SUBJECTIVE EVALUATION AND COMPARISON OF DIGITAL

AND ANALOG MODULATION SYSTEMS

by

RENÉ DOUVILLE

B.A.Sc., University of British Columbia, 1966

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in the Department of

Electrical Engineering

We accept this thesis as conforming to the

required standard

Research Supervisor

Members of Committee

Head of Department

Members of the Department

of Electrical Engineering

THE UNIVERSITY OF BRITISH COLUMBIA

September, 1968

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and Study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of _ELECTRICAL ENGINEERING_

The University of British Columbia
Vancouver 8, Canada

Date _October 7, 1968_

ABSTRACT

The ultimate measure of performance of any communication system is the subjective quality of the received message. In this thesis, the subjective quality of the output of a differential pulse code modulation (DPCM) system was measured as a function of the number of bits of quantization L, the speech bandwidth W, the ratio r of the sampling frequency f_s to the Nyquist frequency 2W, and the number of feedback samples N. For previous-sample feedback (N = 1) the maximum subjective quality was obtained as a function of the bit rate R = 2rWL. The optimum sampling rate was found to be the Nyquist rate; the improvement afforded by increasing f_s over 2W was more than offset by the required increase in bit rate. Noise in the feedback loop caused by dc offset errors and noise present in the output of the feedback coefficient amplifiers prevented a thorough investigation of two- and three- sample feedback, although some results were obtained.

The subjective quality of delta modulated (ΔM) speech was obtained vs r and W, and the quality of amplitude modulated (AM) speech was measured as a function of W and channel signal-to-noise ratio. A technique was then devised to use the AM results to estimate the subjective quality of phase modulated (PM) speech.

A comparison was then made of the capabilities of PCM, DPCM, ΔM, single sideband-AM (SSB-AM), double sideband-AM (DSB-AM), and PM. It was found that when the available channel capacity is small, SSB-AM and DSB-AM are subjectively better than PCM, DPCM, and ΔM. However, for high quality speech communication, DPCM requires less channel capacity than PCM, ΔM, DSB-AM, SSB-AM or PM.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

ACKNOWLEDGEMENT

1. INTRODUCTION

1.1 Communications Systems

A communication system may be considered to be composed of the five parts shown in Fig. 1.1.
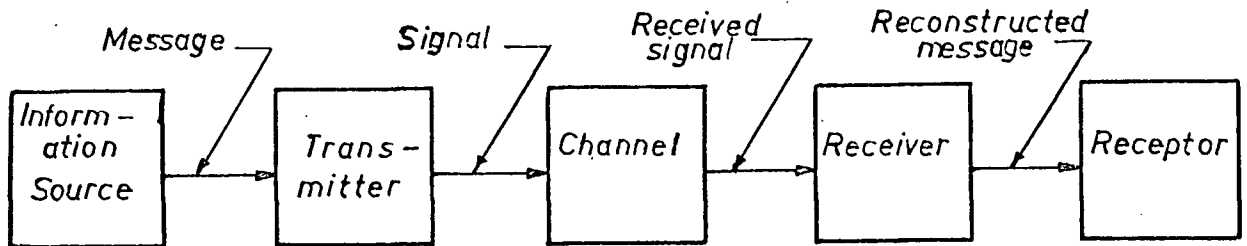
Fig. 1.1 A general communication system

The information source produces the message to be communicated. The transmitter operates on the message to prepare it for transmission over the channel. The channel distorts the signal prior to its reaching the input of the receiver. The receiver reconstructs the message which is then presented to the final receptor. The best performance measure of such a communication system is how satisfac-tory the receptor considers the reconstructed message to be given that he (it) knows exactly the original message.

The operation of the transmitter may be further subdivided into three basic operations: source encoding, channel encoding and modulation. In the case of amplitude and angle modulation systems, the source and channel encoder combination consists of a linear filter. The modulators, in these cases, vary the amplitude or the angle of a carrier according to the variation of the message amplitude.

In a digital communication system, the purpose of the source encoder is, ideally, to encode the message into a sequence of equi-probable, independent, discrete symbols. The channel encoder then

adds, in a way that is optimum for the particular channel and modulation system used, enough redundancy to keep the probability of a transmission error below some specified level. The modulator accepts the channel encoder output and generates from it a signal suitable for transmission over the prescribed channel. The receiver demodulates the received signal, and reconstructs a delayed replica of the input.

1.2   Brief Review of Source Encoding Techniques

An ideal source encoder removes all redundancy from the input signal. In a practical system, only part of the redundancy is removed, since the removal of all redundancy is usually impractical if not impossible. A measure of source encoder efficiency is the amount of redundancy removed.

Much work has been done on the optimization of the source encoder for pulse code modulation (PCM) systems*. Several investigators have been concerned with the optimization of the quantization process alone [1-4] while others have been concerned with optimizing the combined process of quantizing, sampling, and reconstruction [5-9].

Many systems which have been proposed and investigated use feedback around the quantizer to reduce the redundancy in the encoder output. One such system is the differential pulse code modulation (DPCM) system in Fig. 1.2. Delta modulation (ΔM) results when the quantizer in Fig. 1.2 contains two output levels.

Numerous investigations of DPCM have been carried out. Van de Weg [10] derived the signal-to-noise ratio as a function of the number of bits of quantization and the ratio of the sampling frequency to the bandwidth for a system having a bandlimited white noise

---

* A PCM system results if the feedback filters are removed from the source encoder and source decoder in Fig. 1.2.
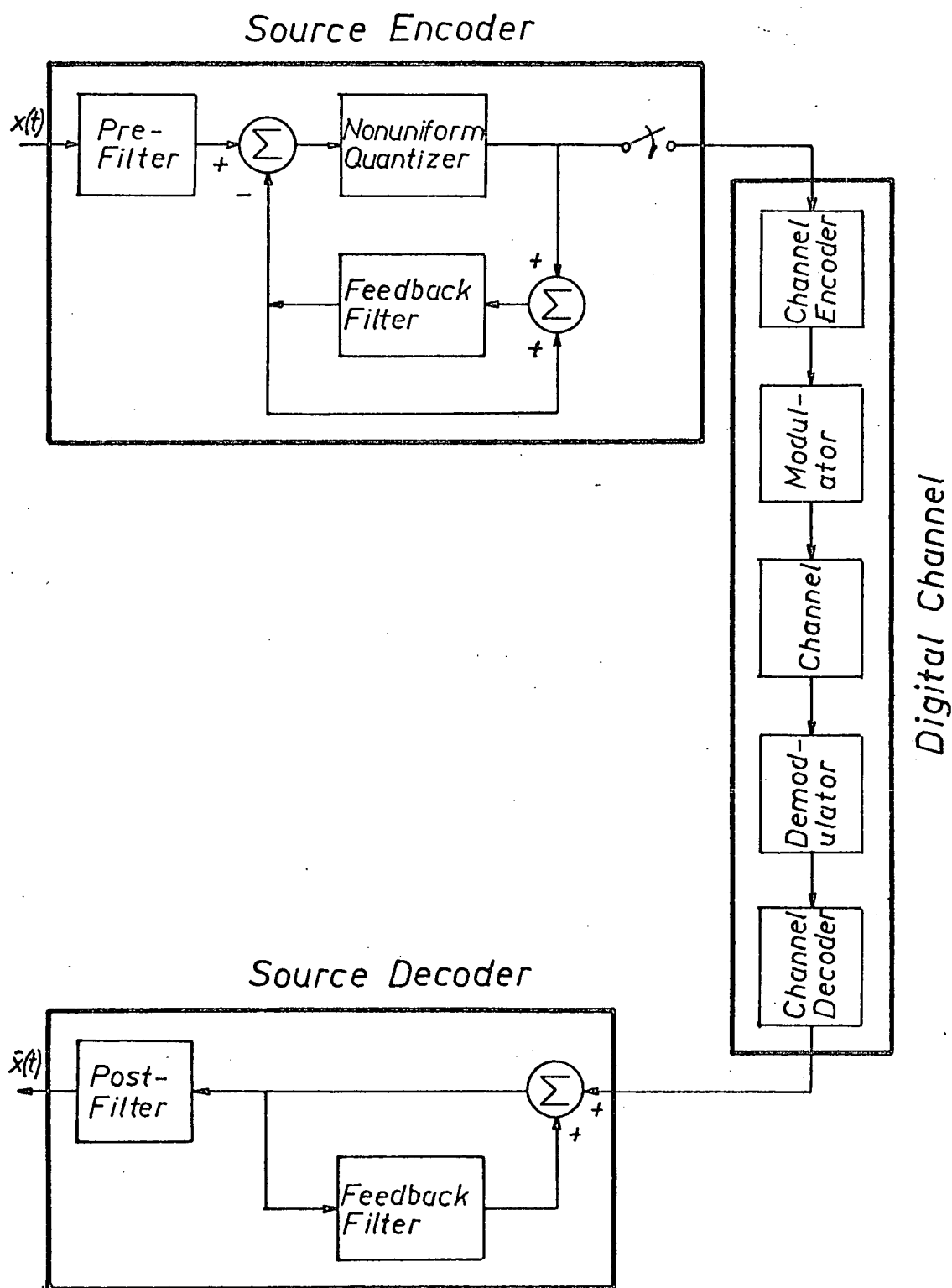
Fig. 1.2 General differential pulse code modulation (DPCM) system

input and a single integrator in the feedback path. Nitadori [11] obtained the quantizer characteristic which minimizes the quantization noise for speech signals when the feedback network is an ideal integrator. O'Neal [12,13] and McDonald [14] derived approximate formulas for the signal-to-noise ratios of DPCM systems having linear predictive feedback networks. O'Neal obtained simulation results for video input signals, while McDonald obtained results for speech input signals. Irwin and O'Neal [15] derived the optimum quantizer and predictor for a Gaussian stationary wide-sense Markov input under the assumption that the quantization noise is Gaussian white. Donaldson and Chan [16] have derived an expression for the signal-to-noise ratio as a function of the bandwidth of the message, the sampling frequency, the quantizer characteristic, the linear prediction coefficients and the statistics of the message and channel noise. O'Neal [17] , Sharma [18] , Hosakawa, Onaga, Katusho and Kato [19] , and Abate [20] , have considered the effects of quantization in ΔM systems.

A complete analysis of the mean square error in the outputs of PCM and DPCM source encoders is still lacking. Even if such a complete analysis were available, the ultimate evaluation of any encoding system requires subjective measurements on real-time systems. Although several investigators [21-23] have made subjective measurements on video systems, very little work has been done on the subjective evaluation of DPCM speech communication systems. Recently, Donaldson and Chan [16] , devised a technique for evaluating as a function of an arbitrary number of system parameters the subjective quality of voice communication systems, and used this technique to measure, as a function of speech bandwidth W and the number of quantization bits L, the subjective quality of PCM and DPCM speech. The sampling rate equalled the Nyquist rate in this investigation.

Few attempts have been made to compare different voice communication systems, either on the basis of mean square difference between the transmitted and received waveforms or on the basis of subjective quality. Comparisons of some particular systems and channels have been carried out by Goblick [24] who compared communication systems on the basis of required channel capacity for the cases of Gaussian inputs with various spectra. He pointed out that the performance of digital systems is limited both by the efficiency of source encoding and the degree of channel interference, whereas the performances of analog systems are restricted only by the channel interference.

1.3 Scope of the Thesis

The work described in this thesis was conducted in order to enable various digital and analog voice communication systems to be compared on the basis of subjective quality. The subjective quality of the output of the system in Fig. 2.1 was obtained as a function of the number of bits of quantization L, the bandwidth W, and the ratio r of the sampling frequency to Nyquist sampling frequency $2W$ for one sample feedback (N = 1 in the feedback loop). It was found that an optimum choice of r, W, and L existed for all bit rates R = 2rWL. It was also found that the optimum sampling rate was the Nyquist rate, and that the improvement afforded by increasing the sampling frequency over the Nyquist rate was more than offset by the required increase in bit rate.

The subjective effect of using more than one sample of feedback was investigated. Noise in the feedback loop caused by dc offset errors in the sample and hold circuits and noise present in the outputs of the feedback coefficient amplifiers prevented a thorough investigation of two- and three- sample feedback.

The subjective value of ΔM speech vs W and r was measured.

The subjective quality of AM communications was measured vs W and channel signal-to-noise ratio, and the maximum subjective quality obtainable for a given channel was determined. A technique was then developed for using these results to estimate the subjective quality of speech transmitted by phase modulation (PM) through a channel of given noise level and bandwidth.

A comparison was made of the subjective performance capabilities of PCM, DPCM, ΔM, DSB-AM, SSB-AM and PM on an important class of communication channels. It was found that except when large channel capacities are available, SSB-SC and DSB-SC effectively outperform ΔM, PCM, and DPCM. However, when large channel capacities are available, the performance of DPCM is significantly better than that of DSB-SC or SSB-SC. The results also indicate that, for high quality speech communications, DPCM requires less channel capacity than does PM.

2. SIGNAL-TO-NOISE RATIO IN PCM, DPCM, AM, FM AND PM COMMUNICATION SYSTEMS

2.1 Signal-to-Noise Ratio in DPCM

A practical DPCM system appears in Fig. 2.1. In the following analysis, the digital channel noise is assumed to be negligible*. The sampler is represented as a product modulator in which the input is multiplied by an infinite sequence of pulses of width Δ and unit amplitude.

The encoder transmits message

$$s(t) = p(t) \cdot (e(t) + q(t))$$
$$= p(t) \cdot (x_1 + q) * (\delta - h) \tag{2.1}$$

where the symbol * denotes convolution. The decoder receives $s(t)$ (for $n(t) = 0$) yielding an output

$$\hat{x} = r * f * g_o \tag{2.2}$$

where $f(t)$ is the impulse response of the system having transfer function $1/(1-H(f))$ and $g_o(t) = (2WB/f_s\Delta)(\sin 2\pi Wt/2\pi Wt)$ is the impulse response of the receiver lowpass filter. Combining equations (2.1) and (2.2) yields

$$\hat{x} = \left\{ p \cdot \left[(x_1 + q) * (\delta - h) \right] \right\} * f * g_o$$

Since $p \cdot (x_1 * (\delta - h)) * f * g_o = x_1/A$ for $AB = 1$ and $f_s \geqslant 2W$,

$$\hat{x} = x_1/A + \left[p \cdot (q * (\delta - h)) \right] * f * g_o \quad . \tag{2.3}$$

The mean square error in the output is therefore

---

* Although the removal of redundancy in the transmitted bit stream makes the received message more susceptible to channel noise, in a properly designed system, channel errors can be made arbitrarily small [25,26] .
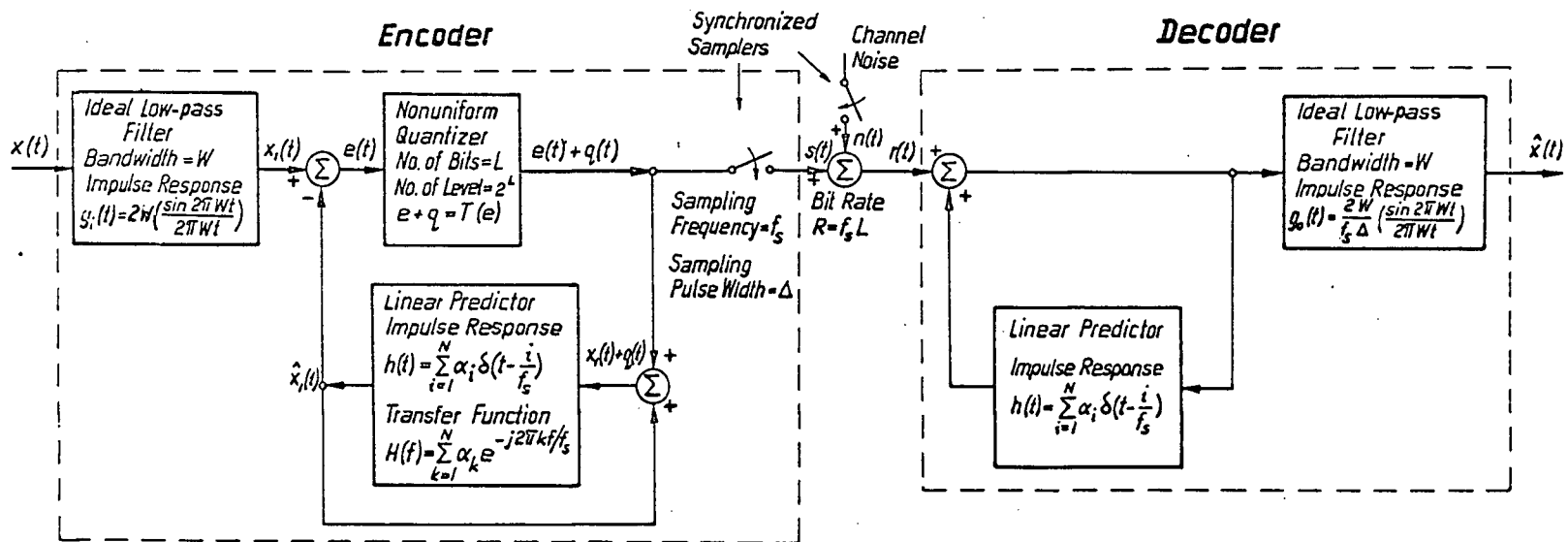
2.1 A practical DPCM system.

$$\overline{\varepsilon^2} = \overline{(x - \hat{x})^2}$$

$$= \overline{(x - x_1/A)^2} + \overline{(\hat{x} - x_1/A)^2} - \overline{2(\hat{x} - x_1/A)(x - x_1/A)} \qquad (2.4)$$

The last term in (2.4) equals zero and

$$\overline{(x - x_1/A)^2} = 2 \int_W^\infty X(f)df \quad .$$

The power density spectrum of $q*(\delta - h)$ equals $Q(f)|1 - H(f)|^2$ and

the spectrum of $p \cdot (q*(\delta - h))$ equals $\sum_{k=-\infty}^{\infty} |p_k|^2 Q(f-kf_s)|1 - H(f-kf_s)|^2$.

Since the transfer function of the decoder equals $B/(f_s \Delta(1 - H(f)))$

for $-W < f < W$, and since $H(f-kf_s) = H(f)$ for all k, by assuming impulse

sampling ($|p_k/f_s| = 1$) for all k, (2.4) becomes

$$\overline{\varepsilon^2} = 2 \int_W^\infty X(f)df + B^2 \sum_{k=-\infty}^{\infty} \int_{-W}^{W} Q(f-kf_s)df \qquad (2.5)$$

In Figure 2.1

$$q = T(e)-e$$

$$= T\left[x_1*(\delta - h)-q*h\right] - \left[x_1*(\delta - h)-q*h\right] \qquad (2.6)$$

If $q_o(t)$ is a solution to (2.6) for a given $x_1(t)$ and if this $x_1(t)$

is replaced by $Kx_1(t)$ with the quantizer characteristic being scaled

such that $q = KT(e/K)$, then the new value of $q(t)$ which satisfies

(2.6) is $Kq_o(t)$. Since $q(t)$ is therefore proportional to $x_1(t)$ and

since $e(t) = x_1*(\delta - h)-q*h$ is proportional to $x_1(t)$, $q(t)$ is pro-

portional to $e(t)$ and therefore $Q(f)$ is proportional to $\overline{e^2}$ for all f.

Thus,

$$\overline{n_q^2} \triangleq B^2 \sum_{k=-\infty}^{\infty} \int_{-W}^{W} Q(f-kf_s)df = B^2 f_Q \overline{e^2} \qquad (2.7)$$

where $\overline{n_q^2}$ is the total noise power in the received signal due to quan-

tization, f_Q is a constant of proportionality which depends on W,

f_s, function T and on the second order amplitude probability density

of $e(t)$. From Figure 2.1 for $\alpha_o = -1$

$$\overline{e^2} = \overline{(x_1*(\delta - h) - q*h)^2}$$

$$= \overline{\left[x_1 - \sum_{i=1}^{N} \alpha_i x_1(t - i/f_s)\right]^2} + \overline{\left[\sum_{i=1}^{N} \alpha_1 q(t - i/f_s)\right]^2}$$

$$- 2\overline{\left[\sum_{i=1}^{N} \alpha_i q(t - i/f_s)\right]\left[x_1 - \sum_{i=1}^{N} \alpha_i x_1(t - i/f_s)\right]}$$

$$= \sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \alpha_j R_{x_1}((i-j)/f_s) + \sum_{i=0}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j R_q((i-j)/f_s)$$

$$+ 2\sum_{i=0}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j R_{x_1 q}((i-j)/f_s) \tag{2.8}$$

where $R_{x_1}(\tau)$ and $R_q(\tau)$ are the autocorrelation functions of $x_1(t)$ and $q(t)$ respectively and $R_{x_1 q}(\tau)$ is the cross-correlation of $x_1(t)$ and $q(t)$. Let $\phi_{x_1}(\tau) = R_{x_1}(\tau)/\overline{x_1^2}$, $\phi_q(\tau) = R_q(\tau)/\overline{x_1^2}$, $\phi_{x_1 q}(\tau) = R_{x_1 q}(\tau)/\overline{x_1^2}$ and $X*(f) = \dot{X}(f)/\overline{x^2}$. Therefore, equation

(2.4) becomes,

$$\frac{\overline{\varepsilon^2}}{\overline{x^2}} = 2\int_W^\infty X*(f)df + f_Q\int_{-W}^W X*(f)df\left\{\sum_{i=0}^{N} \sum_{j=0}^{N} \alpha_i \alpha_j \phi_{x_1}((i-j)/f_s)\right.$$

$$\left. + \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \phi_q((i-j)/f_s) + 2\sum_{i=0}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j \phi_{x_1 q}((i-j)/f_s)\right\} \tag{2.9}$$

When $L \geq 3$, $\overline{q^2} \ll \overline{e^2}$ and $|\phi_q(\tau)|$ and $|\phi_{x_1 q}(\tau)|$ are much smaller than $|\phi_{x_1}(\tau)|$ for most τ. Under this constraint, if f_Q is not strongly dependent on the α_i's, then $\overline{\varepsilon^2}/\overline{x^2}$ is minimized with respect to the α_i's by choosing them to be a solution of the following set of linear equations [29].

$$\phi_{x_1}(i/f_s) = \sum_{j=1}^{N} \alpha_j \phi_{x_1}((i-j)/f_s) \qquad (2.10)$$

This choice of the α_i's results in

$$\frac{\overline{\varepsilon^2}}{\overline{x^2}} = 2 \int_{W}^{\infty} X*(f)df + f_Q \left(\int_{-W}^{W} X*(f)df \right) \left(1 - \sum_{i=1}^{N} \alpha_i \phi_{x_1}\left(\frac{i}{f_s}\right) \right) \qquad (2.11)$$

When the input to the quantizer is a speech waveform, the quantizer is usually constrained to be logarithmic. When μ, the logarithmic quantizer parameter, is large compared to the ratio of the peak to root-mean-square value of the input, the distortion is largely independent of the input signal statistics [4] . Under this constraint $f_Q \simeq \frac{1}{3} [\log(1+\mu)]^2 4^{-L}$ for $L \geq 3$ and $f_s = 2W$. Therefore equation (2.11) becomes

$$\frac{\overline{\varepsilon^2}}{\overline{x^2}} = 2 \int_{W}^{\infty} X*(f)df + (1/3)[\log(1+\mu)]^2 \int_{-W}^{W} X*(f(df(1+\sum_{i=1}^{N} \alpha_i \phi_{x_1}(i/f_s))4^{-L}$$

$$(2.12)$$

When the optimum $L \lesssim 3$, the assumptions $\left| \phi_{x_1 q}(\tau) \right| \ll \left| \phi_{x_1}(\tau) \right|$ and $\left| \phi_q(\tau) \right| \ll \left| \phi_{x_1}(\tau) \right|$ no longer apply with the result that evaluation of $\overline{\varepsilon^2}$ becomes difficult.

2.2 The Effect of Sampling at Higher than the Nyquist Rate

Sampling at frequencies greater than the Nyquist rate has three effects: Aliasing errors and idle channel noise [14] caused by non-ideal lowpass filters are reduced; an improved prediction of the sampled input results; and the amount of quantizer noise lying in the passband of the receiver lowpass filter is reduced. By restricting the minimum sampling frequency to be 2.2 times the 3 db cutoff frequency of the input lowpass filter, the effects of aliasing

errors and idle channel noise are largely eliminated.

To see that an increase in sampling frequency increases the predictability of the sampled input, consider the normalized autocorrelation function of speech in Figure 4.2. It is seen that an increase in sampling frequency results in an increase in correlation between the input and its delayed replica provided $f_s \gtrsim$ 1KHz, which is virtually always the case. In Appendix I it is shown that such an increase in correlation results in an improved prediction of the input signal. This correspondingly means a reduction in mean square input to the quantizer which, by equation (2.7), indicates a reduction in received quantization noise.
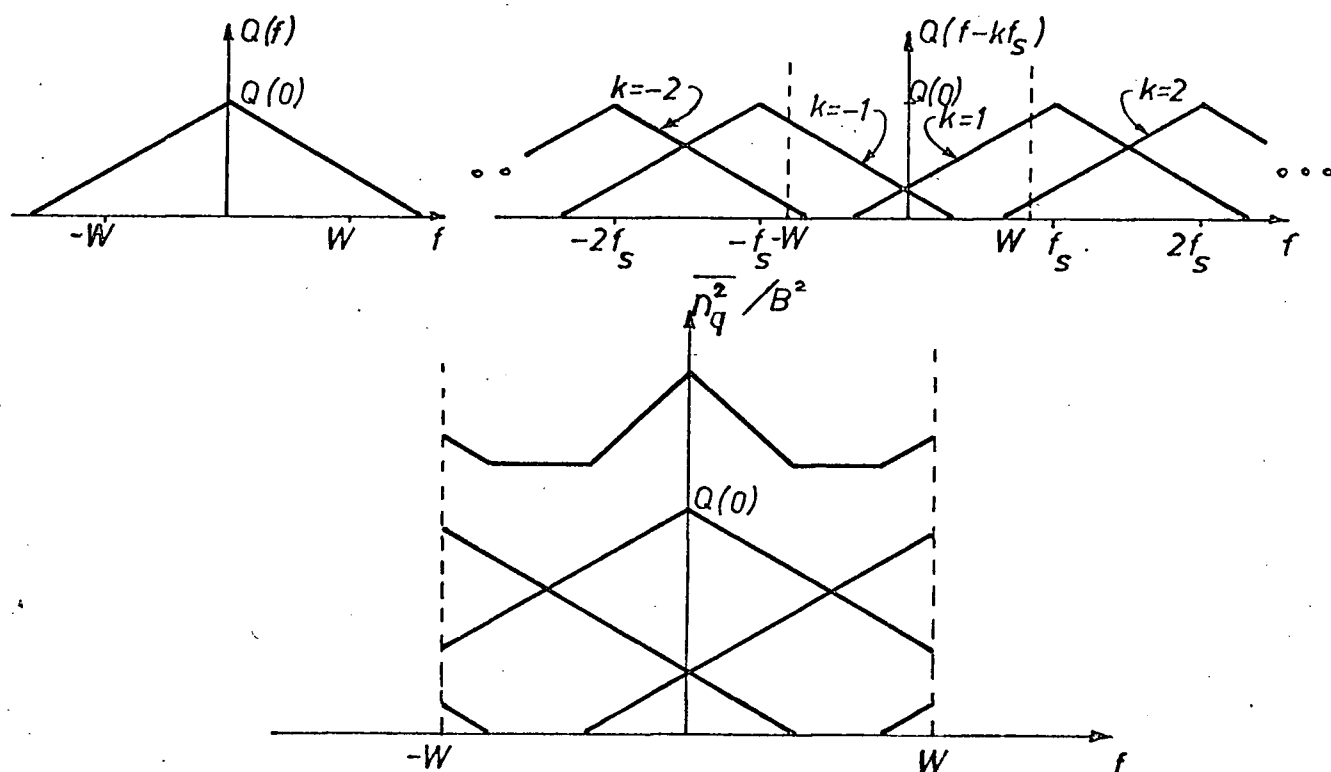
Fig. 2.2 (a) Spectrum of quantization noise. (b) Quantizing noise aliasing due to sampling. (c) Output quantization noise. Note that increasing f_s decreases the contributions due to aliasing.

