Institute of Acoustics

Oversampling and noise shaping for digital to analogue conversion

Dr Malcolm J. Hawksford

University of Essex, Wivenhoe Park, Colchester, Essex, England

Introduction

In this paper we present results of a computer simulation of a digital to analogue conversion scheme which uses a combination of oversampling and noise shaping. The rationale for this approach is, first, the ability to implement more optimum linear phase reconstruction filters offering exact time synchronisation; second, to implement schemes where all the non-ideal behaviour of the digital to analogue converter (DAC) is translated to benign noise and, thirdly, simplification of analogue circuitry which should lead to greater signal transparency.

Present methods of oversampling in digital audio are extensions and adaptations of the work of Philips [1] where ratios of x2, x4 and x16 have emerged. The advantage of improved signal reconstruction using digital interpolation is evident in these schemes, where substantial filtering is performed in the digital domain. As such, the filter response can offer both a rapid attenuation with frequency as well as near ideal group delay characteristics. Also, where 16 bit DACs are used with basic noise shaping, improvements in low-level resolution can be attained yielding a closer approximation to the bound dictated by source data quantisation.

However, there are further advantages to be gained by using even higher oversampling ratios, for example, in the region of x256. In such a scheme the number of samples per Nyquist sample is dramatically extended, allowing further anhancements in reconstruction filtering with corresponding simplifications in analogue design.

To realise the full advantage of oversampling, the process must be combined with that of noise shaping [2]. As such, it becomes feasible to both reduce substantially the number of bits necessary in the DAC and to induce high-level DAC activity such that large scale errors translate to noise. Under such operation the DAC can be viewed as a wide band, spectrally weighted, noise source and therefore takes on a more benign character.

Oversampling and noise shaping for digital to analogue conversion

The process of oversampling and noise shaping can be applied hierarchically to the implementation of DAC schemes. At one end of the range, there is the x1 to x16 oversampling ratios still retaining 14 or 16 bit DACs, while at the other extreme a 1 bit DAC is feasible where the performance is similar to delta-sigma modulation (DSM) [3,4,5]. Indeed, the methods presented here can also be viewed as a more optimum form of PCM to DSM conversion [6].

The method presented in this paper is to investigate the combined processes of oversampling and noise shaping in conjunction with an imperfect DAC using computer simulation. As such, a range of DAC distortion mechanisms can be included, such as both large and small level errors, rise time error, slew rate error and glitch error.

1. Principles of noise shaping and data compaction

The DAC system investigated in this paper requires the joint application of oversampling and noise shaping. Effectively, source data $\{M., f_{s1}\}$ is converted to output data $\{N, f_{s2}\}$ where f_{s1}, f_{s2} are the respective input and output sampling rates ($f_{s2} \gg f_{s1}$) and M bit and N bit are input and output word lengths. Because the input information is restricted to a Nyquist bandwidth of $f_{s1}/2$ Hz and substantial oversampling is employed, information theory enables M<N without loss of source data. Indeed, in the examples presented here, f_{s1} = 44.1KHz, f_{s2} = 5.12MHz, M = 16 bit and N ≈ 4 bit.

The basic processes are shown in Fig.0(a), where a digital low pass filter forms the sampling rate conversion and a noise shaper or level compaction algorithm reduces the number of quantisation levels required in the output code. In this paper, we shall assume ideal characteristics for the low pass filter function and concentrate on the performance of the noise shaper in association with a non-ideal DAC as it is the characteristic of this sub-system which enables substantial error decorrelation of output DAC non-ideality. Since the results are derived by time domain simulation, the elimination of the oversampling filter is trivial as code words of

length M bit at a sampling rate f_{s2}Hz (M = 16 bit, f_{s2} = 5.12MHz) can be generated from within the program. It is also important to note that the N bit wide output DAC is driven open loop from the output of the noise shaper as shown in Fig. 0(a).

The structure of an R^{th} order compaction algorithm is shown in Fig. 0(b) and 0(c). Essentially, it consists of R cascaded digital integrators of transfer function 1/(Z-1), together with (R-1) feedforward paths in order to achieve closed loop stability. The combined outputs of the cascaded integrators are then truncated by quantiser Q, where Q is also matched to the linear output DAC such that 2^R = (number of levels in Q).

Though this processor operates in discrete time at a sampling rate f_{s2}Hz, to appreciate its operation it can be compared with an analogue circuit of a feedback power amplifier with a non-linear output stage. If the forward gain of the feedback amplifier is of a high value, then the input-output transfer characteristic will be less dependent upon output stage non-linearity.

In a similar manner, the effect of cascading integrators is to produce a loop with potentially very high loop gain, particularly at frequencies well below the sampling rate. Consequently, the error waveform generated by output quantiser Q becomes both reduced and frequency shaped by the loop transfer function. For example, for an R^{th} order topology, the noise spectrum follows a curve approximating to 6R dB/octave.

In order to maintain loop stability with a number of integrators, the (R-1) feedforward paths within the noise shaper are paramount, where the structures of Fig. 0(b), (c) can be shown to agree with Tewksbury's [2] estimate for optimum noise shaping. Indeed, simulation confirms a non-divergent response providing the quantiser Q does not saturate [7].

In order to appreciate further how large scale DAC errors can be controlled using this system, the following scaling of signals will prove a useful aid: assume that the quantisation levels of Q are spaced at

intervals of 1 volt (i.e. Q = 1 volt), while the input data of interval δ is quantised to 16 bit (where q = 2^{-16} volt). Also, because the output of Q is fed back directly to the input, the input message (excluding quantisation noise) is of the same scale as the output range. Consequently, for low-level input data, the message signal is well below an output quanta, while for a peak to peak message signal of -1V to +1V, the signal compares with the lower range of the Q quantiser. This observation may at first seem unworkable, but it should be noted that the effect of noise shaping will configure output data sequences which, because of oversampling, will average to represent an accurate estimate of the input message. At this stage it is useful to make comparisons with delta-sigma modulation [3,4] where similar principles can be observed.

To illustrate the time domain behaviour of the noise shaper, Fig. 1 shows four output data rasters for noise shapers of orders R = {1,2,3,4}. The input is set at 0dB ($\sqrt{2}$ volt amplitude) at 8kHz and the sampling rate is 5.12MHz. As the order R of the noise shaper is advanced, more signal activity occurs at the quantiser output, where for R = 4 (Fig. 1(d)), the activity almost appears random and apparently independent of the input. However, computer simulation reveals that by appropriate filtering of the output data sequence, a close approximation to the input is achieved, where the message input-output error reduces as R is advanced.

For interest, the corresponding level and adjacent sample difference histogram distributions of quantiser activity are shown in Fig. 5 (a), (b), for a range of R = {1,2,3,4}. Results show that as R is advanced, the output activity spans more output quanta, where for R = 4 the range is approximately -8 to 8, corresponding to M = 4 bit. Because the system is attempting to resolve information beyond the static constraints of the output quantiser, the high level, high frequency output activity is an essential by-product of the noise shaper.

Computer results also reveal that for $R \geqslant 4$, the envelope and statistical distribution of output quantiser levels are not strongly input signal dependent providing the input is restricted to a range $\approx \pm 1$ volt

(± 1Q). It is this mechanism that allows large-scale DAC errors to be translated to a more benign noise. Effectively, the signal activity acts as a high-level, high frequency intelligent dither that enables all the DAC quantisation levels to participate in the conversion process.

In the next Section, the results of computer simulations are extended to include a range of DAC non-ideality and it is shown that for R = 4, decorrelation is complete, where no harmonic distortion is evident within the output spectral plots.

2. DAC non-ideality and computer generated results

In this Section we illustrate the performance of the oversampled and noise shaped DAC using example results generated by a simulation program of the noise shaper and DAC. The results are accurate predictions of performance as a range of DAC impairment can be included. In presenting the results we emphasise that the actual DAC sub-system is external to the noise shaper feedback path and therefore operates open loop as shown in Fig. 0(a). A perfect DAC would have analogue reconstruction levels that match exactly the levels corresponding to the Q-quantiser (see Fig. 0(b), (c)). However, in practice, some deviation is to be anticipated and the consequences of these deviations are presented in the following results.

The results have been generated using a dedicated transient analysis procedure where time sequences of data are sequentially computed. It is therefore possible to interrogate these sequences (as discussed in Section 1 with reference to Fig. 1 and 5(a)) and by using the Fast Fourier Transform (FFT) to compute the spectral distribution of the noise shaper output, the DAC error spectrum (see Fig. 1(a) for representation of DAC error) and the output DAC spectrum.

