Towards a Definitive Analysis of Audio System Errors

3137 (P-1)

Chris Dunn and Malcolm Hawksford
University of Essex
Colchester, Essex CO4 3SQ, UK

5M/P-1

Presented at
the 91st Convention
1991 October 4–8
New York

AES

AN AUDIO ENGINEERING SOCIETY PREPRINT

TOWARDS A DEFINITIVE ANALYSIS OF AUDIO SYSTEM ERRORS

Chris Dunn and Dr Malcolm Hawksford
Audio Research Group
Department of Electronic Systems Engineering
University of Essex
Wivenhoe Park
Colchester CO4 3SQ
United Kingdom

ABSTRACT

Non-linear audio system errors can be analysed by separating the composite error characteristic into linear and non-linear components and then modelling linear artifacts using MLS techniques. Flexible choice of measurement nodes and the use of music signals as system input excitation results in a digital difference test which is capable of real-time non-linear error analysis and overcomes limitations in previous measurement strategies.

0 INTRODUCTION

Audio system evaluation has become polarised between 'objectivist' and 'subjectivist' schools, objectivists relying upon traditional laboratory measurements to rank component quality while subjectivists claim that the ear should be the final judge of accuracy. A large part of this rift is due to the absence of any definitive laboratory test which can be performed upon audio components and whose results accurately predict the listening experience. For example, amplifier tests commonly include an evaluation of total harmonic distortion (THD), yet there appears to be little to link an amplifier's THD performance (in isolation) with any sonic signature that the component may possess [1]. Similarly, THD tests have been used to 'prove' that cables and passive components used in typical audio systems are essentially linear in isolation [2], even though they may well interact with other components in the system causing non-linear artifacts in musical programme reproduction [3]. However, the fact that results from commonly used testing methods apparently fail to correlate well with the subjective experience should *not* be taken as a justification to eschew objective tests altogether; a scientific approach to evaluating audio system errors will always involve some form of repeatable objective test.

A common criticism of traditional tests such as THD evaluation is that the test signals applied to the device under test (DUT) are steady state and unlike music in character, hence several researchers have attempted to develop test signals that more accurately represent music, but with the advantages of typical steady-state test signals (i.e. predictability and repeatability). Belcher [4] input maximum length sequences (MLS) rich in harmonics to the DUT and processed the output through comb filters, revealing distortion components attributable to the DUT. This test is claimed to yield results which correlate more closely to listener preferences than do normal sine wave tests. Similarly Hirata [5] used compound pulse signals in an attempt to dynamically exercise amplifiers. However, although such test methods claim to approximate typical music excitation, they will always attract criticism as far as their validity is concerned because they use test signals which are artificially generated and ultimately different from the signals that are encountered in reproducing music. Thus testing audio systems using real music signals should represent the most relevant method available to engineers. Such is the thinking behind 'difference' tests such as the Quad Null Test [6] which is effectively a method of examining audio component errors when the DUT is subject to musical excitation.

The following sections will describe such difference tests and their shortcomings, followed by a new method of implementing the difference test in the digital domain, termed a **'digital difference test'**. This new and flexible testing strategy overcomes many of the problems inherent to existing difference tests and, it is hoped, will provide information which will help to bridge the objective-subjective gulf in audio component evaluation. Although the discussion will concentrate upon audio amplifier testing, the methods involved are equally applicable to other components in the audio chain.

1 DIFFERENCE TESTS

Errors generated by audio components can be grouped into two classes; linear and non-linear. Linear errors are those that are due to deviations from a perfectly flat steady-state frequency response, hence a linear error contains only signal components which are present at the input to the device. Conversely non-linear errors appear as additional signals not present in the original input signal, caused by noise and distortion mechanisms within the DUT. Although the importance of linear errors should not be overlooked [7], it is non-linear errors that difference tests attempt to isolate for examination.

