ERROR CORRECTION FOR PCM DIGITAL AUDIO SYSTEMS

Authors: David Shenton, Erik DeBenedictis, Bart Locanthi
Pioneer North America
Pasadena, California
U.S.A.

INTRODUCTION—A review of digital audio

The objective of digital audio is to produce music with sufficient fidelity that it cannot be distinguished from the original. It is hoped that digital audio systems therefore will exceed the limits of human auditory perception. The recording and playback equipment used for these processes should not introduce noticeable degradation. Furthermore, an understanding of the human auditory process is necessary in order to design these components. Some of the more relevant characteristics of the auditory system are sensitivity, frequency response, and distortion. Digital errors produced by playback equipment can introduce clicks, or other degradations.

This paper will analyze the effects of the human auditory system on the design of digital audio equipment.

Frequency Response

Many studies have been done on the frequency response of the human ear. A large majority of these studies indicate that the ear's response becomes negligible above 20KHz. However, a recent study has been performed where subjective changes in real music processed by digital means were tested. This study involved application of a very high-order, zero-phase, digital low-pass filter to music[1].

There are three major proposals for digital audio sampling rates. The European Broadcast Union has an established standard of 32K samples/second. This standard is in use for in-studio and inter-studio communication. A second standard is 44.1K samples/second. This standard is in use on some existing professional digital audio equipment and is proposed for consumer use. A third standard of 50K samples/second is in use in the United States and in Europe for professional mastering. With suitable digital-to-analog interfaces these are capable of maximum cutoff frequencies of 16, 22, and 25KHz respectively.

Sensitivity

Many complicated models of the ear's sensitivity have been proposed. Such models indicate that the differential sensitivity is reduced when the sound pressure level is high, resulting in loud volume masking the lower volume noise. Refinements of this model also indicate that noise within certain frequency bands can be masked by signals at higher levels in other frequency bands. However, the overall volume level may not have to be very high in these cases. Neglecting these special instances, a simple and conservative model defines a critical signal-to-noise ratio which is the difference between the threshold of pain and human audibility. Music produced with a signal-to-noise ratio of 118db will always exceed the ear's sensitivity to noise.

The European Broadcast Union's standard employs a 14 bit linear quantization. This provides an 86db signal-to-noise ratio. Almost all recent proposals use 16 bit linear quantization, providing a 98db signal-to-noise ratio[2]. These figures are exclusive of some coherent perturbations introduced by the sampling system. These additional perturbations are not being discussed in the industry, probably due to their marginal audibility.

Distortion

Simple implementations of the digital/analog interface may produce noticeable aliasing distortion, phase distortion, and noise. These implementations often utilize only an A-D converter and a low pass filter. A steep filter falloff with a very high attenuation in the stop band is necessary to prevent frequency components above half the sampling frequency from appearing in the audio spectrum below this frequency. On the other hand, the filter must be of linear phase to prevent transient distortions from reaching the listener. A ninth order active filter has been proposed for production[3].

A more sophisticated implementation has also been proposed[1]. In this proposal the sampling rate is twice that normally used (100KHz). The input signal is acted upon by a linear phase Bessel lowpass filter with a cutoff frequency equal to the original nyquist rate, 50KHz. The result is a digital signal with no phase distortion and no frequency components above 50KHz. This digital signal is then operated upon by a linear phase lowpass digital filter with a cutoff frequency of 25KHz. The sampling rate is then reduced to 50KHz, resulting in a digital signal having no transient defects.

ERROR CORRECTION IN DIGITAL AUDIO

The performance of recording media and the detectability of the ear to errors clearly define the arena for the error-correcting system. We will discuss these two issues in detail before presenting proposals.

A single incorrect bit in digital music is often noticeable. An incorrect bit will produce a momentary excursion of the sound, the magnitude depending on the position of the bit. Unless the bad bit is one of the least significant bits, this is heard as a click. If an hour of digital music is to be played with no detectable errors, the bit error rate must be less than 10^{-10}. The authors know of no such presently-available recording medium, regardless of cost, that can accomplish this. Plans for digital audio are to put low-cost equipment onto the consumer market with this level of performance.