2.1 Time domain output sequences for low-level input signal

In Fig. 2, the noise shaper output time domain activity is shown for a sinusoidal input -80dB below $\sqrt{2}Q$ and frequency 8kHz using a sampling rate of 5.12MHz, where all initial conditions are set to zero before

commencement of the input sequence. The results show that for R = 1 (first order, i.e. single integrator in loop) the system fails to respond; for R = 2 there is slight but intermittent activity, while for R = 3,4 the activity and coding accuracy progressively improves. Indeed, it will be shown that for R = 4, the coding performance is acceptable for digital audio applications. Also for R = 4, once initiated, the dither activity is always sustained, while for R < 4 it can intermittently cease, having converged to zero. This characteristic would be unacceptable and lead to low-level burst noise rather akin to low-order delta-sigma modulation [7].

A further low-level time domain response is shown in Fig. 3(a), (b). Here R = 4 and the input is -96dB. To show the coding performance, assuming initially a perfect DAC, the sequence in Fig. 3(a) was filtered by a simple 6th-order coincident pole filter for f_{3dB} = 30kHz. The time domain response of the coded sine wave computed over 512 samples is shown in Fig. 3(b) together with the corresponding spectrum in Fig. 4. In this case the input data is not quantised to 16 bit, to allow only the noise due to the noise shaper to be shown. As can be seen there is a recognisable sine wave of amplitude < q, with superimposed random noise, while the spectrum shows no evidence of harmonic distortion, revealing minimal evidence of discontinuous low-level distortion.

One aspect of performance that is important to observe and relates to loop stability is the requirement for quantiser Q not to be saturated. Simulations have shown that, providing Q does not saturate, the loop activity will not diverge even for an impulsive input. However, should the range of Q be restricted, then once saturation occurs, irreversible instability results. This characteristic is illustrated in the two impulse responses shown in Fig. 3(c), (d).

In practice the induction of instability would be circumvented by four strategies:

(a) Set all variables to zero on start up.
(b) Ensure the range of Q is adequate even if this significantly exceeds the DAC range (this is of little consequence as, under normal operation,

the ranges are matched).

(c) the oversampling low pass filter would ensure minimal hf signal components, which could otherwise instigate instability if saturation occured in Q.

(d) Include saturation/instability detection circuitry, to reset coder in the error event of instability.

Interestingly, it is the wide range of Q that enables R > 2 to be used. If the range of Q is restricted to, say, two levels as in delta-sigma modulation, then high order loop filters cannot be used. This observation was made in an earlier paper [7], where techniques are described which show how the output of an oversampled and noise shaped system can be converted to a 1 bit code, thus allowing delta-sigma modulation to operate with high order systems without instability.

2.2 Static, high-level DAC non-linearity

To determine the performance degradation of static DAC errors, deliberate displacements of DAC reconstruction levels were introduced. The displacements were of two types: a large scale sinusoidal error, and a low-level random error. The composite selected error is illustrated in Fig. 5(c) where level displacement (related to q) is plotted against output level (related to Q).

To demonstrate the effect of this static error, the DAC error waveform (see Fig. 0(a)) was analysed using the FFT and the result recorded in Fig. 5(d) (where a linear frequency scale up to $f_{s2}/2$ is used). An interesting characteristic of this spectrum is that the error takes on a noise weighted characteristic that is partially related to the noise weighted spectrum generated at the output of the noise shaper. A separate study has shown that the noise shaping of the DAC error is responding to the sinusoidal large scale DAC displacements and not the lower-level superimposed random error. This weighting can be predicted by observing that the noise shaper produces mainly high frequency noise.

In this example the 8kHz input is set at 0dB; there is again little evidence of harmonic distortion within the DAC error waveform, decorrelation appears to have been achieved for $R \geqslant 4$.

2.3 Ouput noise spectra for ideal DAC with high level input

Exercising the system with an 8kHz, 0dB signal for R = {1,2,3,4} produces the output error spectra shown in Fig. 6. Results show that for R = 1, the spectrum is a line spectrum with relatively high distortion. Indeed, this was evident in the time domain waveforms. However, as R is increased, improvements can be readily observed, where for R = 4, an excellent target performance is predicted when using a sampling rate of 5.12MHz.

In Fig. 7, the same results are shown, but in this instance the output signal of the DAC is bandlimited by an (R + 1) order coincident pole filter where f_{3dB} = 30kHz. This filter is well suited to the noise shaped distortion and reveals what is achievable with a relatively simple reconstruction filter in the analogue domain. With an (R + 1) order filter, the hf noise should fall approximately at 6dB/octave, though the inband noise below 30kHz is determined by the noise shaper.

2.4 DAC error spectra for high-level input using static non-linearity

In these results, the DAC had the error described in Section 2.2, Fig. 5(c), though in this case the 8kHz input is set to 0dB. The computed spectra are shown for a bandwidth of 100kHz (linear scale) in Fig. 8. Interestingly, for R = 2,3, the 2nd harmonic (16kHz) is just evident. However for R = 4 decorrelation is again complete where the low frequency DAC error spectrum (see Fig. 0(a) for definition of DAC error) is white. The DAC error, therefore, is no longer a coherent distortion and appears as more benign noise.

In Fig. 9, the overall output noise spectra are shown for R ={1,2,3,4} where both the noise shaper and DAC error are combined.

2.5 Glitch distortion for segmented and parallel DAC architectures

Glitch distortion in DACs is a dynamic distortion which depends upon signal transitions from one level to another. Indeed, the characteristic of glitch will depend upon the actual levels involved in the transition. Therefore, it can be said to have a dynamic characteristic. Simplistically, glitch arises due to the small time-response differences that occur within the DAC on each of the effective data lines. For example, when the most significant bit settles, this may not be synchronised with that of, say, the least significant bit. Thus, the output is momentarily in error until both events have settled. This error is termed a glitch. Hence, we can relate one source of glitch to internal timing; consequently, the resulting error is a function of the type of DAC and its internal functionality.

In this study, we included basic models for two types of DAC:

(i) segmented DAC, where each binary input is weighted 2^0, 2^1, 2_2 ... etc., and differential timing errors are introduced on a per bit basis;

(ii) a parallel DAC, where a coded word is loaded to a parallel register and each output is given equal weight.

Because the sampling rate is high (5.12MHz), the model need only predict the error in sample area due to glitch and add the appropriate error to the sample value. The exact shape of the glitch is of little consequence, because the glitch duration is only a small fraction of a Nyquist sample period.

For each of the two DAC structures, the program selected random pulse delays up to a maximum of 150ns. For the parallel DAC the chosen example maximum glitch area = 1.240 nV-s while for the segmented DAC 0.770 nV-s. The program then interrogated output code for level transitions, introduced the selected time differentials and computed the corresponding glitch error.

Obviously, the data selected is contrived and chosen only by way of example. Nevertheless, the DAC error and overall spectral plots, shown on a logarithmic scale in Fig. 10, clearly illustrate the distortion trends. Because the gain is slightly changed due to glitch error, small input signal residuals remain. However, this is of little consequence. Again, the results show a smooth spectral weighting, where the error characteristics are once more of minimal consequence.

Fortunately, because a parallel DAC for R = 4 only requires 16 levels, a fast parallel structure can be conceived where differential timing delays are small. As a result, the need for a glitch suppression circuit can be eliminated, and the weighted DAC output can feed directly to a passive reconstruction filter.

2.6 Rise time and slew-rate distortion

Possibly the most significant distortion is that due to slew rate [8] where either finite logic rise time or slewing in subsequent stages results in an error in the area of each sample. This distortion is both a function of the logic speed, the samping rate and intersample differences, where the more frequent the data transitions, the more significant the distortion. Thus the distortion degrades further with an increase in sampling rate. In Fig. 12, a pulse rise time is shown that has both a slew limited region and a linear region (exponential in this example), where the slew-rise time T_s and linear rise time T_r are also defined.

It is clear that the greater ΔV_o, the greater is the loss of area SA due to slew rate. The inter-relationship between slew rate limited and linear rise time then produces an error that is non-linearly related to ΔV_o, where the number of "error areas" also increases with the number of transitions.

In the computer program the error area was calculated and the pulse value correspondingly modified. In Fig. 11, results are shown for a R = 4 system, where the slew rate is set to the maximum slope of a 16V pulse

with a 170ns exponential rise time between 10% and 90% limits. This slew rate is then fixed for all examples in Fig. 11. In Fig. 11(a) we see DAC error for only linear rise time errors, while in Fig. 11(b) the linear rise time is zero and there is only slew rate limiting, whereby the DAC error is greater. Fig. 11(d) shows a 0 to 100kHz spectrum of the DAC error in Fig. 11(b). Finally, in Fig. 11(c), the variation of DAC output integrated noise 0 to 20kHz is plotted as a function of $100\ (\frac{T_s}{T_r + T_s})$ where T_s remains constant.

However, it should be noted that although the level of DAC error in Fig. 11(b), (d) is unacceptable, it still retains a noise-like characteristic with no evidence of harmonic distortion.