Difference tests essentially involve applying a music signal to the DUT and comparing the 'normalised' input with output. The gain characteristics of the DUT are taken into account by normalising or scaling the input, hence a subtraction of the normalised input signal from the output reveals non-linear error signals introduced by the DUT, plus a linear error component due to nulling limitations in the normalisation procedure. The 'Quad Test' is probably the most well known difference test (figure 1), involving an inverting power amplifier whose normalised input is compared to output using a precision bridge; the bridge components mimic the linear frequency response of the DUT. The overall error-null can be reduced to more than 60 dB below the output when testing a 'well designed' amplifier, and hence is difficult to hear when auditioned in isolation within a normal listening environment. Proponents of this test claim that if the non-linear error signal is virtually inaudible in isolation then it will certainly be inaudible when reproduced in the presence of the primary un-distorted signal, due to masking phenomena [8]. A more stringent variation on the Quad test is due to Hafler [9], where the DUT input is compared to output with mid-band gain normalisation but no allowance is made for frequency response variations across the audio frequency range in the test device. Only amplifiers with both low non-linear distortion characteristics *and* very flat frequency response will produce a low level null signal under this test.

In principle, difference tests present an elegant way of determining non-linear error performance in audio components. However, upon closer examination there appear to be two flaws inherent in existing difference tests.

1.1 Masking

The Quad null test assumes that if the error signal can only just be heard in isolation then it will be inaudible when heard in the presence of the primary signal. This assumption is based upon masking phenomena, i.e that a small signal which is audible in isolation may not be audible in the presence of larger signals at similar frequencies. However, the validity of masking theory when examining non-linear error audibility has recently been questioned by Stuart [10,11,12], who suggests that some forms of error may be aurally more significant in the presence of the primary signal than in isolation. If this is true then the relevance of auditioning the error in isolation is questionable, and the assumptions intrinsic to the Quad test may well be invalid. Another factor, mentioned by Gerzon [13], is that most masking data has been derived using *mono* test signals, while audio system errors typically occur in *stereo*. The error characteristic due to any audio component will hence be perceived in two-dimensional space, and since the primary and error signals may occupy different co-ordinates within the stereo space then masking of errors by the primary signal may not be so effective. These misgivings suggest that a more conservative criterion may be required when evaluating errors in isolation, for example that the error can only be judged to be absolutely inconsequential if it can be shown to be below the threshold of audibility when reproduced in a typical audio system. An alternative and less stringent evaluation strategy requires the error waveforms to be mapped against the primary signal using masking theory to judge the relevance of the error. Unfortunately any such artefact appraisal requires the non-linear error signal to be isolated from linear errors to a degree that is not possible with conventional difference tests because of difficulties in reducing the linear error null to much below 70 dB.

1.2 Global System Errors

A second problem concerning difference tests involves the nature of the isolated error from a system perspective. Difference tests measure the difference between the signals at the amplifier's input and output - *they do not take into account how the DUT might interact with other components in the system*, i.e global system errors. An example of such a phenomenon is an amplifier with a non-linear input impedance but linear in all other respects. Such a device would indicate zero non-linearity in a null test, and yet in a real world situation would cause distortion when interfaced to other components. Similarly an amplifier which presents a completely linear input impedance at audio frequencies but suffers from non-linearities at RF can cause errors which will evade detection in a conventional difference test [3]. Another situation in which a component's real performance would contradict null-test findings is one which effects other devices in the replay chain via mechanisms remote from the signal path, for example common power supply variations.

These misgivings suggest that to obtain a near-inaudible null when performing a conventional difference test is not necessarily a sufficient condition to ensure a component's adequacy. If the linear error null could be decreased in order to examine very low-level non-linearities, and a way found of determining a component's characteristic at a system level, then the method would be more appropriate as an absolute test of reproduction fidelity.

2 DIFFERENCE TESTS IN THE DIGITAL DOMAIN

The discussion has so far concentrated upon deficiencies in conventional difference tests, i.e. limited linear-error null and restrictions in analysing global rather than local system errors. Both of these problems can in theory be eradicated by performing the difference test in the digital domain.