Unlike other forms of data, errors in digital audio can often be successfully concealed. Therefore, when a correction technique leaves behind a known error, it may be made considerably less bothersome by error concealment. In error concealment, known incorrect samples are approximately reconstructed from neighboring, good samples. A common and very effective technique is to replace a sample with the average of its neighbors (See figure 1). A limited number of listening tests have been conducted on error concealment by the authors and others. The results generally indicate that error concealments on actual music, even at a high rate (1000 per min.), are not noticeable. However, some sounds, such as sine waves, have been found in which concealments can be readily noticed. Multiple recording and playback operations using error concealment will cause a continuous increase in the number of concealments. This is disturbing because it is the only aspect of digital audio which results in the continuous loss of fidelity with use.

Optical disc technology

Optical discs consist of a thin layer of aluminum evaporated onto the underside of a transparent plastic material. Information is encoded as pits pressed into the underside of the plastic. The depth of the pits is chosen to be ¼ wavelength of the recording laser light. The pits cause the optical distance through the surface to change, and are therefore detectable by a simple interferometer.

Since the signal is on the underside of the plastic, the system is somewhat resistant to dust and scratches on the top side. This is understood with the aid of figure 2. The thickness of the plastic is about 1.2 mm, and the lens rides about one millimeter above the surface. The numerical aperture of the lens is about 0.5, causing the optical area of the reading laser beam at the surface to be 1.7 mm in diameter. Dust particles and scratches on the surface less than 1.7 mm will not totally obscure the signal.

Manufacturing imperfections can be a much more serious cause of errors. Dust on the die, air bubbles, and particles suspended in the plastic have all been observed to interfere with information retrieval.

These phenomena have been investigated optically and electronically by Pioneer. Optical investigation reveals the causes of some of the errors and provides some intuitive understanding of the processing difficulties. Electronic investigation produces more accurate statistics.

Optical Investigations

Several optical discs were examined by Pioneer North America, through the plastic, using a camera-equipped microscope with a long working distance objective.

Figure 3 shows a section from a commercially available video disc. This is an extended-play, or constant linear velocity disc, and has a high information density. Microscopic examination revealed very few irregularities, one of which is shown.

Figure 4 shows a video disc manufactured by Pioneer Electronic Corporation. This is a constant angular velocity disc, which allows stop and slow motion of the television signal. The irregularity shown causes the signal to disappear over an area many times the size of a bit. A burst error occurs each time the irregularity passes under the stylus. Since the irregularity covers a dozen or so tracks, there would be that number of error bursts spaced by the revolution period of the disc.

Figure 5 shows a constant angular velocity video disc. This disc was purchased in a retail store in the vicinity of the laboratory for $3.98 plus tax. Considering the price of the disc and the information potential, this disc represents a phenomonally low price of 10^{-7} cents per bit! This disc has many errors, however. Microscopic examination always revealed dozens of small errors in each field of view. Rough visual examination indicated that the bit error rate is around 5×10^{-3}. The irregularities on this disc are characteristically smaller than those seen on other discs. This would correspond to a smaller average burst length, or a more gaussian distribution of errors.

Figure 6 shows an experimental digital audio disc manufactured by Pioneer Electronic Corporation. Simple inspection will reveal that a different modulation scheme is used on this disc. Video discs use frequency modulation of the pit lengths to generate a composite video signal. In this scheme, the size of the pits and the spacing between pits are used to determine the actual data.

While investigating these discs, a large particle about 0.5mm in length was observed in the high density region of the disc. A portion of this particle may be seen in figure 7. The particle shown is so large, that the tracking mechanism fails as the particle passes under it.

Optical observations demonstrate that dirt and dust are major factors in the quality of optical discs. The number of these imperfections is largely dependent on the cleanliness of the manufacturing facilities. High quality optical discs can only be manufactured in a cleanroom environment. The entertainment industry has not previously required the use of cleanrooms, which are very expensive. The necessity of a large capital investment in the manufacturing facilities will increase the price of discs and may limit the number of entries into this manufacturing market. This could, in turn, reduce the amount of program material available to the consumer and may impact the public acceptance of digital audio discs.

Electronic Investigations

Pioneer Electronic Corporation has performed extensive investigations of its own disc manufacturing capabilities. These investigations involved making constant angular velocity discs with known recorded signal patterns. These discs can then be played and the reproduced patterns compared with the original ones. Statistics gathered by these laboratory investigations indicate that the bit error rate is highest at the center of the disc, where the data packing density is the highest. These bit error rates have been measured to be as high as 10^{-4}. As the packing density is reduced, the error rate is reduced, resulting in better performance. By visual comparison of the constant angular velocity discs (figure 4) and the constant linear velocity discs (figure 3), it can be readily seen that the spacing between data bits in the constant linear velocity discs has been reduced to the minimum value used in the constant angular velocity discs. The industry should then expect their bit error rates throughout the entire disc to be the same as those of the constant angular velocity discs at their centers.

