ATR

AUSTRALIAN TELECOMMUNICATION RESEARCH

VOL. 19, No. 2, 1985

ATR is published twice a year (in May and November) by the Telecommunication Society of Australia. In addition special issues may be published.
ATR publishes papers relating to research into telecommunications in Australia.

CONTRIBUTIONS: The editors will be pleased to consider papers for publication. Contributions should be addressed to the Secretary, ATR, c/- Telecom Australia Research Laboratories, 770 Blackburn Rd., Clayton, Vic., 3168.

RESPONSIBILITY: The Society and the Board of Editors are not responsible for statements made or opinions expressed by authors of articles in this journal.

REPRINTING: Editors of other publications are welcome to use not more than one third of any article, provided that credit is given at the beginning or end as: ATR, the volume number, issue and date. Permission to reprint larger extracts or complete articles will normally be granted on application to the General Secretary of the Telecommunication Society of Australia.

SUBSCRIPTIONS: Subscriptions for ATR may be placed with the General Secretary, Telecommunication Society of Australia, Box 4050, G.P.O., Melbourne, Victoria, Australia, 3001. The subscription rates are detailed below. All rates are post free. Remittances should be made payable to the Telecommunication Society of Australia, in Australian currency and should yield the full amount free of any bank charges.

ATR

AUSTRALIAN TELECOMMUNICATION RESEARCH

ISSN 0001-2777

VOLUME 19, NUMBER 2, 1985

Contents

Challenge • • •

Telecommunications is a key industry throughout the world. It is one of the largest and fastest growing industries, with revenues of service providers growing at 12% per annum and likely to reach $US200 billion by 1985 while the world equipment market is expected to exceed $US100 billion by 1992. Telecommunications form a vital and growing part of the infrastructure of all developed and industrialising nations. The ever-expanding range of information industries are all increasing their dependence on new telecommunications services and technologies.

Against this background, where is Australia placed both now and in the future for its telecommunications services and equipment industry? How do we compare with other developed countries with whom we normally compare ourselves? More importantly, why change and if so, how?

By world standards, Australia's telecommunications services are generally accepted as being amongst the top six in the world taking into account the availability of services and their cost of delivery. However, Australia's telecommunications equipment industry when considering its contribution to Australian technology and/or balance of trade is one of the worst amongst the countries of the OECD. For example considering trade, the Australian export to import ratio for telecommunications products has fallen from 0.16 in 1979/80 to 0.11 in 1983/84 and compares with a figure of 0.15 for Greece or 4.26 for Sweden. Research and development in the industry has fallen from 7.6% of sales in 1976 to 3.6% of sales in 1982/83 which compares with a usual world industry figure of about 8% of sales. Considering one measure of comparative advantage of invention across industries in various countries, Australia's performance in communications is comparable with that of Japan for farm and garden machinery. Some could argue this is a note of warning for Japan's economic future!

The consequences for Australia of this continued situation are the growing dependence on imported technology and increasing foreign debt. For the service providers such as Telecom Australia, a declining capability of local industry will reduce the service provider's capability to provide world competitive telecommunications services. Unlike retailing, telecommunications services and telecommunications technology are synergistic. Companies and countries with the right technology can, given the commercial capability, deliver the best services. This maxim is best demonstrated by service providers in various countries which are vertically integrated with technology development.

In Australia, as in any country, an integrated telecommunications industry entails a healthy service industry and equipment industry without one dominating the other. However, in Australia Telecom alone currently purchases 75% of the industry's output which is both a symptom and a cause of the industry being essentially inward-looking servicing domestic requirements within tariff and other import barriers. Correspondingly Telecom has traditionally been a conservative follower of overseas service developments, not seeing itself as a 'leading edge Telco' as the North Americans would say.

Thus any proposal to change this position MUST involve the service providers in such a way as to maximise the use of their technological expertise and their position in the world market. However, so as not to degrade service provision by way of increased cost or less competitive services, this activity would be best done in a commercially separate organisation. This is truly a challenge for Government, Telecom and industry!

DR. R.P. COUTTS

Performance Estimates for Heterodyne and Homodyne Optical Fibre Communication Systems

G. NICHOLSON and J.C. CAMPBELL

Telecom Australia Research Laboratories

The receiver sensitivity of optical communication receivers is substantially improved by the use of heterodyne or homodyne detection when compared with that obtained with direct detection receivers. This paper outlines simple analytical expressions for the error probability of heterodyne and homodyne digital transmission systems. The concepts of the "shot noise limit" and "quantum noise limit" are discussed. Finally, a brief review of some notable experimental heterodyne and homodyne digital transmission systems is presented.

KEYWORDS: Optical Communications, Coherent Optical Systems

1. INTRODUCTION

Optical fibre digital communication systems used in installations around the world to date are based on the principles of intensity modulation of the optical source and direct detection of the optical intensity at the receiver. Intensity modulation/direct detection (IM/DD) systems transmit digital information by the simple on-off keying of an optical source. The optical frequency band of the source is generally much greater than the bandwidth of the modulating signal. The performance of an IM/DD system is typically of the order of 20 dB worse, in terms of the minimum required received power for a specified bit error rate, than that given by the quantum noise limit (Ref. 1).

The use of heterodyne or homodyne optical communication systems offers the promise of receiver performance up to 10-20 dB better than with IM/DD systems and performance approaching the quantum noise limit (Ref. 1). The improved receiver performance could be translated into significantly longer repeater spans or increased bit rates on existing repeater spans. The performance improvement is better in the longer wavelength region of 1.3-1.6μm than at 0.85μm. Heterodyne and homodyne optical systems are also attractive in the context of using optical amplifiers and frequency-division multiplexing of several signal channels (Ref. 2).

Heterodyne and homodyne optical communication systems are frequently referred to in the literature as "coherent" optical communication systems (Refs.1,2). The use of the term coherent usually refers in this context to the large temporal coherence of the optical sources used, in comparison with the optical source in an IM/DD system. However, the term coherent may also be applicable to the electrical demodulation technique in the receiver for a heterodyne or homodyne system. Consequently, to avoid confusion the use of coherent to generally describe these systems is avoided in this paper.

The paper provides a set of analytical expressions which allow estimation of the performance of heterodyne and homodyne systems, considering the various digital transmission options. The following Section outlines the key components in a heterodyne and a homodyne system. Section 3 contains the derivation of the expressions for evaluating system performance, and concludes with a brief comparison of the digital transmission options. Section 4 discusses the difference between the shot noise limit of performance achievable with a homodyne system and the quantum noise limit of performance for an IM/DD system. Finally, the performances reported for experimental optical fibre systems are reviewed.

2. SYSTEM CONFIGURATIONS

A block diagram of a heterodyne or homodyne optical transmission system is given in Fig. 1. The optical source at the transmitter is a very-narrow linewidth laser, usually a semiconductor laser diode (LD) with appropriate bias current and temperature stabilisation. The optical modulator, which may or may not be separate from the optical source, is the mechanism by which the electric field of the optical signal from the source is modulated. The three basic techniques of modulation are to modulate the amplitude, phase or frequency of the electric field. The modulation techniques are frequently described as amplitude shift keying (ASK), phase shift keying (PSK) and frequency shift keying (FSK) respectively. This paper restricts itself to the consideration of binary modulation schemes, and in this context ASK is modulation by the on-off keying of the optical source. Although this is the same

Paper received 5 March 1985.
Final revision 4 April 1985.

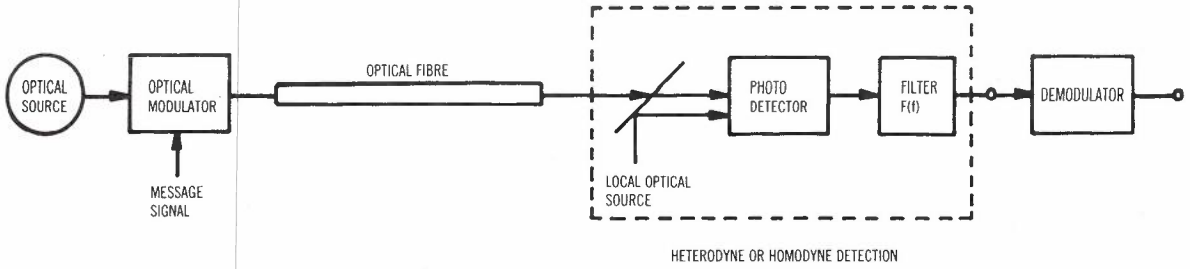


Fig. 1 - Block Diagram of a Heterodyne or Homodyne Optical Transmission System

modulation format as that used for IM/DD systems, as is described below, the laser sources for use in heterodyne/homodyne systems have stringent requirements placed on their spectral characteristics.

Single mode optical fibres are the only practical transmission media for heterodyne or homodyne optical fibre communication systems (Ref. 2). Multimode optical fibres are subject to modal dispersion and speckle pattern problems. In a heterodyne/homodyne receiver the received optical field is combined with a local optical field and the combined field falls on the photodiode. As is described in the next Section, the photodiode responds to the magnitude squared of the complex envelope of the combined field. Because of this relation between the electric field falling on the detector and the resultant output current, the photodiode responds to the difference frequency signal between the received and local optical fields. For a useful signal output containing the transmitted message signal, this difference frequency must be within the electrical frequency response range of the photodiode.

The "beating" between the received and local optical fields in the receiver is the key element in heterodyne and homodyne optical communication systems. Block diagrams of a heterodyne and homodyne optical receiver are given in Figs. 2 and 3. In a homodyne system, this "beat" or difference frequency is zero and the locally generated optical field must closely track the phase of the received field. The message signal is then recovered directly at baseband. In a heterodyne system, the difference or intermediate frequency (IF) is made at least twice the modulation bandwidth of the baseband message signal. The desired message signal can subsequently be recovered

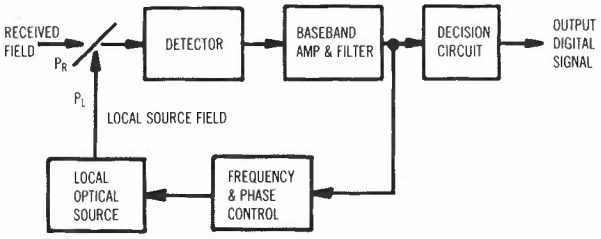


Fig. 3 - Homodyne-Type Optical Receiver

from this IF signal in the electrical domain (as in the case for demodulation of microwave radio signals).

An optical source is not monochromatic or single frequency, but its output has a narrow spectrum of optical frequencies. In the case of a single-longitudinal mode LD, the spectral range is typically of the order of tens to hundreds of MHz depending on the particular device. Heterodyne/homodyne detection requires a very high degree of temporal (i.e. time) coherence between the "carrier" of the received field and the locally generated field at the receiver. This requirement is primarily to ease the task of the local field tracking the "carrier" of the received field, in particular frequency tracking for heterodyne detection, and frequency and phase tracking (i.e. phase locking) for homodyne detection. The task of achieving high temporal coherence, or equivalently a narrow spectral width, of the optical sources is the major problem to be solved before heterodyne and homodyne systems are implemented on a commercial basis.

To demonstrate the high temporal coherence required of the optical sources, consider a heterodyne system with LDs for the transmit and local sources operating at 1.55 μm. This equates to an optical frequency of 1.9 x

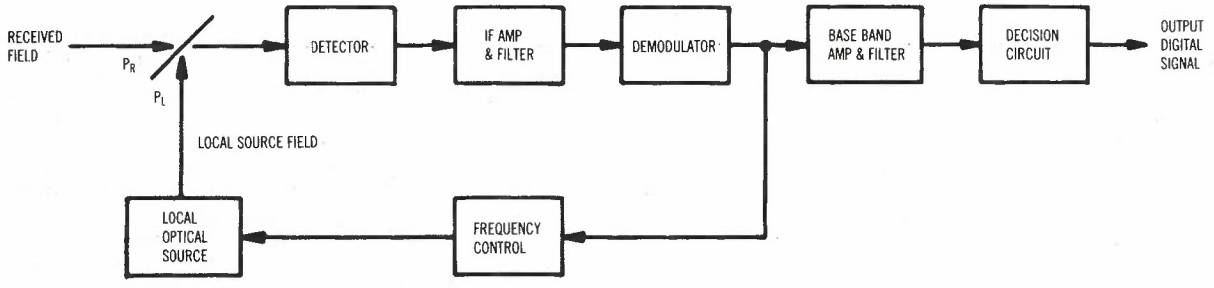


Fig. 2 - Heterodyne-Type Optical Receiver

10^{14} Hz. For an ASK heterodyne system operating at 140 Mbit/s a spectral width less than say 1 MHz may be appropriate. This corresponds to a spectral width of 8×10^{-6} nm in the LD optical output. In contrast, the spectral width of LDs presently used in 1.3 μm IM/DD systems is typically 2-3 nm for multi-longitudinal mode LD's, or about 0.02 nm for single-longitudinal mode LD's (such as the distributed feedback and cleaved-coupled-cavity versions). The subject of achieving very narrow spectral widths is not discussed in this paper; one review of the techniques available is given in Ref. 2.

Heterodyne systems are likely to be implemented before homodyne systems because only frequency difference matching of the fields is required in the optical domain (see Fig. 2). Phase locking of the IF signal is more easily achieved in the electrical domain with well-developed coherent (i.e. synchronous demodulation) techniques. Alternatively, at the expense of some degradation in receiver performance, noncoherent demodulation techniques can be used on the electrical IF signal. One example is the use of an envelope detector in the case of ASK transmission.

In heterodyne and homodyne systems the electric field is modulated at the transmitter and because of the mixing process in the photodiode, the output signal current from the photodetector is directly related to the electric field modulation. Since Maxwell's equations of electromagnetic theory are linear in terms of the electric field, it follows that the system is linear. This is in contrast to an IM/DD system which is, strictly-speaking, nonlinear. Because any time variation in the parameters of the optical sources and photodiode is taken to be small, a reasonable assumption is that the system is time-invariant. Heterodyne and homodyne systems are modelled as linear time-invariant in Section 3 to evaluate the system performance. The digital transmission options that are considered for heterodyne and homodyne receivers are listed in Fig. 4. Note that FSK transmission with a homodyne receiver is not a practical solution because the output signal current from the photodiode is at baseband. In theory, a FSK transmission scheme that may be applicable is based on unequal frequency deviations from the optical frequency carrier for the binary message signals.

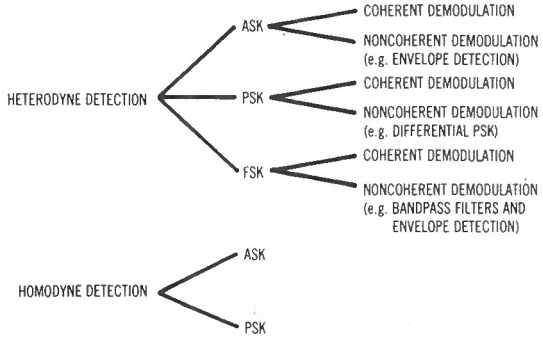


Fig. 4 - Digital Transmission Options

3. EVALUATION OF SYSTEM PERFORMANCE

3.1 Formulation of the Basic System Equations

To describe the time and spatial variations of the electric field associated with an electro-magnetic wave propagating down an optical fibre, the time and spatial variations of the electric field can be assumed to be separable (Ref. 3), that is, any time variation imposed on the electric field will not affect the spatial field pattern associated with the travelling electro-magnetic wave. To describe the time variations associated with the received and locally generated electric fields, let :

Received electric field ≙

$$E_R(t) = \text{Real}\{a_R(t)e^{j2\pi f_R t}\} \quad (1)$$

Local electric field ≙

$$E_L(t) = \text{Real}\{a_L(t)e^{j2\pi f_L t}\} \quad (2)$$

where:

$$a_R(t) = |a_R(t)|\, e^{j\theta_R(t)}$$

$$a_L(t) = |a_L(t)|\, e^{j\theta_L(t)} \quad (3)$$

define the complex envelopes of the respective electric fields, and f_R and f_L are the respective centre frequencies. Now, for each point on the surface of the photo-detector, the electric field is given by the addition of the received and local fields. Again, the time and spatial variations of the field can be separated (Fig. 3), where the time variation is given by :

$$E_T(t) = E_R(t) + E_L(t) \quad (4)$$

and where the combined field pattern is given by the addition at each point on the photo-detecting surface of the received and local source fields. To achieve high output efficiency in the detection process, it is important that the spatial field patterns associated with both these fields be similar and closely aligned on the photo-detecting surface. Spatial misalignment, which can arise through transverse or angular (polarization) misalignments of the fields, will result in an interference field pattern on the photo-detecting surface such that the output efficiency is reduced. However, as the electric fields in single mode optical fibres are closely linearly polarized, the problem of aligning the respective fields reduces to the problem of simply matching the respective polarizations. Experimental systems have demonstrated that this can be achieved by simple polarization tracking techniques (Ref. 4).

The intensity of the resultant incident field is given by (Ref. 3):

$$I(t) = \frac{1}{2Z_w} \mid a_T(t) \mid^2 \qquad (5)$$

where Z_w is the wave impedance of the transmission medium and $a_T(t)$ is the complex envelope of the field $E_T(t)$. From equations (1)-(5) :

$$I(t) = \frac{1}{2Z_w}|a_R(t)|^2 + \frac{1}{2Z_w}|a_L(t)|^2 + \frac{1}{Z_w}|a_L(t)|.$$

$$|a_R(t)|\cos[2\pi(f_R-f_L)t + \theta_R(t)-\theta_L(t)] \qquad (6)$$

With heterodyning (or homodyning), it will be shown that the best receiver sensitivity is achieved when the locally generated field, which is assumed to be of constant envelope, is made much stronger than the received field. This condition occurs naturally in a communications system as the received signal power is small due to the transmission link attenuation. Thus for $|a_L| >> |a_R(t)|$:

$$I(t) = \frac{1}{2Z_w}|a_L|^2 + \frac{1}{Z_w}|a_L||a_R(t)|.$$

$$\cos\,[2\pi(f_R - f_L)t + \theta_R(t) - \theta_L(t)] \qquad (7)$$

3.1.1 The Optical Detection Process. For a simple PIN detector, the output current process is described by a filtered Poisson process whose mean count intensity is proportional to the power incident on the photodetecting surface (Ref. 3). Thus:

$$\text{Mean Count Intensity} \triangleq \lambda(t) = \frac{\eta}{h\nu}.I(t).A + \lambda_d \qquad (8)$$

where η is the quantum efficiency of the photodetector, h is Planck's constant, ν is the optical frequency (Hz) and λ_d is the dark current mean count intensity. A is a parameter (with units of area) whose value is determined by the spatial integration of the field intensity over the photo-detecting surface.

Neglecting the dark current component (an assumption which is reasonable in practice):

$$\lambda(t) = \frac{\eta}{h\nu} = I(t).A = \alpha A I(t) \qquad (9)$$

where for convenience $\alpha = \eta/h\nu$. This paper only considers the use of simple PIN detectors, since the avalanche gain noise introduced by APDs would lead to an increase in the shot-noise-to-signal ratio and thus degraded system performance. The only situation which would warrant the use of an APD is that in which the local source power available is insufficient to make the system "shot noise limited" (Ref. 5). Now define:

$$\text{Local source power} \triangleq P_L = \frac{1}{2Z_w} A \mid a_L \mid^2 \qquad (10)$$

$$\text{Received signal power} \triangleq P_R(t) = \frac{1}{2Z_w} A|a_R(t)|^2 \qquad (11)$$

Then it follows that:

$$\lambda(t) = \alpha P_L + 2\alpha \sqrt{P_R(t)P_L} \cdot \cos\,[2\pi(f_R - f_L)t + \theta_R(t) - \theta_L(t)] \qquad (12)$$

At this point it is appropriate to distinguish between optical heterodyning and homodyning. In optical heterodyning, the local source is controlled to achieve some desired intermediate frequency $f_{IF} = f_R - f_L$ for subsequent demodulation, the intermediate frequency (IF) typically being in the order of a few hundred MHz up to around 2 GHz. However, in optical homodyning, the local source is phase-locked to the received optical carrier. The desired message signal is thus demodulated directly to baseband. As will be shown later, optical homodyning leads to a 3 dB improvement in receiver sensitivity over the corresponding heterodyne scheme.

3.1.2 The Output Current Process. The various signal components that appear at the output of the photodetector can be identified by considering equation (12). Without loss of generality, it is assumed that the photodetector is ideal in that the current pulses produced by the generation of a single electron are ideal impulses, i.e. $e\delta(t)$ where e is the charge of a single electron, and that any bandlimiting produced by the photodetector is included in the following filter F(f). Under this assumption, the power spectrum of the current process output from an ideal photodetector is given by (Ref. 3):

$$S(f) = e^2 \{\alpha^2 P_L^2 \delta(f) + 4\alpha^2 A P_L\, S_m(f) + \alpha P_L\} \qquad (13)$$

where $S_m(f)$ is the power spectrum (suitably defined) of:

$$s(t) = |a_R(t)|\cos[2\pi(f_R-f_L)t + \theta_R(t) - \theta_L(t)] \qquad (14)$$

The first term in (13) can be recognised from (12) as a zero frequency component corresponding to the time average of the mean count intensity $\lambda(t)$. In optical homodyning, the zero frequency component falls directly in the middle of the wanted signal spectrum, a problem which does not arise with optical heterodyning. The second term in (13) represents the desired signal's power spectrum of spectral density $4e^2\alpha^2AP_LS_m(f)$.

The output current process is described by a Poisson process, with mean count intensity given by (12), so that in conjunction with the two components already identified in the power spectrum there is a third component due to the random generation of electrons. This component is known as the shot noise. Because in a heterodyne or homodyne receiver the local optical source power is made much greater than the received signal power ($P_L >> P_R$), it can be readily seen from (12) that the shot noise is essentially determined by only the local optical power P_L. It can therefore be shown (Ref. 3) that the power spectral density of the shot noise is $e^2\alpha P_L$, which is the third component in (13).

Finally, a thermal noise component will also arise at the output of the heterodyne or homodyne receiver, this being primarily due to the electrical amplification and filtering of the photodiode current. However, as the local source power is increased, the shot noise dominates over any thermal noise and the thermal noise can be ignored. Under such a condition the performance of the system is limited only by the shot noise, and the system is said to be "shot noise limited". This is the key point developed in subsequent Sections, which can be exploited for substantial performance improvements over an intensity modulation/direct detection system. The power of the shot noise at the output of the heterodyne (or homodyne) receiver is simply determined by the bandwidth of the filter F(f). It is worth noting from (13) that both the desired signal and shot noise powers are proportional to the local optical power P_L. The signal-to-shot-noise ratio is therefore independent of P_L.

3.2 ASK and PSK Optical Heterodyne Systems

The error probability associated with ASK and PSK optical heterodyne digital transmission can be evaluated by considering the heterodyne system illustrated in Fig. 5. The baseband digital signal to be recovered is given by the complex envelope of the received electric field $E_R(t)$ and will be written as:

$$a_R(t) = \sum_n c_n\, g(t-nT) \qquad (15)$$

where T is the system baud period, $\{c_n\}$ is the transmit symbol sequence and g(t) is the transmit pulse shape.

For the transmit pulse shape g(t), it is assumed that $G(f)^2 = H(f)$, where G(f) is the

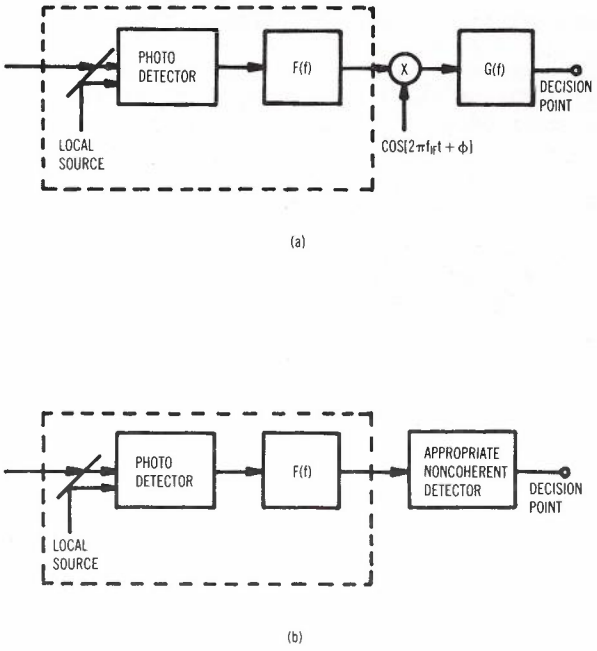


Fig. 5 - Optical Heterodyne Receiver with Coherent and Non-Coherent Demodulation

a. Heterodyne Receiver with Coherent Demodulation

b. Heterodyne Receiver with Non-Coherent Demodulation

Fourier transform of g(t) and H(f) satisfies the Nyquist criterion for zero intersymbol interference. By sharing the pulse shaping equally between the transmitter and receiver, optimum system performance is attained. Also, the normalization of

$$\int_{-\infty}^{\infty} H(f)df = 1$$

is used for computational convenience. The total receiver filtering at baseband is equal to G(f), which has a bandwidth denoted by B. The receiver pulse shaping can equivalently be done either at IF, or at baseband, or with some contribution from both. In this analysis, the pulse shaping is assumed to be performed at IF, i.e. in F(f).

From equations (12) and (13), the desired signal and the shot noise variance that appear at the output of the filter F(f) are given by:

$$\text{Desired signal} \triangleq s(t)$$

$$= 2e\alpha\sqrt{\frac{AP_L}{2Z_w}}\,|a_R(t)|\cos[2\pi f_{IF}t + \theta_R(t) - \theta_L(t)] \qquad (16)$$

$$\text{Shot noise variance} \triangleq \sigma_s^2 = 4e^2\alpha P_L B \qquad (17)$$

The problem now remaining is the

demodulation of the intermediate frequency signal s(t) in the presence of shot noise. Because the shot noise originates from a significant local source power (typically greater than -15 dBm), the statistics of the shot noise can be considered as Gaussian. Two receiver structures for the demodulation of the signal s(t) are of interest:

(i) Coherent demodulation with the baseband filter matched to the transmitted pulse shape (Fig. 5(a)).

(ii) Non-coherent demodulation (if applicable) (Fig. 5(b)).

It is emphasized that the terms coherent and non-coherent refer to the method of electrical demodulation of the intermediate frequency signal s(t).

Finally, for the results that follow, it is convenient to define P_s as the time averaged received signal power. Assuming that the transmitted symbol sequence $\{c_n\}$ is random, then:

$$P_S = E[P_R(t)] = \frac{A\ E[c_n^2]}{2TZ_w} \tag{18}$$

because of the normalization for $|G(f)|^2$.

3.2.1 Heterodyne Performance with Coherent Demodulation. Referring to Fig. 5(a), the desired signal at the output of the baseband filter G(f) (i.e. the decision point) is given by:

$$z(t) = e\alpha\sqrt{\frac{AP_L}{2Z_w}}\ \sum_n c_n\ h(t-nT) \tag{19}$$

for coherent demodulation and using the normalization of h(o) = 1. Considering the shot noise, the noise power spectrum at the output of the filter F(f) is, from equation (13):

$$S_n(f) = \begin{cases} e^2\alpha P_L, & |f| \leq f_{IF} \pm B \\ 0, & \text{elsewhere} \end{cases} \tag{20}$$

Thus the noise power at the decision point is:

$$\sigma_n^2 = \frac{1}{2}\int_{-\infty}^{\infty} S_n(f)|G(f)|^2\ df = \frac{e^2\alpha P_L}{2} \tag{21}$$

Using equations (19) and (21), the performance of ASK and PSK systems can be determined. For example, consider an optical heterodyne receiver with coherent demodulation of a 2PSK signal. At the optimum sampling time the signal at the decision point will assume the values $\pm de\alpha\sqrt{AP_L/2Z_w}$ where the symbols c_n take on the values $\pm$ d. Assuming the decision threshold is set at the optimum of zero (for 2PSK), the probability of error P_E is given by:

$$P_E = Q\left[\frac{de\alpha\sqrt{AP_L/2Z_w}}{\sigma_n}\right] \tag{22}$$

where Q(x) is the Gaussian error function (area under Gaussian tail). Substituting gives:

$$P_E = Q\left[\left(\frac{2\eta P_S T}{h\nu}\right)^{1/2}\right] \tag{23}$$

TABLE 1 - Summary of the Expressions for the Probability of Error of Homodyne and Heterodyne Systems.

Modulation / Receiver	ASK	FSK	PSK
Homodyne	$Q[(\frac{2\eta P_S T}{h\nu})^{1/2}]$	-	$Q[(\frac{4\eta P_S T}{h\nu})^{1/2}]$
Heterodyne + Coherent Demodulation	$Q[(\frac{\eta P_S T}{h\nu})^{1/2}]$	$Q[(\frac{1.21\eta P_S T}{h\nu})^{1/2}]$	$Q[(\frac{2\eta P_S T}{h\nu})^{1/2}]$
Heterodyne + Non-Coherent Demodulation**	$0.5\exp\frac{-\eta P_S T}{2h\nu}$	$0.5\exp\frac{-\eta P_S T}{2h\nu}$	$0.5\exp\frac{-\eta P_S T}{h\nu}$ *

* Probability of Error for DPSK

** Results apply for P_e small, i.e. less than about 10^{-5}

Table 1 presents the expressions for the error probabilities for ASK (i.e. on-off keying of the optical carrier) and that for binary PSK.

3.2.2 Heterodyne Performance with Non-Coherent Demodulation. For each of the digital heterodyne systems presented in the previous Section, a corresponding heterodyne system exists but with non-coherent demodulation of the intermediate frequency signal s(t). The calculation for the error probability is more complex than the corresponding calculation with coherent demodulation and will not be presented in this paper. However, the method of calculation can be found in standard communication text books (Ref. 6) and used in conjunction with the expressions derived in the previous Sections. The results for the corresponding PSK system (DPSK (Ref. 6)) and ASK system (which employs envelope detection) are also presented in Table 1.

3.3 FSK Optical Heterodyne Systems

Binary FSK involves the transmission of two different frequencies for the transmission of the "marks" and "spaces". Of the various receiver structures possible, this paper considers the optimum receiver, which consists of two receivers, one tuned to the "mark frequency" and the other to the "zero frequency". Also, the frequency spacing is assumed to be the optimum (where the cross-correlation of the two waveforms is negative) rather than the case where the signal waveforms are orthogonal. However, the penalty in orthogonal waveforms over the optimum is less than 1 dB (Ref. 6). Finally, the tuned receivers can employ either coherent demodulation or non-coherent demodulation (e.g. envelope detection) of the intermediate frequency signal.

The error probability associated with the FSK optical heterodyne digital transmission systems can be inferred from the expressions derived earlier (Ref. 6). The results are also presented in Table 1 for comparison with the ASK and PSK homodyne and heterodyne systems. The comparison demonstrates that with ASK and FSK (for non-coherent demodulation) the same probability of error is achieved with the same average transmitted signal power. An advantage of FSK is that the optimization of the decision level is straightforward. However, FSK requires two tuned receivers whereas ASK requires only one.

3.4 Optical Homodyne Digital Transmission Systems

As defined previously, in optical homodyning the local source field is controlled to match in both frequency and phase the received optical carrier. The structure of a homodyne receiver thus takes the form illustrated in Fig. 6.

From equations (15, 16), the desired signal at the output of the filter G(f) (i.e. the decision point) is given by (for ASK and PSK):

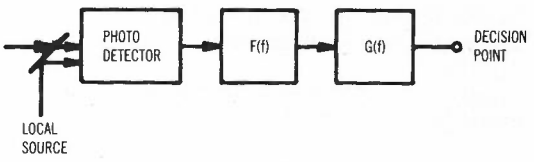


Fig. 6 - Optical Homodyne Receiver

$$z(t) = 2e\alpha\sqrt{\frac{AP_L}{2Z_w}} \sum_n c_n\, h(t-nT) \qquad (24)$$

The shot noise power at the decision point is from (13) given by:

$$\sigma_n^2 = \int_{-\infty}^{\infty} S_n(f)|G(f)|^2\, df = e^2\alpha P_L \qquad (25)$$

Using these equations the performance of homodyne ASK and PSK can be determined. Homodyne FSK is not considered because of the additional receiver complexity involved. The calculations for the respective error probabilities are performed in a similar manner to the example presented in the previous Section, and the results are also presented in Table 1. Of interest is the fact that optical homodyning yields a 3 dB improvement in receiver sensitivity when compared to the corresponding heterodyne system with coherent demodulation of the intermediate frequency signal.

A simple physical explanation for this 3 dB difference is as follows. At the output of the filter following the photodetector, the signal to noise ratio (SNR) with homodyne detection is better by a factor of 6 dB than that with heterodyne detection. This factor arises because:

(i) The desired signal power with homodyne detection is 3 dB greater than that which arises with heterodyne detection (compare equations (16) and (24)).

(ii) The noise power with homodyne detection is 3 dB less than that which arises with heterodyne detection because the noise bandwidth is one-half that with heterodyne detection (compare (21) and (25)).

For a homodyne receiver the output of the filter following the photodetector is the decision point. However, with a heterodyne receiver, in demodulating the intermediate frequency signal an improvement factor of 3 dB can be achieved if coherent demodulation is used. Thus the performance difference between homodyne detection and heterodyne detection with coherent demodulation of the IF signal is a factor of 3 dB in required received power

(or equivalently in SNR at the decision point for the same received power).

3.5 Comparison of Optical Heterodyne and Homodyne Detection

To provide a comparison of the optical heterodyne and homodyne systems, Table 2 lists the systems in order of decreasing performance based on the average received power required for a probability of error of 10^{-9}. Also given is the degradation in receiver performance relative to the optimum binary system which is homodyne PSK.

TABLE 2 - Systems in Order of Decreasing Performance Based on Required Average Received Power

System	Receiver Sensitivity Relative to Homodyne PSK at $P_e = 10^{-9}$ dB
Hom. PSK	0.0
Hom. ASK	3.0
Het. PSK coherent	3.0
Het. DPSK	3.5
Het. FSK coherent	5.2
Het. ASK coherent	6.0
Het. FSK non-coherent	6.5
Het. ASK non-coherent	6.5

With the constant envelope transmission systems of PSK and FSK (actually, only approximately constant envelope as the bandlimiting introduced does give rise to variations in the signal envelope), the required average power equals the required peak power. However, with ASK the required peak power is twice the required average power. Thus for systems which are peak power limited, ASK will suffer a further 3 dB degradation relative to the (approximately) constant envelope digital transmission methods.

Finally, Table 3 presents the receiver sensitivities (for $P_e = 10^{-9}$ and assuming $\eta = 0.6$) that are achieved with homodyne PSK for some likely combinations of bit rate and operating wavelength. The results of Table 2 can be used with those of Table 3 to find the actual receiver sensitivities for heterodyne and homodyne digital transmission systems.

TABLE 3 - Receiver Sensitivities for Homodyne PSK in dBm

Wavelength \ Bit Rate	140 Mbit/s	1.2 Gbit/s
1.3 μm	-64.9	-55.6
1.55 μm	-65.7	-56.4

4. THE SHOT NOISE LIMIT AND THE QUANTUM NOISE LIMIT

This Section considers the limits to performance achievable with uncoded binary optical digital transmission systems. Two concepts are of interest, these being the shot noise limit and the quantum noise limit.

4.1 The Shot Noise Limit

In calculating the probability of error for the homodyne and heterodyne systems considered earlier, the assumption was made that the shot noise power introduced by the local optical source dominates over other noises, which would include the dark current noise from the PIN detector and any thermal noise introduced in the receiver amplifiers. A system which satisfies this condition is said to be "shot noise limited" or operating at the "shot noise limit". In practice, such a condition is realized by supplying sufficient local optical source power. The shot noise limit represents the ultimate performance achievable and is a function of the modulation method and receiver structure, as well as other system parameters.

4.2 The Quantum Noise Limit

When quantum effects are significant, as is the case when transmitting signals at optical frequencies, the limit to performance possible is in many situations achieved by simply counting photons at the receiver (Ref. 7). Optimum performance with photon counting is achieved for the case where the transmitter sends some finite energy E for the transmission of a "mark", and zero energy for the transmission of a "space". An ideal photon counter can detect the incidence of a single photon, and the impairments of dark current and thermal noise are neglected. Now, with photon counting, it is only necessary to observe a single photon in the signalling interval T to decide on a "mark" being transmitted. However, during the transmission of a mark, the number of photons which arrive at the receiver follows a Poisson probability distribution (Ref. 7). The only event which will give rise to an error is that where a mark is transmitted, but zero photons are observed. The probability of this event is $e^{-\Lambda}$, where Λ is the mean number of photons arriving during the signalling interval T. Assuming equal probability of transmitting a mark or space, the mean probability of error is:

$$P_e = \frac{e^{-\Lambda}}{2} = \frac{1}{2}\exp\left[\frac{-2\eta P_{av} T}{h\nu}\right] \qquad (26)$$

where P_{av} denotes the average received signal power.

The minimum average power required for a given probability of error can be evaluated, and this minimum average power is referred to as the "quantum noise limit". It is evident that as the energy per quantum ($h\nu$) increases, then so does the probability of error for a given received average power. In

fact, it is the quantum nature of the transmitted energy that gives rise to the possibility of any error, hence the name quantum noise limit.

In considering the limits to performance achievable with uncoded binary optical digital transmission systems, the question arises as to which is better, photon counting or homodyne PSK (the optimum homodyne system - see Table 1). For homodyne PSK and assuming P_e is small (less than 10^{-5}):

$$P_e \approx \frac{1}{2}\sqrt{\frac{h\nu}{2\eta P_{av} T}}\exp\left[\frac{-2\eta P_{av} T}{h\nu}\right] = \frac{1}{2\sqrt{\pi\Lambda}}\,e^{-\Lambda} \qquad (27)$$

From equations (26) and (27) it is seen that the probability of error for homodyne PSK is $1/\sqrt{\pi\Lambda}$ smaller than that for photon counting. Thus in simple binary data transmission, ideal homodyning performs slightly better in terms of required average received power when compared with ideal photon counting. For example, at 140 Mbit/s and a wavelength of 1.55 μm, the difference in P_{av} is about 0.5 dB for $P_e = 10^{-9}$.

It should be noted that the above comparison is based on a "classical" approach and, to properly investigate the difference in performance between the shot and quantum noise limited situations, a full quantum theoretical approach is required.

5. REVIEW OF EXPERIMENTAL SYSTEMS

A number of laboratories, notably in the United Kingdom, France and Japan, have demonstrated experimental optical fibre systems with heterodyne or homodyne detection. Table 4 summarises the performance of some of the more significant experimental systems. It also demonstrates the substantial performance improvement achieved to date in comparison with intensity modulation/direct detection (IM/DD) systems.

TABLE 4 - Some Notable Experimental Systems

Experimental System	Optical Source and Wavelength	Modulation Technique	Average Received Optical Power, $P_E \sim 10^{-9}$		Comments
			in dBm	Degradation in dB from theory	
FSK Heterodyne 200 Mbit/s	LD with External Cavity 830 nm	LD Bias Current	-44	†	ECL, Japan [5]. Performance Limited by LD Quantum Noise
DPSK Heterodyne 50 Mbit/s	LD with External Cavity 850 nm	$LiNbO_3$ Waveguide Phase Modulator	†	†	CNET France [8] Performance Limited by LD Quantum Noise
ASK Homodyne 140 Mbit/s	LD with Injection Locking 1520 nm	LD Bias Current	-59	4	BTRL, UK [9]
DPSK Heterodyne 140 Mbit/s	LD with External Cavity 1520 nm	$LiNbO_3$ Waveguide Phase Modulator	-56	6	BTRL, UK [9]
ASK Heterodyne 100 Mbit/s	DFB LD 1300 nm	$LiNbO_3$ Waveguide Phase Modulator	-56	6	NEC Japan [10]
PSK Homodyne 140 Mbit/s	H_eN_e Laser 1520 nm	$LiNbO_3$ Waveguide Phase Modulator	-62	4	BTRL, UK [11] Use Optical Phase Locked Loop
IM/DD System* 140 Mbit/s (* for comparison)	LD 1520 nm	Intensity Modulation (i.e. OOK)	-45		with PIN-FET Receiver [9]

† Not available

The performance of the more recent long-wavelength systems is within a few dB of that predicted from theory, i.e. from the analytical expressions giving the optical receiver sensitivity as derived in Section 3. The factors contributing to the difference in performance from theory include intensity and phase noise of the optical sources, various imperfections in the experimental setup, small errors in polarisation matching, etc. (Refs. 9,10). The PSK Homodyne system (Ref. 11) is particularly noteworthy because it is the first use of an optical phase-locked loop to control the phase of the local optical source.

6 CONCLUSIONS

This paper has derived a set of expressions for the error probability of various heterodyne and homodyne optical fibre digital transmission systems. The results from these expressions indicate the substantial performance improvements that can be achieved in comparison with IM/DD systems. The optimum binary system is PSK with homodyne detection. An example of the possible improvement achieved by homodyne PSK is 20 dB for operation at 1.55 μm and 140 Mbit/s.

The ultimate performance of heterodyne and homodyne systems is achieved under the condition that the shot noise dominates over all other noise sources. This performance is compared with that achieved with ideal photon counting in a direct detection receiver, i.e. the quantum noise limit.

7. REFERENCES

1. Okoshi, T. et al., "Computation of Bit-Error Rate of Various Heterodyne and Coherent-Type Optical Communication Schemes", J. Opt. Commun., Vol. 2, 1981, pp. 89-96.

2. Kimura, T. and Yamamoto, Y., "Review: Progress of coherent optical fibre communication systems", Optical and Quantum Electronics, Vol. 15, 1983, pp. 1-39.

3. Gagliardi, T.M., and Karp, S., Optical Communications, John Wiley and Sons, New York, 1976.

4. Hodgkinson, T.G., Smith, D.W. and Wyatt, R., "1.5μm Optical Heterodyne System Operating over 30 km of Monomode Fibre", Electronics Letters, Vol. 18, No. 21, Oct. 1982, pp. 929-930.

5. Saito, S., Yamamoto, Y. and Kimura, T., "S/N and Error Rate Evaluation for an Optical FSK-Heterodyne Detection System Using Semiconductor Lasers", IEEE Journal of Quantum Electronics, Vol. QE-19, Feb. 1983, pp. 180-193.

6. Schwartz, M., Bennett, W.R. and Stein, S., Communication Systems and Techniques, McGraw-Hill, 1966.

7. Pierce, J.R. and Posner, E.C., Introduction to Communication Science and Systems, Plenum Press, New York and London, 1980.

8. Favre, F. and Le Guen, D., "Effect of Semiconductor Laser Phase Noise on BER Performance in an Optical DPSK Heterodyne-Type Experiment", Electronics Letters, Vol. 18, No. 22, Oct. 1982, pp. 964-965.

9. Stanley, I.W. and Midwinter, J.E., "A Review of Techniques for Long Wavelength Coherent Optical Fibre Systems", IEEE GLOBECOM '83 Proceedings, San Diego, Nov. 1983, pp. 21.1.1-5.

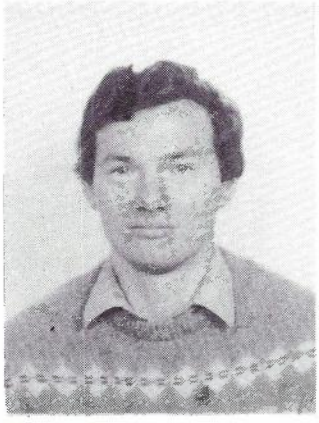
10. Shikada, M. et al., "100 Mbit/s ASK Heterodyne Detection Experiment Using 1.3 μm DFB-Laser Diodes", Electronics Letters, Vol. 20, No. 4, Feb. 1984, pp. 164-165.

11. Malyon, D.J., "Digital Fibre Transmission Using Optical Homodyne Detection", Electronics Letters, Vol. 20, No. 7, March 1984, pp. 281-283.

BIOGRAPHIES

GRANT NICHOLSON was educated at the University of Tasmania, Hobart, Tasmania, where he was awarded the degrees of Bachelor of Engineering with Honours (Electrical) in 1976 and Master of Engineering Science in 1979. He commenced work as an Engineer with Telecom Research Laboratories, Line and Data Systems Section in 1978, working in the area of PCM digital transmission systems. Since 1981 he has been in the Optical Systems Section.

