Acoustic characteristics of an electrodynamic planar digital loudspeaker using noise shaping technology

Atsushi Hayama and Kenji Furihata[a)]
Department of Electrical and Electronic Engineering, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan

David K. Asano
Department of Information Engineering, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan

Takesaburo Yanagisawa
Department of Electrical and Electronic Engineering, Faculty of Engineering, Shinshu University, 4-17-1 Wakasato, Nagano 380-8553, Japan

(Received 7 July 2004; revised 26 January 2005; accepted 3 February 2005)

The present study extends our previous work [Furihata *et al.*, J. Acoust. Soc. Am. **114**, 174–184 (2003)] by investigating our electrodynamic planar loudspeaker when driven by a 12 bit digital signal with noise shaping. Changing the structure of the loudspeaker can lead to improvement, but in this paper improvements that can be made using signal processing are investigated. Results show that the digital loudspeaker demonstrated good linearity over its 84 dB dynamic range from 40 Hz to 10 kHz. This shows that a 12 bit digital loudspeaker which is equivalent to a 16 bit one is possible. [DOI: 10.1121/1.1887025]

PACS numbers: 43.38.Dv, 43.38.Ja [AJZ] Pages: 3636–3644

I. INTRODUCTION

In a recent study,[1] an electrodynamic planar digital loudspeaker was presented. In order to achieve an adequate sound pressure level and bit resolution, the loudspeaker was made with a diaphragm molded in plastic, which allows the creation of streamlined sections in order to suppress divided vibrations. Two voice coils were rolled together in one section of the diaphragm, resulting in a total of 22 voice coils. Also, 11 permanent magnets were arranged under the diaphragm so that the poles of adjacent magnets were opposite to each other. The suspension between the diaphragm and the frame was made of a piece of handmade Japanese paper.

The acoustic responses, such as frequency and distortion, were affected by the selection of the voice coils. Therefore, in order to improve the responses, (a) the combination and (b) the arrangement of the voice coils were discussed. As for (a), pairs of voice coils were chosen by analyzing the results of each acoustic response when the loudspeaker was driven by an analog signal. With regards to (b), each bit was assigned to a voice coil based on the result of the best combination in (a) resulting in an asymmetric arrangement. This asymmetric arrangement was designed to obtain as flat a frequency response to an analog signal as possible. This asymmetric arrangement was compared with a symmetric one and the results showed that the flatness of the frequency responses around 1 and 4 kHz were improved and that the sound reproduction band was from 40 Hz to 10 kHz.

In order to compare the digital loudspeaker with the above-assigned asymmetric arrangement to one with a symmetric arrangement, experiments were performed in which the loudspeakers were driven by a weighted discrete voltage with a maximum amplitude of 16 $V_{p\text{-}p}$ and a resolution of 16 bits, i.e., the same number of bits used in CD (compact disc) and DAT (digital audio tape) recording. In these experiments, the output wave form, frequency response, linearity, total harmonic distortion, distribution of sound intensity level, and directional pattern of both arrangements were evaluated. The results showed that the digital loudspeaker can reproduce pure sound with a distortion factor less of than about 5% from 40 Hz to 10 kHz, exceeding this value only in a narrow frequency band near 4 kHz. This digital loudspeaker was demonstrated to have good linearity over its dynamic range of 84 dB.

From contour maps of the sound intensity level, it was found that the diaphragm movement changes from piston motion at 3.15 kHz to divided vibration mode motion at 4 kHz. In particular, the sound intensity is highest, 86.4 dB, on the lower right surface, and lowest, 80.3 dB, on the upper left surface.

In this paper, in order to improve the acoustic responses further, all voice coils are considered when selecting an arrangement and combinations. However, the number of bits has to be reduced to 12, or the dynamic range will deteriorate because pulse code modulation cannot resolve details smaller than the least-significant bit. Hence, some interesting ideas have been proposed to try to maximize the human-auditory potential. One idea is noise-shaping. Noise-shaping was first proposed by Michael Gerzon and Peter Craven in 1989,[2] and successfully embodied in Meridian's 618,[3] 518,[4] and also in Sony's Super Bit Mapping.[5] Therefore, noise shaping technology is applied to solve this problem.