This manner of investigation is limited to Pioneer discs because recording service is not generally offered by Pioneer's competitors. The facility that made the discs that were analyzed was an experimental one, and the discs made cannot be considered representative of discs available to the consumer.

Error Correction for Optical Disc Systems

The proper choice of an error correction system is a tradeoff between cost of the electronics, correctability, and the expected defects on the disc. Added expense for the correction system will allow the use of less expensive discs with their associated higher error rates, the use of scratched discs, and will reduce the number of concealments used in playback. Consumer market economics further dictate that the error correction be implemented with only a few integrated circuit chips. This constraint rules out may of the more successful advances in coding theory, and is forcing engineers to seek simplifications and combinations of conventional codes.

Manufacturing defects, as we have seen, will cause burst-type errors with characteristic lengths broadly between 1 and 100 bits. Surface scratches will not affect the signal unless they are of the order of 1 mm in size. Such scratches will produce bursts of the order of 1000 bits. A combination of these effects will certainly be present.

CIRC Proposal

Sony and Phillips have jointly proposed a scheme based upon double application of Reed-Solomon codes[5]. Each application of the Reed-Solomon code utilizes only a fraction of the available correcting ability, to reduce the amount of hardware required. The two simple systems complement each other very well, and the resultant system is quite powerful.

The reader is referred to figure 8 in the following explanation. The code consists of two Reed-Solomon codes, one (32, 28) over $GF(2^8)$ and one (28, 24) over the same field. The multiple application is such that the 28 data symbols of the first code are the 28 total symbols (24 data symbols and 4 check symbols) of the second code. The codes are fully interleaved to make the nested code words independent and to increase the burst error correcting capability. Both Reed-Solomon codes have the capability of correcting two incorrect bytes, correcting four marked incorrect bytes, or correcting one incorrect byte and detecting two or more incorrect bytes. Each of these requires different amounts of hardware and time.

The first decoder (32, 28) performs single error correction and two or more error detection. This decoder will generate an erasure flag with each byte that indicates whether or not the byte is known to be reliable. If two or more errors are detected in the data block, then all the data bytes are flagged as bad. The second decoder (28, 24) performs four byte correction on the data bytes of the first decoder. If more than four erasure flags are set, then the decoder will not perform any correction and will mark all the data bytes in that code word as bad, and concealment will be performed on all data bytes that have received flags from both decoders.

The hardware to implement this scheme consists of two chips. One chip is a custom microcoded processor element, and the other is a 1K x 8 ram. The random error performance of this system is 1000 concealments/minute and 1 click/750 hours, at an input bit error rate of 10^{-3}. The system can correct bursts up to 4000 bits with full correction, and up to 14,000 with concealment.

Convolutional Codes

Among other error correction schemes, the authors have investigated convolutional codes to perform error correction. The authors feel that error concealment should be used only as a last resort, and convolutional codes have sufficient correctability that concealment will seldom be necessary. The authors furthermore feel that the requirement for low bit-error-rate discs may not be reasonable. By using a suitable interleaver, bursts up to 1000 bits may be handled. The penalty for using a rate 1/2 code is a reduction of recording time of 33% from a rate 3/4 code.

For price and speed reasons the decoder should be of the threshold type. The code used should be self-orthogonal, and of the class proposed by Robinson and Bernstein[6]. The code could be simply interleaved to gain burst correctibility. The proposed class contains codes with different constraint lengths and correspondingly different correctabilities. A tradeoff can be made between the amount of hardware required to implement the shift register and the desired correctibility.

Several combinations of code rate and constraint length have been evaluated. The most important factor in choosing constraint length is the state of MOS memory fabrication technology. Shift registers of less than 10,000 bits can be reasonably implemented on one chip. Given these factors, a rate 1/2 code with 16 shift register taps is proposed. The total amount of memory, including some interleaving, is 20,000 bits. The simulated performance of this system is 1 miscorrection per hour at an input bit error rate or 3×10^{-3} with a burst correctability of 1000 bits.

