplitude fluctuation passes through C1 before being envelope detected in circuit D1/C2/R3. C2/R3 have a cut-off frequency of approximately 500kHz, allowing detection of the amplitude signal while blocking the 53 MHz carrier. The bias circuit R1/R2 and D2 provide D1 with a slight forward voltage, to allow detection of very low amplitude signals, such as those occurring when the system is in a ramping mode. The low frequency gain of the first op-amp stage is defined by R4/R3, and DC offset introduced by RV1 to set the required envelope depth on the output signal. The second op-amp stage provides further amplification of the rectified signal, with gain being adjusted by RV2. The output of this drives two emitter follower stages to give the required current gain for driving the PA stage.

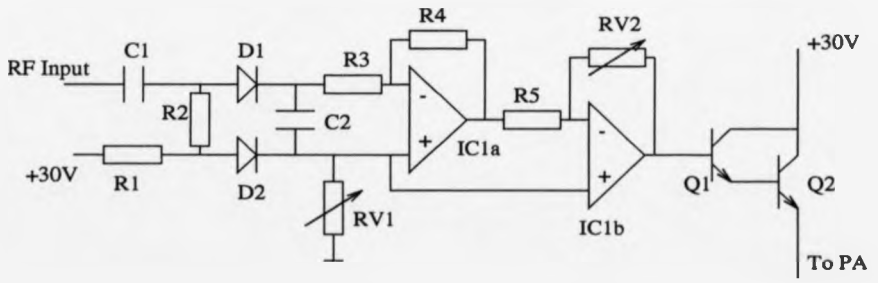


Figure 6.12: Envelope detector and high-level modulator circuit

6.3.5 Envelope feedback

As shown in the above results, the voltage supply-output amplitude response is not perfectly linear, and hence AM-AM distortion results. Use of a relatively simple feedback network, however, can reduce this effect to a very low level. The actual network used is shown in fig 6.13. Here, two identical envelope detector circuits feed

a differential amplifier. The high-level input, i.e. that from the PA output, is set to the required level by an external potential divider network, which behaves as a power control. The differential amplifier compares the envelope of the PA output with that of the low-level signal (taken from the output of the pre-limiter stage), and any difference is amplified and used to vary the PA supply to correct the error. Therefore, the circuit ensures that the PA output has an envelope identical to that of the low-level signal.

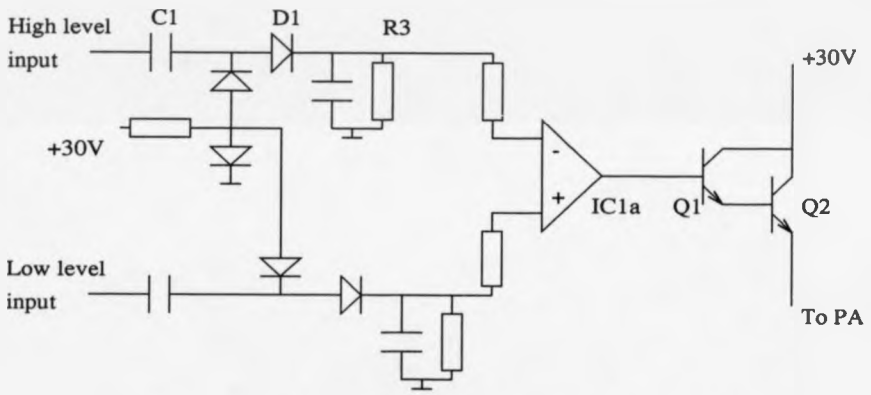


Figure 6.13: Envelope feedback and modulator circuit

6.3.6 Phase feedback

Linearization of the phase characteristic of the PA should be possible using a feedback network with a phase comparator and a simple phase shifter (e.g. varactor controlled tuned circuit). Work has begun on the implementation of this, and it is anticipated that it will be completed successfully in the very near future.

6.3.7 Practical spectra in the EER system

The hardware described in this chapter takes in the 10.7 MHz $\pi/4$-DQPSK signal, up-converts it to 53 MHz and then amplifies to around 10 W peak output power. Burst mode operation was successfully achieved, and it was found that spectral spreading did not occur (in ramp-up mode - ramp-down was not implemented fully).

Several spectra are shown. The first is that at the output of the up-converter (i.e. the low-level signal at 53MHz), fig 6.14. The signal at the same stage, only in ramp mode with the spectrum obtained by gating the burst edge, is shown in fig 6.15. It can be seen that since in ramp mode not all phase transitions are possible, the spectrum is asymmetric about the centre frequency, as would be expected. After being amplitude limited, the signal in fig 6.16 results, and the severe spectral spreading is clearly seen here. Two typical examples of PA outputs are given. Fig 6.17 is one of the better spectra obtained by careful tuning, and it can be seen that the resulting adjacent channel interference is a little below 40 dB. Fig 6.18 shows one of the slightly asymmetric spectra, the asymmetry being caused by phase shift in the modulator circuit. The spectrum is still respectable, however.

The above show that the EER technique can work in practice. Although there is obviously more work to be done on producing a cleaner spectrum, and on achieving a high modulator efficiency, the results are very encouraging, and certainly show that high-efficiency amplification of $\pi/4$-DQPSK is a practical proposition using relatively simple hardware.