At present, his main interests are in the areas of analysis of optical fibre communication systems and the characterisation of single mode optical fibres.



JOHN C. CAMPBELL was born in Melbourne in 1954. He received the B.E. and M.Eng.Sc degrees in Electrical Engineering from the University of Melbourne in 1976 and 1978 respectively.

In 1978 he joined Telecom Australia, and since working in the Research Laboratories of Telecom Australia, he has been involved with the transmission of digital data signals over both line transmission systems and digital radio systems. Currently, he is an Engineer Class 3 in the Optical Systems Section of the Transmission Branch and is concerned with the application of digital systems to optical bearers in the Australian network.

Modelling and Simulation of Digital Satellite Links

B.R. DAVIS

Department of Electrical & Electronic Engineering
University of Adelaide

Satellite transponders are usually non-linear when operated under conditions which lead to maximum efficiency. The non-linearity of the transponder makes the calculation of link performance much more difficult than in the linear case. This paper considers models and techniques which can be used to determine the signal and noise characteristics of non-linear digital satellite links. Because existing analytical methods are complex and limited in application, the main emphasis is on simulation methods and in particular the use of importance sampling to reduce the time and cost of low probability of error simulations.

KEYWORDS: Satellite Communications
Simulation
Importance Sampling
Digital Communications

1. INTRODUCTION

Digital satellite links are subject to additive noise in both the up link and down link communication paths. In communication satellites such as AUSSAT the satellite is simply a transponder and, while the signals are shifted in frequency and amplified, there is no attempt made to regenerate the digital baseband signal. Hence the noise present in the up link path is amplified and retransmitted by the satellite, and at the ground receiver both the up link and down link noise contribute to the degradation of system performance. In a digital system this degradation is in the form of errors in the demodulated digital data.

If the satellite transponder is carrying frequency division multiple access signals then there is intermodulation interference which will also degrade the system performance. In calculating the performance of the system it may be sufficiently accurate to model the intermodulation interference as noise.

In linear systems it is fairly straightforward, at least in principle, to calculate the effects of signal pulse shape, bandwidth, channel characteristics, and non-ideal filtering on the received signal. The results for digital systems can be presented graphically as eye diagrams in which intersymbol interference and timing accuracy requirements are clearly visible. However more statistical data about the received signal can be gathered than is apparent from the eye diagram. In particular, the probability density function of the signal at the instant the received signal is sampled enables the probability of error with Gaussian noise to be calculated.

In linear systems the effects of additive noise are easily calculated if the noise is Gaussian, since only the mean and autocorrelation function of the noise need be known to fully characterise it, and these quantities are readily calculated for linear systems. Fortunately Gaussian noise is a situation which is often closely approximated in practice. The situation with non-Gaussian or impulse type noise is less straightforward, but simplifying assumptions often lead to tractable and reasonably accurate results.

Most transponders in satellites of the AUSSAT type have travelling wave tube power amplifiers which, while linear at low signal levels, are significantly non-linear when operated at the higher signal levels necessary for efficient operation. Efficiency is an important consideration in the satellite, where the total DC power is limited to what can be provided by solar panels of reasonable dimensions.

The non-linear behaviour of the satellite complicates the analysis of both the signal and noise behaviour of the system to such an extent that a purely analytical approach becomes intractable. In this situation approximate methods or computer simulations become the only viable alternatives. However, these depend on how accurately the system can be modelled and the ability to perform the necessary calculations or simulations in a reasonable time at reasonable cost.

The following Sections consider methods of determining the signal and noise characteristics of a non-linear digital satellite link. The emphasis is on simulation methods since these seem to be the most efficient means of determining system performance when significant non-linearities are present in the satellite transponder.

Paper received 11 June 1985.
Final revision 17 July 1985.

2. SATELLITE SYSTEM MODEL

A model of a digital satellite link is shown in Fig. 1. This model is suitable for either analytical or simulation methods, although the details of each block will usually differ depending on the method.

The digital data input is usually modelled as independent equally likely binary impulses and realised in simulations by pseudo-random sequence generators or pseudo-random number generators.

The transmitter pulse shaping filter generates the baseband pulse shape from an impulse input. In a real system the baseband pulses will not usually be generated by this method, but this model is useful both analytically and for simulation purposes. The transmitted signal consists of the baseband pulse modulated on to a carrier frequency. Whether the modulation method is linear or non-linear, the modulated signal itself is rarely used directly in analysis or simulations. Because the carrier as such carries no information, it is convenient and economical to describe the modulated signal by its complex envelope (or phasor) representation.

The up link noise as received by the satellite transponder is usually modelled as white Gaussian noise, and this is a fairly accurate description of the practical situation. The down link noise is similarly modelled.

The satellite transponder is usually quite complex in practice, involving circulators, amplifiers, filters, frequency changers etc. From a modelling point of view, the frequency change can be ignored since this only affects the carrier frequency, and the satellite transponder can be described as a non-linear amplifier with linear filters before and after it.

If the filters are broadband they may be ignored or merged into the transmitter and receiver filters for the purposes of analysis or simulation.

The non-linearity of the amplifier due to the non-linear characteristics of the travelling wave tube power amplifier is more complex, but is usually modelled as a memoryless non-linearity operating on the complex envelope. The envelope and phase at the output are expressed in terms of the instantaneous values of the envelope and phase at the input. This model has been used extensively (Refs. 1,2) and gives an accurate description of the effect of transponder non-linearities in most situations. Effectively it is a describing function technique which does not attempt to model or explain the actual source of the non-linearity in the device. The characteristics of a typical amplifier are shown in Fig. 2. (Ref. 1).

The demodulator transforms the modulated signal back to a baseband signal. The receiver filter ideally is a filter matched to the transmitted pulse shape. The output of this filter is sampled and converted into the digital output data stream.

3. SIMULATION OF SIGNAL COMPONENTS

This Section deals with the simulation of the system under noise-free conditions. The effects of noise will be considered in Section 4.

3.1 Equivalent Baseband Signals

If v(t) is a radio frequency signal then it can be written in the form:

$$v(t) = R_e\,[\tilde{v}(t)\, e^{j\omega_0 t}] \qquad (1)$$

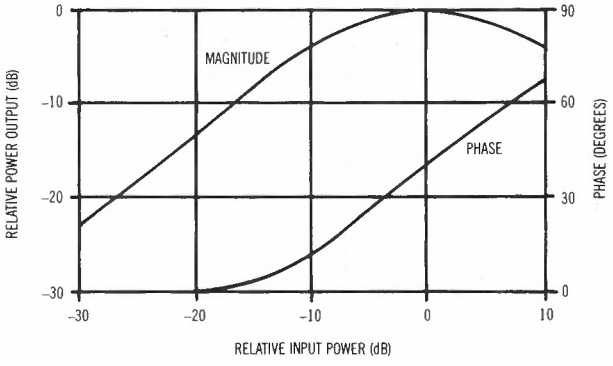


Fig. 2 - Nonlinear Characteristics of a Typical Satellite Amplifier

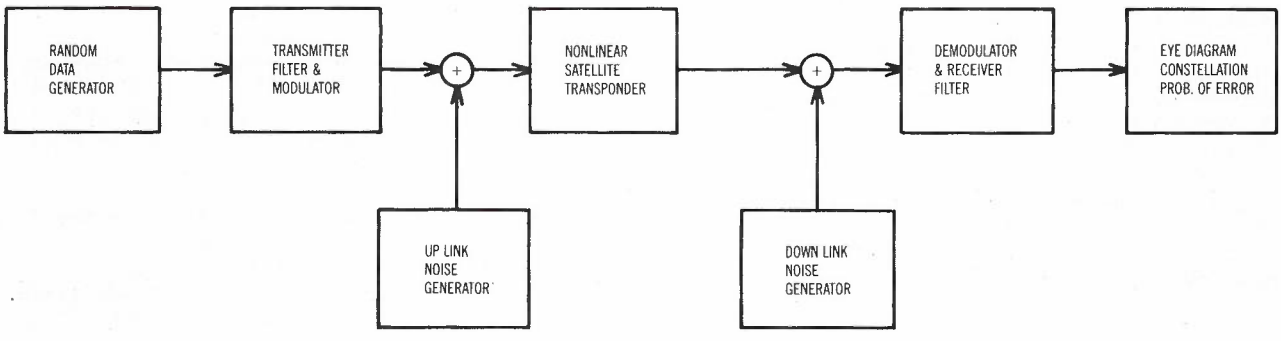


Fig. 1 - Model of Satellite Link Used in Simulations

where $\omega_0 = 2\pi f_0$ is an arbitrary reference frequency and $\tilde{v}(t)$ is the complex envelope or phasor representation of $v(t)$. The complex envelope can be expressed in terms of $v(t)$ and its Hilbert transform $\hat{v}(t)$, viz.

$$\tilde{v}(t) = [v(t) + j\,\hat{v}(t)]\, e^{-j\omega_0 t} \qquad (2)$$

The signal $\tilde{v}(t)$ is an equivalent baseband signal of the bandpass signal $v(t)$, and any filtering operation on $v(t)$ can always be expressed as an equivalent filtering operation on $\tilde{v}(t)$.

The use of the equivalent baseband signal minimises sampling rate requirements and avoids simulating the actual RF signal which is rarely of interest anyway.

3.2 Discrete Time

An actual system operates in continuous time, but a digital simulation operates in discrete time. If the sampling interval is Δ, then the highest frequency which can be sensibly interpreted is the Nyquist frequency equal to $1/2\Delta$ Hz.

In a digital data system with an interval T between pulses, the minimum bandwidth of the baseband signal is $1/2T$ Hz and usually exceeds this. If the number of samples per pulse is K (i.e. $T = K\Delta$), then the Nyquist frequency is $K/2T$. In the simulations a value of $K = 8$ was used, this being a compromise between how accurately the continuous time system is represented and the computation time required for the simulation (Ref 3.).

3.3 Data Generation

Two methods of generating random data were used. The first used a seven bit shift register sequence to generate a pseudo-random binary sequence of period $127 = 2^7-1$. The preliminary simulations used 1024 samples corresponding to 128 data samples. It has been reported (Ref.4) that pseudo-random sequences of length 63 or 127 are sufficient for most simulations.

The second method simply used the random number subroutine available in the computer library. This produces uniformly distributed random numbers in the range $0 \leq x < 1$ and these are easily converted to random binary data by using a threshold of 0.5. This second method was used with later simulations where noise was included and the number of samples used was 16384.

3.4 Modulation and Pulse Shaping

The simulations described here are for coherent and offset quaternary phase shift keying (QPSK) using ideal Nyquist pulse shaping (Ref. 5) with a 50% excess bandwidth. Nyquist pulses with excess bandwidths of 0%, 50% and 100% are shown in Figure 3.

In an ideal communication system the required pulse shape should appear at the receiver output just prior to sampling. It is known that the optimum way to distribute the required filtering is to have transmitter and receiver filters with identical amplitude responses and complementary phase responses (Ref 5.).

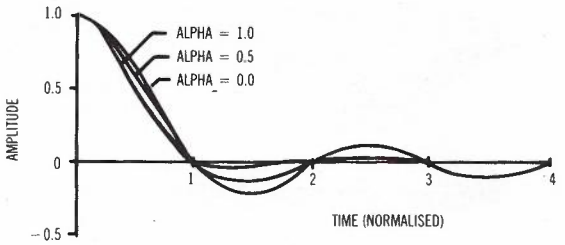


Fig. 3 - Shapes of Nyquist Pulses

In the simulations the filtering was done in the frequency domain by using a discrete Fourier transform (DFT) to convert the complex sequence $\tilde{v}(n)$ into the frequency domain, applying the filtering and then converting back to the time domain. The discrete Fourier transform was implemented by a fast Fourier transform (FFT) algorithm (Ref. 6).

3.5 Transponder Characteristics

The transponder characteristic used in the simulations is shown in Fig. 2. It is assumed that the receiver carrier recovery circuit compensates for the average phase shift through the transponder.

3.6 Eye Diagrams

The eye diagrams for the in-phase and quadrature components of the receiver filter output are plotted by showing transitions which span an interval of two data periods. Those for offset QPSK have the quadrature diagram offset by one half a data period.

Typical eye diagrams are shown in Figs. 4 and 5 for various transponder back-offs. As expected, the eye height is reduced when the satellite transponder is operated in the non-linear region, although the absolute eye opening is greater.

3.7 Signal Constellations

By plotting the locus of the complex envelope corresponding to the receiver output, one obtains the signal constellations shown in Figs. 6 and 7. These correspond to the same parameters as used in Figs. 4 and 5. The reduced envelope variations in offset QPSK are shown clearly in these diagrams, although this does not seem to significantly affect the eye diagrams.

3.8 Power Spectra

The power spectrum at the output of the transponder will be bandlimited if the transponder is operated in the linear region, since the transmitted pulses are bandlimited. Transponder non-linearities will cause spectrum spreading, which will result in

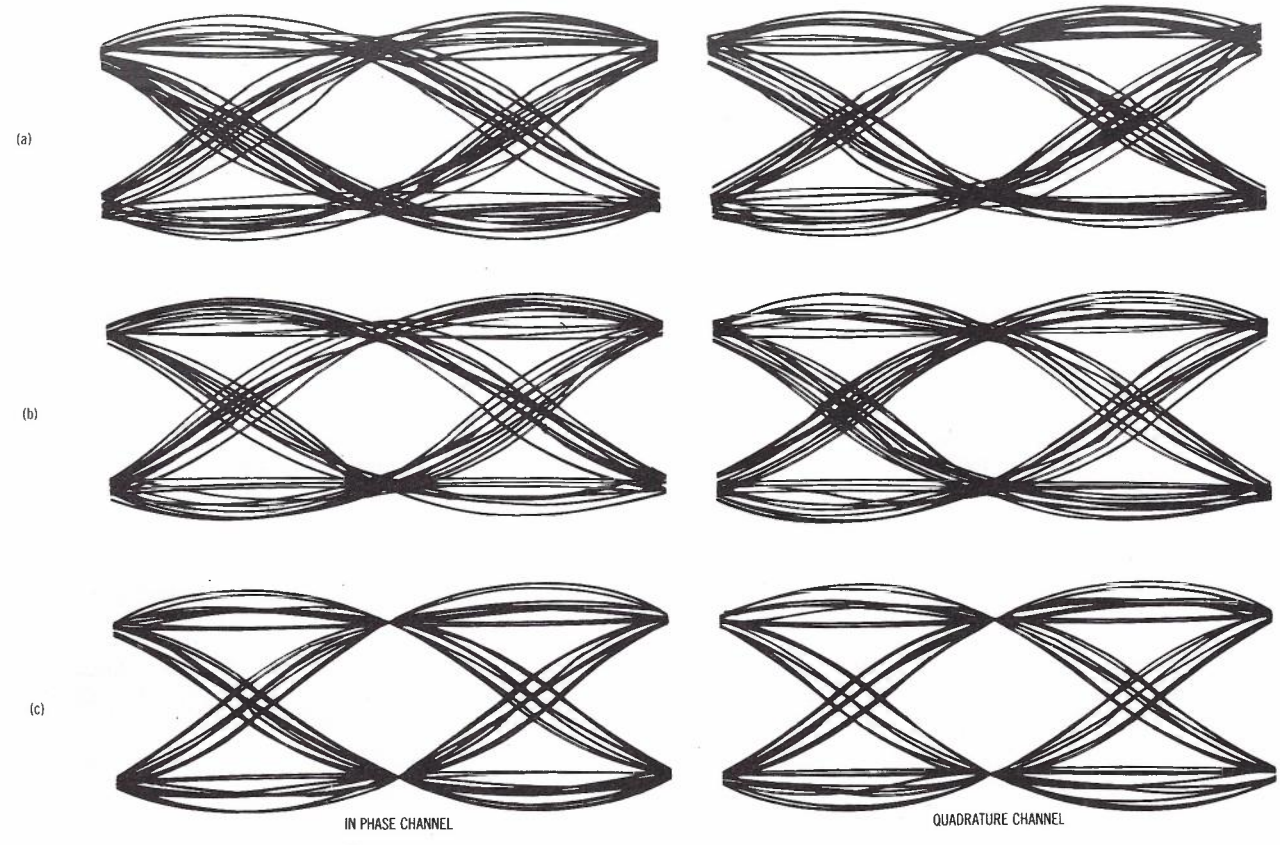


Fig. 4 - Eye Diagrams for Ordinary QPSK

a. Input Power Backoff = 0dB
b. " " " = 10dB
c. " " " = 20dB

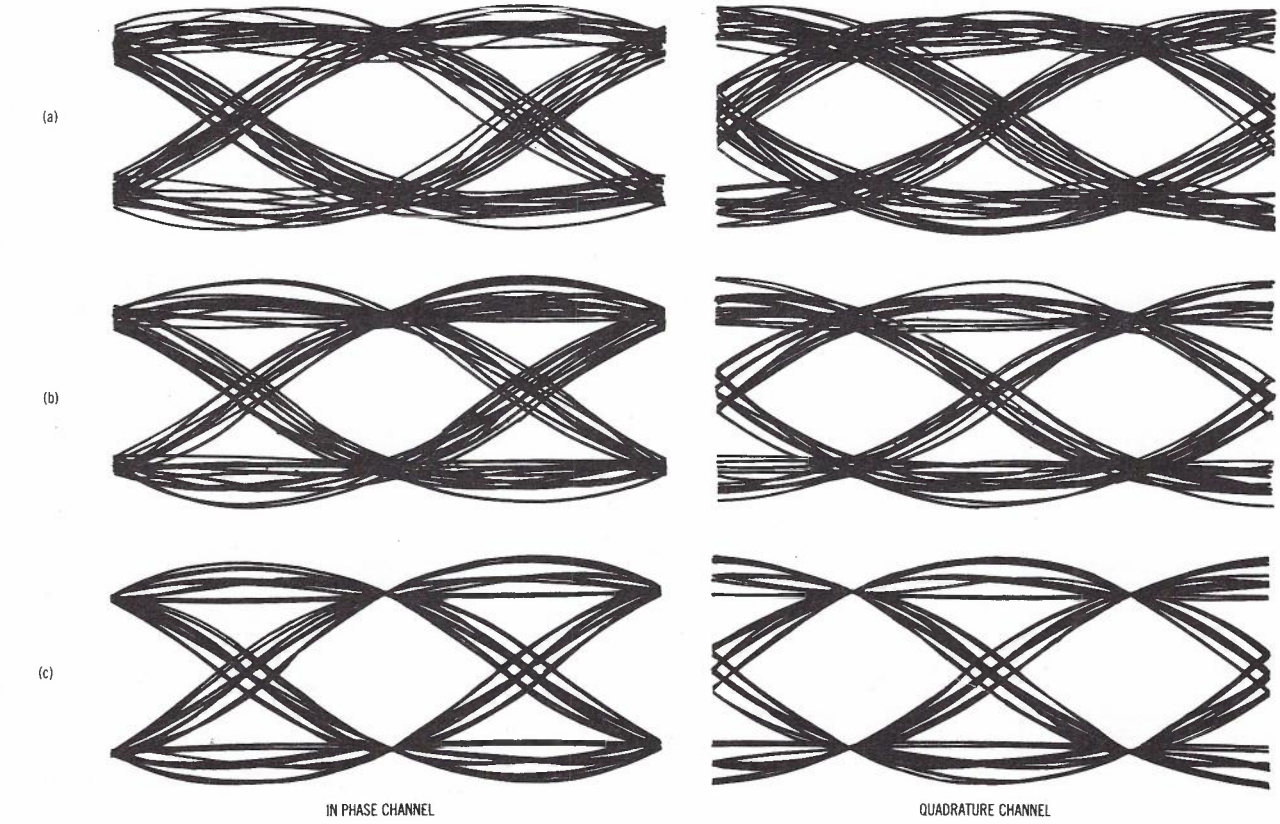


Fig. 5 - Eye Diagrams for Offset QPSK

a. Input Power Backoff = 0dB
b. " " " = 10dB
c. " " " = 20dB

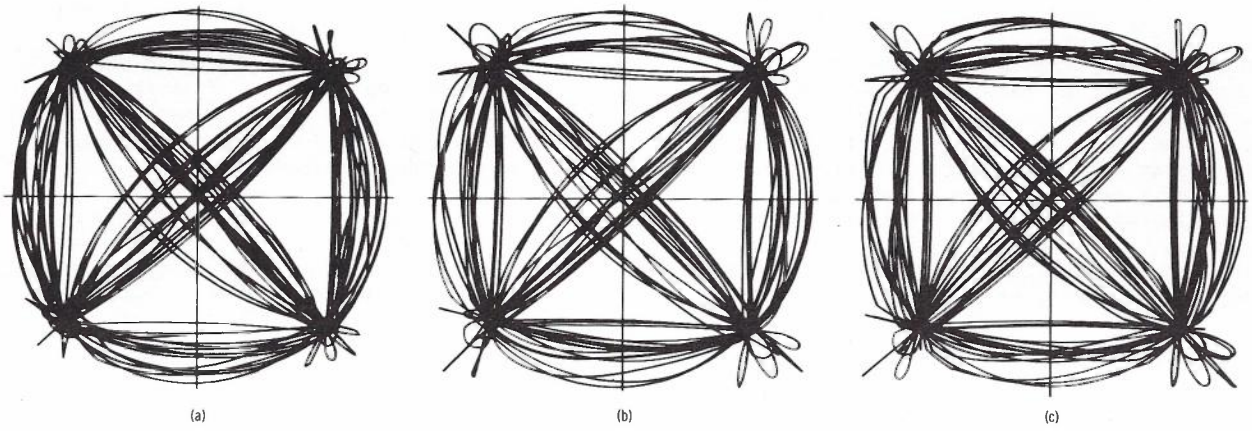


Fig. 6 - Constellations for Ordinary QPSK

a. Input Power Backoff = 0dB
b. " " " = 10dB
c. " " " = 20dB

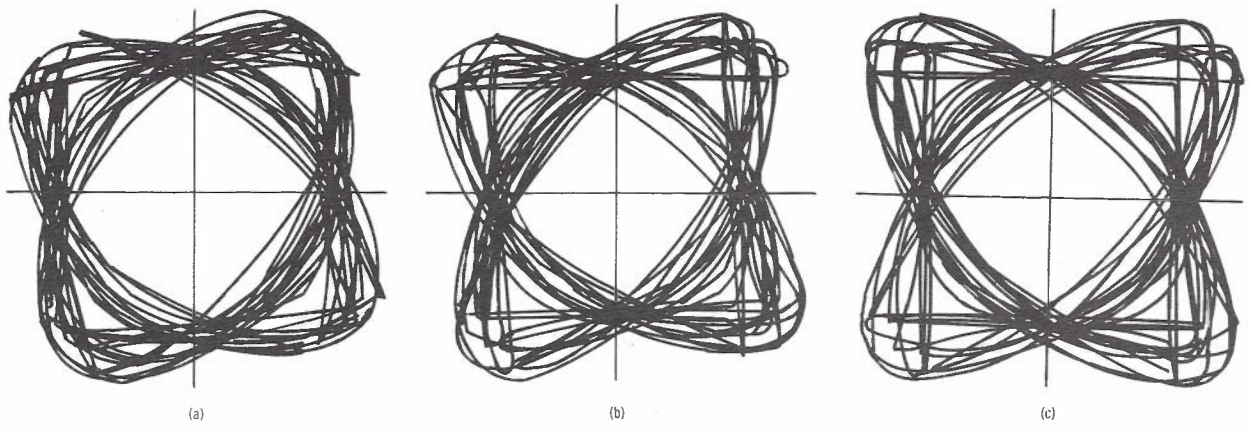


Fig. 7 - Constellations for Offset QPSK

a. Input Power Backoff = 0dB
b. " " " = 10dB
c. " " " = 20dB

adjacent channel interference in an actual system.

In the simulation, the power spectrum of the transponder output can be obtained at the point where the transponder output is converted into the frequency domain in preparation for receiver filtering.

The spectrum is smoothed by a rectangular window ±8 samples either side of the frequency of interest. This yields a standard deviation to mean ratio of about 0.25 (Ref. 6).

Figs. 8 and 9 show the power spectra obtained for various parameters. The vertical scales are normalised and have no absolute value significance. It is apparent that offset QPSK causes marginally less adjacent channel interference, but for both coherent and offset QPSK the level of adjacent channel interference increases with increasing saturation in the transponder. On the frequency axis, a normalised frequency of unity corresponds to the minimum bandwidth for sinc data pulses (i.e. 1/2T Hz). With 8 samples per pulse, the Nyquist frequency (one half the sampling frequency) corresponds to a normalised frequency of 8.

4. MODELLING OF THE EFFECTS OF NOISE

In assessing the effects of noise on a digital satellite link there are clearly three different possibilities. Firstly the problem may be amenable to analysis, either exact or approximate. This will usually be the computationally most efficient, provided the problem is sufficiently tractable.

The second alternative is pure simulation, where all signals present in the actual system are modelled and exist in the simulation. In the case of random signals such as noise, these simulations are usually described as being of the Monte Carlo type.

The third method is a hybrid technique where some of the calculation is done by simulation and some by analysis. The case of down link noise in a satellite system can be approached

this way. Signal simulations of the type described in the previous Section can be used to obtain the statistics of the signal components at the receiver output. Since the system is linear beyond the point where the down link noise is introduced, the effect of this noise is easily calculated provided it is assumed to be Gaussian, a situation well approximated in practice. This method however does not look as promising in the case of up link noise, since the non-linear transponder complicates the noise analysis.

4.1 Analytical Methods

There have been a number of attempts at analytical solutions to the signal and noise characteristics at the output of a non-linear transponder (Refs 2,7).

Shimbo (Ref. 7) derives expressions for intermodulation, AM-PM conversion and noise characteristics of signals passing through a non-linear transponder. These results are expressed in the form of infinite series, which can be simplified in the case of small distortion and small noise. However, his results are basically aimed at deriving the autocorrelation function (and hence power spectrum) of the transponder output.

Hetrakul and Taylor (Ref. 2) derive the probability density function of the noise output from a non-linear transponder using a Gram-Charlier series expansion. However the calculation of the probability density function of the voltage at the output of the receiver filter can not be deduced from this alone because the transponder output noise is non-gaussian. They derive the error probability for a demodulator which consists of taking multiple samples and using a majority logic decision, but this is not the same as for a matched filter receiver.

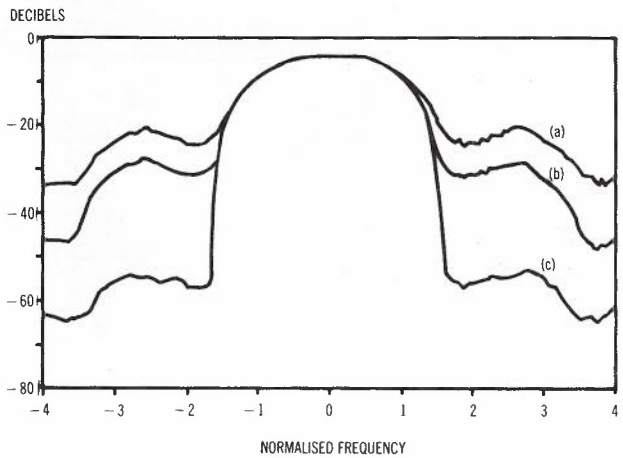


Fig. 8 - Power Spectra for Ordinary QPSK

a. Input Power Backoff = 0dB
b. " " " = 10dB
c. " " " = 20dB

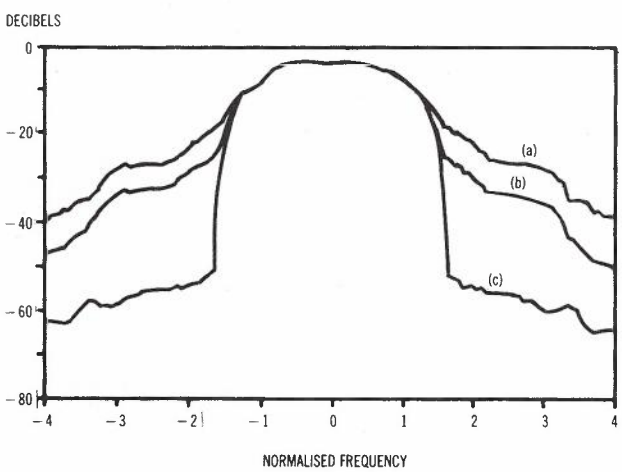


Fig. 9 - Power Spectra for Offset QPSK

a. Input Power Backoff = 0dB
b. " " " = 10dB
b. " " " = 20dB

4.2 Simulation Methods

A paper by Jeruchim (Ref. 4) provides an excellent coverage of the various simulation techniques available for determining the bit error rate in digital satellite systems. Here we will concentrate on Monte Carlo methods and its extensions (such as importance sampling).

4.2.1 Generation of Random Variables In the Monte Carlo simulation the system is modelled as accurately as possible and all random signals are generated by pseudo-random number generators. In a digital communication system both the data and the additive noise are random processes. The data can be generated by pseudo-random binary shift register sequences or by a random number generator of the congruential type which is usually available as a library routine on most computers.

The additive noise is usually required to be Gaussian and there are a number of ways of generating Gaussian random variables. Most computer random number generators produce pseudo-random numbers with a uniform distribution between 0 and 1.

One method of generating Gaussian random variables is by using a non-linear transformation. The accuracy of the distribution of the transformed variable is then only dependent on how close to uniform the original distribution is.

If the Q(x) function is defined as:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(-t^2/2)\, dt \qquad (3)$$

then the variable $y = Q^{-1}(x)$ has a Gaussian distribution of zero mean and unit variance if x is uniformly distributed over (0, 1). This method was used in the computer simulations as an efficient algorithm for computing $Q^{-1}(x)$ was available (Ref 8.). Fig. 10 shows the distribution obtained over 10^5 samples.

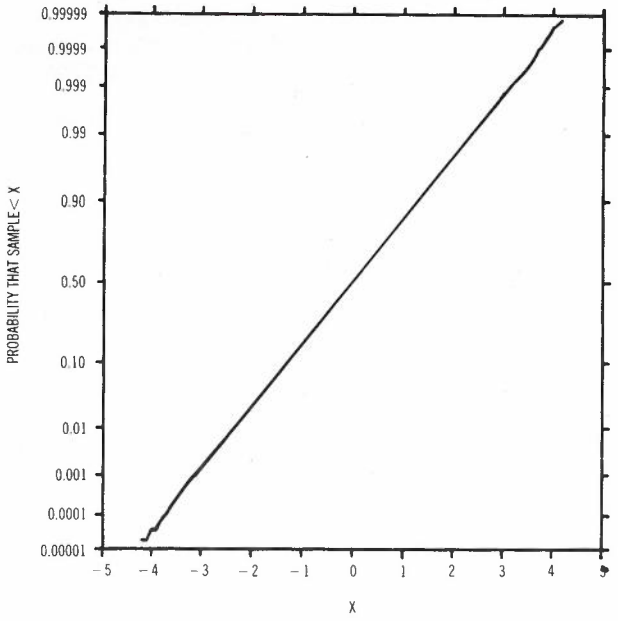


Fig. 10 - Distribution Function of 10^5 Samples

Ideally the variables should be independent. For Gaussian variables it is sufficient that they be uncorrelated since this implies independence. The correlation between the Gaussian samples depends on the correlation properties of the uniform random number generator. Fig. 11 shows the normalised correlation of the Gaussian samples when the standard uniform random number generator in VAX 11/780 FORTRAN was used. For uncorrelated Gaussian variables the normalised correlation has zero mean and 0.003 standard deviation, so Fig. 11 is consistent with the samples being uncorrelated.

4.2.2 Estimation of Error Probability. In a computer simulation the probability of error is estimated by counting error events. If in N total trials, n are in error then the error probability P is estimated as:

$$\hat{P} = n/N \qquad (4)$$

If the N trials are independent, the mean and variance of $\hat{P}$ are:

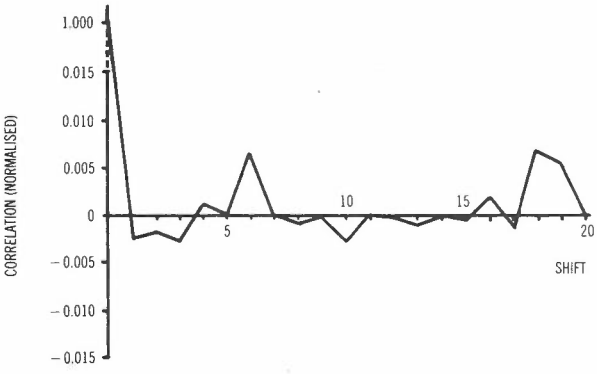


Fig. 11 - Correlation Function of Gaussian Noise Over 10^5 Samples

$$E\{\hat{P}\} = P \qquad (5)$$

$$V\{\hat{P}\} = E\{(P-\hat{P})^2\} = \frac{1}{N} P(1-P) \qquad (6)$$

An accuracy parameter ε may be defined as:

$$\varepsilon^2 = V(\hat{P})/E^2\{\hat{P}\} \qquad (7)$$

For $P \ll 1$, the number N of trials necessary to achieve a specified ε is therefore:

$$N \approx \frac{1}{P\varepsilon^2} \qquad (8)$$

The smaller the value of ε, the more accurate the result. The 95% confidence limits are approximately $\pm 2\varepsilon P$.

Unfortunately this result indicates that as P decreases, the value of N increases inversely. Hence for $\varepsilon^2 = 0.1$, $N = 10^5$ is required if $P = 10^{-4}$, but $N = 10^7$ if $P = 10^{-6}$. This makes the simulation of low error probabilities very costly in computer time, quite apart from the difficulty in ensuring accurate Gaussian random number generators when such small error probabilities are involved.

Another problem which occurs is that of correlation. If the results of one trial are not independent of the results of other trials then the variance increases. This can be thought of as being due to a reduction in the effective number of independent variables. It can be shown (Ref. 9) that an estimate of the effective number of independent variables for a Gaussian process with an ideal low pass spectrum of bandwidth B is given by:

$$N_e \approx 1.14\ a\ BT \qquad (9)$$

where B = bandwidth of process
T = total observation time
$P = P_r(x>a) \ll 1$
$N_e \ll N$

This number tends to be larger than 2BT which is the number which applies when estimating the mean of a process.

The variance of $\hat{P}$ is then $V(\hat{P}) \approx P/N_e$, and hence $N_e = 1/\varepsilon^2 P$. The actual number of trials N will be larger still.

4.2.3 Importance Sampling. This method was described by Shanmugam and Balaban (Ref. 10). The essentials of the method are reproduced here in a slightly different form to the original paper. Importance sampling is also discussed by Jeruchim (Ref. 4).

The idea behind importance sampling is to bias the generation of noise samples towards

those values which cause errors. The generation of small values which cannot cause errors is essentially wasted computational effort. Hence the simulation should concentrate on the generation of "important" events, in this case the occurrence of errors. In order to obtain an unbiased estimate of the true error probability, the artificially enhanced error events must be weighted by a factor less than one.

The probability of error in a digital communication system can usually be expressed as:

$$P = \int_a^\infty p(x)\,dx \qquad (10)$$

where an error occurs if $x > a$ and $p(x)$ is the probability density function of x. This can also be written as:

$$P = \int_a^\infty \frac{p(x)}{q(x)}\,q(x)\,dx$$

$$= \int_{-\infty}^{\infty} w(x)\,q(x)\,dx \qquad (11)$$

where $q(x)$ is any arbitrary probability density function
$w(x) = u(x-a)\,p(x)/q(x)$
= weighting function.

In trials for which x_i, $i = 1, 2 \ldots N$ are generated with a probability density function $q(x)$, the value of P can be estimated as:

$$\hat{P} = \frac{1}{N}\sum_{i=1}^{N} w(x_i) \qquad (12)$$

The mean and variance of $\hat{P}$ are:

$$E\{\hat{P}\} = P \qquad (13)$$

$$V\{\hat{P}\} = \frac{1}{N}\int_{-\infty}^{\infty} [w(x)-P]^2\,q(x)\,dx \qquad (14)$$

By proper choice of $q(x)$ the variance can be made much smaller than for a normal Monte Carlo simulation. The variance can in principle be reduced to zero if $w(x) = P$ for $x > a$, but since this presupposes a knowledge of P it is not of practical use.

A reduction in variance is obtained if $w(x) \ll 1$ for $x > a$. In simulations of digital satellite links where $p(x)$ is usually Gaussian, this is most conveniently achieved by making $q(x)$ Gaussian also, but with a larger variance achieved by multiplying the variable by a scaling factor s.

If $p(x)$ corresponds to a Gaussian distribution of unit standard deviation, then $q(x)$ is a Gaussian distribution of standard deviation s.

The variance of $\hat{P}$ is therefore:

$$V\{\hat{P}\} = \frac{1}{N}\int_a^\infty p^2(x)/q(x)\,dx - P^2/N$$

$$= \frac{s}{N\sqrt{\gamma}}\,Q(a\sqrt{\gamma}) - P^2/N \qquad (15)$$

where $\gamma = 2 - 1/s^2$.

This result is most easily obtained by recognising that $p^2(x)/q(x)$ is $s/\sqrt{\gamma}$ times the probability density function of a Gaussian variable of variance $1/\sqrt{\gamma}$.

However when filters are present in a system, the error will generally depend on a number M of the input random variables. This number M is related to the system memory time. Where M independent input variables contribute to the error probability, the error probability P is given by:

$$P = \int_R p(x_1)\,p(x_2)\,\ldots\,p(x_M)\,dx_1\,\ldots\,dx_M \qquad (16)$$

where R is the region of M dimensional space which corresponds to values of $x_1, \ldots x_M$ which result in an error.

The importance sampling estimate is therefore:

$$\hat{P} = \frac{1}{N}\sum_{i=1}^{N} w(x_{1i})\,w(x_{2i})\,\ldots\,w(x_{Mi}) \qquad (17)$$

where $w(x_{1i}) = u(\underline{x} \in R)\,p(x_{1i})/q(x_{1i})$ and similarly for others. The function $u(\underline{x} \in R)$ is 1 if an error occurs and 0 otherwise. If the sample values for the N trials are independent, then the mean of $\hat{P}$ is $E\{\hat{P}\}=P$ as before and the variance is given by:

$$V\{\hat{P}\} = \frac{1}{N}\int_R \frac{p^2(x_1)}{q(x_1)}\,\frac{p^2(x_2)}{q(x_2)}\,\ldots\,\frac{p^2(x_M)}{q(x_M)}$$

$$dx_1\,\ldots\,dx_M - P^2/N \qquad (18)$$

It is important that M is large enough to include all the significant components of the system memory. It does not matter if M is greater than necessary except that the variance of $\hat{P}$ increases with M. However if M is smaller than the actual system memory then

$\hat{P}$ will be biased, usually being higher (Ref. 10).

The calculation of the optimum scaling factor and the resulting variance reduction factor is only tractable for a linear system, but should give reasonably accurate results for systems with smooth non-linearities.

We suppose that an error occurs if:

$$\sum_{i=1}^{M} h_i x_i > a \qquad (19)$$

where the x_i before scaling are assumed to have unit variance and the weights h_i are normalised so that:

$$\sum_{i=1}^{M} h_i^2 = 1$$

The quantity on the left hand side of equation 19 is therefore Gaussian of unit variance.

The variance of $\hat{P}$ is now:

$$V(\hat{P}) = \frac{1}{N} \left(\frac{s^2}{\gamma}\right)^{M/2} Q(a\sqrt{\gamma}) - P^2/N \qquad (20)$$

To find the minimum we put $\frac{dV}{ds} = 0$. This gives:

$$M(\gamma s^2 - 1)\, Q(a\sqrt{\gamma}) + a\sqrt{\gamma}\, Q'(a\sqrt{\gamma})) = 0 \qquad (21)$$

For x large, $xQ'(x)/Q(x) = -x^2 - 1 + O(1/x^2)$ so this becomes:

$$M(\gamma s^2 - 1) = \gamma a^2 + 1 \qquad (22)$$

If s is large then $\gamma \approx 2$ and this gives:

$$s \approx \left[\frac{a^2}{M} + \frac{M+1}{2M}\right]^{1/4} \qquad (23)$$

A more accurate solution to equation 22 gives an alternative form:

$$s = (A + \sqrt{A^2 - B})^{1/4} \qquad (24)$$

where

$$A = 0.5 \left[1 + \frac{a^2 + 1/2}{M}\right]$$

$$B = a^2/2M$$

For $P \ll 1$ the P^2/N term in equation 20 is usually negligible. The variance for normal Monte Carlo simulation corresponds to having s=1, and the variance with any other value s is reduced by a factor r where:

$$r = \left(\frac{\gamma}{s^2}\right)^{M/2} \frac{Q(a)}{Q(a\sqrt{\gamma})}$$

$$\approx \frac{\gamma^{(M+1)/2}}{s^M} \cdot \exp\{a^2(\gamma - 1)/2\} \qquad (25)$$

Typical scaling factors and reduction factors are shown in Table I.

TABLE I: Scaling and Variance Reduction Factors

	P	10^{-4}	10^{-5}	10^{-6}
	a	3.71902	4.26489	4.75342
M=1	s	3.856	4.384	4.860
	r	317.2	2466	20130
M=10	s	1.451	1.583	1.710
	r	9.294	31.83	127.7

Hence, particularly for small P, the variance of the estimate $\hat{P}$ can be significantly reduced, or conversely for the same variance the number of trials can be reduced by the same factor.

The variance reduction ratio decreases as M increases. Hence if M is much greater than the system memory then $\hat{P}$ will be unbiased but the advantages of importance sampling reduced. However, use of a value of M smaller than the system memory will result in $\hat{P}$ being biased, although its variance may be smaller. In most cases the system memory will not be finite in extent, and the optimum value of M which achieves a compromise between the above factors may have to be chosen experimentally by increasing M until consistent values for $\hat{P}$ are obtained.

4.2.4 Assessment of Importance Sampling. To assess the ability of importance sampling to produce results of reasonable accuracy from a reduced number of trials, it was simulated for simple systems where the result could also be computed analytically for comparison.

The first experiment simulated a unit variance Gaussian distribution obtained by summing M Gaussian variables as in equation 19. The probabilities of exceeding threshold values which would theoretically give error probabilities of 10^{-4} and 10^{-5} respectively were determined by simulation using importance sampling. The results for $N = 10^4$ are shown in Table II.

Various values of M were used with all the h_i in equation 19 being equal. The results indicate that importance sampling does work well. The variance of the sum (as

TABLE II - Simulations Using Importance Sampling

M	Variance (of sum)	Threshold a Theoretical P	3.719 10^{-4}	4.265 10^{-5}
1	0.9797	Scaling factor s Variance factor r Simulation $\hat{P}$	3.856 317.4 1.011×10^{-4}	4.384 2465 1.042×10^{-5}
2	0.9758	Scaling factor s Variance factor r Simulation $\hat{P}$	2.778 136.4 0.9331×10^{-4}	3.144 939.4 0.9411×10^{-5}
5	0.9613	Scaling factor s Variance factor r Simulation $\hat{P}$	1.862 30.73 0.8864×10^{-4}	2.078 154.7 0.8834×10^{-5}
10	0.9253	Scaling factor s Variance factor r Simulation $\hat{P}$	1.451 9.296 0.9730×10^{-4}	1.583 31.82 0.8860×10^{-5}

measured over 10^4 samples) tends to vary a little, probably due to correlation in the random number generator.

It should be noted that the effective number of independent samples is less than N when M > 1 because the output Gaussian variable is generated by a sliding rectangular window over the independent input variables. The output samples are therefore correlated. Using equation 9 the estimated values of N_e for M = 10 are 4100 for a = 3.719 and 4700 for a = 4.265.

The second test involved filtering the data with a Nyquist filter with response as shown in Fig. 3 for $\alpha = 0.5$. Since the impulse response has significant values out to about ± 8 samples from the peak, a value of M = 17 was tried, as well as a smaller value M = 9. A value of N = 16384 was used. The results are shown in Table III.