In order to illustrate the third effect consider equation (2.7) reproduced here in part for convenience.

$$\overline{n_q^2} = B^2 \sum_{k=-\infty}^{\infty} \int_{-W}^{W} Q(f - kf_s)df \qquad (2.13)$$

Since the bandwidth of the quantization noise exceeds the signal bandwidth [28,29], an increase in f_s results in a reduction in the total quantization noise distortion in the system output (see Figure 2.2). Note that if $f_s = 2W$, equation (2.13) becomes

$$\overline{n_q^2} = B^2 \int_{-\infty}^{\infty} Q(f)df = \overline{q^2}B^2$$

2.3 Signal-to-Noise Ratio in AM Communication

The essential features of a double-sideband suppressed carrier amplitude modulation system (DSB-SC) appear in Figure 2.3(a). The lowpass filter $H(f)$ is assumed to be ideal with characteristic as illustrated in Figure 2.3(b). The modulating frequency $f_1 = \omega_0/2\pi$ is restricted to being greater than $2W$. The channel noise $n_w(t)$ is assumed to be additive white Gaussian with power density spectrum

$$S_{n_w}(f) = N_0/2 \qquad -\infty < f < \infty$$
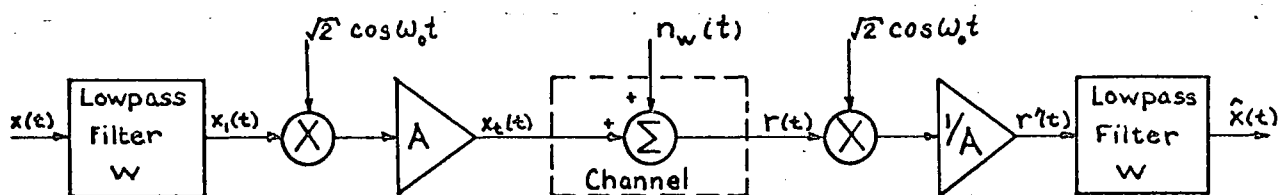
(a)
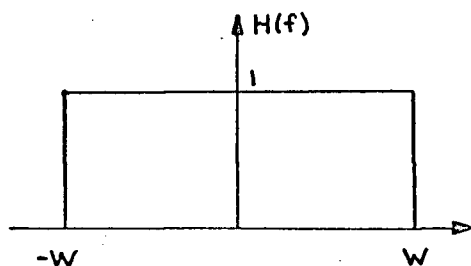
(b)
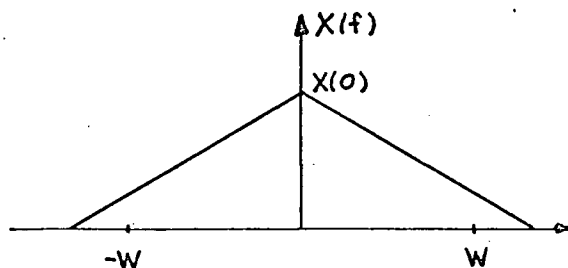
(c)

Fig. 2.3 (a) A double sideband suppressed-carrier amplitude modulation system. (b) Ideal lowpass filter characteristic (c) Assumed power spectrum of the baseband input signal.

Let $X(f)$, the power spectrum of the baseband input signal $x(t)$, be as shown in Figure 2.3(c).

The mean square error in the output is

$$\overline{e^2} = \overline{(x(t) - \hat{x}(t))^2}$$

$$= \overline{(x - x_1)^2} + \overline{n^2} - \overline{2(x - x_1)n} . \qquad (2.14)$$

The last term in (2.14) is zero. Also

$$\overline{(x - x_1)^2} = 2 \int_W^\infty X(f)df . \qquad (2.15)$$

The power spectrum of $2\cos\omega_0 t \cdot n_w(t)$ is the convolution of $S_{n_w}(f)$ with two impulses of magnitude $1/2$ placed at $\pm f_0$, and is equal to $S_{n_w}(f)$. Therefore, the second term on the right side of (2.14) becomes

$$\overline{n^2(t)} = \frac{1}{A^2} \int_{-W}^W \frac{N_0}{2} df = \frac{N_0 W}{A^2} \qquad (2.16)$$

Combining (2.14), (2.15) and (2.16) yields

$$\overline{e^2(t)} = 2 \int_W^\infty X(f)df + \frac{W N_0}{A^2} \qquad (2.17)$$

Since the transmitter signal power is

$$P_s = \overline{x_t^2} = \int_{-\infty}^\infty X_t(f)df$$

$$= A^2 \int_{-W}^W X(f)df$$

the inverse signal-to-noise ratio for AM communications becomes

$$\left(\frac{N}{S}\right)_{AM} = \frac{\overline{e^2}}{\overline{x^2}} = 2 \int_W^\infty X*(f)df + \frac{N_0 W}{P_s} \int_{-W}^W X*(f)df \qquad (2.18)$$

where $\displaystyle\int_{-\infty}^{\infty} X^*(f)df = 1.$

2.4 Signal-to-Noise Ratio in Angle Modulation Systems*.

Angle modulation systems are divided basically into two categories; phase modulation (PM) and frequency modulation (FM). In the following analysis, phase modulation is presented as a special case of frequency modulation. The difference between the two is illustrated in Figure 2.4.

(a)

(b)
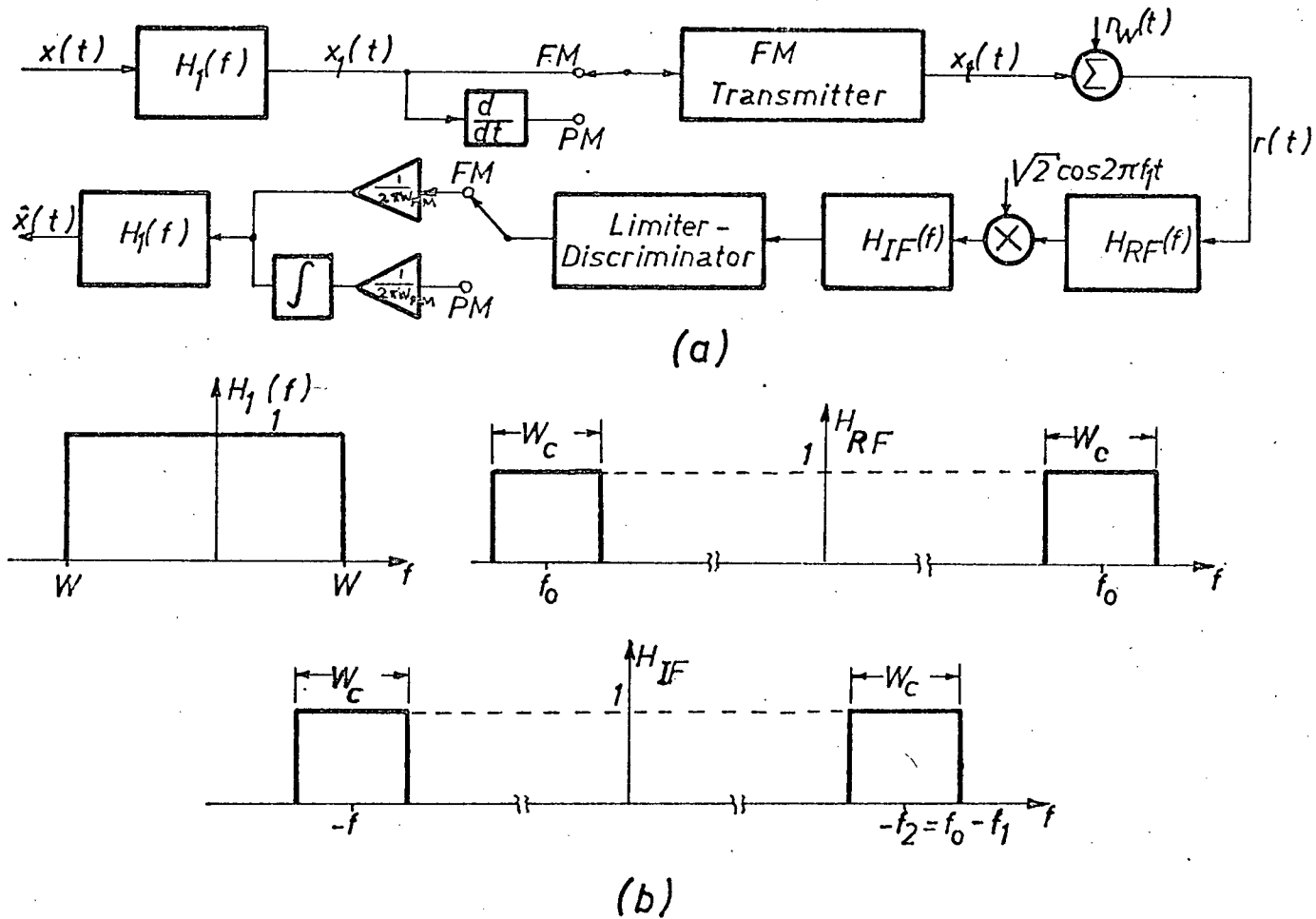
Fig. 2.4   (a) Angle modulation systems. (b) Characteristics of the ideal RF (H_{RF}), IF (H_{IF}) and lowpass (H_1) filters.

The frequency characteristics of the lowpass filter $H_1(f)$, the radio frequency (RF) bandpass filter $H_{RF}(f)$ and the intermediate frequency (IF) filter $H_{IF}(f)$ are also shown in Figure 2.5. Again,

---

* The following analysis is necessarily very simplified. For a more complete analysis of angle-modulation systems the reader is referred to Sakrison [31].

the filters are assumed ideal, and the channel noise is assumed to be additive with a spectrum which is uniform over the channel bandwidth. The FM transmitter generates from the lowpassed input an output

$$x_t(t) = A\sqrt{2}\cos 2\pi(f_o t + W_{FM}\int x_1(t)dt)$$

where $W_{FM}|x_1(t)|_{max}$ is the peak frequency deviation. The received signal $r(t)$ is first bandpassed through $H_{RF}(f)$, multiplied by $\sqrt{2}\cos 2\pi f_1 t$ and then passed through $H_{IF}(f)$. In the absence of noise, the input to the limiter-discriminator is then

$$f_1(t) = A\cos 2\pi\left[f_2 t + W_{FM}\int x_1(t)dt\right] \quad .$$

With $r_1(t)$ as its input, the limiter-discriminator, a device which extracts and differentiates the instantaneous phase of its input, produces an output

$$r_2(t) = d/dt\left[2\pi W_{FM}\int x_1(t)dt\right]$$

$$= (2\pi W_{FM})x_1(t) \quad .$$

After passing through an attenuator of gain $1/2\pi W_{FM}$, the output of the demodulator for noise-free transmission is $x_1(t)$, the lowpassed input signal.

If now the assumption of noiseless transmission is removed, and if the input to the frequency modulator is assumed constant for a short period of time*

$$x_1(t) = x_o, \qquad\qquad -1 < x_o < 1$$

then the input to the limiter-discriminator becomes

$$r_1(t) = a(t)\cos(2\pi(f_2 + W_{FM}x_o)t + \Theta(t))$$

---

* This amounts to approximating the message by a series of rectangular pulses of width Δ. As Δ becomes small compared to $1/2W$ the approximation becomes more exact.

where $a(t) \triangleq \sqrt{(A + n_c(t))^2 + n_s^2(t)}$

and $\theta(t) \triangleq \tan^{-1}(-n_s(t)/(A + n_c(t)))$.

In these equations $n_c(t)$ represents that component of the noise in phase with $\cos(2\pi(f_2 + W_{FM}x_o)t)$ and $n_s(t)$ that in phase with $\sin(2\pi(f_2 + W_{FM}x_o)t)$. The power density spectra of $n_c(t)$ and $n_s(t)$ are equal to the even part of $S_{n_1}(f - f_2 - W_{FM}x_o)$ [25]. If one assumes that $2N_oW \ll A^2$, then the approximation $\theta(t) \cong -n_s(t)/A$ is valid except for certain improbable and hence infrequent instances of time. Recalling the effect of the limiter-discriminator, it may now be seen that the output of the system after attenuation and low-pass filtering is

$$\hat{x}(t) = x_1(t) + n_o(t) \ .$$

The power density spectrum of $n_o(t)$ is given by

$$S_{n_o}(f) = (1/2\pi A W_{FM})^2 S_{n_s}(f)\left|H_1(f)\right|^2\left|j2\pi f\right|^2 \ . \tag{2.19}$$

It follows that the noise power is concentrated in the higher frequencies. Consider now the power spectrum of the noise output of a phase modulation system;

$$S_{n_{oPM}}(f) = (1/2\pi A W_{FM})^2 S_{n_s}(f)\left|H_1(f)\right|^2\left|j2\pi f\right|^2\left|1/j2\pi f\right|^2 \tag{2.20}$$

Note that this power spectrum is uniform across the message bandwidth. From (2.19) and (2.20), the output noise powers of FM and PM are

$$\overline{n_o^2} = (1/W_{FM})^2(N_oW/A^2)(W^2/3)$$

and $$\overline{n_{oPM}^2} = (1/2\pi W_{PM})^2(N_oW/A^2) \ .$$

The transmitter power is given by

$$P = \overline{x_t{}^2(t)} = A^2$$

Combining the above equations yields the inverse signal-to-noise ratios for FM and PM:

$$\left(\frac{N}{S}\right)_{FM} = 2\int_W^{\infty} X^*(f)df + \left(\frac{1}{\overline{x^2}}\right)\left(\frac{1}{3}\right)\left(\frac{W}{W_{FM}}\right)^2 \left(\frac{N_o W}{P}\right)\int_{-W}^{W} X^*(f)df \qquad (2.21)$$

$$\left(\frac{N}{S}\right)_{PM} = 2\int_W^{\infty} X^*(f)df + \left(\frac{1}{\overline{x_1^2}}\right)\left(\frac{1}{2\pi W_{PM}}\right)^2 \left(\frac{N_o W}{P}\right)\int_{-W}^{W} X^*(f)df \qquad (2.22)$$

where again $\int_{-\infty}^{\infty} X^*(f)df = 1$.

The channel bandwidth W_c required for angle modulation communication may be written approximately as [30]

$$W_c = |\Delta f|_{max} + 2W \qquad (2.23)$$

where $|\Delta f|_{max}$ is a measure of the maximum instantaneous frequency deviation. The instantaneous frequency deviation for FM is

$$(\Delta f)_{FM} = W_{FM} x_1(t)$$

and for PM is $\qquad (\Delta f)_{PM} = W_{PM} x_1'(t)$.

Since the probability that $|x_1(t)|$ exceeds $c\sqrt{\overline{x_1{}^2(t)}}$ may be made small by choosing the peak factor c large, an approximate expression for W_{cFM} is

$$W_{cFM} = cW_{FM}\sqrt{\overline{x_1{}^2}} + 2W$$

and for PM $\qquad W_{cPM} = cW_{PM}\sqrt{\overline{x_1'{}^2}} + 2W.$

Also,

$$\sqrt{\overline{x_1'{}^2}} = \left(4\pi^2\int_{-\infty}^{\infty} f^2 X_1(f)df\right)^{\frac{1}{2}}$$

$$= 2\pi\sqrt{\overline{x_1{}^2}}\sqrt{\overline{f_m{}^2}}$$

where

$$\sqrt{\overline{f_m^2}} = \left(\frac{\int_{-\infty}^{\infty} f^2 X_1(f) df}{\int_{-\infty}^{\infty} X_1(f) df} \right)^{\frac{1}{2}} \triangleq \text{rms bandwidth of the message}$$

Therefore, (2.21) and (2.22) may be rewritten

$$\left(\frac{N}{S}\right)_{FM} = 2 \int_{W}^{\infty} X^*(f) df + \frac{c^2 W^2}{3(W_{cFM} - 2W)^2} \left(\frac{N_O W}{P}\right) \int_{-W}^{W} X^*(f) df \qquad (2.24)$$

$$\left(\frac{N}{S}\right)_{PM} = 2 \int_{W}^{\infty} X^*(f) df + \frac{c^2 \overline{f_m^2}(W)}{(W_{cPM} - 2W)^2} \left(\frac{N_O W}{P}\right) \int_{-W}^{W} X^*(f) df \qquad (2.25)$$

3. REAL TIME DPCM AND AM SYSTEMS

3.1 A Real Time DPCM System

If it is assumed that the channel noise in Fig. 2.1 is negligible, then the output x(t) of the system shown in Fig. 3.1 is equivalent to the output x(t) of the system in Fig. 2.1. A practical realization of the system in
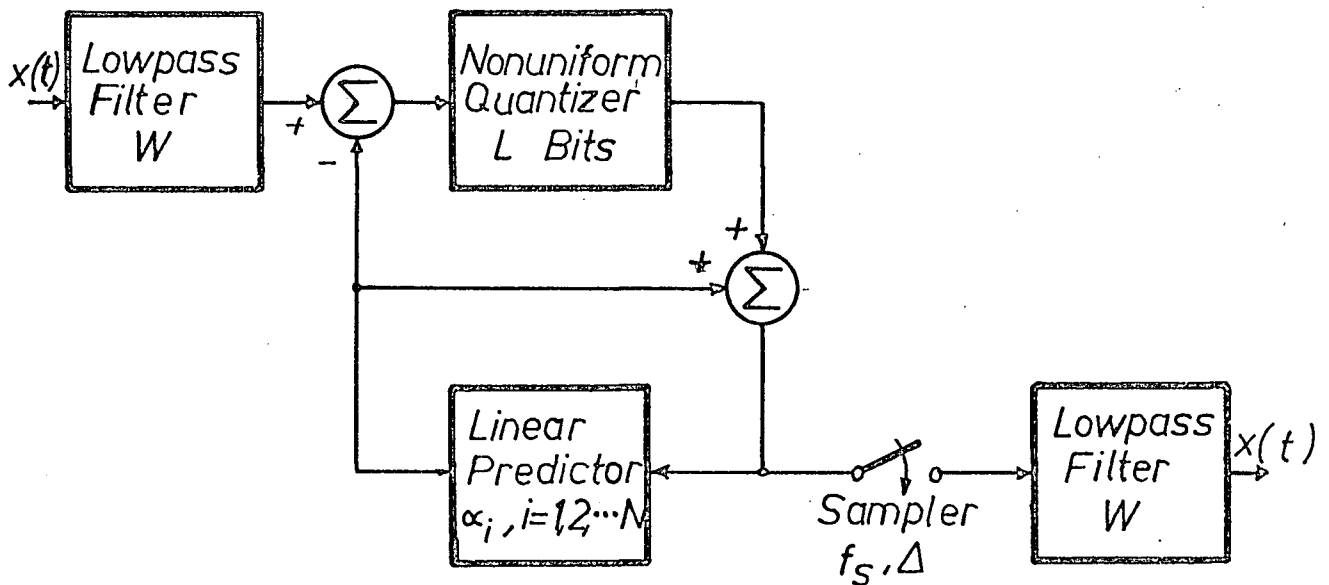
Fig. 3.1 A system equivalent to Fig. 2.1 when channel noise is neglected.

Fig. 3.1 for N = 3 is shown in Fig. 3.2. Only the general opera-tion of this system will be described here. For details of the cir-cuits used, the reader is referred to Chan [32].

Assume the system is initially in its quiescent state. The arrival of a clock pulse initiates the timing sequence; reset, pulse 1, pulse 2, ... pulse 8 (only to pulse 3 if $f_s \gtrsim 21.5$ KHz).

The reset pulse actuates the input sample and hold (S & H) and the output feedback S & H's. The input S & H samples the low-pass filtered speech and stores the current value. At the same
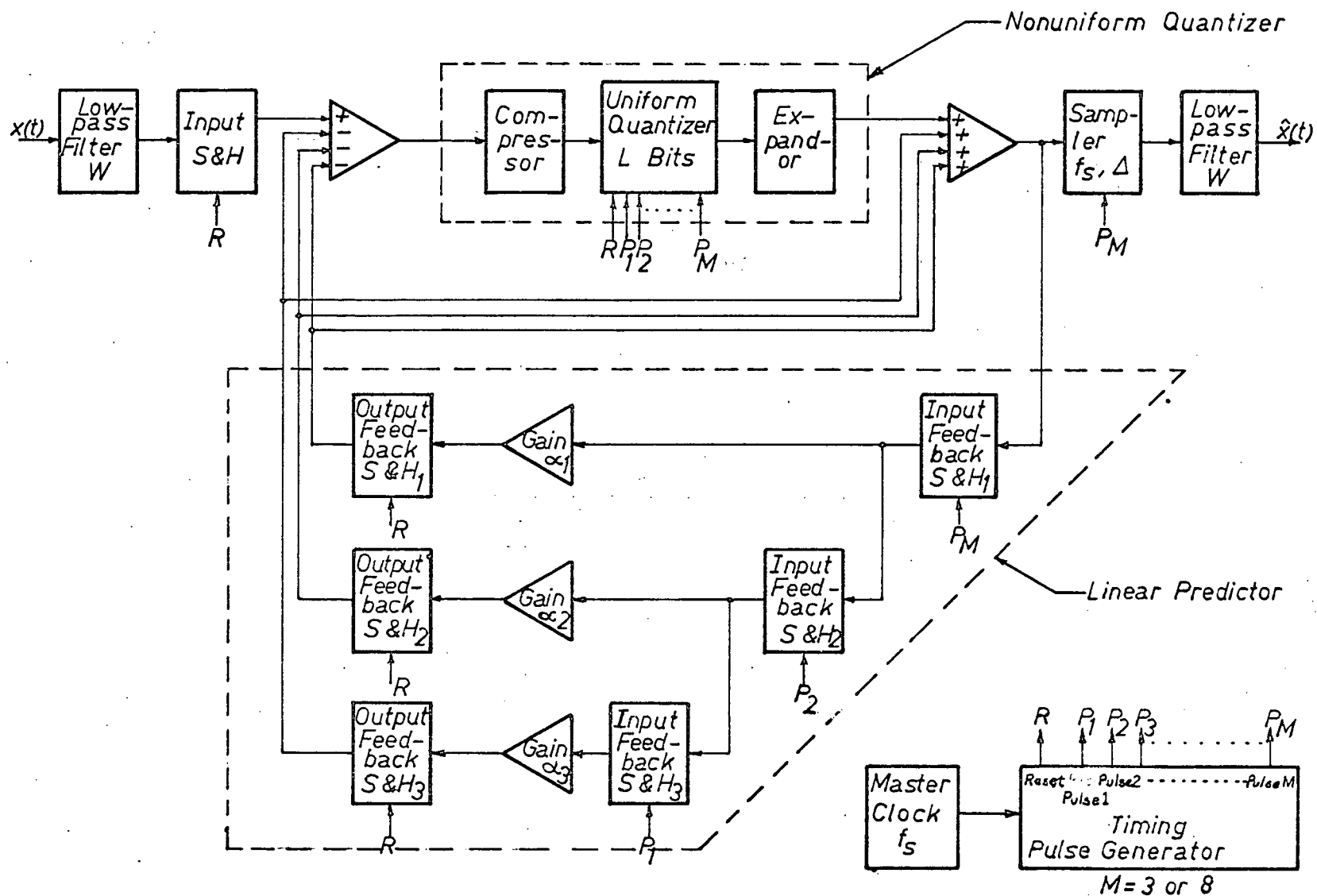
Fig. 3.2   An implementation of Fig. 3.1 for N = 3.

time, the values currently being held by the input feedback S &
H's are multiplied by the prediction coefficients α_1, α_2, and α_3
(see Appendix I), sampled and stored in the corresponding output
feedback S & H's. The sum of the outputs of the three output feed-
back S & H's is now an estimate of the current input sample.

The actual and predicted values are now subtracted and
their difference nonuniformly quantized. The values currently being
held in the first and second of the string of output feedback S & H's
are now shifted to the second and third feedback S & H's respec-
tively. The quantized difference signal is added to the predicted
value and the sum is stored in the first of the string of input
feedback S & H's. The sum is also sampled and lowpass filtered.
The resultant lowpass filter output $x(t)$ is then a replica of the
original speech signal $x(t)$.

The system parameters W, f_s, L and $\alpha_i (i = 1,2,3)$ are
variable, as are the compressor and expandor characteristics.