[a)]Electronic mail: kennfur@gipwc.shinshu-u.ac.jp

 0001-4966/2005/117(6)/3636/9/$22.50

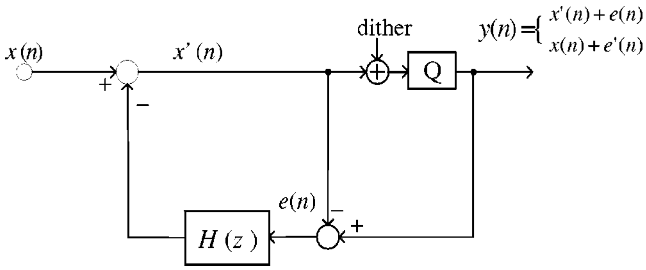

FIG. 1. Block diagram of the noise shaping filter with dithered requantization and error feedback filter $H(z)$. The requantization error is $e'(n)$.

II. NOISE SHAPING

A. Principle (Ref. 6)

Signal requantization is applied in digital audio systems whenever the word-length of audio samples needs to be reduced. This is the case for instance when an audio signal has to be stored on a CD. In entertainment audio, requantization to 8 or 12 bits could be an economically interesting alternative to other forms of data compression because requantized data can be sent directly to a D-A (digital-to-analog) converter, while encoded data needs to be decoded first. Since decoding can be computationally intensive, this could be an important advantage for using requantized data in our digital loudspeaker.

While a white noise dither signal can already improve the quality of low level requantized signals, noise shaping can additionally be applied in order to make the requantization error minimally audible.[7] Figure 1 shows the general scheme for signal requantization with noise shaping. In this scheme, Q represents the quantizer and $H(z)$ is the error feedback filter. Due to the requantization error $e(n)$, the output $y(n)$ differs from $x'(n)$ and from $x(n)$. The error feedback filter has to be controlled such that the difference between $y(n)$ and $x(n)$ becomes minimally audible.

With signals defined as shown in Fig. 1, and using z transforms, we have

$$X'(z)=X(z)-H(z)E(z), \tag{1}$$

$$X'(z)=X(z)+E'(z)-E(z), \tag{2}$$

where $E(z)$ represents quantizer Q's error signal and $E'(z)$ is the additive quantization distortion at the output of the noise shaping requantizer. Subtracting Eq. (2) from Eq. (1), we find that

$$E'(z)=\{1-H(z)\}E(z). \tag{3}$$

The requantization error $e'(n)$ therefore has a power spectrum given by

$$P_{E'}(e^{j\omega})=\|1-H(e^{j\omega})\|^2 P_E(e^{j\omega}). \tag{4}$$

As we can see, the quantizer's error spectrum gets shaped by a noise shaping filter that depends on the error-feedback filter $H(z)$.

This result was obtained without reference to the quantizer Q. Therefore, it applies to any type of quantizer, whether it is dithered, linear, nonlinear, uses rounding, or uses truncating.

In order to achieve minimal audibility of the requantization error, $H(e^{j\omega})$ can be designed to minimize the total amount of perceptually weighted noise power N_ω:

$$N_\omega=\int_{-\pi}^{+\pi} P_{E'}(e^{j\omega})W(\omega)d\omega, \tag{5}$$

where $W(\omega)$ is a perceptual weighting function that approximates the relative audibility of noise power at the different frequencies.

Super Bit Mapping (SBM)[5] introduces a clever trick to design a minimum phase FIR (finite impulse response) noise shaping filter with a given power spectral shape. Note that in order to avoid delayless loops in Fig. 1, it is required that the FIR noise shaping filter can be written as

$$1-H(z)=\sum_{n=0}^{M} a(n)z^{-n}$$

where

$$a(0)=1. \tag{6}$$

It was observed in Ref. 5 that an Mth-order inverse LPC (linear predictive coding) filter is minimum phase (guaranteed if obtained from the autocorrelation formulation[8]) and satisfies Eq. (6). Thus, the required minimum phase FIR noise shaping filter can be obtained by approximating the inverse of the desired noise shaping spectrum with a LPC synthesis filter and inverting the result.