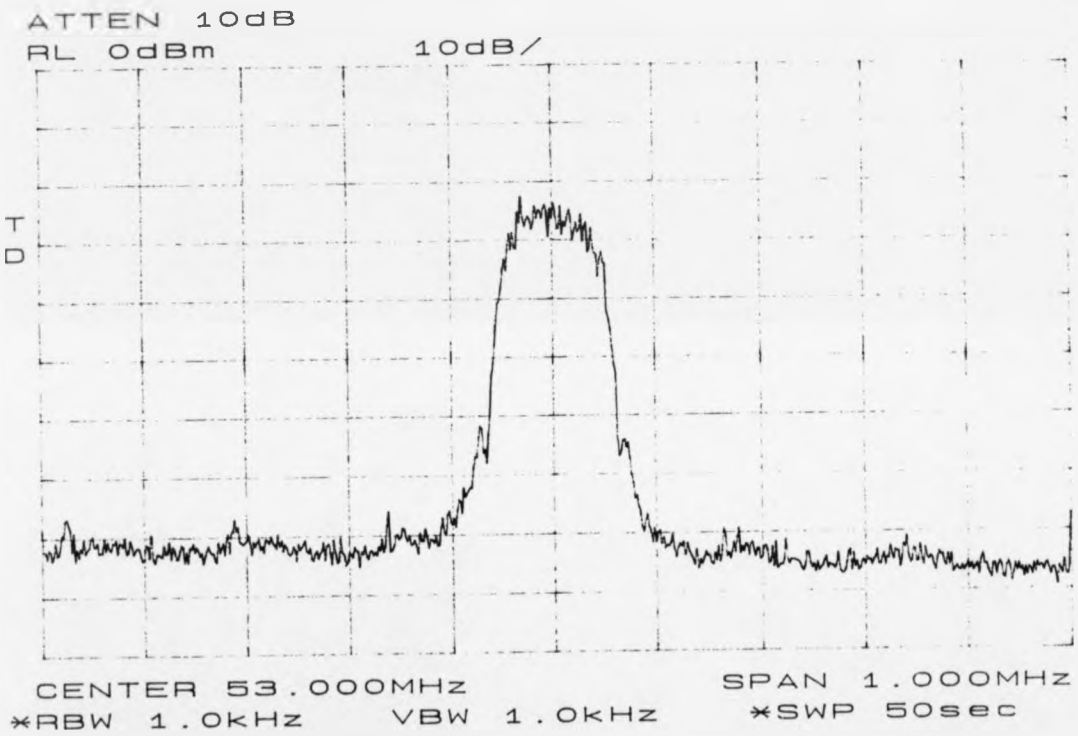


Figure 6.14: Spectrum of low-level 53 MHz $\pi/4$-DQPSK

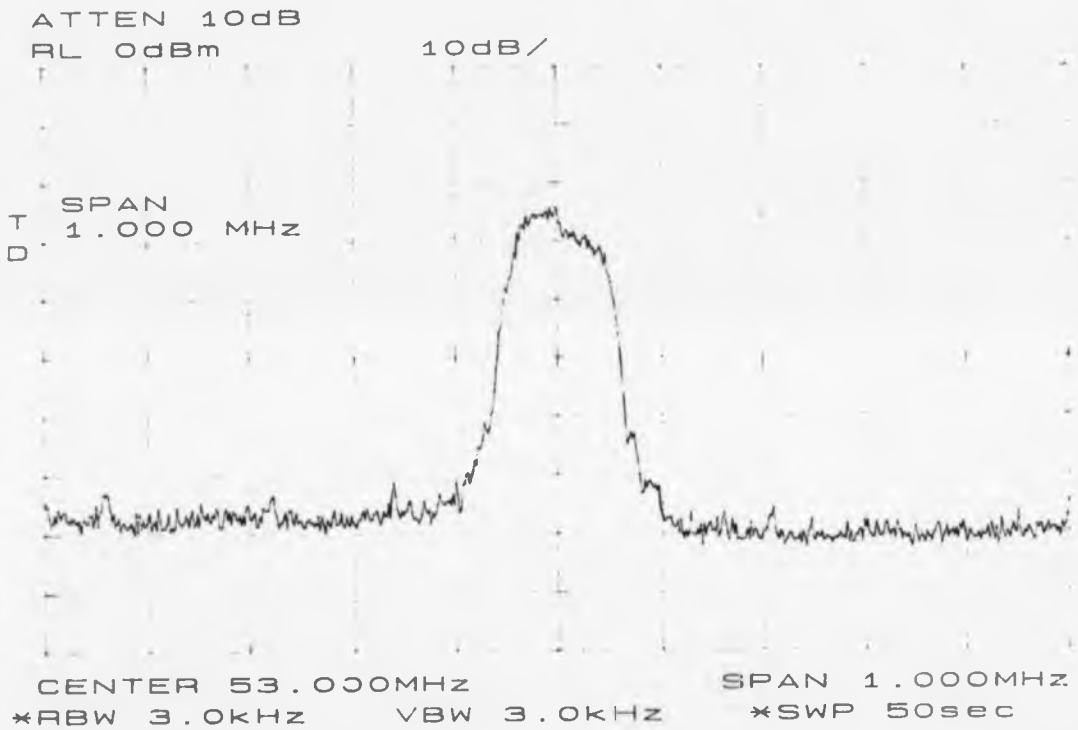


Figure 6.15: Spectrum of low-level 53 MHz burst data

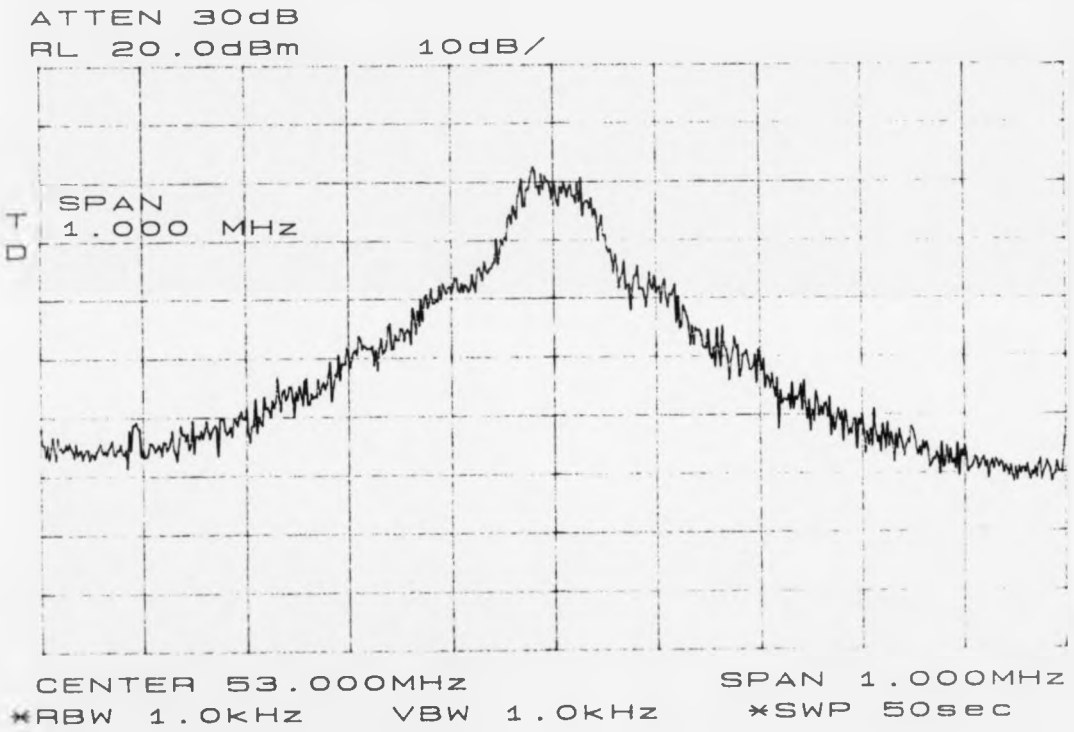


Figure 6.16: Spectrum of limited 53 MHz $\pi/4$-DQPSK

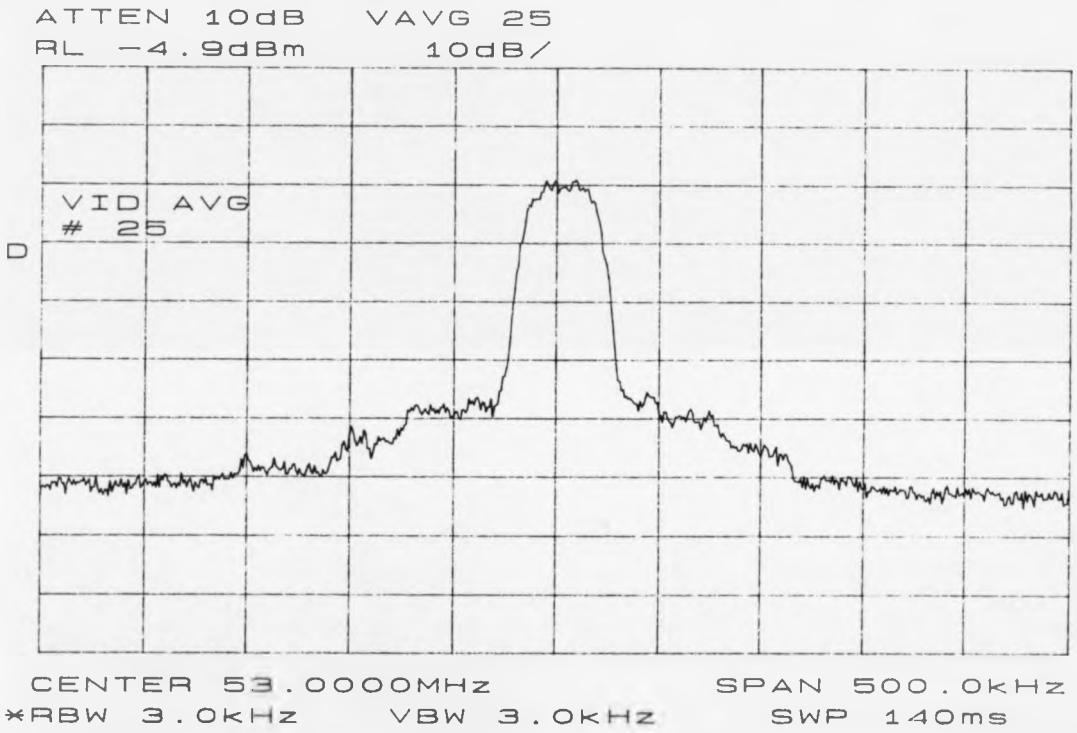


Figure 6.17: Spectrum of PA output (1) 53 MHz $\pi/4$-DQPSK

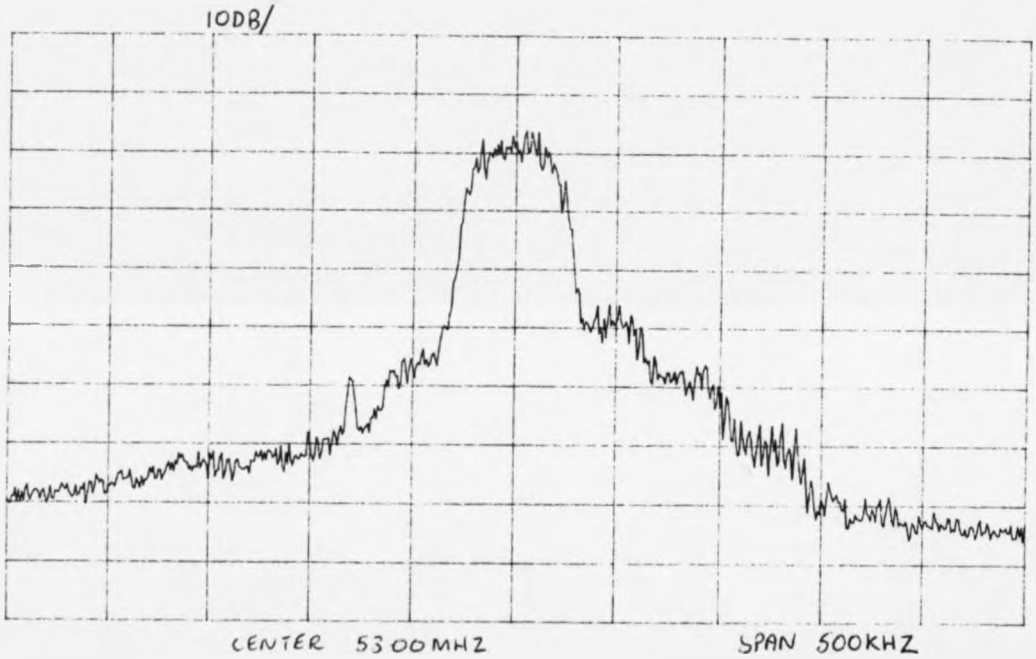


Figure 6.18: Spectrum of PA output (2) 53 MHz $\pi/4$-DQPSK

Chapter 7

Non-linearities in Envelope Elimination and Restoration Systems

7.1 Non-linearities in Envelope Elimination and Restoration Systems

This chapter is concerned with the imperfections of the Envelope Elimination and Restoration technique, particularly those associated with non-linearities in the PA stage due to high-level modulation. Of particular importance are AM-AM and AM-PM distortion due to the dynamically varying supply voltage to the PA. These effects are simulated and compared with distortions found in practice.

