THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 59, No. 9, 1980
Printed in U.S.A.

Direct Digital-to-Analog Conversion of Acoustic Signals

By J. L. FLANAGAN

(Manuscript received December 5, 1979)

We report initial experiments on voice-band acoustic transducers (receivers) designed to accept digital (PCM) signals and directly convert the received bits into analog acoustic output. We describe experimental designs for 4, 5, and 6-bit digital receivers and for an acoustic system that functions as the desampling filter for the decoded signals. We report measurements on the frequency response, amplitude linearity, and signal-to-quantizing noise for the digital receivers and show that these quantities fall into expected design ranges. Though relatively primitive in their initial forms, the transducers accomplish direct digital conversion with acceptable fidelity and suggest a further means for minimizing per-line equipment complexity in digital voice systems.

I. ACOUSTIC TRANSDUCERS FOR DIRECT DIGITAL CONVERSION OF SOUND

To utilize the advantages of digital techniques in the transmission and processing of acoustic signals, especially speech, it is necessary to convert between the analog and digital forms of the information. The traditional means for converting the analog sound wave of a talker into a binary electrical signal involves the separate processes of (*i*) linear transduction of the analog acoustic wave into an analog electrical facsimile, (*ii*) analog electrical filtering to bandlimit the signal and prevent spectral aliasing when the signal is sampled, (*iii*) sampling at or above the Nyquist rate, (*iv*) quantizing the signal amplitude to a prescribed number of approximation levels (2^N, where N is the number of bits), and (*v*) generating binary numbers or code words that represent the quantized samples. The traditional means for converting the electrical digital signal into an analog sound wave involves the separate processes of (*i*) transformation of the binary words into discrete pulses

whose amplitudes equal the quantized levels (that is, a pulse-amplitude-modulated [PAM] representation of the signal), (*ii*) analog electrical filtration to the original frequency range to de-sample the time discrete signal and recover the analog electrical facsimile, and (*iii*) linear transduction of the analog electrical signal to an analog acoustic wave for radiation into the air.

With current circuit technology, most digital electrical processes can be integrated into single units, leaving the linear acoustic transducer and the analog filtration as separate items. A reasonable question, therefore, might be "Is it possible for all the conversion processes to be designed into a single transducer element?" That is, Can one make for the transmitter a single package that accepts the input analog sound wave and provides as output the electrical binary code words? Similarly, Can one make for the receiver a single envelope that accepts the raw binary electrical signal and radiates the reconstructed analog sound wave? In other words, can the traditional analog linear transducers be replaced by "digital" transducers that can convert directly between digital electrical signals and analog sound waves?

In this paper, we describe and demonstrate a solution for the digital receiving transducer (digital to analog). We do not, at the present stage, have implementation for the digital transmitting transducer (analog to digital). In concluding the report, however, we offer some speculations about design possibilities and the outlook for the complementary transmitting transducer.

II. DIGITAL ACOUSTIC RECEIVER

We require an acoustic receiver (earphone) that can be actuated directly by a digital electrical signal. The binary numbers, representing quantized samples of the original signal, must be acoustically decoded to PAM form and desampled to recover the analog acoustic output. To accomplish this, we make use of the fact that a condenser transducer exhibits a diaphragm displacement proportional to the electrical voltage applied across it. We require that the binary electrical signal be presented in sign-magnitude bit-parallel form. In the transducer, we associate independent sound-radiating segments with each magnitude bit. If a bit is true (logical 1), its sound radiator must produce a pulse of sound pressure whose amplitude is proportional to the significance of the bit. The pulse of sound pressure must be a condensation (+) or rarefaction (−), depending upon the sign bit (SGN) of the binary word. Bit pulses are, of course, generated simultaneously at the sampling rate for the digital word. Acoustic superposition, or summation, of the radiated bit pulses is then required to produce the PAM form of the acoustically decoded signal. Acoustic low-pass filtering applied to this

sound wave desamples the time-discrete pulses and recovers the analog waveform.

As indicated above, the pulse of sound generated by the radiating segment associated with a given bit must have an amplitude corresponding to the significance of that bit. A practically useful way to achieve this is to scale the area of each radiating element in accordance with the significance of its bit. Each magnitude bit may then simply gate a constant voltage E to its radiating segment using a polarity + or −, depending upon the SGN bit. The displacement amplitude of each radiating element is proportional to the constant voltage E, and the acoustic volume displacement of each element is therefore proportional to the significance of its bit. The economy and good acoustic performance of the electret condenser element[1] and the fact that it does not require an external polarizing voltage make it appealing for an experimental fabrication. We therefore consider a digital receiving transducer embodying the components shown in Fig. 1. The illustration is made for a 4-bit transducer of circular shape.

The backplate conductor is segmented into separate annular regions equal in number to the number of magnitude bits. The annular regions are insulated from one another and, in this case, are scaled in their areas by powers of 2, which is appropriate for decoding a linear pulse-code-modulated (PCM) signal. For the circular geometry, it is convenient to have the peripheral area correspond to the most significant magnitude bit, and the center area correspond to the least significant bit. The insulating material between the metallized annular regions protrudes slightly to provide small ridges across which a single electret foil can be laid. The metallized side of the electret foil is toward the

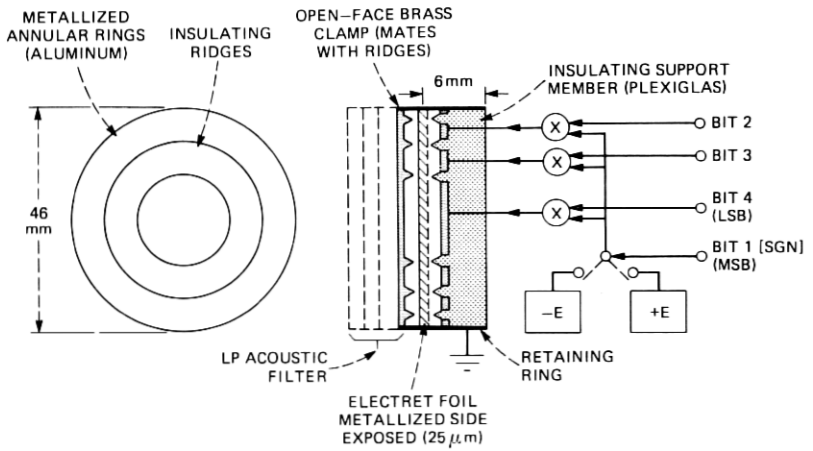

Fig. 1—Design details for the digital receiver. The illustration is for a 4-bit linear PCM receiver.

outside and is normally at ground potential. To insure that the annular areas act as relatively independent radiators, a brass clamping cover mates with the insulating ridges. The cover is slotted widely and hence is acoustically transparent.

The independent radiators, so configured, must work into a cavity that can sum the individual acoustic outputs, and the time-discrete summated signal must be acoustically low-pass filtered to obtain the desampled analog output. We now consider the design of this device.

III. ACOUSTIC DESAMPLING FILTER

In the present work, we are interested in an earphone design that typically must radiate into a closed cavity composed of the telephone earpiece and the human ear canal. We therefore restrict our present experiments to signals produced in such a closed cavity. (Actually, this places more restrictions, in impedance level, upon the design of the acoustic desampling filter than if the device were radiating into open air.) We therefore set down a design for the desampling filter that can function as a telephone receiver. In previous work,[2] we have developed the necessary design techniques, and we utilize these results.

3.1 Equivalent acoustic circuit

As indicated, the necessary filter structure must include a summing cavity containing the active sound radiators for each bit, a low-pass desampling filter of sufficiently high order (we select sixth order, consistent with telephone D-channel bank guidelines), and a final earpiece cavity in which we wish to measure the recovered analog acoustic signal. By proper choice of the filter geometry, the summing cavity can also act as the first element of the low-pass filter. The equivalent lumped-constant circuit for this acoustic system therefore appears as in Fig. 2, where each energy-storage element is expected to exhibit some small loss (owing to viscosity, heat conduction, and radiation resistance). The acoustic volume velocities (currents) circulating in each circuit loop are denoted $U_1, \cdots, U_4$. The input volume velocity U_4 is the superposition of the individual bit sources. The output sound pressure in the earpiece cavity is P_0. We can further simplify this circuit by exploiting our freedom to choose the complex parallel elements to be the same ($z_B = z_2$).*

The loop equations describing this network are therefore

$$\begin{aligned} U_1(z_1 + 2z_2) - U_2 z_2 &= 0 \\ -U_1 z_2 + U_2(z_1 + 2z_2) - U_3 z_2 &= 0 \\ -U_2 z_2 + U_3(z_1 + 2z_2) - U_4 z_2 &= 0. \end{aligned} \tag{1}$$

* This will lead to an earpiece cavity volume of about 6 cc, a realistic value for coupling to the human ear.