The results of this Section indicate the need for a DAC with a fast rise time. In practice, as only ≈ 16 levels are required ECL logic or, indeed, other structures using, for example, fast current steering circuitry to reduce the effect of slew distortion can be used.

3. Conclusion

This paper has presented a technique of digital to analogue conversion that exploits the joint application of oversampling and noise shaping in conjunction with a DAC that exhibits several classic error mechanisms.

It has been demonstrated that, using a 4th order noise shaper, virtually complete decorrelation of DAC error with input signal is achieved for all types of non-linearity. The technique exploited the presence of a high level, high frequency intelligent dither at the output of the noise shaper to fully exercise the DAC, where the large scale statistics of the dither did not change with signal, assuming a suitable scaling of input signal to DAC range. In fact, the recommended loading is that the peak input signal does not exceed 1.5Q.

It is believed that the application of these techniques in association with optimal hardware can yield a viable alternative to the present generalisation of DACs for digital audio applications, and is a more optimum

strategy for exploiting oversampling. Results would suggest that, with appropriate DAC design, a resolution significantly in excess of 16 bit could be achieved, together with a performance that exhibits exceptional low-level linearity, that is primarily limited only by source data quantisation.

A technique where accurate and time synchronised recovery filters can be implemented in the digital domain is also welcome, together with the removal of deglitch circuitry, virtual-earth amplifiers processing wide band signals, sampling rate jitter related to a per-Nyquist sample base and the numerous operational amplifiers and filter circuits that are often used in present systems.

At a philosophical level, the data conversion methods presented are part of a broader continuum where in the limit a 1-bit code format is feasible. Indeed, the method allows high-order noise shaping to be introduced to delta-sigma modulation and discrete time pulse width modulation.

There are still several areas to investigate in this research before viable hardware can be produced. Nevertheless, at this interim stage the techniques appear viable and some encouraging predictions can be made on performance potential.

Acknowledgement

The results presented in this paper are generated by a program written Basil McCrea, a recent M.Sc. research student in the Audio Research Group at the University of Essex, and his assistance is gratefully acknowledged.

MJH/lf

October 1987

4. References

1. Goedhart, D., van de Plassche, R.J., and Stikvoort, E.F., "Digital-to-analogue conversion in playing compact disc", *Philips Tech. Rev.* **40**, No.6, 174-179, 1982

2. Tewksbury, S.K. and Hallock, R.W., "Oversampled, Linear Predictive and Noise Shaping Coders of Order N>1", *IEEE Transactions on Circuits and* Systems, **CAS.25/7**,436-447, July 1978

3. Inose, H.L. and Yasuhiko, Yasuda Y., "A unity bit coding method by negative feedback", Proc. Inst. Elec. Electron. Eng., **51**, 1524-1535, 1963

4. Steele, R., *Deltamodulation Systems*, Pentech Press, ISBN 0727304011, 323-327

5. Candy, J.C., "A use of double integration in sigma delta modulation", *IEEE Trans. Comms.* **COM-33**, 249-258, March 1985

6. Naus, P.J., Dijkmans, E.C., Stikvoort, E.F., McKnight, A.J., Holland, D.J., and Bradinal, W., "A CMOS stereo 16-bit D/A converter for digital audio", *IEEE J. Solid-State Circuits*, **SC-22**, No.3, June 1987

7. Hawksford, M.J., "N^{th}-order Recursive sigma-ADC machinery at the analogue-digital gateway", *78th AES Convention, Anaheim*, Preprint **2248 A-15.**

8. Freeman, D.M., "Slewing distortion in digital-to-analogue conversion", *JAES*, **25**, No. 4, 178-183, April 1977

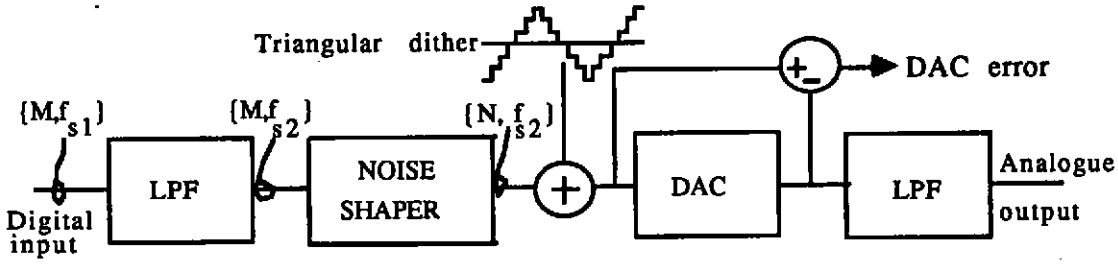


Fig.0a Basic oversampling scheme with level compaction and additive triangular wave

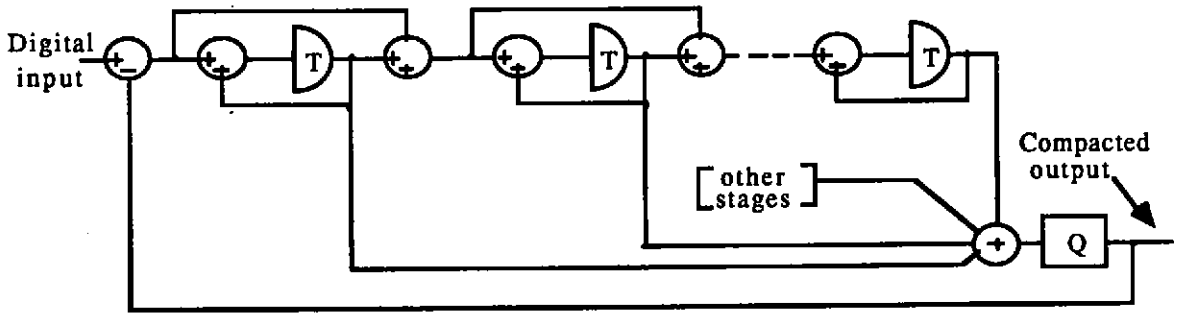


Fig 0b R-th order noise shaping filter for level compaction in DAC applications

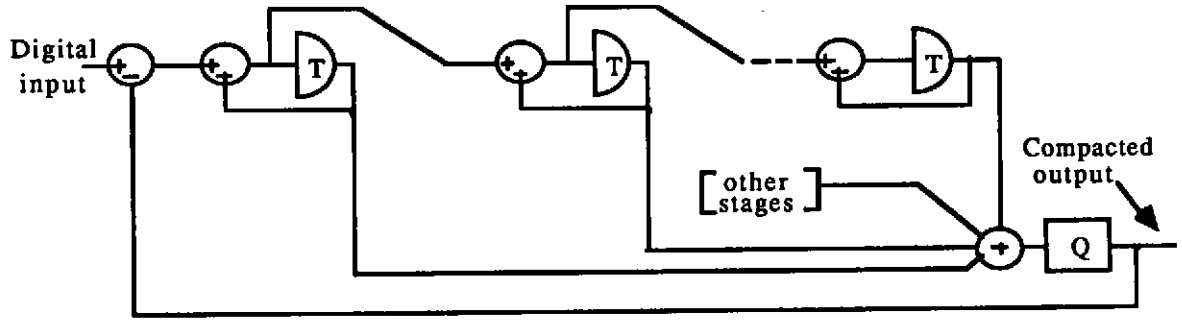


Fig 0c Simplified R-th order noise shaping filter (derived from Fig 0a)