From equation 9 the effective number of samples are $N_e \approx 4300$ and 4900 for threshold values of 3.719 and 4.265 respectively. The smaller values of M may result in $\hat{P}$ being biased, but the variance factor is larger. For M = 17 the number N = 16384 is really too small to give accurate estimates for $\hat{P}$ since, using the values of N_e estimated above, the accuracy parameter ε^2 is 0.54 and 1.9 respectively for the two threshold values. Even with M = 9 the values of ε^2 are 0.21 and 0.51, indicating that runs larger than N = 16384 may be necessary. With importance sampling we calculate ε^2 from the relation:

$$\varepsilon^2 = \frac{1}{r\, N_e\, P} \tag{33}$$

4.3 Simulation Results for Up Link Noise

The simulation structure used was identical to that described previously in Section 3 except that the number of samples used was 16384 and up link noise was added ahead of the transponder.

For these simulations the random data was generated using a random number generator, and the noise samples were generated by the Q^{-1} transformation method described in Section 4.2.1. The simulations were run 10 times with the same parameters but different random number seeds, giving effective run lengths of 163840 samples. Of course since there are 8 samples per data pulse, and the filter output at the receiver is only sampled once for every data pulse, the number of probability of error trials is one eighth of this, or 20480. However, since the noise contributions to each sampling point will not be significantly correlated, it is expected that the effective number of independent data samples will not be much less than this value. Each 16384 sample run required about one minute of CPU time on a VAX 11/780 computer.

In using importance sampling, memory lengths of 9 and 17 were used. In computing the probability of error of the in-phase signal it was assumed that the quadrature

TABLE III - Importance Sampling with Nyquist Filter

M	Variance (of sum)	Threshold a Theoretical P	3.719 10^{-4}	4.265 10^{-5}
9	1.0034	Scaling factor s Variance factor r Simulation $\hat{P}$	1.500 11.05 0.707×10^{-4}	1.644 40.38 1.069×10^{-5}
17	1.0034	Scaling factor s Variance factor r Simulation $\hat{P}$	1.260 4.327 0.317×10^{-4}	1.343 10.54 0.387×10^{-5}

noise had no effect and vice versa. This is clearly true when the transponder is well below saturation but, as the satellite is operated nearer to saturation, the quadrature noise will have a small effect on the in-phase channel. Inclusion of the quadrature noise effect would mean doubling the value of M and a consequent reduction in the variance factor. It is known that in importance sampling the weights of small contributors to the performance can be ignored without serious error (Ref. 10), and this is supported by the investigations in Section 4.2.4.

Figs. 12 and 13 show the results of simulations for an up link noise level which would result in an error probability of 10^{-4} in a linear system. These results indicate that the error probability increases as the satellite power input is increased to near the saturation level. The value of ε^2 for $P = 10^{-4}$ is 0.044 for M = 9 and 0.11 for M = 17. The bias in the results for M = 9 can be seen when compared with those for M = 17. Even for M = 17, the results may be biased when the transponder is near saturation because the effect of the quadrature noise has been ignored.

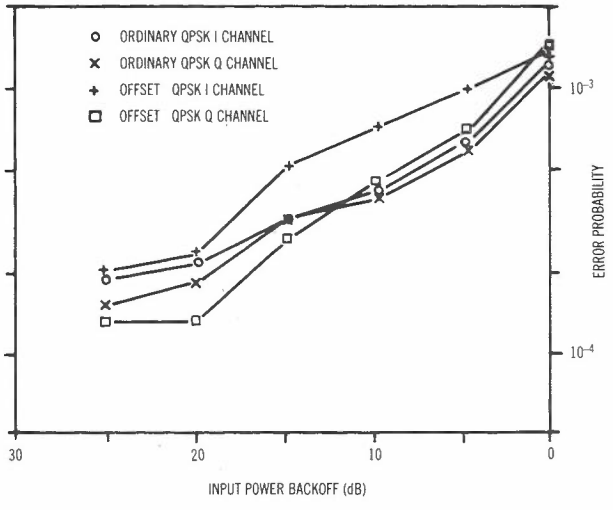


Fig. 12 - Simulation Results Using Importance Sampling and Memory = 9

4.4 Hybrid Methods

Hybrid methods are best used when the noise is introduced at a point in the system where the subsequent processing is linear. This would apply to down link noise in a QPSK digital satellite system. The signal and up link noise behaviour could be determined by simulation, the various statistics measured, and the effect of down link noise calculated from these statistics and the known statistics of the down link noise.

Unfortunately importance sampling is not very useful in obtaining complete probability density functions since the variance reduction achieved in one part of the distribution is balanced by a variance increase elsewhere. However it can tend to make the relative variance of probability estimates more uniform, rather than increasing significantly

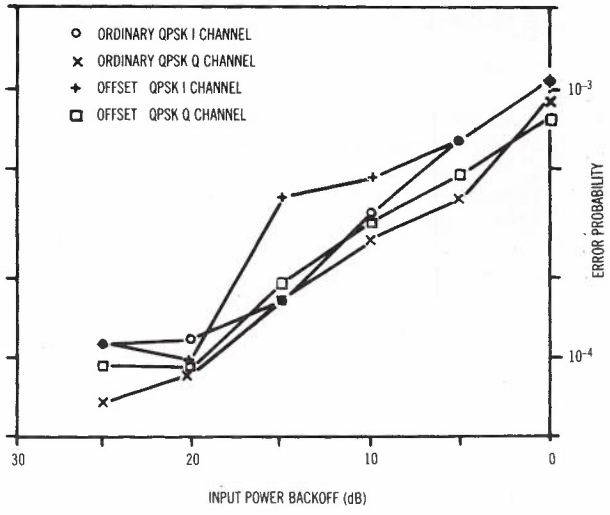


Fig. 13 - Simulation Results Using Importance Sampling and Memory = 17

at the tails of the distribution as occurs with the normal Monte Carlo method.

It may be easier to include the down link noise in the simulation and use importance sampling to reduce the variance of the probability of error estimates. However the runs will have to be longer because the effective value of M will be roughly doubled, and the variance factor will be smaller.

5. SUMMARY AND CONCLUSIONS

The determination of the performance of digital satellite links is relatively straightforward if the system is linear, and analytical methods can be applied.

However, as soon as a non-linear satellite transponder is considered, the analysis becomes very involved and of limited application. The only alternative is simulation and, because of the complexity of the system and the low error rates which are of interest, these simulations can be very expensive in terms of time and money.

Importance sampling has the potential to achieve significant savings in the computer time necessary to obtain reliable simulation results. The disadvantages are that parameters such as memory length and scaling factor have to be chosen correctly, and the accuracy of the simulation results is more difficult to establish. However for small error probabilities the prospect of direct Monte Carlo simulation becomes such a remote possibility that importance sampling is the only viable alternative. It is at these low error probabilities that the method achieves its maximum gains.

6. ACKNOWLEDGEMENT

The work described in this paper was supported by a grant from Telecom Australia.

7. REFERENCES

1. M.A. Marsan, et al, "Digital Simulation of Communication Systems with TOPSIM III", Journal on Selected Areas in Communications, Vol. SAC-2, No. 1, Jan. 1984. pp. 29-41.

2. P. Hetrakul and D.P. Taylor, "The Effects of Transponder Nonlinearity on Binary CPSK Signal Transmission", IEEE Trans. on Communications, Vol. COM-24, No. 5, May 1976. pp. 546-553.

3. L.C. Palmer, "Computer Modelling and Simulation of Communications Satellite Channels", Journal on Selected Areas in Communications, Vol. SAC-2, No. 1, Jan. 1984. pp. 89-101.

4. M.C. Jeruchim, "Techniques for Estimating the Bit Error Rate in the Simulation of Digital Communication Systems", Journal on Selected Areas in Communications, Vol. SAC-2, No. 1., Jan. 1984. pp. 153-170.

5. S. Haykin, "Communication Systems", John Wiley, 1983, pp. 470-476.

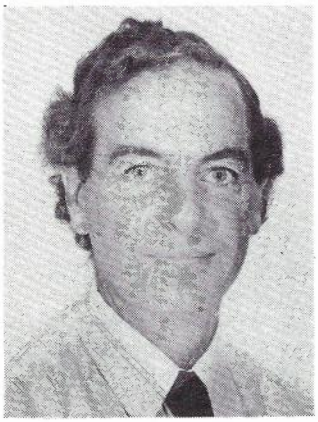
6. D.E. Newland, "Random Vibrations and Spectral Analysis", Longman, 1975, p.142, 220.

7. O. Shimbo, "Effects of Intermodulation, AM-PM Conversion and Additive Noise in Multicarrier TWT Systems". Proc. IEEE, Vol. 59, No. 2, Feb. 1971. pp.230-238.

8. M. Abramowitz and I.A. Stegun, "Handbook of Mathematical Functions", Dover 1965, p. 933.

9. B.R. Davis, "The Effect of Correlation on Probability of Error Estimates in Monte Carlo Simulations". Submitted to IEEE Transactions on Communications.

10. K.S. Shanmugam and P. Balaban, "A Modified Monte Carlo Simulation Technique for the Evaluation of Error Rate in Digital Communication Systems", IEEE Trans. on Communications, Vol. COM-28, No. 11, Nov. 1980. pp.1916-1924.



BIOGRAPHY

BRUCE R. DAVIS has been with the University of Adelaide since 1964 and at present is a Senior Lecturer in Electrical Engineering. His research interests are in the field of communication systems. During 1970 he was with Bell Laboratories, Holmdel, New Jersey, studying various aspects of mobile radio communications, and again in 1977 when he was involved in satellite system research. He is a member of the IEEE and holds the following degrees: B.E. 1960, B.Sc. 1963, Ph.D. 1969, University of Adelaide.

A Numerical Method to Evaluate Blocking Probability for the M(t)/M/N Loss System

A.J. COYLE and M.N. YUNUS

Department of Applied Mathematics
University of Adelaide

A numerical procedure for evaluating the blocking probability for the M(t)/M/N loss system is given. This procedure is based on a method which discretizes continuous arrival rate functions. By applying this discrete method to a matrix differential equation corresponding to the birth-and-death equations, we derive the blocking probability as a function of matrix exponentials. This paper gives an efficient numerical scheme to compute this function and hence the blocking probability.

KEYWORDS: Non-Stationary, Erlang Loss System, Blocking Probability

1. INTRODUCTION

We first consider the birth-and-death equations, a set of differential-difference equations (d.d.e.), associated with the M(t)/M/N loss system (that is with time-varying arrival rates but constant service rates). These d.d.e.'s are derived from the Chapman-Kolmogorov equations. If $p_i(t,N)$ is the probability that at time t, i channels out of a total of N channels in the system are occupied and $\lambda(t)$ is the arrival rate function, we have, after normalizing the mean service rate to make it unity,

$$\dot{p}_0(t,N) = -\lambda(t)p_0(t,N) + p_1(t,N)$$

$$\dot{p}_i(t,N) = \lambda(t)p_{i-1}(t,N) - [\lambda(t) + i]p_i(t,N) + (i+1)p_{i+1}(t,N) \quad \text{for } 0<i<N$$

$$\dot{p}_N(t,N) = \lambda(t)p_{N-1}(t,N) - Np_N(t,N) \qquad (1)$$

To obtain the blocking probability for any initial condition one needs to solve the above set of d.d.e.'s. An earlier paper (see Ref. 2) explains why simple general solutions cannot be obtained. In Ref. 1, Jagerman obtains an integral equation for the blocking probability which has a very complicated form. Therefore it is desirable that a numerical scheme for the above d.d.e.'s be devised. The above set of d.d.e.'s can be written as a matrix d.e., namely

$$\dot{\underset{\sim}{p}}(t,N) = A(t)\underset{\sim}{p}(t,N) \qquad (2)$$

Paper received 30 October 1984.
Final revision 24 April 1985.

where

$$\underset{\sim}{p}(t,N) = [p_0(t,N)\ p_1(t,N)\ \ldots\ p_N(t,N)]^T$$

with $p_N(t,N)$ being the blocking probability, and

$$A(t) = \begin{bmatrix} -\lambda(t) & 1 & 0 & \cdots & 0 & 0 \\ \lambda(t) & -\lambda(t)-1 & 2 & \cdots & 0 & 0 \\ 0 & \lambda(t) & -\lambda(t)-2 & \cdots & 0 & 0 \\ 0 & 0 & \lambda(t) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & N-1 & 0 \\ 0 & 0 & 0 & \cdots & -\lambda(t)-(N-1) & N \\ 0 & 0 & 0 & \cdots & \lambda(t) & -N \end{bmatrix}$$

To solve the problem a solution to the above matrix d.e. must be obtained. However, at the moment, there is no general method for solving matrix d.e.'s with a time-varying coefficient matrix, A(t). We therefore use the approach given in Ref. 2, where the arrival rate function $\lambda(t)$ is discretized into a series of step functions made up of 'bursts' of constant traffic of different magnitude, as shown in Fig. 1. The time subinterval we use is Δt. Therefore we have at $t=\Delta t$

$$\dot{\underset{\sim}{p}}(\Delta t,N) = A_1.\underset{\sim}{p}(\Delta t,N) \qquad (3)$$

where A_1 is the matrix A(t) but with $\lambda(t)$ replaced by λ_1. Similarly, let A_k, $k=2,3,\ldots$ be the same matrices but with λ_1 replaced by λ_k. We have as a solution at $t=\Delta t$

$$\underset{\sim}{p}(\Delta t,N) = \exp(A_1\Delta t).\underset{\sim}{p}(0,N) \qquad (4)$$

where $\underset{\sim}{p}(0,N)$ is the initial probability distribution vector. Likewise for $t=k\Delta t$, we have

$$\dot{\underset{\sim}{p}}(k\Delta t,N) = A_k.\underset{\sim}{p}(k\Delta t,N)$$

$$\underset{\sim}{p}(k\Delta t,N) = \exp(A_k\Delta t).\underset{\sim}{p}((k-1)\Delta t,N)$$

$$= \exp(A_k\Delta t).\exp(A_{k-1}\Delta t).\underset{\sim}{p}((k-2)\Delta t.N)$$

$$= \ldots$$

$$= \left\{\prod_{j=0}^{k-1} \exp(A_{k-j}\Delta t)\right\}.\underset{\sim}{p}(0,N) \qquad (5)$$

Therefore to calculate $\underset{\sim}{p}(k\Delta t,N)$ we need to be able to calculate the matrix exponentials; $\underset{\sim}{p}(0,N)$ can be set as required.

2. THE NUMERICAL PROCEDURE

Since A is a tridiagonal matrix there is a stable, fast method for calculating $\exp(A\Delta t)$, which gives accurate results. This method applies the implicit QL algorithm described in Ref. 3 which is an algorithm for determining the exponential of a symmetric tridiagonal matrix.

Thus, if we can find a diagonal matrix D and a symmetric tridiagonal matrix S such that $A=D^{-1}SD$ then

$$\exp(A) = D^{-1}\exp(S)D$$

where $\exp(S)$ can be calculated using the QL algorithm. Matrices D and S are obtained by solving the system $DA = SD$. Let

$$D = \begin{bmatrix} d_1 & & & & 0 \\ & d_2 & & & \\ & & d_3 & & \\ & & & \ddots & \\ 0 & & & & d_{N+1} \end{bmatrix}, \quad S = \begin{bmatrix} s_{11} & s_{12} & & & 0 \\ s_{12} & s_{22} & s_{23} & & \\ & s_{23} & s_{33} & \ddots & \\ & & \ddots & \ddots & s_{N,N+1} \\ 0 & & & s_{N,N+1} & s_{N+1,N+1} \end{bmatrix}$$

$$DA = \begin{bmatrix} -\lambda d_1 & d_1 & 0 & \cdots & 0 & 0 \\ \lambda d_2 & -(\lambda+1)d_2 & 2d_2 & \cdots & 0 & 0 \\ 0 & \lambda d_3 & -(\lambda+2)d_3 & \cdots & 0 & 0 \\ 0 & 0 & \lambda d_4 & \cdots & 0 & 0 \\ \vdots & & & & & \\ & & & & (N-1)d_{N-1} & 0 \\ & & & & -(\lambda+N-1)d_N & Nd_N \\ 0 & 0 & 0 & \cdots & \lambda d_{N+1} & -Nd_{N+1} \end{bmatrix}$$

and

$$SD = \begin{bmatrix} d_1s_{11} & d_2s_{12} & 0 & \cdots & 0 & 0 \\ d_1s_{12} & d_2s_{22} & d_3s_{23} & \cdots & 0 & 0 \\ 0 & d_2s_{23} & d_3s_{33} & \cdots & 0 & \\ \vdots & & & & & \\ & & & & d_Ns_{N-1,N} & 0 \\ & & & & d_Ns_{N,N} & d_{N+1}s_{N,N+1} \\ 0 & 0 & 0 & \cdots & d_Ns_{N,N+1} & d_{N+1}s_{N+1,N+1} \end{bmatrix}$$

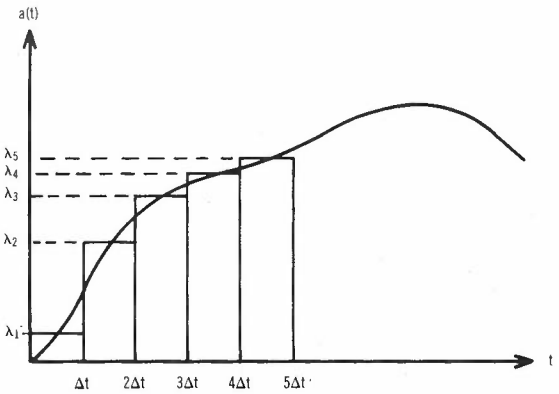


Fig. 1 - How the Continuous Arrival Stream Function is discretized.

If one solution for this system is found then possible solutions are available by multiplying D by a constant c, say, and dividing S by c^2. We shall take d_1 to be one. Since

$$(-\lambda-i+1)d_i = d_is_{ii},\ i<N$$

$$-Nd_{N+1} = d_{N+1}\,s_{N+1,N+1}$$

then

$$-\lambda-i+1 = s_{ii},\ i<N$$

$$-N = s_{N+1,N+1}$$

and all diagonal elements of S can be found. Also

$$d_1 = d_2s_{12}\ ,\ d_1s_{12} = \lambda d_2$$

then

$$d_2 = 1/s_{12} = s_{12}/\lambda$$

$$\therefore\quad s_{12}/\lambda = 1/s_{12}$$

$$\therefore\quad s_{12} = \sqrt{\lambda} \text{ and } d_2 = 1/\sqrt{\lambda}$$

Also,

$$2d_2 = d_3s_{23},\quad \lambda d_3 = d_2s_{23}$$

$$\therefore\ d_3 = \frac{2(1/\sqrt{\lambda})}{s_{23}} = \frac{s_{23}(1/\sqrt{\lambda})}{\lambda}$$

$$\therefore\ s_{23} = \sqrt{2}\sqrt{\lambda} \text{ and } d_3 = \frac{\sqrt{2}}{\lambda}$$

and so on until we obtain

$$D = \begin{bmatrix} 1 & & & & \\ & \frac{\sqrt{1}}{\sqrt{\lambda}} & & & \\ & & \frac{\sqrt{2!}}{(\sqrt{\lambda})^2} & & \\ & & & \ddots & \\ & & & & \frac{\sqrt{N!}}{(\sqrt{\lambda})^N} \end{bmatrix}$$

$$S = \begin{bmatrix} -\lambda & \sqrt{1}\sqrt{\lambda} & & & \\ \sqrt{1}\sqrt{\lambda} & -\lambda-1 & \sqrt{2!}\sqrt{\lambda} & & \\ & \sqrt{2!}\sqrt{\lambda} & -\lambda-2 & \ddots & \\ & & \ddots & \ddots & \sqrt{N!}\sqrt{\lambda} \\ & & & \sqrt{N!}\sqrt{\lambda} & -N \end{bmatrix}$$

The results obtained using the implicit QL algorithm are to machine accuracy, and so the method for finding $\exp(A\Delta t)$ will also be very accurate. Writing $\underset{\sim}{p}_k = \underset{\sim}{p}(k\Delta t.N)$ we have

$$\underset{\sim}{p}_{k+1} = \exp(A_{k+1}\Delta t).\underset{\sim}{p}_k$$

$$= D_{k+1}^{-1} \exp(s_{k+1}\Delta t) D_{k+1} D_k^{-1} \exp(s_k\Delta t) D_k \ldots D_1^{-1} \exp(s_1\Delta t) D_1 . \underset{\sim}{p}_o \qquad (6)$$

Due to the high accuracy with which we can compute the exponential of the matrix A, the errors involved in calculating the state probabilities for the nonstationary process are mainly due to the discretization of the nonstationary arrival processes.

We can get an upper and lower bound for the computed blocking probabilities by using the maximum and minimum values of $\lambda(t)$ during every time interval $(t,t+\Delta t)$ as in Fig. 2. The calculated value of $P_N(t,N)$, the probability of blocking, will be at a maximum when the maximum value of $\lambda(t)$ is used in the procedure to calculate the new probability distribution vector during each interval, and will be at a minimum when the minimum value of $\lambda(t)$ is used.

Let $A_{k,max}$ be the matrix A(t) when $\lambda(t)$ is a maximum on the interval

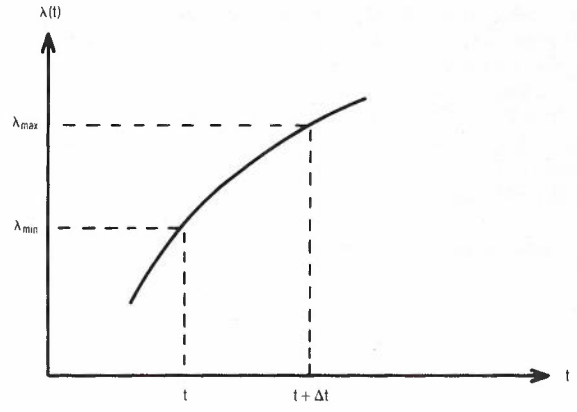


Fig. 2 - The Values of λ we use to Obtain the Upper and Lower Bounds of the Blocking Probability.

$((k-1)\Delta t, k\Delta t)$ and $A_{k,min}$ be the matrix A(t) when $\lambda(t)$ is a minimum on the same interval. Let our initial state probability vector be $\underset{\sim}{p}_{o,max} = \underset{\sim}{p}_{o,min} = \underset{\sim}{p}(0,N)$. Then writing, for $k \geq 1$,

$$\underset{\sim}{p}_{k,max} = \exp[A_{k,max}\Delta t]\underset{\sim}{p}_{k-1,max}$$

$$\underset{\sim}{p}_{k,min} = \exp[A_{k,min}\Delta t]\underset{\sim}{p}_{k-1,min} \qquad (7)$$

we have $b_{1,max} = (p_{1,max})_{N+1}$, $b_{1,min} = (p_{1,min})_{N+1}$ as two approximations to the exact blocking probability, b_1, at time Δt. The exact probability at time Δt would lie between the two approximations, i.e.

$$b_{1,min} \leq b_1 \leq b_{1,max}$$

If $\lambda(t)$ is constant then $b_{1,min} = b_{1,max}$ and the approximations so calculated are correct to machine accuracy. In general, if $t = k\Delta t$ then

$$b_{k,max} = (\underset{\sim}{p}_{k,max})_{N+1}$$

$$b_{k,min} = (\underset{\sim}{p}_{k,min})_{N+1} \qquad (8)$$

and

$$b_{k,min} \leq b_k \leq b_{k,max}$$

By taking smaller time intervals we can force these upper and lower bounds for the blocking probability to be closer to one another. For some cases the upper and lower bounds for the probability are very far apart indicating the possibility of an inaccurate result. This can happen if the traffic, $\lambda(t)$, the number of channels, N, or the time interval, Δt, is too large.

3. NUMERICAL EXAMPLES

A computer program incorporating the above

ideas was developed and several arrival patterns were tested with a system consisting of 10 channels. The time interval used in all the examples was 0.1, i.e. $\Delta t = 0.1$. The numerical results at specific time t were plotted with graphs of exact blocking probability for the same arrival patterns and number of channels found in Ref. 1.

The graphs and the corresponding arrival rate functions are:

Fig. 3(a): $\lambda(t) = 7$

Fig. 3(b): $\lambda(t) = 5 + 2\sin(0.05\pi t)$

Fig. 3(c): $\lambda(t) = 5 + 2\sin(0.1\pi t)$

Fig. 3(d): $\lambda(t) = 5 + 2\sin(0.5\pi t)$

The above arrival rates represent functions with rapid and slow oscillations. The continuous lines represent the probability obtained using the above procedure. The crosses are actually the mean of the upper and lower bounds of the probability. However, the upper and lower bound points for these examples are so close together that on the graphs they cannot be distinguished from one another and the exact results. Fig. 4 is a table of these upper and lower bounds for $\lambda(t) = 10(1-e^{-(t+1)})$ with different Δt's.

As can be seen from the above graphs, the numerical method gives very good results.

4. ERROR AND TIME ANALYSES

In this Section we define the norms of a matrix A, and a vector $\underset{\sim}{x}$, as

$$\|A\| = \max_{1\le i\le N+1} \sum_{j=1}^{N+1} |a_{ij}|$$

$$\|\underset{\sim}{x}\| = \max_{1\le i\le N+1} |x_i| \qquad (9)$$

Then it can be verified that (Ref. 4)

$$\|AB\| \le \|A\|\,\|B\|$$

$$\|A\underset{\sim}{y}\| \le \|A\|\,\|\underset{\sim}{y}\|$$

$$\left\| \sum_{n=1}^{\infty} \underset{\sim}{x}_n \right\| \le \sum_{n=1}^{\infty} \|\underset{\sim}{x}_n\| \qquad (10)$$

Let

$$\underset{\sim}{d}_k = \underset{\sim}{p}_{k,max} - \underset{\sim}{p}_{k,min}$$

Then for any definition of vector norm, we have

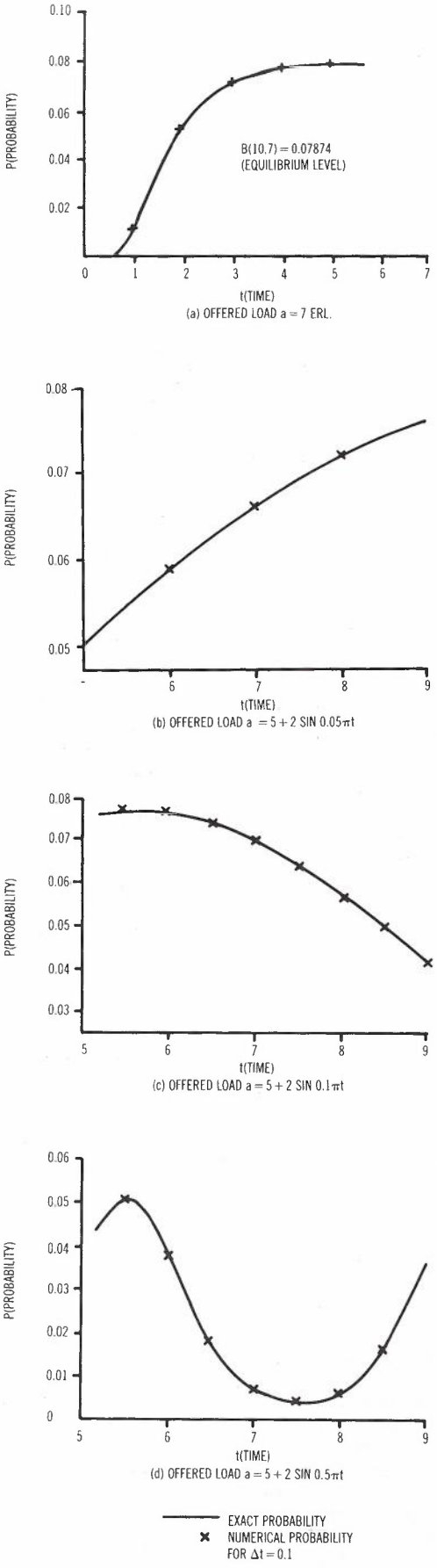


Fig. 3 - Blocking Probabilities for Various Arrival Patterns, for a Group of 10 Trunks.

$$\|\underline{d}_{k+1}\| = \|\underline{p}_{k+1,max} - \underline{p}_{k+1,min}\|$$
$$= \|\exp(A_{max}\Delta t)\underline{p}_{k,max} - \exp(A_{min}\Delta t)\underline{p}_{k,min}\|$$
$$= \|\exp(A_{min}\Delta t + E\Delta t)\underline{p}_{k,max} - \exp(A_{min}\Delta t)\underline{p}_{k,min}\| \quad (11)$$

where

$A_{max} = A_{k+1,max}$

$A_{min} = A_{k+1,min}$,

$E = A_{max} - A_{min}$,

so that matrix E has the following form:

$$E = \begin{bmatrix} -\varepsilon & 0 & 0 & & & \\ \varepsilon & -\varepsilon & 0 & & 0 & \\ 0 & \varepsilon & -\varepsilon & \ddots & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & -\varepsilon & 0 \\ & 0 & & & \varepsilon & 0 \end{bmatrix}$$

where $\varepsilon = \lambda_{max} - \lambda_{min}$.

If we assume $\|A_{max}\Delta t\| \leq 1$ for all t, then

$$\|\underline{d}_{k-1}\| = \|\sum_{i=0}^{\infty} \frac{(A_{min}\Delta t + E\Delta t)^i}{i!}\underline{p}_{k,max} - \sum_{i=0}^{\infty} \frac{(A_{min}\Delta t)^i}{i!}\underline{p}_{k,min}\|$$
$$= \|(I + (A_{min}\Delta t + E\Delta t) + \frac{(A_{min}\Delta t + E\Delta t)^2}{2!} + \frac{(A_{min}\Delta t + E\Delta t)^3}{3!} + \ldots)\underline{p}_{k,max}$$
$$- (I + A_{min}\Delta t + \frac{A_{min}^2\Delta t^2}{2!} + \frac{A_{min}^3\Delta t^3}{3!} + \ldots)\underline{p}_{k,min}\|$$
$$\leq \|\exp(A_{min}\Delta t) \cdot (\underline{p}_{k,max} - \underline{p}_{k,min})\|$$
$$+ \left\{\|F\Delta t\| + \|\frac{(A_{min}\Delta t + E\Delta t)^2 - A_{min}^2\Delta t^2}{2!}\|\right.$$
$$\left. + \|\frac{(A_{min}\Delta t + E\Delta t)^3 - A_{min}^3\Delta t^3}{3!}\| + \ldots\right\}\|\underline{p}_{k,max}\| \quad (12)$$

Now

$$\|(A_{min}\Delta t + E\Delta t)^{k+1} - A_{min}^{k+1}\Delta t^{k+1}\|/(k+1)!$$
$$= \|\frac{(A_{min}\Delta t + E\Delta t)}{k+1} \cdot \frac{(A_{min}\Delta t + E\Delta t)^k}{k!} - \frac{A_{min}\Delta t}{k+1} \cdot \frac{A_{min}^k\Delta t^k}{k!}\|$$
$$= \|\frac{A_{min}\Delta t}{k+1}\left[\frac{(A_{min}\Delta t + E\Delta t)^k}{k!} - \frac{A_{min}^k\Delta t^k}{k!}\right] + \frac{E\Delta t}{k+1}\frac{(A_{min}\Delta t + E\Delta t)^k}{k!}\|$$
$$\leq \|\frac{A_{min}\Delta t}{k+1}\|\cdot\|\frac{(A_{min}\Delta t + E\Delta t)^k}{k!} - \frac{A_{min}^k\Delta t^k}{k!}\| + \|\frac{E\Delta t}{k+1}\|\cdot\|\frac{(A_{min}\Delta t + E\Delta t)^k}{k!}\|$$
$$\leq \frac{1}{k+1}\|\frac{(A_{min}\Delta t + E\Delta t)^k}{k!} - \frac{A_{min}^k\Delta t^k}{k!}\| + \|\frac{E\Delta t}{k+1}\|\frac{1}{k!} \quad (13)$$

	$\Delta t = 0.01$		$\Delta t = 0.001$	
t	Upper Bound	Lower Bound	Upper Bound	Lower Bound
0.1	1.0520E-6	6.7276E-7	5.9719E-6	7.2291E-6
0.2	6.0522E-6	6.5468E-6	1.0253E-5	1.1283E-5
0.3	9.6405E-5	9.3232E-5	8.8732E-5	9.0763E-5
0.4	6.3152E-4	6.1272E-4	6.1175E-4	6.1149E-4
0.5	2.3589E-3	2.2991E-3	2.3236E-3	2.3192E-3
0.6	6.2180E-3	6.0826E-3	6.1519E-3	6.1425E-3

Fig. 4 - Upper and Lower Bounds for the Blocking Probability Obtained Using this Method.

If we let

$$x_k = \frac{\|(A_{min}\Delta t + E\Delta t)^k - A_{min}^k\Delta t^k\|}{k!} \quad (14)$$

then (13) is of the form, for $k \geq 1$,

$$x_{k+1} \leq \frac{x_k}{k+1} + \frac{x_1}{(k+1)!} \quad (15)$$

It can be shown by induction that, for $k \geq 1$,

$$x_{k+1} \leq \frac{1}{(k+1)!}[2x_2 + (k-1)x_1] \quad (16)$$

(12) then becomes

$$\|\underline{d}_{k+1}\| \leq \|\exp(A_{min}Wt)\cdot(\underline{p}_{k,max} - \underline{p}_{k,min})\| + \left\{x_1 + \sum_{k=1}^{\infty}\frac{1}{(k+1)!}[2x_2 + (k-1)x_1]\right\}\|\underline{p}_{k,max}\| \quad (17)$$

The coefficient of $\|\underline{p}_{kmax}\|$ can be simplified, i.e.,

$$x_1 + \sum_{k=1}^{\infty}\frac{1}{(k+1)!}2x_2 + \sum_{k=1}^{\infty}\frac{1}{(k+1)!}(k-1)x_1$$
$$= x_1 + 2x_2(e-2) + x_1(3-e)$$
$$= 2x_2(e-2) + x_1(4-e) \quad (18)$$

Therefore

$$\|\underset{\sim}{d}_{k+1}\| \leq \|\exp(A_{min}\Delta t).(\underset{\sim}{p}_{k,max} - \underset{\sim}{p}_{k,min})\|$$

$$+ \{(e-2)\|(A_{min}\Delta t+E\Delta t)^2 - A_{min}^2\Delta t^2\| + (4-e)\|E\Delta t\|\}\|\underset{\sim}{p}_{k,max}\|$$

$$\leq \exp(\|A_{min}\|\Delta t).\|\underset{\sim}{d}_k\| + \{(e-2).(2\|A_{min}\|\,\|E\|\Delta t^2 + \|E\|^2\Delta t^2)$$

$$+ (4-e).\|E\|\Delta t\} \qquad (19)$$

since $\|\underset{\sim}{p}_{k,max}\| \leq 1$.

Using (9) we have $\|A_{min}\| = 2\lambda+2N-1$, and $\|E\| = 2\varepsilon$. Thus (19) becomes

$$\|\underset{\sim}{d}_{k+1}\| \leq e^{(2\lambda+2N-1)\Delta t}\,\|\underset{\sim}{d}_k\|$$

$$+ (e-2).(2(2\lambda+2N-1).2\varepsilon\Delta t^2 + 4\varepsilon^2\Delta t^2)$$

$$+ (4-e).2\varepsilon\Delta t \qquad (20)$$

As an example, let N=10, λ=7, Δt=0.001 and ε=0.01. (20) now becomes

$$\|\underset{\sim}{d}_{k+1}\| \leq 1.0336\|\underset{\sim}{d}_k\| + (2.6583 \times 10^{-5})$$

The accuracy of the results obtained is dependent on four things. As the time step, Δt, decreases the errors also decrease and in most cases the difference between the maximum arrival rate and minimum arrival rate over this time interval, ε, will also decrease. It is by shortening the time interval used that more accurate results can be obtained, although this leads to a longer running time.

As the value of arrival rate at time t, $\lambda(t)$, or the number of channels in the system, N, increases the error will also increase. If $(\lambda+N)\Delta t$ becomes too large the error may increase very rapidly until the results obtained are useless. The difference between the lower and upper bounds obtained is a true indication of the errors involved. These actual errors are nearly always much less than the errors obtained using the above analysis which nevertheless indicates in what way error bounds will change as the various parameters involved either increase or decrease.

The number of operations required to calculate the congestions depends on two factors. The first of these is the time step, Δt, or the number of times, $1/\Delta t$, the process described above must be repeated. If Δt is halved then the CPU time required to find the state probabilities at a given time will double.

The operations required to find $\exp(A\Delta t)\underset{\sim}{p}$ are of the order N^2, that is as the number of circuits, N, involved increases the CPU time needed will also increase, as a quadratic function of N. If N is doubled the CPU time will increase by a factor of 4, if N is trebled the CPU time will increase by a factor of 9, and so on.

N	Δt	CPU Time (in sec.)	
10	0.1	9.2	constant N, varying Δt
10	0.01	85.0	
10	0.001	820.0	
5	0.1	2.3	varying N, constant Δt
10	0.1	9.0	
20	0.1	42.0	
40	0.1	272.0	

Fig. 5 - The CPU time used when Running this Method.

The table in Fig. 5 shows the CPU time needed to calculate, to a given accuracy, some results using different values of Δt and N. The test arrival function used was $\lambda(t)=10(1-e^{-(t+1)})$.

5. CONCLUSION

This numerical method is efficient as long as λ, N and Δt are not too large. All three quantities are interrelated in the error analysis as shown by (20). The choice of Δt would depend on the arrival rate $\lambda(t)$ (the steeper its slope, the smaller Δt should be) and the restrictions on λ and N will vary accordingly. Unfortunately in real-life, traffic may be greater than a few hundred erlangs in parts of the network and thus, in this case, the method becomes very inefficient. However, in systems with an offered load below 20 erlangs this method works effectively.

Application of this method to arrival rates consisting of steps (as considered in Ref. 2) is straightforward as it considers traffic with constant intensity over a fixed subinterval, Δt, in its calculation of blocking probabilities.

This method can also be adapted to the case of the M(t)/M(t)/N loss system which has the following set of d.d.e.'s:

$$\dot{P}_0(t,N) = -\lambda(t)P_0(t,N) + \mu(t)p_1(t,N)$$

$$\dot{P}_j(t,N) = \lambda(t)p_{j-1}(t,N) - (\lambda(t)+i\mu(t))p_j(t,N)$$
$$+ (i+1)\mu(t)p_{j+1}(t,N), \text{ for } 0<i<N$$

$$\dot{p}_N(t,N) = \lambda(t)p_{N-1}(t,N) - N\mu(t)p_N(t,N)$$
$$(21)$$

where $\lambda(t)$ and $\mu(t)$ are the arrival and service rates respectively. On rescaling the time axis by introducing a new variable

$$\tau = \int_0^t \mu(s)ds$$

in place of t, (21) is transformed into a set of equations similar to (1). $P_i(\tau,N)$ now replaces $P_i(t,N)$ and we can then apply this method to the new set of d.d.e.'s.

6. ACKNOWLEDGEMENTS

We would like to thank the reviewer and Dr. L.T.M. Berry for their helpful suggestions in the course of the preparation of this paper. This research was supported by the Commonwealth Postgraduate Research Awards (AJC) and Universiti Sains Malaysia (MNY). Fig. 3 is reprinted with permission from the Bell System Technical Journal (Copyright 1975, AT&T).

7. REFERENCES

1. Jagerman, D.L., "Nonstationary Blocking in Telephone Traffic", BSTJ, Vol. 54, No. 3., March 1975, pp. 625-661.

2. Yunus, M.N., "A Method for Determining Blocking Probability in the M(t)/M(t)/1 Loss System", ATR, Vol.18, No.2, Nov. 1984. pp. 13-17.

3. Martin, R.S. and Wilkinson, J.H., "The Implicit QL Algorithm", Numer. Math., Vol. 12, 1968, pp.377-383.

4. Bellman, R., Introduction to Matrix Analysis, 2nd. Ed., New York,McGraw-Hill, 1970.

BIOGRAPHIES

ANDREW COYLE received his Honours Bachelor of Science degree from the University of Adelaide in 1983. He is currently working towards a PhD degree at the University of Adelaide studying teletraffic theory under the supervision of Les Berry.

NAIM YUNUS is a PhD student in Applied Mathematics at the University of Adelaide where he completed his Honours degree in 1982. His research, under the supervision of L.T.M. Berry, is on the congestion analysis of telecommunications networks under nonstationary arrival conditions. A Malaysian, he is currently holding a scholarship awarded by Universiti Sains Malaysia in Penang.

SPECIAL SECTION — ISDN TRANSMISSION

Introduction

A. JENNINGS

The present era of rapid progress towards an implementation of an Integrated Services Digital Network (ISDN) is unique in the history of communications. Here we have the development of a network concept largely in the international forum of the CCITT, prior to any large scale equipment or network implementation. This is especially true in the sphere of transmission. A comprehensive set of Recommendations by the CCITT has been developed to support the ISDN network concept but gives little consideration to the problem at the basis of an ISDN implementation - how to get information from the customer to the local exchange. In terms of the ISDN objectives, this requires full duplex tranmission at 144 kbit/s over a single pair of wires, and this is by no means a trivial problem. In fact, in transmission terms it is probably one of the most challenging problems of the 1980's.

To meet the challenge of providing 144 kbit/s over existing cable networks, a number of groups throughout the world have been developing widely different techniques. In the era leading to ISDN implementations, it is likely that there will be a period of completion between these techniques for the prize of market share in the growing national ISDN's. So in this era it becomes important to be able to assess objectively the merits of the competing transmission schemes for 144 kbit/s to the customer. But surprisingly, there are few results available in this area of assessment of transmission systems for ISDN transmission over existing metallic pair networks. It is hoped that this Special Section will redress the balance in this area and spur further work on the assessment of ISDN transmission.

The first paper is by N. Demytko, B.M. Smith, G.J. Semple and P.G. Potter and serves as an introduction to the vital issues of ISDN transmission. It introduces the various approaches to full duplex transmission over a single metallic pair (i.e. burst mode and echo cancellation) and gives an exposition of the different approaches to echo cancellation. The various interference processes which limit the performance of ISDN transmission are then discussed: crosstalk between identical systems sharing the same cable, compatibility between different systems and most importantly impulsive noise interference from analogue services sharing the same cable with ISDN services. The second paper by B.R. Clarke and G.J. Semple discusses the issue of compatibility in more depth. A modification of the NEXT noise figure technique is introduced which allows a rapid resolution of compatibility issues. This paper provides a useful summary of the critical compatibility issues in ISDN transmission. The third paper by G.J. Semple considers the effect of bridged taps on the performance of ISDN transmission schemes. Although the use of decision feedback equalisers overcomes these problems in some cases, this paper provides a complete evaluation of the effects in those situations where a decision feedback equaliser is not employed. The final paper by A.J. Jennings provides an analysis of the effects of the severe line distortion of the copper pair on the operation of a memory compensation principle decision-feedback equaliser.

A.J. Jennings

Transmission Considerations for ISDN Basic Access Systems

N. DEMYTKO, B.M. SMITH, G.J. SEMPLE and P.G. POTTER

Telecom Australia Research Laboratories

This paper gives a broad introduction to the developments in local digital reticulation in telecommunication networks. In particular, two competing digital transmission techniques for the provision of basic access to an "integrated services digital network" (ISDN) are assessed: the adaptive echo cancellation technique and the time compression multiplex technique. The assessment is based on statistical analyses of the interference processes which affect these systems in the local distribution plant. Crosstalk and impulsive noise are considered as well as compatibility between various digital systems sharing the same cable. Considerable advantages in noise immunity (and hence achievable system range) are shown to ensue from the use of an echo cancellation system operating in conjunction with a narrowband line code and receiver filter. In addition, two forms of echo canceller structure which are adapted according to either the stochastic iteration or sign algorithms are examined and compared: those employing conventional transversal filters and those employing the relatively new "memory compensation" filters. The comparison is in terms of adaptation speed, residual mean square cancellation error and the ability to cope with non-linear echo paths. The paper also outlines a range of possibilities that face telecommunications administrations in adopting a transmission interface.