3.2 Disadvantages of the System Used

Consider the integrator network in Fig. 3.3 where $n_1(t)$
is a noise input which may be considered to represent noise caused
by the delaying and attenuating circuitry. For $y(t) = 0$, the power
spectrum of the output is given by

$$X_0(f) = \frac{N_1(f)}{1 - \alpha^2} \, ,$$

where $N_1(f)$ is the power spectrum of $n_1(t)$. It follows that
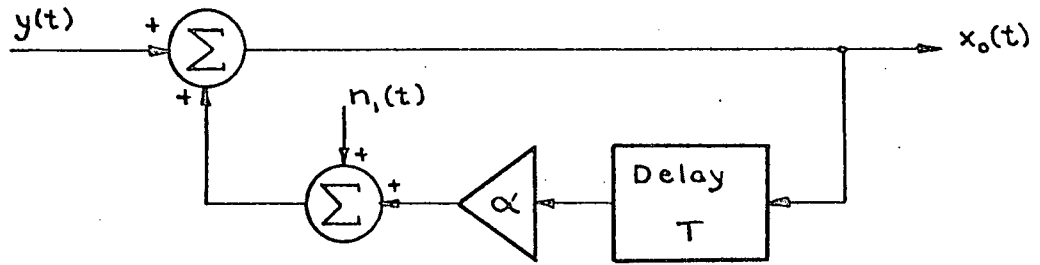
Fig. 3.3 An integrator network.

for α approaching unity, any small amount of noise generated within the predictor will be amplified greatly.

Consider now the complete system of Fig. 3.2. As shown in Appendix I, the sum of α_1, α_2, and α_3 approaches unity. Except for the fact that the noise is now being generated by three attenuation and delay networks, the situation is very similar to that discussed above when $N = 1$, except that the predictor noise, after being quantized, should be subtracted from the output of the output summing amplifier. If the quantizer has very few bits, the quantization noise may be very large, and virtually uncorrelated with the predictor noise. As a result, the feedback noise will still be heard at the output, along with large amounts of quantizer and idle channel noise [14] . Normally, a small DC bias is added to the input signal with the resultant idle channel dominant mode oscillation being removed by the output lowpass filter. However, if the quantization is fine, the small DC bias is overcome by predictor noise, with the result that the idle channel noise may not be eliminated by the output lowpass filter.

As a consequence of this noise buildup, it was extremely difficult to obtain sufficiently reliable results when two and three

samples of feedback were used. Therefore, the experiments described later were conducted for only the previous sample feedback case.

3.3 Measurement of AM Signal-to-Noise Ratio

The system shown in Fig. 3.4 was used to simulate an Am communications system. For $n_W(t)$ white Gaussian noise, the output
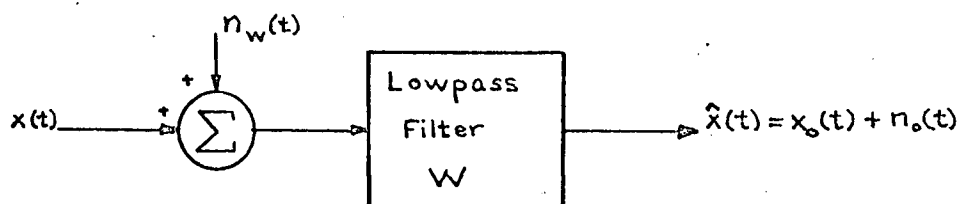
Fig. 3.4 A system whose output is equivalent to the DSB-SC system in Fig. 2.3(a) for A = 1.

of this system is identical to the output of the DSB-SC system shown in Fig. 2.3 with, of course, an appropriate scale factor to account for the gain A in Fig. 2.3.

The signal-to-noise ratio at the output of this system may be written

$$\frac{S}{N} = \frac{\overline{x_o^2(t)}}{\overline{n_o^2(t)}}$$

where $\overline{}$ denotes a time average over all time. The speech sample used for the tests described in Chapter IV may be written

$$x(t) = \begin{cases} m(t) & 0 \leqslant t \leqslant T \\ 0 & \text{elsewhere} \end{cases}$$

Therefore, the time average power of the received message is approximated as follows:

$$\overline{x_o^2(t)} = \tfrac{1}{T} \int_0^T m^2(t)\, dt.$$

A similar equation may be used to estimate the received noise power. Although the definition of the mean specifies an integral over all time, restricting the integral to a finite interval of time does not introduce a large error if the interval is large·compared to the duration of the autocorrelation function of the input [33].

The calculation of the time average power was performed on an EAI PACE 231R Analogue Computer using the circuit shown·in Fig. 3.5. The output of this system is

$$v_o(t) = 10,000\alpha^2 \int_0^t v_i^2(\lambda)d\lambda$$

The mean square value of the input is then given by

$$\overline{v_i^2} = \frac{v_o(T)}{10,000\alpha^2 T}$$
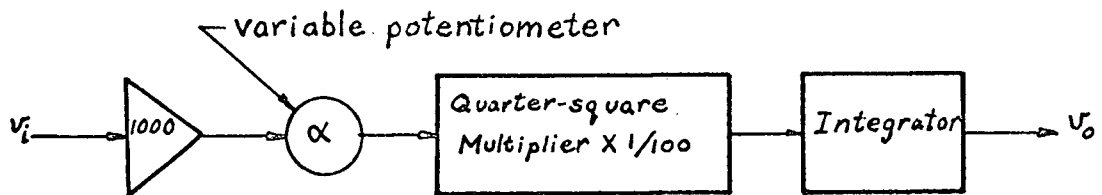
Fig. 3.5 Circuit for measurement of mean square power.

Since the speed of tape on which the sentence was recorded was $7\tfrac{1}{2}$ inches per second, the duration of the speech sample T_s could be calculated by measuring the length of the tape on which the sample was recorded. This length was found to be $15^{\pm}.5$ inches resulting in $T_s=2^{\pm}.06$ seconds. The

measurements in all cases were averaged over five repetitions to minimize the effect of any random disturbances or timing inaccuracies.

The resultant signal-to-noise ratio was calculated as follows:

$$\frac{S}{N} = \frac{v_{os}(T_s)}{10,000\alpha_s^2 T_s} \cdot \frac{10,000\alpha_n^2 T_n}{v_{on}(T_n)}$$

$$= 7.5 \pm 6\% \; \frac{\alpha_n^2}{\alpha_s^2} \; \frac{v_{os}(2)}{v_{on}(15)}$$

where the subscripts s and n refer to the speech and noise respectively.

4. SUBJECTIVE TEST PROCEDURE

4.1 Introduction

Many methods have been developed to quantitatively scale perceptual stimuli [34]. The results presented in Chapter 5 are based on a modification of the paired comparison method [35]. This method was used to obtain equal preference (isopreference) contours, and these contours were then scaled using a version of the subjective-estimate method.

Although either method could have been used to both determine and to rate the isopreference contours, more reliable results are obtained by using a composite of the two. A rating scale derived by use of the paired comparison method is unreliable whenever the variability of the listeners' judgements is not substantial. The subjective-estimate method requires listeners to judge how much better one stimulus is than another. This method also requires the experimenter to extrapolate between the scale values of the test points, since the rated points may not necessarily lie on any particular isopreference contour. For these two reasons, the subjective-estimate method does not yield reliable isopreference contours.

The master sentence used throughout the tests was recorded in an anechoic chamber using a General Radio Type 1560-P3 PZT microphone and a Tandberg 64X tape recorder. The sentence "Joe took father's shoe bench out" was chosen as the master sentence since it contains most of the phonemes and has a frequency spectrum which is representative of conversational speech [36]. The sentence was spoken by a 28 year old male with a Western Canadian accent. An estimate of the speaker's spectrum obtained using the method described in Section 3.3 appears in Figure 4.1.
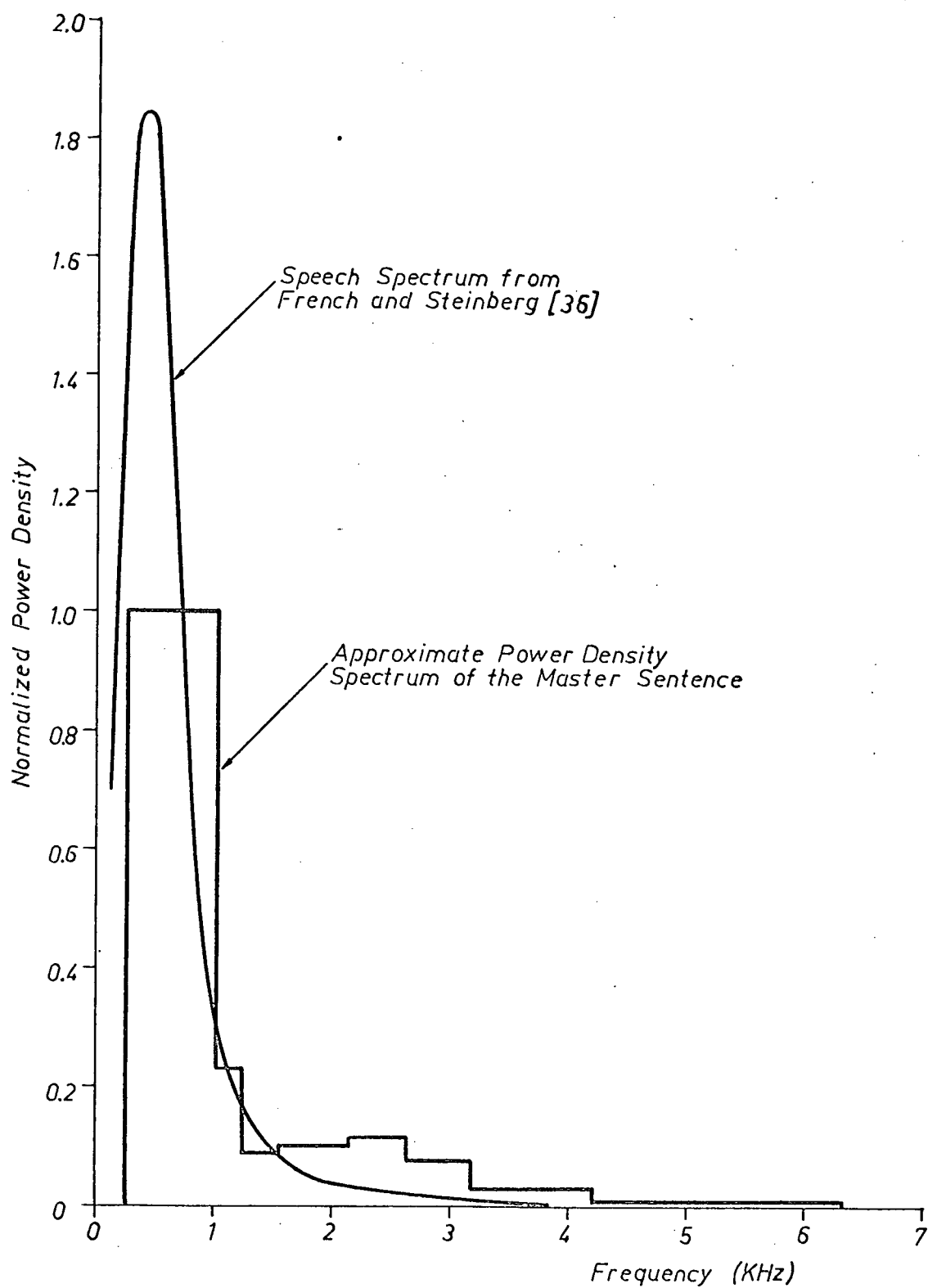
Fig. 4.1 Power density spectra of speech.

4.2   Preparation of Speech Samples

The speech samples were obtained by playing back the master sentence through the appropriate system (either DPCM or AM) and re-recording on a second Tandberg 64X tape recorder. The samples were then spliced together along with appropriate lengths of non-magnetic tape. In order to eliminate hum present at the tape recorders' outputs, all tape playbacks were high-pass filtered to approximately 200 Hz.

4.2.1   DPCM Samples

In the preparation of the DPCM samples, the following assumptions and restrictions were imposed:

(1) The digital channel noise shown in Figure 2.1 equalled zero.

(2) The nonuniform quantizer was constrained to be logarithmic with $\mu = 100$. Panter and Dite[4] have shown that with logarithmic compression, the resultant distortion is relatively independent of the input statistics assuming the peak value of the input does not exceed the maximum quantizer input.

(3) The amplitude of the input signal was always adjusted until the quantizer input occupied the full available range of the quantizer.

(4) For minimum mean square error, the prediction coefficient α_1 was determined from equations (2.10) to be equal to the normalized autocorrelation of the speech signal evaluated at the sampling period. However, equations (2.10) presuppose relatively fine quantization as well as an exact knowledge of the autocorrelation function of the speech sample being processed. As a result, the approximate relationship for α_1 shown in Fig. 4.2 was used. Also shown in this figure are autocorrelation func-
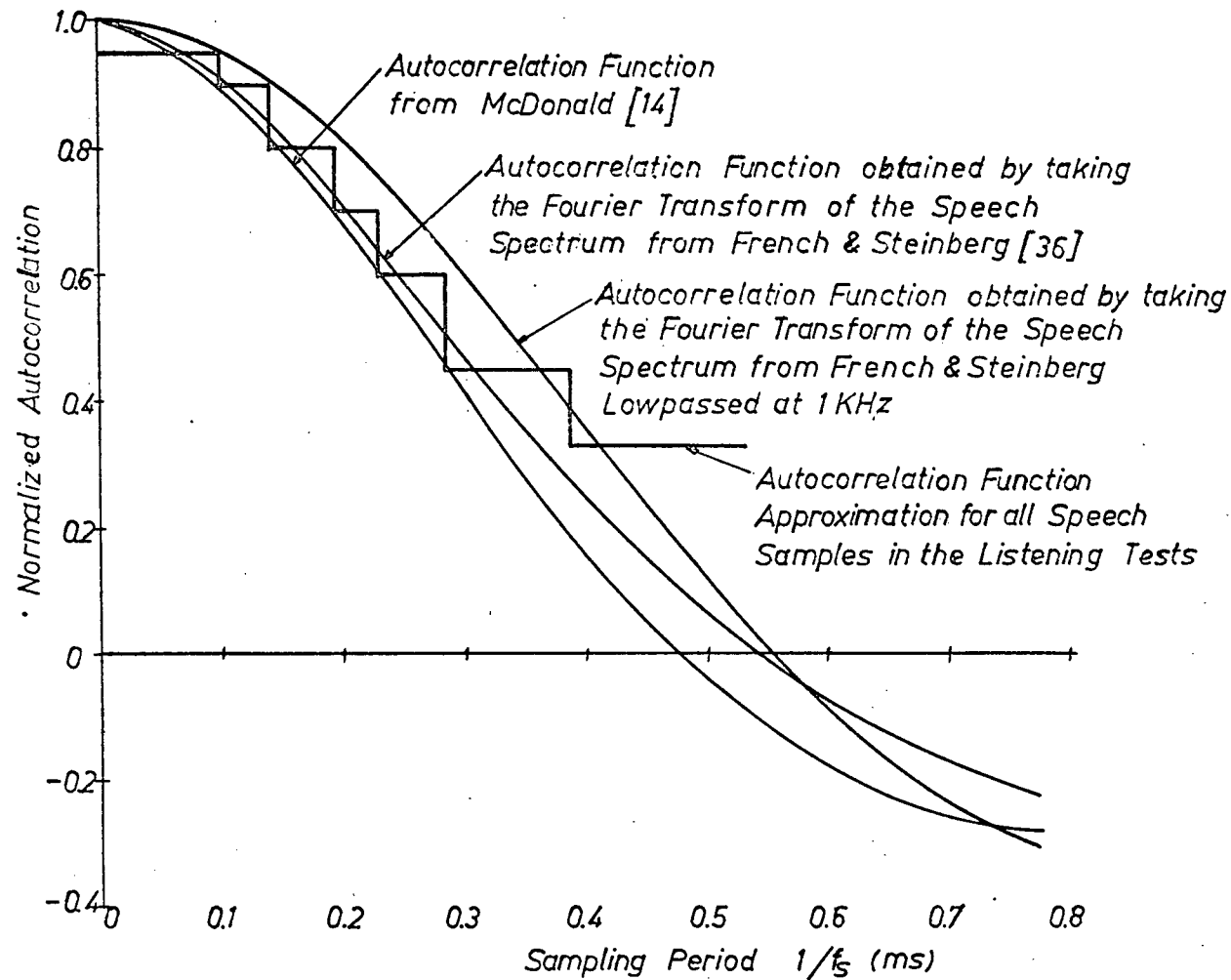
Fig. 4.2 Autocorrelation functions of lowpass filtered speech versus the sampling period $1/f_s$.

tions of speech based on other data, as well as an estimate of the autocorrelation function of the master sentence used. An analysis of the error encountered using the approximation in Figure 4.2 appears in Appendix I.

(t) In order to eliminate aliasing errors and idle channel oscillations [14], the minimum permitted sampling frequency was restricted to 2.2 times the 3 db cutoff of the lowpass filter.

(6) The speech bandwidth W could assume only the discrete values of 1.01, 1.21, 1.55, 2.12, 2.63, 3.17, 4.2 and 6.3 KHz.

4.2.2 AM Samples

In the preparation of the AM samples, it was assumed that the operations of modulation and demodulation were completely noiseless and that the channel noise was flat. The spectrum of the output of the Grason-Stadler noise generator used was measured on a wave analyzer and was found to be flat to within ±0.4 db over the frequency range 200 Hz $\leqslant$ f $\leqslant$ 6.3 KHz. The magnitude of the master sentence input was adjusted so that the resultant integrated value of the output of the system introduced in Section 3.3 was equal to 100 $\pm$1 volts.

4.3 Paired Comparison Tests for Determining Isopreference Contours
4.3.1 Paired Comparison Tests

The paired comparison tests were conducted over periods spanning five days for DPCM and two days for AM. Two sessions took place each day, one in the morning and one in the afternoon. The ten listeners present during any given session were selected from a group of 28 graduate students. Most listeners sat for fewer than eight sessions. The subjects' ages ranged from 21 to 30 with a mean of approximately 25. The tests were conducted in a quiet room

using binaural hearing with Pioneer model SE-1 stereo headphones.

At the beginning of each session, the listeners were given response forms with the following set of written instructions:

"In this test, you will hear pairs of sentences; each pair is separated by a 5 second rest period. After listening to a pair, specify which sentence you would prefer to hear. If both sentences sound equally good, make an arbitrary choice. The first sentence of each pair is sentence A, and the second, sentence B." Sentences A and B were separated by a one second silence. During each session the listeners were required to compare between seventy and eighty pairs of sentences with a two to three minute rest period after every twenty comparisons. The entire set of comparisons (AM or DPCM) were heard in random order. Each comparison appeared twice, with the order of the comparison in the second test being reversed from that in the first. Most listeners were acquainted with the speaker of the sentence.

4.3.2   The method of Paired Comparisons

The method of paired comparisons is based solely on the subjects' ability to judge which of two conditions he prefers. The method makes four basic assumptions: the sample of subjects is chosen from a normal population; any previous paired comparison tests have negligible effect on the test in progress; the variable parameter is available as an underlying continuum; and the judgements are transitive. The latter property is discussed in Section 4.5.

As an example of the use of the method of paired comparisons to develop isopreference contours, consider the points marked A and B in Fig. 4.3. The sample corresponding to point A is compared to samples corresponding to the points marked b_i (i = 1,2,3,4), each of which has the same abscissa value X_B. The results of the compar-

isons are plotted on a psychometric curve as shown in Fig. 4.4(a).
The ordinate of this figure shows the percentage of judgements pre-
ferring sample A to samples having a value of the X parameter equal
to X_B and the values of Y indicated in the plot. The Y_i (i = 1,2,3,4)
are selected such that the percentage preferring A varies from 0 to
100%. A smooth curve is then drawn through the experimental points
and the 50% or equal preference point Y_B plotted as the ordinate of
point B.
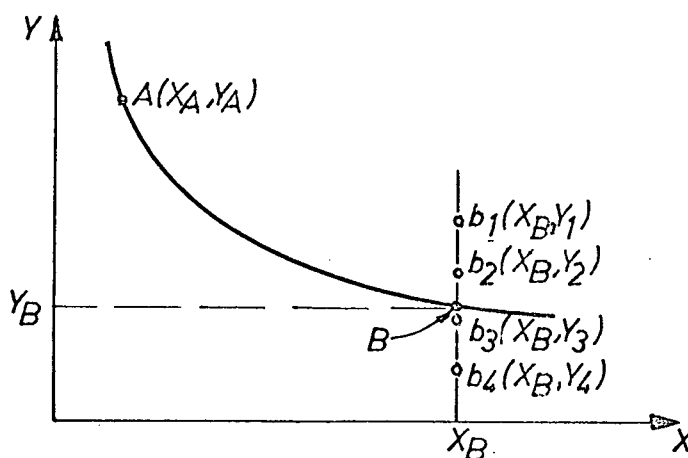
Fig. 4.3   An isopreference curve. Points b_i (i = 1,--4) are
compared to point A and the results of the comparisons
used to determine point B.

Points A and B are now assumed to be equal in preference.
Either point could therefore be used to obtain further isopreference
points. However, the discrete nature of the values of parameters
(W and L) used in the listening tests dictated the use of the same
reference point to obtain all of the isopreference points on any
one contour.

In plotting the psychometric curves, it was found that a
normal distribution curve fitted the data points. The proportion
of listeners preferring the reference sentence was therefore con-
verted to unit normal deviates. Since unit normal deviates corres-
ponding to 0 and 100% are infinite, these values were changed to
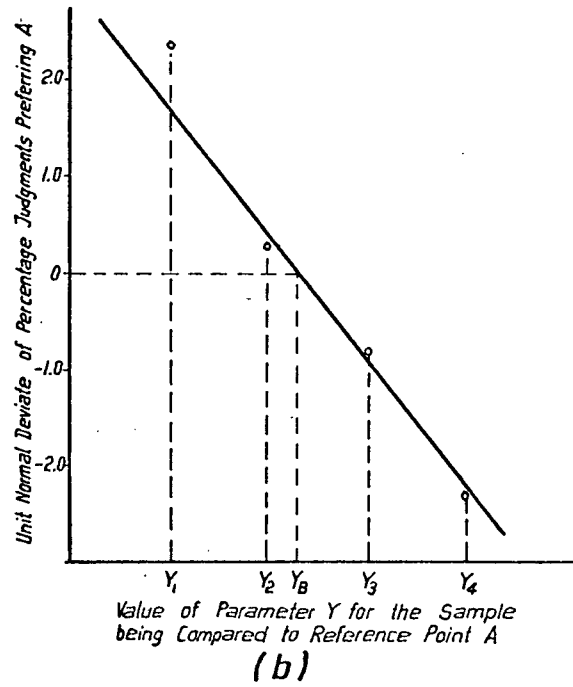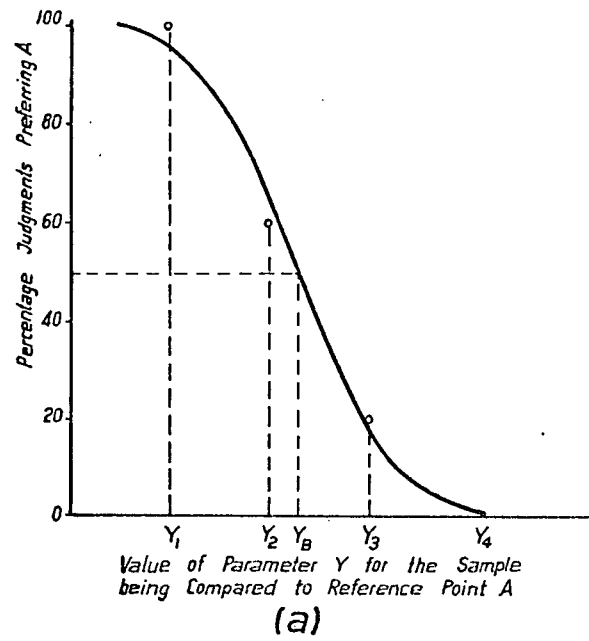
Fig. 4.4   (a) A psychometric curve corresponding to the
derivation of point·B in Fig. 4.3.
(b) The curve of (a) with the ordinate in unit
normal deviates.

0.5 and 99.5% before being converted. Using a weighted least squares technique, a straight line was fitted to the data points. The weight attached to each deviate Y_i was given by [16]

$$W_i = N_i e^{-Y_i^2} / 2\pi p_i (1-p_i)$$

where N_i is the number of judgements on which Y_i is based and p_i is the proportion of judgements preferring the reference sentence. The psychometric curve of Figure 4.4(a) is shown plotted in Figure 4.4(b) in unit normal deviates. The 50% point of Figure 4.4(a) corresponds to zero unit normal deviates. The reciprocal of the slope of the line is equal to the standard deviation σ of the points fitted by the line. The standard deviation associated with each point obtained is indicated by the length of the straight line through the point.

4.3.3 Selection of Test Points and Derivation of Isopreference
Contours

Since, in the DPCM case, the number of available independent parameters was three, it was necessary to obtain isopreference surfaces. The obvious approach was to select a series of planes perpendicular to one of the axes, and to obtain isopreference contours within these planes. These curves could then be extrapolated from plane to plane to obtain isopreference surfaces. However, this procedure has the disadvantage that, for each isopreference surface, only one contour passes through the reference point. Since reference points are the most reliable points available on any contour, it is also desirable that the reference point be situated towards the midpoint of each contour to eliminate the effect of "pivoting" a contour about one of its end points.

By deriving contours in the five planes indicated in

Figure 4.5, the above disadvantages were largely eliminated. The five planes are defined by the following equations:

plane A:- $r = \text{const} = r_o$

plane B:- $W = \text{const} = W_o$

plane C:- $L = \text{const} = L_o$

plane D:- $r = kW \qquad k = r_o/W_o$

plane E:- $r = mW + b \qquad b = r_o - mW_o$
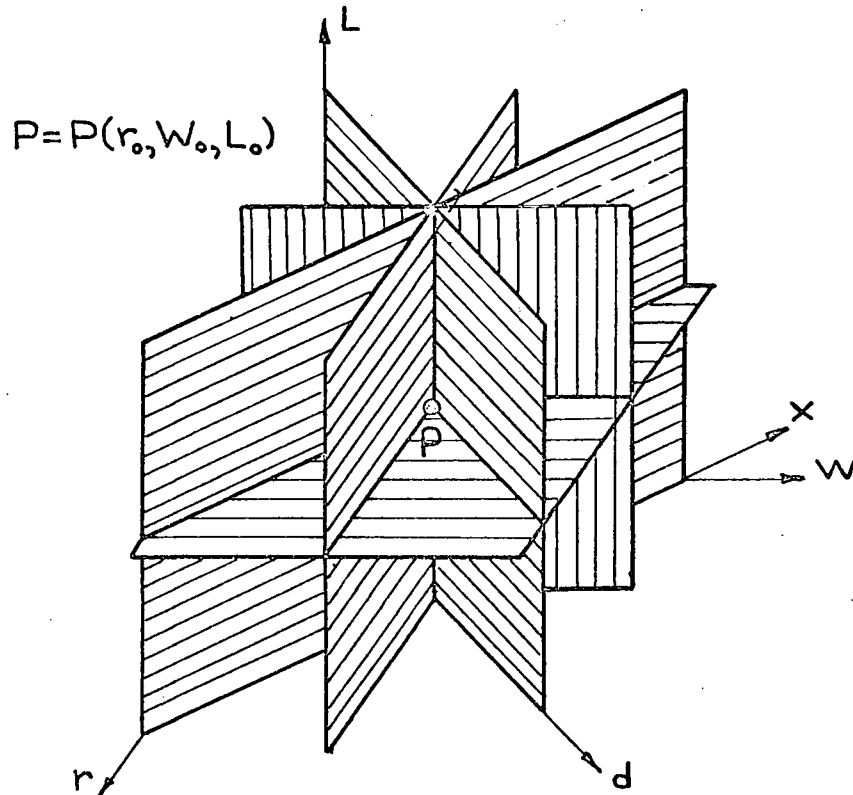
Fig. 4.5 Five planes passing through a reference point (P).