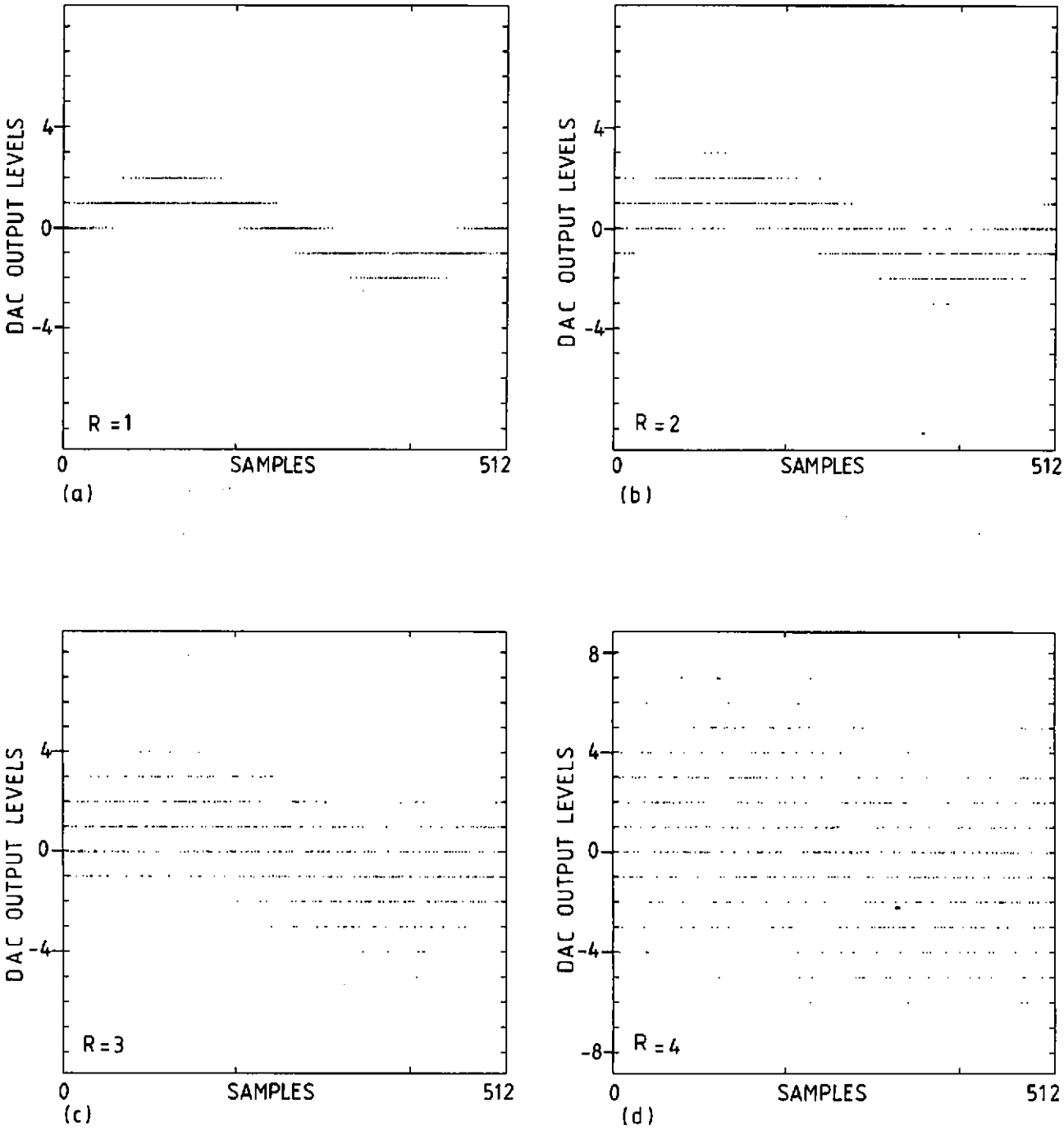


Fig.1 Noise shaper time domain outputs for $R = \{1, 2, 3, 4\}$
Input 0dB (corresponding to $\pm\sqrt{2}\,Q$)
f_s = 5·12MHz

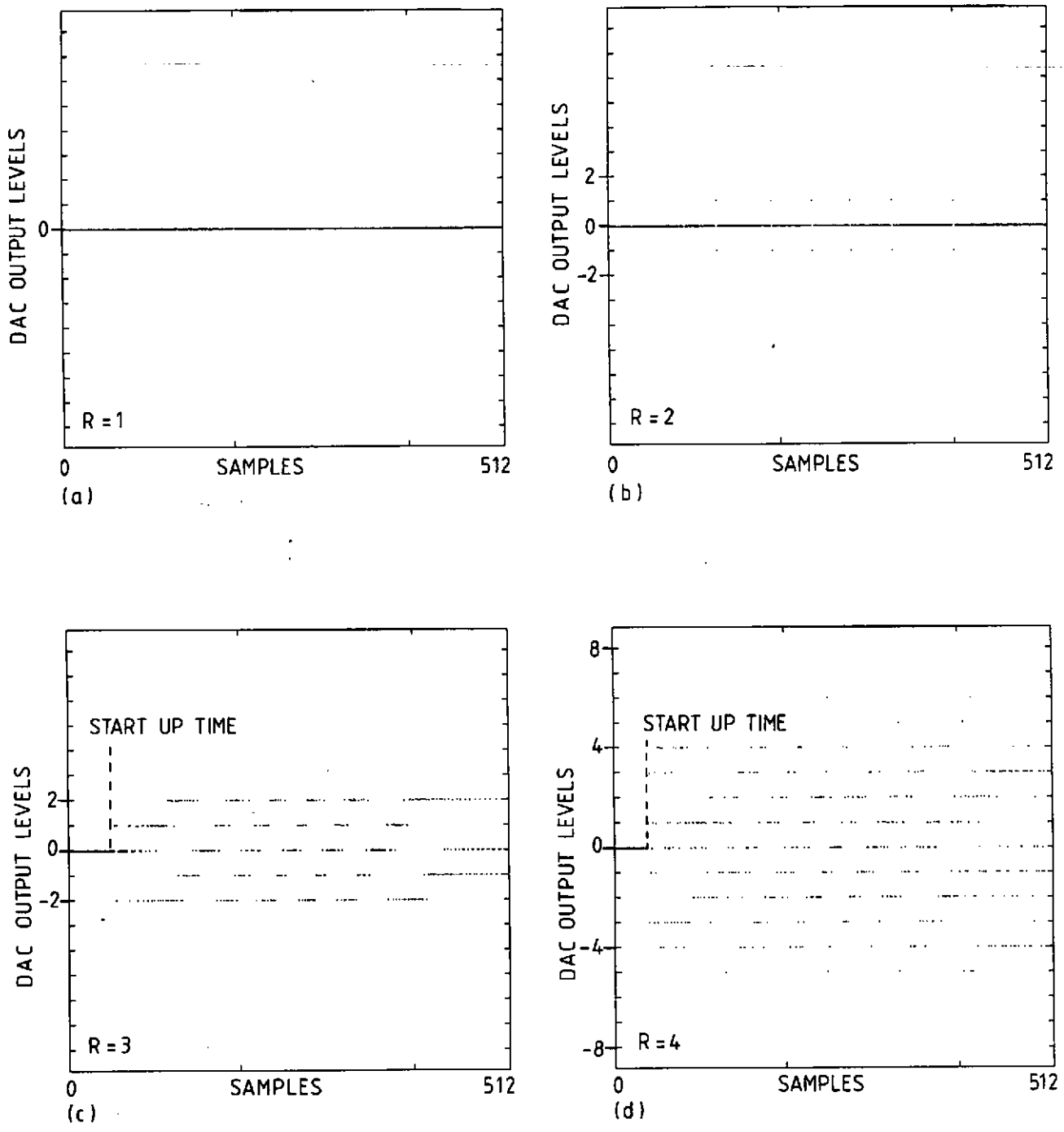
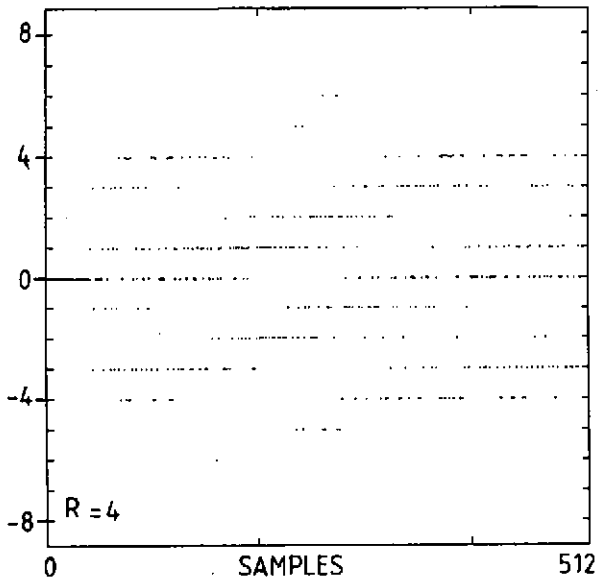
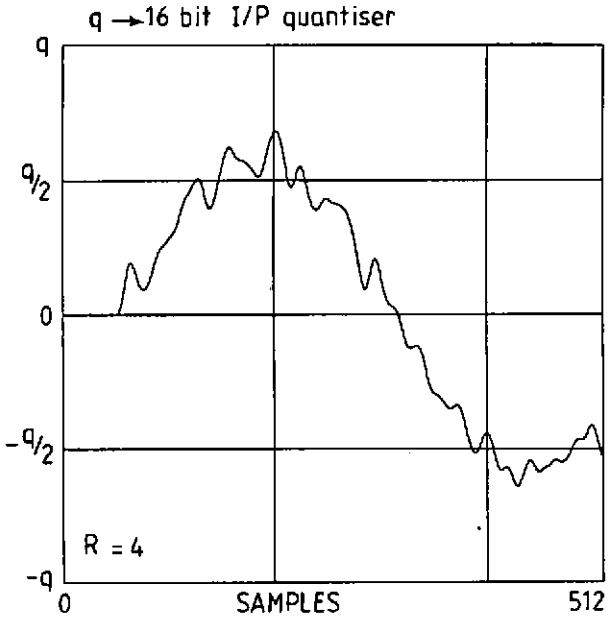