KEYWORDS: ISDN, Crosstalk, Impulsive Noise, Echo Cancellation, Compatibility, Interfaces

1. INTRODUCTION

For more than a decade telecommunication networks around the world have been undergoing a digital evolution. Digitalisation of these networks began with the introduction of cost effective digital transmission systems designed to convey digitally encoded voice signals over the twisted-pair cables between telephone exchanges. Initially it was necessary to convert the digital signals back into analogue form for switching through the exchange. However, this is now becoming unnecessary with the introduction of digital switching, which, together with the introduction of digital trunk transmission systems (initially using microwave radio and coaxial cables and later optical fibres), leads to the development of what is known as an "integrated digital network" (IDN). Although an IDN provides end-to-end digital transmission paths between the customers' local exchanges, the customer access network between the customer and his local exchange still remains analogue. Nevertheless, the potential advantages of having customer-to-customer digital transmission have been recognised. In particular, it allows telecommunications administrations to introduce a network, termed an "integrated services digital network" or ISDN, which can support a wide range of new customer services. This development is being facilitated by the International Telegraph and Telephone Consultative Committee (CCITT) which is actively developing ISDN standards.

About 30-40% of the current investment in the telecommunications network goes into the metallic multipair cable-based local distribution or customer access area. Whilst in the longer term an increasing application of optical fibres may be expected in this area (especially if there is an increased demand for wideband services by business and residential customers), the existing local cable network cannot be replaced in the short to medium term on a widescale basis. Consequently, it is clear that, in the foreseeable future, techniques must be found to utilise the existing metallic twisted-pair local distribution cable network if telecommunications administrations are to provide ISDN services widely and economically. The CCITT has adopted an ISDN basic access capacity of two 64 kbit/s "B" channels for voice and/or data services plus a 16 kbit/s "D" channel for signalling and low speed data. In addition to this 2B+D access, extra capacity must also be allocated for framing and maintenance which leads to a total effective digital line rate of about 160 kbit/s.

The distribution network already carries digital signals. Voice frequency data modems have been in use for about 20 years, whilst baseband data services use line transmission rates up to 72 kbit/s to customers. In

Paper received 20 July 1985.
Final revision 5 August 1985.

addition, some customers have 2 Mbit/s links back to their exchange for digital PABXs. However, a key aspect of the distribution network is that the ordinary customer has only one cable pair (2 wires) whilst the higher rate systems mentioned above (e.g. 72 kbit/s and 2 Mbit/s) require two cable pairs (4 wires). Clearly if telecommunications administrations are to have a widespread and economic introduction of the ISDN, which implies utilising the existing cable network (and not the installation of additional cable pairs), then some technique to achieve two-wire, bi-directional (full-duplex) transmission rates of about 160 kbit/s is required. Three techniques have been proposed to separate the two directions of transmission:

(a) Time Compression Multiplex (TCM) (also referred to as Burst or Ping-Pong).

This technique achieves the separation by increasing the original transmission rate and then time-sharing the cable pair between the two directions of transmission. To achieve this the transmission rate must be increased by a factor of about 2.5 (the extra factor of 0.5 provides a guard band to prevent overlap of signals due to transmission delays). As a consequence, the resulting increase in the bandwidth of the receiver filters leads to a significant increase in the susceptibility of the system to noise and interference.

(b) Frequency Separation

This technique achieves the separation by employing a separate frequency band for each direction of transmission. It is an extremely simple technique but is not preferred because of the high cost of the filters required.

(c) Adaptive Echo Cancellation

Whilst this technique (described in section 2) is relatively complex, it offers the best performance in terms of immunity to noise and interference. Basically, the technique can be divided into two main forms: those involving conventional transversal filters and those involving the relatively new "memory compensation" filters. These are discussed and compared, in sections 2, 3, 4, 5 and 6, in terms of adaptation speeds (for two different adaptation algorithms), residual mean square cancellation error and the ability to cope with non-linear echo paths.

As mentioned above, basic access ISDN transmission systems, at least initially, may have to share a given cable (in the local distribution plant) with a number of other already existing systems. Consequently, due to random electromagnetic coupling between the twisted pairs in the multipair cable, each system will experience interference from all other systems in that cable. In particular, interference will be induced into a given ISDN system from the same type and possibly other types of ISDN systems, and from digital systems with line transmission rates of, for example, 72 kbit/s and 2 Mbit/s. Impulsive noise from rotary dialling on analogue telephone circuits in the same cable and other causes, is another significant source of interference. Conversely, interference from the ISDN systems into other systems sharing the same cable must also be considered. These aspects are discussed in more detail in sections 7, 8, 9 and 10. It should be noted that a different mix of systems may be found in different cables and that new systems, such as digital outdoor extensions from digital PABXs and video systems, may also be introduced into the local distribution plant in the not too distant future.

Section 11 briefly discusses a key issue in relation to large scale integration (LSI) of ISDN basic access equipment. Section 12 introduces the subject of ISDN network interfaces and outlines a range of possible options that telecommunications administrations face in adopting a transmission interface.

2. ADAPTIVE ECHO CANCELLATION TECHNIQUES

The basic echo canceller configuration, as used in a full-duplex transmission system, is shown in Fig.1. It is assumed that all signals are sampled at T second intervals. The signal, $a(k) \triangleq a(kT)$, is transmitted to receiver B, while simultaneously receiver A attempts to receive the signal, $u(k) \triangleq u(kT)$, which is an attenuated and dispersed version of the signal, $b(k) \triangleq b(kT)$, sent from transmitter B. Due to reflections along the cable pair and the non-ideal hybrid, a filtered version of the transmitted signal, a(k), leaks through the hybrid to form an interfering "echo" signal, $y(k) \triangleq y(kT)$, which is added to the received signal, u(k). The purpose of the adaptive digital filter is to produce an estimate, $\hat{y}(k)$, of the echo, y(k), using the transmitted signal, a(k), as input. The signal, $r(k) \triangleq r(kT)$, that remains after subtraction of $\hat{y}(k)$ from $u(k) + y(k)$ is used to form an estimate, $\hat{b}(k)$, of the symbol, b(k), originating from transmitter B, and also to serve as a control signal for the adaptation of the adaptive filter coefficients. It should be noted that the filter must be adaptive because in most applications the echo path characteristics are unknown and must be derived automatically.

Because of the difficulty in adequately matching the hybrid to the variety of possible cable configurations over the required frequency range, the leakage through the hybrid can be quite high and result in an unwanted echo power at the receiver which is attenuated by as little as, say, 5 dB. For successful reception of the signal transmitted from the far end, the unwanted echo power should be about 20 dB less than the received signal power. To satisfy this condition, after taking into account the total transmitted power loss through the hybrids and over the line (which may be as high as 45 dB), the echo power must be reduced by up to

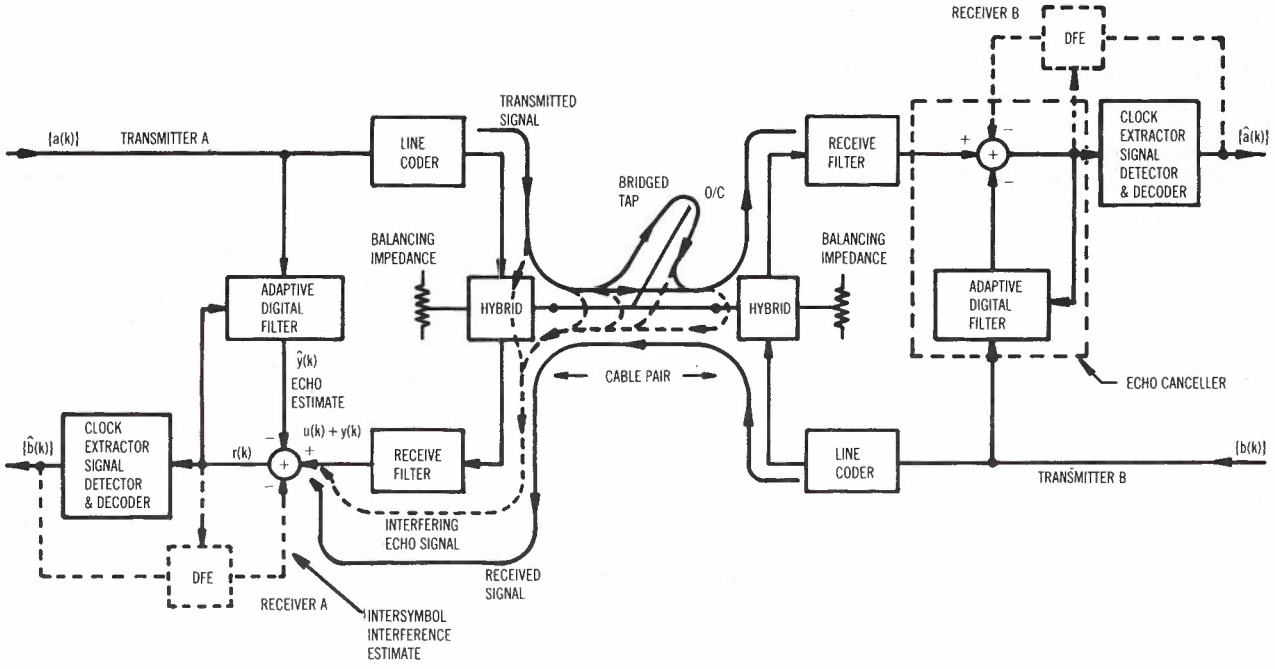


Fig. 1 - Adaptive Echo Cancellation System Providing Bi-Directional Transmission on a single cable pair.

45+20-5 = 60 dB. Another way of stating this is that the successful introduction of basic access ISDN on the metallic cable network requires the cancellation of complicated and unknown (to the echo canceller) waveforms by a factor of up to about 1000 in voltage. This degree of cancellation places severe linearity constraints on echo canceller systems employing conventional transversal filters.

In addition to the dispersion caused by the two hybrids, the receive filter and possibly the line coder in the forward transmission path (A to B or B to A), the transmitted signals are further dispersed by discontinuities along the cable pair and the presence of gauge changes and open circuit "bridged taps" or stubs which frequently occur in the local distribution plant. The dispersion caused by bridged taps and cable pair gauge changes in the forward transmission direction are often referred to as "forward echoes" (see Fig.1). Typically, dispersion caused by the forward transmission channel occurs over several symbol or sampling intervals and leads to overlapping of symbols or "intersymbol interference" at the receiver. This intersymbol interference may be removed by the inclusion of a decision feedback equaliser (DFE), (Refs. 1,2) as shown (dashed) in Fig.1.

If a DFE is included, a new signal, r'(k), which remains after the estimate of the intersymbol interference is subtracted from r(k), is used to control the adaptation of both the echo canceller and the DFE. As with the echo canceller, the DFE must be made adaptive because the forward transmission path characteristics are usually unknown.

Some systems also employ an adaptive reference to the received signal level, not shown in Fig.1, in the above control loops (Ref. 3). In this case an estimate of the received signal level, after intersymbol interference and echo components have been removed, is subtracted from r'(k), and is used to control all adaptive loops. However, clock extraction and signal detection and decoding are still performed on r'(k). In addition, adaptive timing recovery may also be required, and hence it is possible to have up to four complex and interactive adaptive loops that must converge before successful transmission can be achieved.

This paper discusses the performance of echo canceller structures only. To simplify the discussion, it is assumed that intersymbol interference is negligible and can be ignored, there is no adaptive reference control, receiver timing is known and, in sections 3 to 5, any noise which may be added to the transmission channel is negligible.

Until the last few years, transversal or "tapped delay line" (TDL) filters have generally been used to form the echo estimate, $\hat{y}(k)$, as shown in Fig.2(a), where it has been assumed that the echo path impulse response has a finite duration. A relatively new method, known as the memory compensation or "table look-up" technique, works on the principle that the echo, y(k), and the corresponding echo estimate, $\hat{y}(k)$, are a function of only the last N transmitted data symbols, a(k), and consequently only a finite set (2^N for binary signalling) of possible echo estimate values are required at the sampling instants. These can easily be stored in a digital memory and the appropriate echo estimate selected by means of an address formed by the last N transmitted data symbols

as shown in Fig.2(b). However, it should be noted that the set of echo estimates grows exponentially with N, and therefore the technique is of practical use only if the echo path impulse response duration can be made reasonably short (less than about ten bit intervals) by the selection of a suitable line code and an appropriate receive filter characteristic.

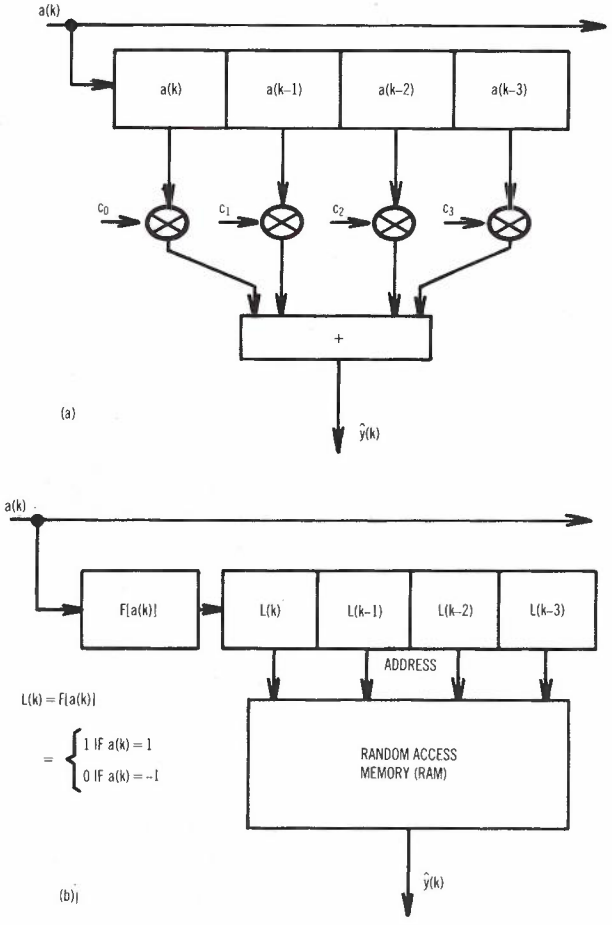


Fig. 2 - (a) Four Tap Transversal Filter. (b) Four Bit Memory Compensation Filter.

The tap weights, c_i, of the transversal filter are an estimate of the echo path impulse response. When convolved with the transmitted sequence, it yields the echo estimate, $\hat{y}(k)$. Hence the transmitter, which is part of the echo path impulse response, must be linear (for example, transmitted pulse asymmetry must be kept extremely low). Furthermore, non-linearities in hybrids (e.g. transformer saturation) and in A/D and/or D/A converters which may form part of the echo or control paths, must also be maintained at a very low level in order to obtain the high degree of echo suppression required for ISDN applications. However, in the memory look-up case, all possible echo estimates are stored directly in memory (i.e. linear operations are not required to form echo estimates). Consequently, the technique is suitable for all non-linearities, including those produced by non-linear line codes.

3. BASIC ECHO CANCELLER EQUATIONS

The following vectors are introduced for the subsequent discussion of the echo canceller performance:

the near-end data vector

$$A(k) = [a(k), a(k-1), \ldots, a(k-N+1)]^T$$

the far-end data vector

$$B(k) = [b(k), b(k-1), \ldots, b(k-I+1)]^T$$

the echo path vector

$$G = [g(0), g(1), \ldots, g(N-1)]^T$$

the transversal filter coefficient vector

$$C(k) = [c_0(k), c_1(k), \ldots, c_{N-1}(k)]^T$$

the forward transmission channel vector

$$H = [h(0), h(1), \ldots, h(I-1)]^T$$

where T denotes vector transposition.

The signal from the distant transmitter, u(k), the echo signal, y(k), the echo estimate, $\hat{y}(k)$, and the receiver input signal, r(k), are then given by:

$$u(k) = B^T(k).H \qquad (3.1)$$

$$y(k) = A^T(k).G \qquad (3.2)$$

$$\hat{y}(k) = A^T(k).C(k) \qquad (3.3)$$

$$r(k) = u(k) + y(k) - \hat{y}(k)$$

$$= u(k) + A^T(k).[G - C(k)] \qquad (3.4)$$

The following definitions are also used:

$E\{x\}$ = expected value of x

the mean square (MS) value, s_u^2, of the received signal, u(k), which is assumed to be stationary

$$s_u^2 = E\{u^2(k)\} \qquad (3.5)$$

the residual echo signal, e(k), given by

$$e(k) = y(k) - \hat{y}(k) \qquad (3.6)$$

the MS value, $s_e^2(k)$, of the residual echo at time instant k

$$s_e^2(k) = E\{[y(k) - \hat{y}(k)]^2\} \qquad (3.7)$$

The performance of an echo canceller is generally analysed by considering the convergence curve, R(k), as a function of k, where R(k) is defined by:

$$R(k) = s_e(k)/s_u \qquad (3.8)$$

The function, R(k), strongly depends upon the adaptation algorithm that defines the updating of the tap weights, $c_i(k)$, and the echo estimates, $\hat{y}(k)$, of the transversal and memory compensation filters, respectively. R(k) also depends on the type of filter employed. In this paper, the stochastic iteration algorithm (SIA) and the sign algorithm (SA) are examined for the two filter techniques described above.

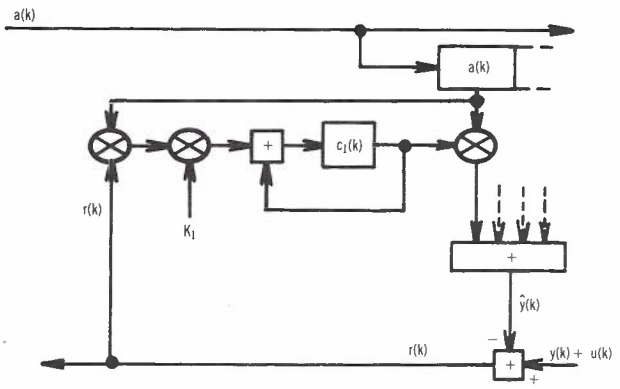


Fig. 3 - Adaptation of the Transversal Filter Using the Stochastic Iteration Algorithm.

4. THE STOCHASTIC ITERATION ALGORITHM (SIA)

4.1 Transversal Filter

The stochastic iteration algorithm for the case of the transversal filter (TF) has been analysed in great detail (Refs. 4,6) and is described by:

$$C(k+1) = C(k) + K_1 r(k)A(k) \qquad (4.1)$$

where K_1 is a gain constant. Fig.3 illustrates one coefficient, $c_1(k)$, being updated according to the SIA. In Ref. 4 the analysis of the TF employing the SIA is based on the following assumptions.

A: a(k) and u(k) are zero-mean, wide-sense stationary, mutually independent stochastic processes.

B: The probability density function of a(k) is binary, i.e. $Pr[a(k) = \pm 1] = 1/2$.

C: a(k) is a white process, i.e. $E\{a(k)a(k+n)\} = \delta(n)$

D: The misalignment vector, G - C(k), is statistically independent of both a(k) and u(k).

Using these assumptions and equations 3.1 to 3.8 and 4.1, it can be shown that the convergence of the residual echo power is governed by the following equations (Ref. 4).

$$R_1^2(k) = p_1^k[R_1^2(0) - R_1^2(\infty)] + R_1^2(\infty) \qquad (4.2)$$

where $$p_1 = 1 - 2K_1 + K_1^2 N \qquad (4.3)$$

and $$R_1^2(\infty) = K_1 N/(2 - K_1 N) \qquad (4.4)$$

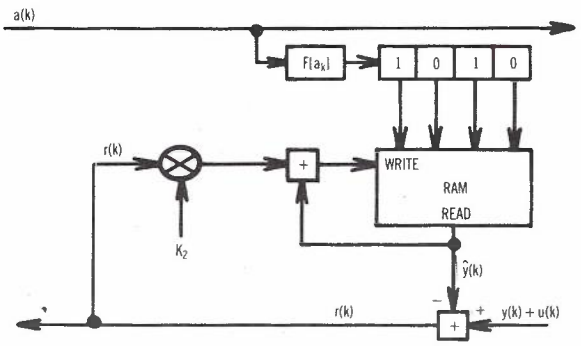


Fig. 4 - Adaptation of the Memory Compensation filter Using the Stochastic Iteration Algorithm.

4.2 Memory Compensation Filter

In the memory compensation filter (MCF), each cell or register is only updated when a given address (fixed sequence of N bits) is generated. This is illustrated in Fig.4 where the memory cell corresponding to the 4 bit address, 1010, is shown being updated according to the SIA. It should be noted that the echo value, y(k), corresponding to a given address (e.g. 1010) is a constant (say Y), i.e. the echo value, Y, will always occur every time the given address (e.g. 1010) is, generated. A detailed analysis of the MCF employing the SIA is given in Ref. 6. In this paper a simpler approach to the analysis is given, which, nevertheless, yields the same results.

If the transmitted sequence {a(k)} is scrambled, it can be assumed that all addresses are generated with equal probability. Thus for an N bit sequence, there are 2^N different address states and each address is generated once, on average, for every $M = 2^N$ bits that are transmitted. In addition, each address state corresponds to a fixed value of echo. The M possible echo states, corresponding to address states 1 to M, will be denoted by Y_1 to Y_M. Note also that $Pr\{y(k) = Y_i\} = 1/M$ for i = 1 to M.

The adaptation of a single cell (say cell 1) according to the SIA is described by:

$$\hat{Y}_1(m+1) = \hat{Y}_1(m) + K_2 r(m) \qquad (4.5)$$

where K_2 is a gain constant and m is the number of times a given cell is addressed in k bit intervals. Each memory cell operates independently of all other memory cells and therefore only one cell needs to be analysed.

Subtracting (4.5) from the corresponding fixed echo value, Y_1, and using (3.4) and (3.6), gives:

$$E_1(m+1) = (1 - K_2)E_1(m) - K_2 u(m) \qquad (4.6)$$

where $E_1(m) = Y_1 - \hat{Y}_1(m)$

Using assumptions A, B and C given in section 4.1, squaring both sides of (4.6) and taking the expectation over the ensemble of received signal samples, u(m), results in the following recurrence relation

$$\bar{E}_1^2(m+1) = (1 - K_2)^2 \bar{E}_1^2(m) + K_2^2 s_u^2 \qquad (4.7)$$

where $\bar{E}_1^2(m)$ = expectation of $E_1^2(m)$ over the ensemble, u(m).

It follows that:

$$\bar{E}_1^2(m) = p^m \bar{E}_1^2(0) + K_2^2 s_u^2 (1 - p^m)/(1 - p) \qquad (4.8)$$

where $p = (1 - K_2)^2 \qquad (4.9)$

Then, noting that the number of times that a given cell is addressed in k bit intervals is given by the binomial distribution and taking the expectation over the possible residual echo states in cell 1 yields:

$$\bar{E}_1^2(k) = p_2^k \bar{E}_1^2(0) + K_2^2 s_u^2 (1 - p_2^k)/(1-p) \qquad (4.10)$$

where $\bar{E}_1^2(k) = \sum_{m=0}^{k} \bar{E}_1^2(m) C_m^k (1/M)^m (1 - 1/M)^{k-m}$

$\bar{E}_1^2$ = expectation over the ensemble, $\bar{E}_1^2(m)$, at time instant k,

and $$p_2^k = \sum_{m=0}^{k} p^m C_m^k (1/M)^m (1 - 1/M)^{k-m}$$

$$= (1 - 2K_2/M + K_2^2/M)^k \qquad (4.11)$$

and $$C_m^k = \frac{k!}{m!\,(k-m)!}$$

Finally, taking the expected value over all the memory cells and dividing by s_u^2 gives:

$$R_2^2(k) = p_2^k [R_2^2(0) - R_2^2(\infty)] + R_2^2(\infty) \qquad (4.12)$$

where $R_2^2(k) = \sum_{i=1}^{M} \bar{E}_i^2(k)/M/s_u^2$

and $R_2^2(\infty) = K_2/(2-K_2) \qquad (4.13)$

5. THE SIGN ALGORITHM (SA)

If the receive filter characteristic and line code (see Fig.1) are suitably chosen, u(k) will have a probability density function that is close to binary (i.e. $u(k) \simeq \pm s_u$) over practical transmission distances (ignoring bridged taps). In this case, it can be shown that the convergence of an echo canceller, which is adapted according to the sign algorithm, will terminate as soon as the magnitude of the residual echo error, e(k), becomes permanently smaller than the minimum magnitude of u(k) (Ref. 5,7). This limitation is eliminated, however, if the SA is operated in conjunction with a "dithering" signal. Consequently, for the following analysis, it is assumed that a dithering signal, d(k), is added to the control signal, r(k), for adaptation purposes only. It should be noted that the signal, r(k), which is input to the receiver, remains unchanged.

The dithering signal is taken to be uncorrelated with both a(k) and u(k) and to have a uniform probability density in the interval $(-s_u, s_u)$. Furthermore, in order to simplify the analysis and comparison of the two filter techniques, it is assumed that $u(k) = \pm s_u$ and that the positive echo states, denoted by $e(k) = E_i(k)$ (i = 1 to M/2), are clustered around $+s_e$ and the negative states, $e(k) = E_{M/2+i}(k) = -E_i(k)$, (i = 1 to M/2), are clustered around $-s_e$. This is reasonable for the receive filter, line code and transmission distances assumed above, where the number of echo path impulse response coefficients, N, is small and the magnitude of the first coefficient, |g(0)|, is, in general, much larger than the magnitude of the remaining coefficients.

Under the above assumptions, it follows that

$$m_e(k) \simeq s_e(k) \simeq |E_i(k)| \quad (i = 1 \text{ to } M) \qquad (5.1)$$

where $m_e(k) \triangleq E\{|e(k)|\} = 2 \sum_{i=1}^{M/2} E_i(k)/M \qquad (5.2)$

and that the new control signal, r'(k), may be written as

$$r'(k) = r(k) + d(k) = v(k) + e(k) \qquad (5.3)$$

where $v(k) = u(k) + d(k) \qquad (5.4)$

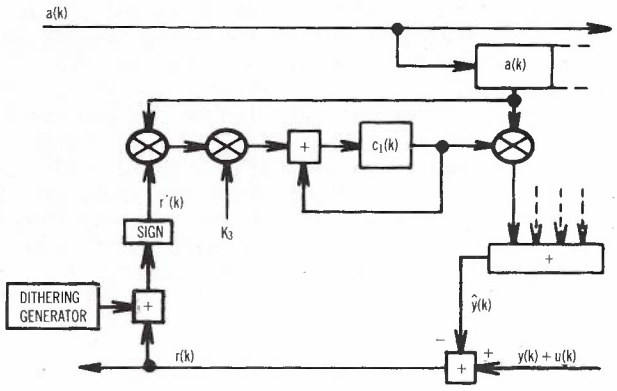


Fig. 5 - Adaptation of the Transversal Filter Using the Sign Algorithm.

has a uniform probability density in the interval $(-2s_u, 2s_u)$, see Fig.7.

5.1 Transversal Filter

The transversal filter which is updated according to the SA, see Fig. 5, is described by

$$C(k+1) = C(k) + K_3 \text{sign}[r'(k)]A(k) \qquad (5.5)$$

where K_3 is a gain constant. From the assumptions in section 4.1 and equations 3.2, 3.3, 3.6 and 3.7, it follows that:

$$s_e^2(k) = E\{[G - C(k)]^T[G - C(k)]\} \qquad (5.6)$$

Subtracting (5.5) from G, using (5.6), and dividing by s_u^2 then yields

$$R_3^2(k+1) = R_3^2(k) - 2K_3W(k)/s_u^2 + K_3^2\, N/s_u^2 \qquad (5.7)$$

where $W(k) = E\{e(k)\text{sign}[e(k) + v(k)]\}$ (5.8)

This case has been analysed in Ref. 5 under the assumption that the probability density of e(k) does not change with time and is gaussian. Proceeding along similar lines to that given in Ref. 5 and using the assumptions of section 5 (which do not assume that the probability density of e(k) does not change with time) it can be shown that

$$W(k) = -2 \sum_{i=1}^{M} E_i(k)F_v[-E_i(k)]/M \qquad (5.9)$$

where $F_v(-x)$ is the distribution function of v(k) as shown in Fig.7.

From (5.9) and Fig.7

$$W(k) = \begin{cases} m_e(k) \simeq s_e(k), & s_e(k) >> 2s_u \\ s_e^2(k)/(2s_u), & s_e(k) << 2s_u \end{cases} \qquad (5.10)$$

and, from (5.7), it follows that:

$$R_3^2(k+1) \simeq \begin{cases} [R_3(k)-z/N]^2+z^2(1-1/N)/N, & R_3(k)>2 \qquad (5.11) \\ (1-z/N)R_3^2(k)+z^2/N, & R_3(k)<2 \qquad (5.12) \end{cases}$$

where $z = K_3N/s_u$ (5.13)

Neglecting the second term on the RHS of (5.11), since z is small, taking the square root of the remaining term and applying the resulting recurrence relation, and the recurrence relation of (5.12), yields

$$R_3(k) \simeq R_3(0) - k(K_3/s_u),\ R_3(0,k) > 2 \qquad (5.14)$$

$$R_3^2(k) \simeq p_3^k[R_3^2(0) - R_3^2(\infty)] + R_3^2(\infty), \quad R_3(0,k) < 2 \qquad (5.15)$$

where $p_3 = (1 - K_3/s_u)$ (5.16)

and $R_3^2(\infty) = z = K_3N/s_u$ (5.17)

5.2 Memory Compensation Filter

The adaptation of a single cell (cell 1) according to the SA, see Fig.6, is described by (Ref. 7):

$$\hat{Y}_1(m+1) = \hat{Y}_1(m) + K_4\text{sign}[r'(m)] \qquad (5.18)$$

where K_4 is a gain constant and m is the number of times that cell 1 is addressed in k bit intervals, as before. Proceeding in the same manner as in section 4.2, using the assumptions of sections 4.2 and 5, and noting that $E_1^2(m) \simeq E_1(m)$ if k is small, results in the following recurrence relation:

$$E_1^2(m+1) \simeq E_1^2(m) - 2K_4E_1(m)W'(m) + K_4^2 \qquad (5.19)$$

where $W'(m) = E\{\text{sign}[E_1(m) + v(m)]\}$

$$= 1 - 2F_v[-E_1(m)] \qquad (5.20)$$

From (5.20) and Fig.7:

$$W'(m) \simeq \begin{cases} \text{sign}[E_1(m)] & \text{if } |E_1(m)| > 2s_u \\ E_1(m)/(2s_u) & \text{if } |E_1(m)| < 2s_u \end{cases} \qquad (5.21)$$

and from (5.19) and (5.21) it follows that

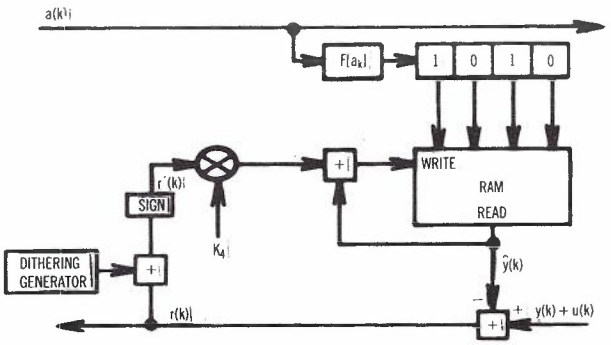


Fig. 6 - Adaptation of the Memory Compensation Filter Using the Sign Algorithm.

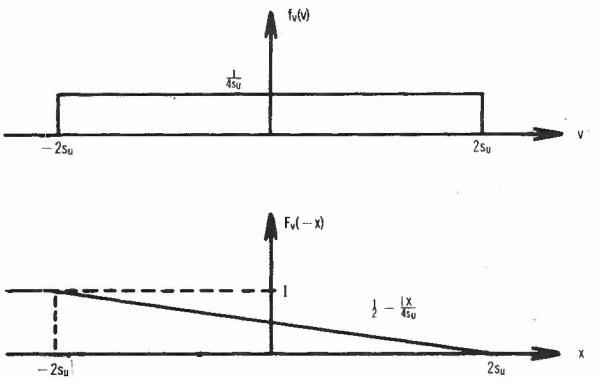


Fig. 7 - Probability Density and Distribution Functions of v(k).

$$E_1^2(m+1) = \begin{cases} [|E_1(m)| - K_4]^2 & , \quad s_e >> 2s_u \\ (1 - K_4/s_u)E_1^2(m) + K_4^2 & , \quad s_e << 2s_u \end{cases} \qquad (5.22)$$

Then using the procedure given in section 4.2, yields:

$$R_4(k) \simeq R_4(0) - k(K_4/M/s_u), \qquad R_4(0,k) > 2 \qquad (5.23)$$

$$R_4^2(k) \simeq p_4^k[R_4^2(0) - R_4^2(\infty)] + R_4^2(\infty), \quad R_4(0,k)<2 \qquad (5.24)$$

where $p_4 = (1 - K_4/M/s_u)$ (5.25)

and $R_4^2(\infty) = K_4/s_u$ (5.26)

6. COMPARISON OF FILTER TECHNIQUES

From (4.4) and (4.13) it can be seen that the ratio of residual echo power to received signal power for both the TF and MCF, using the SIA, are the same if $K_2 = K_1N$. Furthermore, from (4.3) and (4.11), $p_1 \simeq (p_2)^{M/N}$, if K_2 is small, and it follows from (4.2) and (4.12) that the convergence of the TF using the SIA is approximately M/N times faster than the MCF using the SIA (see Fig.8, where M/N = 4). Identical conclusions can also be reached for the two filters employing the SA if $K_4 = K_3N$ and K_4 is small (see equations 5.14 to 5.17 and 5.23 to 5.26). The convergence speed of the MCF can be increased, however, by dividing the memory into a number of smaller parts and cascading them, but only at the expense of losing the general ability of the MCF technique to cope with non-linear echo paths. This "multiple memory" technique is not discussed in this paper. However an account of the technique may be found in Ref. 8.

Although the MCF is slower to adapt, it should be noted that it is easier to implement than the TF, if N is small (see Figs. 3,4,5 and 6), and it has the ability to cope with non-linear echo paths (see section 2). Another point worth noting is that the residual echo power is proportional to s_u^2 for the filters employing the SIA, whereas it is only proportional to s_u for the filters employing the SA (see equations 4.4, 4.13, 5.17 and 5.26). Thus, if the power in the received signal, u(k), is assumed to (notionally) increase by a factor of 2 (or 3 dB) as a direct result of noise or interference from other systems, then the residual echo power will also increase by 3 dB for the filters employing the SIA. However, it will only increase by approximately 1.5 dB for the filters employing the SA. That is, filters employing the SA are less sensitive to interference, in relation to residual echo power, than those employing the SIA. In addition, in order to obtain adequate echo cancellation, dithering is required in conjunction with the SA (this is not the case, however, if an adaptive reference control is employed), whereas it is not required with the SIA.

Fig.8 shows the typical convergence behaviour for the 4 cases considered in this

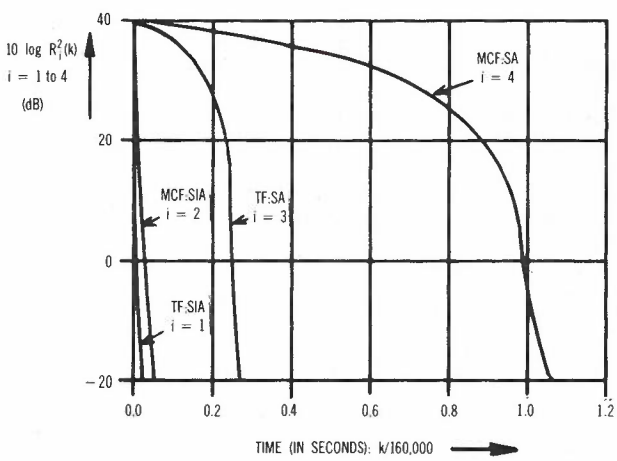


Fig. 8 - Typical Adaptation Curves for the Transversal Filter (TF) and the Memory Compensation Filter (MCF) Employing the Stochastic Iteration Algorithm (SIA) and the Sign Algorithm (SA).

paper. For the purpose of comparison it has been assumed that N = 4, the transmit data rate is 160 kbit/s, the line loss is 40 dB and the residual echo power is 20 dB below the receive signal power (corresponding to the conditions stated in section 2). From the figure, it can be seen that adaptation of either the TF or MCF using the SA is approximately 25 times slower than that employing the SIA.

Finally, it may be noted that a major consideration in the design and performance of echo cancellation systems is the choice of line code. Among other reasons, line coding is used to shape the frequency content of the line signal and is therefore a significant factor in determining the level of interference produced by these systems. Better performance in relation to interference (see sections 7 to 10) can be achieved by the lower bandwidth line codes, but at the expense of longer echo path impulse response and intersymbol interference durations, which lead to greater echo canceller and DFE complexity. For this reason, echo cancellers based on the MCF are only practicable when used with relatively high bandwidth line codes, such as conditioned diphase or Walsh type 1 (WAL1).

7. INTERFERENCE PROCESSES

The three most important interference processes which affect ISDN basic access systems are crosstalk from other identical ISDN basic access systems, impulsive noise due to crosstalk from line signalling events on adjacent circuits carrying analogue telephony, and compatibility with other types of system in the cable due to crosstalk into or from those systems. The three kinds of interference are illustrated in Fig.9. All result from irregularities in twisting of cable pairs which result in small random electromagnetic couplings between circuits in the multipair cable. The aggregate of all of the small couplings along the cable may be described by the near end crosstalk (NEXT) and far end crosstalk (FEXT) transfer functions between the disturbing and the disturbed pairs.

Because the crosstalk couplings are modelled statistically (Ref. 9), system designs are usually based on a criterion of the following form:

"For a digital system on a cable pair chosen at random, the probability that a system of maximum design length or loss fails to meet the error performance objective should be less than 0.01."

The performance objective in this paper is based on an interpretation of the requirements of CCITT Recommendation G.821 (Ref. 10), and states that the long term mean bit error ratio (LTMBER) should be less than 1×10^{-7}. With appropriate classification of the crosstalk paths (into those pair combinations with both pairs within the same unit, WU, for example) the attenuation of the crosstalk path is approximately normally distributed with mean μ_0 and variance

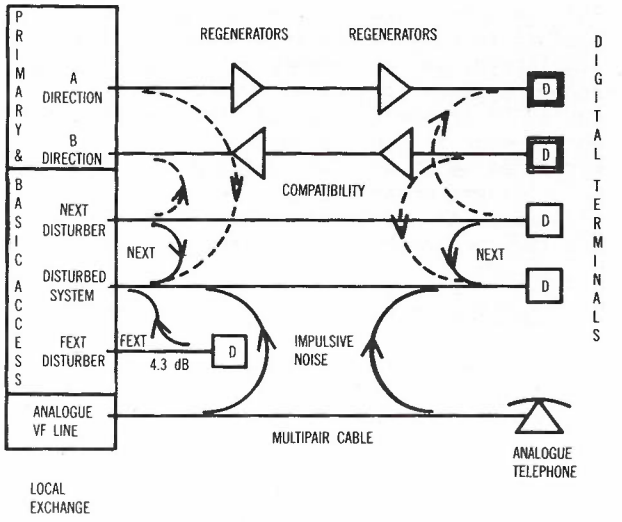


Fig. 9 - Crosstalk Between Basic Access ISDN Systems, Compatibility with Primary Access (2 Mbit/s) systems and Impulsive Noise from Line Signalling Events on Analogue Telephone Circuits.

σ_0^2 at frequency f_0 and truncated at 2.5 standard deviations from the mean. The analyses which follow are based on the most severe crosstalk class, WU, which for small size cables with 10 pair units allows a maximum of 9 disturbing circuits.

The system performance in the presence of the three interference processes depends on several critical system design choices. The most fundamental is the choice between the echo canceller (EC) and the time compression multiplex (TCM) techniques. As outlined in section 1, the TCM system has a higher line rate (Ref. 11) and, for the purpose of the comparisons in this paper, is assumed to be 2.5 times the line rate of the corresponding EC system (with a duty cycle of 0.4 in either direction).

The choice of line code also influences the achievable noise immunity of a system; as indicated in the previous section, this noise immunity is greater for narrow bandwidth codes. Three codes which have been proposed for use in local digital transmission systems are conditioned diphase (or WAL1), alternate mark inversion (AMI) and 4B3T which is an alphabetic block code (Ref. 12). The last two have full width transmitted pulses. Equalisation and filtering at the receiver should be designed for effective noise immunity on a range of cable configurations incorporating different cable gauges and types. It may include adaptive or fixed linear filtering as well as nonlinear decision feedback equalisation. For the general comparisons to be made here, an adaptive linear equaliser is assumed to produce a 100% raised cosine (frequency domain) pulse shape at the receiver decision point for each rectangular segment of the transmitted code.

The immunity of the system to noise from other types of system and to impulsive noise also depends directly on the transmit level.

This is best chosen as high as possible, subject to constraints from power feeding limitations and on interference from ISDN basic access systems into other systems in the cable. A transmit level of +10 dBm into 120 ohm is used for the following analyses. A margin of +6 dB is also included in all of the calculations which follow. This is an allowance for non-ideal receiver implementation and for the uncertainty of cable parameters.

8. CROSSTALK BETWEEN ISDN BASIC ACCESS SYSTEMS

The range of EC systems may be limited by both NEXT and FEXT, whereas the range of TCM systems in which the bursts in each direction are synchronised is only limited by FEXT. The crosstalk noise figure technique (Ref. 13) is used to determine length limits for the idealized systems under the conditions described above.

The crosstalk theory used here is based on the observations (Ref. 9) that the total noise from N crosstalk disturbers is a normal process with variance v_T^2 which is itself lognormal over the ensemble of crosstalk paths. Hence $[v_T^2] = 10 \log v_T^2$ (for notational convenience) is $N(\mu_T, \sigma_T^2)$ where for the NEXT case:

$$\mu_T = [N] + 0.115\sigma_o^2 - \mu_o + [T] + [I_N] - 0.5[K] \quad (8.1)$$

$$\sigma_T^2 = 4.343\,[K] \quad (8.2)$$

$$K = 1 + \{U.\ 10^{0.023\sigma_o^2} - 1\}/N \quad (8.3)$$

where U and T are truncation factors depending on σ_o and the assumed truncation of the normal NEXT at 2.5 standard deviations from the mean.

$$I_N = \int_0^\infty \frac{f^2\alpha(f_o)}{f_o^2\alpha(f)} \left|\frac{D(f)}{G(f)}\right|^2 W(f)\, df \quad (8.4)$$

reflects the noise amplification by the receiver and depends on the cable attenuation constant $\alpha(f)$ dB/km, the equalised pulse shape D(f) (100% raised cosine), the cable transfer function G(f) and the power spectral density W(f) of the code impulse train.

N_o is defined as the mean-square NEXT noise at the decision point which just satisfies the performance objective of LTMBER $< 10^{-7}$. The value of N_o used here is based on the minimum eye opening (maximum intersymbol interference). The design objective of 1% failure probability is satisfied by:

$$\mu_T + 2.33\,\sigma_T < [N_o] \quad (8.5)$$

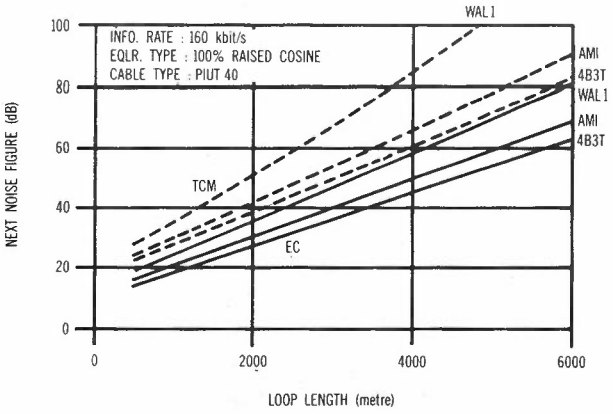


Fig. 10 - Near End Crosstalk (NEXT) Noise Figure at 100 kHz for Echo Canceller (EC) and Time Compression Multiplex (TCM) Systems for three Line Codes.

This design rule may be rearranged to isolate the receiver dependent terms on the RHS.

$$\mu_o - S(\sigma_o, N, c) > [R_N(f_o, L)] = [I_N] - [N_o] \quad (8.6)$$

where

$$S(\sigma_o, N, c) = .115\sigma_o^2 + [N] + [T] - .5[K] + 4.86\sqrt{[K]} \quad (8.7)$$

R_N is the NEXT noise figure referred to f_o. Its dependence on cable length L is illustrated in Fig.10 for EC and TCM systems with each of the three line codes on 0.4 mm paper-insulated, unit-twin (PIUT40), copper cable with NEXT attenuation mean = 76.2 dB and standard deviation = 8 dB at 100 kHz. The LHS of (8.6) depends on the mean NEXT attenuation μ_o and on the value of S which is plotted in Fig.11 as a function of N for a range of σ_o values.