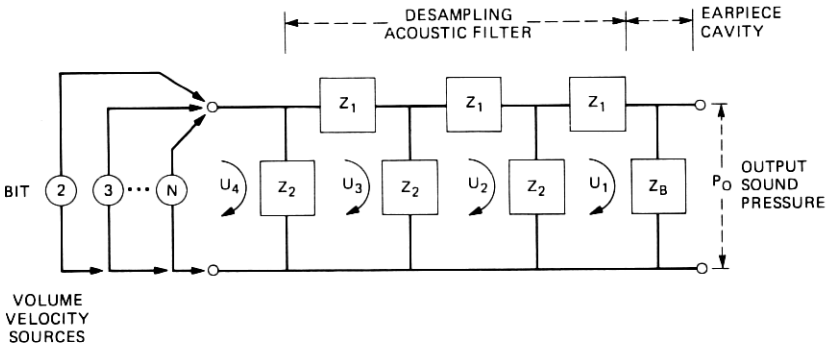

Fig. 2—Equivalent circuit for the acoustic desampling filter and earpiece cavity.

The response of the ladder network to sinusoidal excitation is easily calculated by assuming a unit value for the current U_1 and writing the equation set as the recursion

$$U_{i+1}(j\omega) = \left[\left(2 + \frac{z_1}{z_2}\right) U_i(j\omega) - U_{i-1}(j\omega)\right], \tag{2}$$

$$i = 1, 3 \text{ and } U_0(j\omega) = 0$$

$$U_1(j\omega) = (1 + j\,0),$$

where ω is the radian frequency of the sinusoidal excitation.

Our interest, of course, is in the output sound pressure P_0 developed in the final (front) cavity in response to an electrical voltage applied to the transducer. Because, for a condenser element, the diaphragm displacement is proportional to the applied voltage e, the input volume velocity to the acoustic filter produced by electrical excitation of any transducer is $U_4(j\omega) \sim j\omega e(j\omega)$. The output sound pressure is $P_0(j\omega) = z_2 U_1(j\omega)$, so that the transfer function of interest is proportional to

$$(P_0/e) = j\omega z_2 (U_1/U_4). \tag{3}$$

Observing that for our calculation $U_1(j\omega) = (1 + j0)$ and that in simplest form and neglecting losses, $z_B = z_2 = 1/j\omega C$, where C is the acoustic compliance of the earpiece cavity, then

$$(P_0/e) \sim (C\,U_4)^{-1}. \tag{4}$$

3.2 *Computed responses of acoustic system*

To construct the desampling filter and acoustic system of Fig. 2, it is convenient to take the serial and parallel impedances as

$$z_1 = (R_1 + j\omega L)$$

$$z_2 = \left(R_2 + \frac{1}{j\omega C}\right). \tag{5}$$

We make use of simple relations established earlier[2] for the reactive acoustic elements and, to first order, represent the acoustic inertance (inductance) of a thin circular aperture and the acoustic compliance (capacitance) of a right circular cylinder as

$$L = \rho/d \qquad \text{(cgs acoustic henries)}$$

$$C = \frac{Al}{\rho c^2}, \qquad \text{(cgs acoustic farads)} \quad (6)$$

where ρ is the air density, c the velocity of sound, d the diameter of a circular aperture, and A and l the cross-sectional area and length, respectively, of a cylindrical cavity. We will be satisfied to estimate the resistive losses R_1 and R_2 from past experience with damping material (cotton, felt, or silk screen).

To calculate the expected response of the acoustic system, we wrote Fortran programs to compute the quantities given in (3) and (4), utilizing the relations in (5) and (6), and to display the amplitude and phase responses of the complete acoustic system.

Using these programs, several trial calculations easily establish values for L, C, and their associated losses, that lead to practical and desirable values for d, A, and l.* Particularly useful values are

$$\begin{aligned} L &= 3. \times 10^{-3} && \text{cgs acoustic henries} \\ R_1 &= 1 - 10 && \text{cgs ohms} \\ C &= 4. \times 10^{-6} && \text{cgs acoustic farads} \\ R_2 &= 1 - 10 && \text{cgs ohms.} \end{aligned} \quad (7)$$

The computed amplitude and phase responses for the acoustic system are shown in Fig 3. The responses are calculated for two values of resistive loss—a low value $R_1 = R_2 = 1$, which represents underdamped behavior (and which puts several poles in evidence), and a high value of $R_1 = R_2 = 10$, which represents a response that is relatively heavily damped. The total network is seventh order. One sees, too, that the cut-off frequency of the acoustic system is appropriate for voice-band signals.

3.3 Effects of cavity leaks

As part of the design, it is of importance to examine the effects of air leaks in the cavities producing the compliances C. (The simplest and most convenient mechanical fabrication of these chambers makes it difficult to insure airtight fits.) A puncture (or leak) in each cavity allowing connection to the outside air is, to first order, represented by a lossy inertance L_p in parallel with C, making the acoustic system more like a bandpass filter. If we make a simple modification to the

* The physical implementation of the acoustic system will incorporate methods for control of short-wave-length cross-mode effects. These are discussed subsequently.

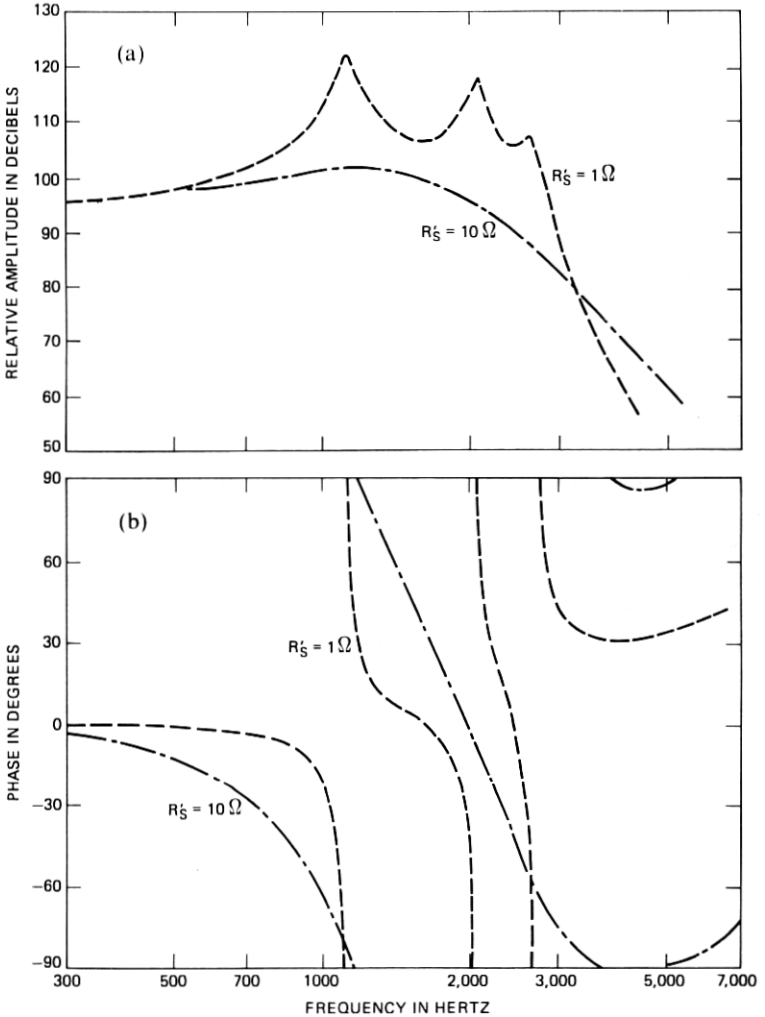

Fig. 3—Computed responses for the acoustic network; (a) log-amplitude (dB) and (b) phase (degrees) versus log-frequency (Hz). Calculated values are given for two values of loss: $R_1 = R_2 = 1$ cgs acoustic ohm and $R_1 = R_2 = 10$ cgs acoustic ohms.

response program, i.e., let

$$z_2(j\omega) = \left[R_2 + j \left(\frac{L_p}{1 - \omega^2 L_p C} \right) \right],$$

we can examine the effects of leaky cavities.

A single puncture or leak in each cavity of about 1 mm diameter corresponds approximately to $L_p = 12. \times 10^{-3}$ acoustic henries. With this size inertance communicating to the outside air, the amplitude

response for the acoustic system is shown in Fig. 4. The response is illustrated for a low value of loss R's = 1., and a relatively large value of loss R's = 10. In addition to modifying the mode structure through the addition of reactive shunt elements, cavity leaks clearly tend to diminish the low-frequency response and produce a bandpass characteristic. As the leaks diminish in size, i.e., as $L_p \rightarrow \infty$, the response approaches that shown previously in Fig. 3.