Conclusions

The potential that digital audio offers is the ability to record and play back audio signals repeatedly without degradation. The human ear was used to define the parameters which the digital audio industry has accepted as their standards. It has been shown in this paper that it is possible to produce digital music that meets these requirements.

The industry must now produce digital discs and, in the near future, tapes that store this vast quantity of information. However, since recording, playback and especially manufacturing are not perfect processes, an error correction scheme needs to be employed to insure the high quality of audio that this medium can offer. Selection of an appropriate error correction system for digital audio will affect the potential lifetimes of the discs, as well as their overall quality. Furthermore, the expense of manufacturing might also limit the selection of available discs. These factors will all affect the acceptance of digital discs in the consumer market.

REFERENCES

1. Korte, Blüthgen, Swientek, "Investigation to obtain perception limits of filter related sonic degradations in digital audio source encoding." AES 67th convention, No. 1717 (H-2).

2. Blesser, "Digitization of Audio, a Comprehensive Examination of Theory, Implementation, and Current Practice," Journal of the AES, Vol. 26, October 1978.

3. Lagadec, Weiss, Greutmann, "High Quality Analog Filters for Digital Audio," AES 67th convention No. 1707 (B-4).

4. Pioneer technical report, internal document.

5. Vries, Immink, Nijboer, Hoeve, Doi, Odaka, Ogawa, "The Compact Disc Digital Audio System: Modulation and Error-Correction," AES 67th convention No. 1674 (H-8).

6. J. Robinson and A. Bernstein, "A Class of Binary Recurrent Codes with Limited Error Propagation," I.E.E.E. Transactions on Information Theory, Vol. 13, January 1967.