For given crosstalk statistics and number of disturbers N, an upper limit on the permissible NEXT noise figure is obtained. This is reduced by the 6 dB margin to provide a horizontal line on Fig.10 and hence the maximum range for each line code is obtained. The NEXT limits for EC and unsynchronised TCM systems with N=9 disturbers are summarised in Table 1.

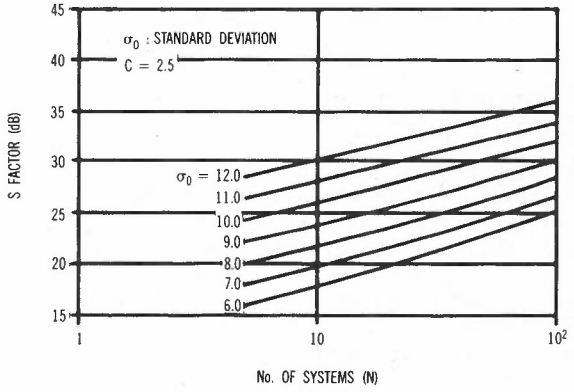


Fig. 11 - $S(\sigma_o, N, c)$ Factor with $c = 2.5$.

Table 1 - Near End Crosstalk (NEXT), Worst Case Far End Crosstalk (FEXT) and Impulsive Noise Range Limits (in km)

LINE CODE	NEXT LIMIT		FEXT LIMIT		IMPULSIVE NOISE	
	EC	TCM unsynch.	EC	TCM	EC	TCM
4B3T	4.3	2.8	4.5	3.6	3.4	2.4
AMI	3.9	2.5	4.3	3.5	3.2	2.2
WAL1	3.1	1.8	4.0	3.2	2.9	1.9

The analysis of FEXT is based on assumptions similar to those for NEXT. The quantity of interest is the far end signal to crosstalk ratio (FESXTR) which, for systems with equal transmit level at either end, is lowest when the disturbing system has a critical loss of 4.3 dB as illustrated in Fig.9. The resulting worst case FEXT noise figure $[R_F'(f,L)]$ is incorporated in a design equation of the same form as (8.6),

$$\mu_o - S(\sigma_o,N,c) > [R_F'(f_o,L)] = [I_F'] - [N_o] \qquad (8.8)$$

where μ_o and σ_o are here the mean and standard deviation of the equal level FEXT ratio for a 1 km length of cable. Table 1 also includes FEXT limits for both EC and TCM systems with 9 disturbers on 0.4 mm paper-insulated, unit-twin, copper cable with FEXT ratio mean = 73.7 dB and standard deviation = 8 dB at 100 kHz.

The limitation on system range imposed by crosstalk is well correlated with the effective line rate of the transmission technique and line code. Since even worst case FEXT can be shown to be less severe than NEXT on average, the NEXT limit is dominant for EC systems. For TCM systems the FEXT limit permits ranges similar to those for EC systems, and these FEXT limits are easily extended by a factor of nearly two by adjustment of the transmit level at the customer end (Ref. 14). Unfortunately, some network arrangements, such as those for baseband data services and the possible use of digital outdoor extensions to PABXs, would not be able to satisfy the condition of synchronised bursts with TCM; in these cases NEXT would limit the range even more severely than for EC systems.

9. COMPATIBILITY

The previous section on crosstalk assumes that the disturbing systems are identical to the disturbed system. However, with the evolution of the network, it is possible that different digital systems may be called upon to co-exist within the same cable unit. In this section, the crosstalk compatibility of TCM and EC systems with each other and with existing digital systems using the local loop is briefly considered. A full discussion is given in Ref. 15. The existing systems that are analysed are a 72 kbit/s baseband data system which uses a WAL1 (conditioned diphase) line code transmitting at +6 dBm and a 2048 kbit/s primary rate access system which employs a HDB3 line code (Ref. 12) and transmits at +16 dBm. The ISDN basic access systems to be considered are 160 kbit/s WAL1 and 160 kbit/s 4B3T echo canceller systems and a time compression multiplex system using the AMI line code and a burst line rate of 400 kbit/s. All of these basic access systems are considered to have a transmit level of +10dBm.

In section 8 it was mentioned that if the bursts of a TCM system were synchronised then the transmission span of this system would be limited by FEXT rather than NEXT. However if any of the other systems mentioned above is allowed to exist in the same cable as the TCM system then it has been found that NEXT becomes the dominant crosstalk. Consequently, only NEXT is considered here.

The compatibility studies use the NEXT Noise Figure outlined in section 8 but with some modifications. The most significant change is that the I_N term becomes

$$I_N = \int_0^\infty \frac{f^2\alpha(f_o)}{f_o^2\alpha(f)} \left| \frac{D_2(f)\ \ S_1(f)}{G(f)\ \ S_2(f)} \right|^2 W_1(f)\ df \qquad (9.1)$$

where the subscript 1 represents the disturbing system while the subscript 2 represents the disturbed system. S(f) is the transmit pulse shape transform. The additional term in (9.1) compared to (8.4) is the ratio of the transmit pulse shape spectra. This accommodates different pulse shapes and amplitudes in the disturbing and disturbed systems.

From (9.1) it is seen that the transmit spectrum must be specified when examining compatibility. It is assumed that all systems transmit approximately rectangular pulses with linear rise and fall times lasting 10% of the pulse duration. The results are summarised in Table 2. The vertical columns contain the transmission spans of a given disturbed system for the different disturbing systems. From this table the following points may be noted.

1. The echo canceller systems are effectively compatible with all other systems considered except for the TCM system. However, using EC systems employing different line codes in the same cable unit reduces the transmission span of the EC system employing the line code with the superior crosstalk performance.

2. The transmission spans of the existing digital systems are not reduced by either the TCM or the EC systems, hence re-engineering will be unnecessary.

3. The NEXT interference into a synchronised TCM system from all other systems is significant. Unless all other significant disturbing systems can be removed from the cable unit, then the transmission span of the synchronised TCM system may be reduced from 3.5 km (AMI FEXT limit in Table 1) to only approximately 2.7 km. This is a significant impairment to the use of synchronised TCM systems for basic customer access to an ISDN.

10. IMPULSIVE NOISE DUE TO ANALOGUE TELEPHONY

Loop disconnect signalling events on adjacent analogue telephone circuits in the cable may be coupled through the NEXT path to produce impulsive noise in digital systems. The signalling events may be modelled as step changes in level, $V_o\,u(t)$, (Ref. 16), so that:

$$V(f) = \frac{V_o}{j2\pi f} \qquad (10.1)$$

where V_o is typically 50 V. As well as the transmit level of the digital system and the other system parameters discussed previously, the error rate also depends on the rate of occurrence of the disturbing events (telephony traffic) on the nearby cable pairs.

The most severe coupling occurs when the noise sources are co-located with the digital receiver; in this case the noise voltage e(t) at the decision point may be expressed in terms of the crosstalk coupling C(x) per unit length.

$$e(t) = \int_0^{\ell} C(x)\, b(x,t)dx \qquad (10.2)$$

where b(x,t)dx is the time response at the receiver decision point to a step disturbance coupled through a differential element dx at x, and

$$b(x,t) \xrightarrow{F.T.} B(x,f) = j2\pi f\, V(f)\, e^{-2\gamma(f)x}\, E(f) \qquad (10.3)$$

$\gamma(f)$ is the cable propagation constant, and E(f) is the transfer function from the receiver input to the decision point. Since C(x) is assumed to be a zero-mean normal process over the ensemble of cables and, for any given pair combination m, is uncorrelated with length,

$$E\{C(x)C(y)|\, k_m\} = k_m\, \delta(x-y) \qquad (10.4)$$

It follows that e(t) is $N\{0, r^2(t)\}$ where

$$r^2(t) = k \int_0^{\ell} b^2(x,t)\, dx \qquad (10.5)$$

The noise voltage, λ, required to cause an error is assumed to be half of the minimum eye opening. The expected numbers of errors from an event, averaged over all possible relative phasings of the noise event and the digital signal, is ε.

$$\varepsilon = \frac{\zeta\nu}{T} \int_{-\infty}^{\infty} Q\left\{\frac{\lambda}{\sqrt{k}\, r(t)}\right\} dt \qquad (10.6)$$

where ζ is the duty cycle (1 for EC), ν is a coding factor (1.31 for 4B3T, 1.5 for AMI, 1 for WAL1), and Q is the area in the tail of the standard normal probability density function.

Table 2 - Crosstalk Compatibility Transmission Spans (in km)

DISTURBING SYSTEM	DISTURBED SYSTEM				
	72 kbit/s WAL1	160 kbit/s WAL1	160 kbit/s 4B3T	400 kbit/s AMI	2048 kbit/s HDB3
72 kbit/s WAL1	3.7	3.7	4.4	3.3	1.7
160 kbit/s WAL1	3.8	3.1	4.3	2.7	1.1
160 kbit/s 4B3T	4.2	3.6	4.3	3.2	1.2
400 kbit/s AMI	4.0	2.9	4.4	2.5	1.0
2048 kbit/s HDB3	5.6	3.8	6.0	2.9	1.0

For the rates of occurrence of events observed in the telephone network, the probability of overlap of the responses from events on different disturbers is negligible. Hence the overall error rate is the sum of the error rates from each disturber alone.

$$\text{LTMBER} = \frac{\zeta\nu}{T} \sum_{i=1}^{N} z_i \int_{-\infty}^{\infty} Q\left\{\frac{\lambda}{k_i^{1/2} r(t)}\right\} dt \quad (10.7)$$

where z_i is the average number of events per second on the i-th disturber. Alternatively, the margin [M^2] dB against the performance objective of LTMBER = 10^{-7} can be obtained from:

$$10^{-7} = \frac{\zeta\nu}{T} \sum_{i=1}^{N} z_i \int_{-\infty}^{\infty} Q\left\{\frac{\lambda}{Mk_i^{1/2} r(t)}\right\} dt \quad (10.8)$$

The rate of occurrence of events may be estimated from the average teletraffic ρ (in erlang) by assuming an average call length of 3 minutes and an average of 35 events per call from the customer end (mostly dial pulses).

$$z_i = \frac{35\ \rho_i}{180} \quad \text{event/sec} \quad (10.9)$$

Equation (10.8) may be solved for a fixed set of average traffic levels corresponding to a particular network environment. Alternatively the ρ_i may be generated at random from a population with given statistics; the beta distribution with second parameter equal to 2 adequately represents the distribution of the traffic levels over the analogue circuits (in the local distribution network).

$$f(\rho) = a\ (a+1)\ \rho^{a-1}\ (1-\rho) \quad (10.10)$$

where $a = 2\ E(\rho)/\{1-E(\rho)\}$ is obtained from the mean traffic per line, $E(\rho)$, assumed equal to 0.1 erlang for a busy urban area in the busy hour.

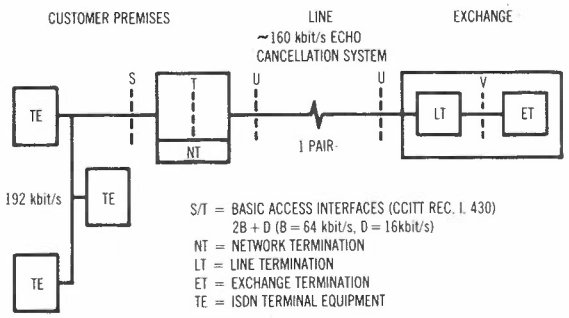


Fig. 12 - Customer and Network Interfaces for Basic ISDN Access.

The k_i are also generated at random from a lognormal density. Indeed [k] is distributed as $N(2.5-\mu_0-[\pi^2 f^2_0/\alpha_0],\ \sigma^2_0-31)$, (Ref. 16). For each set of N randomly generated pairs of k_i and ρ_i a sample margin [M^2] is obtained by iteratively solving (10.8). This process is repeated a thousand times to generate a set of sample margins from which, after sorting, the first percentile corresponding to the design objective is obtained. The line length is iteratively adjusted until this first percentile of the margins is equal to the design margin of 6 dB. The achievable ranges on 0.4 mm paper-insulated, unit-twin, copper cable are included in Table 1.

As for crosstalk, the ranges increase with decreasing effective line rate (or receiver bandwidth). For the chosen transmit level of +10 dBm, the impulsive noise limits are a little more severe than the NEXT limits for EC systems, but are much more severe than the FEXT limits for synchronised TCM systems. The differences in performance between the line codes is less marked than for the crosstalk case; this is because the impulsive noise energy spectrum is almost flat whereas the power spectrum of the crosstalk noise is biased towards the higher frequencies.

11. LSI IMPLEMENTATION

The discussion in the previous sections has been mainly concerned with the performance and network aspects of these systems. An equally important issue that must be resolved if ISDN basic access systems are to be economically introduced is the integration of the transmission functions onto one or a small number of LSI chips. As was briefly indicated in section 1, the functions to be performed at the receiver are particularly difficult with adaptive filtering, equalisation and echo cancellation of complex waveforms to a high degree of accuracy being required as well as other digital operations. A key issue is whether analogue and digital functions should be combined on the one chip or whether all operations should be digital and hence lead to an all-digital chip implementation. Some studies carried out in this area (Ref. 17) have identified that the all digital implementation may offer the following advantages:

* a lower design and development time and cost,
* higher performance,
* a higher yield due to smaller chip size and lower sensitivity to processing parameter variation, and
* lower power consumption.

12. ISDN INTERFACES

A major issue facing telecommunications administrations when developing their ISDN networks is the use of interfaces. Fig.12

shows some of the existing or proposed interfaces. The customer interfaces, S or T, are the subject of CCITT Recommendations and are not considered further in this paper. The U and V interfaces are not currently the subject of CCITT Recommendations but are important to the administrations in that they can permit a flexibility in the network development and equipment supply for the ISDN.

It is not the intention of this paper to discuss the advantages and disadvantages of these interfaces in detail; rather some of the different interface options, as they impact on the transmission aspects of ISDN access, are discussed to illustrate some of the issues that need to be addressed.

A: With a U and a V Interface

* A full specification of both the U and V Interfaces would allow different manufacturers (to the exchange equipment manufacturers) to independently develop and supply the NTs and LTs. In order to ensure correct interworking of equipment, the U Interface specification would need to include the line code, pulse shape, transmit level, immunity to impulsive and crosstalk interference, activation/de-activation procedures, frame format including maintenance and synchronisation functions, and power feeding arrangements. Ideally the specification would leave open the choice of the techniques and technology to be used in the implementation of the system, provided some required transmission performance and range is achieved.

* Alternatively, the use of a common line transmission chip(s), ideally available from several sources, could be specified for use by the manufacturers of the NTs and LTs. The use of such a chip would ensure correct transmission between LT and NT without the need for a full U Interface specification as stated above.

Note that the specification of a V Interface allows the full separation of LT and ET functions and thereby allows LT and exchange equipment to be provided by different manufacturers.

B: With a U and without a V Interface

* This approach is the same as in A above, except that the LT would, by necessity, be provided by the exchange manufacturer.

C: Without a U and with a V Interface

* This approach would allow different manufacturers to supply the combined LT-NT equipment, including the transmission system. Although the LTs from one manufacturer would, in general, not interwork with the NTs from another manufacturer there would be the benefits of the competitive supply of transmission equipment. The disadvantage of this approach is that the administration must ensure that the NT at the customer's premises is of the same type as the LT at the exchange.

* Another more limited option is to allow only one LT-NT transmission supplier who could be different from the switching supplier. This approach avoids the interworking requirement above but does not give any competitiveness in the supply of transmission systems.

D: No U or V Interfaces

* The exchange and associated transmission equipment (i.e., ET, LT and NT) must be provisioned from the same supplier or licensees of this manufacturer.

In each of the cases, A to D, the following requirements must be met.

1. The transmission systems must provide adequate capacity for maintenance information.

2. In general, the transmission systems must serve a high proportion of the network's customers. This implies the use of the echo cancellation technique and an appropriate narrowband line code. However, in the first approach under C, an administration might choose to serve the majority of customers with shorter range systems for, say, cost reasons and to use long reach systems to serve the remaining customers with long loops. This of course means that the NT must match the corresponding LT at the exchange for each customer.

3. The transmission systems must share the cable with other digital systems both ISDN and non-ISDN and hence be compatible with respect to coupled crosstalk interference. This rules out the use of the TCM (Burst) technique.

Another point to be noted is that the interfaces may be either specified internationally by, for example, the CCITT or nationally by either the appropriate national standards setting body or, in some countries, by the telecommunications administration. In some cases the interface standards could be set on a regional basis by the relevant organisation; e.g. Conference of European Post and Telecommunications Administrations (CEPT).

13. CONCLUSIONS

Two different bi-directional digital transmission techniques are candidates for the widescale and economic provision of ISDN basic access to customers over a single pair of wires. Basic access refers to an information capacity of 144 kbit/s, comprising two "B" or 64 kbit/s channels and one "D" or 16 kbit/s channel. The two techniques are known as the time compression multiplex (TCM) technique (requiring a burst line transmission rate of about 400 kbit/s) and the adaptive echo

cancellation (EC) technique (requiring a continuous line transmission rate of about 160 kbit/s). A third candidate, known as the frequency separation technique, is not a contender on the basis that it is too costly to implement.

Impulsive noise probably provides the most severe constraint on the transmission range of the above systems, particularly those based on the TCM technique. In addition, the incompatibility of TCM systems with other digital systems sharing the same cable results in the conclusion that echo cancellation systems are to be preferred. In fact most telecommunications administrations and system designers are now basing their system developments on this technique.

For high immunity to noise and interference, the better line codes are those with lower effective line rates (narrowband transmit power spectral densities) operating in conjunction with narrowband receiver filters. This is reflected in some EC system developments in Europe, where rate reduction codes such as 4B3T, 3B2T and quaternary line codes have been used. EC systems based on these codes should be able to service nearly 100% of urban customers without repeaters or concentrators. Other less optimum line codes, such as conditioned diphase (WAL1), which must operate in conjunction with wider band receiver filters, can lead to simpler realisations, but fail to service as high a percentage of urban customers.

For successful bi-directional transmission, adaptive or automatic cancellation of complicated and unknown waveforms (due to the presence of cable gauge changes, bridged taps and the variety of cable configurations in the local distribution network) by a factor of up to 1000 in voltage is required. This degree of cancellation places severe linearity constraints on echo canceller systems employing linear filters such as the conventional transversal or "tapped delay line" filter (TF). That is, non-linearities in hybrids, analogue to digital and digital to analogue converters, which may form part of the echo or echo control paths, as well as transmitted pulse asymmetry must be kept extremely low. On the other hand, echo canceller systems employing the relatively new "memory compensation" filter (MCF) have the ability to cope with non-linear echo paths but are very slow to adapt to new or changing echo path conditions if the echo path impulse response duration is too long. They also become uneconomic or difficult to realise if the impulse response becomes greater than about 10 bit intervals in duration, as occurs with rate reduction codes such as 4B3T and 3B2T.

For equivalent residual echo power the MCF, employing the stochastic iteration algorithm (SIA) for control purposes, is approximately $2^N/N$ (where N is the duration of the echo path impulse response in bit intervals) times slower to adapt than the TF employing the SIA. This factor is the same if the two filter techniques employ the sign algorithm (SA) instead of the SIA. However, the MCF or TF employing the SA (and incorporating a "dithering signal") can be as much as 25 times slower than the corresponding filter employing the SIA. Filters employing the SA are nevertheless easier to implement than those employing the SIA (because multiplication operations may be replaced by additions and subtractions).

The adaptation speed of echo cancellers employing the SA can be increased significantly by starting the adaptation at a more significant bit in the coefficient register (in the case of the TF) or the memory cell (in the case of the MCF), and then switching to the least significant bit after a certain number of iterations. This corresponds to increasing the gain constant (K_3 or K_4) in order to obtain fast initial adaptation, and then decreasing it in order to obtain a low final value of residual echo power (at the lower adaptation rate). In addition, in the local distribution plant, the line and therefore echo path characteristics do not change dramatically from one call to another. Consequently, the contents of the memory (MCF) or the registers (TF) can be stored from call to call, thus eliminating the long adaptation times of echo cancellers using the SA, once initial adaptation has been established.

The adaptation speed of the MCF can also be increased by dividing the memory into a number of smaller memories and cascading them, but only at the expense of losing the general ability of the MCF technique to cope with non-linear echo paths. This multiple memory technique is not discussed in this paper. However, a good account of the technique may be found in Ref. 8.

The best choice of system, of those considered, for basic access to an ISDN would therefore appear to be the echo cancellation system employing a TF which is adapted according to the SIA and uses a narrowband line code such as 3B2T. Nevertheless a TF which is adapted according to the SA is also attractive because it is easier to realise and has the added advantage that the residual echo power is less sensitive to noise and interference than the TF employing the SIA, even though it is slower to adapt. However, the realisation of the linearity required in the echo and echo control paths, to obtain the high degree of echo cancellation required, on one or a few LSI chips is a very complex task, and one that must first be overcome before these systems become a viable proposition. Nevertheless, it is a task that the major telecommunications manufacturers expect to achieve in the next few years.

The use of the existing metallic cable distribution network need not be the only method of providing basic access to the ISDN. The use of remote multiplexers and concentrators in the local distribution network which are fed from the exchange by digital carrier systems operating on either metal or optical fibre cables, is being adopted by some

Administrations to provide a flexible infrastructure for local distribution. In these cases the provision of basic ISDN access on the relatively shorter pair cable lengths from the remote multiplexer or concentrator to the customer would be much easier. ISDN Basic Access can also be expected to be provided on other media such as, for example, optical fibres and radio including digital mobile radio systems.

The provision of appropriate network interfaces in the ISDN provides a telecommunications administration with flexibility in network development and equipment provision. Whilst there has been minimal progress, in for example the CCITT, in developing international standards for the interface between the transmission equipment at the exchange and that at the customers premises (the U Interface), the issue is still being actively discussed in various areas. In particular it is quite likely that a U Interface standard may emerge from the United States where the Federal Communications Commission (FCC) has defined the "copper termination" as the boundary between the network and the customers' equipment.

In summary, the successful marrying of the latest communications and LSI techniques to a traditional transmission medium (the twisted pair cable network) can be expected, in the next few years, to provide the necessary transmission for ISDN basic access.

14. REFERENCES

1. R.J. Keeler, "Construction and Evaluation of a Decision Feedback Equalizer", Conference Record, 1971 IEEE International Conference on Communications, Montreal, Canada, pp. 8-13, June 1971.

2. P.F. Adams, S.A. Cox and P.J. Glen, "Long Reach Duplex Transmission Systems for ISDN Access", British Telecom Technology Journal, Vol.2, No.2, April 1984.

3. D.D. Falconer, "Adaptive Reference Echo Cancellation", IEEE Trans. Commun., vol. COM-30, pp. 2083-2094, Sept. 1982.

4. N.A.M. Verhoeckx, et. al., "Digital Echo Cancellation for Baseband Data Transmission", IEEE Trans. Acoust., Signal Processing, vol. ASSP-27, pp. 768-781, Dec. 1979.

5. N.A.M. Verhoeckx and T.A.C.M. Claasen, "Some Considerations on the Design of Adaptive Digital Filters Equipped with the Sign Algorithm", IEEE Trans. Commun., vol. COM-32, pp. 258-266, March 1984.

6. P.J. Van Gerwen, N.A.M. Verhoeckx and T.A.C.M. Claasen, "Design Considerations for a 144 kbit/s Digital Transmission Unit for the Local Telephone Network", IEEE Journal on Selected Areas in Communications, vol. SAC-2, pp. 314-323, March 1984.

7. N. Holte and S. Stueflotten, "A New Digital Echo Canceller for Two-wire Subscriber Lines", IEEE Trans. Commun., vol. COM-29, pp. 1573-1580, Nov. 1981.

8. F.G. Bullock, "Echo Canceller Structures for Digital Loop Access Systems", A.T.R., vol. 19, No. 1, pp. 23-41, 1985.

9. A.J. Gibbs and R. Addie, "The Covariance of Near End Crosstalk and its Application to PCM System Engineering in Multipair Cable", IEEE Trans. Commun., vol. COM-27, pp. 469-477, Feb. 1979.

10. CCITT, Red Book, vol. 3, Fascicle 3.3, Recommendation G.821.

11. T.Soejima, T. Tsuda and H. Ogiwara, "Experimental Bidirectional Subscriber Loop Transmission System", IEEE Trans. Commun. vol. COM-30, pp. 2066-2073, Sept. 1982.

12. N.Q. Duc and B.M. Smith, "Line Coding for Digital Data Transmission", A.T.R., vol. 11, pp. 14-27, 1977.

13. A.J. Gibbs, "Measurement of PCM Regenerator Crosstalk Performance", Electron. Lett., vol.15, pp. 82-83, 1979.

14. P.G. Potter, "Reduction of Crosstalk Interference in Local Digital Transmission", Electron. Lett., vol. 17, pp. 423-425, 1981.

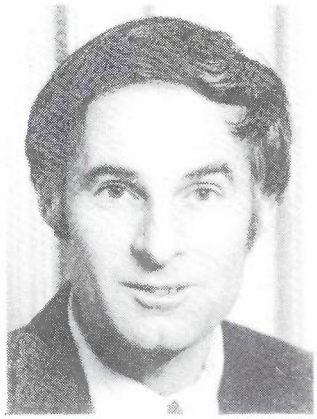
15. B.R. Clarke and G.J. Semple, "Crosstalk Compatibility in the Local Loop", A.T.R., this issue, pp.

16. P.G. Potter and B.M. Smith, "Statistics of Impulsive Noise Crosstalk in Digital Line Systems on Multipair Cable", IEEE Trans. Commun., vol. COM-33, pp. 259-270, Mar. 1985.

17. C.D. Rowles, private communication.

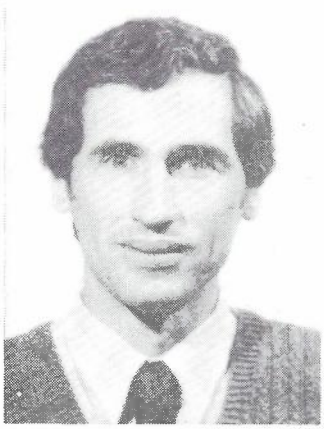
BIOGRAPHIES

NICHOLAS DEMYTKO was born in Kiel, Germany in 1946. He received the B.Sc. and B.E.(Hons). degrees from the University of Adelaide in 1968 and 1969, respectively, and the M.Admin. degree from Monash University in 1979. He was the recipient of a Post Office Cadetship in 1969 and was appointed to the A.P.O. Research Laboratories in 1970. His early work was in the field of echo control on long-distance telephone circuits and the design and development of high speed adaptive digital echo cancellers with the ability to cope with the time varying characteristics of these circuits. He is currently with the Line and Data Systems Section of the Telecom Australia Research Laboratories, where he is engaged in the study of systems and techniques to achieve bi-directional digital transmission, at rates of up to 160 kbit/s, utilising the existing two-wire cable network between customers and their local exchanges.



BERNARD M. SMITH was born in Mount Gambier, South Australia in 1943. He received the B.E. and Ph.D. degrees in electrical engineering from the University of Adelaide in 1964 and 1969 respectively. He joined the Research Laboratories of Telecom Australia in 1968 and has worked on various problems in digital data transmission over the analogue telephone network, digital networks and line transmission systems both for interexchange and customer access applications. He was closely involved with the development of a course on digital transmission systems theory run jointly with Monash University. His current position is Head of the Line and Data Systems Section, Transmission Systems Branch. He is a Senior Member of the Institute of Electrical and Electronics Engineers.

JOHN SEMPLE received the Degrees of Bachelor of Engineering (Electrical) with honours and a Master of Engineering Science from the University of Melbourne in 1966 and 1968 respectively. He commenced work as an Engineer with the then APO Research Laboratories, Pulse Systems Division in January 1967, where he was concerned with problems associated with primary level PCM and high capacity digital transmission systems. In 1972 he joined the Line and Data Systems Section, which is now part of the Transmission Systems Branch of the Telecom Research Laboratories, where he continued work associated with the application of PCM transmission in the inter-exchange cable network. In about 1980 Mr. Semple commenced the investigation of local digital reticulation systems, and is currently concerned with Integrated Services Digital Network (ISDN) developments and the digital two-wire full duplex transmission systems required to provide customer access to such a network.



PHILIP G. POTTER received the B.E. and Ph.D. degrees in electrical engineering from Monash University, Melbourne, Australia in 1974 and 1979, respectively.

Since joining the Research Laboratories of Telecom Australia in 1979, he has worked on various aspects of digital transmission, with particular emphasis on statistical modelling of crosstalk and impulsive noise impairments in digital systems on multipair cable. He is currently with the Wideband Systems Section, Telecom Australia Research Laboratories, Melbourne, and is concerned with transmission and media access techniques in local area networks.

The Effect of Bridged Taps on Local Digital Reticulation Systems

G.J. SEMPLE

Telecom Australia Research Laboratories

In this paper the effect that bridged taps in the Telecom Australia customer access network could have on the transmission performance of local digital reticulation systems is investigated. The range of systems considered includes baseband data modems operating at transmission rates up to 72 kbit/s, the types of systems being developed for future ISDN basic access and 2 Mbit/s primary access line systems. The paper initially discusses the occurrence of bridged taps in the customer distribution network and examines their effect on the frequency response of the cable loops involved. A reflected wave model is developed to provide basic insight into the effects produced by bridged taps and the relative importance of the various parameters involved. The impact of these effects on the performance of the particular local digital reticulations systems considered is then investigated with the aid of computer simulations. For each case the receiver eye diagram is computed and the eye closure produced by the bridged tap is used to quantify the corresponding reduction in noise margin at the system decision point.

The results of the study indicate that the higher the line rate of the digital system the greater the potential interference produced by a bridged tap. For example, the satisfactory operation of a 2 Mbit/s line system on a cable loop with a bridged tap cannot be assured unless the length of the tap is restricted to less than about 25 metres. With ISDN basic access systems bridged taps are not likely to be of major concern if decision feedback equalisation is employed. Where this is not the case their operation on loops with bridged taps can be significantly restricted depending on the length and gauge of the bridged tap. The operation of a 72 kbit/s baseband data modem is not significantly degraded by a bridged tap, provided its length is less than about 300 metres. As most bridged taps tend to be shorter than this length, the majority of cable pairs with bridged taps should be usable with 72 kbit/s modems, provided acceptable levels of noise apply. For baseband data transmission rates below 72 kbit/s, bridged taps are of no concern.

KEYWORDS: Bridged Taps, Tees, Multiples, Digital Transmission, Local Digital Reticulation

1. INTRODUCTION

Past cable layout and installation practices have resulted in a significant number of cable pairs in the Telecom Australia local distribution cable network with one (or more) additional cable pair(s) bridged across some point along their length. These additional pairs, referred to in this paper as "bridged taps" (also known as "tees" or "multiples") may or may not be used at some future time depending on where the demand for new cable pairs occurs. In any case, only one of the outgoing pairs from the bridging point is likely to be used at any given time; the end of the other pair(s) being left open circuited.

Although bridged taps cause no perceptible interference to voice signals they can degrade the transmission performance of digital signals. The amount of degradation depends both on the characteristics of the bridged tap and the type of digital system involved. For example, the voice frequency carrier data modems which have been in use in the local network for about 20 years are not affected by bridged taps. However, the operation of higher speed digital transmission systems which have been introduced into the local network over recent years (such as 72 kbit/s baseband data modems and 2 Mbit/s digital line systems) is susceptible to bridged taps. Since the presence of bridged taps can significantly degrade the operation of these systems, cable pairs with bridged taps have, in the past, been avoided for the provision of these services. However, as the demand for these services increases it will become increasingly difficult to continue to avoid pairs with bridged taps and, since it is expensive and time consuming to find and remove them, it is desirable to know under what circumstances (if any) they can be tolerated.

In the future, with the introduction of an integrated services digital network (ISDN), two-wire bi-directional digital transmission

Paper received 25 July 1985.
Final revision 16 August 1985.

systems which operate at about 160 kBaud will be introduced to provide customer access to the network. Two alternative types of systems based on the burst mode and echo cancellation techniques are being developed to achieve the bi-directional operation. These systems are inherently complex and may include adaptive decision feedback equalization (DFE) (Ref. 1) to accommodate the signal dispersion associated with the received signal. When this is the case the forward signal echoes produced by bridged taps can generally also be accommodated. However, some telecommunication administrations do not use bridged taps in their network and some ISDN basic access system developments do not include a DFE. Before these later types of systems can be considered for use in the Telecom Australia network, a knowledge of the effect of bridged taps is required.

Other digital reticulation systems under development around the world include pre-ISDN integrated customer access systems designed to provide customers with simultaneous voice and data on a single cable pair, and four-wire digital transmission systems for digital PABX extensions. Since these systems are generally being developed without DFEs, a knowledge of the limitations imposed by bridged taps on their operation is also essential if they are to find application in the Telecom customer access network. The paper also examines the effect of bridged taps on the operation of 2 Mbit/s primary access systems.

In this paper the effect of bridged taps on the above systems is considered. A reflected wave model is used to provide basic insight into the effects caused by bridged taps and to establish the relative importance of the various parameters involved. The results of computer simulations are then used to illustrate the effect of bridged taps on particular systems. The reduction in noise margin is used as a measure of the performance degradation.

2. THE OCCURRENCE OF BRIDGED TAPS

Due to variations in the local cable installation practices which have occurred over time and between regions, the occurrence of bridged taps varies from one area of the Telecom network to another. Bridged taps can occur both in the main and distribution cable segments which service a particular customer area. They occur along the cable at points where flexibility is required to accommodate uncertainty in future service demand. In main cables, about 5% of the cable pairs might typically contain bridged taps which generally terminate on a pillar or cabinet. In the distribution area, a "tailing on" practice leads to possibly more than 50% of all cable pairs with one or more bridged taps. With this practice a pair of wires in a distribution or branch cable is intercepted with a drop wire to provide a customer with a service; the redundant or excess wire remains attached for the remaining part of the distribution run. A common practice allocates 4 drop wires per pole; these cable pairs could be said to have a single appearance with a tail (i.e. bridged tap). In addition, 2 cable pairs in every 10 pairs are allocated for future use, such as the supply of a second service to a given customer's premises. These cable pairs have so-called "multiple appearances"; they may or may not have drop wires connected (i.e. multiple bridged taps may occur).

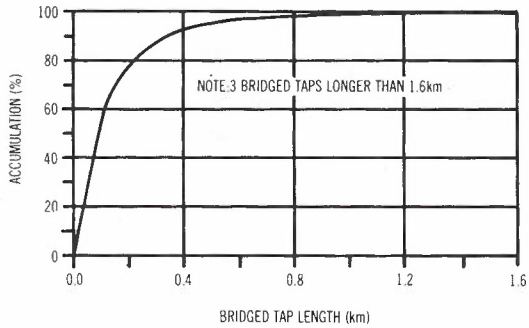


Fig. 1 - Cumulative Distribution of Bridged Tap Lengths (421 Samples, All States)

In Fig.1 a cumulative distribution of bridged tap lengths is given. This information is based on an analysis of data obtained from a customer loop survey carried out by Telecom Australia (Refs. 2, 3). It is seen that about 90 % of all bridged taps have lengths less than 350 m. It is also seen that briged tap lengths in excess of 1.6 km occasionally occur.

3. THEORETICAL ANALYSIS

3.1 Chain matrix representation

A bridged tap, which from an electrical point of view acts as an open circuited stub, is shown diagrammatically in Fig.2. The cable segments on each side of the bridged tap are labelled 1 and 2, and the bridged tap is labelled 3. The cable parameters associated with each cable segment are identified by appropriate subscripts. The cable is assumed to be driven from a voltage source $V_S(f)$ through a source impedance, $Z_S(f)$. The output signal $V_L(f)$ appears across a load impedance $Z_L(f)$.

If each cable segment is represented by the four terminal network shown in Fig.3, then the input/output voltage/current conditions of each segment can be related by the ABCD chain matrix such that (Ref. 4),

$$\begin{bmatrix} V_I \\ I_I \end{bmatrix} = \begin{bmatrix} A_i & B_i \\ C_i & D_i \end{bmatrix} \begin{bmatrix} V_{I+i} \\ I_{I+i} \end{bmatrix} \qquad (1)$$

For a transmission line the elements of the chain matrix have the values

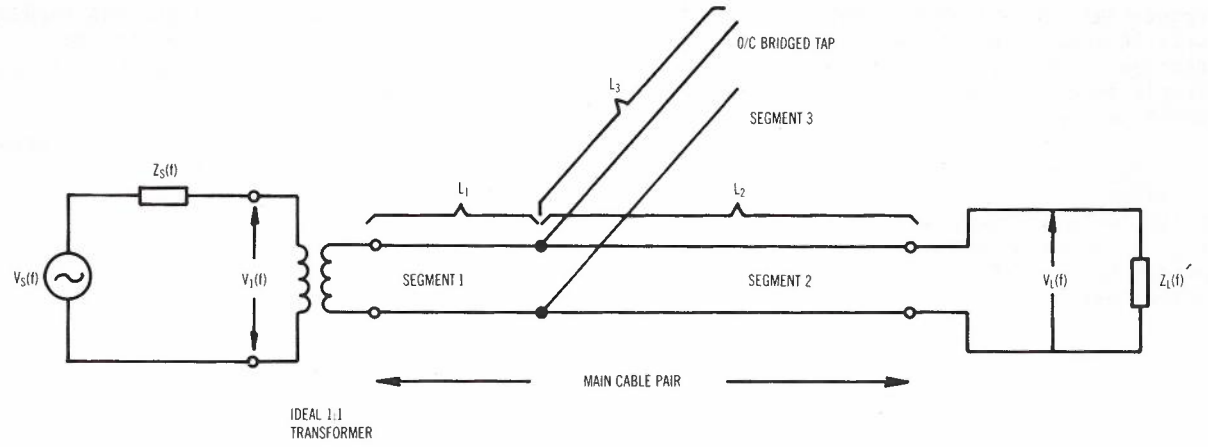


Fig. 2 - Model of Cable Loop with Single Bridged Tap

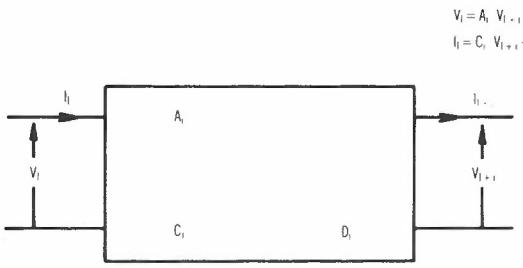


Fig. 3 - Four Terminal Network Representation of a Cable Segment

$$A_i = D_i = \cosh \gamma_i L_i$$

$$B_i = Z_{0i} \sinh \gamma_i L_i \qquad (2)$$

$$C_i = \sinh \gamma_i L_i / Z_{0i}$$

where in this case

i = (1,2,3)

$\gamma_i = \alpha_i + j\beta_i$: cable propagation constant (per unit length)

α_i = attenuation per unit length (neper/km)

β_i = phase shift per unit length (radians/km)

Z_{0i} = characteristic impedance of cable segment i

L_i = i th cable segment length (km)

Obviously, γ_i, α_i, β_i and Z_{0i} are functions of frequency.

By chain matrix manipulation it can be shown that the overall input/output characteristics of the cable loop shown in Fig.1 are given by,

$$\begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ I_2 \end{bmatrix} \qquad (3)$$

where

$$A = A_1A_2 + B_1C_2 + B_1A_2C_3/A_3$$

$$B = A_1B_2 + B_1D_2 + B_1B_2C_3/A_3$$

$$C = C_1A_2 + D_1C_2 + D_1A_2C_3/A_3$$

$$D = C_1B_2 + D_1D_2 + D_1B_2C_3/A_3$$

By further matrix manipulation it can be shown that (Ref. 4)

$$\frac{V_L}{V_S} = T(f) = \frac{1}{Z_LC + A + DZ_S/Z_L + B/Z_L} \qquad (4)$$

and

$$\frac{V_L}{V_1} = G(f) = \frac{1}{A + B/Z_L} \qquad (5)$$

where G(f) is the voltage gain of the loop (all variables are obviously functions of frequency).

To allow computer simulations to be performed, the above approach has been extended in a computer program to accommodate any arbitrary cable loop configuration. The computer program calculates G(f) and T(f) from cable data stored on computer files (Ref. 3). Actual values of γ_i and Z_{0i} at a large number of appropriately spaced

frequencies for all of the commonly used cable types in the Telecom network have been recorded from very accurate open and short circuit impedance measurements on a 5 metre length of each type of cable.

The program also allows the key parameters of any arbitrary transmission system to be included so that the reduction in receiver noise margin produced by a bridged tap can be quantified. Results for particular systems are discussed in Section 5.

3.2 Reflected Wave Model

An alternative way to establish the effect of bridged taps is to adopt a reflected wave approach. With this approach the behaviour of the signal is considered as it propagates along the various cable segments, and reflection (ρ) and transmission (η) coefficients are used to build up the overall response. For example, with reference to Fig.2, assume a signal $V_1(f)$ is applied to the input of the cable loop. The response of this signal after it has propagated a distance of x km along the cable is given by $V_x(f)$, where

$$V_x(f) = V_1(f)\, e^{-\gamma_i(f)x} \qquad (6)$$

When the signal $V_x(f)$ hits the impedance discontinuity at the junction of the bridged tap, part of it is reflected with a reduced amplitude given by the voltage reflection coefficient ρ_{23}, where

$$\rho_{23} = \frac{(Z_{02}//Z_{03}) - Z_{01}}{(Z_{02}//Z_{03}) + Z_{01}} \qquad (7)$$

with the symbol (//) indicating impedances in parallel. The remaining part of the signal propagates down both the bridged tap (cable segment 3) and cable segment 2 with the reduced amplitudes given by η_{13} and η_{12} respectively.

As a result of voltage continuity across the junction it is seen that

$$\eta_{13} = \eta_{12} = 1 + \rho_{23} = \frac{2(Z_{02}//Z_{03})}{(Z_{02}//Z_{03}) + Z_{01}} \qquad (8)$$

The signal which propagates down the bridged tap is reflected at the open circuit with a reflection coefficient of 1. On return to the junction of the bridged tap it suffers a similar fate to the original signal with reflection and transmission coefficients ρ_{12} and η_{32}, η_{31} respectively, where

$$\rho_{12} = \frac{(Z_{01}//Z_{02}) - Z_{03}}{(Z_{01}//Z_{02}) + Z_{03}} \qquad (9)$$

and

$$\eta_{32} = \eta_{31} = 1 + \rho_{12} \qquad (10)$$

Forward signals transmitted past the the bridged tap junction towards the output of the cable loop experience a further impedance discontinuity at Z_L. This causes further reflected signals with their levels attenuated by the reflection coefficient ρ_L, where

$$\rho_L = \frac{Z_L - Z_{02}}{Z_L + Z_{02}} \qquad (11)$$

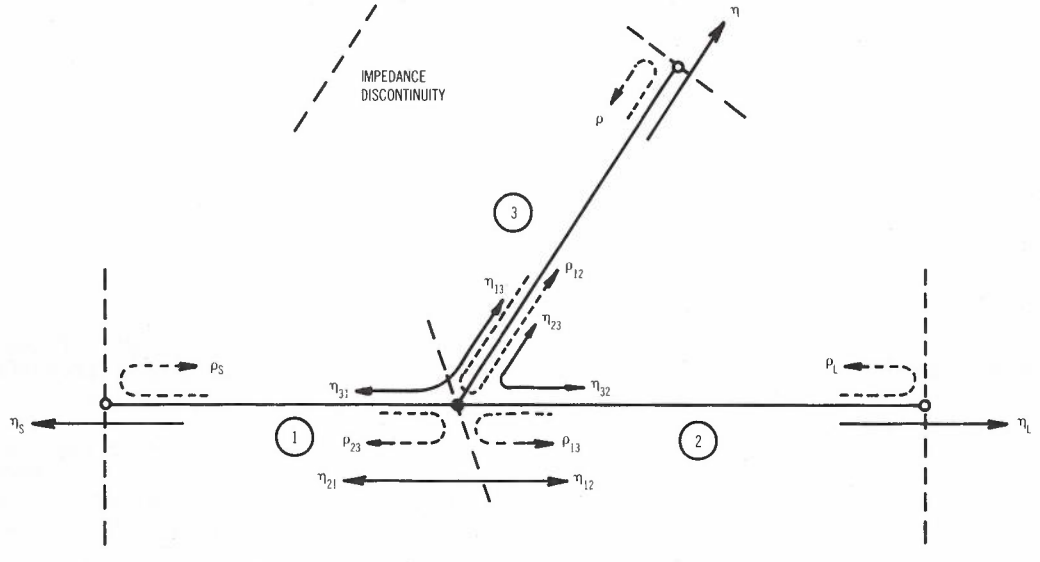


Fig. 4 - Transmission and Reflection Coefficients Associated with Cable Loop and Bridged Tap

The signals absorbed in the terminating impedance are attenuated by the transmission coefficient η_L, where

$$\eta_L = 1 + \rho_L \qquad (12)$$

Similarly, the transmission and reflection coefficients at all the remaining points of impedance discontinuity for the cable loop configuration shown in Fig.2 can be calculated. The complete set are shown in Fig.4. The magnitude of these coefficients with all impedances matched (which is often a reasonable approximation to practical situations), that is, for $Z_{01} = Z_{02} = Z_{03} = Z_S = Z_L$, are such that

$$\rho_{12} = \rho_{13} = \rho_{23} = -1/3$$

$$\rho_S = \rho_L = 0$$

$$\eta_S = \eta_L = 1$$

$$\eta_{12} = \eta_{13} = \eta_{32} = \eta_{31} = \eta_{21} = \eta_{23} = 2/3$$

By using these transmission and reflection coefficients the transfer function of the subscriber loop can ideally be directly written down. However, in practice the output is produced by an infinite sum of signal waves and it is not be possible to directly write down the complete response as given by (4) or (5). Nevertheless, for a number of specific useful cases, it is possible to directly write down a transfer function which closely approximates the actual transfer function.