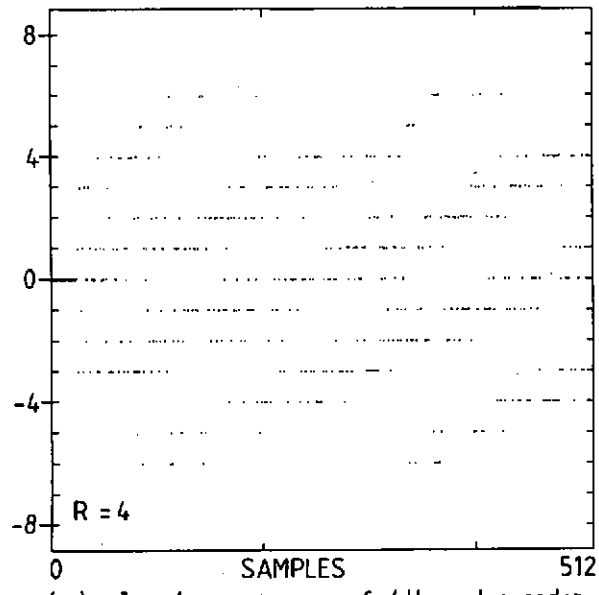
Fig. 2 Noise shaper time domain outputs for R = {1,2,3,4,}
Input - 80dB (below $\sqrt{2}$Q)
f_s = 5·12 MHz



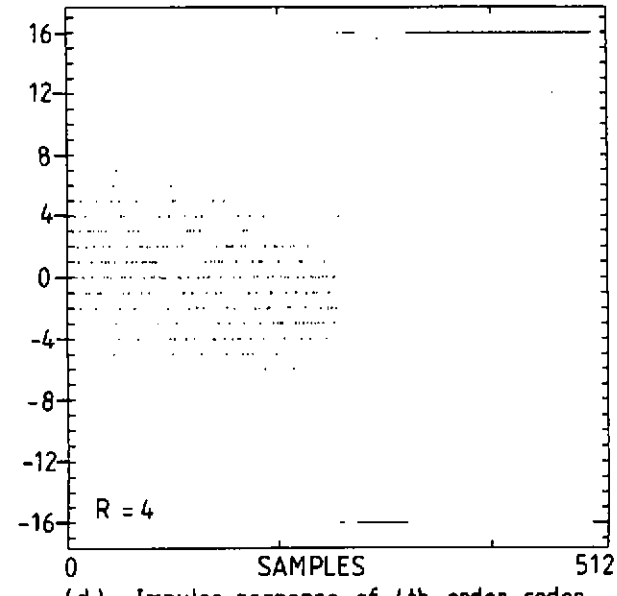
(a) Input signal -96dB
Output activity for low level input



(b) Filtered version of (a)
6th-order, coincident pole, 30kHz



(c) Impulse response of 4th order coder (100q), non-saturating



(d) Impulse response of 4th order coder, saturating (non-linear instability)

Fig. 3 4th order noise shaper time domain response examples f_s =5·12MHz

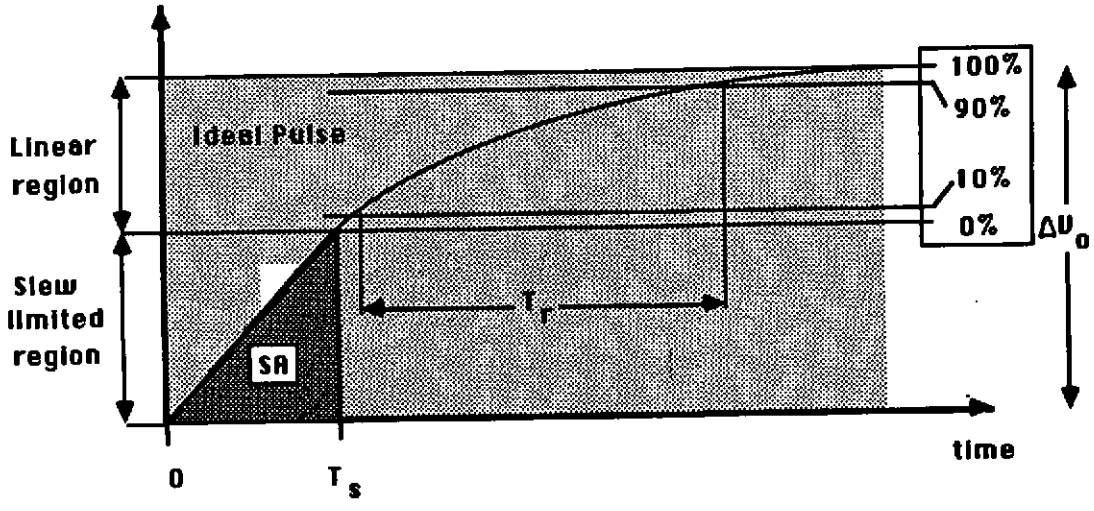


Fig.12 Slew limited and linear rise time regions of DAC transitions

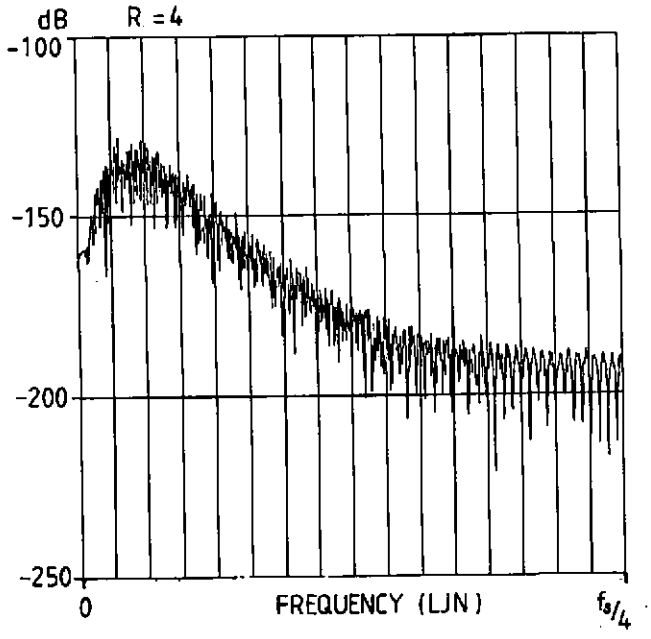


Fig.4 Spectrum of waveform of Fig.3(b).
I/P -96dB, 8kHz
Output filter 6th order coincident pole, 30kHz (Perfect DAC)