In SBM, the desired noise shaping spectrum is taken to be the hearing threshold when no audio is present. Although SBM can be successfully applied to make the quantization error minimally audible in the absence of audio, spectral shaping that minimizes the audibility of the requantization error in the presence of the actual audio is preferable.

B. Noise shaping filter designed procedure

The noise shaping filter was designed as follows.

(1) In order to take into account all voice coils when selecting the arrangement and the combinations, the number of bits used in the dithered quantizer (Q) shown in Fig. 1 was set to 12.

(2) The noise shaper is unstable if the noise transfer function gain in the stop band is too large. Therefore, a software version of the noise shaping requantizer (Fig. 1) was implemented. A 48 kHz sampling frequency and 12th order FIR designs for $H(z)$ were used to avoid oscillation.[5]

(3) The filter coefficients $a(n)$ of Eq. (6) were chosen experimentally to make the quantization error minimally audible in the absence of audio with the aid of MATLAB. The coefficients obtained were: $a(1)=-0.996$, $a(2)=0.196$, $a(3)=0.112$, $a(4)=0.144$, $a(5)=-0.396$, $a(6)=-0.020$, $a(7)=0.208$, $a(8)=-0.032$, $a(9)=-0.140$, $a(10)=-0.048$, $a(11)=0.152$, and $a(12)=-0.100$.

(4) Figure 2 shows the weighted noise power N_ω of Eq. (5) (0 dB reference at the noise spectral level of no noise shaping). From this figure, it can be seen that although

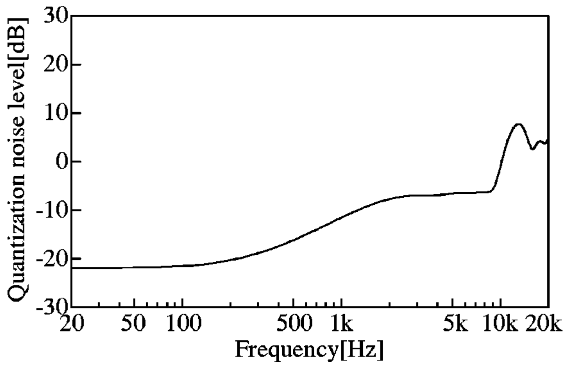

FIG. 2. Frequency response of the noise shaper transfer function.

the curve was different from our hearing threshold or equi-loudness curve, it features a reduced amount of low-frequency noise below 8.75 kHz (−6 dB), but high-frequency noise in the audible range is significantly increased. The most audible noise is concentrated at 12 kHz (6.8 dB) around the second spectral peak of the human ear's sensitivity, but the loudspeaker acts as a low-pass filter with a cutoff frequency of around 10 kHz and a slope of about −18 dB/octave.

C. Noise shaping experiment

In our experimental setup, a software version (C++ programming language) of the noise shaping requantizer was implemented with a computer (NEC: LW500J/2).

Figure 3 shows the frequency response and linearity characteristics of the noise shaper to each digital signal. Comparing Fig. 3(b) with Fig. 3(c), although both inputs are 12 bit digital signals, the results show that the quantization noise is reduced at frequencies less than 10 kHz by the noise shaping filter and that the linearity characteristic is improved to a point where it is almost equal to the characteristics for the 16 bit digital signal shown in (a). Therefore, we can see that the noise shaper acts correctly.

From the results shown in Fig. 3(b), the sound reproduction band of the loudspeaker was chosen to be from 40 Hz to 10 kHz. It is assumed that this characteristic does not significantly change even when noise shaping is applied. The quantization noise will be reduced at frequencies less than 10 kHz compared with a typical 12 bit digital signal.

III. ELECTRODYNAMIC PLANAR DIGITAL LOUDSPEAKER

A. Equivalent electric circuit model

An electrodynamic planar digital loudspeaker[1] driven by a signed magnitude binary (SMB) signal is effective. SMB is the simplest and one of the most obvious methods of encoding positive and negative numbers. The most significant bit (MSB:b_{12}) is assigned to be the sign bit. If the sign bit is 0, this means the number is positive. If the sign bit is 1, then the number is negative. The remaining 11 bits (b_{11}, b_{10}, b_9, ..., b_2, b_1) are used to represent the magnitude of the binary number in the unsigned binary notation. Therefore, the formula to convert a 12 bit input signal with a length of 12 bits is

$$V_p(t) \propto (1-2b_{12}) E_0 \sum_{i=1}^{11} 2^{(i-11)} b_i(t), \tag{7}$$

where V_p is the total velocity amplitude of the panel surface, and E_0 is 16.0 V. As the diaphragm operates in the mass-controlled region, the force applied to it should, ideally, be able to operate with both polarities to reproduce the original sound pressure field correctly.