3.3 Illustration of Reflected Wave Model

3.3.1 Matched Terminating Impedances. If $Z_S = Z_{01}$ and $Z_L = Z_{02}$, then

$$\eta_S = \eta_L = 1, \text{ and } \rho_S = \rho_L = 0$$

Therefore, the cable loop transfer function can be directly written down since the only impedance discontinuities occur at the junction of the bridged tap and at its remote end. That is,

$$G(f) = e^{-\gamma_1 L_1}\, \eta_{12}\, e^{-\gamma_2 L_2} + e^{-\gamma_1 L_1}\, \eta_{13}\, e^{-2\gamma_3 L_3}\, \rho_{12}\, e^{-2\gamma_3 L_3}\, \eta_{32}\, e^{-\gamma_2 L_2}$$

$$\vdots$$

$$+ e^{-\gamma_1 L_1}\, \eta_{13}\, \rho_{12}^{n-1}\, e^{-2n\gamma_3 L_3}\, \eta_{32}\, e^{-\gamma_2 L_2}$$

$$= e^{-\gamma_1 L_1}\, e^{-\gamma_2 L_2} \left[\eta_{12} + \frac{\eta_{13}\,\eta_{32}}{\rho_{12}} \sum_{n=1}^{\infty} \left(\rho_{12}\, e^{-2\gamma_3 L_3} \right)^n \right] \qquad (13)$$

However, since $\left| \rho_{12}\, e^{-2\gamma_3 L_3} \right| < 1$ and $n = \infty$,

$$G(f) = e^{-\gamma_1 L_1}\, e^{-\gamma_2 L_2} \left[\eta_{12} + \frac{\eta_{13}\,\eta_{32}\, e^{-2\gamma_3 L_3}}{(1 - \rho_{12}\, e^{-2\gamma_3 L_3})} \right] \qquad (14)$$

If the length of the bridged tap is short or γ is small, i.e.

$$e^{-2\gamma_3 L_3} = 1,$$

then

$$G(f) = e^{-\gamma_1 L_1}\, e^{-\gamma_2 L_2} \left[\eta_{12} + \frac{\eta_{13}\,\eta_{32}}{(1 - \rho_{12})} \right] \qquad (15)$$

If $Z_{01} = Z_{02}$ and $\gamma_1 = \gamma_2 = \gamma$,

$$G(f) = e^{-\gamma(L_1 + L_2)} = e^{-\gamma L} \quad \text{as expected.} \qquad (16)$$

If the length of the bridged tap is very long or γ_3 is large, ie.

$$e^{-2\gamma_3 L_3} \approx 0,$$

then

$$G(f) = \eta_{12}' \, e^{-\gamma_1 L_1} \, e^{-\gamma_2 L_2}$$

If all cable segments have the same gauge, i.e.

$$Z_{01} = Z_{02} = Z_{03},$$

then

$$G(f) = 2/3 \, e^{-\gamma(L_1 + L_2)} \qquad (17)$$

That is, the bridged tap causes a flat loss of 2/3 or 3.5 dB.

Between the above extremes the effect of the bridged tap is described by H(f), where

$$G(f) = H(f) \, e^{-\gamma_1 L_1} \, e^{-\gamma_2 L_2} \qquad (18)$$

with

$$H(f) = \eta_{12} + \frac{\eta_{13} \, \eta_{32}}{(1 - \rho_{12} \, e^{-2\gamma_3 L_3})} \qquad (19)$$

The form of H(f) is illustrated in Fig.5 and Fig.6 for a 200 m and 1200 m length bridged tap composed of 0.64 mm PIUT cable. It can be seen that the bridged tap imposes a cyclic ripple in both the attenuation and phase response of the cable loop (Ref. 6). The repetition interval of the ripple depends on the length (propagation delay) of the bridged tap, with the first and most significant dip in the attenuation response occurring at the frequency at which the time taken for the signal to propagate along the bridged tap and return corresponds to a half period of the signal frequency. The magnitude of the envelope, which bounds the extremes in the attenuation and phase ripples exponentially, reduces with increasing frequency and is dependent on the gauge of the bridged tap cable; the heavier the gauge the larger the envelope of the ripple.

3.3.2 Mismatched Terminating Impedances. If the source and load terminations of the cable loop are not matched to the adjacent cable segments, signal reflections occur at these points. For example, a signal reflected from an impedance mismatch at the output will travel back to the junction of the bridged tap where part of it will again be reflected back to the terminating load and so on. If only reflections between the bridged tap and the ends of the three cable segments are considered it is readily seen that the cable loop transfer function is given by:

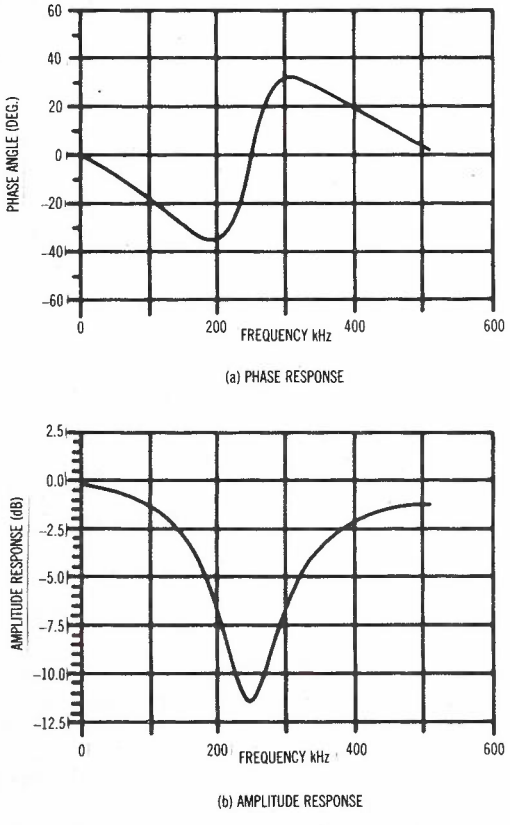


Fig. 5 - Response Imposed on Cable Transfer Function H(f) by Addition of a Bridged Tap.
Cable Type is 0.64mm PIUT
Bridged Tap Length = 200m

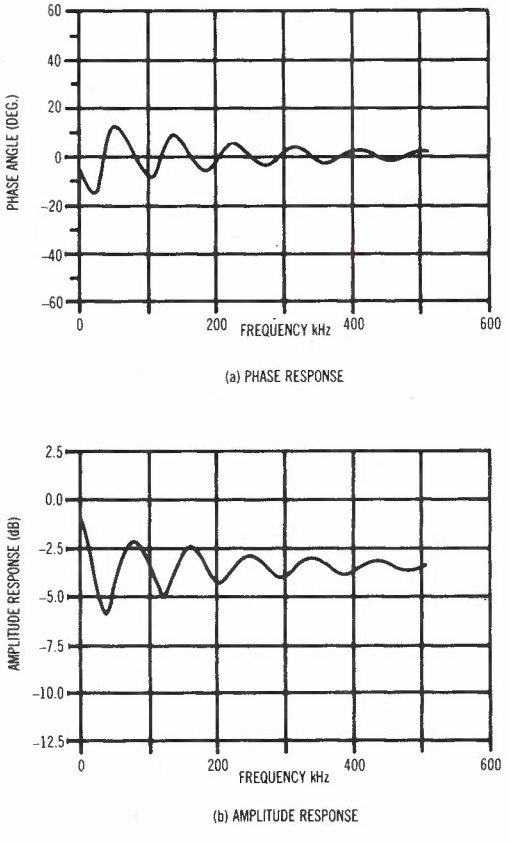


Fig. 6 - Response Imposed on Cable Transfer Function H(f) by Addition of a Bridged Tap.
Cable Type is 0.64mm PIUT
Bridged Tap Length = 1200 m

$$G(f) = \eta_L e^{-\gamma_1 L_1} e^{-\gamma_2 L_2} \cdot \left[H(f) + \eta_{12} \sum_{n=1}^{\infty} (\rho_{23} \rho_S \rho_L e^{-2\gamma_1 L_1})^n + \eta_{12} \sum_{n=1}^{\infty} (\rho_{23} \rho_L e^{-2\gamma_2 L_2})^n \right] \quad (20)$$

If only the dominant first echo components are included and the case where

$$Z_{01} = Z_{02} = Z_{03} \neq Z_S \text{ or } Z_L$$

is considered, then

$$G(f) = \eta_L e^{-\gamma_1 L_1} e^{-\gamma_2 L_2} \left[2/3 + 4/9\, e^{-2\gamma_3 L_3} - 2/9 \, (\rho_L e^{-2\gamma_2 L_2} + \rho_S e^{-2\gamma_1 L_1}) \right] \quad (21)$$

However, since the terminating impedances would generally be selected such that ρ_L and ρ_S would be less than 0.1, the echo component caused by the bridged tap is generally significantly larger (>20 dB) than the echo components produced by mismatched terminations.

3.3.3 The Effect of the Position of the Bridged Tap. If all impedances are matched then it is seen from (18) that G(f) is independent of the position of the bridged tap. However, if the load and source impedances do not match the characteristic impedance of the adjacent cable segments, then G(f) is given by (21) and is dependent on the position of the bridged tap.

If the bridged tap is some distance from the mismatched terminations then from (21) it is seen that

$$G(f) = \eta_L e^{-\gamma_1 L_1} e^{-\gamma_2 L_2} H(f) \quad (22)$$

If the three cable types are identical, then (22) simplifies to

$$G(f) = 2 \eta_L e^{-\gamma_1 L_1} e^{-\gamma_2 L_2} \left[\frac{1 + e^{-2\gamma_3 L_3}}{3 + e^{-2\gamma_3 L_3}} \right] \quad (23)$$

Furthermore, if $\gamma_1 = \gamma_2 = \gamma$, and $L_1 + L_2 = L$, then

$$G(f) = 2 \eta_L e^{-\gamma L} \left[\frac{1 + e^{-2\gamma_3 L_3}}{3 + e^{-2\gamma_3 L_3}} \right] \quad (24)$$

or, if only the first echo term is considered,

$$G(f) = \frac{4a}{9(a+1)} e^{-\gamma L} (3 + 2e^{-2\gamma_3 L_3}) \quad (25)$$

where $a = Z_L/Z_0$

If the bridged tap is at the end of the cable loop, the expression for G(f) becomes

$$G(f) = \eta_S e^{-L} \left[\eta_{1L} + \frac{\eta_{13} \eta_{3L} e^{-2\gamma_3 L_3}}{(1+\rho_{1L} e^{-2\gamma_3 L_3})} \right] \quad (26)$$

where

$$\eta_{13} = \eta_{3L} = \eta_{1L} = \frac{Z_L//Z_0 - Z_0}{Z_L//Z_0 + Z_0} + 1$$

$$= \frac{2a}{(2a + 1)}$$

$$\rho_{1L} = \frac{-1}{(2a + 1)}$$

or, if only the first echo component is considered,

$$G(f) = \frac{2a}{(2a + 1)} e^{-\gamma L} \left[1 + \frac{2a}{(2a + 1)} e^{-2\gamma_3 L_3} \right] \quad (27)$$

From a comparison of (26)(27) and (24)(25) it is seen that the interference caused by a bridged tap will vary slightly, in general, depending on its position, except for the specific case when a = 1 (i.e. when all impedances are matched).

3.3.4 Multiple Bridged Taps. If a cable pair has multiple bridged taps then the reflected wave approach can be used to build up a picture of the effect that they would have on the transmission performance of a digital signal. Although extra reflections occur between the junction points of the various bridged taps these can generally be ignored.

If only the main echo components are considered, the loop transfer function $G_2(f)$ of a cable loop of length L with bridged taps of length L_1 and L_2 can be directly written down for the case of matching cable termination impedances as

$$G_2(f) = \frac{2}{3}\frac{2}{3} e^{-\gamma L} \cdot \left[1 + \frac{2}{3} e^{-\gamma L_1} + \frac{2}{3} e^{-\gamma L_2} + \frac{2}{3}\frac{2}{3} e^{-\gamma(L_1+L_2)}\right] \quad (28)$$

Similarly the loop transfer function $G_3(f)$ for three bridged taps of length L_1, L_2 and L_3 would be

$$G(f) = \left(\frac{2}{3}\right)^3 e^{-\gamma L}\left\{1 + \frac{2}{3}\sum_{i=1}^{3} e^{-\gamma L_i} + \left(\frac{2}{3}\right)^2\left[e^{-\gamma(L_1+L_2)} + e^{-\gamma(L_1+L_3)} + e^{-\gamma(L_2+L_3)}\right] + \left(\frac{2}{3}\right)^3 e^{-\gamma(L_1+L_2+L_3)}\right\} \quad (29)$$

and so on.

Each bridged tap produces transmission loss and echoes from the various reflection points. A particular case of interest is where there are multiple bridged taps of the same length. For this case it can be seen that the magnitude of the composite echo produced from the first echo produced by each bridged tap can exceed the amplitude of the original signal. That is, it is possible for

$$\frac{2}{3}\sum_{i=1}^{n} e^{-\gamma L_i} > 1 \quad (30)$$

This situation is illustrated in Ref. 7 for a case of four bridged taps. In general this situation would prevent digital transmission unless sophisticated echo cancellation techniques were used. However, since multiple bridged taps occur infrequently in the Telecom Australia customer loop network they are not considered further in this paper.

4. SIMULATION OF THE EFFECT OF BRIDGED TAPS ON DIGITAL TRANSMISSION SYSTEMS

The previous section modelled the effect of bridged taps in the frequency domain. The discussion was independent of particular system parameters but illustrated the way backward and forward signal echoes are produced at points of impedance discontinuity. In the time domain, with digital signal transmission the backward and forward signal echoes produce intersymbol interference in the receive signal which reduces the system noise margin. The degree of degradation in noise margin depends on the particular system (e.g. its line rate, line code, etc.) and on the length and gauge of the bridged tap cable pair. The reduction in noise margin can also depend on where the noise (i.e. crosstalk and/or impulsive noise) is coupled into the disturbed cable pair relative to the position of the bridged tap as discussed in Section 4.3.

In this section the results are given for a computer simulation of the effect of bridged taps on the types of digital customer access systems listed below, viz:

(i) Baseband Data Modems

(ii) Basic ISDN access

(iii) Primary (2 Mbit/s) access.

Results are given for various bridged tap lengths and cable types. However, the paper only considers systems that use an AMI or Diphase (WAL1) line code.

4.1 System Model

The results in this paper specifically relate to baseband digital transmission systems which only use linear signal equalization, since these systems are particularly vulnerable to the effects of bridged taps. In contrast, digital baseband transmission systems which include adaptive decision feedback equalization (DFE) can generally accommodate bridged taps provided the initial level of eye closure that is produced, particularly by bridged taps around the critical length, does not prevent the DFE adaptation process from operating correctly. The paper does not specifically investigate the number of baud intervals over which significant echo signals persist as a result of a bridged tap, since the number of DFE taps required in systems which inherently rely on a DFE for normal operation on cable loops without bridged taps (e.g. systems which use rate reduction codes such as 4B3T (Ref. 8), 3B2T (Ref. 9) and 2B1Q (Ref. 10), far exceeds the number of DFE taps required to accommodate the effects of a bridged tap.

The system model used for the computer simulations is shown in Fig.7. A pseudo random test sequence is applied to an AMI or Diphase (WAL1) line coder which drives a transmit pulse generator which produces half width transmit pulses S(f). The voltage transfer function of the cable loop, G(f), is computed as discussed in Section 3.1. The system model includes a fixed receiver equalizer, E(f), which compensates for the response of the cable loop with the bridged tap removed. A receiver pulse shaping network, R(f), is included which, together with S(f), produces a 100 % raised cosine signal, D(f), relative to the effective baud rate of the line signal.

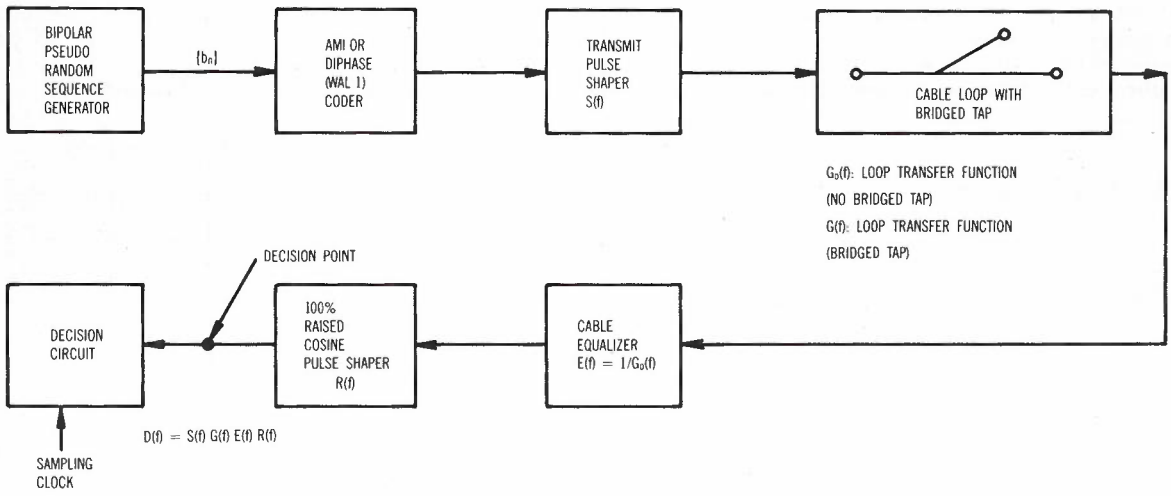


Fig. 7 - System Model Used for Computer Simulation

Although the computer simulation model uses a fixed receiver equalizer the results are also applicable to systems which use adaptive Bode type receiver equalization, as discussed below.

Bode equalizers are designed for a given cable attenuation characteristic and are automatically adjusted for a given cable loss by a control signal generally derived from the peak voltage of the equalized signal (i.e., at the receiver decision point). The Bode equalizer compensates for both loss and frequency response in a manner which maintains a constant peak signal voltage at the decision point. Consequently, it might appear that the response of the Bode equalizer could be significantly affected by the attenuation and intersymbol interference introduced by a bridged tap and therefore a fixed equalizer model may not be adequate. However, this is not the case.

For small changes in cable loss the major change in the Bode equalizer response is a compensating change in equalizer gain; the relatively small change in frequency response and any consequential change in the amount of intersymbol interference in the receive signal is insignificant. This change in equalizer gain produces a corresponding reduction in noise margin. For this case the Bode equalizer is hence adequately modelled by a fixed equalizer, where the reduction in noise margin is reflected by the reduction in signal level produced by the extra loop loss introduced by the bridged tap.

As will be illustrated, bridged taps can produce excessive intersymbol interference in the received signal of a digital system, causing a significant reduction in noise margin. However, when excessive intersymbol interference is produced, the signal attenuation introduced by the bridged tap tends to be offset by the intersymbol interference such that the peak level of the signal at the output of a fixed equalizer tends to remain relatively constant. Consequently, the response of a Bode equalizer is relatively unchanged irrespective of whether the bridged tap is present or not, even though a significant reduction in noise margin is produced when the bridged tap is present. That is, the fixed equalizer model again adequately reflects the behavour of a Bode equalizer.

The flat 3.5 dB attenuation produced by a long bridged tap is compensated by a corresponding increase in the gain of a Bode equalizer; the change is not exactly 3.5 dB due to the small change in the frequency response which in turn produces a small change in the level of intersymbol interference (i.e. on the peak signal level) and hence on the amount of gain adjustment. Although with the fixed equalizer model the signal is attenuated by the 3.5 dB (exactly), producing a corresponding reduction in noise margin, the difference between the two cases is small.

4.2 Reduction in Noise Margin Measure

If the response of a cable loop with a bridged tap removed is denoted by $G_o(f)$, then

$$E(f) = 1/G_o(f) \qquad (31)$$

By design

$$S(f)R(f) = \begin{cases} 0.5aT\left[1-\sin(\pi afT-0.5\pi)\right] & : \ 0<f<1/aT \\ 0 & : \ f>1/aT \end{cases} \qquad (32)$$

where: $a = 1$ for AMI
$= 0.5$ for Diphase (WAL1)

The Fourier transform of the pulse response d(t) at the decision point is given by D(f), where

$$D(f) = S(f)\ G(f)\ E(f)\ R(f)$$
$$= S(f)\ G(f)\ R(f)/G_o(f) \qquad (33)$$

The simulation program computes D(f) and then takes its inverse Fourier transform to compute d(t). The receive code symbol response y(t) depends on the line code and is given by

$$y(t) = (d(t - a/T) - d(t + a/T))/(2a) \qquad (34)$$

where the (a) factor in the denominator is included to maintain a constant peak to peak voltage for both the AMI and Diphase line codes.

For a random bipolar (NRZ) transmit data sequence b_n, where $b_n = (1,-1)$, the receive signal z(t) is given by

$$z(t) = \sum_{n=-\infty}^{\infty} b_n \, y(t - nT) \qquad (35)$$

The computer program computes the receive signal eye diagram from z(t). The reduction in noise margin produced by a bridged tap is quantified from the peak eye closure (i.e. the minimum eye opening) at the optimum sampling time (t_0) relative to the peak eye closure with the bridged tap removed.

The eye diagrams produced by z(t) for AMI and Diphase line codes are shown in Figs. 8(a) and 9(a) respectively, for the case of a 144 kbit/s line signal transmitted over 5000m of 0.90mm PIUT cable loop with no bridged taps. Figs. 8(b) and 9(b) show the corresponding eye diagrams produced by these two codes when a 500m bridged tap of 0.90 mm PIUT cable is connected across the centre of the main cable loop. (Not all signal transitions are shown - the dotted lines indicate the inner and outer boundaries of the eye patterns.) It can be seen that the bridged tap causes signal attenuation and introduces intersymbol

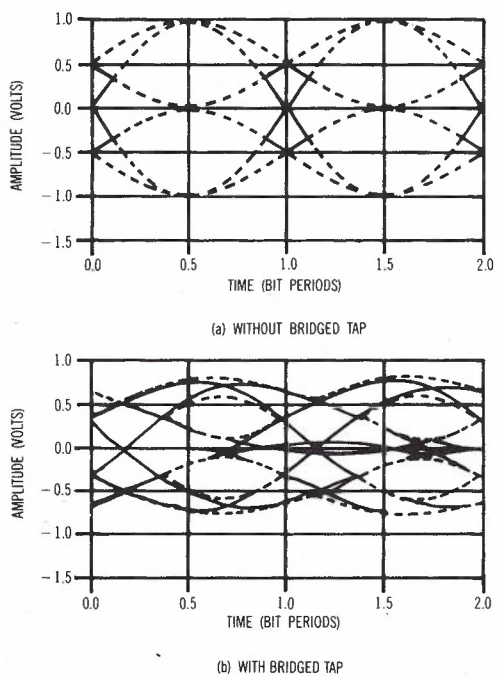


Fig. 8 - Eye Diagram Produced by AMI Line Code with 100% Raised Cosine Equalisation

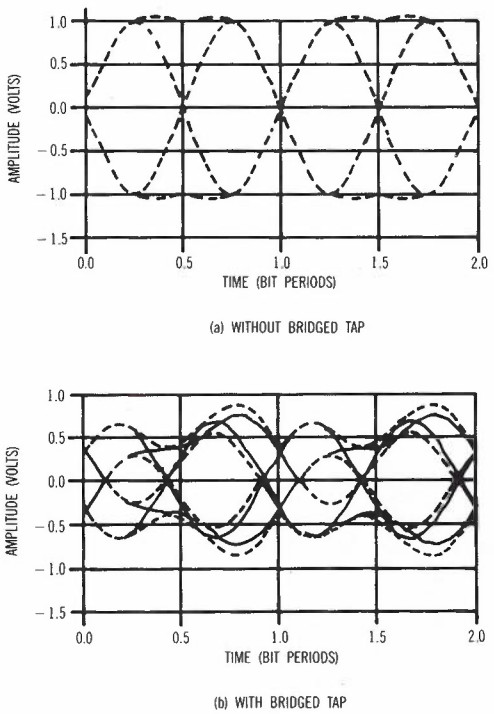


Fig. 9 - Eye Diagram Produced by WAL1 Line Code with 100% Raised Cosine Equalisation

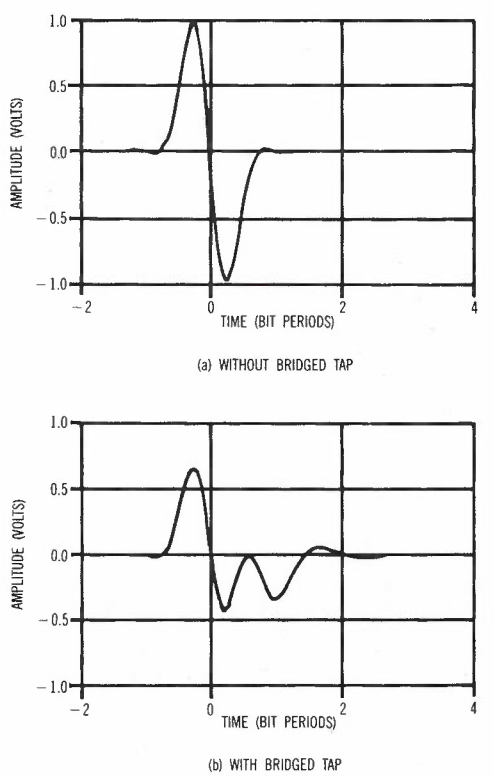


Fig. 10 - Pulse Response Produced by WAL1 Line Code with 100% Raised Cosine Equalisation

interference. The receive pulse responses, y(t), produced by the basic code symbol for the cases illustrated in Fig. 9 are shown in Fig. 10. The effect of the echo components on the received code symbol response is clearly seen in Fig. 10(b).

For a WALI (binary) code the peak eye closure is given by E_{WALI}, where

$$E_{WALI} = 2|y(t_o)| - \sum_{n=-\infty}^{\infty} |y(t_o - nT)| \qquad (36)$$

For an AMI code it can be shown that if the decision level is set at the middle of the eye opening the peak eye closure is given by E_{AMI}, where

$$E_{AMI} = 1.5\,(|y(t_o)| + |y(t_o + T)|) - 0.5\,(||y(t_o)| - |y(t_o + T)||) - \sum_{n=-\infty}^{\infty} |y(t_o - nT)| \qquad (37)$$

4.3 Other Factors Which Affect Noise Margin

A further factor which must be considered when assessing the performance of a digital transmission system and the reduction in noise margin produced by a bridged tap is the type of interference. If the interference is coupled into the disturbed cable loop via a near end crosstalk (NEXT) path then the interference it causes can also depend on the position of the bridged tap. For example, if the bridged tap is at the far end or on the far end side of the near end cable segment along which the dominant NEXT coupling occurs, then the level of the noise appearing at the input to the receiver (i.e. before the input equalizer) will be independent of the bridged tap. Any errors the noise may produce are therefore only dependent on the characteristics of the noise and on the receiver sensitivity which, as seen earlier, depends on the impairment produced by the bridged tap. However, if the bridged tap is at the near end, it will cause transmission loss to not only the transmitted signal but also to the coupled interference. This will tend to reduce the interference produced by the noise. For example, with a very long bridged tap located at the input of a receiver, both the noise and the signal will be attenuated by 3.5 dB and therefore the noise will have the same effect on the receiver as it would have had, had the bridged tap not been present. If the same bridged tap was located on the far side of the NEXT coupling path, the signal would still experience the 3.5 dB attenuation, but the noise would not. That is, the signal to noise ratio would be 3.5 dB worse than in the previous case. If the bridged tap occurs within the cable segment along which significant crosstalk coupling occurs, the interference produced will be somewhere between the cases cited above.

Short bridged taps have minimal effect on the noise and signal irrespective of where they occur. However, the overall effect caused by a medium length bridged tap located at the near end to a receiver is not so clear. Under these circumstances the bridged tap attenuates the main noise component but also causes a delayed and further attenuated component of the same noise to appear at the receiver input. The combined interference caused by the direct and delayed noise components will depend on their relative amplitudes and the amount of delay (i.e. on the amount of correlation between the two noise components at the decision point). A detailed analysis of this latter case is beyond the scope of this paper.

In this paper, the computed reductions in noise margin only refer to the eye closure produced in the received signal and do not include the above effects.

5. SIMULATION RESULTS

5.1 Baseband Data Modems

Four-wire baseband modems which operate at 3.2, 6.4, 12.8 or 72.0 kbit/s are used to provide access to the Telecom Australia Digital Data Service (DDS). All modems use the Diphase line code and their receiver pulse shaping produces an output signal closely represented by a 100% raised cosine.

Since the 72 kbit/s modem is the one most sensitive to the effects of bridged taps the computer simulations are restricted to this speed (Ref. 11). The effect of bridged taps on the lower speed modems is estimated by appropriately scaling the effects experienced at 72 kbit/s.

The ideal eye diagram produced by a 100 % raised cosine Diphase signal is shown in Fig. 9(a). With this code there are two eyes per Baud interval. In some systems only one of these eyes is sampled to determine the transmitted data symbols b_n. However, in other systems both eyes are sampled and compared to provide continuous in-service error monitoring. In general, the intersymbol interference introduced by a bridged tap reduces the noise margin associated with each eye by different amounts as illustrated in Fig. 9(b). Consequently, depending on the amplitude probability distribution of the noise source, the error probabilities associated with the two samples can be significantly different.

As mentioned earlier, the intersymbol interference produced by a bridged tap depends on its gauge and length as well as on the line code of the transmission system. For an AMI signal the worst case occurs when the transmission delay up and back the bridged tap corresponds to one Baud interval. For the Diphase line code used in the DDS modems the worst interference occurs when the delay corresponds to half a Baud interval.

With pair cable the absolute delay is not strongly dependent on frequency and, depending on the cable type, typically has a value

between 5 and 10 microseconds per km (μs/km) in the frequency band of interest. For 0.64 mm PIUT cable the absolute delay is about 5.3 μs/km at 54 kHz (i.e. at the frequency where the PSD of a 72 kbit/s Diphase line signal peaks). Therefore for this type of cable the length of bridged tap which causes maximum interference is of the order of l_{crit} where,

$$l_{crit} \approx 0.65 \text{ km}$$

Although the absolute delay of pair cable is not strongly dependent on cable gauge, its attenuation is; hence, the heavier the gauge of the bridged tap cable the greater the level of interference it produces (i.e. the less loss experienced by the echo). The attenuation A_{crit} experienced by the echo signal as it travels up and back a bridged tap of critical length is

$$A_{crit} = 2 A_o l_{crit}$$

where A_o is the cable attenuation at an appropriate reference frequency f_o (e.g. the frequency where the PSD of the transmitted line signal peaks). For 0.64 mm PIUT cable

$$A_{crit} \approx 5.6 \text{ dB}$$

Since l_{crit} is inversely proportional to the Baud rate of the line signal and hence proportional to f_o^{-1}, and A_o (at f_o) is proportional to f_o^a (where $a \approx 0.5$), it is seen that A_{crit} is proportional to $f_o^{-½}$ (approximately). That is, the relative magnitudes of the echo components which produce intersymbol interference and hence a reduction in noise margin relative to the direct signal, increase as the transmission rate increases. Consequently, the higher the Baud rate of a baseband digital transmission system the more susceptible it is to the effects of a bridged tap.

For the DDS 72 kbit/s modem case, the computed reduction in noise margin, for bridged taps of three different cable gauges, is shown in Fig. 11. It shows the reduction in noise margin associated with both Diphase eyes for bridged taps located both in the centre and at the end of 5022 metre cable loops of 0.64 mm cable. The eye opening associated with the eye after the guaranteed zero crossing is denoted by E2, and the eye before the guaranteed zero crossing is denoted by E1. The results are essentially independent of the length of the main cable loop.

From Fig. 11 the variation in noise margin that can occur due to the position of the bridged tap is clearly seen. Typically a bridged tap at the end of a cable loop produces more eye closure than that by the same length bridged tap located some distance from either end. It is also seen that the relative noise margins associated with the two Diphase eyes can vary significantly. The peak reduction in noise margin occurs at a length of about 0.65 km irrespective of the gauge of the bridged tap, as predicted earlier. The peak reductions

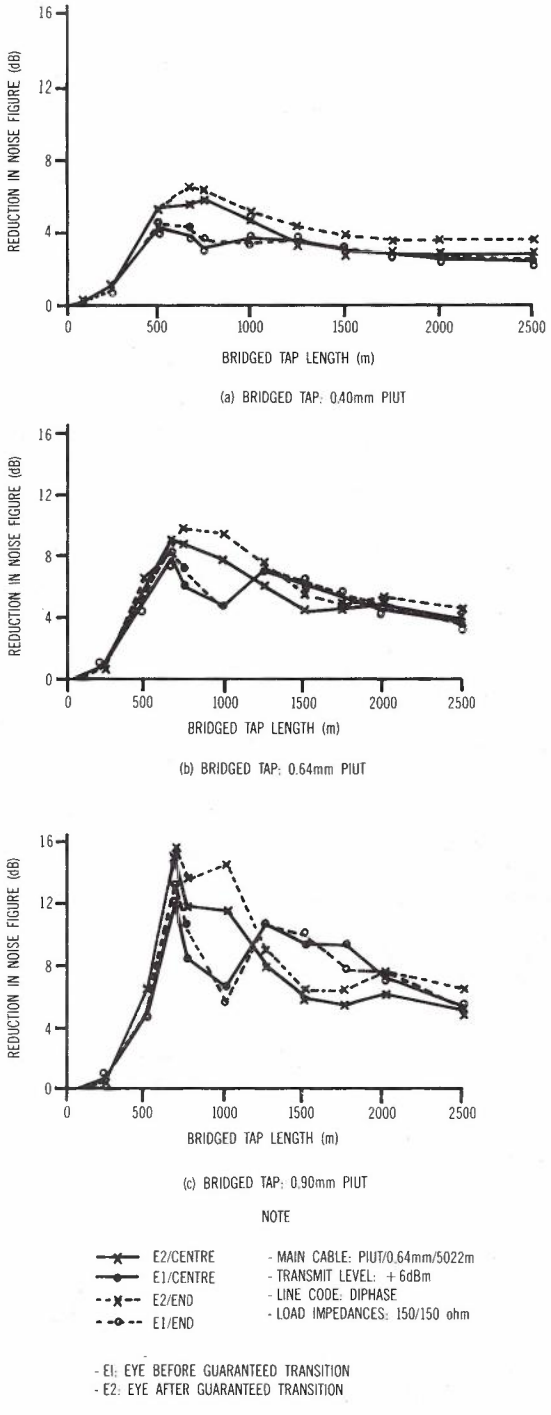


Fig. 11 - Computed Reduction in Noise Figure of TELETTRA 72 kbit/s Baseband Data Modem Due to Various Length Bridged Taps -- For Two Different Tap Positions and Several Gauges of Tap Cable

in noise margins were of the order of 7, 10 and 16 dB for the bridged tap cable gauges of 0.40, 0.64 and 0.90 mm respectively.

To illustrate the effect of line code, the simulated results for a 72 kbit/s AMI signal are shown in Fig. 12. It is seen that the peak reduction in noise margin occurs at about 1.3 km as predicted earlier. The peak reductions in noise margin were of the order of 6, 10 and 40 dB for the bridged tap cable

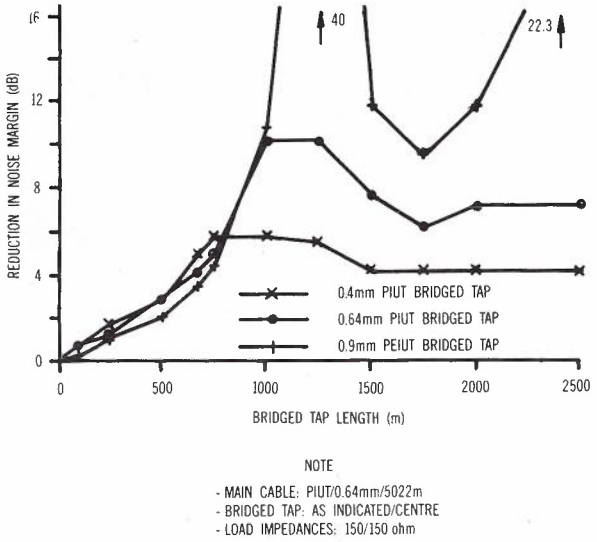


Fig. 12 - Computed Reduction in Noise Margin for a 72 kbit/s AMI Signal with 100% Raised Cosine Equalisation, for Various Length Bridged Taps

gauges of 0.40, 0.64 and 0.90 mm respectively. From the various cases considered it is seen that certain bridged tap conditions can cause a significant reduction in the noise margin of a 72 kbit/s modem. However, since these conditions do not occur very often in the network (e.g. from Fig. 1 it is seen that bridged taps with lengths greater than about 350 metres do not occur very often), bridged taps should not significantly restrict the application of 72 kbit/s DDS modems in the network.

The effect of bridged taps on the transmission performance of the lower speed modems is readily established by scaling the 72 kbit/s results. For example consider the data shown in Table 1. In this table the attenuation and delay characteristics of three different cable types are tabulated at a reference frequency appropriate to the different transmission rates. The corresponding length of bridged tap which would produce an echo with a half Baud period delay (i.e. l_{crit}) is also shown, together with the signal attenuation experienced by the echo signal as it travels up and back the bridged tap.

It is clearly seen from Table 1 that the signal attenuation (at the reference frequency) produced by bridged taps of critical length for the lower data rates is considerably higher (i.e. by about 8 to 16 dB) than that for 72 kbit/s operation. Consequently, the peak reduction in noise margin produced by critical length bridged taps for the lower speed modems will be considerably smaller. Furthermore, the critical bridged tap lengths for the lower data speeds are significantly longer, to the point where bridged taps of these lengths are unlikely to occur in the Telecom Australia network (see Fig.1). Consequently, bridged taps are not likely to have any perceptible effect on the operation of 3.2, 6.4 and 12.8 kbit/s DDS modems.

5.2 Basic ISDN Access

The baseband DDS data modem discussed above achieves bi-directional data transmission

TABLE 1 - Bridged Tap Cable Data

Cable Type	Data Rate (kBaud)	Ref. Freq. (kHz)	Cable Attenuation (dB/km) (A_o)	Cable Delay (μs/km)	L_{crit} (km)	A_{crit} (dB)
0.40 PIUT	72	54	8.9	6.1	0.57	10.15
	12.8	9.6	4.9	10.7	1.83	17.93
	6.4	4.8	3.6	14.9	2.62	18.86
	3.2	2.4	2.6	20.9	3.74	19.45
0.64 PIUT	72	54	4.3	5.34	0.65	5.57
	12.8	9.6	2.8	7.48	2.61	14.62
	6.4	4.8	2.1	9.9	3.95	16.59
	3.2	2.4	1.8	13.5	5.79	20.84
0.90 PEIUT	72	54	2.2	5.2	0.67	2.95
	12.8	9.6	1.56	6.0	3.26	10.17
	6.4	4.8	1.3	7.3	5.35	13.91
	3.2	2.4	1.1	9.5	8.22	18.08

by the use of two cable pairs; one pair is used for each transmission direction. This approach is acceptable for the DDS because of the relatively small number of DDS customers compared to the number of ordinary telephone subscribers. However, since the local cable distribution network is essentially dimensioned on the basis of one cable pair per customer, then to obtain widespread and economical access to an integrated services digital network (ISDN) of the future, a viable digital transmission technique which achieves bi-directional transmission on a single cable pair, at a transmission rate of about 160 kbit/s, is required.

The integrated customer access systems and the digital PABX extension systems mentioned in the introduction are not specifically investigated in this paper. However these systems use similar transmission techniques and system parameters to those being developed for ISDN basic access. Consequently many of the results and conclusions reached in this Section will also be generally applicable to them.

To achieve ISDN basic access two main transmission techniques have been proposed (Ref. 12); viz, the time compression multiplex (TCM) (also referred to as the burst or ping pong) and the echo cancellation (EC) techniques. The TCM technique achieves the bi-directional transmission by time sharing the cable pair between the two directions. This is achieved by increasing the transmission rate by a factor of about 2.5 or more. In this paper computer simulations for this type of system were performed for a rate of 432 kBaud. The rate expansion factor is chosen to allow for the propagation delay of the cable and provide a guard time between the Go and Return bursts. The backward echoes produced by a bridged tap place a constraint on the minimum acceptable guard time. However, this aspect is not considered in this paper.