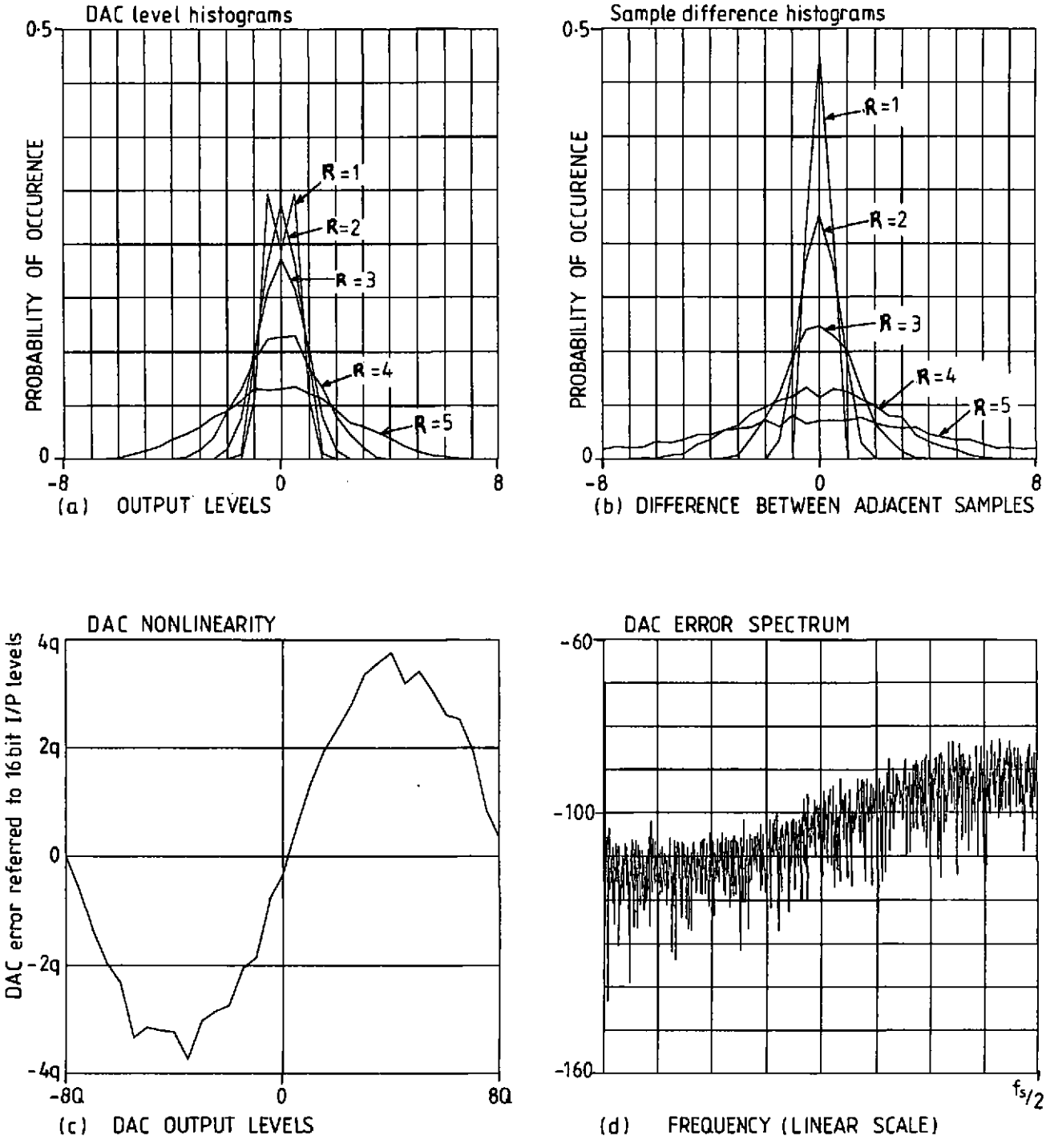


Fig. 5 a,b DAC statistical behaviour for 0dB, 8kHz input
c,d DAC nonlinearity and corresponding DAC error spectrum for 0dB, 8kHz input
f_s = 5·12MHz

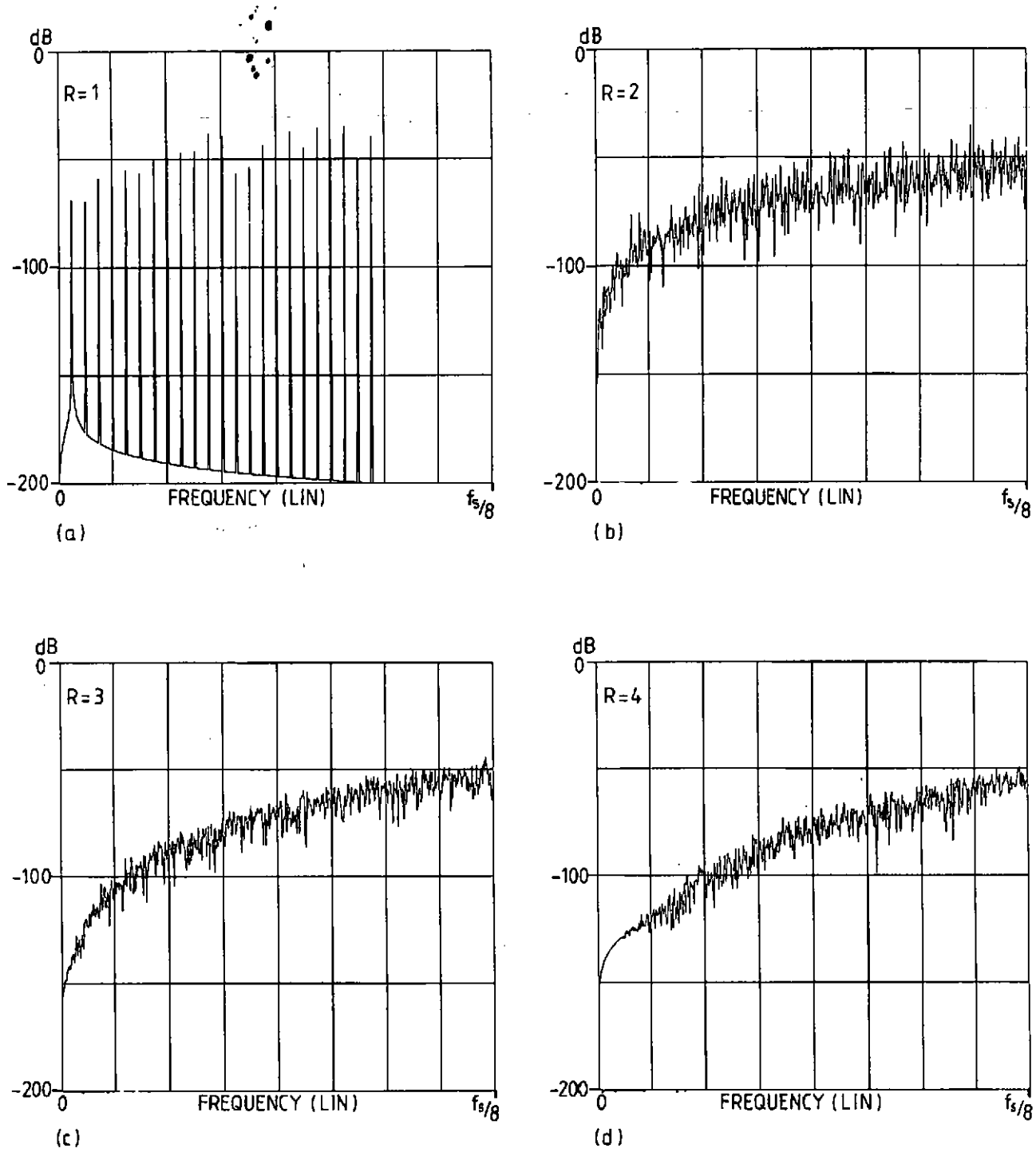


Fig.6 Output noise spectrum for ideal DAC
Input 0dB, 8kHz, R = {1, 2, 3, 4,}
f_s = 5·12MHz