The abscissas in planes D and E are indicated in the figures with $d = W\sqrt{1 + k^2}$ and $x = W\sqrt{1 + m^2}$. Once the reference points were selected, the planes could then be defined and the required comparisons determined.

In order to obtain an estimate of the shape of the iso-preference surfaces, some pilot tests were conducted using the subjective-estimate method to be described in Section 4.5. These results were then used as guidelines in deciding which parameter should be the variable in the paired comparison tests. If the pilot curves indicated that by varying parameter A the number of listeners preferring the reference point would vary more rapidly than by varying any other parameter, the parameter A was varied.

4.4 Scaling of Isopreference Contours

A test based on the subjective estimate method was used to assign a meaningful value to each of the derived isopreference contours. The method permits the derivation of a scale based directly on the listeners' own quantitative estimates of the quality of a sentence.

The tests were conducted in two parts. At the beginning of the first part, the listeners received the following written instructions:

"In this test you will hear pairs of sentences; each pair is separated by a 5 second rest period. If zero denotes a sentence which is just unintelligible, and 10 denotes the first sentence (sentence A), rate the second sentence (sentence B) on an equal interval (0 to 10) scale on the basis of overall quality." A one second silence occurred between sentences within a pair. Sentence A was chosen to be the master sentence bandlimited to 6.3 KHz and re-recorded. In the second part of the tests, the order of presentation of the sentences within each pair was reversed in order to eliminate listener bias. In the second part, the listeners received instructions similar to those in the first part.

The rating tests were conducted in groups of 30 to 40 samples with two to three minute rest intervals at the end of each ten samples. Prior to each session, the listeners were asked to rate a series of five sentences spanning approximately the full range of scale values. This served to familiarize the listeners with the range of quality to be expected.

The tests were conducted in a quiet room using binaural listening with stereo headphones. The listeners were selected from the same 28 graduate students used in the paired comparison tests. The number of subjects used varied from 10 to 15 with 5 subjects listening at one time.

Rating tests were performed in four groups. The first group consisted of 116 samples used to obtain pilot curves for the ensuing DPCM isopreference tests. These were followed by the rating of 23 DPCM samples some of which were later used as reference points. These tests also included the rating of some two and three sample feedback DPCM sentences. The third group, an AM pilot run, consisted of 32 ratings. On the basis of these, 40 more AM samples were rated in order to obtain reference points for the AM isopreference contours. Some DPCM points were also rated in this last group for purposes of comparison. The results of all rating tests are tabulated in Appendix II.

The scale value assigned to each rated sample was taken as the mean of the listeners' ratings. The sample standard deviation was used as a measure of the variability of the obtained scale value.

4.5 Transitivity Checks

If a point A is judged to be equal in preference to a point B and to a point C, then the method of isopreference testing presup-

poses that if B is compared directly to C, it will be found equal in preference to point C. Although extensive tests were not conducted to check this assumption, an indication of this property of transitivity was obtained in the following way.

Some samples corresponding to AM reference points were used to obtain equal preference DPCM points. The scale values of the AM reference points used were then compared to the scale values of the DPCM reference points associated with the isopreference contours closest to the derived DPCM point. The similarity of these AM and DPCM scale values indicates the transitivity of the results. In a similar way, some samples corresponding to DPCM reference points were used to obtain equal preference AM points and the corresponding scale values were then compared.

5. RESULTS, EVALUATION AND CONCLUSIONS

5.1 Results of the DPCM Subjective Tests

5.1.1 Determination of the DPCM Isopreference Surfaces

The data obtained from the DPCM subjective tests are shown in Figures 5.1 to 5.5. The standard deviation associated with each data point equals the length of the line drawn through the point parallel to the axis along which it is measured. The scale values of the reference points (indicated by solid markers) are shown. Also shown are the standard deviations associated with each of the scale values. Each contour is assigned the scale value of the reference point through which it passes.

Figures 5.1 to 5.5 correspond to the sets of planes mentioned in Section 4.3.2. The five contours appearing in each figure do not all lie on a constant-parameter plane and therefore will not necessarily have the same general shape. This is illustrated in Figure 5.5 in which the wide range of shapes is due to a large variation in the orientation of each of the planes with respect to the r and W axes.

The curves drawn in these figures should be considered as first iterations to the contours obtained by the intersection of the isopreference surfaces with the above mentioned planes, since each of the sets of curves, as well as being self-consistent, must also be consistent with the other four sets of curves. To develop this consistency, Figures 5.1 to 5.5 were used to derive isopreference contours in planes of constant W, L and r. The principle advantage of choosing planes for which one parameter is held constant is that the shape of any one contour supplies an estimate of the general shape of the adjacent contours within that plane. Also, at least one of the contours appearing in Figures 5.1 to 5.3 appears in each of the
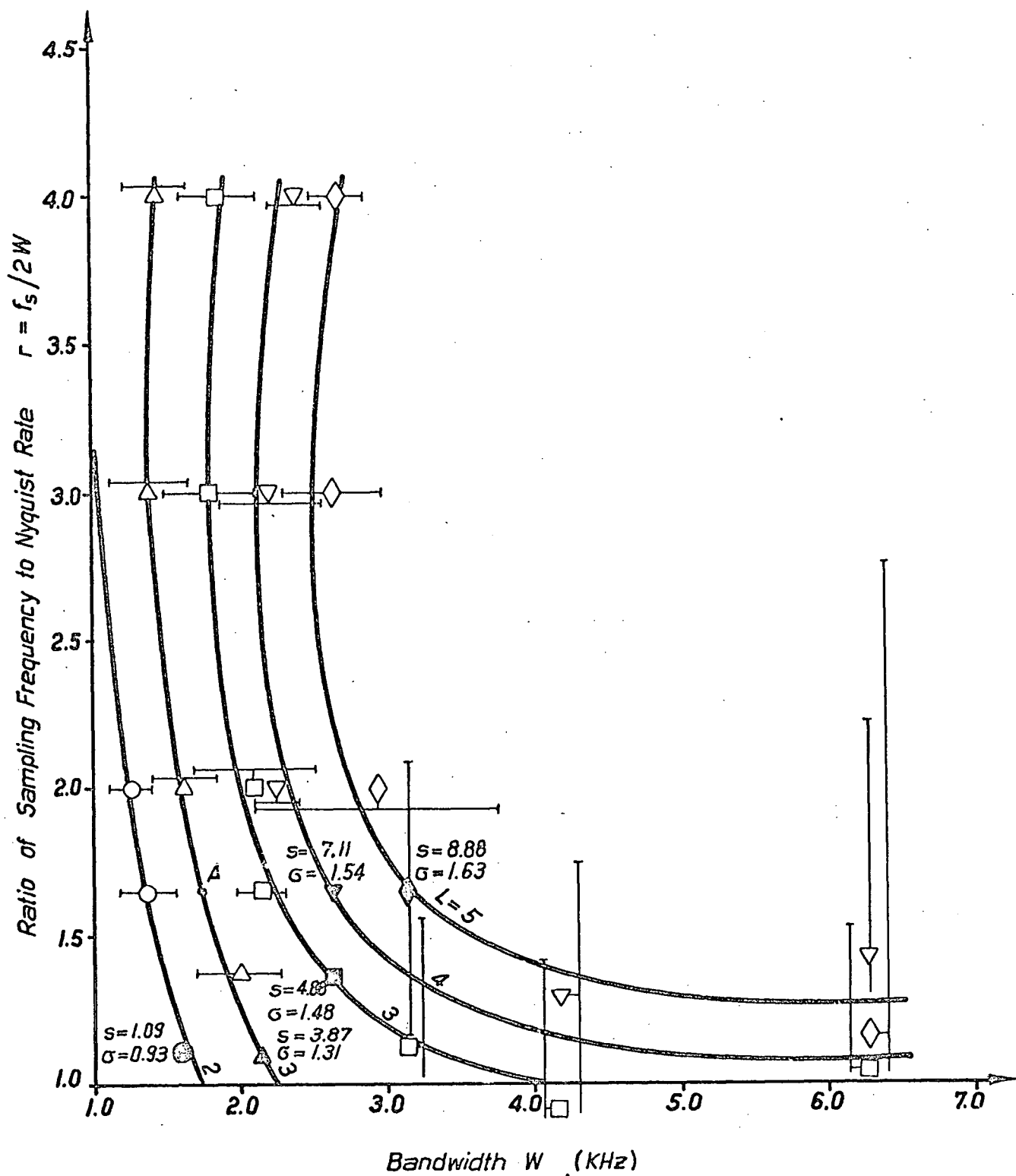
Fig. 5.1   Isopreference contours in L = constant planes (L is the number of bits of quantization). The scale value s and standard deviation σ associated with each contour is shown next to the reference point through which that contour passes.
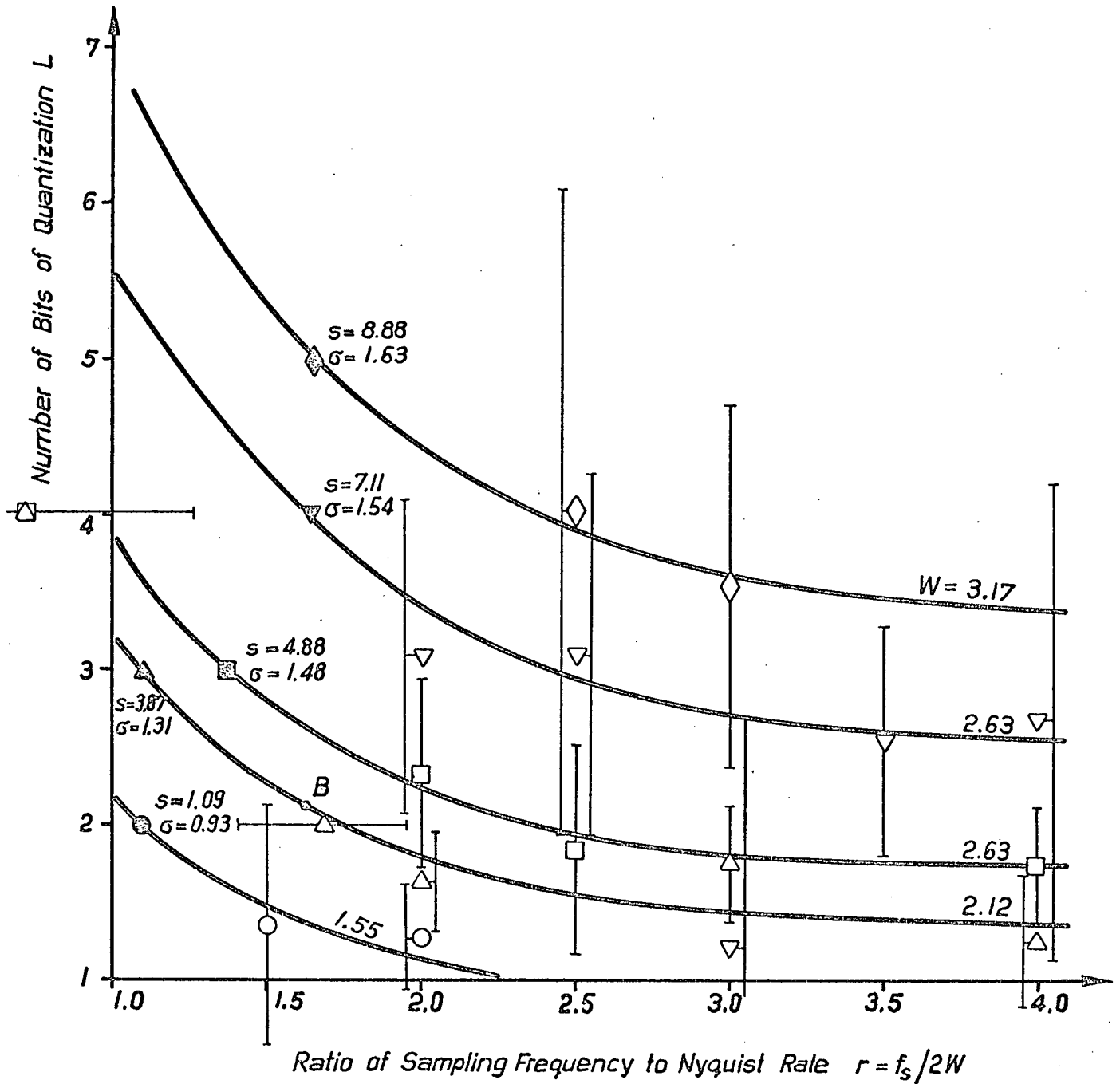
Fig. 5.2 Isopreference contours in W = constant planes (W is the bandwidth). The scale value s and standard deviation σ associated with each contour is shown next to the reference point (drawn solid) through which that contour passes.
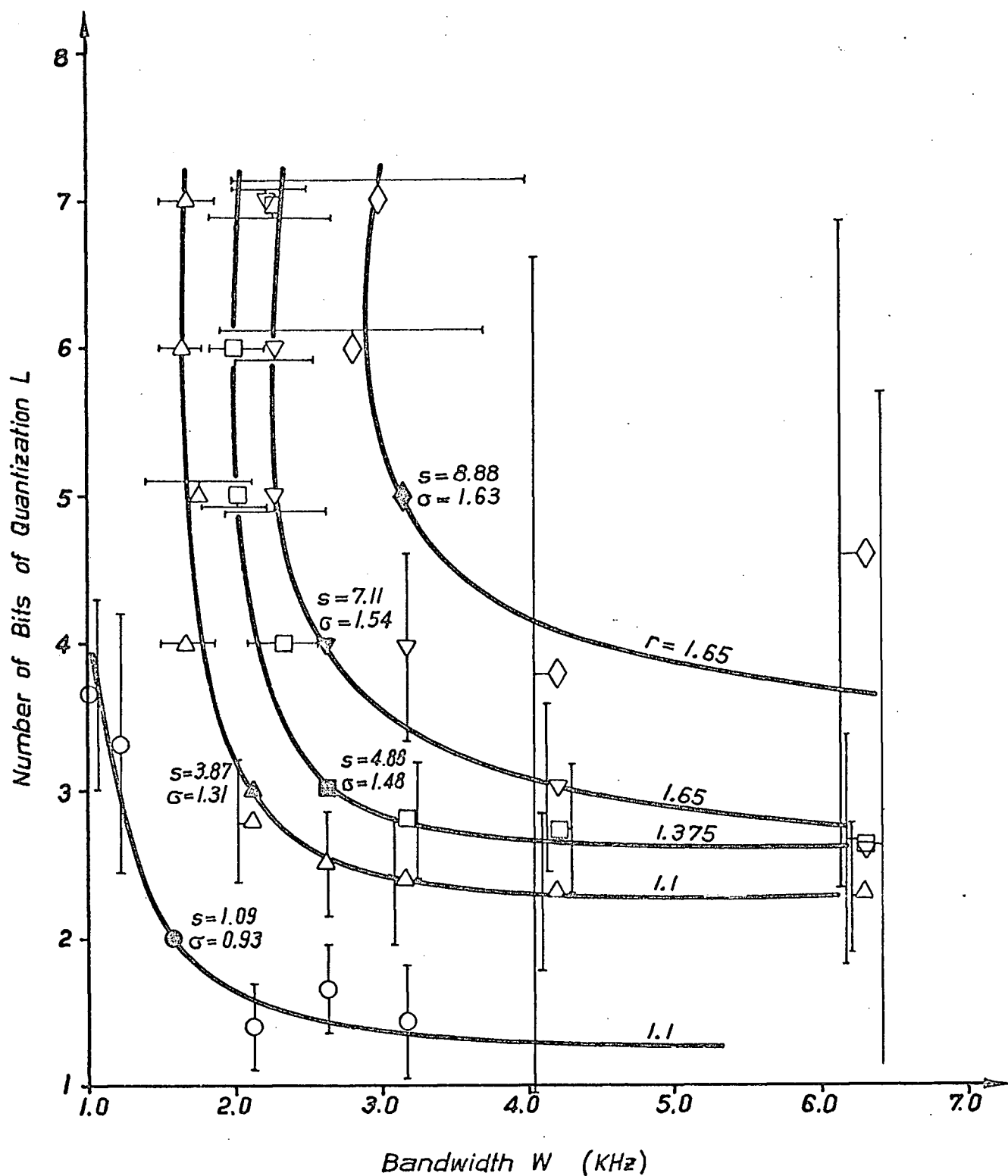
Fig. 5.3 Isopreference contours in r = constant planes (r is the ratio of the sampling frequency f_S to the Nyquist rate 2W). The scale value s and standard deviation σ associated with each contour is shown next to the reference point (drawn solid) through which that contour passes.
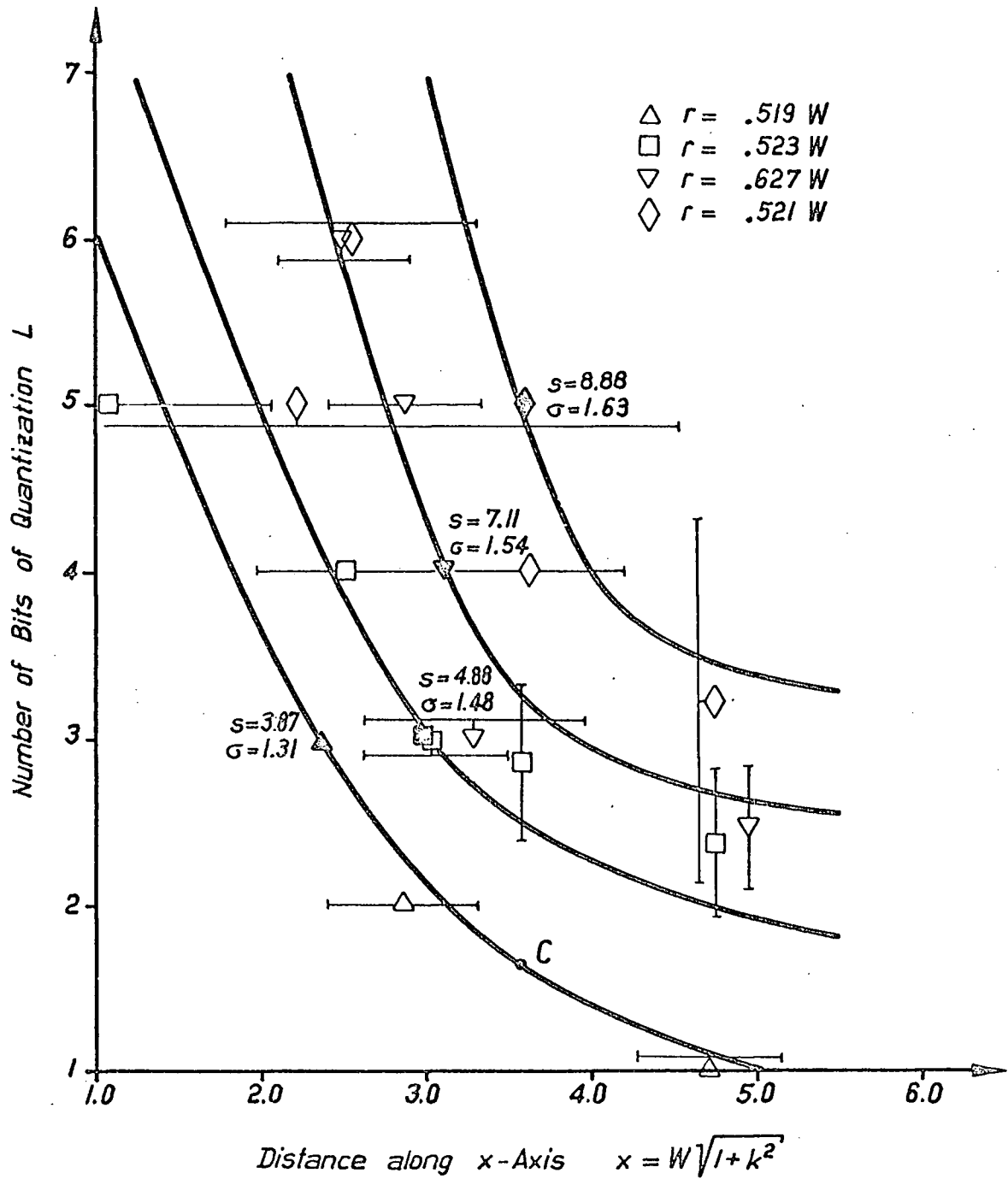
Fig. 5.4 Isopreference contours in planes defined by equations
of the form r = kW. The scale value s and standard
deviation σ associated with each contour is shown
next to the reference point (drawn solid) through
which that contour passes.
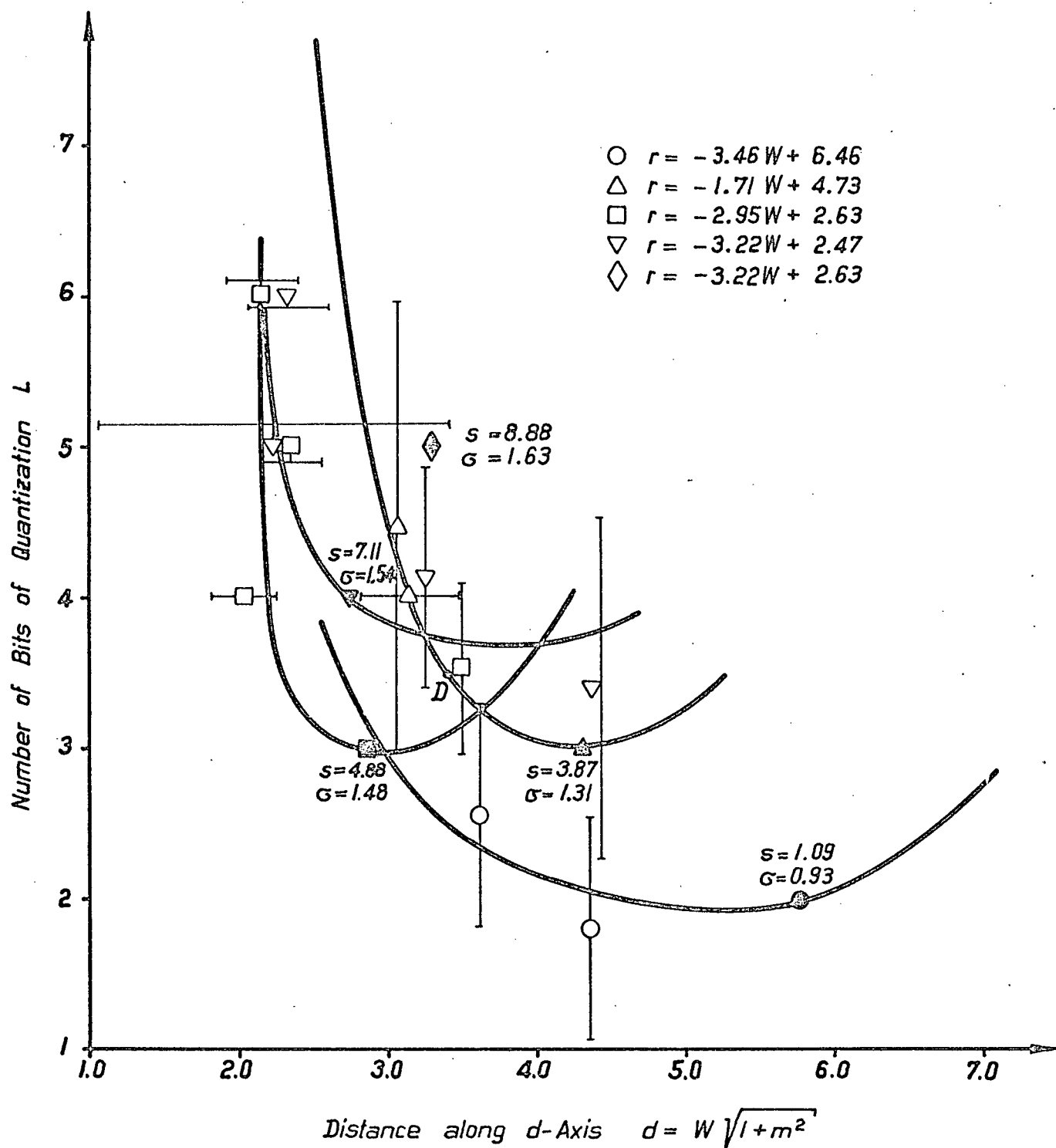
Fig. 5.5 Isopreference contours in planes defined by equations
of the form r = mW + b. The scale value s and standard
deviation σ associated with each contour are shown next
to the reference point (drawn solid) through which that
contour passes.

particular planes chosen. Since such contours are defined directly by data points, they are given the most weight in developing the general shape of curves appearing in that plane.

The method of deriving a curve in one plane from curves appearing in other planes is as follows. Assume it is desired to obtain a contour of scale value s = 3.87 in a plane for which the ratio of sampling frequency to Nyquist rate is held constant, say r = 1.65. The method is based on finding the intersection points of the r = 1.65 plane with the s = 3.87 contours in each of the planes of Figures 5.1, 5.2, 5.4 and 5.5. These points of intersection are indicated by the points labelled A,B,C,D in the four planes and are plotted in Figure 5.8.

Isopreference curves in each set of planes corresponding to W, L or r should be sufficient to define an isopreference surface. However, although the contours within any one plane should now be consistent within that plane, they may not necessarily be consistent with adjacent planes, and with isopreference curves in the other two sets of planes. This difficulty is overcome by using two of the three planes to iteratively update the curves in the third plane. A final set of contours so obtained is shown in Figures 5.6 to 5.9. In Figures 5.10 and 5.11, the intersections of the isopreference surfaces with planes defined by W = 2.12 KHz and L = 4 are shown. Figure 5.12 shows an isopreference surface.