2.1 Increasing the Linear Error Null

The primary problem with conventional difference tests lies in the limited linear error null. In the Quad test this null is bounded by several factors; the complexity of the passive network used to approximate the linear transfer function of the DUT, drift in this network and drift in the DUT (due to temperature variations and other environmental factors). This error null can be improved by performing the difference measurement in the digital domain; figure 2 presents a block diagram of the proposed measurement strategy. The input and output of the DUT are each connected to a channel of a two-channel analogue to digital convertor (ADC). The input to the DUT is then presented with a music signal from a compact disc (CD) player, and the digitised versions of input and output signals from the ADC stored in computer memory. After the programme measurement has been performed an MLS is applied to the DUT, the input and output sequences again digitised using the same ADC channels and stored in the computer; cross-correlating the output and input maximum-length sequences will then reveal the linear impulse response of the DUT [14]. If the music recording of the input is now convolved with the impulse response of the DUT, the resultant convolved input sequence is equivalent to the signal that would be obtained at the output of the passive network that imitates the frequency response of the DUT in conventional difference tests, i.e we have effectively performed the input signal de-normalisation in the digital domain (because each ADC channel is used in both a music recording as well as an MLS recording, then the linear characteristics of the ADC and MLS used only effect the accuracy of this digital normalisation to a secondary degree). Subtracting the convolved input record from the DUT output record in the computer now yields a difference signal consisting purely of non-linear errors attributable to the DUT. The accuracy of the normalisation and subtraction procedures using this method are effectively limited by ADC accuracy and the length of the MLS sequence used to determine the impulse response (i.e. amount of memory available in the computer to record the input-output MLS). However, the null obtained using this method is inherently more accurate than that realised in a conventional difference test, because the frequency response of the DUT is measured rather than approximated. The digital difference test procedure may be summarised as follows:

(i) Record DUT input and output music sequences MI,MO.

(ii) Record DUT input and output maximum length sequences MLSI, MLSO.

(iii) Cross-correlate MLSO with MLSI to obtain the DUT impulse response IR.

(iv) Convolve MI with IR to obtain the normalised input MIC.

(v) Subtract MIC from MO to obtain non-linear DUT errors.

The flexibility inherent to a digital difference test should not be underestimated; because the input and output signals have been recorded in the computer, then it is possible to examine input, output, linear error and non-linear error signals as well as perform useful processing operations on the recorded data (such as cross-correlation) *after* the recording has taken place. Thus a complete picture can be built up of the test device's performance under music excitation. This is in stark contrast to conventional difference tests where only the non-linear error can be auditioned in real time as the music signal plays through the test device.

An example of such a difference test in the digital domain is indicated in figures 3 to 7. This example shows measurements made on a test circuit which exhibits a controlled non-linearity in its transfer function (figure 3); the ADC used to record the measurements possesses a dynamic range of approximately 14 bits. For the sake of clarity this example uses sinusoidal signals output from a CD player as system excitation. Figure 4 indicates the test circuit input signal, essentially a pure tone at 1 kHz, -10 dB. Figure 5 shows the output spectrum; the fundamental has increased in level by about 6 dB (the mid band gain of the test circuit), but signal harmonics have also been introduced, mainly 2nd. harmonic at -70 dB. If the input is now subtracted from the output without any processing operation carried out upon the input, then the difference signal will effectively reflect the linear gain of the test circuit plus distortion (i.e. linear *and* non-linear errors). Figure 6 shows such an un-normalised error spectrum; the fundamental is only reduced to -10 dB.

Figure 7 represents the difference between the *normalised* input and output, and a comparison with figure 6 indicates the efficiency of the normalisation procedure. Normalisation involves applying a 16383 point MLS to the test circuit at a low level (where the circuit is effectively linear) to obtain the impulse response of the circuit. Once the input signal has been convolved with the impulse response then the linear transfer function of the circuit has been accounted for; the non-linear error signal indicated in figure 7 includes all distortion harmonics introduced by the DUT (c.f. figure 5) plus a low-level 1 kHz tone. This error signal at the fundamental frequency is not due to inaccurately determining the impulse response of the DUT, but occurs because of a change in non-linear behaviour of the test circuit between the 1 kHz test signal level and the MLS level from which the impulse response is derived. This 1 kHz error tone can thus be thought of as a gain change in the test circuit from the gain at the MLS reference level. Probably the best way to understand this phenomenon is to imagine the input signal traversing an input-output curve that represents the transfer function of the test device (figure 8); the gain (curve gradient) experienced by the low-level MLS signal will be higher than that at the higher level test signal. Note that this type of distortion will not be revealed in traditional THD tests since conventional distortion analysers null out fundamental frequency information irrespective of whether it is present in the input signal or due to a non-linearity.