These computed responses, together with the transducer design notions of Fig. 1, let us proceed to an experimental system.

IV. EXPERIMENTAL RECEIVING TRANSDUCERS

Following the notions outlined in Section II, we constructed experimental receivers for direct conversion of 4-bit, 5-bit, and 6-bit linear PCM sign-magnitude digital signals. For obvious reasons, we consider it desirable to choose an overall shape and size for the experimental transducers to make them comparable to the conventional telephone receiver. We therefore used an overall diameter of 46 mm for the circular backplate that comprises the supporting member and annular conducting segments.

We machined the insulating plexiglas support for the prescribed number of conducting segments, leaving the protruding ridges as shown previously in Fig. 1. All transducers used the peripheral annulus for the most significant magnitude bit and the center conductor for the least significant bit. For linear PCM, the areas scale as powers of 2, and the outside radius of the peripheral annulus is constrained to be

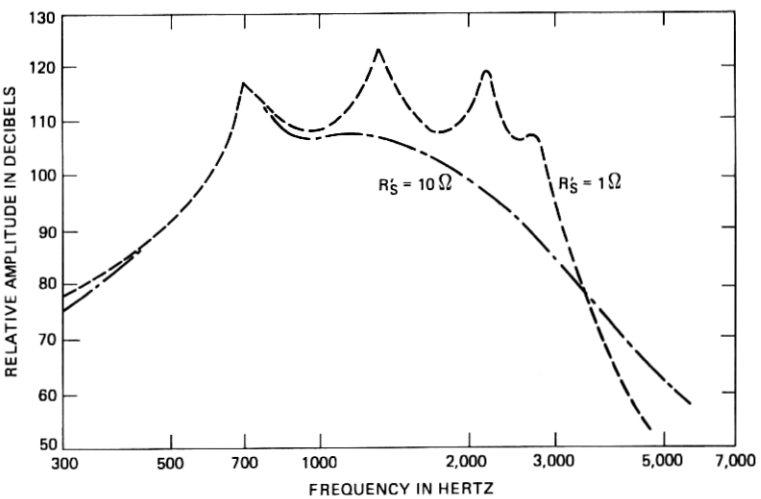

Fig. 4—Computed effects of air leaks in the acoustic cavities $L_p = 12 \times 10^{-3}$ cgs acoustic henries. Responses are calculated for two values of loss.

Table I

Number of bits (N)	Outside radii of contiguous annular segments (in mm)				
	r_1	r_2	r_3	r_4	r_5
4	8.7	15.1	23.0	—	—
5	5.9	10.3	15.7	23.0	—
6	4.1	7.2	10.9	16.0	23.0

the usable radius of the circular transducer $r_{MAX} = 23$ mm. An N-bit transducer has M independent radiating segments, where $M = (N - 1)$, and for this assignment of bit radiators the outside radius of the mth annulus is given by

$$r_m = r_{MAX}\left[\sum_{i=1}^{m} \frac{1}{2^{M-i}} \Big/ \sum_{j=1}^{M} \left(\frac{1}{2^{j-1}}\right)\right]^{1/2}. \tag{8}$$

The dimensions resulting from this design are shown in Table I.

To produce the conducting annuli, we vacuum-deposited approximately 5000 Å of aluminum onto the Plexiglas* conductors and polished the protruding ridge insulators free of the metal. We used silver paste to connect wires (entering through the supporting members) to each metallized annulus. Individual contacts were therefore available for their respective magnitude bits in the digital word. For each transducer, a single electret foil provided the polarized solid dielectric and the external conducting element of the condenser transducer (see Fig. 1). Each diaphragm was composed of a film of Teflon† 25-μm thick, vacuum metallized with aluminum on one side, and given an electrostatic polarization by raster scanning with an electron beam (approximately 20 kev) in vacuum.[1] To facilitate operation of the annular segments as independent sound radiators, each transducer was fitted with a widely slotted brass clamp that mated with the insulating ridges of the back plate as shown in Fig. 1.

V. IMPLEMENTATION OF ACOUSTIC DESAMPLING FILTERS

All the transducers described in Section IV require low-pass desampling filters to recover the analog acoustic output. We make use of the design results of Section III, together with a healthy respect for the short-wavelength effects that can exist at audible frequencies in acoustic structures with dimensions as large as 46-mm diameter.[2]

To mitigate the effects of cross modes, we effectively reduce the dimensions of the acoustic filter by making it a parallel connection of four identical structures. Moreover, we introduce angle elements into each cavity partition to combat cross modes, especially the first radial

* Trademark of Rohm & Haas Co.
† Trademark of E.I. DuPont de Nemours.

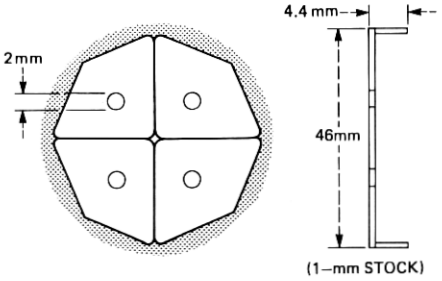

Fig. 5—Physical design of a single section of the acoustic desampling filter.

mode. In accordance with these notions, the design of each section of the desampling filter is given in Fig. 5.

Additionally, as mode control, we utilize irregular inserts into the periphery of the earpiece cavity. With this insert, the cavity volume is approximately 6 cc, and we will utilize this cavity for measurement of the output sound pressure. Still further, we introduce light absorbent cotton patches into each cavity to provide a system damping corresponding approximately to the values $R_1 = R_2 = 5\Omega$. The resulting acoustic filter, earpiece cavity, and the 4-bit and 6-bit digital receivers so constructed are pictured in Fig. 6.*

VI. DIGITAL CIRCUITRY FOR TRANSDUCER DRIVE AND MEASUREMENT

It is of importance to assess the performance not only of the digital transducer element, but also of the complete acoustic system. Relevant issues relate to frequency response, dynamic amplitude range, and the ability to decode direct digital drives. Test signals useful for these assessments include sinewaves, bands of noise, and speech.

A means for directly characterizing the performance of the acoustic filter system is to measure its response when driven by a single-radiator conventional analog transducer. To properly characterize the digital transducer, on the other hand, requires a bit-parallel, sign-magnitude binary signal. We therefore arranged laboratory circuitry to provide these test signals and both analog and digital drives, as illustrated in Fig. 7a.

The acoustic output of the system is the quantity of interest. We measure this output by a pressure pick-up in the 6 cc earpiece cavity. We insert a calibrated ½-inch condenser microphone into the cavity wall with a reasonably tight fit. We amplify this signal for measurement. It is convenient to conduct this preliminary experimentation in an untreated laboratory room, as opposed to a more elaborate (and

* While our experimental acoustic system is machined from brass, an easier and highly economical mass fabrication would utilize injection-molded plastic.