The actual active part of the diaphragm is 13 cm wide, 13 cm high, and 1.1 cm deep. The diaphragm is molded plastic (total mass $m_p = 0.0344$ kg, additional mass $m_{\mathrm{ad}} = 0.000\,88$ kg) and has streamlined sections to suppress the divided vibrations. A double-voice-coil is rolled together in one section of the diaphragm, resulting in a total of 22 voice coils, each of which has a dc impedance of 4 Ω, 22 turns, and $l = 4.142$ m. Also, 11 permanent magnets (each magnet is 7.5 mm wide, 87.25 mm long, and 6.0 mm thick, and each gap is 2.0 mm) are arranged under the diaphragm so that the poles of adjacent magnets are opposite each other. The suspension between the diaphragm and the frame is made of a piece of handmade Japanese paper (0.09 mm thick). The lowest resonant frequency f_0 was measured and found to be 22.7 Hz. Therefore, the stiffness s_p of the suspension is 736 N/m.

The free moving diaphragm generates sound in both the front and back of the element. The normal use of the element is to place it in an enclosure. In our test of the digital loudspeaker we used an enclosure ($V_c = 0.04194\ \mathrm{m}^3$) which was 28.5 cm wide, 54.5 cm high, and 27.0 cm deep. Therefore, the equivalent stiffness $s_c = \rho c^2 (ab)^2 / V_c$ is 549 N/m (20 °C).

The frequency responses of each voice coil show that at 31.0 Hz a small peak appears due to the closed enclosure and the lowest resonant frequency ($f_0 = 22.7$ Hz). From 30 to about 500 Hz, the diaphragm moves as a whole by acting as a piston. Such a loudspeaker is mass-controlled over most of its pass-band. However, when it comes to higher frequencies, the radiating area exhibits more or less strong partial vibrations. The phase does not deviate more than ±35° in the region where the diaphragm acts as a piston,[1] but there are large phase differences, especially at frequencies larger than 500 Hz. Consequently, timing jitter and distortion may result. Therefore, our electro-dynamic planar loudspeaker has some problems that may cause large distortion when driven by a digital signal because each voice coil's response is not the same. Therefore, in order to accomplish D-A conversion at the loudspeaker, it is necessary to flatten each response. One solution is to connect in series a voice coil with a peak in its frequency response and a voice coil with a dip. We found that the acoustic responses can be improved by an appropriate combination of voice coils. Because the flatness of the frequency response from 40 Hz to 10 kHz can be improved only with the asymmetric arrangement of 11 pairs of voice coils, a 12 bit digital signal is best suited to drive the loudspeaker. We call this "the asymmetric arrangement for a 12 bit digital signal" as follows: (b_3,b_1), (b_7,b_6), (b_9,b_8), (b_5,b_4), (b_{11},b_{10}), (b_3,b_2), (b_9,b_8), (b_7,b_6), (b_{11},b_{10}), (b_5,b_4), and (b_2,b_1).