Based on the results of the previous Section, the relatively high transmission rate of a TCM system is likely to make it susceptible to the effects of a bridged tap (Ref. 13). This is illustrated in Figs. 13 and 14 for Diphase and AMI line codes respectively. At 432 kBaud the critical bridged tap length is about 0.12 km and 0.23 km for the Diphase and AMI line codes respectively. For an AMI signal a bridged tap can cause complete eye closure. It is interesting to note that telecommunication

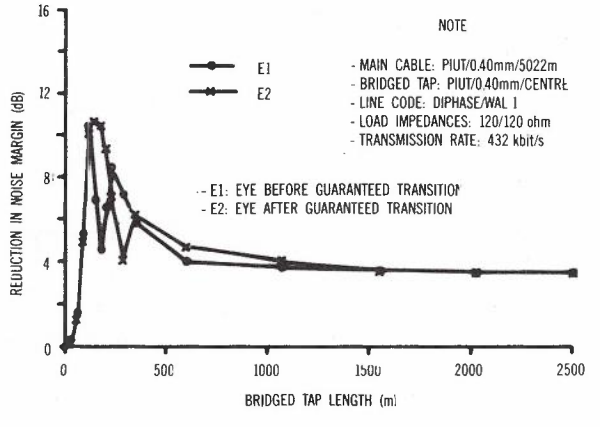


Fig. 13 - Computed Reduction in Noise Margin of 432 kbit/s Diphase (WAL1) Signal Due to Various Length Bridged Taps

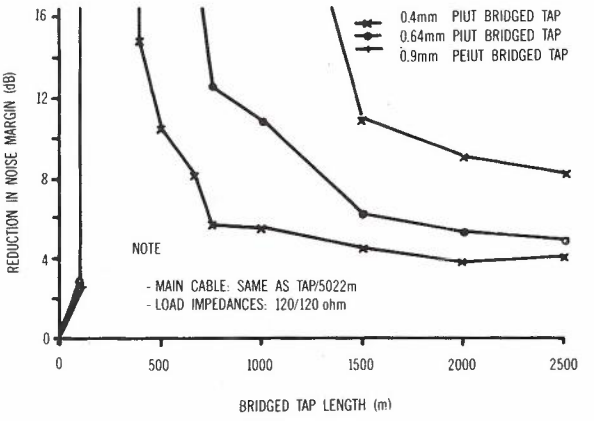


Fig. 14 - Computed Reduction in Noise Margin for a 432 kbit/s AMI Signal with 100% Raised Cosine Equalisation, for Various Length Bridged Taps

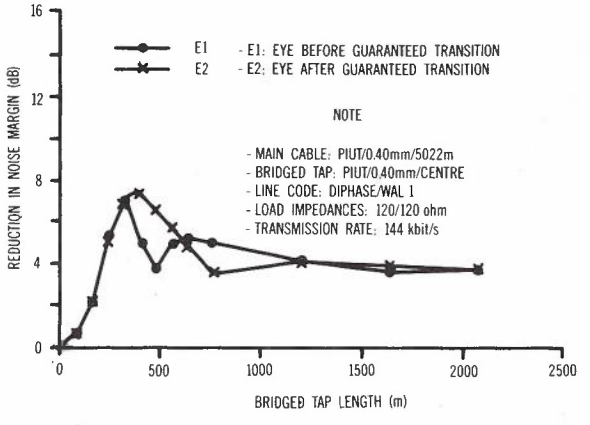


Fig. 15 - Computed Reduction in Noise Margin of 144 kbit/s Diphase (WAL1) Signal Due to Various Length Bridged Taps

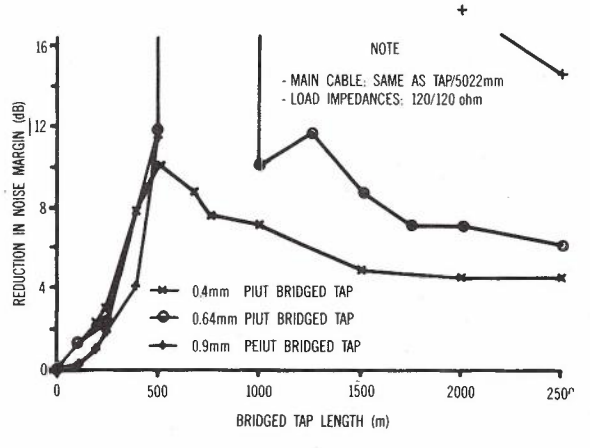


Fig. 16 - Computed Reduction in Noise Margin for a 144 kbit/s AMI Signal with 100% Raised Cosine Equalisation, for Various Length Bridged Taps

networks in many countries do not have bridged taps and consequently a number of the TCM systems which have been developed around the world do not include techniques to accommodate bridged taps (e.g. they do not use a DFE). These systems would therefore, in general, not be suitable for the Telecom Australia network. One TCM system development which does attempt to accommodate bridged taps uses a fixed compromise equalizer (Ref. 14) to minimize their effects. However this approach is not likely to be adequate under the worst case conditions illustrated above.

Unlike TCM systems, EC systems do not increase the transmission rate and therefore are less sensitive to bridged taps. The effect of bridged taps on these systems operating at 144 kbit/s is illustrated in Figs. 15 and 16 for Diphase and AMI line codes respectively.

Some two-wire digital customer access systems (Ref. 15) have been developed with no consideration of bridged taps. The results shown in Figs. 15 and 16 indicate that these systems would perform poorly under certain bridged tap conditions. However, most EC systems currently being developed, particularly those that employ narrowband line codes such as 4B3T, 3B2T and 2B1Q (Refs. 9,10,11) inherently require a DFE (typically with a large number of taps) to accommodate pulse dispersion produced by the cable loop. These systems can adequately accommodate (i.e. remove) the intersymbol interference produced by bridged taps, provided the DFE adaptation circuitry can accommodate the high levels of initial eye closure that may occur with some bridged tap conditions. However, the DFE cannot remove the signal attenuation produced by a bridged tap, which consequently produces some reduction in noise margin (typically less than 3.5 dB).

5.3 Primary Access

Customers that have high capacity data requirements or digital PABXs can be provided with four-wire, 2 Mbit/s access to their local exchange over the existing cable network. The digital line equipment normally associated with 30 channel PCM systems (PCM30) is used for this application. Although these systems use a HDB3 line code, their transmission performance is accurately modelled assuming an AMI line code.

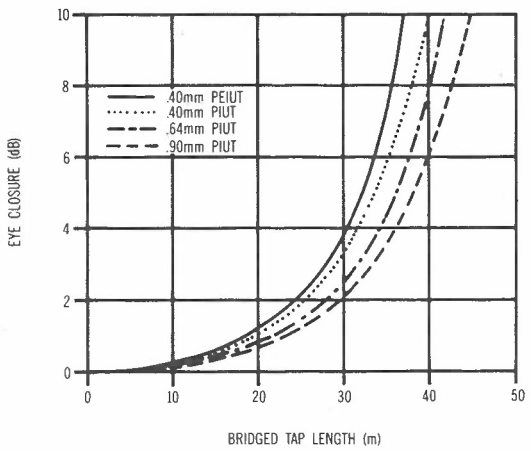


Fig. 17 - Eye Closure Due to Bridged Taps, for 2048 kbit/s and AMI Line Code

The susceptibility of 2 Mbit/s systems to a bridged tap is illustrated in Fig. 17 which displays the computed eye closure as a function of the length of a bridged tap for various cable types. It is seen that the critical bridged tap length is about 50 m. For a range of loop lengths around 50 m, complete eye closure occurs and to ensure a less than, say, 4 dB reduction in noise margin the length of a bridged tap would have to be less than 40 m or very much longer than 50 m (i.e. perhaps greater than 1 km.).

6. CONCLUSION

The effect of bridged taps on the operation of baseband digital transmission systems has been considered. The first part of the paper investigates the effect of a bridged tap on the frequency response of the cable loop alone. In the latter part of the paper, computer simulations are used to illustrate the effect of a bridged tap on particular transmission systems. The systems considered included baseband data modems operating at rates up to 72 kbit/s, the types of systems being developed for basic ISDN access and 2 Mbit/s primary access systems.

A reflected wave model has been developed to provide basic insight into the effects of bridged taps and the relative importance of the various parameters involved. It has been shown that a bridged tap modifies the cable loop response by introducing frequency and tap length dependent attenuation. As the frequency or length of the tap increases this attenuation asymptotes to a flat loss of 3.5 dB (under matched impedance conditions). At lower frequencies or shorter tap lengths, a cyclic ripple is introduced into both the phase and attenuation responses of the cable loop. The repetition interval of the ripple depends on the length of the bridged tap, with the first and most significant dip in the attenuation response occurring at the frequency at which the time taken for the signal to propagate along the bridged tap and return corresponds to a half period of the signal frequency. The magnitude of the envelope which bounds the extremes in the attenuation and phase ripples exponentially reduces with increasing frequency and is dependent on the gauge of the bridged tap; the heavier the cable gauge the larger the envelope of the ripple. The loop response has also been shown to be, in general, dependent on the terminating impedance conditions and on the position of the bridged tap.

An assessment of the effect of a bridged tap on the transmission performance of particular digital baseband transmission systems is based on computer simulations. The impairment introduced is quantified by the reduction in noise margin produced at the system decision point. The results of the study indicate that the higher the

transmission line rate the greater the potential interference produced by a bridged tap (i.e. the shorter the length of bridged tap required to cause maximum interference and the greater the reduction in noise margin caused by that maximum interference). The bridged tap length at which maximum interference occurs is also dependent on the line code used.

The operation of 72 kbit/s baseband modems can be significantly degraded by a bridged tap with a length, for typical cable types, around 650m; however, longer and shorter lengths cause less signal degradation. For example, very long bridged taps can cause a reduction in noise margin which asymptotes to about 3.5 dB, whereas bridged taps less than about 300m in length typically cause less than 2 dB reduction in noise margin. At lower baseband data modem speeds (e.g. 3.2, 6.4 and 12.8 kbit/s), bridged taps cause insignificant impairment. Since most bridged taps are less than 300m in length, the majority of cable pairs with bridged taps will still be usable for baseband data transmission at rates up to 72 kbit/s, provided acceptable levels of noise (i.e. crosstalk and impulsive noise) apply.

The transmission performance of ISDN basic access systems which incorporate a DFE in their design is not likely to be significantly affected by bridged taps provided the DFE adaptation algorithm can accommodate the levels of initial eye closure (including complete eye closure in some cases). However, where this is not the case, the operation of these systems on cable loops with bridged taps will be significantly restricted. In particular, bridged taps with a length which causes a transmission delay of around half a Baud interval for a WAL1 line code and one Baud interval for an AMI code can cause complete eye closure.

Burst mode systems are potentially more likely to be affected than echo cancellation systems because of their higher line rate. For example, a bridged tap length has to be greater than about 400 m to produce a 3.5 dB reduction in noise margin in a 144 kbit/s AMI echo cancellation system, whereas the same level of impairment can be produced in a burst mode system by a tap length of 100 m.

With 2 Mbit/s primary access systems the length of a bridged tap needs to be less than about 25 m if the reduction in noise margin is to be kept below 3 dB. For tap lengths greater than about 35 m, large reductions in noise margin occur, including complete eye closure when the transmission delay up and down the loop corresponds to a Baud interval.

7. REFERENCES

1. Keeler, R. J., "Construction and Evaluation of a Decision Feedback Equalizer", Conference Record, 1971 IEEE International Conference on Communications, Montreal, Canada, June 1971, pp. 8-13.

2. Demytko, N., "The Topology of the Australian Urban Subscriber Loop Network", Telecom Australia Research Laboratories Report 7551, June 1982.

3. Demytko, N. and Webster, R. I., "Extended Survey of the Australian Urban Subscriber Loop Network", Telecom Australia Research Laboratories Report 7692, September 1984.

4. Groenendaal, G. C., "The Effects of a Bridged Tap on Broadband Digital Transmission via Telephone Lines", 1978 International Seminar on Digital Communications, 7th - 9th March 1978, Zurich, Switzerland.

5. Lynch, J. K., "Transmission Characteristics of Symmetric Pair Cables at Carrier Frequencies", Telecom Australia External Plant Information Bulletin No. 60, June 1981.

6. Webster, R. I. and Semple, G. J., "The Effect of Bridged Taps on the Transmission Characteristics of Pair Cable", Telecom Australia Research Laboratories Report 7656, October 1883.

7. Suzuki, T., Takatori, H., Ogawa, M. and Tomooka, K., "Line Equalizer for a Digital Subscriber Loop Employing Switched Capacitor Technology", IEEE Transactions on Communications, Vol. COM-30, No. 9, September 1982.

8. Dierckx, D. and Taeymans, J. R., "System 12 ISDN Line Circuit", Electrical Communication (ITT) Vol. 59, No. 1/2, 1985.

9. ITT Digital Subscriber Loop Carrier Network Interface Specification, Contribution to American Standards Institute Telecommunications Committee (ANSI/T1), Document No. T1D1.3/84-040, November 1984.

10. Adams, P. F., Cox. S. A. and Glen, P. J., "Long Reach Duplex Transmission Systems for ISDN Access", British Telecom Technology Journal, Vol. 2, No. 2, April 1984.

11. Semple, G. J., "The Effect of Bridged Taps on the Transmission Performance of the Baseband Digital Data Modems Used in the Digital Data Network", Telecom Australia Research Laboratories Report 7590, December 1983.

12. Ahamed, S. V., Bohn, P. P. and Gottfried, N. L., "A Tutorial on Two-Wire Digital Transmission in the Loop Plant", IEEE Transactions on Communications, Vol. COM-29, No. 11, November 1981.

13. Semple, G. J., "The Effect of Bridged Taps on the Transmission Performance of Local Digital Reticulation Systems", Telecom Australia Research Laboratories Report 7704, March 1985.

14. Ahamed, S. V., "Simulation and Design Studies of Digital Subscriber Lines", The Bell System Technical Journal (BSTJ), Vol. 61, NO. 6, July-August 1982.

15. Holte, N. and Stueflotten, S., "A New Digital Echo Canceler for Two-Wire Subscriber Lines", IEEE Transactions on Communications, Vol. COM-29, No. 11, November 1981.

BIOGRAPHY

JOHN SEMPLE received the Degrees of Bachelor of Engineering (Electrical) with honours and a Master of Engineering Science from the University of Melbourne in 1966 and 1968 respectively. He commenced work as an Engineer with the then APO Research Laboratories, Pulse Systems Division in January 1967, where he was concerned with problems associated with primary level PCM and high capacity digital transmission systems. In 1972 he joined the Line and Data Systems Section, which is now part of the Transmission Systems Branch of the Telecom Research Laboratories, where he continued work associated with the application of PCM transmission in the inter-exchange cable network. In about 1980 Mr. Semple commenced the investigation of local digital reticulation systems, and is currently concerned with Integrated Services Digital Network (ISDN) developments and the digital two-wire full duplex transmission systems required to provide customer access to such a network.

Convergence of Memory Compensation Principle Decision-Feedback Equalisers with Decision Errors

A.J. JENNINGS

Telecom Australia Research Laboratories

This paper examines the convergence behaviour of decision-feedback equalisers employing the memory compensation principle in the subscriber loop environment. In this situation the line dispersion is typically so severe that the decision-feedback equaliser must operate initially with a high proportion of decision errors. Whilst this is a practical problem encountered in the operation of ISDN digital transmission equipment it has not received any attention in the literature. To understand the behaviour of the decision-feedback equaliser in this situation this paper employs a random walk analysis in the situation of incorrect decisions. Despite the numerous assumptions necessary to make the analysis tractable the analysis compares favourably with the simulated behaviour in the presence of incorrect decisions. The analysis indicates the critical importance of the adaption gain in determining the convergence properties of a memory compensation principle decision-feedback equaliser.

KEYWORDS: ISDN, Decision-feedback Equalisation, Adaptation, Stochastic Models

1. INTRODUCTION

To achieve customer access to an integrated services digital network (ISDN), full duplex digital transmission between the customer and the local exchange must be achieved. For economic reasons it is preferred that this transmission is over existing customer lines at 144 kbit/sec. The most promising approach to achieving full duplex transmission appears to be the echo-cancellation technique (Refs. 1,2,8). This paper deals with one aspect of the use of combined echo-cancellation and decision-feedback equalisation (DFE) in this application -- convergence of the decision-feedback equaliser in the presence of decision errors.

Figure 1 shows a typical arrangement for full-duplex operation over subscriber lines. The transmitted line signal is coupled imperfectly to the two-wire transmission line - a certain proportion of the transmitted signal leaks into the receiver - the echo signal. By forming a replica of the echo path the echo signal can be imitated and subtracted from the received line signal. Intersymbol interference introduced in transmission over the subscriber line is removed by the DFE.

The EC and DFE must both adapt to the particular characteristics of an individual subscriber line. However there is a fundamental difference between the adaption processes in the two cases. In the case of the echo-cancellor the transmitted data are known with certainty whereas in the case of the DFE the received data must be estimated. When there is severe intersymbol interference these estimates will be highly unreliable.

Section 2 introduces the method of operation of a memory compensation principle DFE. The basic principle is that the DFE stores all possible intersymbol interference contributions in different memory addresses. When a particular sequence of symbols is received this sequence is used as an address to locate the relevant intersymbol interference contribution. In this way the effects of intersymbol interference can be removed. However, the memory must adapt to different line conditions, and this adaption is problematic when the estimates of transmitted symbols are in error.

Section 3 discusses the process of adaption in the presence of decision errors. Since the behaviour in this region is complicated, some simplifying assumptions are introduced in order to make the problem tractable. In some cases these assumptions only seem reasonable when taken in the light of the subsequent analysis.

Given these assumptions we can model the adaption of the DFE memory contents as a random walk in the region of incorrect decisions. Essentially the DFE memory contents wander at random until they form into a configuration that allows correct decisions. The analysis of Sections 4-6 gives the statistics of the time it takes to wander into the region of correct decisions. Given the complexity of the problem it is not surprising that extensive use is made of the assumptions introduced in Section 3.

Paper received 28 June 1985.
Final revision 5 July 1985.

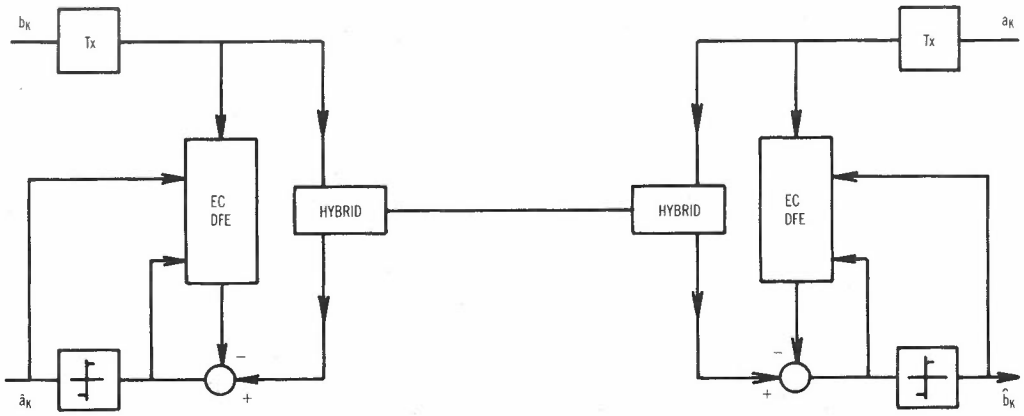


Fig. 1 - Echo-Cancellation and Decision Feedback Equalisation for Full-Duplex Data Transmission Over Subscriber Lines

2. MEMORY-COMPENSATION PRINCIPLE DFE

The structure of Figure 2 allows for simultaneous correction for the echo signal and intersymbol interference by the memory compensation principle (Ref. 3). This principle uses the fact that the echo-signal and intersymbol interference can be considered as a function of a finite number of transmitted or received symbols. By adaptively determining each possible echo signal and storing them in a memory these can be recalled as a function of the transmitted symbols to cancel the echo signal at the receiver.

To consider the operation of the unit described in Figure 2, perfect echo-cancellation will be assumed. As discussed in Ref. 4, it can be arranged that the EC RAM is adapted first to convergence, and only then will the DFE RAM be enabled. If at this stage there is excessive intersymbol interference then the effect of decision errors on the adaption of the DFE RAM must be considered. It seems reasonable then to analyse the operation of the DFE RAM in isolation.

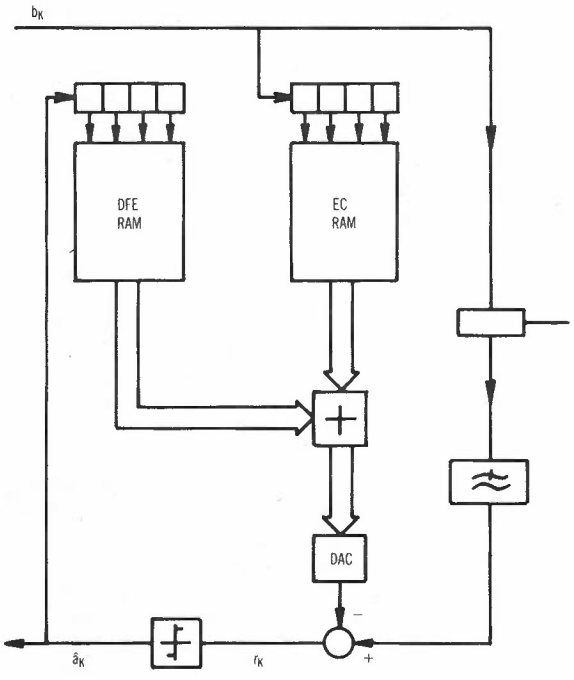


Fig. 2 - Memory Compensation Principle Combined Echo-Cancellation and Decision-Feedback Equalisation

The received signal after compensation then becomes

$$r_k = a_k h_0 + \underline{a}_k^T . \underline{h} - d_{q(\underline{\hat{a}}k)} \qquad (1)$$

where the terms in this equation are:

a_k transmitted symbol $\epsilon(-1,1)$

$\underline{h}^T = [h_1\ h_2\ \ldots\ h_N]$ channel impulse response

$\underline{a}_k^T = [a_{k-1}\ a_{k-2}\ \ldots\ a_{k-N}]$

$\underline{\hat{a}}_k^T = [\hat{a}_{k-1}\ \hat{a}_{k-2}\ \ldots\ \hat{a}_{k-N}]$

$q(\underline{\hat{a}}_k)$ RAM address

d_q RAM contents

and the decision process can be described

$$\hat{a}_k = \text{sign}\ (r_k) \qquad (2)$$

Here the notation $\hat{a}_k$ denotes an estimate of a_k. Throughout this paper the symbol q will always denote $q(\underline{\hat{a}}_k)$, the memory address formed from the estimate of the last N transmitted symbols.

The effect of excessive intersymbol interference is to cause decision errors in (2), which in turn cause errors in the RAM address selection.

In order to adapt to different line conditions the DFE RAM contents are adapted using the sign algorithm. If the contents of the DFE RAM are

$$\underline{d}(k) = (d_1(k)\ d_2(k)\ \ldots\ d_{2N}(k)) \qquad (3)$$

then at time k only the contents of memory at address $q(\underline{\hat{a}}_k)$ are altered.

$$d_i(k+1) = d_i(k), \quad i = 1, \ldots 2^N, \; i \neq q$$

$$= d_i(k) + \Delta d_i(k) \;, \; i = q \qquad (4)$$

where the notation of Ref. 2 is adopted for convenience. The sign algorithm dictates that

$$d_q(k+1) = d_q(k) + \gamma \; \text{sign} \; (e_k) \qquad (5)$$

where

$$e_k = r_k - \hat{a}_k \, d_o \qquad (6)$$

is the error signal and d_o is the adaptive reference signal. If no adaptive reference is present then this term can be neglected. It will be seen that it plays little part in the process of adaption of the DFE RAM contents in the region of incorrect decisions. The adaptive reference is updated by

$$d_o(k+1) = d_o(k) + \gamma \; \text{sign}(e_k) \; \hat{a}_k \qquad (7)$$

When decisions are correct the analysis of adaption can be inferred from standard results (Ref. 3), and it will not be pursued here. A monotonic decrease of the mean-square error over the ensemble $\{a_i \; ; \; i = 1, \ldots \infty\}$ of possible data sequences $E(e_k^2)$ can be readily demonstrated. In the region of incorrect decisions the behaviour of the DFE RAM contents is unknown, and is explored in this paper.

3. DECISION ERRORS

The equation (5) can be rewritten as

$$\Delta d_{q(\underline{\hat{a}}_k)} = \gamma \; \text{sign}(a_k h_o + \underline{a}_k^T . \underline{h} - d_{q(\underline{\hat{a}}_k)} - \hat{a}_k d_o) \qquad (8)$$

and equation (7) as

$$\Delta d_o(k) = \gamma \; \text{sign}(a_k h_o + \underline{a}_k^T . \underline{h} - d_{q(\underline{\hat{a}}_k)} - \hat{a}_k d_o) \; \hat{a}_k \qquad (9)$$

Considering the N-dimensional Euclidean space of possible RAM contents, there will be a region where the RAM contents are such that the only decision-errors are due to random noise. If we call this region of correct decisions region 2 and the region of incorrect decisions region 1, then the process of convergence consists of random motion in region 1 leading to capture in region 2. This process is illustrated in Figure 3 in the two-dimensional case. Since the motion in region 1 is random in nature there is no guarantee of passage to region 2, only a probability of passage to region 2.

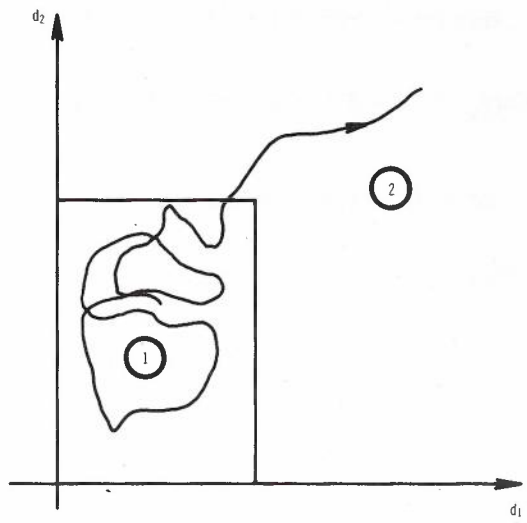


Fig. 3 - Two Regions of Convergence

1. Decision Errors
2. Decisions Correct

In order to analyse the process of adaption in region 1, some simplifying assumptions are necessary. The assumptions are related to those for the transversal filter structure DFE, presented in Ref. 5. For convenience it is assumed that the DFE RAM contents are initialised to zero at the commencement of adaption.

Assumption 1. Symbol decisions in region 1 are not influenced by the contents of the DFE RAM.

$$\hat{a}_k = \text{sign} \; (a_k h_o + \underline{a}_k^T . \underline{h} - d_{q(\underline{\hat{a}}_k)}) \qquad (10)$$

So according to Assumption 1,

$$\hat{a}_k = \text{sign} \; (a_k h_o + \underline{a}_k^T . \underline{h}) \qquad (11)$$

and $\hat{a}_k$ is not influenced by the DFE RAM contents. This assumption rests on the observation that

$$| \; \underline{a}_k^T . \underline{h} \; | > h_o$$

when decisions are mostly incorrect for incorrect decisions in region 1, and that in region 1 the d values are smaller than the intersymbol interference contribution $\underline{a}_k^T . h$.

Assumption 2. The process $a_k h_o + \underline{a}_k^T . \underline{h}$ is Gaussian.

Although this is only asymptotically correct as the number of contributors to $\underline{a}_k^T . \underline{h}$ grow in number, it is a useful approximation for small numbers of contributors. The results to be presented support the use of this approximation.

Assumption 3. When an adaptive reference is used in region 1 it does not influence significantly the adaption of the DFE RAM contents.

Recalling from (8) that

$$\Delta d_{q(\hat{\underline{a}}_k)} = \gamma \, \text{sign}(a_k h_o + \underline{a}_k^T.\underline{h} - d_{q(\hat{\underline{a}}_k)} - \hat{a}_k d_o) \quad (13)$$

Then this assumption implies that

$$\Delta d_{q(\hat{\underline{a}}_k)} \simeq \gamma \, \text{sign}(a_k h_o + \underline{a}_k^T.\underline{h} - d_{q(\hat{\underline{a}}_k)}) \quad (14)$$

and d_o does not influence the adaption of the DFE RAM contents. Recalling (9),

$$\Delta d_o(k) = \gamma \, \text{sign}(a_k h_o + \underline{a}_k^T.\underline{h} - d_{q(\hat{\underline{a}}_k)} - \hat{a}_k d_o) \, \hat{a}_k$$

and it will be seen that Assumption 3 implies that d_0 remains small in comparison with $(d_q, q = 1, \ldots 2^N)$ in region 1. The following results are derived using the method to be outlined in the next section. If

$$a_k h_o + \underline{a}_k^T.\underline{h} \quad : \quad N(0, \| \underline{h} \|^2)$$

$$d : N\left(0, \frac{\|\underline{d}\|^2}{2^N}\right) \quad (15)$$

where

$$\| \underline{v} \| = \sqrt{v_1^2 + v_2^2 + \ldots + v_N^2}$$

are both normal processes with zero mean, then

$$E(d_o) = 0$$

$$\text{var}(d_o) \simeq \frac{\gamma}{2}\sqrt{\frac{\pi}{2}\left[\|\underline{h}\|^2 + \frac{\|\underline{d}\|^2}{2^N}\right]} \quad (16)$$

and this reinforces Assumption 3. We shall see that this is the case. The derivation of equation (16) follows similar lines to the following analysis of the statistics of $(d_q ; q = 1, \ldots 2^N)$ in region 1.

4. INCREMENTAL STATISTICS

We now proceed to examine the statistics of the RAM contents in region 1 to attempt to define the probability of passage to region 2 as the adaption progresses. Equation (13) describes the incremental adjustment of each of the RAM contents. There are two possibilities.

(1) $\hat{\underline{a}}_k = \underline{a}_k \Rightarrow q(\hat{\underline{a}}_k) = q(\underline{a}_k)$

(2) $\hat{\underline{a}}_k \neq \underline{a}_k \Rightarrow q(\hat{\underline{a}}_k) \neq q(\underline{a}_k)$

The RAM address may be correct or incorrect, depending on the past N symbol decisions. If

$$p(\hat{a}_k = a_k) = \xi \quad (17)$$

then using Assumption 1, ξ can be evaluated by considering all possible transmitted symbol sequences. For a particular sequence, if

$$\text{sign } \underline{a}_k^T.\underline{h} = \text{sign } a_k h_o$$

then the decision (11) will be correct. However if the signs are opposing then

$$| \underline{a}_k^T.\underline{h} | < | a_k h_o |$$

is necessary for the decision to be correct.

Then

$$p(\hat{\underline{a}}_k = \underline{a}_k) = \xi^N \quad (18)$$

using Assumption 1. If an address has been selected correctly then (14) implies that adaption is almost always in the correct direction, i.e.

$$p(\text{sign } \Delta d_q = \text{sign } (\hat{\underline{a}}_k^T.\underline{h} - d_q) \mid \hat{\underline{a}}_k = \underline{a}_k) \simeq 1 \quad (19)$$

The term $\text{sign}(\hat{\underline{a}}_k^T.\underline{h} - d_q)$, where $q = q(\hat{\underline{a}}_k)$, is the correct direction for d_q to further reduce intersymbol interference when location q is selected.

From (14) we can see that (19) implies that the intersymbol interference dominates the received signal, an approximation that will be more exact as the proportion of incorrect decisions increases.

However if an address is selected incorrectly then (14) implies that

$$\Delta d_q \quad : \quad N(-d_q, \| \underline{h} \|^2) \quad (20)$$

is a normal random variable with mean $-d_q$ and variance $\| \underline{h} \|^2$. Excluding the case where $\hat{\underline{a}}_k = \underline{a}_k$ then the probability of Δd_q being in the correct direction (over the ensemble of cases where $\hat{\underline{a}}_k = \underline{a}_k$) is derived in Appendix 1. The probability of a correction in the correct direction for the case where this correction is positive (the negative case is similar) is then

$$p(\text{sign } \Delta d_q = \text{sign } (\hat{a}_k^T.\underline{h} - d_q) \mid \hat{\underline{a}}_k \neq \underline{a}_k)$$

$$= (1 - P(d_q/\| \underline{h} \|) - 1/2^N)(2^N/(2^N-1)) \quad (21)$$

where

$$P(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}$$

Expanding P(x) in a Taylor series about the origin and retaining only the first term we find

$$P(x) \simeq 0.5 + \frac{\partial P(x)}{\partial x}\Big|_{x=0} .x$$

so

$$P\left(\frac{d_q}{||\underline{h}||}\right) \simeq 0.5 + \frac{d_q}{\sqrt{2\pi}||\underline{h}||} \qquad (22)$$

and Eq. (21) becomes

$$p(\text{sign } \Delta d_q = \text{sign}(\underline{\hat{a}}_k{}^T.\underline{h} - d_q) \mid \underline{\hat{a}}_k \neq \underline{a}_k)$$

$$\simeq \left(\frac{-1}{2^N} + 0.5 - \frac{d_q}{\sqrt{2\pi}\,||h||}\right)\left(\frac{2^N}{2^N-1}\right) \qquad (23)$$

So combining (19) and (23),

$$p\ (\text{sign } \Delta d_q = \text{sign}\ (\underline{\hat{a}}_k{}^T.\underline{h} - d_q)) =$$

$$\xi^N + \frac{2^N}{2^N-1}\left[(1-\xi^N)\left(0.5 - \frac{1}{2^N}\right) - \frac{d_q(1-\xi^N)}{\sqrt{2\pi}||h||}\right] \qquad (24)$$

If we now recall that d_q can only proceed in discrete steps, i.e.

$$d_q = j\gamma$$

then

$$p_{j,j+1} = \alpha - j\beta$$

$$p_{j,j-1} = (1 - \alpha) + j\beta \qquad (25)$$

are the discrete transition probabilities given that the correct direction is positive, where

$$\alpha = \xi^N + (1-\xi^N)(0.5-1/2^N)(2^N/(2^N-1))$$

$$\beta = \gamma(1-\xi^N)\,\frac{1}{\sqrt{2\pi}}\,\frac{1}{||\underline{h}||}\,\frac{2^N}{2^N-1} \qquad (26)$$

Recall that ξ is the proportion of correct decisions and can be evaluated using Assumption 1 and simply enumerating the 2^{N+1} possible cases of r_k.

5. DISCRETE EVOLUTION EQUATION

From equation (26) we have the incremental statistics of a particular RAM address when that address is selected. From this we will obtain the evolutionary statistics of the entire RAM contents as a function of time in region 1.

Eq. (26) described the incremental statistics of the RAM contents when that address is selected. From Assumption 1 each address is equally likely and hence

$$\begin{aligned} p_{j,j+1} &= (\alpha - j\beta)\ \rho \\ p_{j,j} &= 1 - \rho \\ p_{j,j-1} &= ((1 - \alpha) + j\beta)\ \rho \\ \rho &= 1/2^N \end{aligned} \qquad (27)$$

are the transition probabilities of any particular address contents. Under some conditions some addresses may be more likely than others, but the assumption that they are equally likely holds well for the problems we consider. This issue is discussed further in Section 6. If p_k^{n+1} denotes the probability of state k at adaption step n+1, then

$$p_k^{n+1} = \rho(\alpha-(k-1)\beta)\ p_{k-1}^n + (1-\rho)\ p_k^n + \rho((1 - \alpha) + (k + 1)\beta)\ p_{k+1}^n \qquad (28)$$

is the evolution equation. If

$$\sum_{k=-\infty}^{\infty} p_k^{n+1}\ e^{-k\theta} = M_{n+1}(\theta)$$

is the moment generating function, then (28) can be written (Ref. 6).

$$\sum_{k=-\infty}^{\infty} p_k^{n+1}\ e^{-k\theta} = \sum_{k=-\infty}^{\infty} \rho(\alpha-(k-1)\beta)\ p_{k-1}^n\ e^{-k\theta} + \sum_{k=-\infty}^{\infty} (1-\rho)\ p_k^n\ e^{-k\theta} + \sum_{k=-\infty}^{\infty} \rho((1-\alpha)+(k+1)\beta)\ p_{k+1}^n\ e^{-k\theta} \qquad (29)$$

$$\Leftrightarrow M_{n+1}(\theta) = \rho\alpha e^{-\theta} M_n(\theta) + \rho\beta e^{-\theta} M_n'(\theta) + (1-\rho) M_n(\theta) + \rho(1-\alpha) e^{\theta} M_n(\theta) - \rho\beta e^{\theta} M_n'(\theta)$$

Here $M_n'(\theta)$ denotes the first derivative of $M_n(\theta)$.

$$\Leftrightarrow M_{n+1}(\theta) = (\rho\alpha e^{-\theta} + \rho(1-\alpha)e^{\theta} + (1-\rho)) M_n(\theta) + (\rho\beta e^{-\theta} - \rho\beta e^{\theta}) M_n'(\theta)$$

Differentiating we obtain

$$M_{n+1}'(\theta) = (\rho e^{\theta} - 2\rho\alpha \cosh\theta) M_n(\theta) + [(1-\rho) - 2\rho\alpha \sinh\theta + \rho e^{\theta} - 2\rho\beta \cosh\theta] M_n'(\theta) - 2\rho\beta \sinh\theta M_n''(\theta) \quad (31)$$

and since $M_{n+1}'(0) = -\mu_{n+1}$, where μ_{n+1} is the mean of the process d_i, we obtain

$$\mu_{n+1} = \mu_n (1 - 2\rho\beta) + \rho(2\alpha - 1)$$

$$\Leftrightarrow \mu_n = \frac{2\alpha - 1}{2\beta} [1 - (1 - 2\rho\beta)^n] \quad (32)$$

Differentiating (31) we obtain

$$M_{n+1}''(\theta) = (\rho e^{\theta} - 2\rho\alpha \sinh\theta) M_n(\theta) + (2\rho e^{\theta} - 4\rho\alpha \cosh\theta - 2\rho\beta \sinh\theta) M_n'(\theta) + ((1-\rho) + \rho e^{\theta} - 2\rho\alpha \sinh\theta - 4\rho\beta \cosh\theta) M_n''(\theta) - 2\rho\beta \sinh\theta M_n'''(\theta) \quad (33)$$

and since $M_{n+1}''(0) = \mu_{n+1}^2 + \sigma_{n+1}^2$ we can use (32) to obtain (See Appendix II for derivation of this result):

$$\sigma_n^2 = (1/4\beta)(1 - (1-4\rho\beta)^n) + \left(\frac{2\alpha - 1}{2\beta}\right)^2 [(1-4\rho\beta)^n - (1-2\rho\beta)^{2n}] \quad (34)$$

Here the second term is small except for n very small.

Equations (32) and (34) can now be used to give a complete description of the process $\underline{d}$ in region 1.

6. CAPTURE ANALYSIS

The process $\underline{d}$ in region 1 is of the form

$$d_i : N(\mu_n, \sigma_n^2) \qquad i = 1, \ldots 2^N \quad (35)$$

and (32), (34) show that each d_i is asymptotically stationary as $n \to \infty$. From eq. (1) there is a region where correct decisions are made when the correct RAM address is selected.

$$\hat{a}_k = \text{sign}(a_k h_o + \underline{a}_k^T.\underline{h} - d_{q(\hat{\underline{a}}_k)})$$

$$\Rightarrow | \underline{a}_k^T.\underline{h} - d_{q(\underline{a}_k)} | < | a_k h_o | \quad (36)$$

and the edge of region 2 can be defined:

$$d_i > \eta_i \quad , \quad (\eta_i > 0)$$
$$d_i < \eta_i \quad , \quad (\eta_i < 0)$$
$$i = 1, \ldots 2^N \quad (37)$$

where

$$\eta_i = \min d_i : | \underline{a}_k^T.\underline{h} - d_i | < | a_k h_o |$$

when

$$\hat{\underline{a}}_k = \underline{a}_k \text{ and } i = q(\underline{a}_k).$$

For capture in region 2, all of the d_i must simultaneously be present in region 2. In addition, the first address selection in region 2 must be correct so if P_i is the probability that d_i is in region 2 then

$$P_i = 1 - P\left(\frac{\eta_i - \mu_n}{\sigma_n}\right) \quad (38)$$

and the probability of capture in region 2 is

$$P_c = \prod_{i=1}^{2^N} P_i \quad (39)$$

so the probability of remaining in region 1 is

$$P_f = 1 - \prod_{i=1}^{2^N} P_i \quad (40)$$

The probability that passage to region 2 takes longer than n symbols is then

$$p\,(t_c > n) = 1 - p_f^{\,n} \qquad (41)$$

where it has been assumed that each step of the adaption is independent of any other step. This precludes a 'staged' passage from region 1 to region 2 where a number of RAM addresses go to 2 and then some time passes before the remainder pass to 2. The argument against this possibility rests on the realisation that $\hat{a}_k$ is itself a random process and that unless the entire RAM is in region 2 there will be corruption of the memory contents. So in the 'staged' passage the RAM addresses that pass to 2 will be corrupted and the total passage will fail.

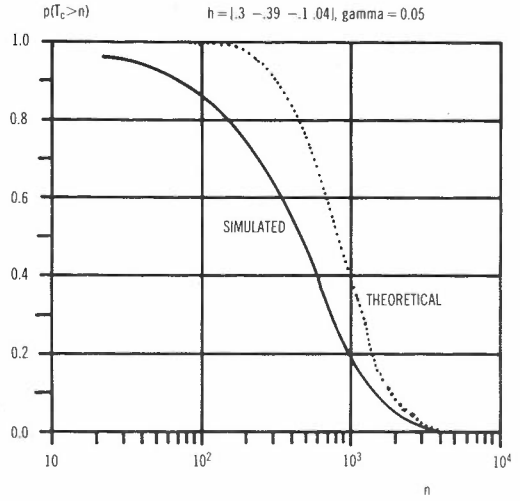


Fig. 4 - Comparison between Theory and Experiment.

Figure 4 shows a comparison between the distribution of time to capture obtained from simulation with that obtained from the above results. The simulations were performed by simulation of equations (1) - (7) with a pseudorandom data source a_k. In simulation, passage to region 2 was detected by the occurrence of 256 consecutive correct decisions. Reasonable agreement between the theory and experiments is obtained. However there is a residual difference between theory and experiment, and this can perhaps be explained by the choice of capture criterion used in the two cases. It is possible that the DFE RAM can pass to region 2 by fortuitous occurrence of data sequences such that even though decision errors could occur, the sign of the intersymbol interference is of the same sign as the data. In our theoretical analysis we require that the contents of the DFE RAM grow to a value where decision errors are impossible. So we would expect that the theory is conservative for small n but asymptotically correct as n is made larger, and this is the case.

7. DESIGN CONSIDERATIONS

The results of the previous section indicate that in region 1 the individual RAM contents are independent, asymptotically stationary and normally distributed. Eqs. (32) - (41) give a method of evaluation of the distribution of time to capture given a particular channel impulse response.

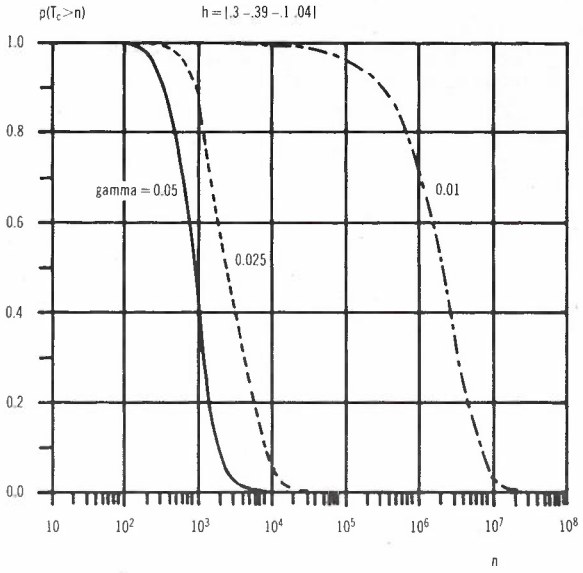


Fig. 5 - Variation in Time to Capture Vs. Adaption Gain

Figure 5 shows the variation in the time to capture versus adaption gain. It is clear that small gains can lead to very long times to capture. This can be illustrated from the previous results by considering the case where a single RAM dominates the convergence. Then (Ref. 7):

$$y = \frac{n_i - \mu_n}{\sigma_n}$$

$$p_f \simeq 1 - \frac{1}{y\sqrt{2\pi}}\, e^{-y^2/2} \qquad (42)$$

and taking the time to 99% probability of capture,

$$\Rightarrow 1 - p_f^{\,n} = 0.01$$

$$\Leftrightarrow n \simeq 0.01\, y\sqrt{2\pi}\, e^{y^2/2} \qquad (43)$$

and the asymptotic behaviour versus the adaption gain is

$$n \simeq \frac{k_0}{\sqrt{\gamma}}\, e^{k_1/\gamma} \qquad (44)$$

which is consistent with the observed strong dependence on the adaption gain.

8. CONCLUSION

The preceding analysis of memory compensation principle DFE's shows that passage to a situation of correct decisions (i.e. progress into region 2) is almost sure. If it is assumed that once this situation is obtained that it persists, then we can state that convergence of the memory compensation principle DFE is almost sure. However the dynamic behaviour in the region of incorrect decisions is strongly dependent on the adaption gain. A design equation for memory-compensation principle DFE's has been presented, allowing the selection of gain settings in region 1 to meet required performance objectives.

9. ACKNOWLEDGEMENTS

The exact derivation of eq. (34) presented in Appendix II is due to Dr. P. Potter.

10. REFERENCES

1. P.F. Adams, S.A. Cox and P.J. Glen, "Long reach full duplex transmission systems for ISDN access", British Telecom Technology Journal, Vol. 2, No. 2, pp. 35-42, April 1984.