5.1.2 Discussion of the DPCM Curves and Surfaces

If in any one of Figures 5.6 to 5.9, the value of L is increased along a line of constant W, a region is reached in which a further increase in L does not result in a substantial increase in scale value. In this region, the quality is primarily determined by speech bandwidth. Note that for a larger value of r, this region
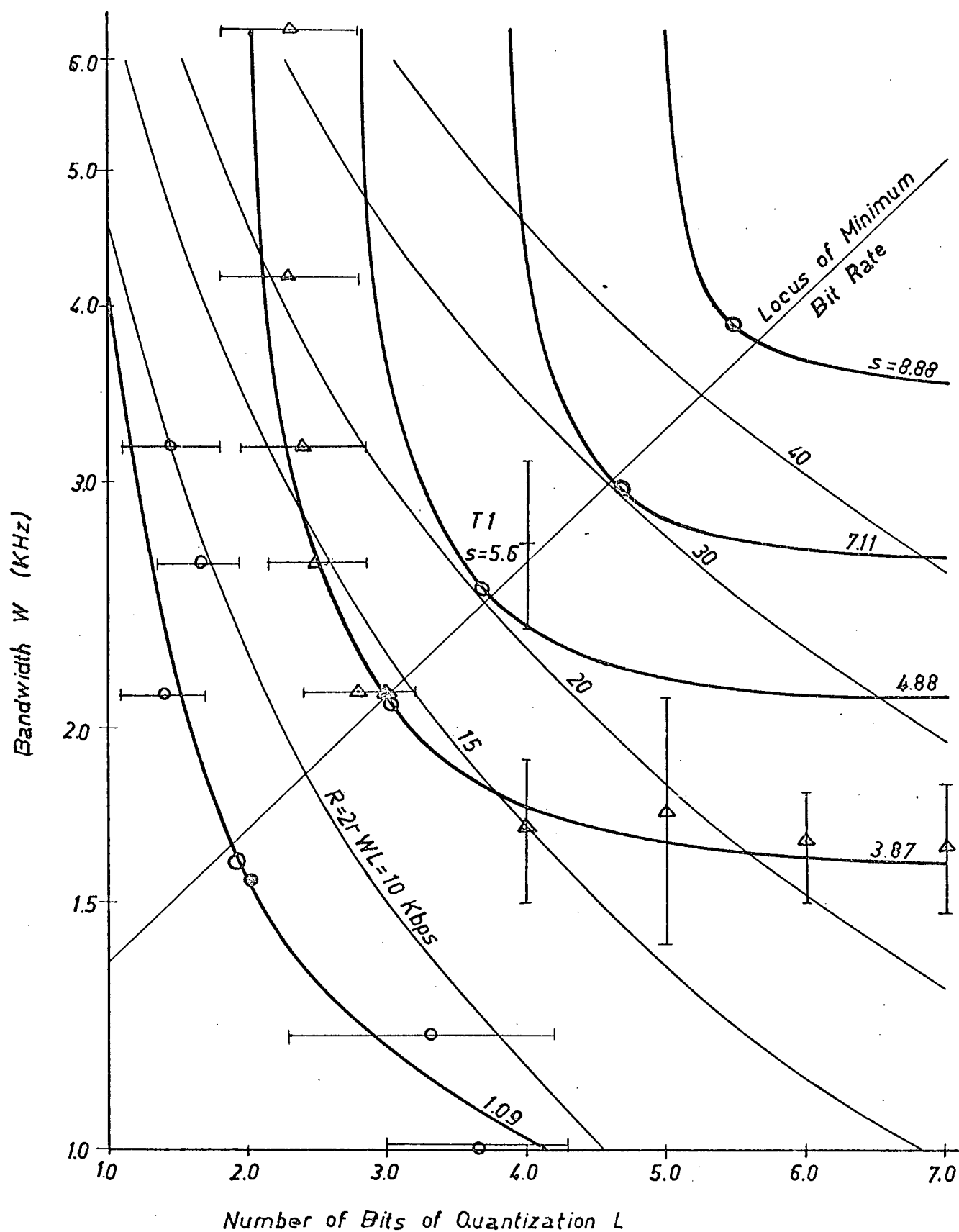
Fig. 5.6 Final isopreference contours and contours of constant
bit rate in the r = 1.1 plane. The scale value s or
bit rate R associated with each contour is shown on
the contour.
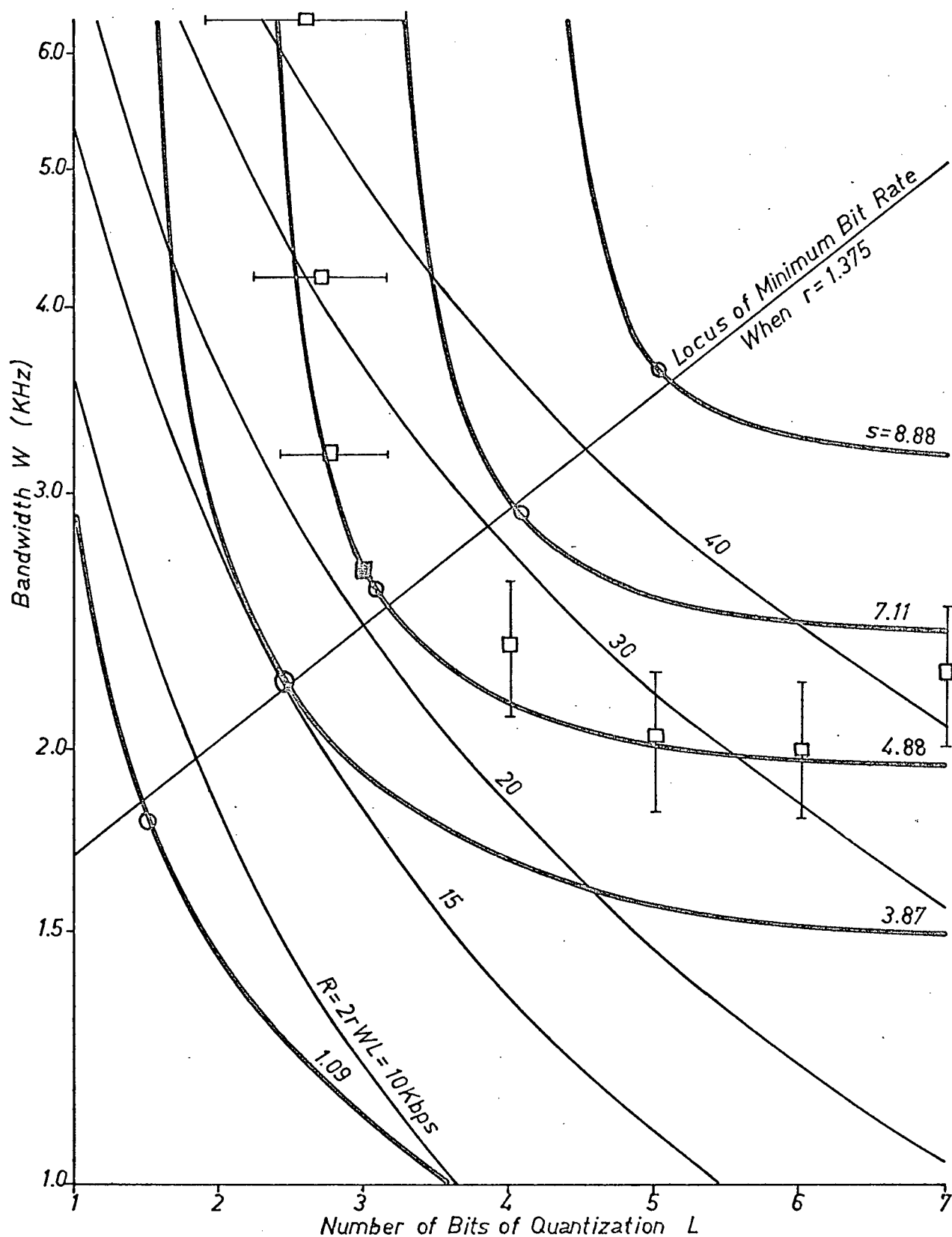
Fig. 5.7 Final isopreference contours and contours of constant
          bit rate in the r = 1.375 plane. The scale value s
          or bit rate R associated with each contour is shown
          on the contour.
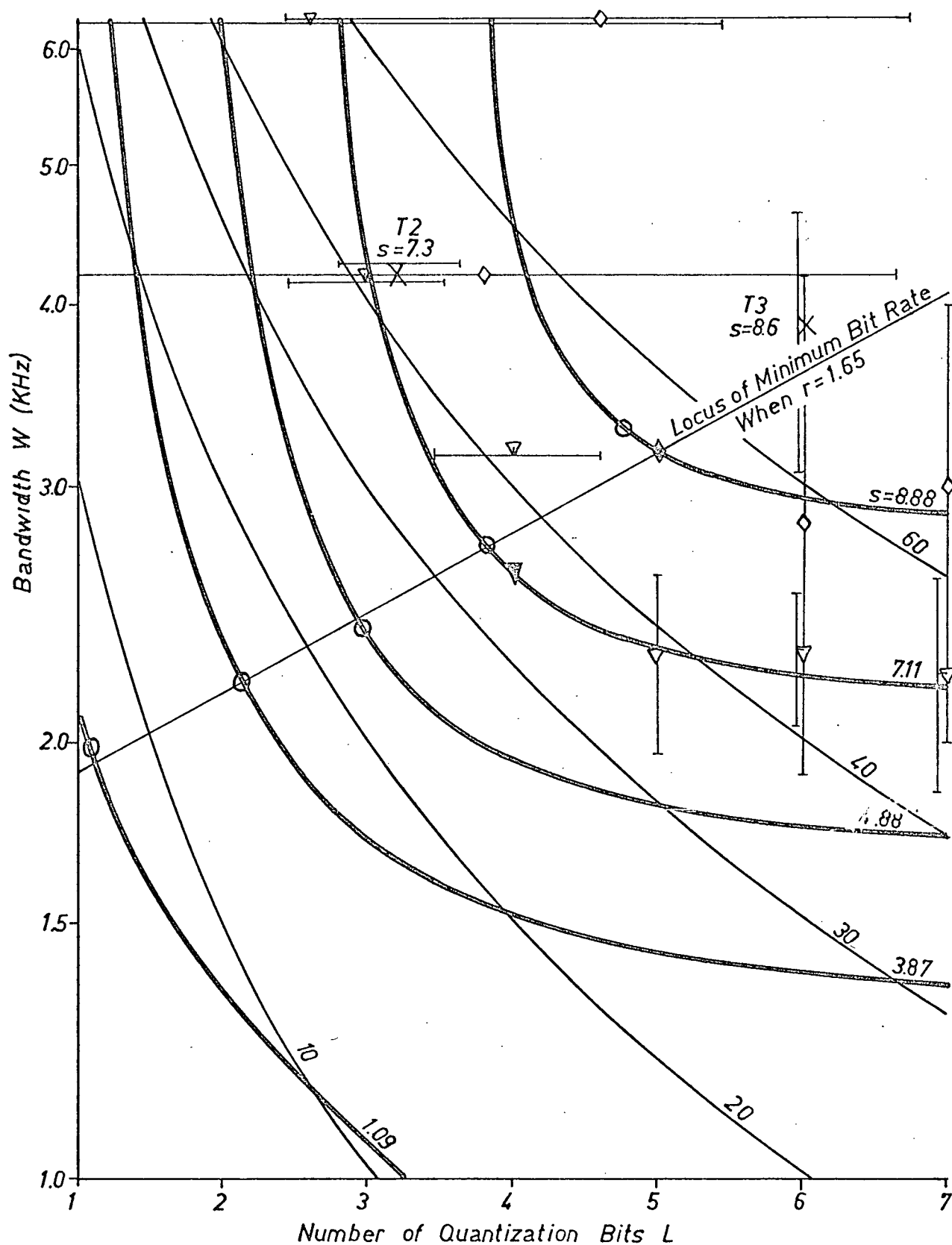
Fig. 5.8   Final isopreference contours and contours of constant bit rate in the r = 1.65 plane. The scale value s or bit rate R associated with each contour is shown on the contour.
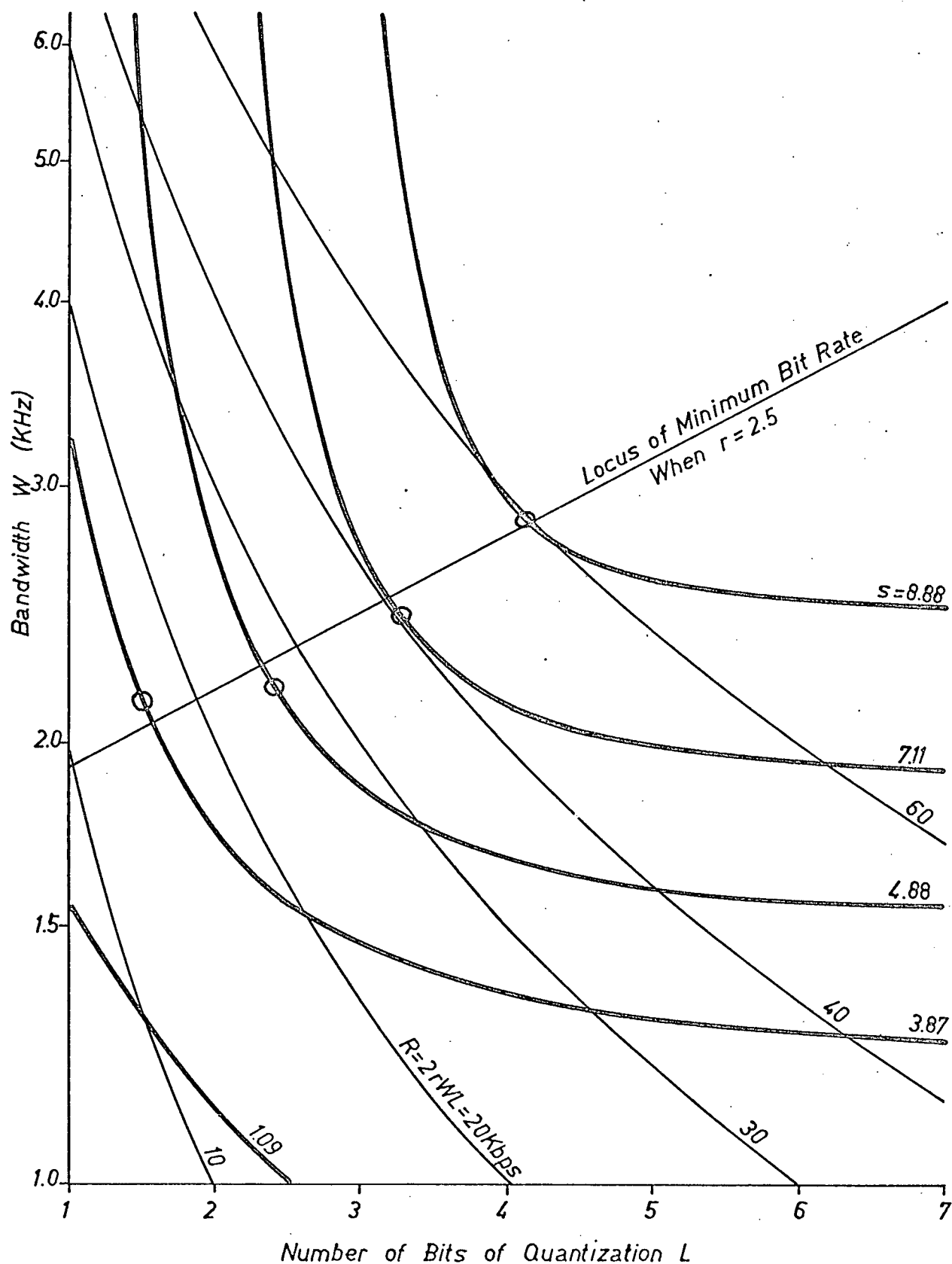
Fig. 5.9 Final isopreference contours and contours of constant
bit rate in the r = 2.5 plane. The scale value s or
bit rate R associated with each contour is shown on
the contour.
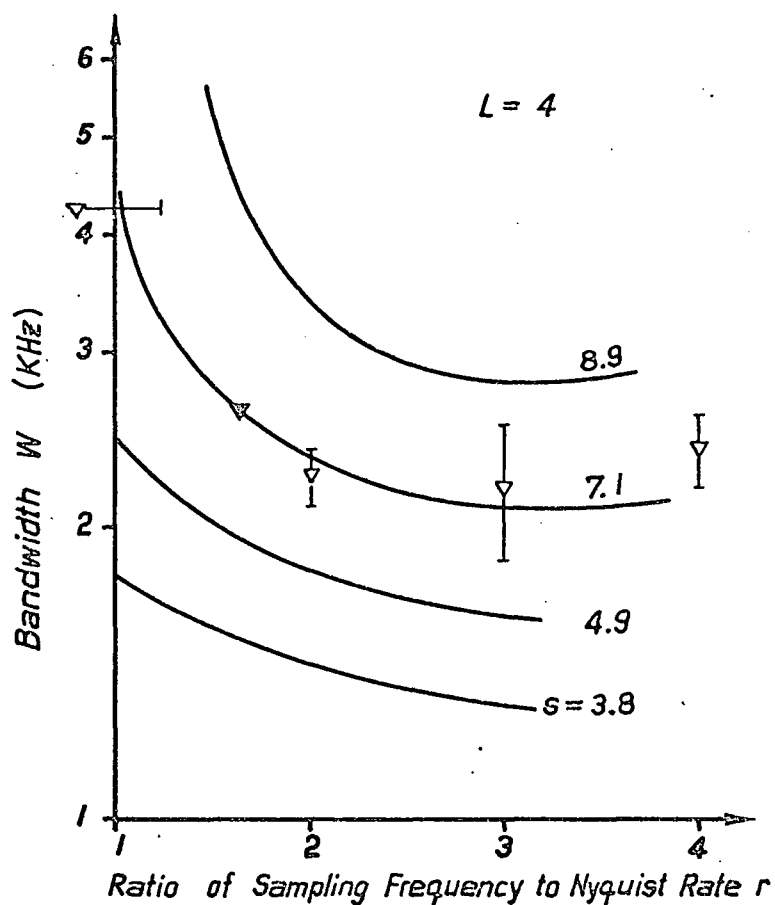
Fig. 5.10 Final isopreference contours in the L = 4 plane. The scale value s associated with each contour is shown on the contour.
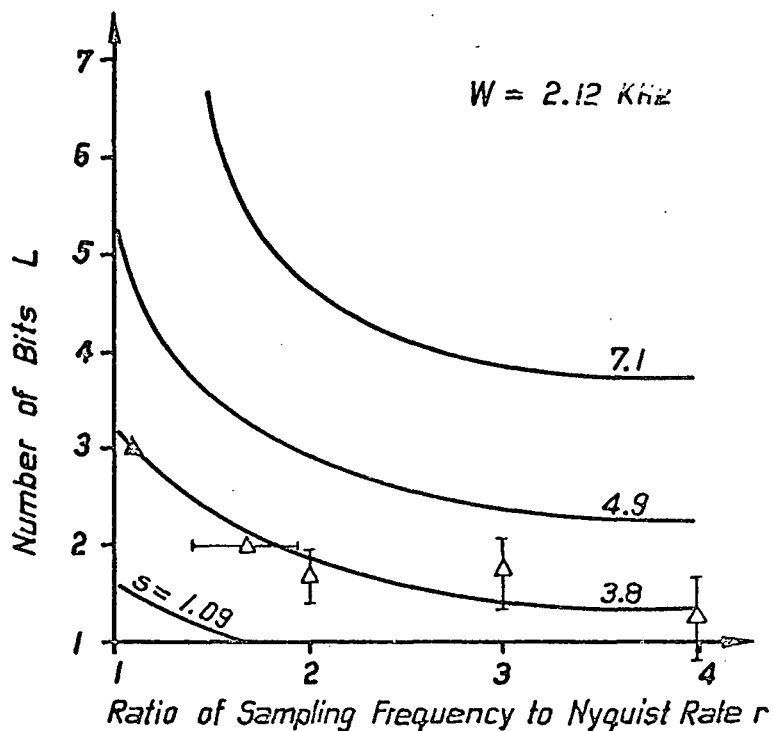
Fig. 5.11 Final isopreference contours in the W = 2.12 KHz plane. The scale value s associated with each contour is shown on the contour.
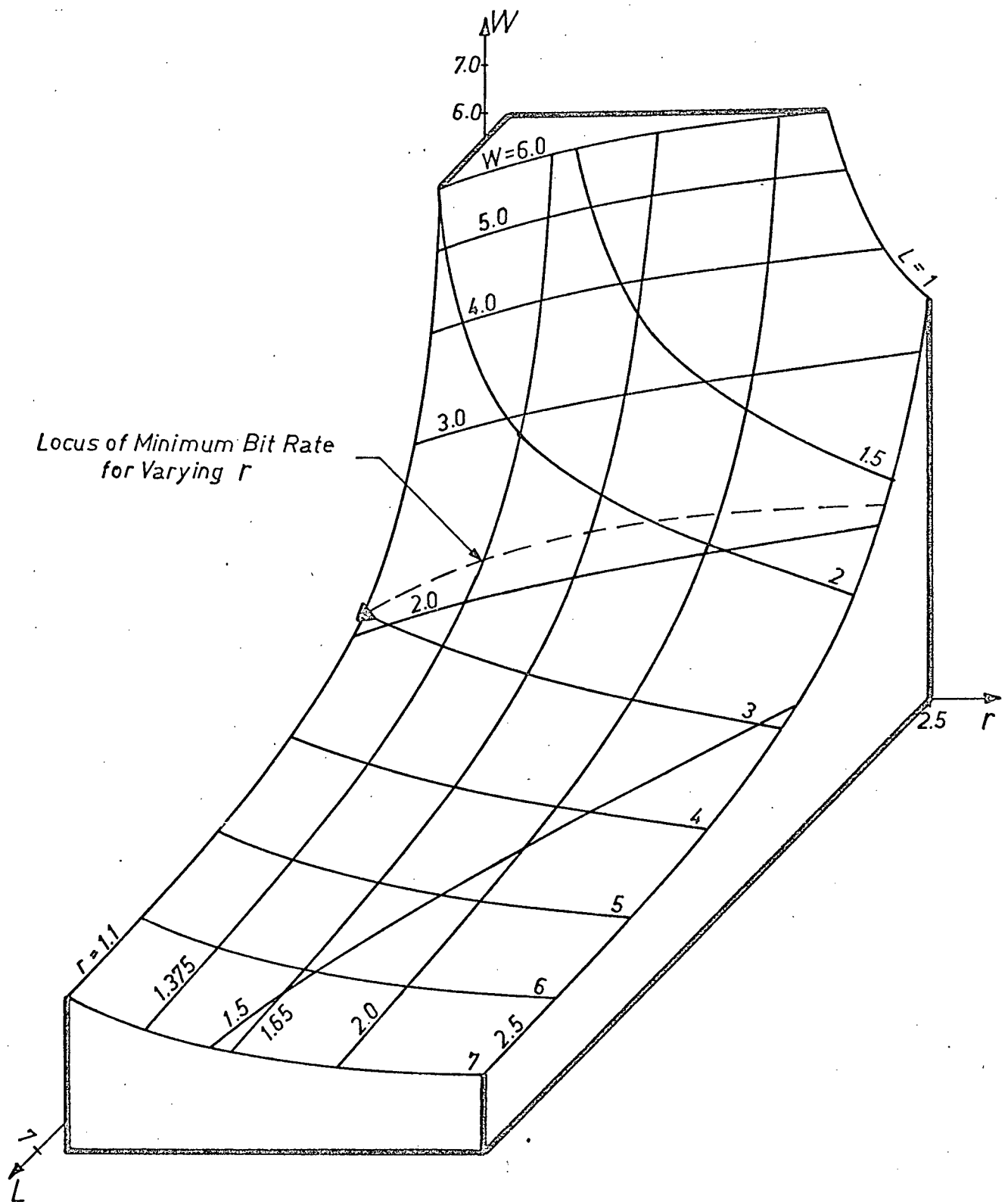
Fig. 5.12   An isopreference surface.   Contours of constant L, W
            and r are shown.

is reached at a smaller value of L. This results because the quantization noise, already made small by a large r, does not require as large a value of L to reduce it to a point beyond which the listener is sensitive only to the loss of naturalness due to lowpass filtering.

If W is increased along a line of constant L, a region is reached in which the quality is primarily limited by quantization noise. The larger the value of r in this region, the greater is the dependence of quality on W, since a small increase in W results in a large increase in sampling frequency when r is large. For reasons given in Section 2.3, quantization noise decreases as r increases.

The effects of increasing r are also apparent in Figures 5.10 and 5.11. For example, in the upper left portion of Figure 5.10, a much larger increase in W is required to yield a fixed increase in scale value than is required in the lower right region. Similar comments apply to Figure 5.11.