7.2 Modelling non-linearities

By measuring the phase and amplitude response of a power amplifier to a given input signal, it is possible to produce a model of the response which can then be used in simulating the amplifiers response to any arbitrary signal. By this process, spectral spreading effects due to the non-linearities can be predicted. If the amplifier can be described as frequency independent over the bandwidth of interest, the modelling can take the form of a relatively simple power series approximation. If, however, frequency independence cannot be assumed, then a more complex procedure must be employed which is usually based upon Volterra series techniques [126]. This work uses the power series approach since the amplifier can be assumed to be frequency independent over the relatively narrow bandwidths employed. A power series model for the appropriate non-linearity will be of the form

$$y(t) = a_0 + a_1 x(t) + a_2 x(t)^2 + a_3 x(t)^3 + a_4 x(t)^4 + \cdots + a_n x(t)^n \tag{7.1}$$

$$= \sum_{i=0}^{n} a_i x(t)^i \tag{7.2}$$

where y is the output, x the input, and the a terms the appropriate coefficients.

7.3 Amplitude modulation (AM) and AM-AM distortion

In an EER system, the envelope of the carrier is first removed, and then re-introduced, typically by means of high-level amplitude modulation. Non-linear power amplifiers, given the appropriate drive level, produce, ideally, an output voltage amplitude directly proportional to the power supply voltage. High level modulation can therefore be used to dynamically vary the output carrier amplitude as required. Unfortunately, practical amplifiers will not exhibit a perfectly linear AM-AM characteristic, however. This effect gives rise to intermodulation distortion, which can produce significant spectral spreading effects. Such spectral spreading implies a degree of adjacent channel interference in multi-channel radio systems, and is therefore of great concern in transmitter design.

Amplitude modulation can be described by the equation below

$$y(t) = A_c[1 + m(t)]\cos\omega_c t \tag{7.3}$$

where A_c is the unmodulated carrier output amplitude, $m(t)$ the modulating signal and $\omega_c t$ the carrier frequency (radian). In an EER system, the term $m(t)$ is introduced by high-level modulation and it is this term which is distorted by any amplitude non-linearity. If, in the simple case that $m(t) = m\cos\omega_m t$, distortion products are produced at multiples of ω_m, resulting from the mathematical expansion of $\cos^k\omega_m t$ terms, where k is an integer value. Practically, the modulating signal will tend to be far more complex than a simple cosinusoid, and the number of distortion terms produced far greater. It is, however, illustrative to consider the effect of the distortion on an input of the form $\cos\omega_{m1}t + \cos\omega_{m2}t$, which is the signal commonly used to measure the performance of amplifiers with regard to intermodulation distortion. It should be noted, however, that the nonlinearities being considered here are of a different form to those encountered in general amplifier

work, since it is the non-linearities produced by high-level modulation rather than those due to a varying input envelope which are of concern. Normally, therefore the input signal will be, for a two-tone test, $\cos(\omega_c t + \omega_{m1} t) + \cos(\omega_c t + \omega_{m2} t)$, as up-conversion of the tones will have already occurred. The following analysis therefore is in contrast to this situation, and it is not possible to ignore even-order effects which could otherwise be done.

First-order: In the simple first order (linear) case the resultant signal is given by:

$$y = A_c[1 + m(\cos\omega_{m1} t + \cos\omega_{m2} t)]\cos\omega_c t \tag{7.4}$$

which, upon expansion produces

$$y = A_c[\cos\omega_c t + \frac{m}{2}\cos(\omega_c \pm \omega_{m1})t + \frac{m}{2}\cos(\omega_c \pm \omega_{m2})t] \tag{7.5}$$

Second order: If $m(t)$ is instead given by $m(\cos\omega_{m1} t + \cos\omega_{m2} t)^2$, then, after expansion of $m(t)$,

$$y = A_c[1 + m(\frac{1 + \cos 2\omega_{m1} t}{2} + \cos(\omega_{m1} \pm \omega_{m2})t + \frac{1 + \cos 2\omega_{m2} t}{2})]\cos\omega_c t \tag{7.6}$$

which results in

$$y(t) = A_c[(1+m)\cos\omega_c t + \frac{m}{2}\cos(\omega_c \pm \omega_{m1} \pm \omega_{m2})t + \frac{m}{4}\cos(\omega_c \pm 2\omega_{m1})t + \frac{m}{4}\cos(\omega_c \pm 2\omega_{m2})t] \tag{7.7}$$

A third order non-linearity will produce sidebands at $\omega_c \pm \omega_{m1}$, ω_{m2}, $3\omega_{m1}$, $3\omega_{m2}$, $2\omega_{m1} \pm \omega_{m2}$, and $2\omega_{m2} \pm \omega_{m1}$.

7.4 Phase modulation (PM) and AM-PM distortion

It was mentioned in the previous section that high-level amplitude modulation of a non-linear PA will produce unwanted sidebands due to intermodulation distortion, if the process is not perfectly linear. It will also result in unwanted phase modulation (AM-PM distortion) which will likewise give rise to intermodulation distortion products. Phase modulation of a carrier may be described by

$$y(t) = A_c \cos(\omega_c t + \theta(t)) \tag{7.8}$$

where $\theta(t)$ is the instantaneous phase of the modulated carrier relative to the unmodulated carrier. $\theta(t)$ can be described, for PM, by

$$\theta(t) = k_p m(t) \tag{7.9}$$

where $m(t)$ is the modulating signal. If the maximum amplitude of m(t) is A_m, then the *phase modulation index*, β, is given by $\beta = k_p A_m$. For a modulating signal of $A_m \cos\omega_m$, the resulting modulated output is given by

$$y(t) = A_c \cos(\omega_c t + \beta \cos\omega_m t) \tag{7.10}$$

Evaluation of the sideband terms of a phase modulated signal generally involves the use of Bessel functions, however in the case where $\beta(t)$ is small ($\ll 1$), the *narrowband* situation, the mathematics is simplified considerably, and the following can be used.

$$\begin{aligned} y(t) &= Re[A_c \exp j(\omega_c t + \theta(t))] \\ &= Re[A_c \exp j\omega_c t \exp j\theta(t)] \end{aligned} \tag{7.11}$$

If β, and hence $\theta(t)$ are appropriately small, then $\exp j\theta(t)$ can be approximated by $1 + j\theta(t)$, which, when substituted into 7.11 gives

$$\begin{aligned} y(t) &= Re[A_c(1 + j\theta(t)) \exp j\omega_c t] \\ &= A_c(1 + j\theta(t)) \cos \omega_c t \\ &= A_c[1 + jk_p m(t)] \cos \omega_c t \end{aligned} \tag{7.12}$$

A comparison of 7.12 with 7.3 shows that the main difference between AM and narrowband PM is the fact that in the latter the sidebands are in phase quadrature with the carrier. Distortion analysis using a two-cosine input signal will produce sidebands at identical frequencies to that produced by AM distortion, however the phase of these sidebands will be different, as shown in the following section.

7.5 Combined AM and PM

If both AM and PM exist in a system as a result of a common modulating signal, say $\cos \omega_m t$ and if the PM is assumed to be narrowband, then analysis of the resulting signal can proceed as follows from 7.3 and 7.8

$$y(t) = A_c[1 + m(t)] \cos(\omega_c t + \theta(t))$$

$$= A_c[1 + m \cos \omega_m t] \cos(\omega_c t + \beta \cos \omega_m t) \tag{7.13}$$

but,

$$\begin{aligned} \cos(\omega_c t + \beta \cos \omega_m t) &= Re[\exp j(\omega_c t + \beta \cos \omega_m t)] \\ &= Re[\exp j\omega_c t \exp j\beta \cos \omega_m t] \\ &= \cos \omega_c t \cos(\beta \cos \omega_m t) - \sin \omega_c t \sin(\beta \cos \omega_m t) \end{aligned}$$

and, for $\beta \ll 1, \sin(\beta \cos \omega_m t) \approx \beta \cos \omega_m t$ and $\cos(\beta \cos \omega_m t) \approx 1$, so that

$$\begin{aligned} y(t) &= A_c[1 + m \cos \omega_m t](\cos \omega_c t - \beta \cos \omega_m t \sin \omega_c t) \\ &= A_c[\cos \omega_c t + \frac{m}{2} \cos(\omega_c \pm \omega_m)t - \frac{\beta}{2} \sin(\omega_c \pm \omega_m)t - \\ &\quad \frac{m\beta}{2} \sin \omega_c t - \frac{m\beta}{4} \sin(\omega_c \pm 2\omega_m)t] \end{aligned} \tag{7.14}$$

It is possible to represent the components of 7.14 in phasor diagram form, taking $\cos \omega_c t$ as the reference, and this, for the first three terms, is shown in 7.1. The sidebands resulting from the amplitude modulation are seen to be symmetrical around $\cos \omega_c t$, while those for phase modulation are symmetrical around $-\sin \omega_c t$. Vector addition of the upper sideband (USB) terms (i.e. those at relative angles of ω_m) and the lower sideband (LSB) terms (those at $-\omega_m$) will produce resultant sidebands of magnitude $\sqrt{(m/2)^2 + (\beta/2)^2}$. It is interesting to note that simultaneous AM and narrowband FM produce sidebands of unequal magnitude, since the sidebands produced by FM are in phase quadrature with those produced by PM [80], producing addition on one sideband but subtraction on the other.