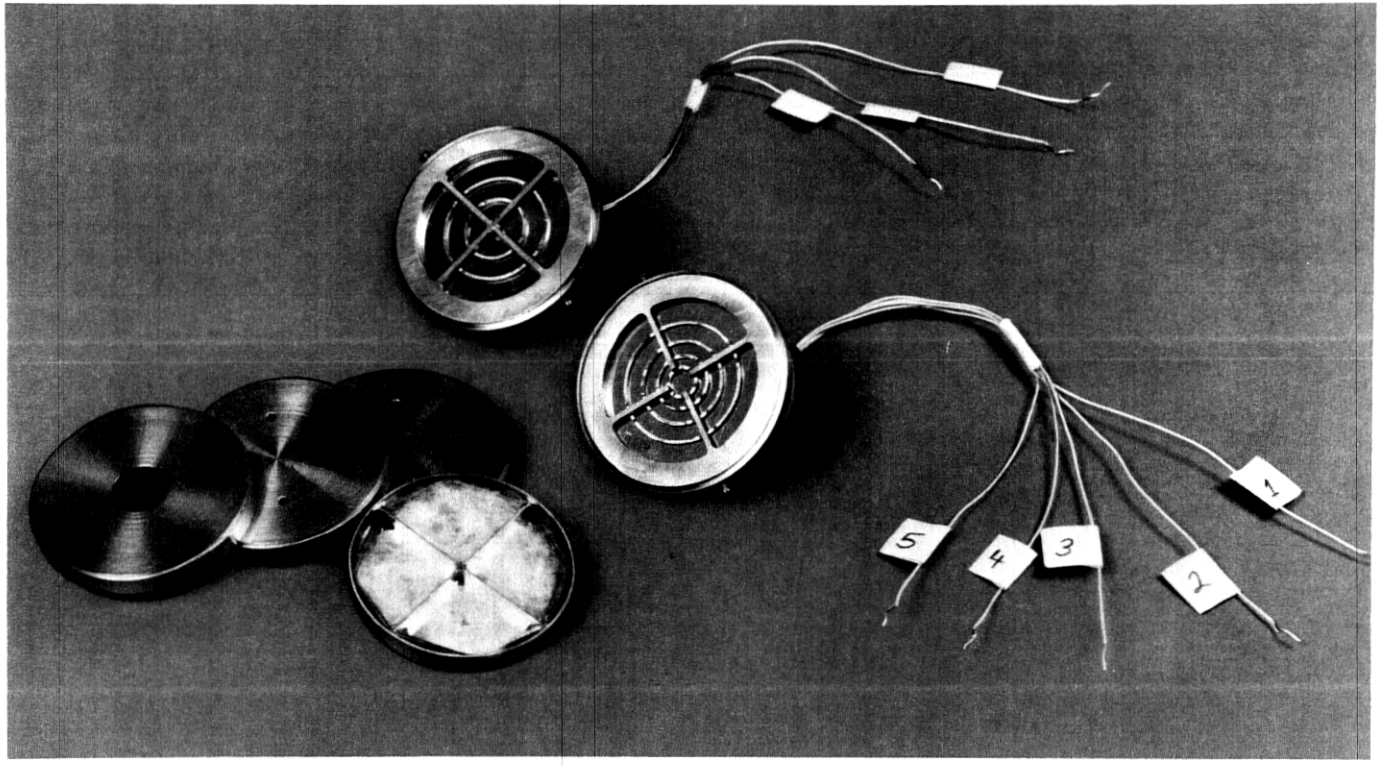

Fig. 6—The experimental 4-bit and 6-bit digital receivers, the acoustic desampling filter, and the earpiece cavity.

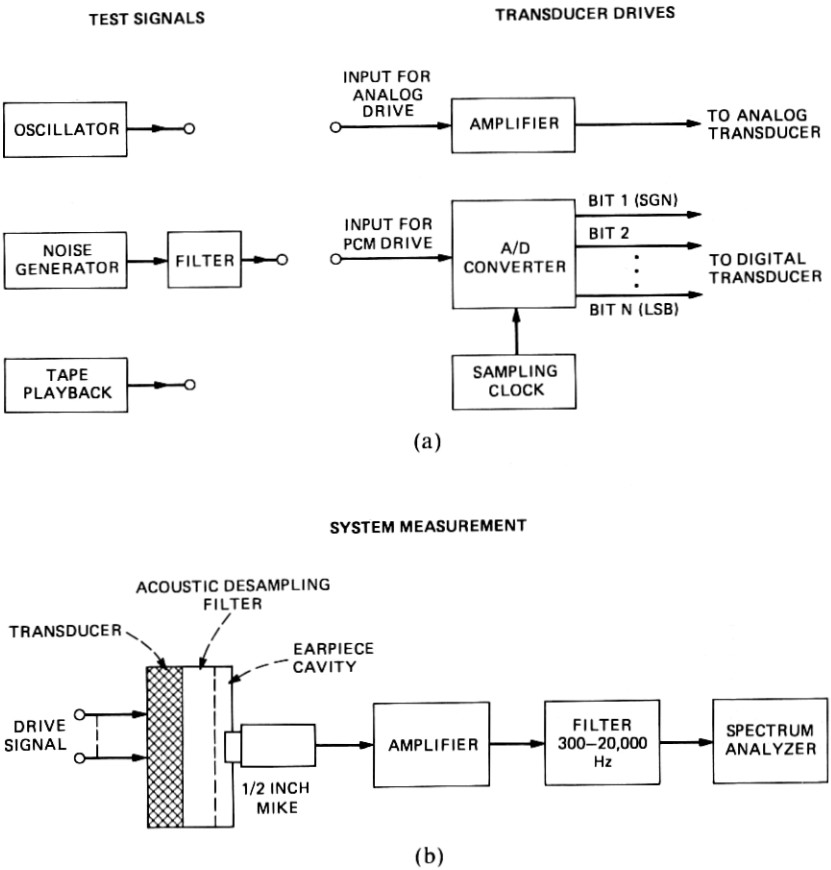

Fig. 7—Block diagram of laboratory equipment for measurements on the digital receivers and acoustic system. (a) Signal sources and transducer drives. (b) Circuit arrangement for measuring acoustically decoded sound output.

remote) set-up in the anechoic chamber. An important factor, of course, is the ambient acoustic noise in the laboratory room environment. But most of the room noise is relatively low in frequency and can be suppressed effectively by a measuring system that has a response that is high-pass at 300 Hz and essentially flat thereafter to above 20,000 Hz. To allow this convenience, we utilize the measuring circuitry shown in the block diagram of Fig. 7b.

Further, it is particularly convenient to measure the frequency responses of the system with a fast Fourier transform (FFT) spectrum analyzer. In all the data we present here, therefore, we obtain the frequency spectrum by having the FFT analyzer average over 128 spectral computations. The scope display of the analyzer shows the complete spectrum to 20,000 Hz and specific frequency values can be read by cursor control.

VII. RESULTS

7.1 Analog excitation; frequency response of the acoustic system

A convenient means to examine the frequency response of the acoustic system is to use wideband random noise delivered as an analog drive to a conventional single-radiator analog transducer. Alternatively, the digital transducer can be used as a single linear element by connecting the individual radiating elements electrically in parallel. The single-element electret transducer was therefore driven with a flat spectrum (300 to 20,000 kHz) noise voltage with 10 V rms amplitude applied to the transducer. The input signal and the observed acoustic response as taken from the scope display of the FFT analyzer are shown in Fig. 8. The vertical cursor of the spectrum analyzer is set here to a frequency of 3000 Hz. In all data presented here, we display responses on a decibel amplitude scale (10 dB/tick) and a linear frequency scale, 0 to 20,000 Hz. The latter wide range is used throughout this paper because we recognize the importance of cross-mode control in the acoustic system, and we wish constant reassurance that the measurements are not being influenced by high-frequency components. In this case, the ambient acoustic room noise is greater than 30 dB down at all frequencies.

As Fig. 8 suggests, the acoustic desampling filter and the summing and earpiece cavities perform largely as designed and calculated. The measurement does not show any significant cross-mode excitation, and such components are typically more than 30 dB down in amplitude.

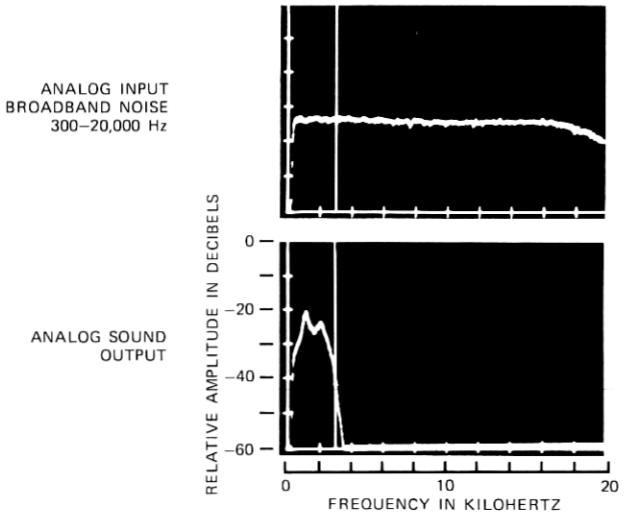

Fig. 8—Spectrum analyzer displays for broadband analog noise excitation of the acoustic system.

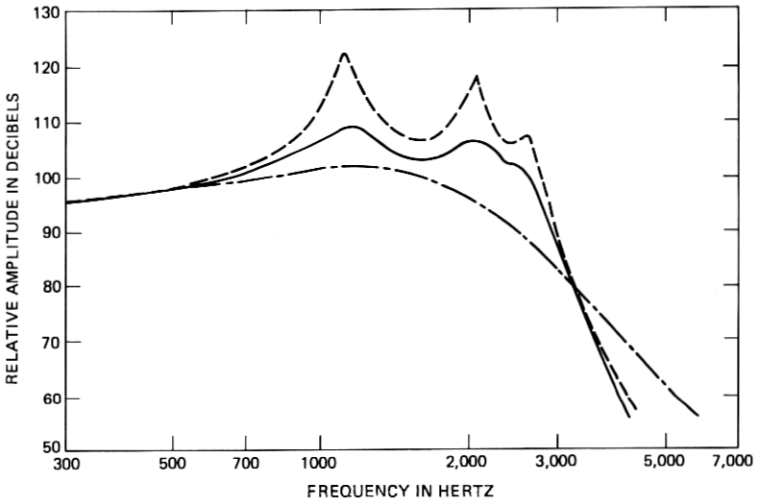

Fig. 9—Measured frequency response of the acoustic system. The calculated responses of Fig. 3 are reproduced for comparison.