5.1.3 Minimum Bit Rate Loci

Also shown in Figures 5.6 to 5.9 are points on each contour for which the bit rate R = 2rWL is a minimum for a fixed value of r. The locus of these point of minimum bit rate was found in each case to be a curve defined by the equation $W = a(r)2^{b(r)L}$. In Figure 5.13, the parameters a(r) and b(r) are plotted as functions of r. The equations for a(r) and b(r) were found to be closely approximated as follows.

$$a(r) = \begin{cases} r & r \leq 1.7 \\ 1.7 & r \geq 1.7 \end{cases}$$

$$b(r) = \begin{cases} -.23r+.57 & r \leq 1.7 \\ .175 & r \geq 1.7 \end{cases}$$

(5.1)

It follows that the equation of the locus of minimum bit rate in a
plane of constant r is

$$W = \begin{cases} r \; 2^{(-.23r+.57)L} & r \lesssim 1.7 \\ (1.7) \; 2^{.175L} & r \lesssim 1.7 \end{cases} \tag{5.2}$$
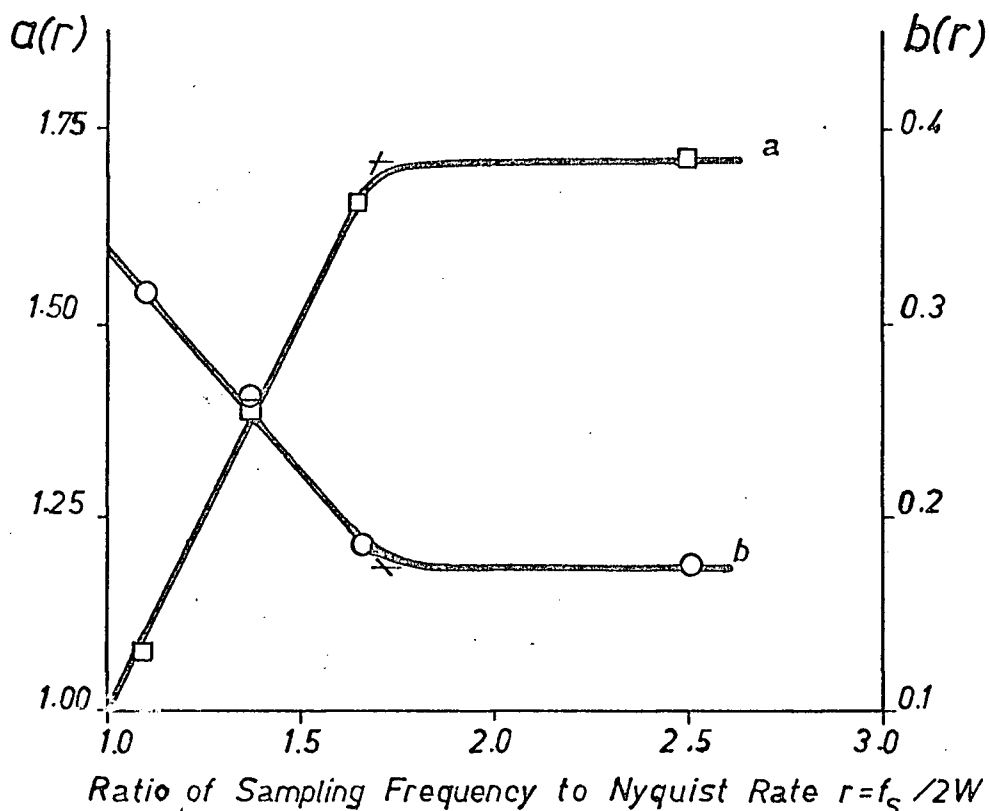
Fig. 5.13 The variation vs r in the parameters
a(r) and b(r) in the equation
$W = a(r)2^{b(r)L}$.

In Figure 5.14, the scale values corresponding to the
intersection points of equation (5.2) with the isopreference contours
in each of Figures 5.6 to 5.9 are plotted in unit normal deviates
versus the required number of quantization bits and the required
bandwidth. For each value of r, the points are fitted well by a
straight line, indicating that the scale values are normally distri-
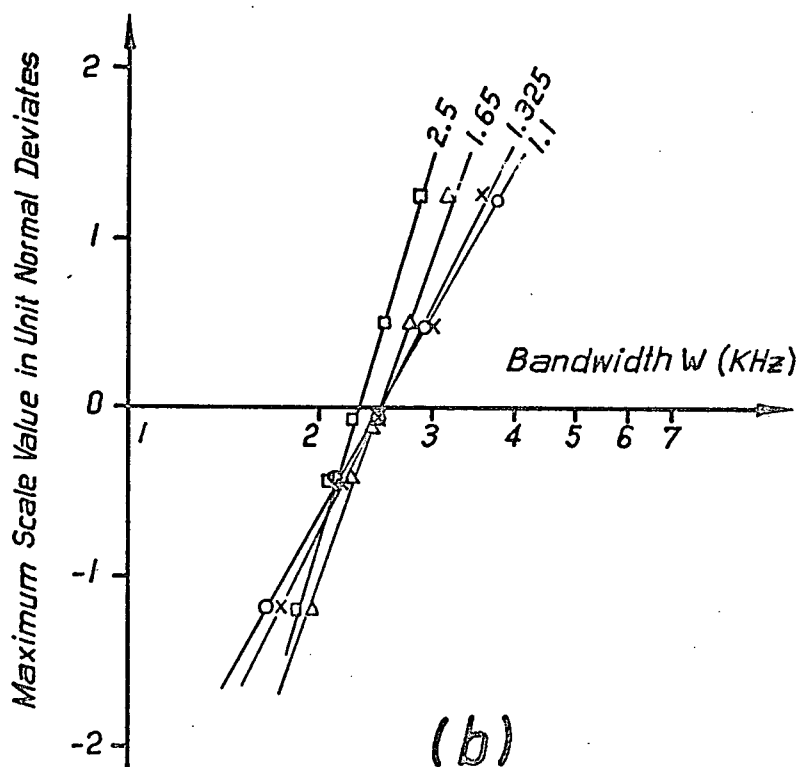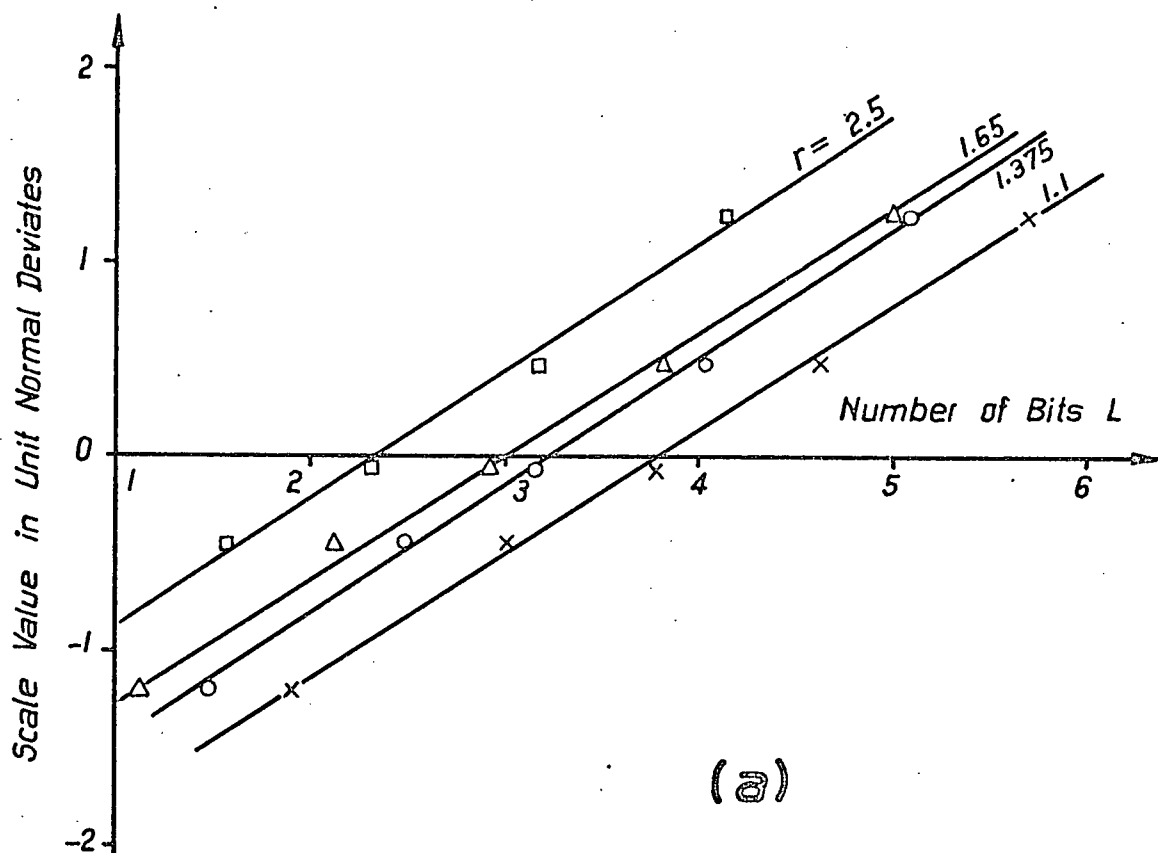
Fig. 5.14 The scale values corresponding to the intersection points
of $W = a(r)2^{b(r)L}$ with isopreference contours in planes
of constant r as functions of (a) the required quantiza-
tion bits (b) the required bandwidth.

buted over L and W. From Figure 5.14(a), it follows that the effect of an increase in r is equivalent to a decrease in L (with the corresponding change in W). Such an effect is suggested by equation (2.12). The parallelism of the lines in Figure 5.14(a) suggests that an increase in r has no effect on the sensitivity of the scale value to changes in L whereas the same increase in r causes the scale value to become more sensitive to changes in W. Since $f_s = 2rW$, a given increase in W is translated into a large increase in f_s for large r. However, unless Wr is of the same order of magnitude as the bandwidth of the quantization noise (controlled primarily by L) [28,29], a change in r will not tend to affect the sensitivity of the scale value to changes in L.

The maximum scale value attainable for a given bit rate R is plotted in Figure 5.15 for various values of r. For all bit rates R, the optimum value of r is r = 1.1. This result indicates that, for a fixed bit rate, the decrease in quantization noise in the receiver passband caused by an increase in r is not sufficient to justify the required decrease in L and change in W. In practical systems however, L is constrained to be discrete. It is therefore of interest to see if the decrease in scale value caused by the use of a discrete L may be compensated for by an increase in r. Table 5.1 shows that even when L is constrained to be discrete, the optimum value of r is 1.1.

Table 5.1 Comparison of Maximum Scale Values Attainable using Discrete and Continuous Values for the Number of Bits of Quantization L.

| s \ r | 1.1 | | 1.375 | | 1.65 | | 2.5 | |
R	Ldis	Lcont	Ldis	Lcont	Ldis	Lcont	Ldis	Lcont
10	2.16	2.2	1.98	1.98	1.98	1.98	1.84	1.70
20	4.54	4.63	4.38	4.4	4.05	4.1	3.8	3.8
30	6.63	6.7	6.43	6.47	5.9	6.0	5.15	5.55
40	8.08	8.2	7.7	7.8	7.5	7.5	6.7	7.0
50	9.05	9.10	8.90	8.95	8.66	8.70	8.08	8.15
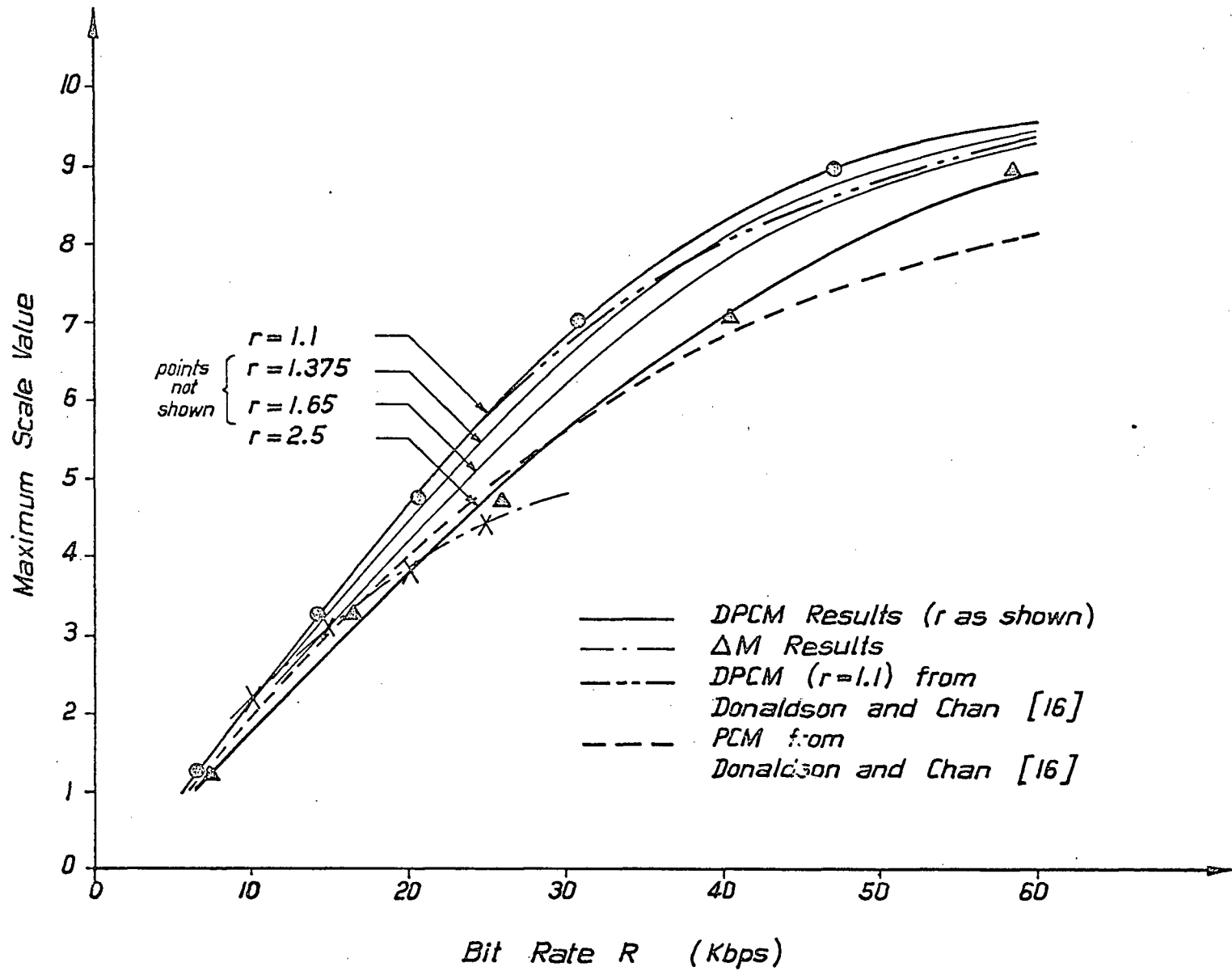
Fig. 5.15 The maximum attainable scale value for a fixed bit rate R plotted
for various values of r. Also shown are results presented by
Donaldson and Chan for DPCM (r = 1.1) and PCM. The points marked
by X correspond to maximum scale values for ΔM.
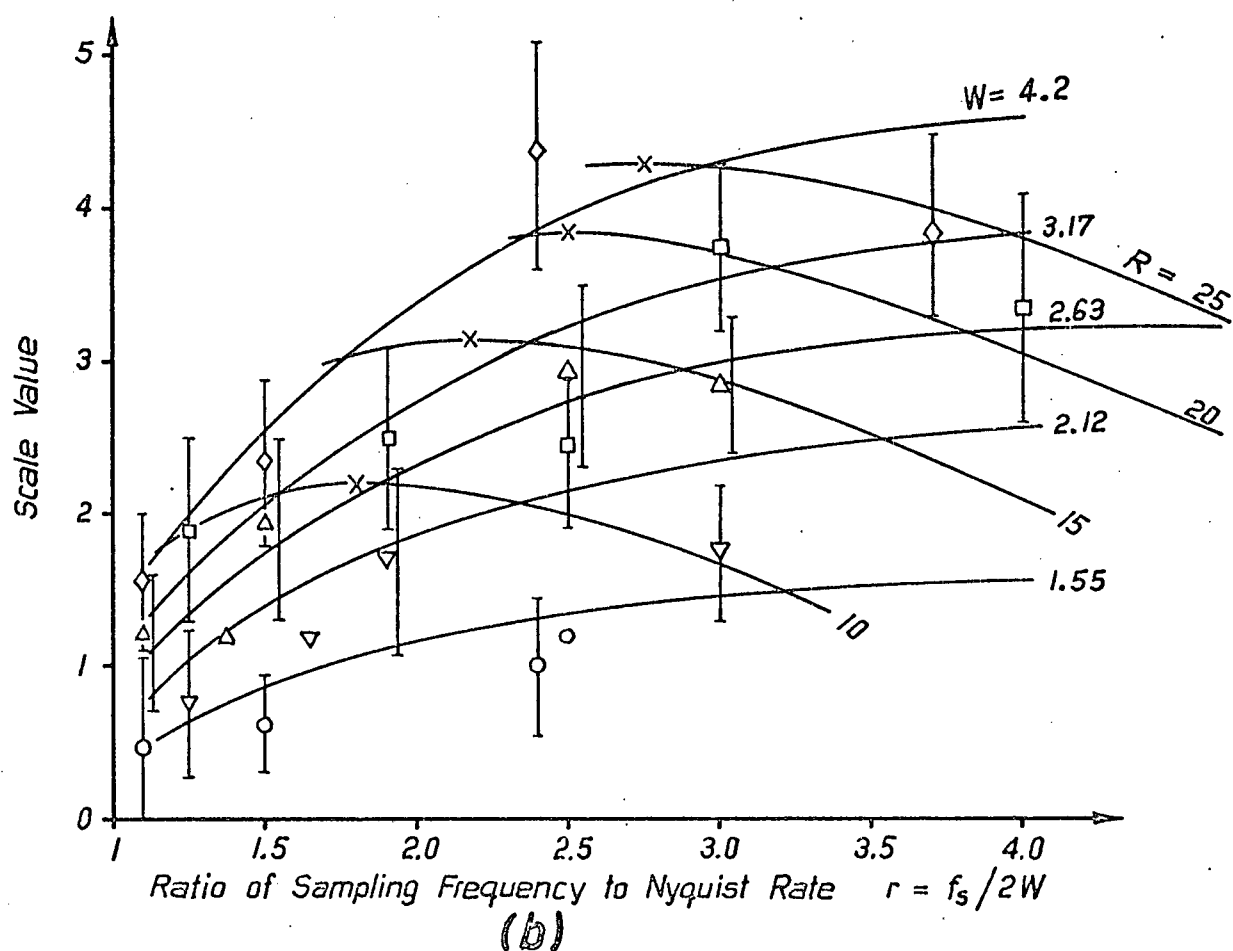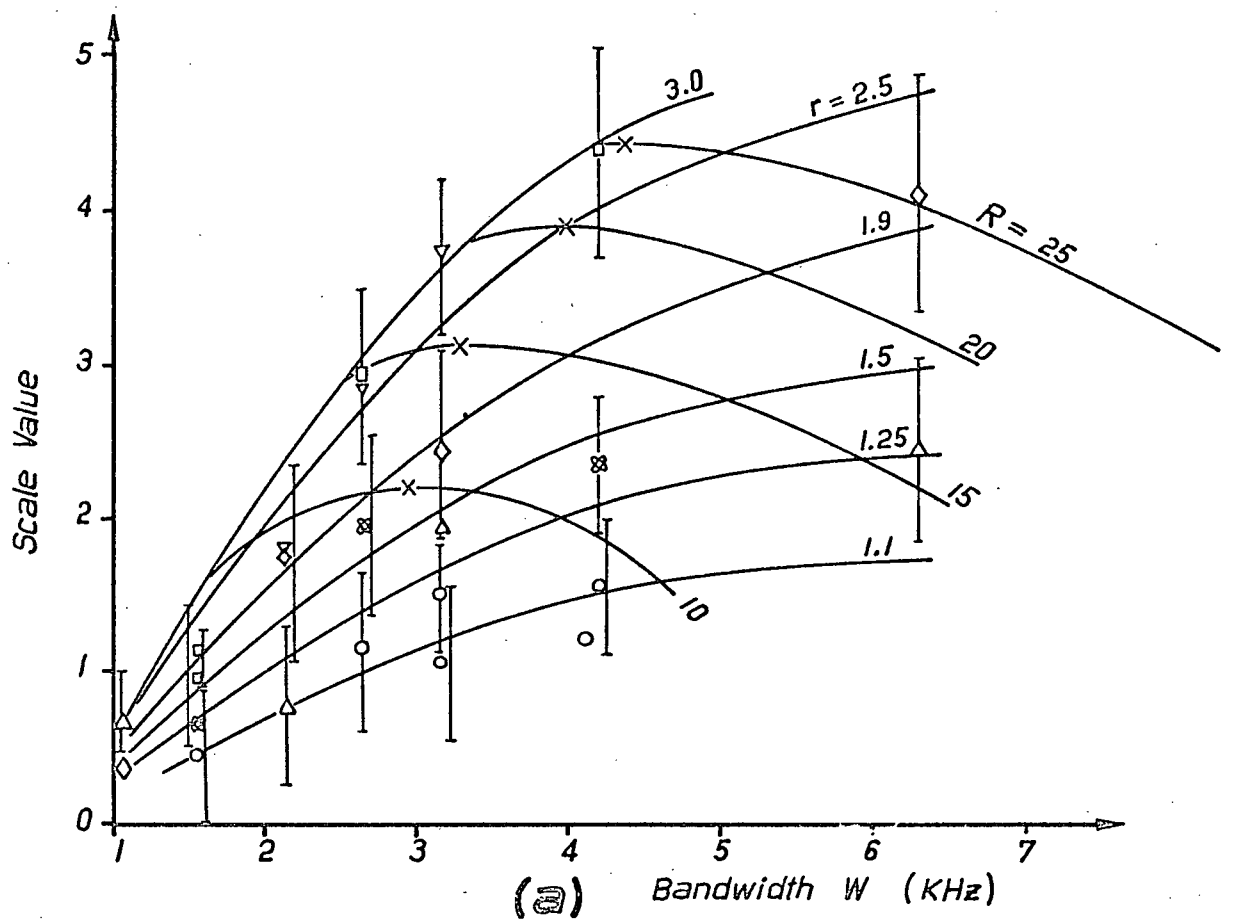
Fig. 5.16 Curves of scale value versus (a) bandwidth W (b) ratio
of sampling frequency to Nyquist frequency for the
special case of L = 1. Also shown are contours of
constant bit rate.

A special case, known as delta modulation (ΔM), occurs for $L = 1$ and $r \gg 1$. Since the quantizer for such a system is very simple, ΔM is comparatively easy to implement. The data available from the isopreference tests were not sufficient to permit the derivation of isopreference contours in the $L = 1$ plane. By using the data from the rating tests (presented in Appendix II) however, the curves of scale value versus W and r for $L = 1$ in Figure 5.16 were obtained. These curves show that for $r \geq 2.5$, the quality of the speech output is not substantially improved by an increase in r.

Also shown in Figure 5.16 are contours of constant bit rate. The maximum attainable scale values for given bit rates are equal to the maxima of these curves. The scale values and parameter values corresponding to these maxima in both Figure 5.16(a) and Figure 5.16(b) were found and appear in Table 5.2. It is seen that the two sets of results agree almost exactly. The average of the two resultant scale values appears as the ΔM curve in Figure 5.16. A comparison of the DPCM and the ΔM curves indicates that only for very low bit rates is the performance of ΔM able to match that of PCM or DPCM.

Table 5.2   Maximum Scale Values for Fixed Bit Rates Obtained From Figs. 5.16(a) and (b).

R	Fig. 5.16(a)			Fig. 5.16(b)			Average
	r	W(KHz)	S_{MAX}	r	W	S_{MAX}	
10	1.7	2.95	2.20	1.8	2.78	2.20	2.20
15	2.03	3.69	3.15	2.17	3.45	3.15	3.15
20	2.5	4.00	3.90	2.5	4.00	3.85	3.93
25	2.86	4.38	4.45	2.75	4.54	4.3	4.38

Also shown in Figure 5.16 are PCM and DPCM minimum bit rate curves obtained by Donaldson and Chan [16]. The DPCM results ($r = 1.1$) are in close agreement, although the DPCM curve presented here indicates larger scale values than obtained by Donaldson and Chan when the bit rate is large. One possible explanation of this

discrepancy is that more steps were used here in the approximation
of the feedback coefficient than were used by Donaldson and Chan.
Another is that the loci of minimum bit rate are not measured as
exactly here as in [16] since the present work yielded less data in
the plane $r = 1.1$ than did [16].

5.2 Results of the AM Subjective Tests

5.2.1 AM Results and Discussion

The results of the AM subjective tests are shown in
Fig. 5.17. The format of the figure follows that used for the DPCM
curves. The points plotted at $S/N = 60db$ were obtained by simply
lowpassing the input sentence witoout the addition of noise. This
was considered permissible since it was expected that as S/N was
increased, the curves would approach asymptotes determined only by the
bandwidth. That this is indeed the case is apparent in Fig. 5.17.

As the value of W is increased along a line of constant S/N,
a region is reached in which the quality is primarily limited by S/N.
However, if W is increased further a point is reached at which the
quality begins to deceease. This behaviour may be traced to the rela-
tive absence of signal power in the high frequency portions of the
speech spectrum (see Fig. 4.1) and to the sensitivity of human hearing
to high frequency noise [38]. Since, as W is increased, more noise
is added to the higher frequency portion of the spectrum, the scale
value decreases.

5.2.2 Extension of AM Results to Angle Modulation

As shown in Section 2.5, the spectrum of the output noise
of an FM System for white Gaussian channel noise varies as the square
of the audio frequency. For this reason, the results will be extended
only to PM for which the output noise spectrum is flat.