7.5.1 Results of non-linearities in the modulation processes

The previous analysis for combined AM and PM modulation assumed a modulating signal of $\cos \omega_m t$, however this took no account of non-linearities in the modulator. A general form of an approximation to a non-linearity by means of a power series was given earlier in this chapter. If the modulating signal is now, in the general case, denoted as $e(t)$, then the expression for the combined AM and PM resulting is as below.

$$y(t) = A_c[\sum_{i=0}^{n} a_i e(t)^i] \cos(\omega_c t + \sum_{k=0}^{l} p_k e(t)^k) \tag{7.15}$$

which, for narrowband PM, results in

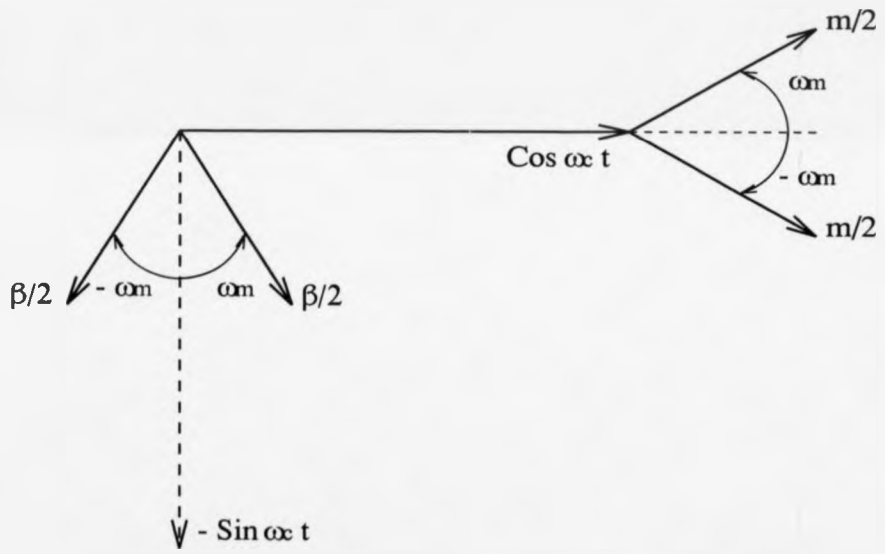


Figure 7.1: Phasor diagram showing AM and PM sidebands

$$y(t) = A_c[(\sum_{i=0}^{n} a_i e(t)^i)(1 + j\sum_{k=0}^{l} p_k e(t)^k)]\cos\omega_c t$$
$$= A_c[\sum_{i=0}^{n} a_i e(t)^i + j\sum_{i=0}^{n}\sum_{k=0}^{l} a_i b_k e(t)^{i+k}]\cos\omega_c t \quad (7.16)$$

For an arbitrary $e(t)$, therefore, it is possible to determine all intermodulation terms if accurate polynomials for the non-linearities are known.

7.6 Modelling the effects of AM-AM and AM-PM distortion on $\pi/4$-DQPSK

A $\pi/4$DQPSK signal can be represented in its baseband form by

$$s(t) = i(t) + jq(t) \quad (7.17)$$

It was shown in the previous chapter that, with the envelope elimination and restoration process, limiting of the above signal takes place (after up-conversion), and the envelope is then re-introduced by high-level modulation in the PA. AM-AM and AM-PM effects caused by PA characteristics then distort the signal. By characterization of the PAs phase and amplitude response, it is possible to produce polynomials to model these responses, and therefore determine the distortion effects. In EER, the signal driving the PA is limited in amplitude, which prevents distortion due to a varying input signal envelope, which is in contrast to the situation most often considered in the modelling of amplifier distortion. With EER it is the time-varying supply voltage that causes the non-linearities, and therefore the relevant amplifier characteristics may be measured by simply feeding an unmodulated carrier into the amplifier and measuring the relative phase and voltage of the output signal for a number of different settings of the DC supply. Application

of these results to practical signals depends upon the amplifier having frequency independent characteristics over the bandwidth of interest, however this can be assumed to be the case in narrowband systems using amplifier circuits with practical values of Q. Measurements of the characteristics of an amplifier were made, and additionally simulation work using PSPICE produced theoretical characteristics for the same amplifier implementation. Plots of the response of the simulated amplifier are shown in fig. 7.2.

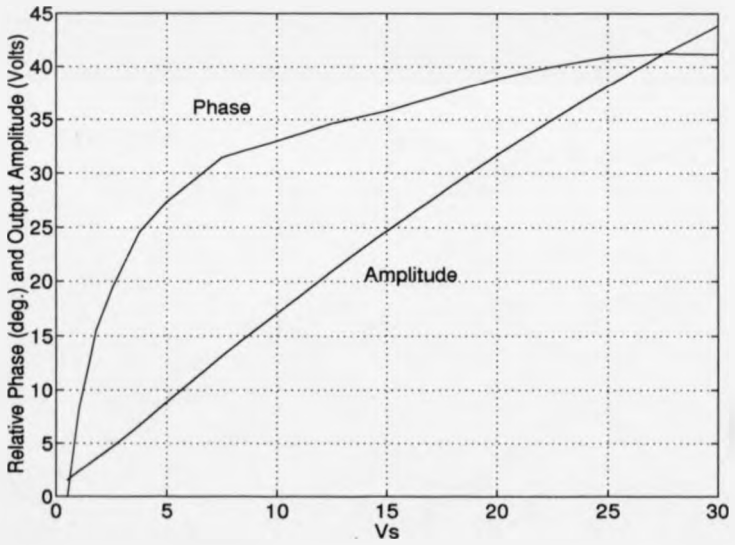


Figure 7.2: Plots of phase and amplitude response vs. Vs

In order to produce an accurate fit to the plots, polynomials of order seven were produced using the curve fitting feature of *Mathematica*. For the simulated amplifier, the following polynomials were produced:

$$p(x) = 0.806 + 0.325\ x - 0.0727x^2 + 0.00910x^3 - 6.51\ 10^{-}4\ x^4 + 2.64\ 10^{-}5\ x^5 - 5.60\ 10^{-}7\ x^6 + 4.87\ 10^{-}9\ x^7 \quad (7.18)$$

$$a(x) = 0.954 + 1.304\ x + 0.113\ x^2 - 0.0155\ x^3 + 0.00105\ x^4 - 3.99\ 10^5\ x^5 + 7.934\ 10^{-}7\ x^6 - 6.49\ 10^{-}9x^7 \quad (7.19)$$

This chapter has so far dealt with the distortion effects on a simple modulating signal, however the theory is equally well applied to more complex modulating signals. However, due to the amount of work involved in the analysis, simulation employing suitable software tools is appropriate for the modelling of the distortion effects on practical signals. The *MATLAB* package was used for this purpose, and suitable code written to simulate filtering, limiting, AM-AM distortion, AM-PM distortion etc. on $\pi/4$-DQPSK. This was described in chapter 4, but ignored effects due to non-linearities. For the simulation of non-linearities, it is necessary to begin with the appropriately filtered I and Q baseband streams. Ideal quadrature modulation and up-conversion are assumed, permitting the analysis to proceed at baseband. A block diagram of the simulation processes is shown in figure 7.3.

Here the filtered I and Q channels feed a limiter, and an envelope generator. The limiter outputs are summed to form the effective PA drive signal of constant envelope. The output of the envelope generator feeds the two non-linearities. The AM-AM nonlinearity generates the PA output amplitude from the envelope by use of the power-series approximation calculated. The resulting amplitude is multiplied by the summed limiter outputs to produce the PA output due to the amplitude non-linearity. The AM-PM non-linearity takes the ideal envelope as its input and generates the appropriate phase shift due to that envelope. This phase shift is

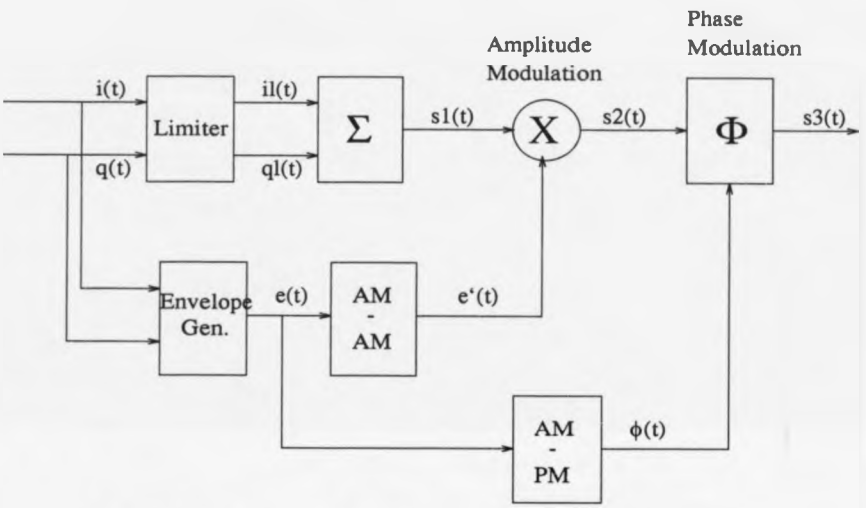


Figure 7.3: Diagram of non-linearity simulation procedure

then mapped onto the amplitude-modulated signal, and the output of this process forms the actual PA output due to high-level modulation and AM-AM and AM-PM conversion effects. The actual power spectral density of the output signal is then calculated by MATLABs Fourier transform algorithm, as can that of any of the signals occurring in the simulation. By this method it is possible to investigate the individual effects of the AM-AM and AM-PM distortion, and hence determine the worth of potential compensatory measures.

The effect of the above process is shown, at one particular point in time, $t1$, in fig 7.4. $s1(t1)$ corresponds to the limited instantaneous signal represented on a constellation diagram. AM on this signal produces $s2(t1)$, and a phase shift of θ then rotates $s2(t1)$ to $s3(t1)$. Of course, the order of the AM and PM can be changed, with the effect that the intermediate point, $s2(t1)$, will change, but the resultant signal $s3(t1)$ will remain the same.