FIGURE 1

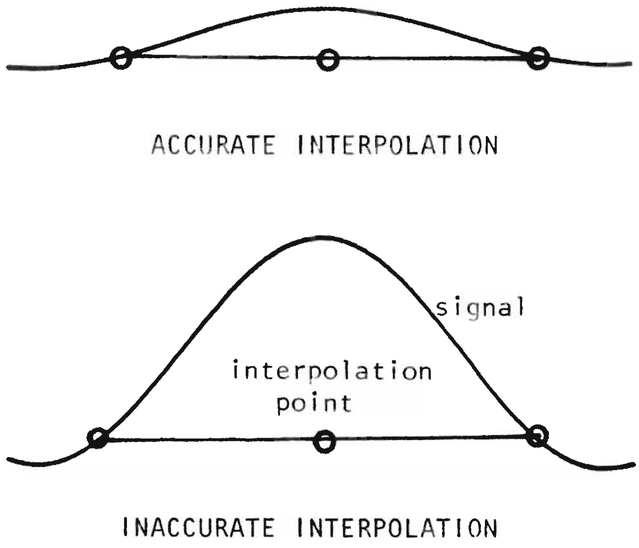

FIGURE 2

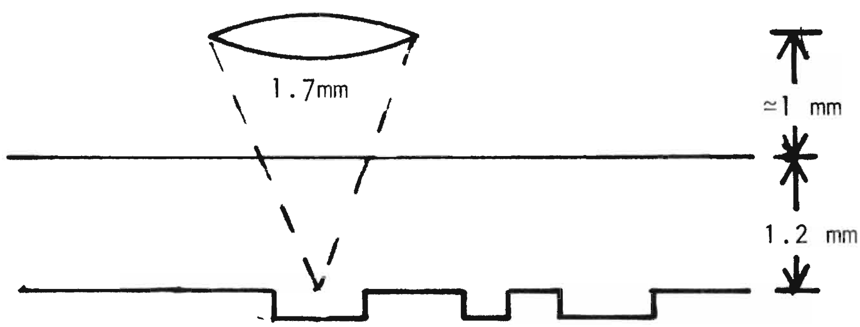

FIGURE 3 — CONSTANT LINEAR VELOCITY DISC