The signal-to-noise ratio obtained experimentally as des-
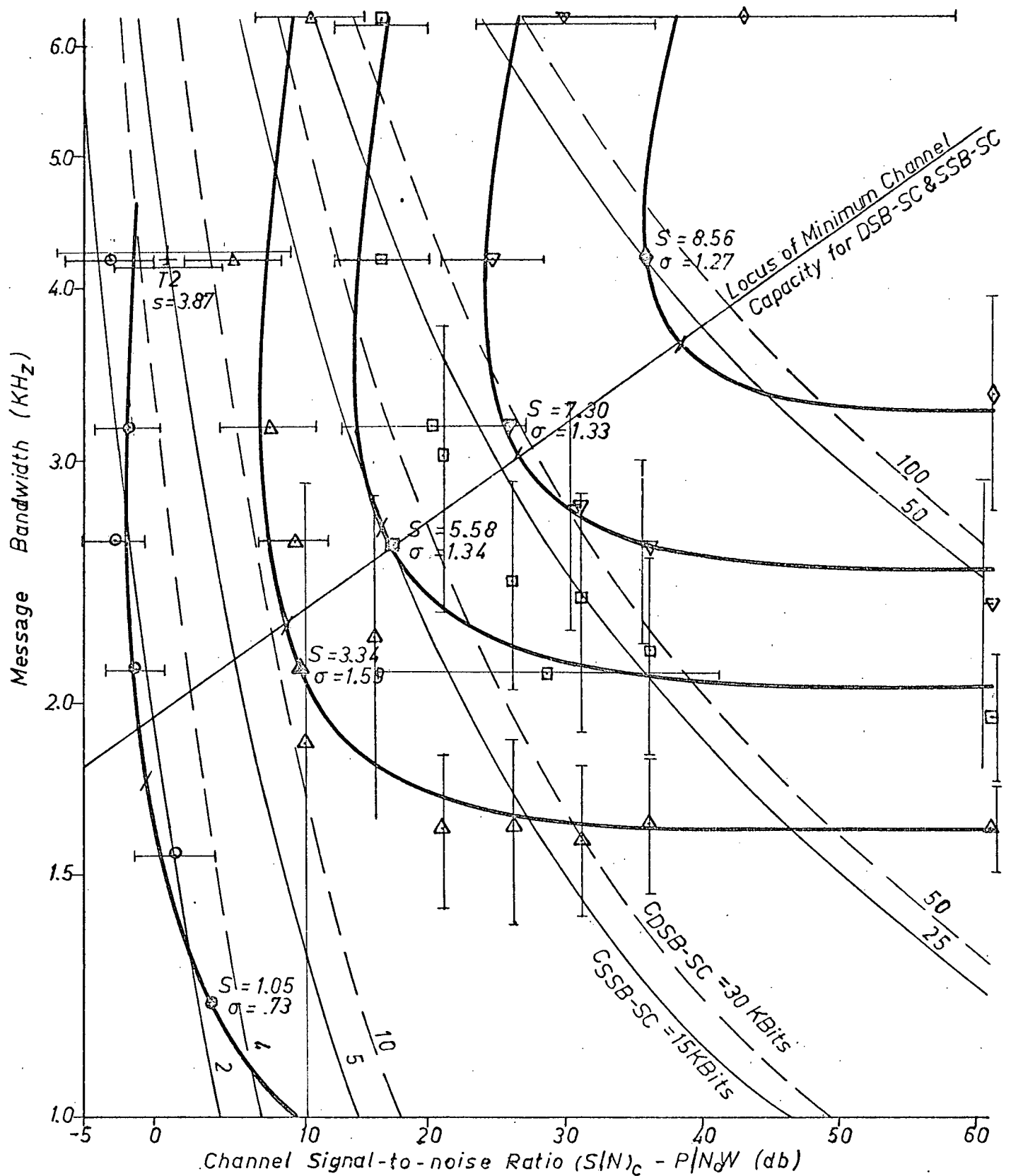
Fig. 5.17 AM isopreference contours. The scale value and standard
deviation associated with each contour is shown next to the
reference point (drawn solid) through which that contour
passes. Also shown are contours of constant channel capa-
city for DSB-SC (C_D) and SSB-SC (C_S) communication systems.

cribed in Section 3.1 equals the channel signal-to-noise ratio of an AM system

$$\left(\frac{S}{N}\right)_{cAM} = \frac{P}{N_oW}$$

The equation for $(S/N)_{cAM}$ may be written

$$\frac{S}{N}_{AM} = \left(2\int_W^\infty X^*(f) + \frac{1}{(S/N)_{cAM}}\int_{-W}^W X^*(f)df\right)^{-1}$$

Similarly, the equation for $(S/N)_{PM}$ may be written

$$\left(\frac{S}{N}\right)_{PM} = \left(2\int_W^\infty X^*(f) + \frac{1}{(S/N)_{cPM}}\int_{-W}^W X^*(f)df\right)^{-1}$$

where

$$\left(\frac{S}{N}\right)_{PM} \triangleq \left(\frac{S}{N}\right)_{cPM} \frac{(W_{cPM}-2W)^2}{c^2\ \overline{f_m^2}\ (W)}$$

and $(S/N)_{cPM} \triangleq (P/N_oW) \triangleq$ the PM channel signal-to-noise ratio. For purposes of comparison, the assumption is made that $(S/N)_{PM}$ and $(S/N)_{AM}$ may be directly associated with the scale value. In order to operate at a common scale value, $(S/N)_{cAM}$ and $(S/N)_{cPM}$ must therefore be equal. For fixed W_{cPM} and W, it is therefore necessary to adjust the channel power-to-noise ratios $(P/N_o)_{AM}$ and $(P/N_o)_{PM}$ until they are equal.

For low quality telephone applications a value of $W_{cPM} - 2W = |\Delta f|_{max}$ of the order of 15 KHz is typical, while for high quality transmission, 75 KHz is typical [39]. Results for these two cases are plotted. In order to compare the AM and PM systems, it remains to estimate $\overline{f_m^2}$ and to choose a value for c^2. Published results indicate that for speech, the probability that the magnitude of the instantaneous input x(t) will exceed $c\sqrt{\overline{x^2}}$ is less than 2.0% for a peak factor $c^2 \geqslant 10$ [40]. Approximate values of $\overline{f_m^2}$ corresponding to eight bandwidths are tabulated in Table 5.3. These values were calculated using the power spectrum of speech from French and Steinberg [36]

shown in Figure 4.1. Using the above approximations, and the equation

$$\left(\frac{S}{N}\right)_{cPM} = \frac{c^2 \overline{f_m^2(W)}}{(W_{cPM} - 2W)^2} \left(\frac{S}{N}\right)_{cAM}$$

the results shown in Figures 5.18 and 5.19 were obtained.

Table 5.3 Mean Square Bandwidth of Speech

W(KHz)	$\overline{f_m^2}$(KHz)
1.0	0.232
1.2	0.272
1.5	0.316
2.0	0.367
2.5	0.422
3.0	0.483
3.5	0.546
4.0	0.633

5.3 Comparison of DPCM, PCM, ΔM, DSB-SC, SSB-SC and PM.

It may be shown that regardless of the communication system used, there is a minimum rate $R(\varepsilon)$ at which information must be transmitted to a receiver in order to specify an analog source to within a mean square error ε [25]. However, for a channel of fixed bandwidth and signal-to-noise ratio, there is an upper limit imposed on the rate at which information may be transmitted without resulting in a high probability of transmission errors. This limit is termed the channel capacity C. For bandlimited white Guassian channels

$$C = W_c \log_2 (1 + P/N_o W_c) \text{(bits/second)} \tag{5.3}$$

where P is the average signal power, $N_o/2$ is the noise power density and W_c is the channel half-bandwidth. It follows that a necessary condition for transmitting an analog waveform to a receiving point via a noisy channel is that

$$R(\varepsilon) \leqslant C \tag{5.4}$$
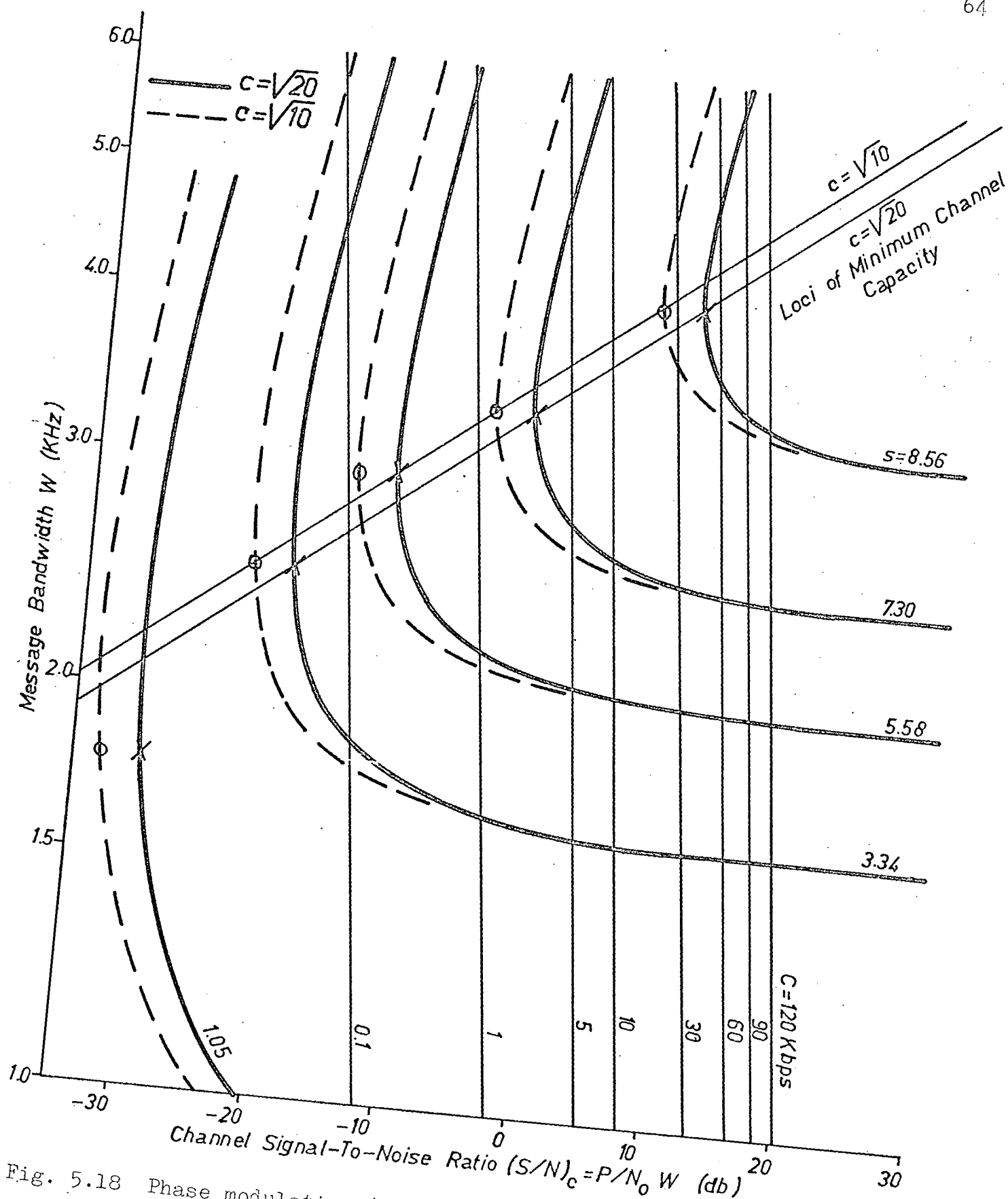
Fig. 5.18 Phase modulation (PM) isopreference contours and con-
tours of constant channel capacity for the peak fre-
quency deviation $|\Delta f|_{max}$ = 75 KHz. The contours are
plotted for peak factors c = $\sqrt{20}$ and c = $\sqrt{10}$.
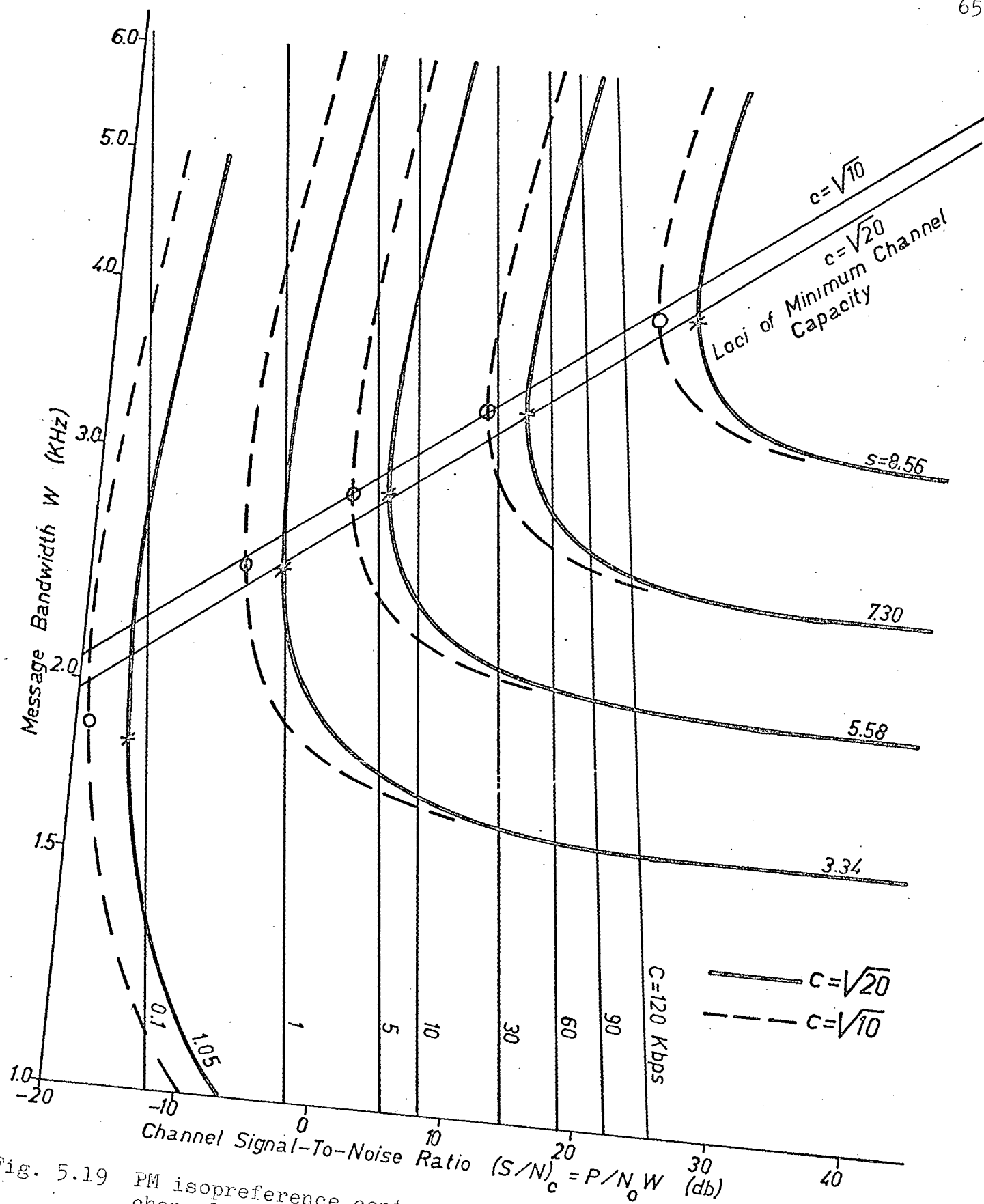
Fig. 5.19 PM isopreference contours and contours of constant channel capacity for the peak frequency deviation $|\Delta f|_{max} = 15$ KHz. The contours are plotted for peak factors $c = \sqrt{20}$ and $c = \sqrt{10}$.

The definition of the rate distortion function $R(\varepsilon)$, does not specify the source of the mean square error ε. In analog communication systems, the distortion of the transmitted waveform by channel noise is the primary cause of error. However, in transmitting information from a discrete source, it may be shown that the probability of occurrence of channel errors may be made arbitrarily small provided that the data rate R_1 of the discrete source satisfy $R(\varepsilon) \leqslant R_1 \leqslant C$. The error in this case originates mainly in the sampling and digitization process in the source encoder. It is therefore of interest to compare the performances of analog and digital communication schemes on the basis of required channel capacity.

The channel bandwidth W_c required by a SSB-SC system is equal to the bandwidth W of the message and that required by a DSB-SC system is equal to twice W. Therefore, for SSB-SC and DSB-SC, (5.3) may be written (assuming that the noise power is constant over the channel bandwidth)

$$C_{SSB-SC} = W \log_2 \ (1 + (S/N)_c)$$
$$C_{DSB-SC} = 2W \log_2 (1 + (1/2)(S/N)_c)$$

$$(5.5)$$

where $(S/N)_c = P/N_o W$ is the channel signal-to-noise ratio. These equations have been plotted in Figure 5.17 for several values of C_{SSB-SC} and C_{DSB-SC}. Marked on the isopreference curves are the points for which C_{SSB-SC} and C_{DSB-SC} are minimum for each scale value. The loci of these points are shown and were used to derive the SSB-SC and DSB-SC curves in Figure 5.18.

The channel bandwidth required by a PM system is given approximately by equation (2.23). The required channel capacity for PM communication is therefore given by

$$C = 2(|\Delta f|_{max} + 2W) \log_2 \left(1 + \frac{1}{2}\left(\frac{S}{N}\right)_C \left(\frac{W}{|\Delta f|_{max} - 2W}\right)\right).$$

This equation was plotted in Figures 5.18 and 5.19 for $|\Delta f|_{max} = 75$ KHz and $|\Delta f|_{max} = 15$ KHz and for $c = \sqrt{20}$ and $c = \sqrt{10}$. The loci of the points of minimum channel capacity were then obtained and used to derive the PM contours in Figure 5.20. However, since as $(S/N)_C$ decreases in a PM system, the approximation $2N_oW \ll A^2$ made in the derivation of (2.25) becomes invalid. This phenomenon known as "thresholding" occurs for $(S/N)_C \approx 25$ db for $|\Delta f|_{max} = 75$ KHz and $(S/N)_C \approx 18$ db for $|\Delta f|_{max} = 15$ KHz*. For $c = \sqrt{20}$, this corresponds to a channel capacity $C \cong 400$ Kbps for $|\Delta f|_{max} = 75$ KHz and $C \cong 70$ Kbps for $|\Delta f|_{max} = 15$ KHz. Although this appears to invalidate the PM results shown in Figure 5.20, the asymptotic behaviour of these curves is still useful in estimating the performance of PM as compared to other communication systems which use large channel capacities.

The minimum bit rate curves shown in Figure 5.15 correspond to the discrete source data rate R_1 mentioned earlier. By letting C_{DPCM}, the DPCM required channel capacity, equal R_1/k, the DPCM curves of minimum channel capacity shown in Figure 5.18 were drawn. The parameter k is a measure of the efficiency of the code used. A value of $k = 1$ implies perfect coding in the sense that negligible channel error occurs without the necessity of increasing the channel capacity. The values of $k = 3/4$ and $k = 1/2$ indicate codes which require channel capacities equal to $(4/3)R_1$ and $2R_1$ respectively, in order to obtain the required negligible probability of a transmission error. The remaining curve in Figure 5.20 is derived from the locus of minimum bit rate for PCM shown in Figure 5.15 and is shown only for perfect coding.

From Figure 5.20 it may be concluded that for low values of

* These values were estimated using equation (7) of Wojnar [41].
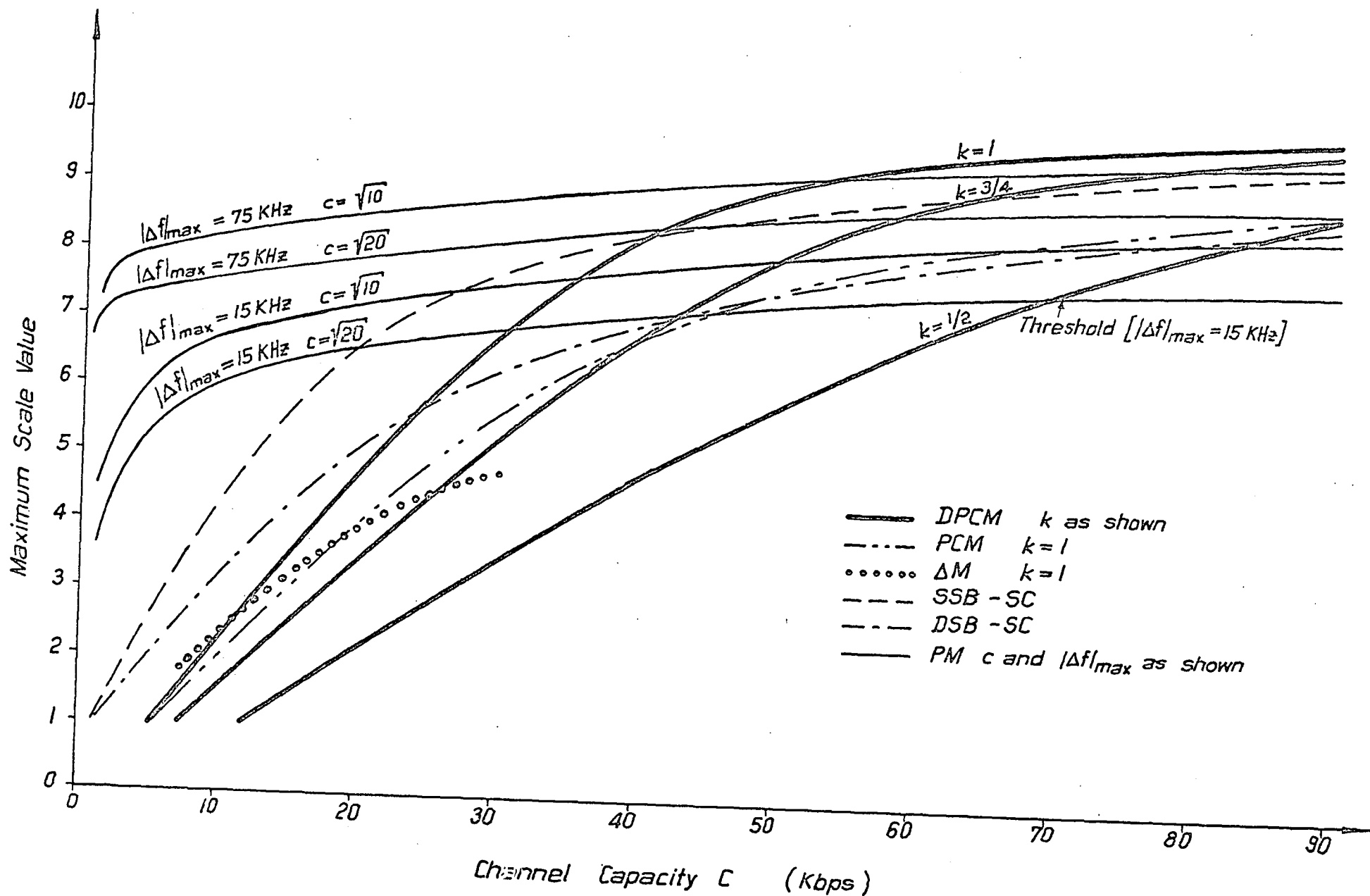
Fig. 5.20 Curves of maximum attainable scale value vs available channel capacity for DPCM, PCM, ΔM, DSB-SC, SSB-SC and PM. The parameter k represents a measure of coding efficiency.

channel capacity, SSB-SC and DSB-SC outperform DPCM, ΔM, and PCM in speech communication. The reason for this may be traced to the distinctly structured quality of the noise created by digital systems, as opposed to the randomness of the noise heard in analog systems. Humans appear to be more annoyed by noise with strong structure than by random noise* and, in the region of the curves under consideration, the noise levels are quite large. As the channel capacity is increased, the PCM and DPCM systems are permitted to quantize more finely and sample more often. The finer quantization causes the granular structure of the digital noise to be less noticeable. As a result, the PCM curve gradually approaches the DSB-SC curve, until for very high quality speech communications, PCM becomes more efficient than DSB-SC (for 100% efficient coding). The fact that DPCM is better than PCM is a result of the increased redundancy reduction capabilities (resulting from an improved prediction) afforded by the finer quantization and higher sampling frequency. In effect R_1 approaches $R(\varepsilon)$ more closely. The results in Figure 5.20 also indicate that for large channel capacities, DPCM outperforms PM for speech communication.

It is interesting to compare these results with the theoretical results presented by Goblick [24]. Goblick concludes that unless the channel signal-to-noise ratio is substantial and, in the case of digital communications, the source digitization efficient, the modulation systems which expand bandwidth cannot be used to good advantage. Digital and angle modulation systems are such systems. It may be seen that the conclusions reached here agree with those of Goblick.

5.4 Transitivity Checks

In Figures 5.6, 5.8, and 5.17 are shown points marked T1, T2, T3 and T4. AM reference points 3, 4, and 5 were used in isopre-

---

* Similar responses have been reported for picture quantization. See for example [22] and [23].

ference tests to obtain the equal preference DPCM points T1, T2 and T3, and DPCM reference point 2 was used to obtain T4. The scale value corresponding to each of these reference points is shown next to the obtained point.

From these points, it appears that the results obtained for T3 and T4 are not as consistent as one might expect. One of the factors contributing to the inconsistency may be the inaccuracies caused by terminating (at 0 and 10) the tails of the distribution of scale values curve. Another factor is the difference in types of noise appearing in the two systems, a difference which is most apparent for the lower scale values. Although this limits to some extent the validity of the comparison of the DSB-SC and DPCM results, it does not necessarily have any bearing on results obtained by comparing one set of curves.

5.5 Concluding Remarks

The results obtained here indicate that for DPCM the sampling frequency should be made as close to the Nyquist rate as possible without causing large aliasing errors. However, restriction (3) in Section 4.2.1 (which limits the amplitude of the input to the quantizer) may not be a practical one with the result that slope overload noise [14] becomes a major problem. Since this form of noise is very dependent on the sampling frequency, it is reasonable to expect that an increase in sampling frequency may be very desirable. A possible approach to this problem would be to use some form of adaptive gain at the input.

APPENDIX I

Feedback Coefficient Optimization

The mean square error $\overline{e^2(t)}$ at the output of the predictor shown in Fig. A.1.1 is given by

$$\overline{e^2(t)} = \overline{(x - \hat{x})^2}$$
$$= \overline{\left[x - \sum_{k=1}^{N} \alpha_k x(t-kT) \right]^2} \qquad (A.1.1)$$

This equation is minimized when the α_k's are solutions to the following set of linear equations [27]

$$\emptyset_j = \sum_{j=1}^{N} \alpha_i \emptyset_{j-i} \qquad (A.1.2)$$

where $\emptyset_k \triangleq R_x(kT)/R_x(0)$ and $R_x(\tau)$ is the autocorrelation function of $x(t)$. The optimum α_k's may be written

$$\alpha_k = \frac{|P_k|}{|P|} \qquad (A.1.3)$$
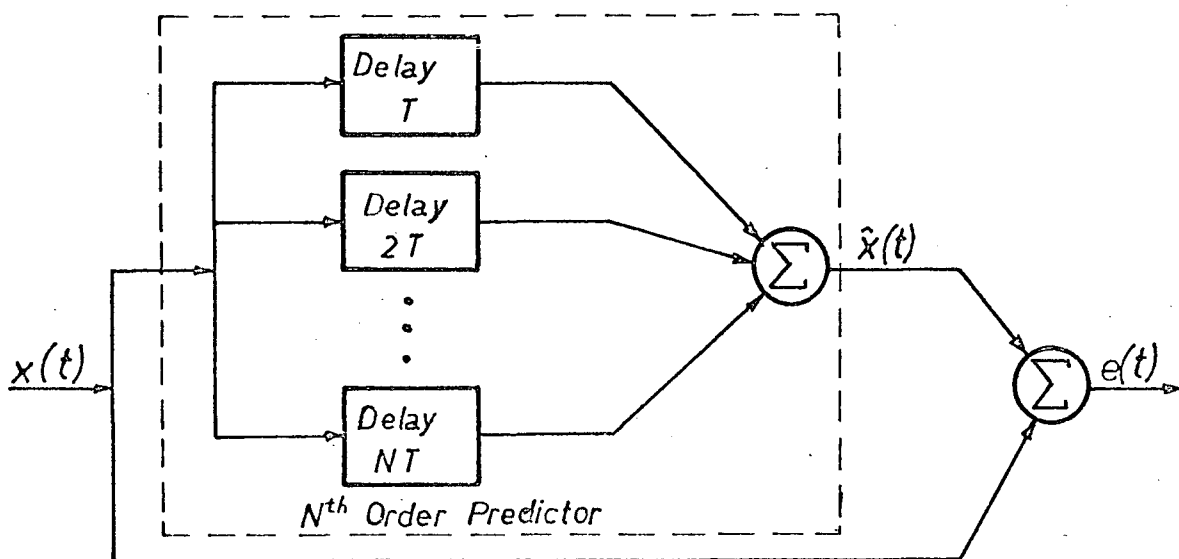
Fig. A.1.1 An Nth order linear predictor.

where $|P|$ is the determinant of the matrix of the coefficients of the α_k's on the right hand side of (A.1.2) and P_k is the determinant of P with its kth column replaced by the column vector $(\emptyset_1, \emptyset_2, \ldots, \emptyset_N)^t$. The resultant mean square error is given by

$$\overline{e_{on}^2(t)} \triangleq \overline{e_o^2(t)} \Big/ \overline{x^2(t)} = 1 - \sum_{i=1}^{N} \alpha_i \emptyset_i \qquad (A.1.4)$$

The optimum α_k's are plotted in Fig. A.1.2, for N = 1, 2, 3 for the particular case of speech of bandwidth greater than 1.5 KH_z.