To help qualify this analysis figures 9 through 11 illustrate another test where the 'DUT' is simply a short circuit between input and output measurement nodes, i.e. the test device should be perfectly linear. Figures 9 and 10 show the input and output signals respectively, while figure 11 illustrates the error signal after normalising the input with the impulse response of the 'system'. Because the measured system now has the same transfer function characteristics at both the test signal and the MLS signal levels then the gain change is effectively zero, and the 1 kHz error is much smaller at about 95 dB below the primary signal. This level of fundamental null along with a benign noise floor approximately 84 dB below full scale represents the measurement accuracy of the experimental system as it presently stands (note that the noise is spread over 4096 frequency bins and consequently is plotted at a lower level in the diagrams). Although these examples were performed using a fairly crude experimental setup they do indicate that a difference test in the digital domain is feasible. Since the ADC is utilised both in recording input and output signals from the DUT as well as determining its impulse response, then ADC accuracy represents the limiting factor in the linear error null achievable in a digital difference test. Using state of the art ADC technology [15] should enable non-linear error analysis well below the threshold of hearing and this is currently being incorporated into the experimental system.

2.2 Analysing Global System Errors

Since music signals can be applied to the DUT and input and output sequences recorded in a computer, then the possibility exists of comparing these signals to some reference sequence recorded when the DUT is not present in the system. To allow a valid comparison there must exist some method of accurately synchronising data capture to the music programme. This is possible if a CD player is used as programme source in the tests, since the digital SPDIF interface available on many integrated and all transport-only CD players contains subcode information including timing data synchronised to audio data on the CD.

Figure 12 indicates a block diagram of a measurement scheme which allows timing synchronisation of data capture. The CD SPDIF digital interface is connected to a timing decoder which triggers data capture at a certain time in the programme. The analogue output of the CD player is connected to the DUT, whose input is digitised by the ADC and recorded in the computer. The DUT is then removed from the system, and the measurement repeated using *exactly the same section of music.* Any linear errors that may be introduced by the DUT in loading the source can be accounted for by modelling the DUT 'input impedance' impulse response using MLS techniques as described above, and then subtracting the two digitised sequences from one another after normalisation. Any residual in this difference signal can be attributed to global errors caused by the DUT which are undetectable using conventional difference tests.

3 CONCLUSIONS

This paper has briefly examined problems with conventional difference tests, namely a practical limit to the level of linear error null that can be achieved and an inability to examine global errors introduced by the DUT. A flexible difference test in the digital domain is proposed as a method of overcoming both of these problems. A crude experimental system has been constructed whose initial results indicate promising performance, while improvements to the measurement system should permit error performance evaluation to well below the threshold of hearing, allowing a definitive analysis of audio component behaviour under musical excitation.

ACKNOWLEDGEMENT

This work is part of ongoing research supported by the Science and Engineering Research Council, U.K.

BIOGRAPHY

Chris Dunn is a second year postgraduate student at the University of Essex, England. Under the supervision of Dr Malcolm Hawksford, he is working within the Audio Research Group on audio amplifier errors and high resolution measurement strategies.

REFERENCES

1. Martin Colloms, "The Sound of Amplifiers," Hi-Fi News and Record Review (HFN&RR), May 1985.

2. D.R.G.Self, "Ultra-Low-Noise Amplifiers and Granularity Distortion," J.Audio Eng.Soc., Vol.35, No.11, November 1987.

3. Paul Miller, "Resonances and Repercussions," HFN&RR, June 1989.

4. R.A.Belcher, "A new Distortion Measurement," Wireless World, May 1978.

5. Yoshimutsu Hirata, "Nonlinear Distortion Measurement Using Composite Pulse Waveform," J.Audio Eng.Soc., April 1981.

6. Peter J.Baxandall, "Audible Amplifier Distortion is Not a Mystery," Wireless World, November 1977.

7. Malcolm Hawksford, "The Essex Echo: on Errors, Low Feedback, and Fuzzy Distortion," HFN&RR, September 1984.

8. Louis D.Fielder, "Evaluation of the Audible Distortion and Noise produced by Digital Audio Converters," J.Audio Eng.Soc, Vol.35, No.7/8, July/August 1987.