Scale: 1 in. = 11.54 microns

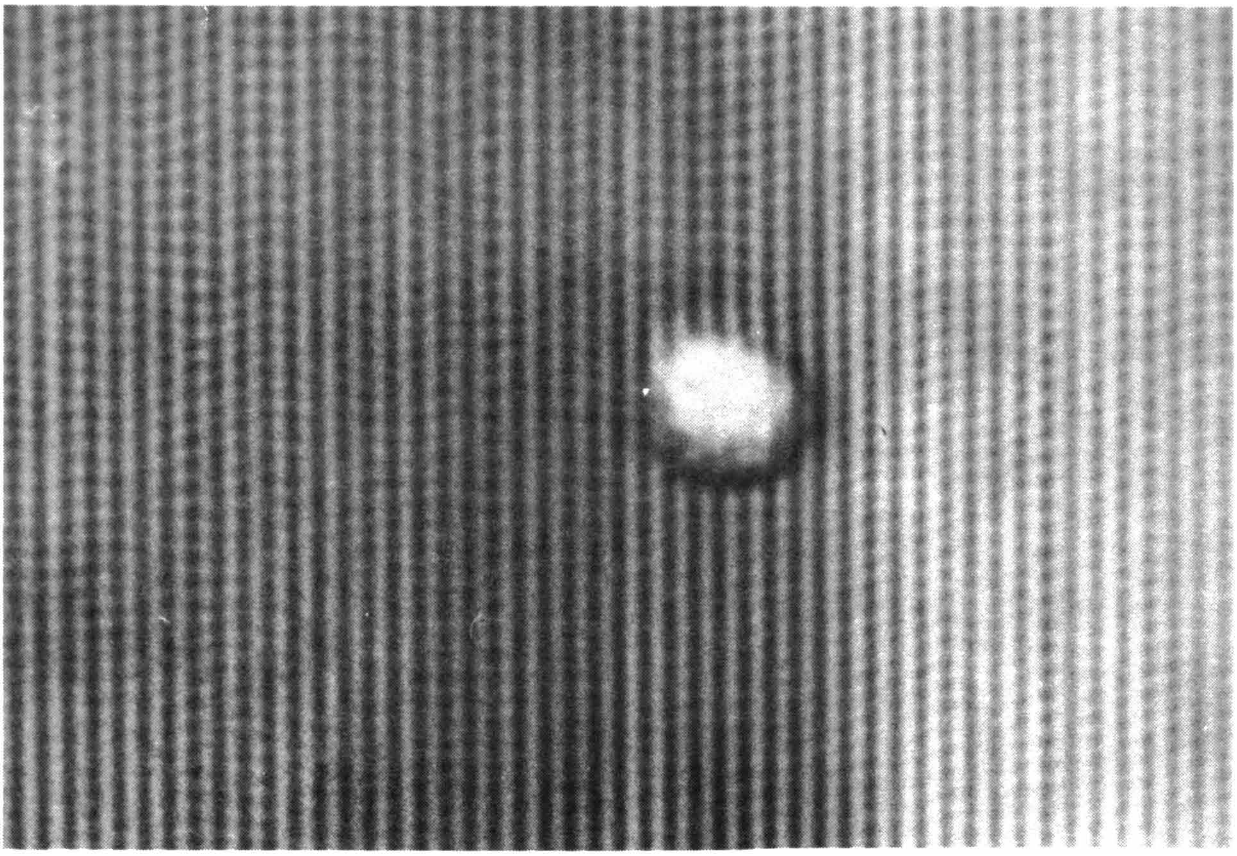

FIGURE 4 — PIONEER VIDEO DISC

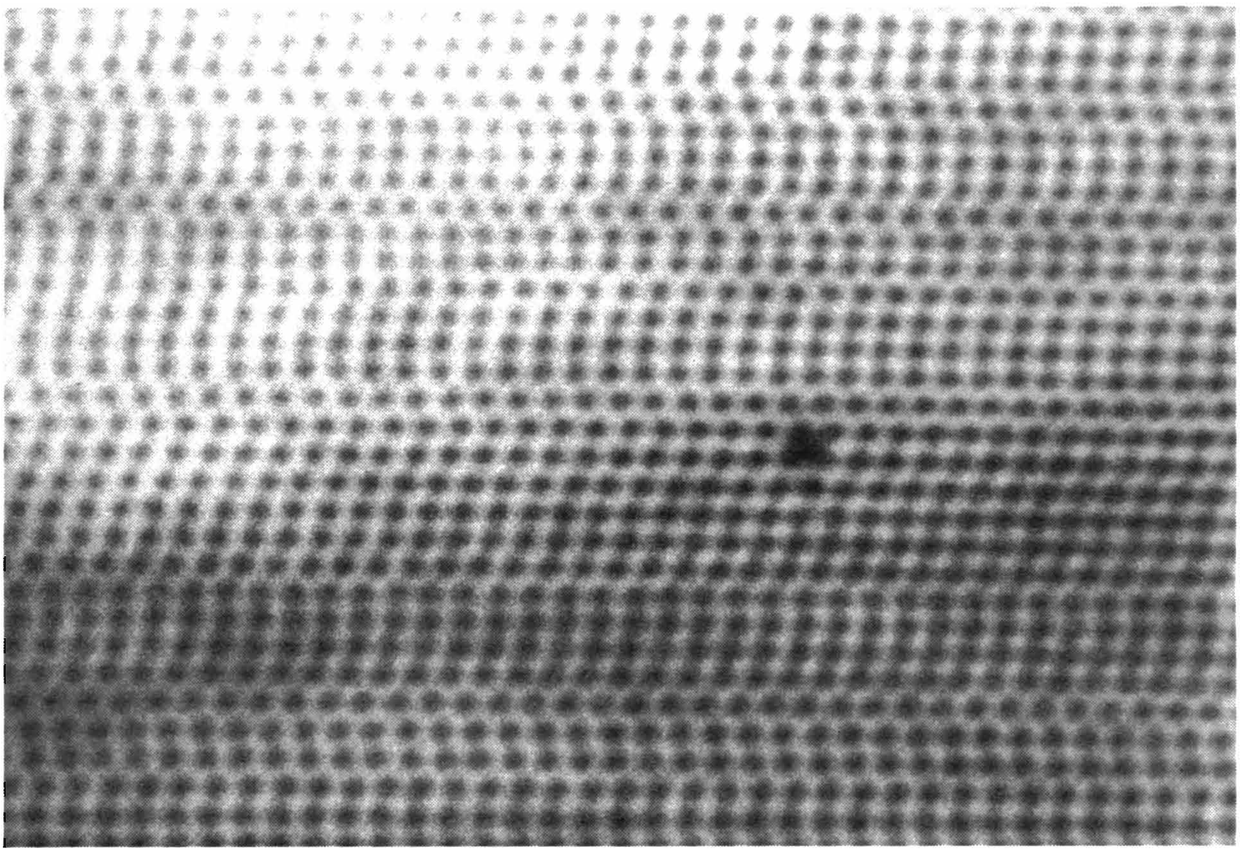