We are therefore reassured that the mode-control design is acceptably effective.

Through its cursor control and digital readout, the spectrum analyzer permits point-by-point measurement of the frequency response, and this measurement is compared with the earlier computer calculation (both lightly damped and heavily damped) in Fig. 9. The measured response is essentially bracketed by the computed curves, suggesting that the estimated damping of $R_1 = R_2 = 5$ cgs ohms is approximately correct. The low frequency response of the system, at least down to about 300 Hz, agrees with the calculated responses suggesting that the fit of the acoustical system is adequate. The voltage output of the calibrated measuring microphone corresponds to a sound pressure level of approximately 90 dB re 0.0002 dynes/cm^2 in the earpiece cavity. (This sizeable pressure for the single element radiator is 5 to 10 dB greater than will be produced by the 4-bit and 6-bit digital transducers.) Considering the parameters involved, the agreement between measurement and calculation seems satisfactory, and the analog response suggests an acceptable characteristic as a desampling filter.

Another important property of the system is its linearity and dynamic range. A noise-band signal 300 to 3000 Hz delivered as an analog voltage to the single-element electret transducer is useful for this measurement. The rms acoustic output and the rms drive voltage are plotted for this excitation in Fig. 10. The acoustic filter is, of course, a substantively linear system for the sound levels of interest here. The single-element analog transducer is also seen to be relatively linear

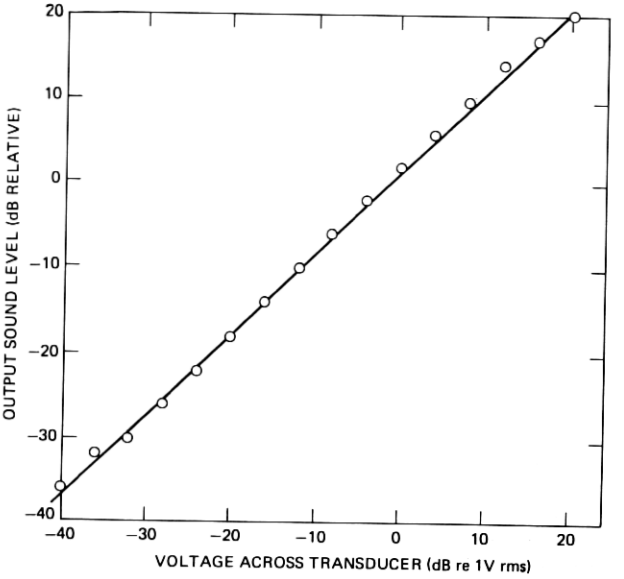

Fig. 10—Amplitude linearity of the analog transducer and acoustic system. The signal is a noise band.

over a range of more than 50 dB. Below −40-dB room noise limits the amplitude measurements.

7.2 Direct digital excitation

Having confirmed that the acoustic system behaves as designed, the next and main interest is in the performance of the digital transducers driven directly from binary PCM signals. Though we have fabricated 4-bit, 5-bit, and 6-bit transducers, it is adequate for our purposes here to describe and compare measurements on only the 4-bit and 6-bit devices. Performance of the 5-bit transducer interpolates properly between the other two.

Direct PCM drive is accomplished by directing individual magnitude bits, polled by the sign bit, to respective individual sound radiators. Bits are presented at the sampling rate of the original A/D conversion. A latch on each bit holds the bit condition for the sampling period. In all cases here, the magnitude of the bit voltage delivered to the radiating segment corresponds to $E = 20$v.

7.2.1 Characteristics of the digital test signals

Convenient signals for measurement include band limited noise and sinewaves. To produce these in digital form, they must be sampled, quantized, and digitally converted. We use a conventional 12-bit A/D converter for this and connect to the number of bits desired. Our acoustic system is designed for the nominal frequency range 300 to

3000 Hz; hence, we focus on signals of this spectral extent. For convenience with existing equipment and to provide a moderate over sampling, we use a sampling rate of 9150 Hz throughout. We can illustrate the characteristics of the test signals produced by this arrangement.

Figure 11 shows the amplitude spectrum for an input analog noise band of 300 to 3000 Hz (limited by a fourth order filter). The cursor of the spectrum analyzer is set at 3000 Hz. Also shown is the spectrum of the sampling signal (45-μsec pulses produced at the 9150-Hz rate) and the spectrum of the sampled noise band. The cursor of the spectrum analyzer marks the 9150-Hz sampling rate. One sees, of course, that the sampled signal is nonbandlimited. As further information, also shown is the spectrum of the sign bit signal (which dictates the polarity of the magnitude bits), the spectrum of the first magnitude bit (bit 2) which connects to the largest radiator segment, and an electrically summated signal for a 4-bit drive (properly scaled for significance electrically). The latter signal is the electrical facsimile of the sound pressure appearing in the summing cavity immediately in front of the sound radiators. In all cases, one can identify the spectral extent of the original input signal.

A similar display for a sinewave input is shown in Fig. 12. The analog sine input is from an oscillator set to 1500 Hz. The input spectrum shows that the oscillator (an RC design) produces a slight amount of odd harmonic distortion. For the sinusoidal input, the spectra of the sampled signals approximate line spectra and the fine structures of the bit signals, i.e., the components arising from quantization, are, of course, directly conditioned by the relationship between the input frequency and the sampling rate.

7.2.2 Measurements on acoustically decoded output

The acoustically decoded sound outputs from the 4-bit and 6-bit transducers are shown in Fig. 13 for the input noise band of 300 to 3000 Hz. In all panels of the figure, the spectrum analyzer cursor is set to 3000 Hz. The ambient acoustic noise for the measurement is shown in the lower panel, suggesting a signal-to-noise ratio of the order of 30 dB for all but the lowest frequencies.

The acoustic output from the digital transducers, as revealed by the noise-band signal, is similar in both cases and is primarily conditioned by the acoustic-filter system. The cursor and digital readout of the spectrum analyzer permit point-by-point measurement along the spectrum, and these output data are compared to the design spectra in Fig. 14. The output sound level in the earpiece cavity for this signal is approximately 80 dB re 0.0002 dynes/cm^2. Also shown for completeness in Fig. 14 is the measured spectrum of the analog input noise band.

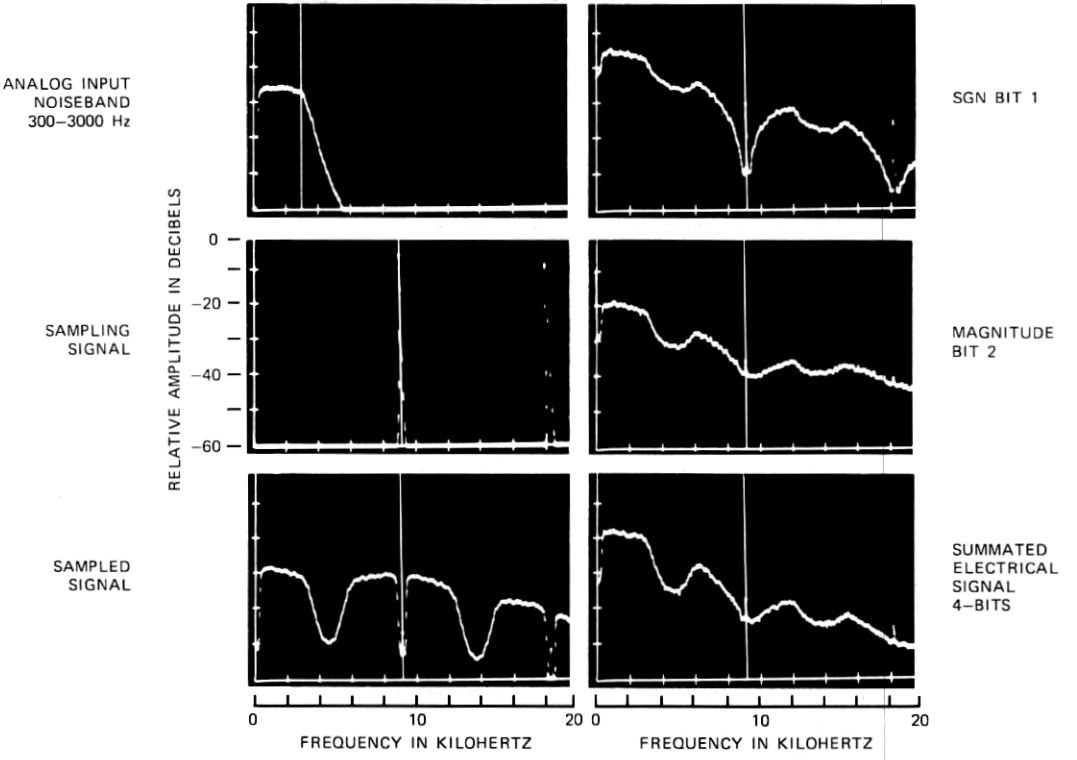

Fig. 11—Spectra for noise band test signal used for measurements on the digital receivers. The cursor is positioned at 3000 Hz in the top left panel and at the sampling frequency 9150 Hz in the remaining panels.