9. David Hafler, "The Designer Series," HFN&RR, November 1986.

10. J.R.Stuart, "The Stuart Hypothesis," (letter), HFN&RR, January 1991.

11. Brian C.J. Moore, "A Null Hypothesis," (letter), HFN&RR, May 1991.

12. J.R.Stuart, Letter, HFN&RR, June 1991.

13. Michael A. Gerzon, "Masking of coding/decoding errors in audio data compression systems", Proceedings of the Institute of Acoustics, November 1990.

14. Douglas D.Rife and John Vanderkooy, "Transfer-Function Measurement with Maximum-Length Sequences," J.Audio Eng.Soc., Vol.37, No.6, June 1989.

15. UltraAnalog Inc., ADC 20048 20 Bit Audio ADC Data Sheet.

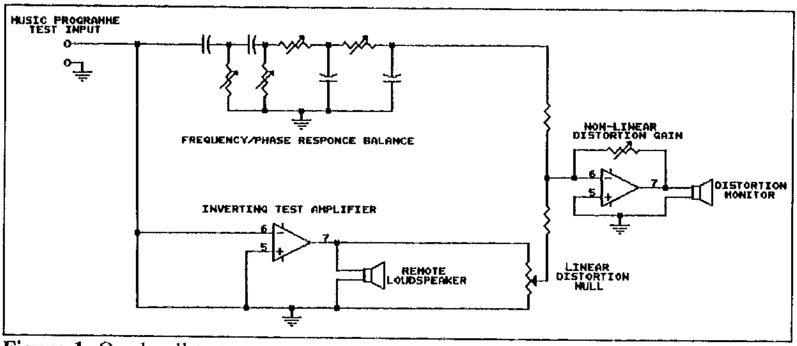

Figure 1: Quad null test.

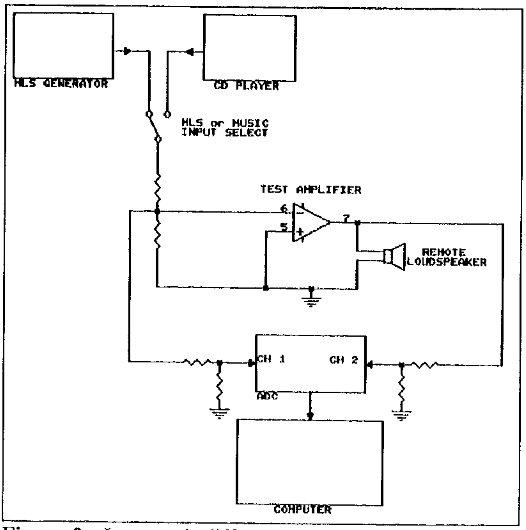

Figure 2: Improved difference test, where nulling is performed in the digital domain.

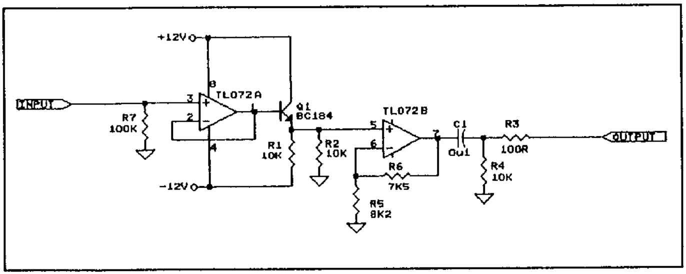

Figure 3: Test circuit with controlled non-linearity.