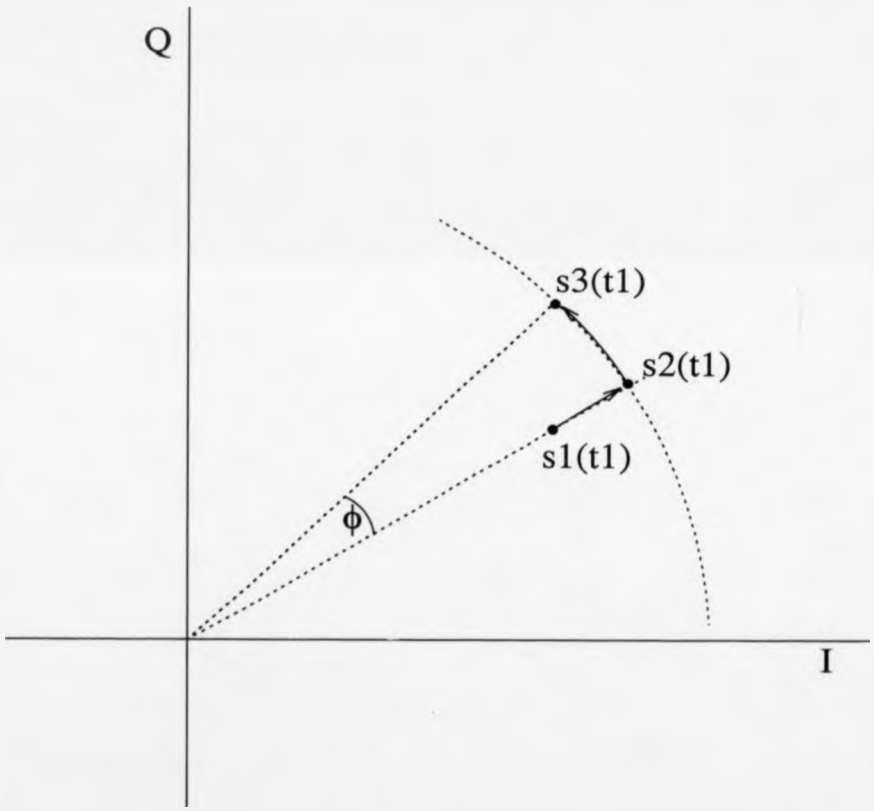


Figure 7.4: Diagram showing AM and PM effects on an instantaneous signal

7.6.1 Results

The plot for the filtered baseband spectrum is shown in fig. 7.5 together with that for the signal after a hard-limiting process, which gives rise to substantial spectral spreading, as can be seen. The AM and PM distortion in the PA are then considered separately, initially, to reveal the relative importance of each of the distortion mechanisms in the spectral spreading at the output of the PA. The resulting spectra are shown in fig. 7.6. Finally, the spectrum resulting from the combined distortion is shown in fig. 7.7. The actual RF spectra would simply be two-sided versions of each of the spectra shown.

These results do agree very well with many of the spectra that have been observed practically, and therefore appear to be very accurate in their prediction of distortion.

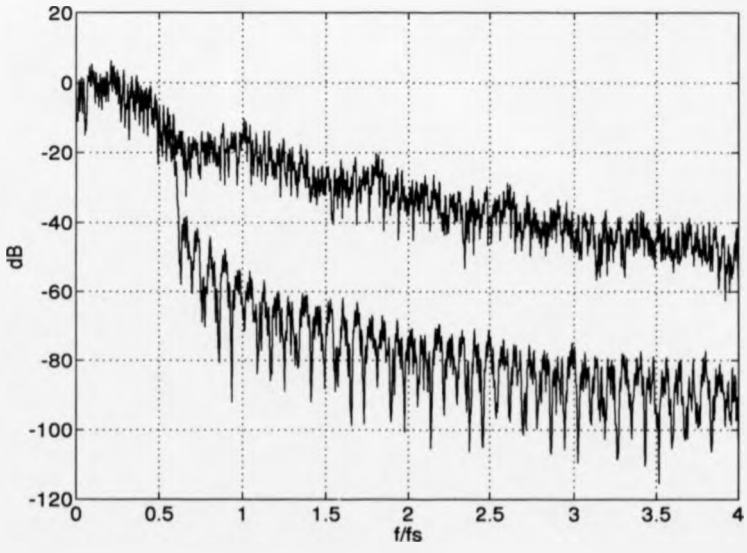


Figure 7.5: Spectra of filtered baseband (lower) and limited (upper) signals

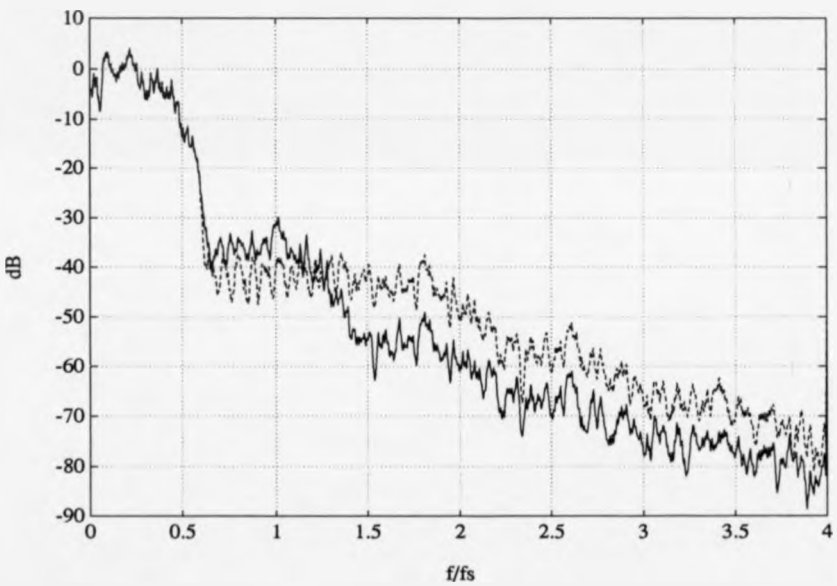


Figure 7.6: Spectra after phase distortion only (dashed) and after amplitude distortion only (solid)

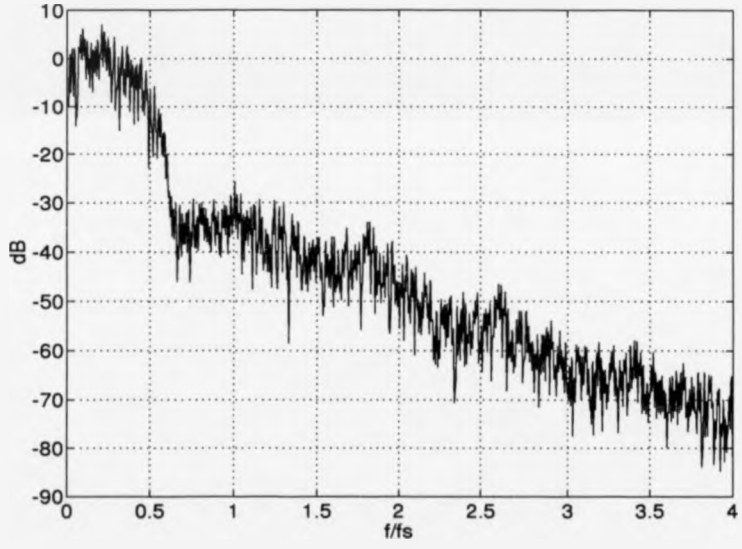


Figure 7.7: Spectrum after both phase and amplitude distortion in PA

Chapter 8

Conclusions

8.1 Discussion of propagation results

The results obtained in Sierra Leone suggested that multipath propagation would be unlikely to cause problems for the great majority of situations encountered there. The only occasions where the multipath appeared to be sufficiently strong to be of concern were those where the direct path was obstructed and there was a suitable 'reflector', such as was the case on the Aberdeen Pt. - Tower Hill path in Freetown. This path did show that delays of at least several microseconds may be found on such links, and remedial measures may be necessary to ensure reliable data transmission. Deep signal fading would be a likely cause of errors, and hence antennae location is important. The inherent route diversity in this system, however, is an effective countermeasure against poor radio channels, since better channels are likely to be selected during call set-up.

On the whole, the work in Sierra Leone was successful. It gave a good indication of signal strengths that could be expected over path the lengths anticipated in the rural radio system, that is, up to approximately 40km. In the Kabala region particularly, some of the transmission loss measurements agreed very well indeed with calculated values based upon knife-edge diffraction models, which did demonstrate their usefulness in predicting the quality of a radio link. Some results did not agree so well, though, and this can be attributed to prediction model inaccuracies, errors in assumed station locations, mapping errors, and to a lesser extent measurement errors.

It was unfortunate that the delay-spread measuring equipment had a number of short-comings, however some reasonable results were obtained, and the experience gained through encountering such problems was invaluable in the planning for the more recent propagation work which was undertaken in Tanzania, at 300 MHz.

Work on some longer paths would have been interesting, up to 100 km. Ideally, such paths would be studied over a long period of time to identify any potential problems due to anomalous propagation. Should ducting or sporadic-E be present

at any time, then it would be expected that interference to the system would result from distant stations. This could cause congestion and blocking.