The electrical impedance will include the effect of the

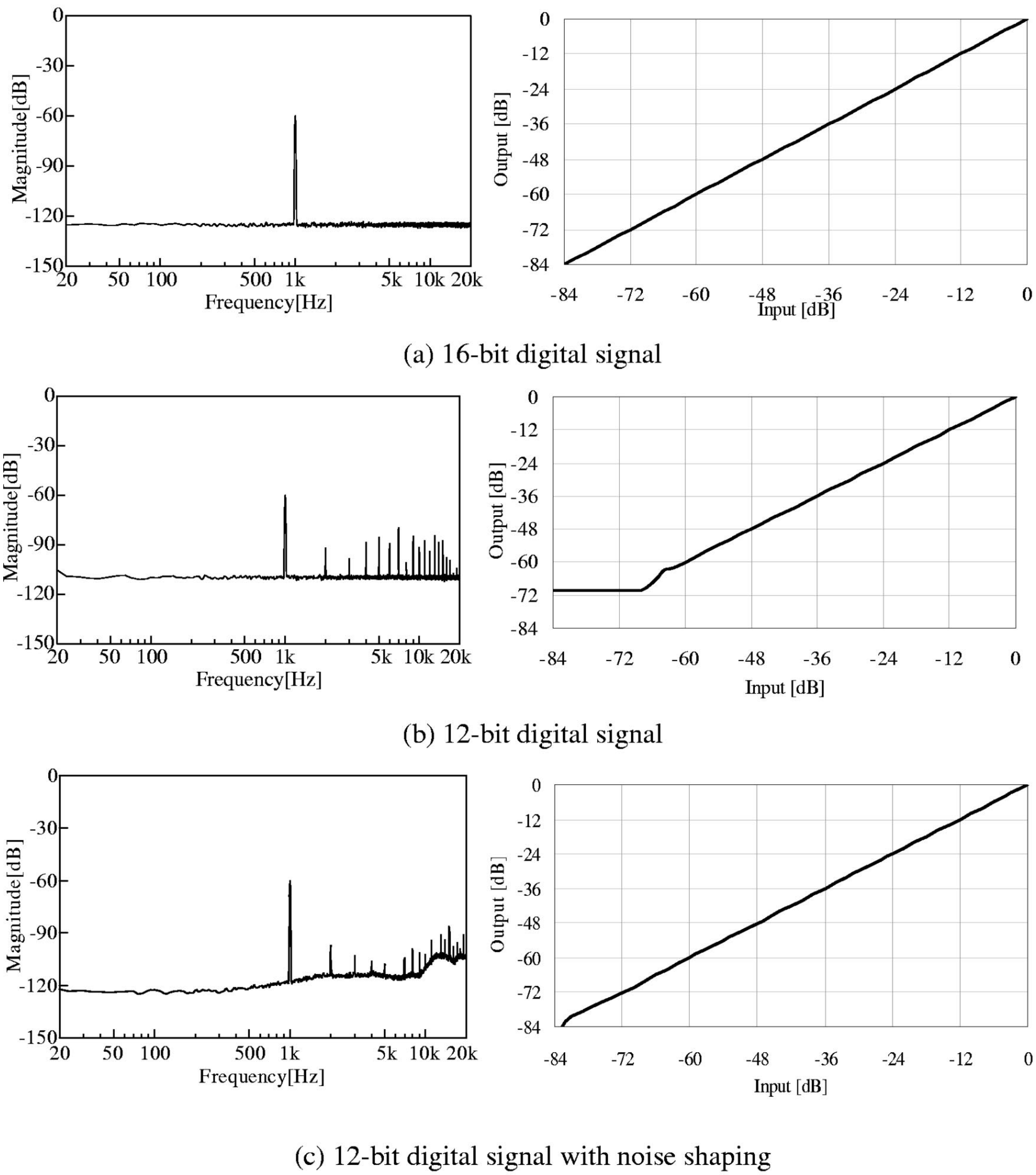

FIG. 3. The frequency response and linearity characteristics of the noise shaper when the following digital signals are input: (a) 16 bit digital signal, (b) 12 bit digital signal, and (c) 12 bit digital signal with noise shaping. The figures on the left are the frequency responses to a 1 kHz pure tone (-60 dB), while the figures on the right are the linearity characteristics at 1 kHz.

mutual coupling between each double-voice-coil: (b_3,b_1), (b_7,b_6), (b_9,b_8), (b_5,b_4), (b_{11},b_{10}), (b_3,b_2), (b_9,b_8), (b_7,b_6), (b_{11},b_{10}), (b_5,b_4), and (b_2,b_1), which are rolled together in one section of the diaphragm. Figure 4 shows a lumped-parameter equivalent circuit model[9] for a multidriver transducer consisting of a panel, a suspension, a frame and parallel exciters with magnets, and eleven double-voice-coils. In this model, R (Ω) and L (H) are the dc resistance and inductance, respectively, of each coil. M (H) represents the mutual inductance between each double-voice-coil. B (Wb/m^2) is the magnetic flux density, l (m) is the length of the ith voice coil, and Bl (N/A) is the force factor.