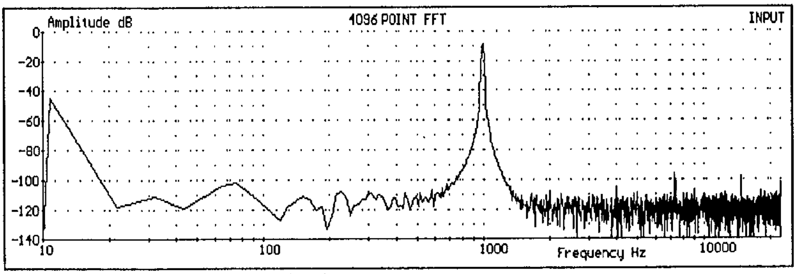

Figure 4: 1 kHz input to test circuit.

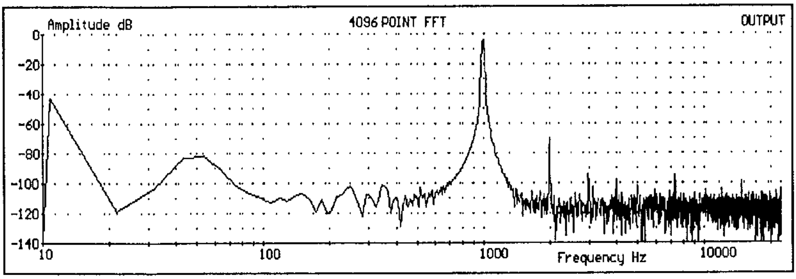

Figure 5: Test circuit output.

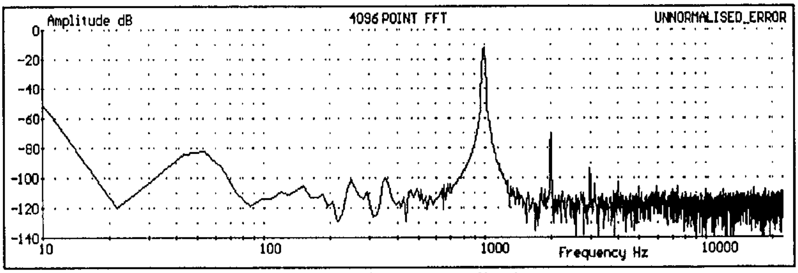

Figure 6: Test circuit error when input is un-normalised.

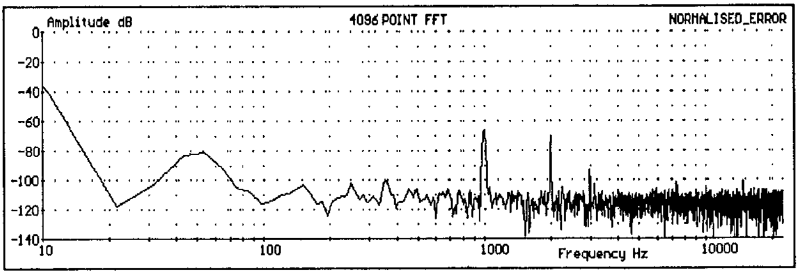

Figure 7: Test circuit error with normalised input.

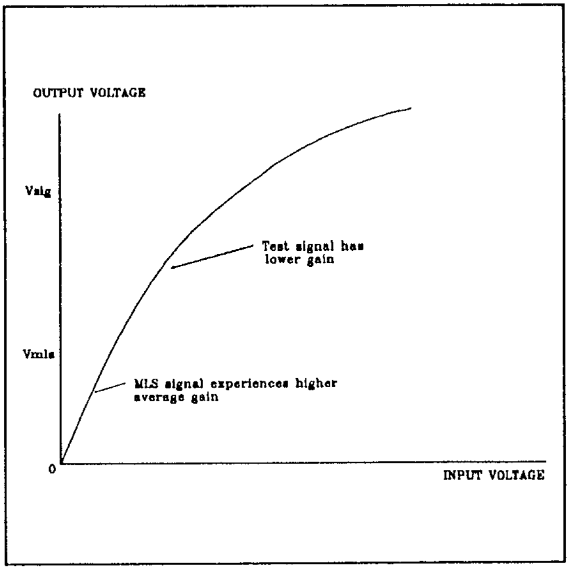

Figure 8: Input-output transfer function indicating gain change due to non-linearity, where Vmls is MLS peak input level and Vsig is peak input test level.