8.2 Discussion of results of the modulation work

The modulation work has been extremely useful in demonstrating that it is quite feasible to produce transmitter hardware which can transmit spectrally efficient signals at good power efficiency, without the need for highly complex architectures.

The EPROM look-up system worked very well. Although the limited filter truncation length is a drawback to this implementation, the use of the Kingsbury filter was found to reduce the spectral spreading that would be found with raised cosine filters truncated to similar lengths. The method is flexible, since it is easy to change the coefficients held within the memory, and is appropriate for implementation on gate-array type structures, to reduce the overall size. Therefore, changes to filter types, and even modulation scheme is practical 'on site'.

Envelope elimination and restoration, as a means of linearization, appears to be generally overlooked somewhat, as it can clearly be capable of good results. This will be especially true if the work aimed at producing a simple phase correction network is successful, as it will be then possible to reduce by a very large extent the distortion present on the transmitted signal. Adjacent channel interference levels approaching 50 dB have been achieved, and it is felt that this can be increased. Also to be of great interest will be the development of the very efficient high-level modulator circuit, as overall amplifier efficiencies could then be very high indeed, of the order of 70 - 80 %. Such efficiencies would greatly exceed those attainable with schemes such as cartesian feedback and adaptive predistortion, since in EER the amplifier can operate in saturation for all time. In addition to this efficiency advantage, the complexity of the system will be far less than that in the other schemes, which is attractive for both size and cost considerations.

Bibliography

[1] STERN, M.: Cellular Technology - Revisited for Rural Service, *Communications International*, Oct. 1987.

[2] PERILLAN, L.B. and WU, W.W.: Digital Satellite Services and Networks Architecture: Intelsat Services and Their Evolution, *Proc. ICC'87*, 1987

[3] BUTLER, R.E.: World Communications Year (WCY), 1983: Development of Communications Infrastructure, *Proc. Pacific Telecom. Conf*, 1983.

[4] HUDSON, H.E.: Overcoming the Barriers of Distance: Telecommunications and Rural Development, *IEEE Technology and Society Mag.*, Dec.1989.

[5] BEAUPRE, D.M. and ZAVITZ, H.J.: TDMA Microwave Radio: The Technology and Its Applications, *Proc. WESCANEX 86*, IEEE 1986.

[6] LE-NGOC, T., MORRIS, M., ZAVITZ, J. and ROUX, M.: A New Point-to-Multipoint Data Polling Radio System, *Proc. ICC'87*, IEEE 1987.

[7] BONNEROT, G., FLOURY, G. and TANGUY, R.: IRT2000 A Versatile Multipurpose Radiosystem for Subscribers, *Proc. GLOBECOM'87*, IEEE 1987.

[8] LILLEY, J.: Transmission Techniques for Rural Networks, *Communications Engineering International*, Feb. 1987.

[9] Technology Forces the Missing Link, *Communications Systems Worldwide*, July/Aug. 1988.

[10] ESA, M. and RAHMAN, S.: Technological Development in Malaysian Rural Environment, *Proc. WESCANEX '91*, IEEE 1991.

[11] CHANDLER, S.A.G., BROWNE, S. and STEALEY, K.: A Distributed Digital Radio Network for Use in Rural Areas of Developing Countries, *Proc. 2nd Int. Conf. Rural Telecom.*, London, Oct. 1990. IEE.

[12] The Missing Link: Report of the Independent Commission for World Wide Telecommunications Development, ITU, 1984, Geneva.

[13] CHANDLER, S.A.G. and KARIMU, J.: A Village Radio Telecommunications System in Sierra Leone, *Proc. Int. Conf. Rural Telecom.*, London, May 1988, IEE.

[14] CHITAMU, P.J. and BRAITHWAITE, S.J.: Non-coherent Detection of Narrow-Band Modulation for Use in an Advanced Rural Radio Telephone Network, *Proc. IEE Colloquium on Radio Systems for Rural Communications*, London, Jan. 1993.

[15] BROWNE, S. and CHANDLER, S.A.G.: Modulation and Radio Transmission Aspects of the Distributed Rural Radio Telephone System, *Proc. IEE Colloquium on Radio Systems for Rural Communications*, London, Jan. 1993.

[16] STEALEY, K.B.: Protocols and Control Algorithms for Distributed Radio Telecommunications for Rural Areas, Thesis for MSc by research, University of Warwick, Department of Engineering, July 1992.

[17] CHANDLER, S.A.G.:Bush Telegraph- a pilot rural radio network in Sierra Leone", *IEE Review*, Vol. 36, no. 2, February 1990.

[18] DEVASIRVATHAM, D.M.J.: Multipath Time Delay Spread in the Digital Portable Radio Environment, *IEEE Comms. Mag.*, June 1987, Vol.25, No.6.

[19] COX, D.C.: Delay Doppler Characteristics of Multipath Propagation at 910 MHz in a Suburban Mobile Radio Environment, *IEEE Trans. Ant. Prop.*, Vol. AP-20, No. 5, Sept. 1972.

[20] SALEH, A.A.M., and VALENZUELA, R.A.: A Statistical Model for Indoor Multipath Propagation, *IEEE J. Sel. Areas Commun.*, Vol. SAC-5, No.2, Feb. 1987.

[21] CHUANG J.C-I.: The Effects of Time Delay Spread on Portable Radio Communications Channels with Digital Modulation, *IEEE J. Sel. Areas Commun.*, Vol. SAC-5, No. 5, June 1987.

[22] SCHMID, H.F.:A Prediction Model for Multipath Propagation of Pulse Signals at VHF and UHF Over Irregular Terrain, *IEEE Trans. Antennas and Propagation*, Vol. AP-18, No. 2, March 1970.

[23] ZOGG, A.: Multipath Delay Spread in a Hilly Region at 210 MHz, *IEEE Trans. Vehicular Technology*, Vol. VT-36, No. 4, Nov. 1987.

[24] DEVASIRVATHAM, D.M.J.: Time Delay Spread and Signal Level Measurements of 850 MHz Radio Waves in Building Environments, *IEEE Trans. Antennas and Propagation*, Vol. AP-34, No. 11, Nov. 1986.

[25] BULLINGTON, K.: Radio Propagation for Vehicular Communications, *IEEE Trans. Vehicular Technology*, Vol. VT-26, No. 4, Nov. 1977.

[26] SILLER, C.A.: Multipath Propagation, *IEEE Communications Mag.*, Vol. 22, No. 4, Feb. 1984.

[27] CHRISTMAS, W.J.: A Method for Measuring Echoes in the Multipath Propagation of VHF Signals, *BBC Rep.* BBC RD 1978/1.

[28] GUIDOTTI, G., LO MUZIO, P. and MONTANARI, M.:Effects of VHF Multipath Propagation on Data Transmission with Spread Spectrum Modulations, *Telettra Rev.*, Vol. 41, 1987.

[29] LO MUZIO, P., GUIDOTTI, G, BENVENUTO, N. and PUPOLIN S.: Experimental Characterization of VHF Propagation in a Mountain Environment, *Alta Frequenza*, Vol. 56, No. 6, Aug. 1987.

[30] BENVENUTO, N.: Distortion Analysis on Measuring the Impulse Response of a System Using a Crosscorrelation Method, *AT & T Bell Labs. Tech.*

Jrnl., Vol. 63, No. 10, Dec. 1984.

[31] TEWARI, K.K., SWARUP, S. and ROY, M.N.: Radio Wave Propagation Through Rain Forests of India, *IEEE Trans. Antennas and Propagation*, Vol. 38, No. 4, April 1990.

[32] HUGHES, K.A.: Propagation Studies for Africa, *News from Rohde & Schwartz*, 125, 1989/11.

[33] CHIA, S.T.S., STEELE, R., GREEN, E. and BARAN, A.: Propagation and Bit Error Measurements for a Microcellular System, *J. IERE*, Vol. 57, No. 6 (Supplement), Nov./Dec. 1987.

[34] NIX, A.R. and McGEEHAN, J.P.: Modelling and Simulation of Frequency Selective Fading Using Switched Antenna Diversity, *Electronics Letters*, 25th October 1990, Vol.26, No.22.

[35] SESAY, S.J.: Propagation Problems of a Tropical Environment in the Region 70-170 MHz, *Proc. EUROCON '77*, Venice, April 1977, Pt II.

[36] RAPPAPORT, T.S., SEIDEL, S.Y. and SINGH, R.: 900 MHz Multipath Propagation Measurements for U.S. Digital Cellular Radiotelephone, *IEEE Trans. Veh. Technol.*, Vol. VT-39, No.2, May 1990.

[37] PRASAD, M.V.S.N., SAIN, M. and REDDY, B.M.: Effect of Obstacles on VHF TV Signal Propagation, *IEEE Trans. Broadcasting*, Vol. 36, No. 3, Sept. 1990.

[38] BRAUN, W.R.: A Physical Mobile Radio Channel, *IEEE Trans. Veh. Technol.*, Vol. VT-40, No. 2, May 1991.

[39] RUMMLER, W.D.: A New Selective Fading Model: Application to Propagation Data, *Bell Syst. Tech. J.*, Vol.58, No.5, May-June 1979.

[40] GIGER, A.J. and BARNETT, W.T.: Effects of Multipath on Digital Radio, *IEEE Trans. Commun.*, Vol. COM-29, No.9, Sept. 1981.

[41] CHUANG, J.C-I., The Effects of Multipath Delay Spread on Timing Recovery, *IEEE Trans. Veh. Technol.*, Vol. VT-35, No. 3, Aug. 1987.