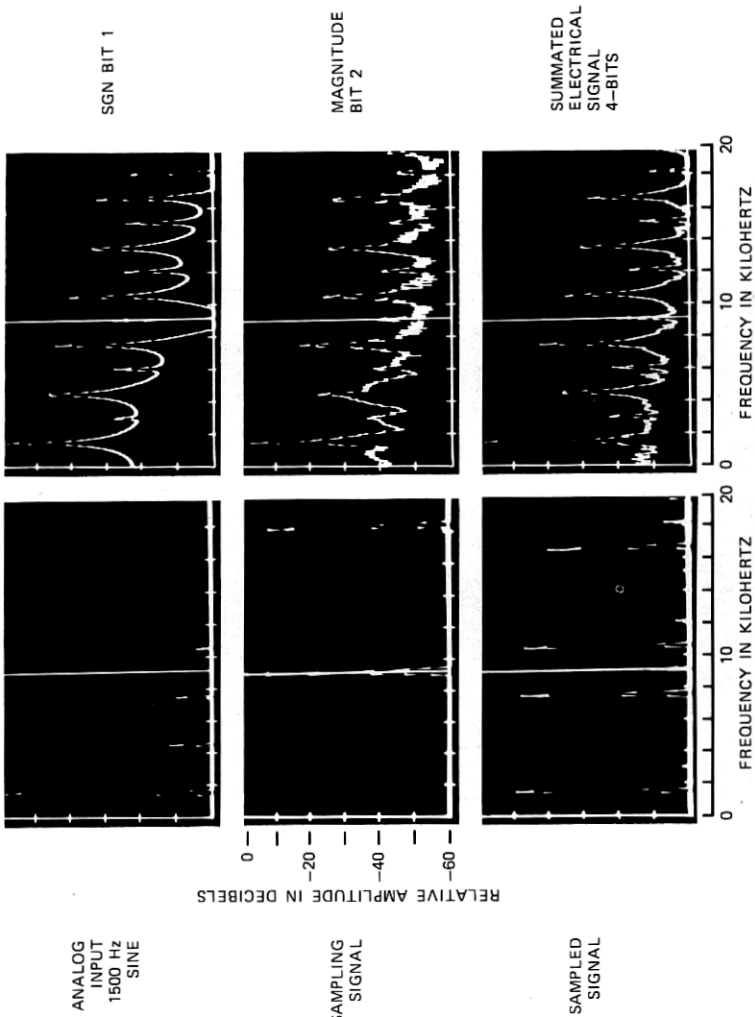

Fig. 12—Spectra for sinewave test signal used for measurements on the digital receivers. The analog input signal exhibits a small amount of odd-harmonic distortion.

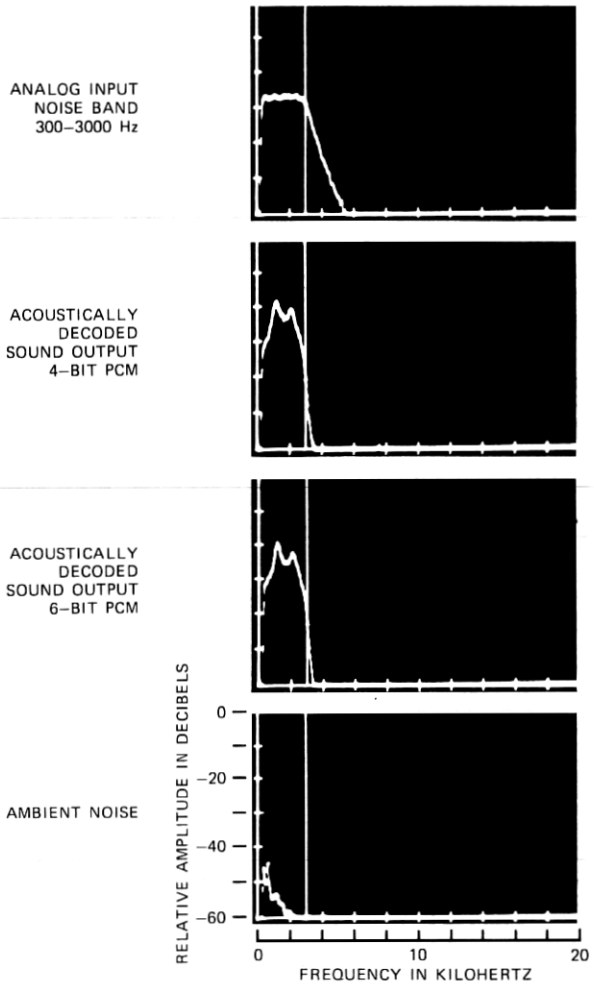

Fig. 13—Spectrum analyzer displays of acoustically decoded outputs for the 4-bit and 6-bit digital receivers. The input signal is a noise band 300 to 3000 Hz. The cursor is positioned at 3000 Hz.

A similar measurement for the sinusoidal input, using a midband frequency of 1500 Hz, is slightly more informative in comparing the 4-bit and 6-bit transducers. These data are shown in Fig. 15. Close comparison shows that, as would be expected, the inband quantizing components are noticeably more emphasized for the 4-bit transducer than for the 6-bit transducer. Especially strong quantizing components are noticeable at several high frequencies. These are typically more pronounced for the 4-bit transducer than for the 6-bit transducer. Some of these components may be slightly favored at cross-mode eigenfrequencies of the acoustic system.

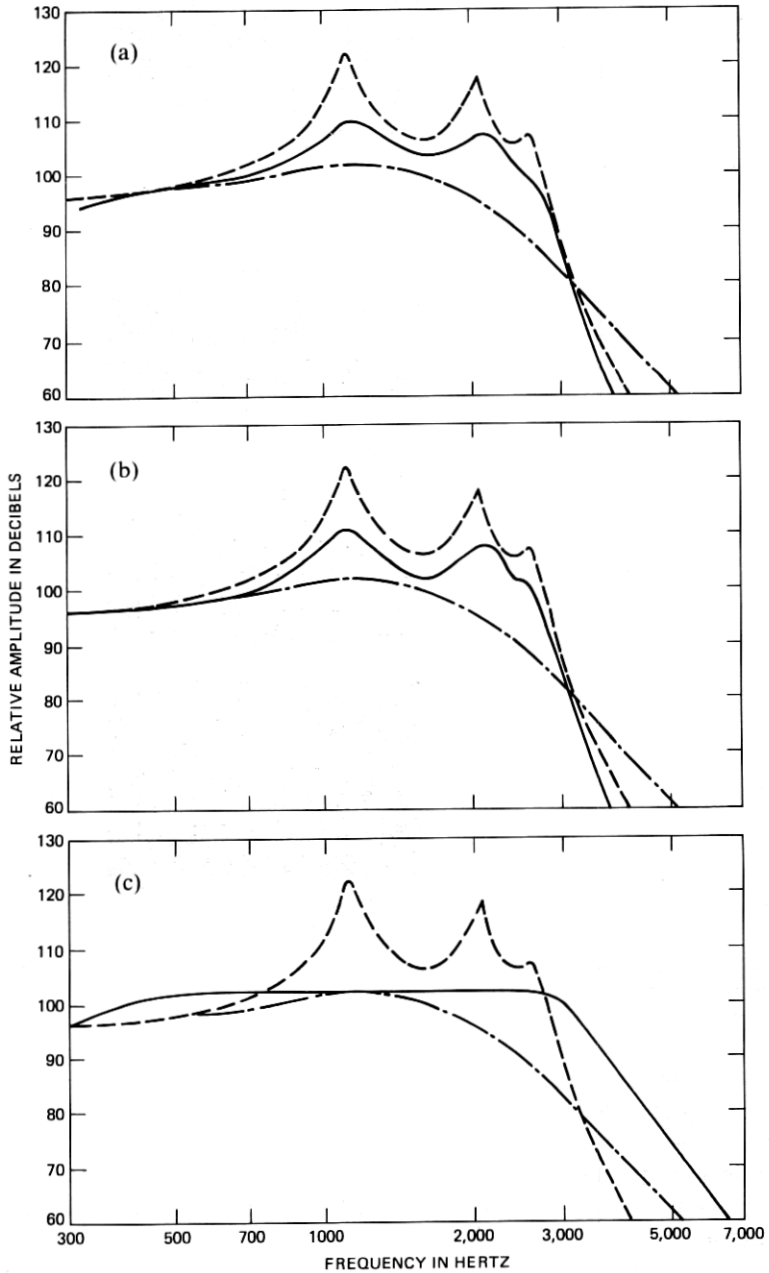

Fig. 14—Measured frequency responses for the acoustically decoded noise band signals. (a) 4-bit receiver. (b) 6-bit receiver. (c) Input noise band spectrum.