Consideration of each coil driven from an alternating current source of amplitude I_0 and frequency ω, and with an output impedance r_o, leads to the following series of equations:

$$r_{o3}I_{03}=(R+r_{o3}+j\omega L)I_3+j\omega MI_1+\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,$$

$$r_{o1}I_{01}=(R+r_{o1}+j\omega L)I_1+j\omega MI_3+\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,$$

$$r_{o7}I_{07}=(R+r_{o7}+j\omega L)I_7+j\omega MI_6+\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,$$

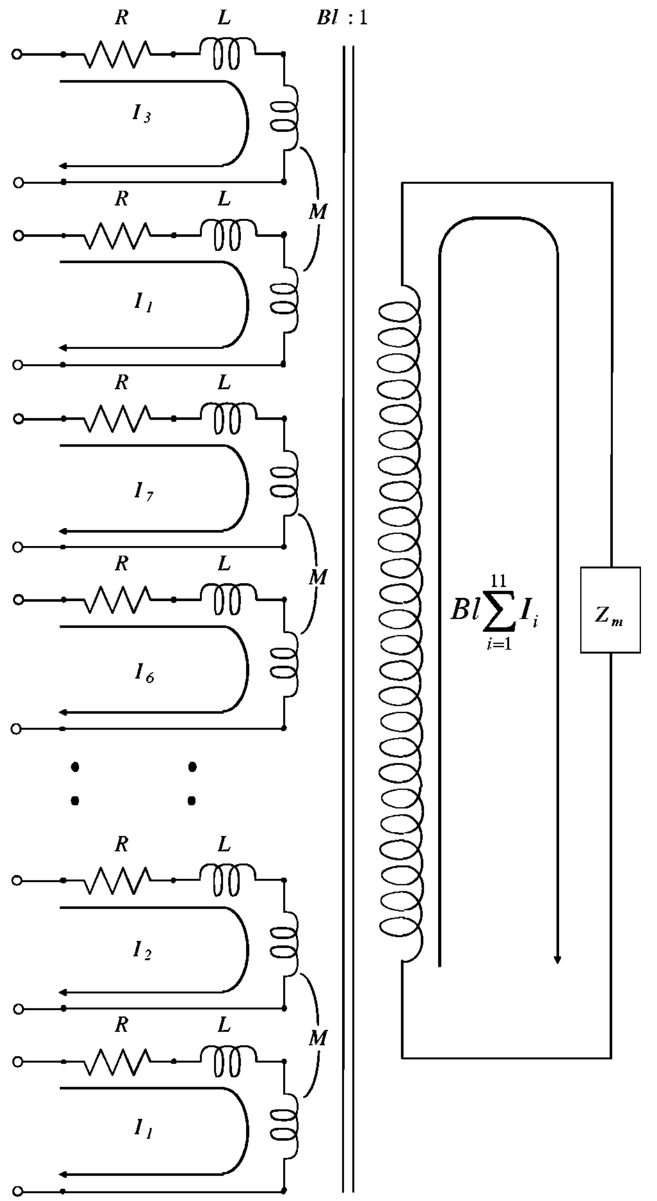

FIG. 4. Equivalent electric circuit of our electrodynamic planar digital loudspeaker.

$$
\begin{aligned}
r_{o6}I_{06} &= (R+r_{o6}+j\omega L)I_6+j\omega M I_7+\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
&\vdots\\
&\vdots\\
r_{o2}I_{02} &= (R+r_{o2}+j\omega L)I_2+j\omega M I_1+\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
r_{o1}I_{01} &= (R+r_{o1}+j\omega L)I_1+j\omega M I_2+\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,
\end{aligned}
\tag{8}
$$

where I_i is the current in the ith coil, and Z_m is the mechanical impedance of the motion system. The asymmetric arrangement for a 12 bit digital signal is connected in series. The voltage at the input terminals of each voice coil may be expressed as