FIGURE 5 — CONSTANT ANGULAR VELOCITY VIDEO DISC

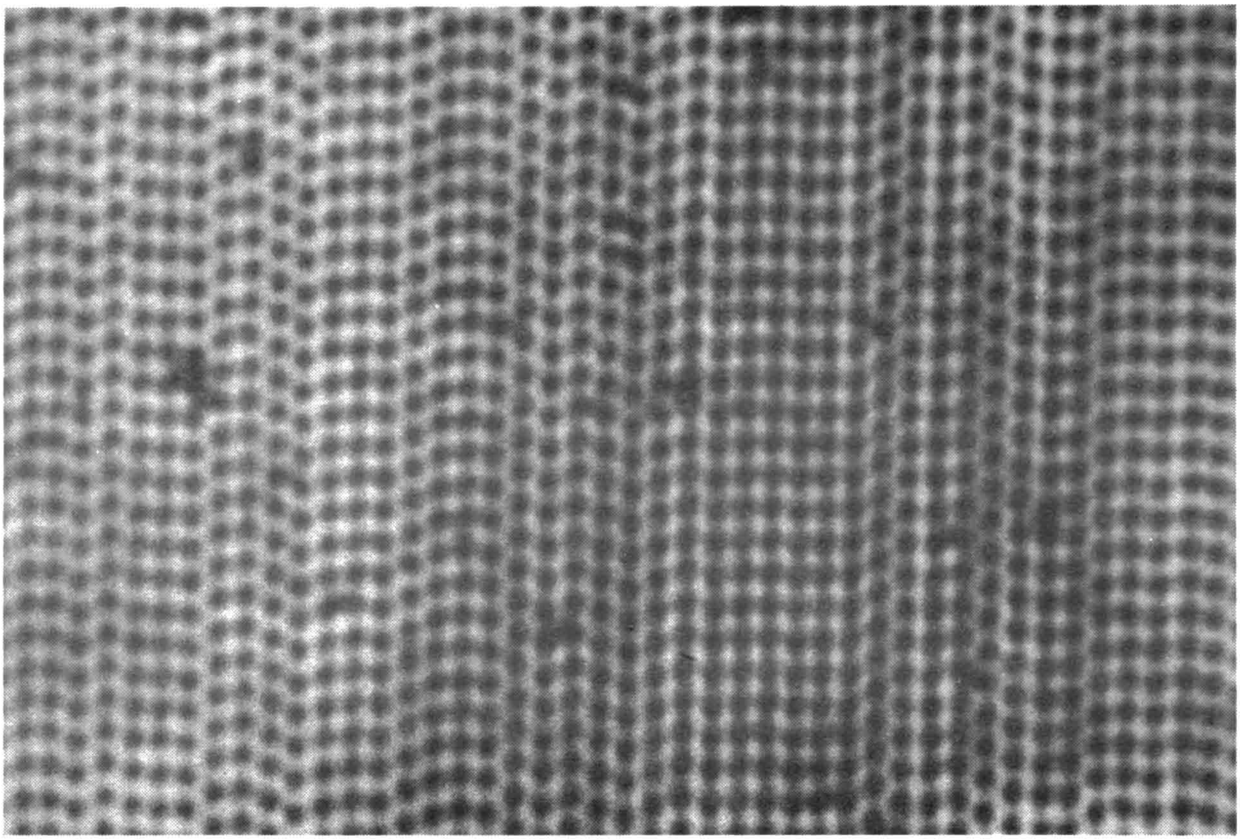

FIGURE 6 — EXPERIMENTAL DIGITAL AUDIO DISC

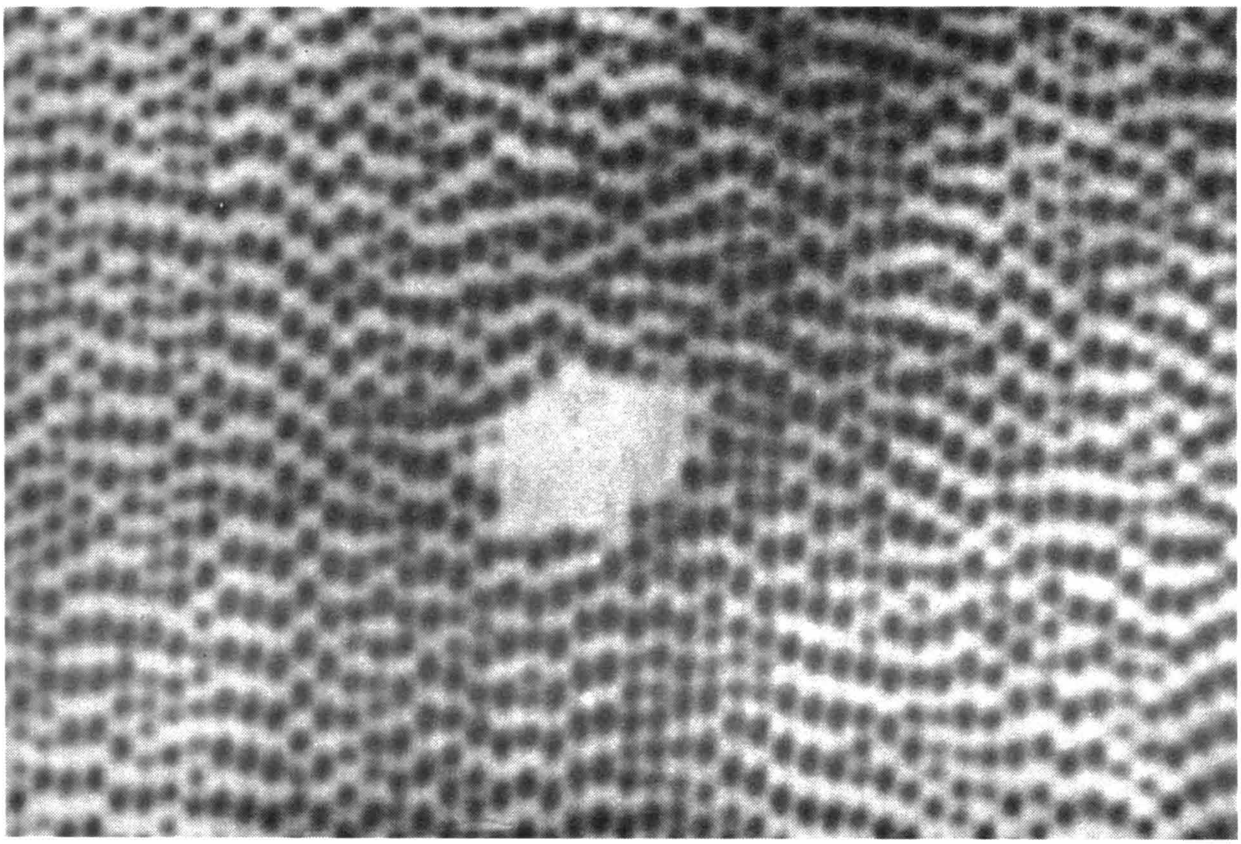