2. P.J. VanGerwen, N.A.M. Verhoeckx and T.A.C.M. Claasen, "Design considerations for a 144 kbit/s digital transmission unit for the local telephone network", IEEE Journal on Selected Areas in Communications, Vol. SAC-2, No. 2, pp. 314-323, March 1984.

3. N. Holte and S. Stueflotten, "A new digital echo canceller for two- wire subscriber lines", IEEE Trans. Commun. Vol. COM-29, pp. 1573- 1581, November 1981.

4. T.A.C.M. Claasen and N.A.M. Verhoeckx, "Convergence slow-down of joint adaptive Filters : interaction of decision-feedback equaliser (DFE) and echo canceller (EC)", Poster presented at l'Aquila Workshop on Digital Signal Processing, l'Aquila, Italy, September 6-8, 1983 (available from authors on request).

5. A. Jennings, "Convergence of decision-feedback equalisers with incorrect decisions", submitted to IEEE Trans. Communications.

6. D.R. Cox and H.P. Miller, "The theory of stochastic processes", Chapman-Hall, London, 1965.

7. M. Abromowitz and I.E. Stegun, "Handbook of mathematical functions", Dover, N.Y., 1970.

8. N. Demytko, B.M. Smith et al, "Transmission considerations for ISDN basic access systems", ATR, this issue.

APPENDIX I

From eq. (13) for a particular q, ($q = q(\underline{\hat{a}}_k)$), and utilising Assumption 2 over the ensemble of possible $\underline{a}_k$,

$$p(\Delta d_q > 0) = 1 - P(d_q / ||\underline{h}||) \qquad (AI.1)$$

However amongst these possibilities there is one where $\underline{\hat{a}}_k = \underline{a}_k$ and for this case the correction is always positive. Eliminating this case reduces the total number of possibilities by 1 so

$$p(\Delta d_q > 0) = ((1 - P(d_q/||\underline{h}||)) - 1/2^N)(2^N/(2^N-1)) \qquad (AI.2)$$

where the correction for the reduction of the ensemble size is included.

APPENDIX II

The iteration (33) when $\theta = 0$ becomes

$$M''_{n+1}(0) = M''_n(0)(1-4\rho\beta) + M'_n(0)2\rho(1-2\alpha) + \rho \qquad (AII.1)$$

and

$$M'_n(0) = -\mu_n = \frac{(2\alpha-1)}{2\beta}[1 - (1-2\rho\beta)^n] \qquad (AII.2)$$

So the recursion can be rewritten with $X_n^2 = M''_n(0)$ and

$$X^2_{n+1} = (1-4\rho\beta)X^2_n + (2\rho-1)^2(\rho/\beta)[1 - (1-2\rho\beta)^n] + \rho \qquad (AII.3)$$

Let

$$A = 1 - 4\rho$$

$$B = \rho$$

$$C = \left(\frac{2\alpha-1}{\beta}\right)^2 \rho$$

$$D = 1 - 2\rho\beta$$

then

$$X^2_{n+1} = AX^2_n + B + C(1-D^n)$$

and

$$x_n^2 = B\,(1 + A + A^2 + \ldots + A^{n-1}$$

$$+ \sum_{i=1}^{n-1} C(1-D^i)A^{n-1-i} \qquad \text{(AII.4)}$$

$$= B\,\frac{1-A^n}{1-A} + C \sum_{i=1}^{n-1} A^{n-1-i}$$

$$- CD^{n-1} \sum_{i=1}^{n-1} \left(\frac{A}{D}\right)^{n-1-i}$$

Finally

$$x_n^2 = B\,\frac{1-A^n}{1-A} + C\,\frac{1-A^{n-1}}{1-A} - CD\left(\frac{D^{n-1} - A^{n-1}}{D-A}\right)$$

(AII.5)

and substituting we obtain

$$\sigma_n^2 = \left(\frac{1 - (1-4\rho\beta)^n}{4\beta}\right)$$

$$+ \left(\frac{2\alpha-1}{2\beta}\right)^2 \left[(1-4\rho\beta)^n - (1-2\rho\beta)^{2n}\right] \qquad \text{(AII.6)}$$

BIOGRAPHY

ANDREW JENNINGS was born in Melbourne in 1952. He graduated with honours in Electrical Engineering from Monash University in 1975, and received the Ph.D degree from the same University in 1979. Since 1980 he has worked for the Telecom Australia Research Department in the Transmission Systems Branch. He has been involved with studies of domestic satellite systems, primary level PCM systems and more recently with local digital reticulation systems for ISDN digital customer access. He is the co-inventor of a method of initial synchronisation for echo cancellation full duplex digital transmission systems. He has published a number of papers in fields related to communication systems, and his current research interests include digital signal processing systems in the digital subscriber loop and multiple access techniques for high rate access to an ISDN. He is a Senior Member of both the IREE and the IEEE, and an active member of the Telecommunications Society.

Crosstalk Compatibility in the Local Loop

B.R. CLARKE and G.J. SEMPLE

Telecom Australia Research Laboratories

The crosstalk compatibility of some possible ISDN digital customer access systems with each other and with existing local loop digital systems is considered. It is found that the performance of a burst-mode customer access system can be significantly degraded by crosstalk from other digital systems. In contrast, echo canceller customer access systems are generally compatible with other systems. In addition the performance of existing digital systems, a 72 kbit/s baseband data service and a 2048 kbit/s primary rate access system, are not significantly degraded by the introduction of customer access systems.

KEYWORDS: ISDN, Crosstalk, Compatibility, Customer Access, Digital Transmission

1. INTRODUCTION

The telecommunications networks of many administrations around the world are evolving towards an Integrated Services Digital Network (ISDN). This network will permit customers access to a wide variety of data services as well as supporting digital voice. For economic reasons, full-duplex customer access to the ISDN will generally be achieved by employing the existing two-wire local loop copper cable. Basic access to the ISDN consists of two 64 kbit/s "B" channels supporting voice and data information and one 16 kbit/s "D" channel for signalling and low speed data. With additional framing and maintenance bits, the data rate is typically about 160 kbit/s. The task of reliably and economically transmitting this digital information in full-duplex over the existing two-wire copper cable, which was designed for low bandwidth analogue speech signals, is clearly a demanding one.

There are three important types of interference that affect the error performance of customer access systems; impulsive noise, crosstalk from other identical customer access systems, and crosstalk from different systems that may co-exist in the same cable with the customer access system considered. In all three interference types, a crosstalk path between cable pairs exists due to electro-magnetic coupling within the cable. In the Australian network the major sources of impulsive noise are the signalling events of analogue telephone services such as seizure, dialling and ringing. Impulsive noise is thus a problem that will exist only in the transition from analogue to digital transmission systems. Nevertheless, impulsive noise is a severe transmission impairment in the mixed environment (Ref.1) and this environment will continue for a considerable length of time. The second type of interference, crosstalk from other identical customer access systems, determines the transmission span of these systems when they occupy all of the adjacent cable pairs. This situation may occur in the future if all the analogue and digital systems that exist in the local loop are replaced by one type of customer access system. Impulsive noise and crosstalk from identical systems are important sources of interference but will not be specifically considered in this paper; the interested reader is referred to Ref. 1 for further details. Instead, this paper examines the third type of interference which is crosstalk between non-identical systems and denotes this aspect of the system's crosstalk interference as "compatibility". Crosstalk compatibility is an important issue since customer access systems will be gradually introduced, necessitating that these systems and the existing systems be compatible (i.e., they can co-exist in the same cable without causing significant mutual interference). In addition, other systems may be introduced into the local cable and these systems should also be compatible with the customer access systems. Before the specific crosstalk compatibility cases examined in this paper are listed, the transmission techniques proposed for the customer access systems are briefly reviewed.

Two main two-wire full-duplex transmission techniques have been proposed for ISDN digital customer access. These are the burst-mode (or time compression multiplex) system (Ref. 2) and the hybrid echo canceller system (Ref. 3). The burst-mode system assembles data into blocks and transmits these blocks at a much higher rate. This rate multiplication factor is typically 2.5 resulting in an effective transmission rate of 160 x 2.5 = 400 kbit/s. Full-duplex operation on a single pair of wires is achieved by time-division multiplexing these blocks. In contrast, the hybrid echo canceller system transmits data in both directions simultaneously. Transmit and receive paths are separated using an analogue hybrid combined with an echo canceller to remove

Paper received 15 June 1985.
Final version 7 July 1985.

residual hybrid leakage and signal components reflected from the customer loop.

This paper considers the crosstalk to and from different systems which may exist in the same cable. Different systems may exist in the same cable for four reasons. Firstly, it is conceivable that both burst-mode and echo-canceller systems may be permitted in the local loop. Secondly, customer access systems with different line codes may be allowed to co-exist. Thirdly, customers will be offered not only basic access (160 kbit/s) but also primary access (2048 kbit/s) and these must therefore be jointly compatible. Fourthly, the ISDN will evolve from the current network and compatibility with existing systems in the local loop is an important issue. Hence an investigation into compatibility is necessary to determine if modifications to existing systems may be required or to determine if any constraints should be placed on either the design or the use of customer access systems. The existing systems that are examined in this paper are a 72 kbit/s baseband data service and a 2048 kbit/s primary rate access system. The latter is also used to give insight into the compatibility of basic and primary access systems. Crosstalk compatibility with existing analogue systems in the local loop is not considered.

In this paper the scope of the investigation is indicated in Section 2. The measure of the effect of crosstalk on the system performance used in the paper is based on the Crosstalk Noise Figure (Refs. 4-8) which is briefly reviewed in Section 3. In Section 4 the crosstalk compatibility results are presented in a graphical form allowing easy evaluation of achievable transmission spans. The important features of these results are also discussed in this Section. Briefly, they indicate that:

1. NEXT from other systems into burst-mode systems is significant and limits the achievable transmission span;

2. Echo-canceller systems are compatible with the existing digital systems including the primary rate access system;

3. Crosstalk from burst-mode or echo-canceller systems into existing digital systems does not appear to be a problem;

4. The types of echo-canceller or burst-mode system line codes allowed in the local loop may have to be restricted, since the co-existence of systems using line codes with poor crosstalk properties and systems employing line codes with superior crosstalk performance may significantly reduce the transmission span of the latter.

Finally, in Section 5, some conclusions are made.

2. SCOPE OF THE INVESTIGATION

The purpose of the investigation was to obtain insight into the significance of crosstalk compatibility in the local loop. To decouple the investigation from actual implementations, the paper considers idealised transmission systems. However, a system margin is later incorporated to account for equipment implementation imperfections.

TABLE 1 - Transmission Systems Investigated for Compatibility.

Transmission System	Effective Transmission Rate kbit/s	Line Code
Baseband Data Service	72	WAL1
Echo-Canceller	160	WAL1
Echo-Canceller	160	4B3T
Burst-Mode	400	AMI
Primary Rate Access	2048	HDB3

The transmission systems considered are listed in Table 1. The interested reader is referred to Ref. 9 for further details on the line codes, but a brief sketch of their main properties is given here. The WAL1 (or conditioned diphase) line code has a guaranteed signal transition in the middle of the symbol interval plus a transition at the start of the interval when a binary "zero" is present. In this line code the effective line rate is double the information rate. Alternate mark inversion (AMI) is a ternary line code which transmits no pulses when a binary "zero" is present and a pulse of alternating polarity otherwise. The HDB3 line code is similar to AMI but has a modification which removes sequences of zeros longer than 3 symbols to avoid timing recovery difficulties. The 4B3T line code is a rate reducing block line code which converts four binary digits into three ternary digits on the basis of a running digital sum.

Only digital systems are considered since the wide bandwidth analogue systems (e.g. 15 kHz sound programme channels) for which crosstalk compatibility may be an issue are relatively rare in the local loop. The 72 kbit/s baseband data service and 2048 kbit/s primary rate access systems were selected because it was considered that the crosstalk compatibility of these systems with ISDN transmission systems may be important. The particular line codes for the echo-canceller system (WAL1, 4B3T) and the burst-mode system (AMI) were chosen because international developments suggest that these are prime line code contenders. Other digital systems that could have been considered include: 1) an 80 kbit/s data above voice pre-ISDN system with WAL1 line code, and 2) a 256 kbit/s digital PABX extension system with WAL1 line code. However, in the first case, the results for the 72 kbit/s system will be approximately applicable. Also, the results for the

160 kbit/s WALl echo system will not be dissimilar to those for the 256 kbit/s PABX extension system. Therefore, although the above systems were not explicitly considered it is expected that the conclusions of this report would not be substantially affected.

Two different transmit pulse shapes have been considered. These are:

1. Rectangular pulse (0% trapezoidal)

2. 10% trapezoidal pulse

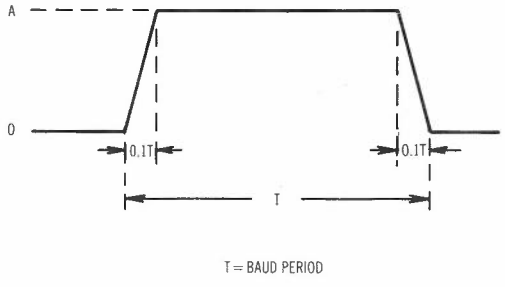


Fig. 1 - A 10% Trapezoidal Pulse

As indicated in Fig. 1, the 10% trapezoidal pulse is essentially a rectangular pulse but with linear rising and falling edges both occupying 10% of the total pulse duration. This pulse shape was included in the investigation because 1) it has reduced spectrum spread compared to the rectangular pulse, and 2) it perhaps better represents the transmit pulse of actual system implementations. The reduction in the spectrum spread may be important since the mean near-end crosstalk (NEXT) attenuation decreases by approximately 15 dB/decade over the frequency range of interest. An additional assumption is that full-width transmit pulses are used for the AMI, 4B3T and HDB3 line codes. While the existing 2048 kbit/s primary rate access systems actually employ half-width pulses, the results contained here for full-width pulses are not substantially different.

The receiver is assumed to contain a linear equalizer which removes all distortion introduced by the cable and results in a 100% raised-cosine pulse shape at the decision point. Perfect timing recovery is also assumed and hence no intersymbol interference is present at the sampling instant. Finally, only homogeneous 0.40mm paper insulated unit twin (PIUT40) cable is considered as this is the most common of the smaller gauge cables in the local network.

3. NEXT NOISE FIGURE

The two principal crosstalk couplings between line systems occupying the same cable are near-end crosstalk (NEXT) and far-end crosstalk (FEXT). However, for all the systems considered in this paper, NEXT is the dominant interference. The reason for this is that, unlike NEXT, FEXT is attenuated by the cable into which it is coupled. It should be mentioned that a burst-mode system allows the possibility of synchronised bursts. In this case, if only burst-mode systems exist on adjacent cable pairs then the transmission span would be FEXT. rather than NEXT, limited. However, if any of the other digital systems outlined in Section 2 are allowed to exist in the same cable as the burst-mode system then it has generally been found that NEXT from these other systems becomes the dominant form of crosstalk into the synchronised burst-mode system. Hence, this paper presents only the NEXT results.

The crosstalk measure used in the investigation was the NEXT Noise Figure (Refs. 1,4-8). This measure has previously been used to determine crosstalk transmission spans for various transmission system proposals for customer access when the disturbing and disturbed systems are identical (Ref. 1). A straightforward extension of this method is used in the present compatibility investigation and so the NEXT Noise Figure is only briefly presented with the emphasis on the unique features of compatibility.

The NEXT Noise Figure is a performance measure ideally suited to compatibility since it arises naturally from system engineering analysis (Ref. 7,8) and allows the loop length limits set by crosstalk to be accurately estimated. The crosstalk performance of a system is determined by both the crosstalk characteristics of the cable and by the system's sensitivity to this crosstalk. The former depends on the cable design and the latter on the particular system parameters. The NEXT Noise Figure allows these two effects to be separately considered. Therefore a direct comparison of the crosstalk compatibility of different systems under a wide variety of crosstalk conditions can be readily made.

The NEXT interference between two different systems occupying the same cable is illustrated in Figure 2. The first block in each system represents the line coder whose output power spectral density is $W_i(f)$, $(i = 1, 2)$, where the subscript 1 represents the disturbing system while the subscript 2 represents the disturbed system. This is followed by a transmit pulse shaping network with transform $S_i(f)$. The cable transfer function is represented by $G(f)$. The receiver contains an ideal linear equaliser with transfer function $E_i(f)$ which removes all cable distortion and presents some desired pulse shape to the decision device. Therefore,

$$E_i(f) = \frac{D_i(f)}{S_i(f)\ G(f)} \quad , \quad i = 1, 2$$

where $D_i(f)$ is the pulse shape transform desired at the decision point. In this paper, the desired pulse shape is chosen to be 100% raised-cosine. The two systems are crosstalk coupled by the NEXT transfer function $X_N(f)$.

It has been found (Ref. 4-6) that for a system disturbed by NEXT to have an error

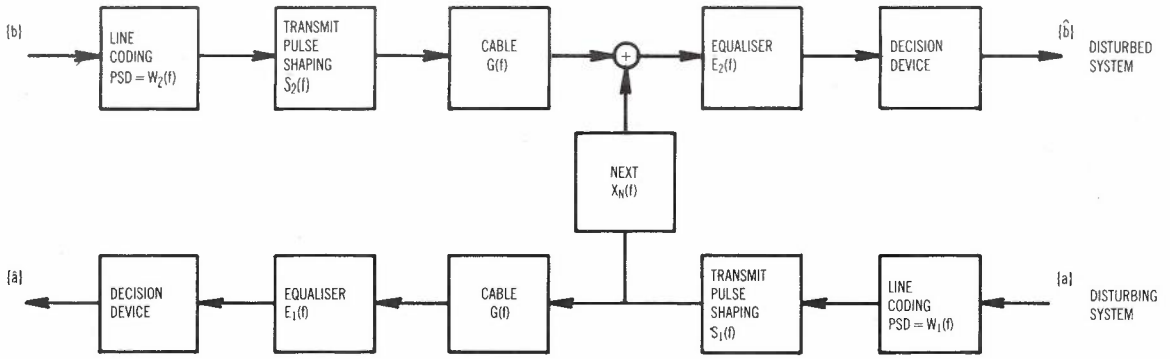


Fig. 2 - NEXT Interference Between Two Different Transmission Systems

probability equal to or better than 10^{-7} with 99% probability, then

$$C_N(\mu_o, \sigma_o, N) > [R_N(f_o, L] \quad (1)^+$$

where μ_o and σ_o represent the NEXT attenuation mean and standard deviation at the reference frequency f_o, N is the number of disturbers, R_N is the NEXT Noise Figure and L is the cable length. The lower the NEXT Noise Figure, R_N, the better the crosstalk performance. The LHS of (1) depends only on the cable characteristics and can be expressed as

$$C_N(\mu_o, \sigma_o, N) = \mu_o - S(\sigma_o, N) \quad (2)$$

where $S(\sigma_o, N)$ is plotted in Figure 3. Typical NEXT attenuation mean and standard deviation values obtained from cable measurements are $\mu_o = 83.6$ and $\sigma_o = 8.0$.

The RHS of (1) is the NEXT Noise Figure which is composed of two terms:

$$[R_N(f_o, L)] = [I_N(f_o, L)] - [N_o] \quad (3)$$

The I_N term reflects the noise amplification introduced by the equaliser and

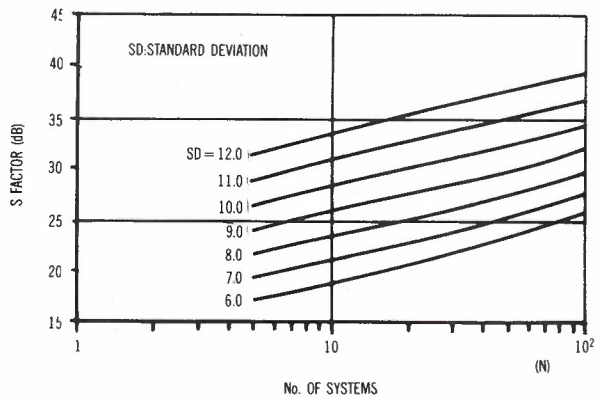


Fig. 3 - S(SD,N) Factor in dB

+ Throughout this paper, quantities in square brackets are in decibels i.e. $[X] = 10 \log_{10}(X)$.

can be expressed as:

$$I_N = \int_0^\infty W_1(f) \, |S_1(f)|^2 \, |X_N(f, f_o, L)|^2 \cdot \frac{|D_2(f)|^2}{|S_2(f)|^2 \, |G(f,L|^2} \, df \quad (4)$$

The case where both systems are identical is a special case of the compatibility study where the transmit pulse shape transforms are identical and cancel in (4).

The N_o term represents the mean-square noise voltage at the decision point required to produce an error rate of 10^{-7}. This term depends on parameters associated with the decision process (e.g. number of code levels) and normally includes the effects of intersymbol interference and sampling offset. However, in this paper, sampling at the optimum instant is assumed and, since equalisation is assumed ideal, no intersymbol interference exists at this instant. Furthermore, it is assumed that the signal level at the decision point is equalised to one volt. Therefore for noise with Gaussian distributed amplitude:

$$[N_o] = \begin{cases} -14.4 \text{ dB for 2 level line codes} \\ \quad \text{e.g. WAL1, WAL2, MWAL2} \\ -20.4 \text{ dB for 3 level line codes} \\ \quad \text{e.g. AMI, 4B3T, HDB3} \end{cases}$$

Two additional terms have been inserted into (1) to yield the following expression:

$$C_N - M - Z > [R_N] \quad (5)$$

M is a design margin to allow for system implementation imperfections and a 6 dB margin has been used in this paper. The term Z is unique to compatibility studies and represents a transmit level factor which allows different system transmit levels to be taken into account:

$$Z = 20 \log \left[\frac{V_1}{V_2}\right]$$

where V_1/V_2 represents the ratio of the disturbing system to the disturbed system transmit peak voltages. The Z term has been extracted from the I_N component of R_N (see (4)). Thus, the R_N are calculated assuming equal transmit levels in the disturbed and disturbing systems. The Z term then accounts for any transmit level differences.

4. RESULTS

Expression (5) is the equation from which the maximum transmission span length is determined such that the performance objective of 99% of systems with an error probability of better than 10^{-7} is satisfied. This maximum span is the length at which the equality in (5) holds. The approach used in this paper to determine this length is to plot the NEXT Noise Figure [R_N] and graphically determine the length at which the equality is satisfied. The chief advantage of this approach is flexibility since the NEXT Noise Figure plots are applicable to a range of crosstalk conditions, so changes in the cable NEXT attenuation characteristics or the number of disturbers or the relative transmission levels can all be accommodated without the need to recalculate the NEXT Noise Figure R_N.

To determine the maximum transmission span, a value for the LHS of (5) must be found. This requires cable data and the relative transmit levels between the different systems to be selected. Reasonable choices are made for insight into the significance of crosstalk compatibility.

For the worst case of the disturbers being within the same cable unit, typical NEXT attenuation mean and standard deviation measured values are $\mu_0 = 83.6$ and $\sigma_0 = 8.0$ respectively (Sec. 3). A cable unit consists of 10 cable pairs and so we assume that the unit contains 9 identical disturbing systems and one, perhaps different, disturbed system. For these values, from Fig. 2 and (2):

$$C_N = \mu_0 - S(\sigma_0, N)$$

$$= 83.6 - 23.1$$

$$= 60.5$$

TABLE 2 - System Transmit Levels

System	Transmit Power (into 120Ω) dBm	Peak Transmit Pulse Voltage Level V^2	dBV
72 kbit/s WAL1	6	0.5	-3.2
160 kbit/s WAL1	10	1.2	0.8
160 kbit/s 4B3T	10	1.7	2.4
400 kbit/s AMI	10	2.4	3.8
2048 kbit/s HDB3	16	3.0	4.8

Table 2 summarises the different system transmit levels that have been used. Using this information the LHS of (5) can now be calculated.

The NEXT Noise Figure plots are contained in Figs. 4-10. There is a separate graph for the crosstalk into each of the different systems. Figs. 4-8 are the Noise Figure results when all systems have a rectangular transmit pulse shape. Figs. 9 and 10 are when

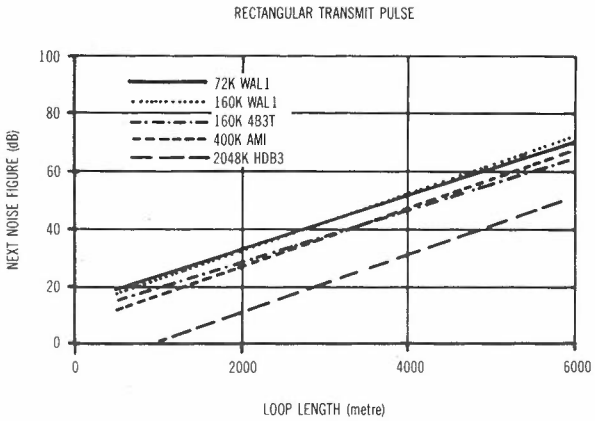


Fig. 4 - Crosstalk Noise Figure At 40 kHz Into 72 kbit/s WAL1

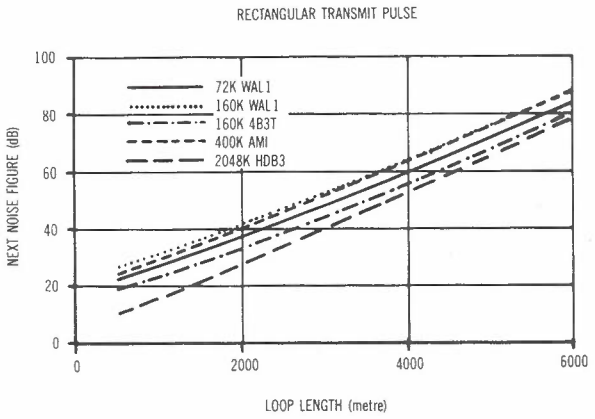


Fig. 5 - Crosstalk Noise Figure at 40 kHz Into 160 kbit/s WAL1

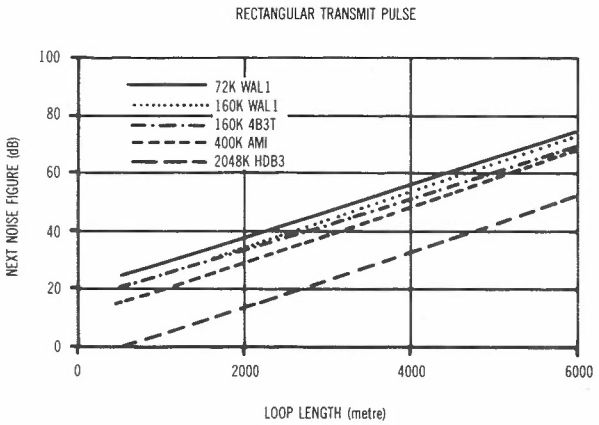


Fig. 6 - Crosstalk Noise Figure at 40 kHz Into 160 kbit/s 4B3T

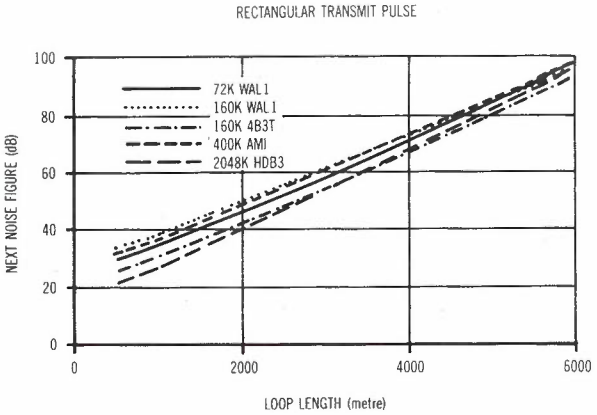


Fig. 7 - Crosstalk Noise Figure at 40 kHz Into 400 kbit/s AMI

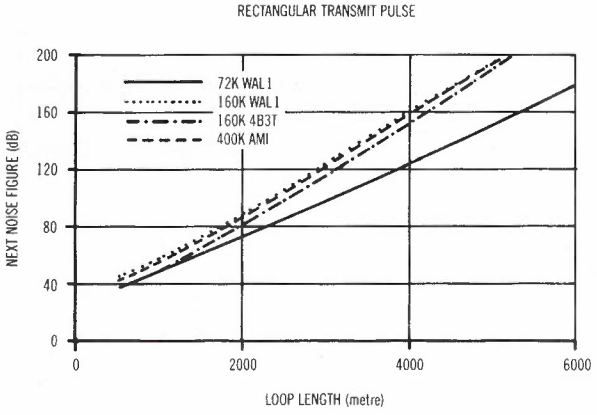


Fig. 8 - Crosstalk Noise Figure at 40 kHz Into 2048 kbit/s HDB3

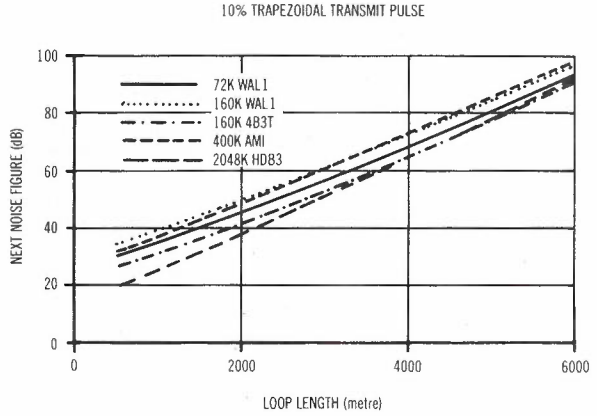


Fig. 9 - Crosstalk Noise Figure at 40 kHz Into 400 kbit/s AMI

all systems possess a 10% trapezoidal transmit pulse shape. For this pulse shape, the crosstalk into the 400 kbit/s AMI system and the 2048 kbit/s HDB3 system only are considered since it is only these that are significantly different from the rectangular pulse shape cases.

To illustrate the determination of the transmission spans from the NEXT Noise Figure graphs, the 72 kbit/s WAL1 system is considered in detail. Table 3 contains the values of the LHS of (5) and the transmission spans for

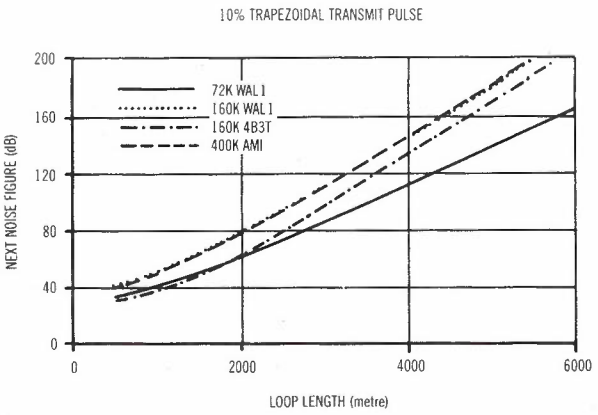


Fig. 10 - Crosstalk Noise Figure at 40 kHz Into 2048 kbit/s HDB3

TABLE 3 - Transmission Spans of the 72 kbit/s WAL1 Baseband Digital System for Various Disturbing Systems. A Rectangular Transmit Pulse is Assumed for all Systems.

Disturbing System	C_N - M - Z	Transmission Span (km)
72 kbit/s WAL1	54.5	4.2
160 kbit/s WAL1	50.5	3.8
160 kbit/s 4B3T	48.9	4.2
400 kbit/s AMI	47.5	4.0
2048 kbit/s HDB3	46.7	5.6

various disturbing systems. The C_N - M - Z term is found using the previously given values for C_N (= 60.5) and M (= 6) and determining Z from Table 2. The spans are found by determining the length at which the NEXT Noise Figure for each system equals the values of C_N - M - Z contained in Table 3. Using the above procedure, the transmission limits imposed by the various transmission systems when they interfere with each system have been determined. These results are summarised in Table 4.

Two systems are considered to be crosstalk compatible when the crosstalk between the systems is insufficient to cause the performance of either system operating at its maximum transmission span to become unacceptable. This maximum span may be determined by either impulsive noise or by crosstalk when disturbing and disturbed systems are identical. For the purpose of this investigation we consider that the maximum spans for the customer access systems are not determined by impulsive noise. Under this assumption it is shown later in this section that the burst-mode system is incompatible with most of the other systems considered. However, if impulsive noise is considered then Ref. 1 indicates that only relatively short maximum transmission spans can be achieved and the burst-mode system then becomes compatibile with the other systems.

TABLE 4 - Transmission Spans for all Combinations of Disturbing and Disturbed Systems.

Disturbing System	Disturbed system transmission span (km)				
	72 kbit/s WAL1	160 kbit/s WAL1	160 kbit/s 4B3T	400 kbit/s AMI	2048 kbit/s HDB3
72 kbit/s WAL1	4.2 (3.7*)	3.7	4.4	3.3	1.7
160 kbit/s WAL1	3.8	3.1	4.3	2.7	1.1
160 kbit/s 4B3T	4.2	3.6	4.3	3.2	1.2
400 kbit/s AMI	4.0	2.9	4.4	2.5	1.0
2048 kbit/s HDB3	5.6	3.8	6.0	2.9	1.0 *

* Note: These are Implemented Maximum Transmission Spans and are Determined by Impulsive Noise.

From the results in Table 4, some observations can be made about the crosstalk compatibility of the different systems. Due to impulsive noise, the current maximum design span for the 72 kbit/s WAL1 system is some 3.7 km for PIUT40 cable. The relative significance of other disturbing systems can be determined by comparison with the case when the identical 72 kbit/s WAL1 systems are the disturbers. From Table 4, 160 kbit/s WAL1 and 400 kbit/s AMI disturbers reduce the transmission span of the 72 kbit/s WAL1 system in comparison with the case of 72 kbit/s WAL1 disturbers. However, since the achievable spans are still greater than the implemented maximum span of 3.7 km, it is believed that the considered systems will not significantly degrade the performance of the 72 kbit/s WAL1 baseband system.

For the case of the 160 kbit/s WAL1 system the minimum transmission span occurs when the 400 kbit/s AMI systems are the disturbers. However, the transmission span is not significantly different from the case when identical 160 kbit/s WAL1 systems are the disturbers. Hence, the presence of other than identical disturbing systems will not significantly degrade the performance of the 160 kbit/s WAL1 system. In other words, the maximum span for a 160 kbit/s WAL1 system will not be limited by interference from the other systems but rather by crosstalk from identical systems.

The lowest transmission span of the 160 kbit/s 4B3T system is imposed not by identical disturbing systems but by the 160 kbit/s WAL1 systems. This is due to the superior crosstalk properties of the 4B3T line code compared with the WAL1 line code (Ref. 1). Whilst the reduction in the transmission span is small for this case, it indicates the more general principle that systems using such superior line codes may not achieve their potential transmission spans if digital systems with less efficient line codes exist in the same cable unit. Hence, where systems with transmission rates of the same order of magnitude are allowed to co-exist within a cable unit, then the span limits set for systems with superior line codes (from a crosstalk point of view) must take into account the crosstalk from less efficient systems. For example, if both 160 kbit/s 4B3T and 160 kbit/s WAL1 echo canceller systems were allowed for ISDN digital customer access, then the transmission span would have to be determined by taking into account the crosstalk interference from 160 kbit/s WAL1 systems.

If only 400 kbit/s AMI burst-mode systems are allowed in the same cable unit and if the bursts are synchronised, the achievable transmission span is approximately 3.5 km (Ref. 1). However, from Table 4, if any of the other systems are allowed to exist in the same unit, then the NEXT from these systems will significantly reduce the transmission span. Thus, for a synchronised AMI burst-mode system it is important either to ensure the unit is clear of all other significant sources of NEXT or the transmission spans may have to be limited to less than approximately 2.7 km.

The final case to consider is the 2048 kbit/s HDB3 primary rate access system. All the transmission span limits imposed by the other systems match or exceed the implemented maximum transmission length of the 2048 kbit/s primary rate access system. Therefore, crosstalk from these other systems is not expected to degrade the performance of existing 2048 kbit/s systems. Also the results indicate that basic and primary rate access systems may co-exist in the same cable unit.

Comparison of the crosstalk interference for rectangular transmit pulses and 10% trapezoidal pulses reveals that the use of transmit pulse shapes with reduced spectrum spread is of limited value. It is only for crosstalk into the relatively wide bandwidth 2048 kbit/s HDB3 system that crosstalk is greatly reduced, but even this is not significant since the crosstalk from other systems into the 2048 kbit/s HDB3 system is not a problem with rectangular transmit pulses. It is interesting to note that, in some cases, reducing the transmit spectrum spread actually increases the crosstalk interference. This is due to the increase in the disturbing system transmit power that is within the disturbed systems receiver bandwidth.

5. CONCLUSIONS

This paper has considered the significance of crosstalk from digital disturbing systems which differ from the disturbed system. Five systems were considered; 72 kbit/s WAL1 baseband data service, 160 kbit/s WAL1 echo canceller, 160 kbit/s 4B3T echo canceller, 400 kbit/s AMI burst-mode, and 2048 kbit/s HDB3 primary rate access. To assess the influence of crosstalk compatibility, the NEXT Noise Figure was modified. It was assumed that the maximum transmission spans of the customer access systems were not determined by impulse noise. The results were presented in graphical form and are quite flexible, allowing a range of crosstalk conditions to be assessed using the same graph. A set of realistic crosstalk conditions were then selected and the transmission spans determined for all combinations of disturbing and disturbed systems. Although specific conditions were selected, it is believed that the following conclusions generally apply. Firstly, the NEXT into the burst-mode system from all other systems is significant. Unless all other significant disturbing systems can be removed from the cable unit, the transmission span will be significantly reduced (2.7 km for the conditions considered). Secondly, the echo canceller systems are essentially compatible with the other systems (apart from the burst-mode system). This includes the compatibility of basic and primary rate customer access systems. Thirdly, the crosstalk from the 3 proposed customer access systems, 160 kbit/s WAL1, 160 kbit/s 4B3T and 400 kbit/s AMI, into the 2 existing systems, 72 kbit/s WAL1 and 2048 kbit/s HDB3, will not significantly degrade the performance of these existing systems. Hence, re-engineering of these systems caused by the introduction of any one of the ISDN digital customer access systems considered will generally be unnecessary. Fourthly, if several line codes, say for an echo canceller type system, are used in the local loop then the achievable span of those systems with superior crosstalk performance may be restricted by crosstalk from the other systems. Alternatively the type of customer access systems with different line codes permitted in a cable unit may have to be restricted. Finally, the use of transmit pulse shapes other than rectangular with reduced spectrum spread results in only insignificant improvements in the transmission span.

6. REFERENCES

1. G.J. Semple, P.G. Potter and B.R. Clarke, "Crosstalk compatibility and impulsive noise in the digital customer loop", IREECON International, 1985, Convention Digest.

2. N. Onoue, R. Komiya and Y. Inoue, "Time-shared two-wire digital subscriber transmission system and its application to the digital telephone set", IEEE Trans. Commun., Vol. COM-29, pp. 1565-1572, Nov. 1981.

3. N.A.M. Verhoeckx, H.C. van den Elzen, F.A.M. Snijders and P.J. van Gerwen, "Digital echo cancellation for baseband data transmission", IEEE Trans. Acoust., Speech and Signal Proc., Vol. ASSP-27, pp. 768-781, Dec. 1979.

4. A.J. Gibbs and R. Addie, "The covariance of near end crosstalk and its application to PCM system engineering in multipair cable", IEEE Trans. on Comm., Vol. COM-27, No. 2, pp. 469-477, Feb. 1979.

5. G.J. Semple and A.J. Gibbs, "Assessment of methods for evaluating the immunity of PCM regenerators to near end crosstalk", IEEE Trans. on Comm., Vol. COM-30, No. 7, pp. 1791-1797, July 1982.

6. L.J. Millott and G. Nicholson, "Aspects of PCM regenerator design for crosstalk limited environments", IEEE Trans. on Comm., Vol. COM-29, No. 9, pp. 1330-1336, Sep. 1981.

7. A.J. Gibbs, "Measurement of P.C.M. regenerator crosstalk performance", Elect. Lett., Vol. 15, No. 3, pp. 82-83, Feb. 1979.

8. A.J. Gibbs, L.J. Millott and G. Nicholson, "Optimum P.C.M. regenerator performance in presence crosstalk", Elect. Lett., Vol. 15, No. 16, pp. 490-492, Aug. 1979.

9. N.Q. Duc and B.M. Smith, "Line coding for digital data transmission", ATR, Vol. 11, No. 2, pp. 14-27, 1977.

BIOGRAPHIES

BRUCE CLARKE received a B.E. and a PhD from the University of New South Wales. The PhD thesis concerned the joint approximation of linear phase and rectangular amplitude by non-minimum phase filters. In 1983 Dr. Clarke joined the Telecom Research Laboratories where he is currently senior engineer in the Line and Data Systems Section of the Transmission Systems Branch. He has been investigating various aspects of ISDN customer access digital transmission systems including crosstalk compatibility, baud-rate timing recovery, adaptive algorithms, architectures and start-up procedures.

JOHN SEMPLE received the Degrees of Bachelor of Engineering (Electrical) with honours and a Master of Engineering Science from the University of Melbourne in 1966 and 1968 respectively. He commenced work as an Engineer with the then APO Research Laboratories, Pulse Systems Division in January 1967, where he was concerned with problems associated with primary level PCM and high capacity digital transmission systems. In 1972 he joined the Line and Data Systems Section, which is now part of the Transmission Systems Branch of the Telecom Research Laboratories, where he continued work associated with the application of PCM transmission in the inter-exchange cable network. In about 1980 Mr. Semple commenced the investigation of local digital reticulation systems, and is currently concerned with Integrated Services Digital Network (ISDN) developments and the digital two-wire full duplex transmission systems required to provide customer access to such a network.

Benefits of Authorship for ATR

- ATR contains papers of a high standard and has extensive international distribution and abstracting.
- ATR facilitates contact with other research workers in the telecommunication field in Australia.
- Published papers will contribute to the viability and vitality of future Australian research and industry.
- Publication is facilitated because ATR arranges drafting of figures at no cost to the author.

Information for Authors

Manuscripts should be written clearly in English. They must be typed using double spacing with each page numbered sequentially in the top right hand corner. The title, not exceeding two lines, should be typed in capital letters at the top of page 1. Name(s) of author(s), in capitals, with affiliation(s) in lower case underneath, should be inserted on the left side of the page below the title. An abstract, not exceeding 150 words and indicating the aim, scope and conclusions of the paper, should follow below the affiliation(s).

ATR permits three orders of headings to be used in the manuscript. First-order headings should be typed in capitals and underlined. Each first-order heading should be prefixed by a number which indicates its sequence in the text, followed by a full stop. Second-order headings should be underlined and typed in lower case letters except for the first letter of each word in the heading, which should be typed as a capital. They should be prefixed by numbers separated by a full stop to indicate their hierarchical dependence on the first-order heading. Third-order headings are typed as for second-order headings, underlined and followed by a full stop. The text should continue on the same line as the heading. Numbering of third-order headings is optional but, when used, should indicate its hierarchical dependence on the second-order heading (i.e. two full stops should be used as separators).

Tables may be included in the manuscript and sequentially numbered in the order in which they are called up in the text. The table heading should appear above the table. Figures must be supplied as, at least, clear unambiguous freehand sketches. Figures should be sequentially numbered in the order in which they are called in the text using the form: Fig. 1. A separate list of figure captions is to be provided with the manuscript. Equations are to be numbered consecutively with Arabic numerals in parentheses, placed at the right hand margin.

A list of references should be given at the end of the manuscript, typed in close spacing with a line between each reference cited. References must be sequentially numbered in the order in which they are called in the text. They should appear in the text using the form: (Ref. 1). The format for references is shown in the following examples.

Harris, R.J., "Comparison of Network Dimensioning Models", ATR, Vol. 18, No. 2, 1984, pp 59-69.

Abramowitz, M. and Stegun, I.A., (Eds), Handbook of Mathematical Functions, Dover, New York, 1965.

Four copies of the paper, together with a biography and clear photograph of each of the authors should be submitted to the secretary (see inside front cover). All submissions are reviewed by referees who will recommend acceptance, modification or rejection of the material for publication. After acceptance and publication of a manuscript authors of each paper will receive 50 free reprints of the paper and a complimentary copy of the journal.

ATR AUSTRALIAN TELECOMMUNICATION RESEARCH
ISSN 0001-2777

VOLUME 19, NUMBER 2,
1985

Titles (Abbreviated)