$$
\begin{aligned}
(1-2b_{12})\times 16.0\times b_{11} &= 2(R+j\omega L)I_{11}+2(j\omega M)I_{10}+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-1}b_{10} &= 2(R+j\omega L)I_{10}+2(j\omega M)I_{11}+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-2}b_9 &= 2(R+j\omega L)I_9+2(j\omega M)I_8+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-3}b_8 &= 2(R+j\omega L)I_8+2(j\omega M)I_9+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-4}b_7 &= 2(R+j\omega L)I_7+2(j\omega M)I_6+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-5}b_6 &= 2(R+j\omega L)I_6+2(j\omega M)I_7+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-6}b_5 &= 2(R+j\omega L)I_5+2(j\omega M)I_4+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-7}b_4 &= 2(R+j\omega L)I_4+2(j\omega M)I_5+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-8}b_3 &= 2(R+j\omega L)I_3+j\omega M I_1+j\omega M I_2+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-9}b_2 &= 2(R+j\omega L)I_2+j\omega M I_1+j\omega M I_3+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,\\
(1-2b_{12})\times 16.0\times 2^{-10}b_1 &= 2(R+j\omega L)I_1+j\omega M I_2+j\omega M I_3+2\frac{(Bl)^2}{Z_m}\sum_{i=1}^{11} I_i,
\end{aligned}
\tag{9}
$$

where each part of Eq. (9) contains two voltage generators, V_m and V_z. $V_m=2(j\omega M)I_i$ represents the induced electromotive force (emf) in any given coil due to the change in current in the pair coil. Likewise $V_z=2(Bl)^2\Sigma_{i=1}^{11}I_i/Z_m$ represents the induced emf in any given coil due to the motion of the diaphragm, which is itself dependent on the total driving force on the motion system and hence on the summation of all the currents in all the voice coils.

Unfortunately the total velocity amplitude of the surface of the panel driven by 11 exciters introduces considerable complexity for two voltage generators.

The advantages of using our electrodynamic planar digital loudspeaker instead of a multiple-voice-coil digital loudspeaker[9] or conventional panel loudspeakers[10] are as follows.

(a) In a multiple-voice-coil digital loudspeaker, the motion system and the radiation impedance are the same as for a conventional loudspeaker. Since the force on the diaphragm needs to be consistent for each voice coil (when they are equally energized), they must be wound in close proximity. The electrical impedance, therefore, will include the effect of the mutual coupling between each of the voice coils. Furthermore since the current weighting in each coil must be accurately controlled, the use of a high-impedance switchable current source driver will overcome the variations in current produced by a conventional voltage source driver.

(b) A panel loudspeaker primarily consists of a panel and an inertia exciter. The exciter is essentially a voice-coil-driver with the coil attached to the panel. The magnet serves as a proof mass to produce inertial force. In lieu of a rigid diaphragm as used in conventional loudspeakers, flexible panels are employed as the primary sound radiators. Resonance of flexural motion is encouraged so that the panel vibrates as randomly as possible. The sound field produced by this type of distributed mode loudspeaker[10] is very diffuse at high frequencies. Therefore, it is difficult for conventional panel loudspeakers to achieve flatness in the frequency response from 40 Hz to 10 kHz.

B. Experiment

First, the electrical input impedance of each pair of voice coils was measured with a precision LCR meter (Agilent: 4284A, an input voltage is 100 mV). Second, when the 11 bit pair of voice coils was inputted with an 8 V_{rms} amplitude signal from 20 Hz to 20 kHz, the induced electromotive force [emf: V_m+V_z (V)] for each pair of voice coils was measured with an ac voltmeter (Kenwood: VT-181E).

C. Results

The performance was measured using the following indices.

1. Input impedance

Figure 5 shows: (a) the measured input impedance $|Z_{\text{in}}|$ (Ω), and (b) the phase (deg) for the 11 and 1 bit pairs of voice coils. We see that (a) small peaks (the 11 and 1 bit pairs of voice coils) at the fundamental resonant frequency of 31.0 Hz, where $|Z_{\text{in}}|=9.05\ \Omega$, (b) both amplitudes are about 8.8 Ω from 20 Hz to 4 kHz and both phases are less than about 10°, and (c) increases due to inductive effects at higher frequencies are from 8.8 to 12.36 Ω.