FIGURE 7 — LOCALIZED ERROR   Partial view of .5 mm error

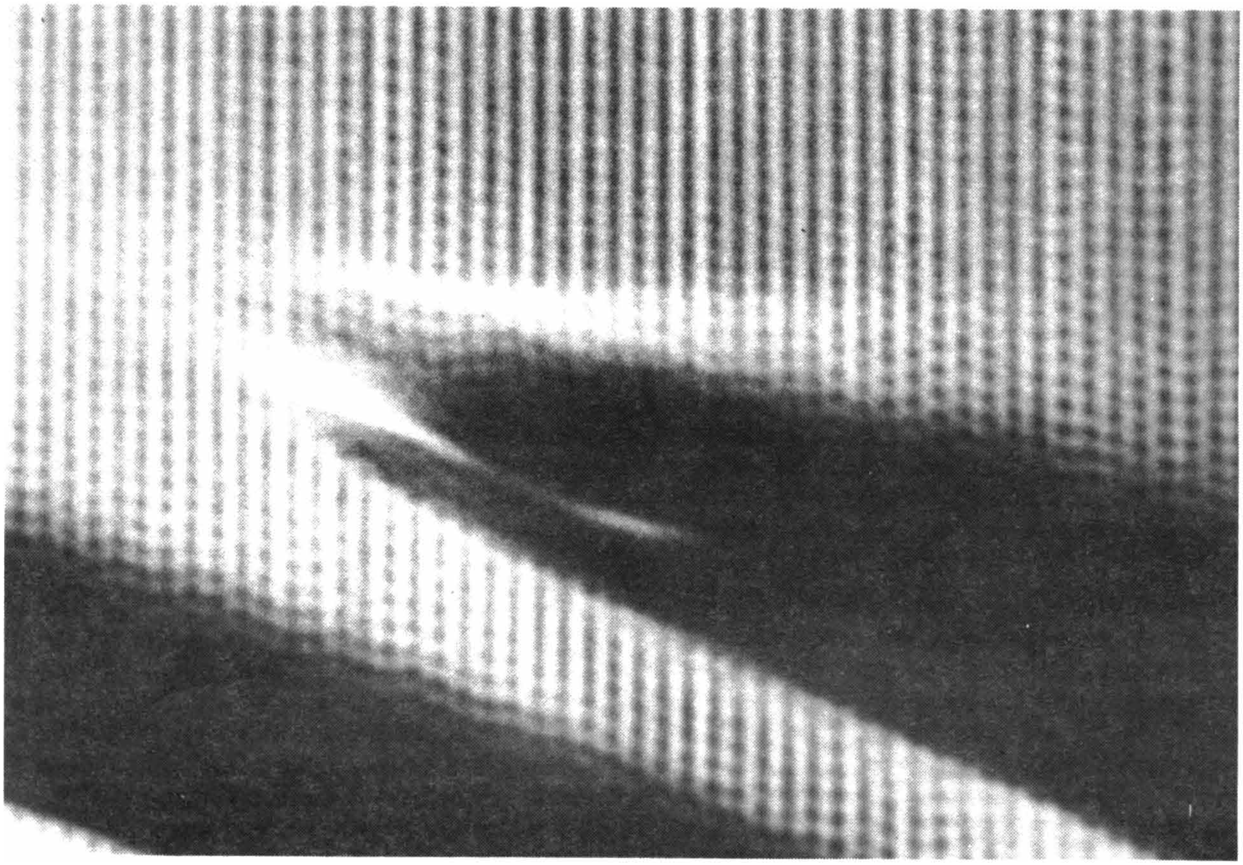

FIGURE 8 — SONY-PHILIPS PROPOSAL