[42] YOSHIDA, S., ONOE, S. and IKEGAMI, F.: The Effect of Sample Timing on Bit Error Rate Performance in a Multipath Fading Channel, *IEEE Trans. Veh. Technol.*, Vol. VT-35, No.4, Nov. 1986.

[43] DICKSON, F.H., EGLI, J.J., HERBSTREIT, J.W. and WICKIZER, G.S.: Large Reductions of VHF Transmission Loss and Fading by the Presence of a Mountain Obstacle in Beyond-Line-of-Sight Paths, *Proc I.R.E.*, Aug. 1953.

[44] GROSSKOPF, R.: Comparison of Different Methods for the Prediction of the Field Strength in the VHF Range, *IEEE Trans. Antennas Propagat.*, Vol. AP-35, No. 7, July 1987.

[45] MEEKS, M.L.: VHF Propagation over Hilly, Forested Terrain, *IEEE Trans. Antennas Propagat.*, Vol. AP-31, No.3, May 1983.

[46] BULLINGTON, K.: Radio Propagation Fundamentals, *Bell Syst. Tech. J.*, Vol. 36, No.3, May 1957.

[47] BULLINGTON, K.: Reflection Coefficients of Irregular Terrain, *Proc. I.R.E.*, Vol. 42, Aug. 1954.

[48] Effects of Tropospheric Refraction on Radiowave Propagation: Recommendations and Reports of the CCIR, *Propagation in Non-Ionized Media*, V, XVI Plenary Assembly, 1986.

[49] PARSONS, D.: Characterisation of Fading Mobile Radio Channels, Chapter 3, Personal and Mobile Radio Systems, Macario, R.C.V. (ed.), Peregrinus,

1991.

[50] DEYGOUT, J.: Multiple Knife-Edge Diffraction of Microwaves, IEEE Trans. Antennas Propagat., Vol. AP-14, No. 4, July 1966.

[51] EPSTEIN, J. and PETERSON, D.W.: An Experimental Study of Wave Propagation at 850 MC, /em Proc. I.R.E., Vol. 41, May 1953.

[52] Report 715-2 - Propagation by Diffraction: Recommendations and Reports of the CCIR, *Propagation in Non-Ionized Media*, V, XVI Plenary Assembly, 1986.

[53] TOWNSEND, A.A.R.: Digital Line-of-Sight Radio Links: a Handbook, Prentice Hall, 1988.

[54] Report 1007: Recommendations and Reports of the CCIR, *Propagation in Non-Ionized Media*, V, XVI Plenary Assembly, 1986.

[55] HOROWITZ, P. and HILL, W.: The Art of Electronics, Cambridge University Press, 2nd ed., 1989, pp 655-660.

[56] MUROTA,K. and HIRADE, K.: GMSK for Digital Radio Telephony, *IEEE Trans. Commun.*, Vol. COM-29, No. 7, July 1981.

[57] CHUNG, K-S.: Generalized Tamed Frequency Modulation and Its Application for Mobile Radio Communications, *IEEE Trans. Veh. Technol.*, Vol. VT-33, No. 3, Aug. 1984.

[58] KINOSHITA, K., HATA, M. and NAGABUCHI, H.: Evaluation of 16 kbit/s Digital Voice Transmission for Mobile Radio, *IEEE Trans. Veh. Technol.*, Vol. VT-33, No. 4, Nov. 1984.

[59] DE JAGER, F. and DEKKER, C.B.: Tamed Frequency Modulation, A Novel Method to Achieve Spectrum Economy in Digital Transmissiom, *IEEE Trans. Commun.*, Vol. COM-26, No. 5, May 1978.

[60] MUROTA, K.: Spectrum Efficiency of GMSK Land Mobile Radio, *IEEE Trans. Veh. Technol.*, Vol. VT-34, No. 2, May 1985.

[61] AULIN, T. and SUNDBERG, C.W.: Continuous Phase Modulation - Part I: Full Response Signaling, *IEEE Trans. Commun.*, Vol. COM-29, No. 3, March 1981

[62] AULIN, T., RYDBECK, N. and SUNDBERG, C.W.: Continuous Phase Modulation - Part II: Partial Response Signaling, *IEEE Trans. Commun.*, Vol. COM-29, No. 3, March 1981

[63] HEATH, M.R. and LOPES, L.B.: Variable Envelope Modulation for Personal Communications, *Proc. IEE Telecomm. Conf. 1989*, pp249-254.

[64] AKAIWA,Y. and NAGATA, Y.: A Linear Modulation Scheme for Digital Radio Communications, *Proc. GLOBECOM '85.*, IEEE 1985.

[65] GRONEMEYER, S.A. and McBRIDE, A.L.: MSK and Offset QPSK Modulation, *IEEE Trans. Commun.*, Vol. COM-24, No. 8, Aug.1976.

[66] MORAIS, D.H. and FEHER, K.: Bandwidth Efficiency and Probability of Error Performance of MSK and Offset QPSK Systems, *IEEE Trans. Commun.*, Vol. COM-27, No. 12, Dec.1979.

[67] AKAIWA, Y. and NAGATA, Y.: Highly Efficient Digital Mobile Communications with a Linear Modulation Method, *IEEE J. Sel. Areas Commun.*, Vol. SAC-5, No. 5, June 1987.

[68] ONO, S., KONDOH, N. and SHIMAKAZI, Y.: Digital Cellular System with Linear Modulation, *Proc. 36th IEEE Conf. Veh. Tech.*, IEEE 1989.

[69] NAGATA, Y.: Linear Amplification Technique for Digital Mobile Communications, *Proc. 36th IEEE Conf. Veh. Tech.*, IEEE 1989.

[70] GUO, Y. and FEHER, K.: Performance Evaluation of Differential $\pi/4$-QPSK Systems in a Rayleigh Fading/Delay Spread Environment, *Proc. 40th IEEE Conf. Veh. Tech.*, IEEE 1990.

[71] JENKS, F.G., MORGAN, P.D. and WARREN, C.S.: Use of Four-Level Phase Modulation for Digital Mobile Radio, *IEEE Trans. Electromagnetic Compatibility*, Vol. EMC-14, No. 4, Nov. 1972

[72] GREENSTEIN, L.J. and FITZGERALD, P.J.: Envelope Fluctuation Statistics of Filtered PSK and Other Digital Modulations, *IEEE Trans. Commun.*, Vol. COM-27. No.4, April 1979.

[73] SUNDBERG, C.E.: On Continuous Phase Modulation in Cellular Digital Mobile Radio Systems, *Bell Syst. Tech. J*, Vol. 62, No. 7, Sept. 1983.

[74] QI, Y., WU, G., YANG, J. and FEHER, K.: Spectral Pollution in TDMA: Digital FM and $\pi/4$-QPSK Experimental Study and CAD Improvement, *Proc. 41st Conf. Veh. Technol*, IEEE 1991.

[75] BUSTILLO, J., RODRIGUEZ-PALANCA, M. and PEREZ, J.: 4-PAM/FM Modulator with DSP: A Solution for ERMES, *Proc. 41st Conf. Veh. Technol.*, IEEE 1991.

[76] FAGUE, D.F. and FEHER, K.: Experimentally Optimized Narrowband FM System for TDMA Operated Mobile Radio Environment, *Proc. 24th Asilomar Conf. on Signals, Systems and Computers*, Nov. 1990.

[77] RAPPAPORT, T.S. and FUNG, V.: Simulation of Bit Error Performance of FSK, BPSK, and $\pi/4$ DQPSK in Flat Fading Indoor Radio Channels Using a Measurement-Based Channel Model, *IEEE Trans. Veh. Tech.*, Vol. VT-40, No. 4, Nov. 1991.

[78] DIVSALAR, D. and SIMON, M.K.: The Power Spectral Density of Digital Modulations Transmitted Over Nonlinear Channels, *IEEE Trans. Commun.*, Vol. COM-30, No. 1, Jan. 1982.

[79] MACARIO, R.C.V.: Modulation Techniques: analog and digital: Chapter 1, Personal Radio and Mobile Communications Systems, Peregrinus, 1991.

[80] STREMLER, F.G.: Introduction to Communications Systems, 2nd ed., Addison-Wesley, 1982.

[81] SCHWARTZ, M.: Information, Transmission, Modulation and Noise, 4th ed., McGraw-Hill, 1990.

[82] GIBSON, J.D.: Principles of Digital and Analog Communications, Macmillan, 1989.

[83] PROAKIS, J.G.: Digital Communications, 2nd ed., McGraw-Hill, 1989.

[84] LIU, C-L. and FEHER, K.: Noncoherent Detection of $\pi/4$-QPSK Systems in a CCI-AWGN Combined Interference Environment, *Proc. 39th Conf. Veh. Technol*, IEEE 1989.

[85] GOODE, S.H., KAZECKI, H.L. and DENNIS, W.: A Comparison of Limiter-Discriminator, Delay and Cohrent Detection for $\pi/4$-QPSK, *Proc. 40th Conf. Veh. Technol*, IEEE 1990.

[86] SENAK,P.: Amplitude Modulation of the Switched-Mode Tuned Power Amplifier, *Proc IEEE*, Oct. 1965.

[87] SOKAL, N.O. and SOKAL, A.D.: Class E - A New Class of High-Efficiency Tuned Single-Ended Switching Power Amplifiers, *IEEE J. Solid-State Circuits*, Vol. SC-10, No. 3, June 1975.

[88] RAAB, F.H.: Effects of Circuit Variations on the Class E Tuned Power Amplifier, *IEEE J. Solid-State Circuits*, Vol. SC-13, No. 2, April 1978.