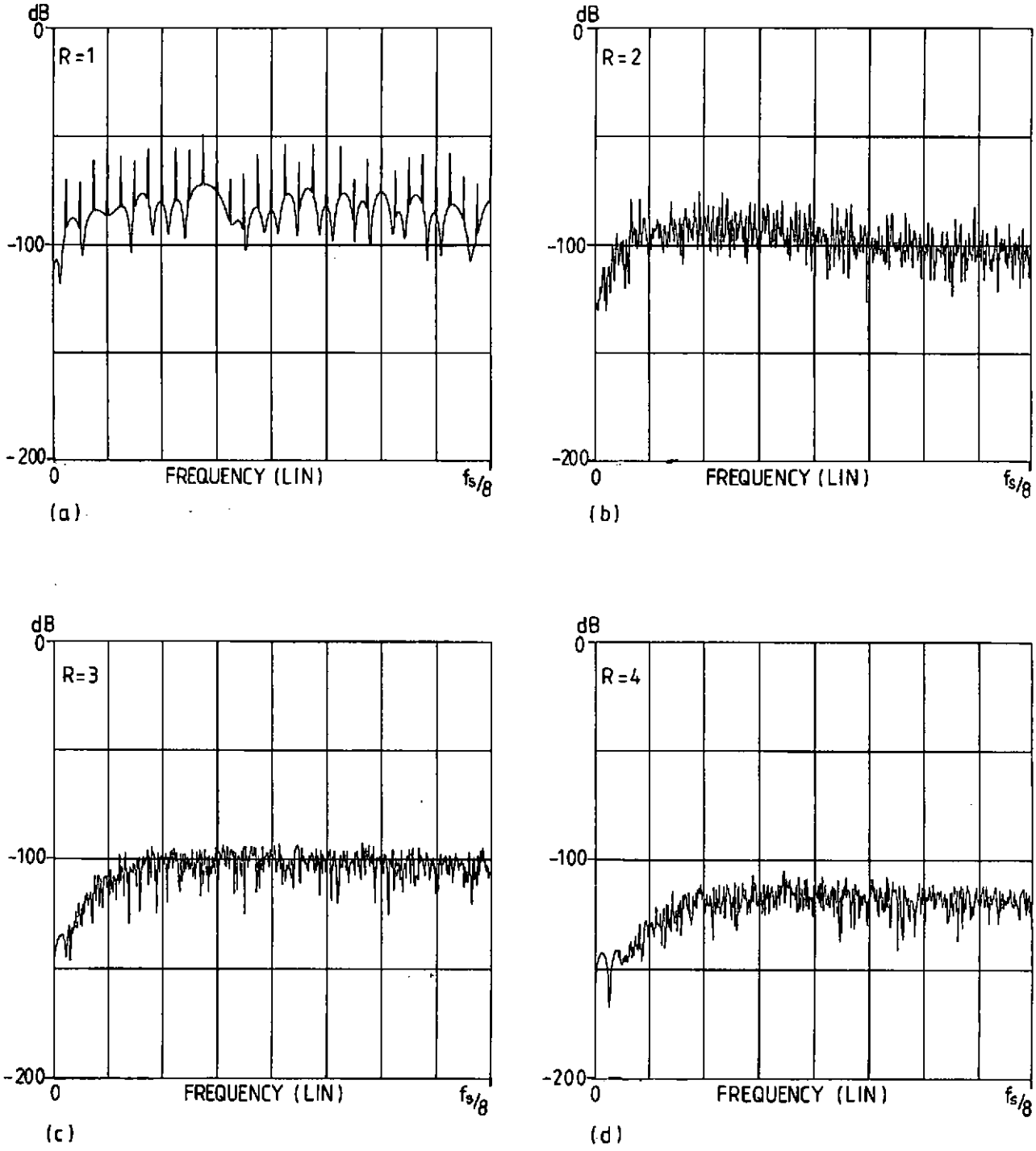


Fig. 7 As Fig.6 but with coincident pole filter of order R+1, 30kHz bandwidth
Input 0dB, 8kHz, R = {1,2,3,4}
f_s = 5·12MHz

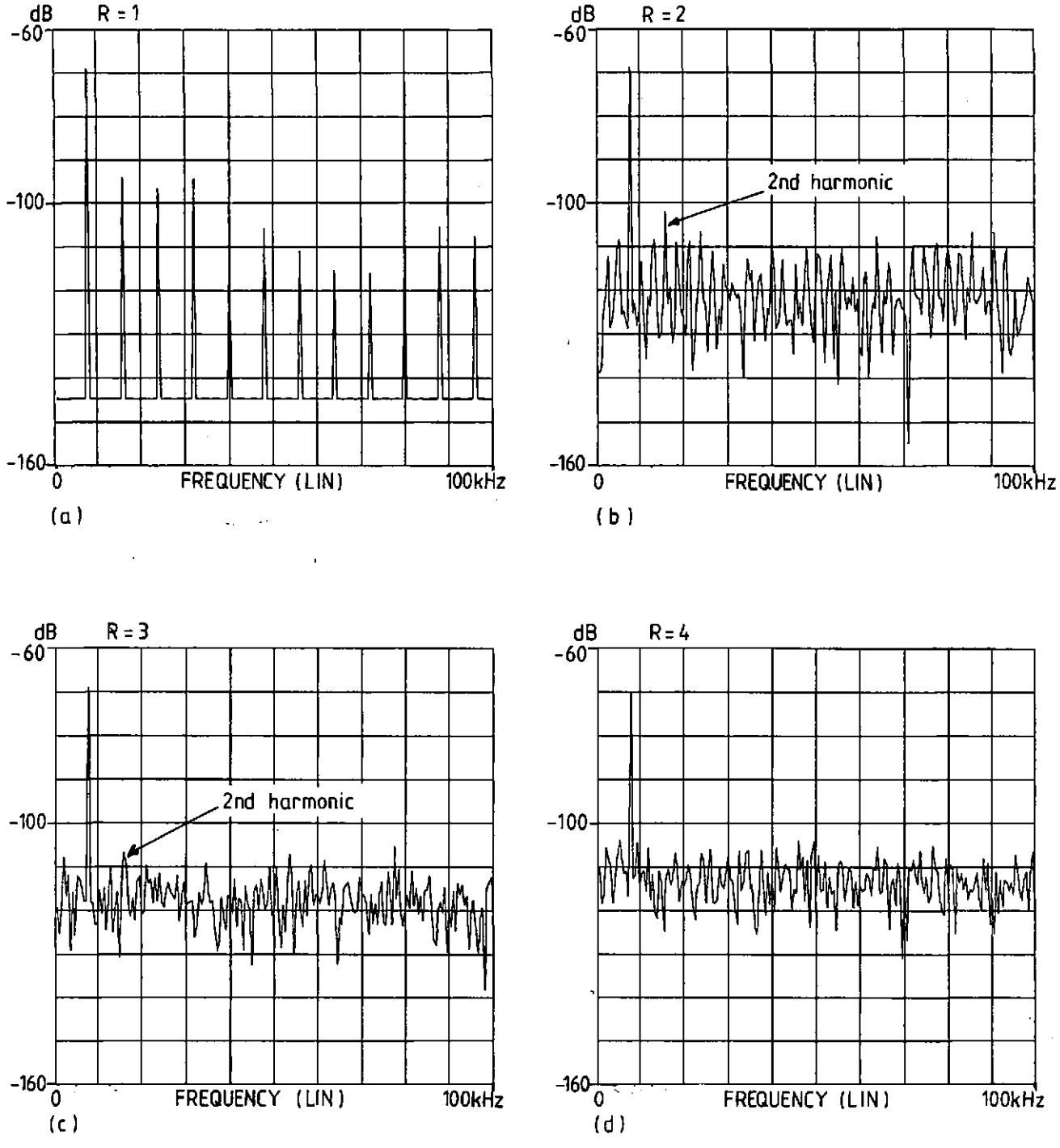


Fig.8 DAC error spectrum using DAC error from Fig. 5
Input 0dB, 8kHz, $R = \{1, 2, 3, 4\}$
f_s = 5·12MHz

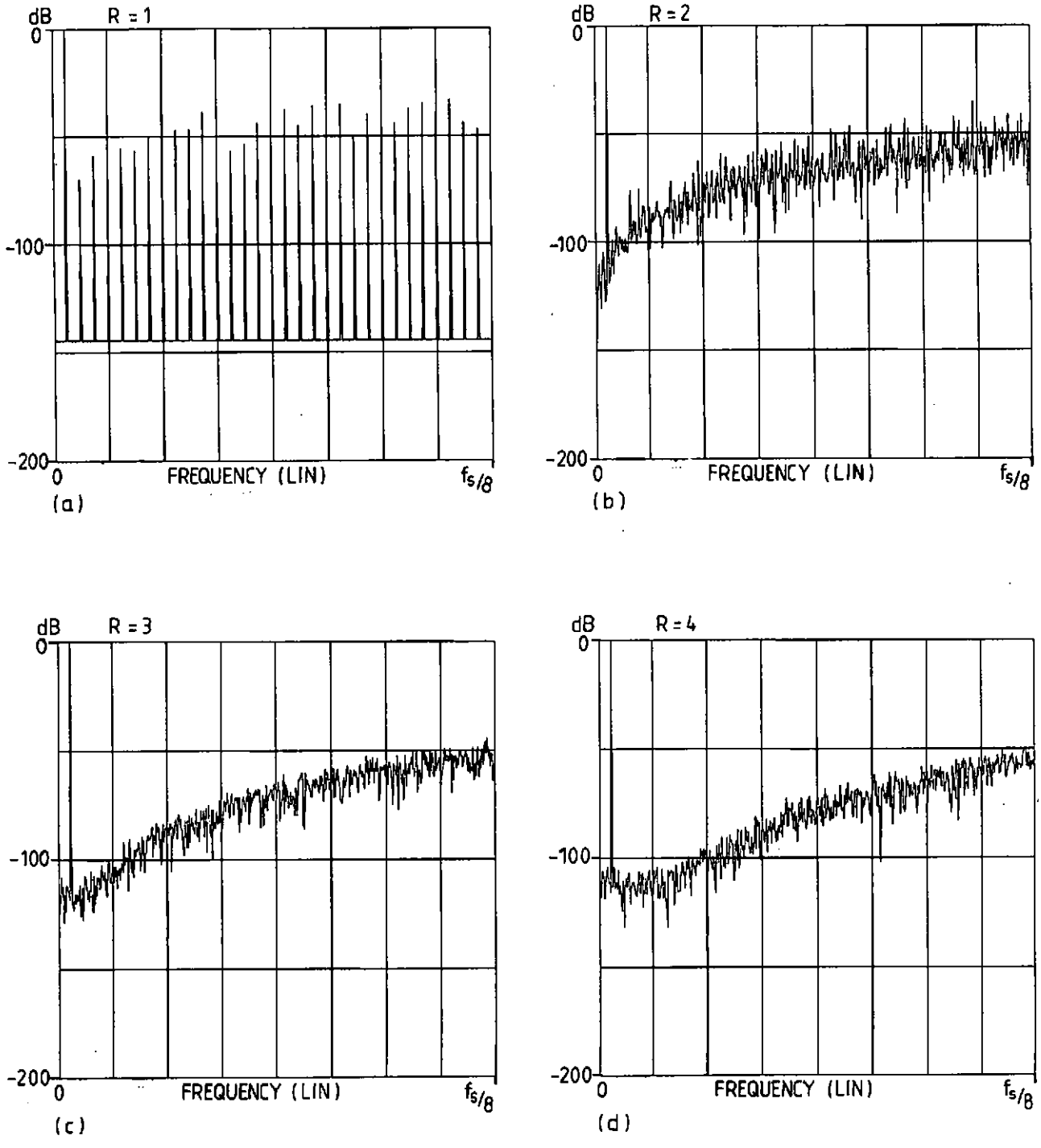


Fig.9 Same conditions as for Fig.8 but results are for overall noise at output of DAC / noise shaper
Input 0 dB, 8kHz, R = {1, 2, 3, 4} , f_s = 5·12MHz