For N > 1 and high sampling frequencies, the evaluation of the optimum α_k's becomes very difficult, since $|P|$ and $|P_k|$ approach zero resulting in a ratio of two very small quantities. Since, for x(t) wide-sense stationary, $R_x(\tau) = R_x(-\tau)$, $\emptyset_k$ may be approximated by a polynomial of the form

$$\emptyset_k \cong 1 - a_1 (kT)^2 - a_2(kT)^4 - \ldots - a_{N-1}(kT)^{2(N-1)}.$$

Using this approximation (A.1.1) may be written

$$\overline{e^2} \Big/ \overline{x^2} = f_o(\vec{\alpha}) + f_1(\vec{\alpha})a_1 T^2 + \ldots + f_{N-1}(\vec{\alpha})a_{N-1}T^{2(N-1)}.$$

where $\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_N)$, and $f_i(\alpha)$ is a function representing the dependence of the ith coefficient on $\vec{\alpha}$. By choosing $\vec{\alpha}$ to be a solution of

$$f_i(\vec{\alpha}) = 0 \qquad (I = 1, 2, \ldots, N) \qquad (A.1.5)$$

$\overline{e^2} \Big/ \overline{x^2}$ is made equal to zero to an $N-1^{th}$ order approximation. The solutions to A.1.5 are represented by the asymptotes of the curves in Fig. A.1.2 as $T \to 0 (f_s \to \infty)$.

The discrete values $(\hat{\alpha}_1)$ of α_1 used in the previous sample feedback (N = 1) experiments are shown in Fig. 4.2. A measure of the

effect of using $\hat{\alpha}_1$ may be obtained by defining the function $r_g = \overline{e_a^2} \Big/ \overline{e_o^2}$ where $\overline{e_a^2}$ is the mean square error obtained by using $\hat{\alpha}_1$. A large value of r_g indicates a bad choice of $\hat{\alpha}_1$. The maximum values of r_g, which occur for the values of time delay T at which $\hat{\alpha}_1$ changes values, are tabulated in Table A.1.1. For $f_s \lesssim 15\text{KHz}(T \gtrsim .067\text{ms})$, it may be seen that the deterioration in performance is less than 0.1 db.

Table A.1   Values of $r_g = \overline{e_a^2} \Big/ \overline{e_o^2}$ corresponding to points of maximum discrepancy between α_1 and its approximation $\hat{\alpha}_1$ (Bandwidth $W \gtrsim 1.5\text{KHz}$)

f_s (KHz)	T(ms)	$\emptyset_1$	$\hat{\alpha}_1$	r_g	r_g(db)
2.56	.39	.38	.33	1.0029*	.001
2.56	.39	.38	.45	1.057*	.025
3.51	.29	.534	.45	1.0099	.043
3.51	.29	.534	.60	1.0061	.026
4.3	.233	.674	.60	1.0100	.043
4.3	.233	.674	.70	1.0012	.005
5.0	.2	.754	.70	1.0068	.029
5.0	.2	.754	.80	1.0049	.021
6.9	.145	.864	.80	1.0162	.070
6.9	.145	.864	.90	1.00512	.022
10.0	.1	.936	.90	1.0105	.045
10.0	.1	.936	.95	1.0016	.007
15.0	.075	.962	.95	1.0013	.006
33.0	.0303	.994	.95	1.162	.652

However, for $f_s \gtrsim 15\text{KHz}(T \lesssim 0.067 \text{ ms})$, stability considerations dictated that $\hat{\alpha}_1$ not be too close to unity. Therefore, $\hat{\alpha}_1$ was limited to 0.95. As a result, as f_s is increased from 15KHz, r_g also increases. The maximum frequency considered in the experiments was 33KHz. The corresponding value of r_g is included in Table A.1.1. Similar results may be derived for the 2 and 3 sample feedback cases, although considerable computation is required.

---

* These values were calculated for $W = 1.0\text{KHz}$ since for $f_s \lesssim 3.3\text{KHz}$ the autocorrelation function for $W = 1.5\text{KHz}$ does not apply.
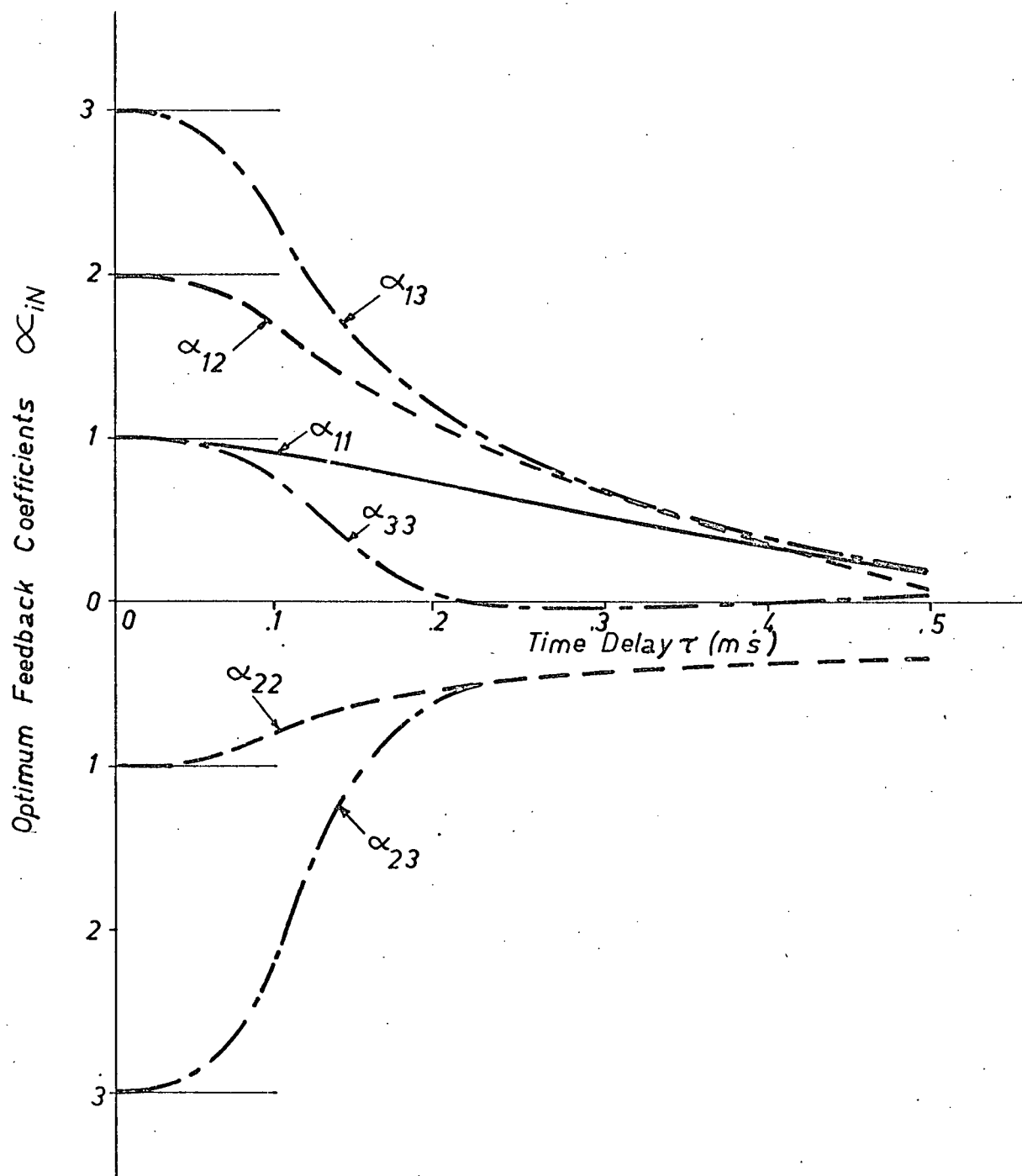
Fig. A.1.2   Optimum predictor coefficients α_{ij}. The
subscript i refers to 1, 2, or 3 intervals
of time delay and subscript j refers to
the order of the predictor.

APPENDIX II

Rating Test Data

A.2.1   Results of DPCM Pilot Rating Tests

L	W(KHz)	r	Sample Mean	Std. Dev.		L	W(KHz)	r	Sample Mean	Std. Dev.
4	1.01	1.1	2.2	1.21		6	2.12	1.1	8.15	1.20
3	3.17	1.25	5.94	1.29		6	1.01	1.1	2.02	.94
3	4.2	1.1	5.37	1.59		4	2.63	3.0	8.15	1.17
5	3.17	3.0	7.68	1.30		5	1.01	1.25	2.40	1.26
3	2.63	1.5	4.92	1.61		3	1.55	1.5	2.60	1.08
5	2.63	2.4	7.82	.97		1	2.12	1.25	.75	1.02
5	2.63	1.5	7.97	1.06		1	2.12	1.9	1.70	1.33
2	1.01	2.4	.85	.60		5	2.12	3.0	6.95	1.50
1	6.3	1.25	2.45	1.26		2	2.12	1.1	2.30	1.24
1	2.63	2.4	2.85	.94		1	3.17	1.9	2.48	1.33
6	2.63	1.25	7.35	1.20		2	2.63	1.25	2.63	1.32
3	2.12	1.25	4.43	1.50		1	1.55	1.1	.43	1.62
2	6.3	2.4	7.12	1.66		6	2.12	1.5	6.78	1.34
6	1.01	1.5	2.5	1.13		4	3.17	2.4	8.58	.98
6	4.2	1.25	8.35	1.13		3	1.01	1.9	2.18	1.08
2	2.63	3.0	6.35	1.30		1	4.2	1.5	2.35	1.00
4	2.12	1.1	4.37	1.29		2	1.01	1.1	.73	.67
2	2.63	1.9	4.68	1.47		1	3.17	3.0	3.72	1.10
2	4.2	1.9	5.33	1.58		4	1.55	3.0	4.78	1.42
1	1.01	3.0	.85	.55		2	2.12	1.5	4.12	1.26
4	2.12	2.4	6.18	1.45		2	2.12	2.4	4.97	1.10
6	1.55	3.0	5.03	1.43		3	1.55	2.4	3.87	1.09
1	1.01	1.25	.65	.63		5	1.55	1.1	3.62	1.23
3	2.63	1.1	3.78	1.20		1	2.63	1.5	1.93	1.15
2	4.2	3.0	7.28	1.61		5	4.2	1.5	8.70	1.37
5	1.01	1.9	2.80	1.38		4	2.12	1.5	6.78	.85
2	1.01	1.5	.93	.75		4	1.01	1.5	2.75	1.05
6	2.12	2.4	6.43	1.57		4	1.01	2.4	2.67	1.08
5	1.01	3.0	3.15	1.28		4	1.55	1.9	4.68	1.36
3	1.01	3.0	2.70	.92		2	3.17	1.1	3.10	1.26
2	3.17	2.4	5.72	1.56		2	6.3	1.5	4.92	1.60
5	1.55	2.4	4.62	1.28		6	2.63	1.9	7.85	1.15
1	3.17	1.25	1.89	1.34		1	4.2	2.4	4.40	1.43
5	2.12	1.25	6.38	1.34		3	4.2	3.7	7.15	1.28
5	3.17	1.25	8.27	.87		1	6.3	1.9	4.10	1.45
3	6.3	1.25	7.08	1.53		2	3.17	1.5	3.68	.98
3	2.63	2.4	7.22	1.34		1	1.55	2.4	.98	.94
4	3.17	1.1	6.75	1.43		4	1.55	1.25	3.48	1.46
6	6.3	1.1	8.60	.91		6	1.01	2.4	2.95	1.66
6	3.17	1.5	7.90	1.12		4	4.2	1.25	7.82	.95
1	1.55	1.5	.62	.72		2	1.55	1.25	1.50	.68
4	3.17	1.5	7.53	1.09		1	2.12	3.0	1.73	.84
3	2.12	3.0	6.05	1.63		3	3.17	3.0	8.15	1.34
3	1.55	1.1	2.13	1.06		4	6.3	1.5	8.78	1.04
6	2.12	1.1	5.85	1.71		5	2.63	1.1	6.98	1.03

Parameters			Sample Mean	Std. Dev.
L	W(KHz)	r		
3	4.2	2.4	7.73	1.24
1	1.01	1.9	.33	.42
2	6.3	1.1	3.95	1.51
3	2.12	1.9	5.05	1.51
3	4.2	1.5	6.12	1.58
5	3.17	1.9	8.73	.96
6	1.55	1.9	3.85	1.25
3	1.01	1.25	1.50	.97
4	6.3	1.1	7.58	1.34
5	2.12	1.9	6.50	1.24
2	3.17	3.7	7.32	1.21
3	3.17	1.9	7.10	1.09
5	4.2	1.1	8.27	1.11
5	1.55	1.5	3.88	1.50
2	4.2	1.25	3.63	1.27
6	1.55	1.25	3.95	1.55
2	1.55	1.9	1.93	.91
1	4.2	3.7	3.88	1.32
4	4.2	1.9	8.57	.96
1	2.63	1.1	1.17	.97
3	6.3	1.9	7.55	1.33
1	4.2	1.1	1.53	.90
2	1.55	3.0	2.83	1.20
4	2.63	1.25	6.53	1.31
4	2.63	1.9	8.25	.97
5	6.3	1.25	9.15	.69

Table A.2.2   DPCM Reference Point Rating Tests

Parameters				Sample Mean	Std. Dev.
N*	L	W(KHz)	r		
3	4	2.12	1.375	4.16	1.31
2	4	2.12	1.375	4.29	1.34
2	3	2.63	1.375	4.55	1.45
2	4	2.63	1.65	7.11	1.54
2	5	3.17	1.65	8.175	1.27
2	2	1.55	1.1	1.09	0.93
2	5	2.63	1.375	7.81	1.62
2	3	2.12	1.1	3.875	1.31
2	3	2.12	1.21	4.275	1.36
2	4	2.12	1.21	3.925	1.44
1	1	3.17	2.5	2.46	1.02

Parameters				Sample Mean	Std. Dev.
N*	L	W(KHz)	r		
1	4	2.12	1.375	4.14	1.35
1	1	3.17	4.0	3.35	1.50
1	4	2.63	1.65	6.91	1.71
1	3	2.12	1.21	4.06	1.42
1	5	2.63	1.375	7.45	1.58
1	3	2.63	1.375	4.88	1.48
1	4	2.12	1.21	3.84	1.42
1	1	3.17	3.5	2.675	1.27
1	2	1.55	1.1	0.81	0.63
1	5	3.17	1.65	8.88	1.63
1	3	2.12	1.1	3.36	1.30
1	1	3.17	3.0	2.48	1.08

* N = number of samples of feedback.

Table A.2.3 AM Pilot Rating Tests

S/N(db)	L	W(KHz)	r	Mean	Std. Dev.	S/N(db)	L	W(KHz)	r	Mean	Std. Dev.
21.1		4.2		7.05	1.09	17.6		1.55		4.13	1.81
20.1		1.01		2.5	1.46	7.03		1.01		1.45	1.36
DPCM	5	2.63	1.375	8.48	1.06	DPCM	5	2.12	1.1	6.8	1.73
21.5		2.63		6.85	1.39	19.1		2.63		6.75	1.60
DPCM	6	3.17	1.375	9.05	.81	DPCM	3	2.63	1.375	6.08	1.17
DPCM	2	3.17	1.375	4.18	1.35	23.3		6.3		7.48	1.4
DPCM	4	4.2	1.1	8.13	1.19	18.2		3.17		6.75	1.18
11.0		1.01		1.63	.97	9.13		2.12		4.2	1.20
DPCM	2	3.17	2.5	7.08	1.23	DPCM	7	2.12	1.65	7.05	1.97
DPCM	2	4.2	1.65	5.18	1.28	DPCM	2	1.55	1.1	1.05	.81
DPCM	3	2.63	2.5	8.45	1.27	DPCM	5	3.17	1.65	9.48	.80
16.3		2.63		5.95	1.16	DPCM	1	3.17	1.65	2.05	1.09
34.2		6.3		9.13	.69	DPCM	1	2.63	2.5	2.9	1.20
DPCM	1	3.17	1.1	1.5	.93	21.7		3.17		7.43	1.04
DPCM	3	2.12	1.1	4.38	1.55	14.6		2.12		5.5	1.48
DPCM	4	2.63	1.65	8.43	1.24	11.4		1.55		3.63	1.32

Table A.2.4 AM Reference Point Rating Tests

S/N(db)	L	W(KHz)	r	Mean	Std. Dev.	S/N(db)	L	W(KHz)	r	Mean	Std. Dev.
9.5		2.12		3.34	1.59	27.0		3.17		3.27	1.44
DPCM	7	2.12	1.65	5.95	1.96	17.8		1.55		4.52	1.71
7.0		1.01		1.69	1.27	10.0		6.3		4.52	1.71
DPCM	6	3.17	1.375	7.72	1.63	24.6		3.17		7.30	1.33
DPCM	3	2.12	1.1	3.31	1.37	DPCM	5	3.17	1.65	8.88	1.25
60.0*		1.55		4.41	1.73	60.0*		2.63		7.03	1.81
-3.0		4.2		1.75	1.14	23.0		3.17		6.89	1.40
18.2		3.17		6.1	1.48	DPCM	3	2.63	1.375	4.72	1.73
DPCM	4	2.63	1.65	6.83	1.85	60.0*		4.2		9.27	.81
23.0		6.3		6.64	1.56	60.0*		2.12		4.66	1.82
DPCM	2	4.2	1.65	4.30	1.75	31.75		1.55		3.81	1.46
33.0		6.3		8.66	.91	24.8		4.2		7.53	1.82
21.1		4.2		6.72	1.42	DPCM	2	1.55	1.1	1.22	1.24
60.0*		1.0		2.5	1.30	60.0*		3.17		8.73	1.33
3.0		1.55		1.63	1.11	27.0		4.2		7.42	1.19
30.8		4.2		8.08	1.23	3.0		1.23		1.05	.73
7.0		1.55		2.33	1.76	34.8		4.2		8.56	1.27
16.3		2.63		5.58	1.44	60.0*		1.23		3.25	1.16
19.0		2.63		4.38	1.52	DPCM	1	3.17	1.1	1.08	1.02
14.0		2.12		6.98	1.29	21.8		2.63		6.45	1.35

* The value of 60 db in these tables corresponds to zero noise added to the speech sample. (See Section 5.2.1)

REFERENCES

1. Max, J. "Quantizing for Minimum Distortion", IRE Trans. on Information Theory, vol. IT-6, pp. 7-12, March, 1960.

2. Williams, G. "Quantizing for Minimum Error with Particular Reference to Speech", Electronic Letters, vol. 3, pp. 134-135, April, 1967.

3. Bruce, J.D. "Optimum Quantization", Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Tech. Rept. 429, March, 1965.

4. Panter, P.F. and Dite, W., "Quantization Distortion in Pulse-Count Modulation with Non-uniform Spacing of Levels", Proc. IRE, vol. 39, pp. 44-48, January, 1951.

5. Donaldson, R.W., "Optimization of PCM Systems Which Use Natural Binary Codes, Proc. IEEE (Corres.), vol. 56, pp. 1252-1253, July, 1968.

6. Liff, A.I., "Mean-Square Reconstruction Error", IEEE Trans. on Automatic Control (Correspondence), vol. AC-10, pp. 370-371, July, 1965.

7. Katzenelson, J. "On Errors Introduced by Combined Sampling and Quantization", IRE Trans. on Automatic Control, vol. AC-7, pp. 58-68, April, 1962.

8. Steiglitz, K. "Transmission of an Analog Signal over a Fixed Bit-Rate Channel", IEEE Trans. on Information Theory, vol. IT-12, pp. 469-474, October, 1966.

9. Goodman, L.M. and Drouilhet, P.R., Jr., "Asymptotically Optimum Pre-Emphasis and De-Emphasis Networks for Sampling and Quantizing", Proc. IEEE (Correspondence), vol. 54, pp. 795-796, May, 1966.

10. Van de Weg, H., "Quantizing Noise of a Single Integration Delta Modulation System with an N-Digit Code", Philips Research Rept., vol. 8, pp. 367-385, 1953.

11. Nitadori, Kazuhiko, "Statistical Analysis of ΔPCM", "Electronics and Communications in Japan, vol. 48, pp. 17-26, February, 1965.

12. O'Neal, J.B., "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals", Bell Sys. Tech. J., vol. 45, pp. 689-721, May-June, 1966.

13. O'Neal, J.B., "A Bound on Signal-to-Quantizing Noise Ratios for Digital Encoding Systems", Proc. IEEE, vol. 55, pp. 287-292, March, 1967.

14. McDonald, R.A., "Signal-to-Noise and Idle Channel Performance of Differential Pulse Code Modulation Systems-Particular Application to Voice Signals", Bell Sys. Tech. J., vol. 45, pp. 1123-1151, September 1966.

15. Irwin, J.D. and O'Neal, J.B., "The Design of Optimum DPCM (Differential Pulse Code Modulation) Encoding Systems Via the Kalman Predictor", Joint Automatic Control Conference, 1968, pp. 130-136.

16. Donaldson, R.W. and Chan, D., "Analysis and Subjective Evaluation of Differential Pulse Code Modulation Voice Communication Systems", To be published in IEEE Trans. on Communications Technology.

17. O'Neal, J.B., "Delta Modulation Quantizing Noise Analytical and Computer Simulation Results of Gaussian and Television Input Signals", Bell Sys. Tech. J., vol. 45, pp. 117-141, January, 1966.

18. Sharma, P.D., "Characteristics of Asynchronous Delta-Modulation and Binary-Slope-Quantized - p.c.m Systems. Second in the Series of Articles on Feedback Modulation Systems", Electronic Engineering (GB), vol. 40, pp. 32-37, January, 1968.

19. Hosakawa, S., Onaga, K., Katusho, O., and Kato, K., "A Type of Companded Delta Modulation Coder", Electronics and Communications in Japan, vol. 50, August, 1967.

20. Abate, J.E., "Linear and Adaptive Delta Modulation", Proc. IEEE, vol. 55, pp. 298-308, March, 1967.

21. Brainard, R.C., "Subjective Evaluation of PCM Noise-Feedback Coder for Television, Proc. IEEE, vol. 55, pp. 346-353, March, 1967.

22. Huang, T.S., and Chikhaoui, M., "The Effect of BSC on PCM Picture Quality", IEEE Trans. on Information Theory, vol. IT-13, pp. 270-273, April, 1967.

23. Schreiber, W.F., "Picture Coding", Proc. IEEE, vol. 55, pp. 320-331, March, 1967.

24. Goblick, J., Jr., "Theoretical Limitations on Transmission of Data from Analog Sources", IEEE Trans. on Information Theory, vol. IT-11, pp. 558-567, October, 1965.

25. Wozencraft, J.M. and Jacobs, I.M., Principles of Communication Engineering, New York: J. Wiley and Sons, 1965, pp. 320-323, pp. 645-654.

26. Mayo, J.S., "A Bipolar Repeater for Pulse Code Signals", Bell Sys. Tech. J., vol. 41, pp. 96, January, 1962.

27. Papoulis, A., Probability, Random Variables, and Stochastic, Processes, New York: McGraw-Hill, 1965, pp. 390-394.

28. Bennett, W.R., "Spectra of Quantized Signals", Bell Sys. Tech. J., vol. 27, pp. 446-472, July, 1948.

29. Widrow, B., "Statistical Analysis of Amplitude Quantized Sampled-Data Systems", Trans. of AIEE vol. 79, pp. 555-567, January, 1961.

30. Plotkin, S.C., "FM Bandwidth as a Function of Distortion and Modulation Index", IEEE Trans. on Communication Technology, vol. COM-15, pp. 467-471, June, 1967.

31. Sakrison, David, Communication Theory: Transmission of Waveforms and Digital Information, New York: J. Wiley and Sons, 1968, pp. 164-192.

32. Chan, Donald, "Analytical and Subjective Evaluation of Differential Pulse Code Modulation Voice Communication Systems", M.A.Sc. Thesis, University of British Columbia, British Columbia, Canada, 1967.

33. Lee, Y.W., Statistical Theory of Communication, New York: J. Wiley and Sons, 1963, pp. 262-264.

34. Torgerson, W.S., Theory and Methods of Scaling, New York: J. Wiley and Sons, 1958.

35. Munson, W.A. and Karlin, J.E., "Isopreference Method for Evaluating Speech-Transmission Circuits", J. Acoust. Soc. Am., vol. 34, pp. 762-774, June, 1962.

36. French, N.R. and Steinberg, J.C., "Factors Governing the Intelligibility of Speech Sounds", J. Acoust. Soc. Am., vol. 19, pp. 90-119, January, 1947.

37. Protonotarios, E.N., "Slope Overload Noise in Differential Pulse Code Modulation Systems", Bell Sys. Tech. J., vol. 46, pp. 2119-2161, November, 1967.

38. Beranek, L., "The Design of Speech Communication Systems", Proc. IRE, vol. 35, pp. 880-890, September, 1947.

39. Terman, F.E., Electronic and Radio Engineering, New York, McGraw-Hill, 1955, pp. 964-965.

40. Purton, R., "A Survey of Telephone Speech-Signal Statistics and Their Significance in the Choice of a P.C.M. Companding Law", Proc. IEE, pp. 60-66, January, 1962.

41. Wojnar, A., "Noise and Threshold in FM Systems", Proc. IEEE (Corres.), vol. 55, pp. 1639-1641, September, 1967.