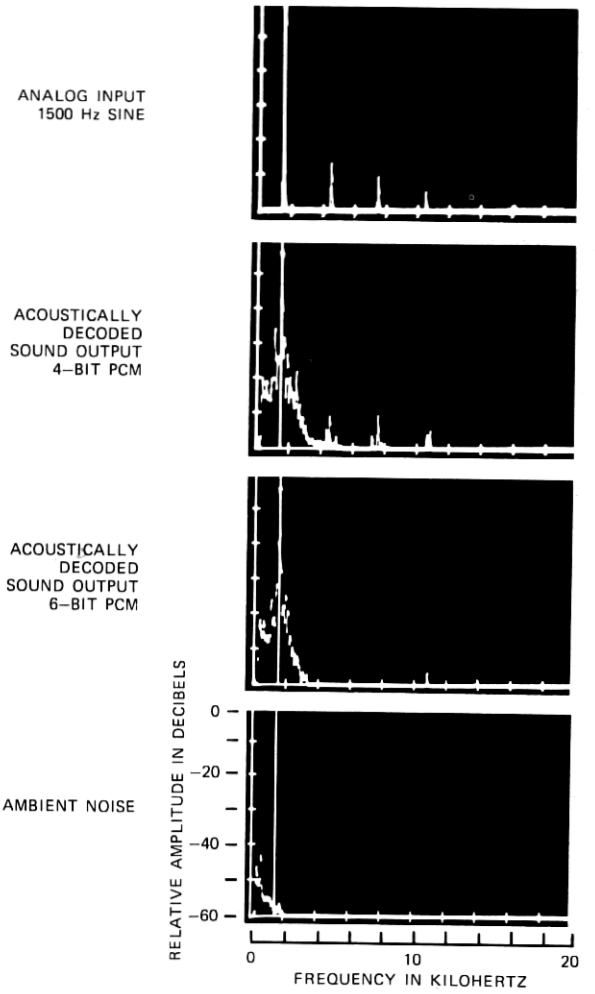

Fig. 15—Spectrum analyzer displays of acoustically decoded outputs for the 4-bit and 6-bit digital receivers. The input signal is a 1500-Hz sine wave. The cursor is positioned at 1500 Hz.

While conventional measurement of signal-to-quantizing noise in the acoustic decoded output is limited by the ambient room noise, we have attempted to estimate this ratio in two ways. For sinusoidal input to the digital receivers, we have made phase-locked cancellations of the 1500-Hz signal component in the output acoustic spectrum, leaving only the quantizing noise and ambient room noise, and using the spectrum analyzer we have made amplitude measurements of the quantizing components in the output acoustic spectrum. When signal and noise powers are compared, the measurements yield signal-to-

noise (s/n) ratios of 19 dB for the 4-bit transducer and 26 dB for the 6-bit transducer. The measurements indicate a 7-dB difference between the two devices, which ideally should be 12 dB (i.e., 6 dB per bit). However, considering the limitations placed upon the measurement by the ambient room noise (which, of course, contributes to the amplitude of the inband distortion components), this preliminary assessment of s/n ratio seems acceptable. Additional measurements with a phase-inverting distortion analyzer give similar results at frequencies of 500, 1000, and 2000 Hz.

The amplitude linearity of the 4-bit and 6-bit transducers is also an issue of interest. The dynamic range of each is set fundamentally by the number of amplitude representation levels in the linear A/D quantization, ±8 levels for the 4-bit transducer, and ±32 levels for the 6-bit transducer. The measured linearity of the acoustic output for a decoded 1500-Hz signal is given for each transducer in Fig. 16. The threshold level for each bit of quantization and the full-scale (clip) level for the digitization are indicated in the figure. One sees that the dynamic range of signal magnitude is approximately 18 dB for the 4-bit transducer and approximately 30 dB for the 6-bit receiver, consistent with the available representation levels. The ambient noise level for the acoustic output measurement in the open laboratory room is also indicated on the graphs.

As a further assessment of the digital receivers, we digitized continuous speech from a tape playback and supplied the digital drive to the 4-bit and 6-bit transducers. The input speech was bandlimited by the same 300 to 3000 Hz filter that produced the noise-band signal of Fig. 11. One can, of course, listen to the decoded output speech simply by putting the receiver to one's ear. A better quantitative display (and one that correlates well with auditory perception) is the sound spectrograph. This display plots time along the abscissa, frequency along the ordinate, and intensity in terms of grey-scale density.

It is informative to compare the acoustic output of the digital transducers to the electrical signal that is decoded by a conventional D/A converter and an electrical desampling filter. Acoustic and electrical decodings are similarly limited in accuracy by the number of bits used for the quantization. In the case of the acoustic decoding, the frequency response of the acoustic system is not ideally flat in-band, with spectral emphasis especially near 1000 Hz, and is limited in cutoff to something less than 3000 Hz (see, for example, Fig. 14). Additionally, the acoustic decoding, as measured here, includes noticeable ambient acoustic noise from the laboratory room. These differences are in fact reflected in the spectrographic comparisons of Fig. 17. The speech utterance is "... source of digital pulses." Even considering these shortcomings and the primitive forms of the experimental devices, the

performance of the acoustic decoding compares relatively favorably with the conventional electrical D/A conversion. This is particularly true for the 6-bit receiver. The spectrographs suggest that the 4-bit quantization has noticeably more in-band quantizing components than the 6-bit quantization, and that significant acoustic room noise "peeks

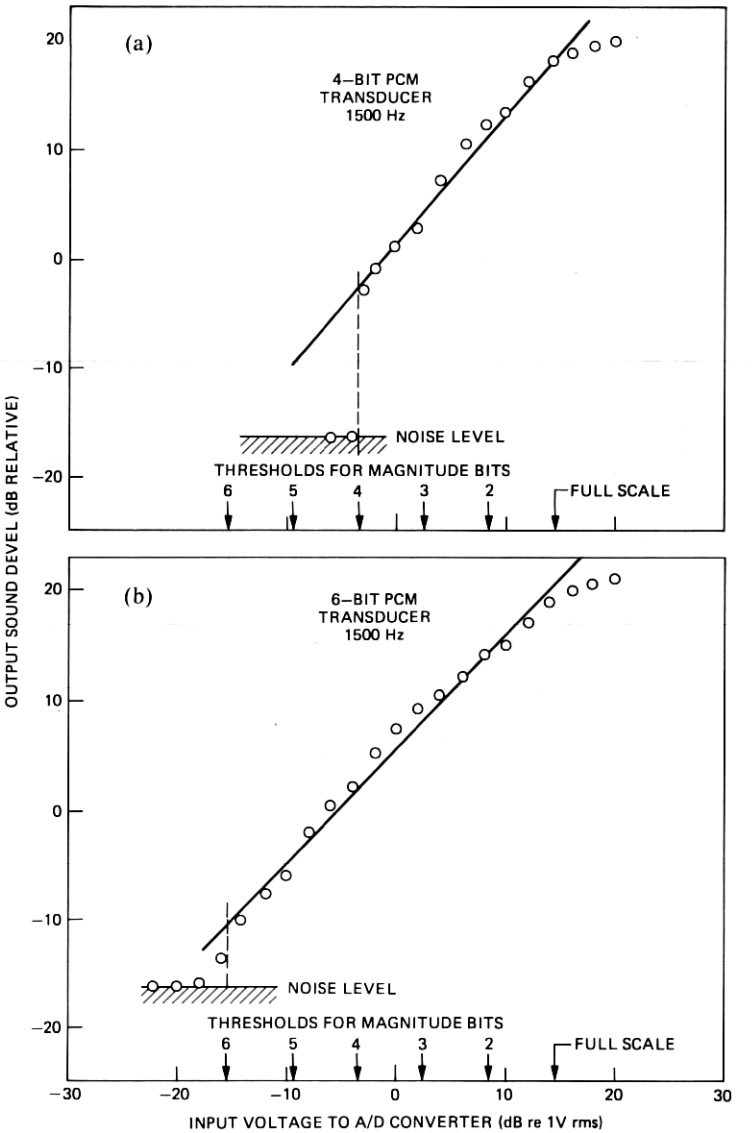

Fig. 16—Dynamic range of signal magnitude for the digital receivers. (a) 4-bit. (b) 6-bit. Threshold levels for each magnitude bit are indicated along the x-axis.