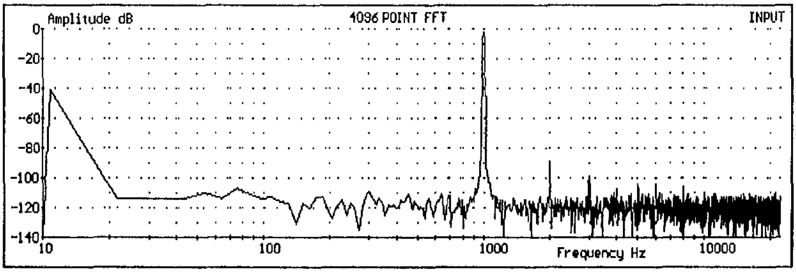

Figure 9: Input signal to short circuit.

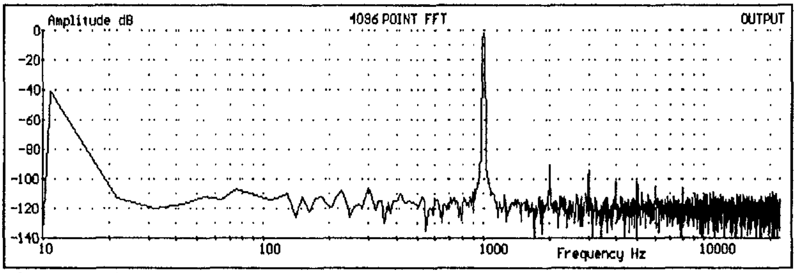

Figure 10: Output from short circuit.

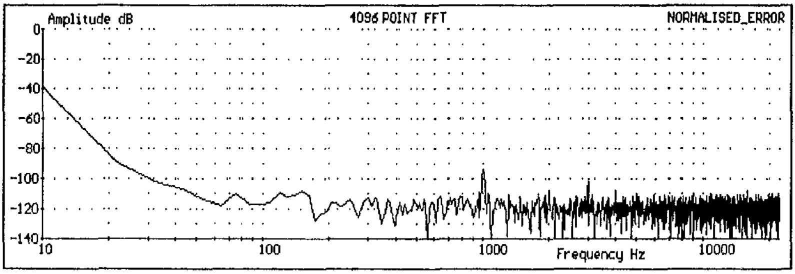

Figure 11: Short circuit error spectrum.

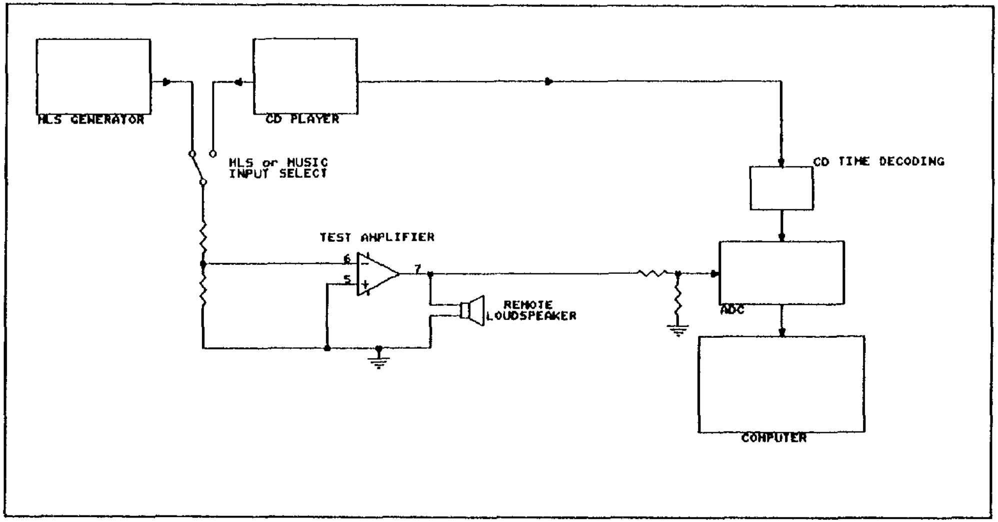

Figure 12: Determining global errors introduced by the DUT.