[89] KAZIMIERCZUK, M.: Collector Amplitude Modulation of the Class E Tuned Power Amplifier, *IEEE Trans. Circuits Syst.*, Vol. CAS-11, No. 6, June 1984.

[90] KRAUSS, H.L. and SANFORD, J.F.: Collector Modulation of Transistor Amplifiers, *IEEE Trans. Circuit Theory*, Vol. CT-12, Sept. 1965.

[91] ROSE, B.E.: Notes on Class-D Transistor Amplifiers, *IEEE J. Solid-State Circuits*, June 1969.

[92] BAILEY, R.L.: Large-Signal Nonlinear Analysis of a High-Power High-Frequency Junction Transistor, *IEEE Trans. Electron Devices*, Vol. ED-17, No. 2, Feb. 1970.

[93] KAZIMIERCZUK, M.K. and TABISZ, W.A.: Class C-E High-Efficiency Tuned Power Amplifier, *IEEE Trans. Circuit Syst.*, Vol. CAS-36, No. 3, March 1979.

[94] PETROVIC, V. and GOSLING, W.: VHF/AM transmitter using VMOS technology, *Electronic Engineering*, June 1978.

[95] PETROVIC, V. and GOSLING, W.: Polar-Loop Transmitter, *Electronics Letters*, Vol 15, No. 10, 1979.

[96] KAZIMIERCZUK, M.K.: Class E Tuned Power Amplifier with Nonsinusoidal Output Voltage, *IEEE J. Solid-State Circuits*, Vol. SC-21, No. 4, Aug. 1986.

[97] RAAB, F.H.: Idealized Operation of the Class E Tuned Power Amplifier, *IEEE Trans. Circuits Syst.*, Vol. CAS-24, No. 12, Dec. 1977.

[98] KAZIMIERCZUK, M.K. and PUCZKO, K.: Exact Analysis of Class E Tuned Power Amplifier at any Q and Switch Duty Cycle, *IEEE Trans. Circuits Syst.*, Vol. CAS-34, No.2, Feb. 1987.

[99] RAAB, F.H.: Effects of Circuit Variations on the Class E Tuned Power Amplifier, IEEE J. Solid-State Circuits, Vol. SC-13, No. 2, April 1978.

[100] KOCH, M.J. and FISHER, R.E.: A High Efficiency 835 MHz Linear Power Amplifier for Digital Cellular Telephony, *Proc. 39th Conf. Veh. Technol.*, IEEE 1989

[101] CASADEVALL, F. and OLMOS, J.J.: On the Behaviour of the LINC Transmitter, *Proc. 41st Conf. Veh. Technol.*, IEEE 1990.

[102] McCUNE, E.W.: Control of Spurious Signals in Direct Digital Synthesis, *Proc. RF Expo 88.*

[103] CAVERS, J.K.: A Linearizing Pre-distorter with Fast Adaption, *Proc. 40th Conf. Veh. Technol.*, IEEE 1990.

[104] DAVARIAN, F. and SUMIDA, J.T.: A Multipurpose Digital Modulator, *IEEE Commun. Mag.*, Feb. 1989.

[105] FEHER, K.: MODEMS for Emerging Digital Cellular-Mobile Radio System, /em IEEE Trans. Veh. Tech., Vol. VT-40, No. 2, May 1991.

[106] SUZUKI, H. and YAMAO, Y.: Design of Quadrature Modulator for Digital FM Signaling with Digital Signal Processing, *Trans. IECE*, Sept. 1982.

[107] KINGSBURY, N.G.: Transmit and Receive Filters for QPSK Signals to Optimise the Performance on Linear and Hard-Limited Channels, *Proc. IEE*, Vol. 133, Pt. I, No. 4, July 1986.

[108] RAAB, F.H.: Envelope-Elimination-and-Restoration System Requirements, *Proc. RF Expo 88.*

[109] FAULKNER, M., MATTSSON, T. and YATES, W.: Adaptive Linearisation Using Pre-distortion, *Proc. 40th IEEE Veh. Tech. Conf*, 1990.

[110] TOMISATO, S., CHIBA, K. and MUROTA, K.: Phase Error Free LINC Modulator, *Electronics Letters*, 27th April 1989, Vol. 25, No. 9.

[111] CHIBA, K., NOJIMA, T. and TOMISATO, S.: Linearized Saturation Amplifier with Bidirectional Control (LSA-BC) for Digital Mobile Radio, *Proc. GLOBECOM 90*, IEEE 1990.

[112] COX, D.C.: Linear Amplification with Nonlinear Components, *IEEE Trans. Commun.*, Vol. COM-22, Dec. 1974.

[113] ARIYAVISITAKUL, S. and LIU, T-P.: Characterizing the Effects of Nonlinear Amplifiers on Linear Modulation for Digital Portable Radio Communications, *IEEE Trans. Veh. Tech.*, Vol. VT-39, No. 4, Nov. 1990.

[114] TAKAMI, T., SAITO, S., TOMISATO, S. and YAMAO, Y.: Nyquist QPSK Transmission using Rational-Function Filter for Mobile Radio, *Proc. 41st IEEE Veh. Tech. Conf.*, 1991.

[115] JOHANSSON, M. and MATTSSON, T.: Transmitter Linearization using Cartesian Feedback for Linear TDMA Modulation, *Proc. 41st IEEE Veh. Tech. Conf*, 1991.

[116] KENINGTON, P.B., WILKINSON, R.J. and MARVILL, J.D.: Broadband Linear Amplifier Design for a PCN Base-Station, *Proc. 41st IEEE Veh. Tech. Conf.*, 1991.

[117] HETZEL, S.A., BATEMAN, A. and McGEEHAN, J.P.: A LINC Transmitter, *Proc. 41st IEEE Veh. Tech. Conf.*, 1991.

[118] KAHN, L.R.: Single-Sideband Transmission by Envelope Elimination and Restoration, *Proc. IRE*, July 1952.

[119] BERMAN, A.L. and MAHLE, E.: Nonlinear Phase Shift in Travelling-Wave Tubes as Applied to Multiple Access Communications Satellites, *IEEE*

Trans. Commun., Vol. COM-18, No. 1, Feb. 1970.

[120] YAMAO, Y., SAITO, S, SUZUKI, H. and NOJIMA, T.: Performance of $\pi/4$-QPSK Transmission for Digital Mobile Radio Applications, *Proc. 39th Conf. Veh. Tech*,IEEE 1989.

[121] PETROVIC, V. and BROWN, A.N.: Application of Cartesian Feedback to HF SSB Transmitters, *IEE Conf. HF Commun. Syst. and Techniques*, IEE Feb.1985.

[122] OHNO, K., IKURA, M., TAKAMI, T. and TANAKA, T.: Signal Transmission Technologies for Digital Cellular Systems, *NTT Rev.*, Vol. 4, No. 1, Jan. 1992.

[123] LUNDQUIST, L., LOPRIORE, M. and GARDNER, F.M.: Transmission of 4 ϕ-Phase-Shift-Keyed Time-Division Multiple Access over Satellite Channels, *IEEE Trans. Commun.*, Vol. COM-22, No. 9, Sept. 1974.

[124] MORAIS, D.H. and FEHER, K.: the Effects of Filtering and Limiting on the Performance of QPSK, Offset QPSK, and MSK Systems, *IEEE Trans. Commun.*, Vol. COM-28, No. 12, Dec. 1980.

[125] EVERARD, J.K.A. and WILKINSON, A.J.: Highly Efficient Class E Amplifiers, *Proc. IEE Colloq. on Solid-State Power Amplifiers*, 1991.

[126] NARAYANAN, S.: Transistor Distortion Anlaysis Using Volterra Series Representation, *Bell Syst. Tech. J.*, Vol. 67, May-June 1967.

[127] ABUELMA'ATTI, M.T. and GARDINER, J.G.: Approximate Method for Hand Computation of Amplifier Intermodulation Product Levels, *IEE Proc.*, Vol. 128, Pt. G, No. 1, Feb. 1991.

[128] EVERARD, J.K.A. and KING, A.J.: Broadband Power Efficient Class E Amplifiers with a Non-linear CAD Model of the Active MOS Device, *J.*

IERE, Vol. 57, No. 2, March/April 1987.

[129] FAULKNER, M: Spectral Sensitivity of Power Amplifiers to Quadrature Modulator Misalignment, *IEEE Trans. Veh. Tech*, Vol. VT-41, No. 4, Nov. 1992.

[130] WILKINSON, T.A.: An Assessment of the Performance of Linearisation Schemes in the Australian Mobilesat System by Simulation, *Proc. 7th IEE Conf. Mobile Radio and Personnal Communications*, 1991.

[131] JONES, A.E., WILKINSON, T.A.H. and GARDINER, J.G.: Effects of Modulator Deficiencies and Amplifier Nonlinearities on the Phase Accuracy of GMSK Signalling, *IEE Proc.*, Vol. 140, Pt. I, No. 2, April 1993.

[132] KRAUSS, H.L., BOSTIAN, C.W. and RAAB, F.H.: Solid State Radio Engineering, John Wiley & Sons., 1980.

TITLE PROPAGATION STUDIES AND MODULATION TECHNIQUES FOR A DISTRIBUTED ARCHITECTURE RURAL RADIO-TELEPHONE SYSTEM

AUTHOR Simon BROWNE

DEGREE Ph.D

AWARDING BODY Warwick University

DATE 1993

THESIS NUMBER DX182463