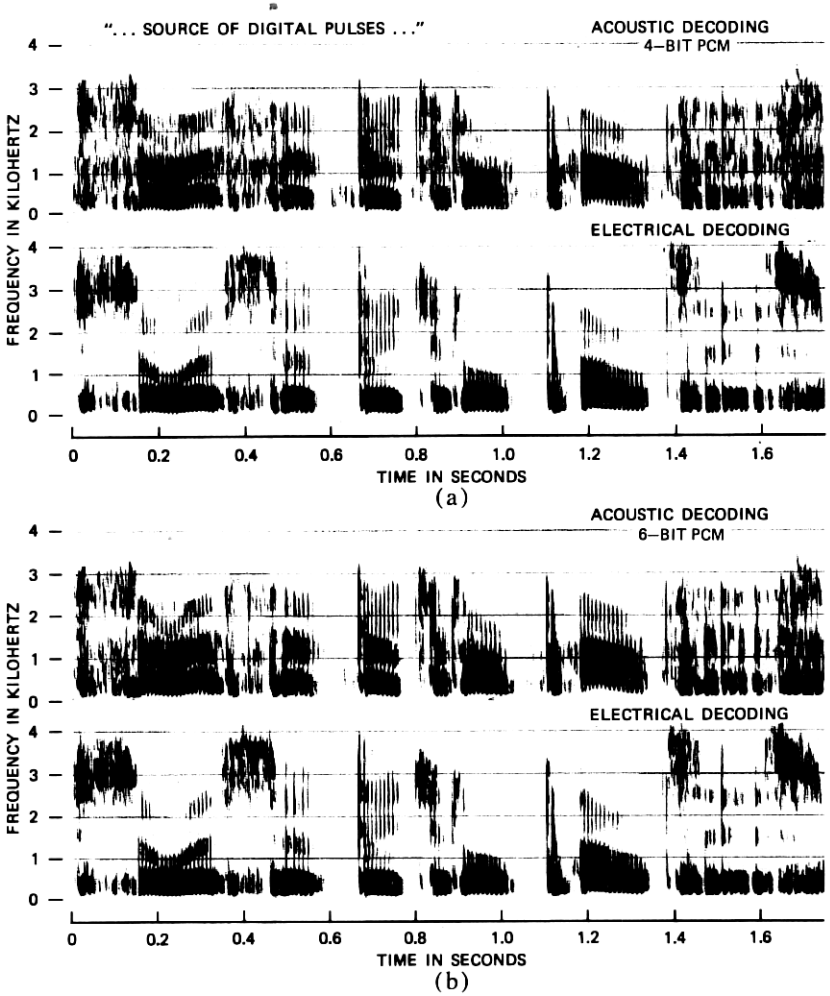

Fig. 17—Sound spectrograms comparing acoustic and electrical decoding of PCM digital speech. (a) 4-bit. (b) 6-bit.

through" the quiet speech gaps in the acoustic decoding. Informal listening assessment and comparisons of the electrical and acoustic decodings also suggest favorable performance of the digital receivers.

VIII. CONCLUSION

This report has summarized an initial examination of the possibility for direct digital-to-sound conversion of acoustic signals. It has described the design, fabrication, and measurement of experimental digital receivers for 4-, 5-, and 6-bit linear PCM signals. While relatively primitive in their first forms, the experimental transducers suggest

that direct conversion of digital signals can be accomplished in the acoustic receiver itself. For transmission systems that utilize digital format out to the station equipment, this possibility may be useful and could help minimize cost and complexity in per-line equipment. The best quality digital receiver discussed here utilizes 6-bit linear PCM. This falls short of good telephone quality, but receivers for seven or eight bits seem possible to fabricate. Accuracies up to 12 to 13 linear bits do not presently seem feasible. However, companded and adaptive quantizations can provide good quality signals with only 4 to 6 bits. While the present results reflect only a first study, they nevertheless suggest reasonable promise for the notion of direct conversion. Much refinement and additional understanding are needed, but both seem attainable through reasonable effort.

The experimental receivers described here utilized beam-charged electret foils as the transducing elements. This fabrication is convenient, inexpensive, and relatively well understood. It is, however, somewhat limited in the output sound levels that can be attained. Other transducer materials might usefully be considered, especially piezoelectric materials. In this direction, we have superficially examined PZT and PVDF, which we have fabricated with segmented electrodes. These elements also seem to hold promise for digital receivers, but this, of course, can only be established in further work.

Our discussion here has focused on direct acoustic decoding of PCM signals. One notices, however, that the combination of transducer and desampling filter is applicable to other signals. A single transducer element, used in a linear analog form with the acoustic desampling filter, can recover the baseband signal from a time-discrete PAM excitation. Also, a single-element digital transducer (capable of representing only two displacement amplitudes, $\pm E$) used in combination with the acoustic desampling filter can accept directly the bit stream of a linear delta modulated signal and provide an acoustically decoded output. Detailed measurements, not reported here, substantiate both these applications.

One reason for interest in digital receivers of few bits, especially 4 bits, is not just the simplicity of construction, but rather the attraction that digital codes such as adaptive differential PCM (ADPCM)[3] have for speech transmission. Speech coded by ADPCM at 32 k bits/s (8 kHz × 4 bits) and at 24 k bits/s (6 kHz × 4 bits) retains usefully high quality and is attractive for voice storage systems. The encoder for ADPCM is typically a linear quantizer whose step-size is changed according to a prescribed algorithm applied to recently past code words. That is, the quantizer step-size at any moment is conditioned by the amplitude statistics of the signal. The decoder, therefore, must be a multiplying D/A converter, where the decoding step-size is controlled by the

received bit stream (according to the same algorithm as that used at the transmitter). The adaptively decoded signal is then delivered to a low-pass (predictor) filter for desampling. These features can be obtained directly in the digital acoustic receiver by letting the sign-polled bit voltage become kE, where k is the step-size multiplier.

The present study has only addressed digital receiving transducers, that is, direct D/A conversion. The complementary issue, of course, is direct digital conversion at the transmitting microphone. Though we have speculated about designs for this elsewhere,[4] the digitally converting microphone must remain a story for a later time. It clearly seems a more difficult problem than the receiver, but not an insurmountable one. One possibility for the A/D converting microphone might be the use of semiconductor gates fabricated onto piezoelectric substrates. The substrates might be cantilever (bender) bimorphs or diaphragms. Also, optical links to actuate the semiconductor sensors of the digital microphone seem attractive possibilities, considering the particle displacements encountered at typical sound levels.

One final point about the digital receiver is perhaps worth while. We have particularized our study to an earphone transducer suitable in size for use as a telephone receiver. There appears to be considerable interest in the possibility of the digital technique for fabrication of loudspeakers. The high-fidelity and recorded music technology has long striven to achieve high-quality sound reproduction at high sound levels. Air suspension of analog speakers and assortments of tweeters and woofers with electrical crossover networks are representative of these efforts. The fundamental problem of course, is that it is difficult to get large excursions of loudspeaker diaphragms to produce high-analog sound levels without nonlinear distortions of the mechanical system. For the digital transducer, on the other hand, each acoustic radiator is either at rest or displaced by a two-valued amount, $\pm X$. As the dimensions of the loudspeaker device become large, one must, of course, consider short wavelength effects, and these issues have to be examined in detail. Preliminary considerations, however, suggest that modest-sized loudspeakers such as those used in speakerphones might be successfully fabricated as directly converting digital transducers.

IX. ACKNOWLEDGMENTS

I am indebted to a number of people for assistance in this study. I wish to thank Robert Kubli of the Acoustics Research Department for his ingenuity in vacuum-plating the multielectrode back plates and in assembling the electret transducers. I wish also to thank him for his aid in mechanical layout and construction of the acoustic desampling filters. Similarly, I wish to thank James Johnston for his aid in

assembling the A/D converter and the bit-parallel, sign-magnitude PCM drive circuit. Further, I am grateful to Trig Meeker and John Hokanson of Bell Laboratories, Allentown, for providing me with experimental PZT devices with multiple electrodes and to George Scalco of the Plastics Chemistry Research Department for providing me with experimental PVDF film metalized with multiple electrodes.

REFERENCES

1. G. M. Sessler and J. E. West, "Applications," Ch. 7 in *Electrets*, G. M. Sessler, Editor, New York: Springer-Verlag, 1980.
2. J. L. Flanagan, "Acoustic Filters to Aid Digital Voice," B.S.T.J. *58*, No. 4 (April 1979), pp. 903–944.
3. P. Cummiskey, N. S. Jayant, and J. L. Flanagan, "Adaptive Quantization in Differential PCM Coding of Speech," B.S.T.J., *52*, No. 7 (September 1973), pp. 1105–1118.
4. J. L. Flanagan, "Direct Digital Conversion in Acoustic Transducers," J. Acoust. Soc. Am., Suppl. 1, *66* (November 1979), p. S54.