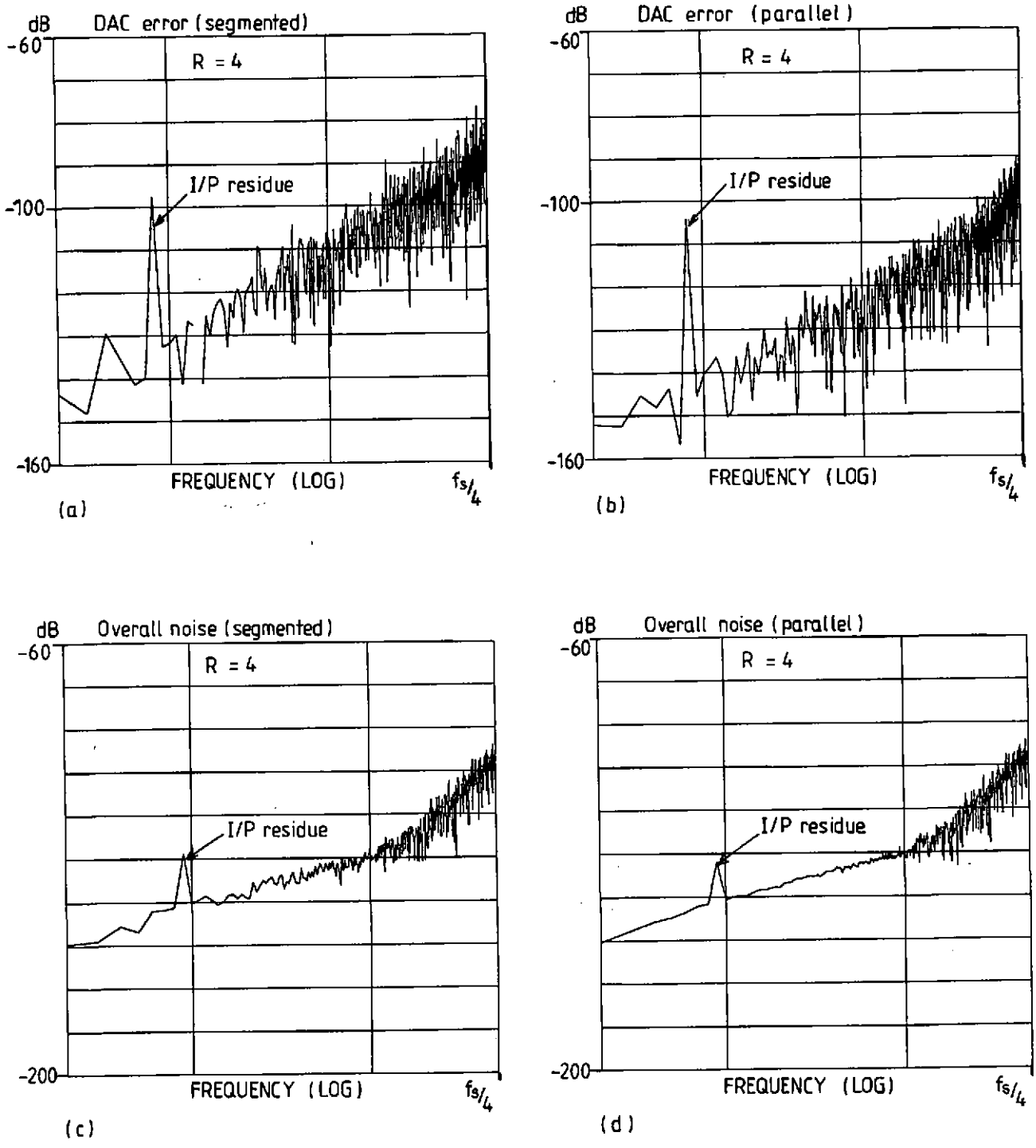


Fig.10 Error spectrums for glitch distortion for both parallel and segmented DACs
actual maximum glitch area : Parallel 1240 pV-s
Segmented 770 pV-s

Max. delay 150 ns
f_s = 5·12MHz Input 0dB, 8kHz

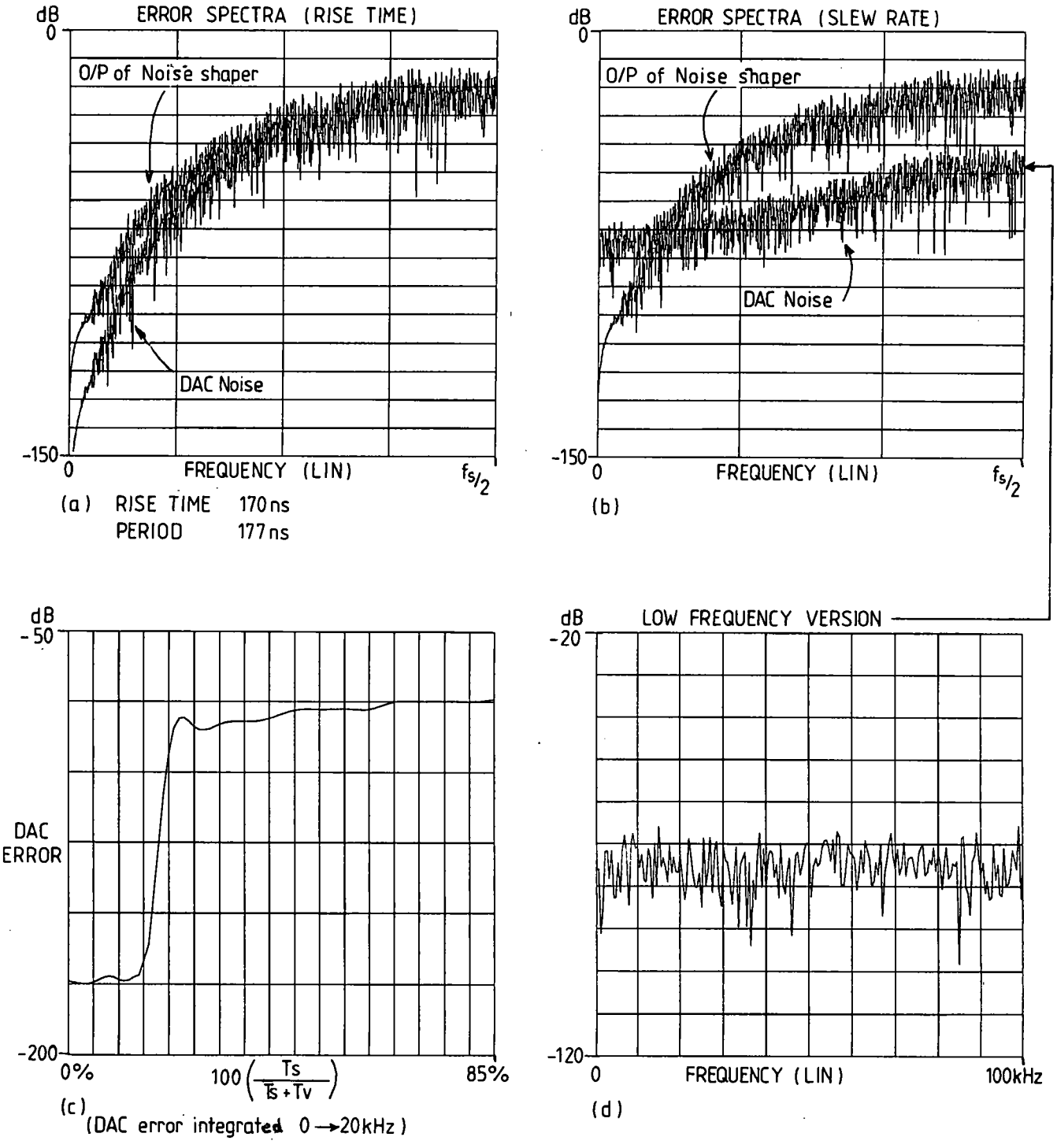


Fig. 11 Slew rate and rise time distortions
Input 0dB, 8kHz, $f_s = 5 \cdot 12$MHz