Table I shows the parameters ($2R$ and $2L$) of Eq. (9) related to the equivalent electric circuit model shown in Fig. 4. From Table I, we see that the parameter ($2R$) of each bit pair of voice coils varies from 8.89 to 10.98 Ω, and $2L$ varies from 0.0657 to 0.0731 mH.

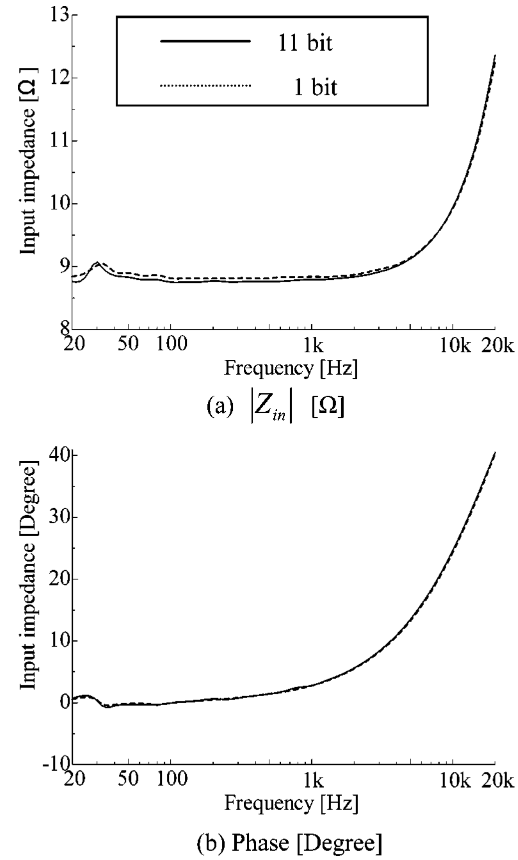

FIG. 5. Measured input impedance (amplitude and phase) of the 11 and 1 bit pairs of voice coils.

2. Induced electromotive force

Figure 6 shows the measured induced electromotive force [emf: V_m+V_z (mV)] for the 10, 8, 6, and 1 bit pairs of voice coils when the 11 bit pair was driven with an 8.0 V_{rms} signal. From the lowest peak voltages of the fundamental resonant frequency, we see that the motional impedance $2\,(Bl)^2/Z_m$ of the 10 bit pair of voice coils is 0.0786 Ω, and $2\,(Bl)^2/Z_m$ of the 1 bit pair is 0.0605 Ω. Therefore, the efficiency (the ratio of the motional impedance to the free impedance of the 10 bit pair of voice coils) in converting electrical power to sound of our electrodynamic planar digital loudspeaker is about 0.87%, because the free impedance is 9.05 Ω. From the emf voltage at 10 kHz for the 10 bit pair of

TABLE I. Measured parameters of our multidriver loudspeaker.

bits	$2R$ (Ω)	$2L$ (mH)	bits	$2R$ (Ω)	$2L$ (mH)
11	9.15	0.0659	5	9.16	0.0678
10	9.02	0.0731	4	9.36	0.0707
9	10.34	0.0677	3	10.98	0.0677
8	9.03	0.0685	2	8.89	0.0657
7	9.26	0.0684	1	9.08	0.0676
6	9.07	0.0673			

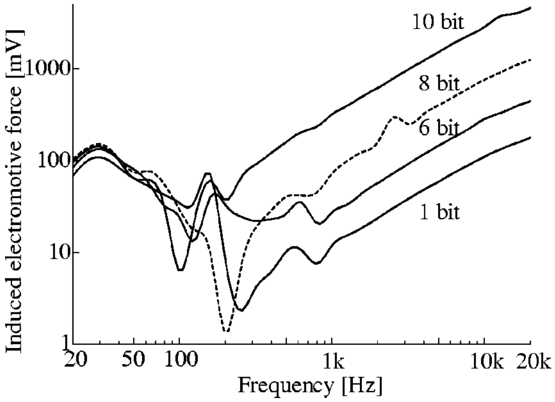

FIG. 6. Measured induced electromotive force of the 10, 8, 6, and 1 bit pairs of voice coils, when the 11 bit pair was driven with 8 V_{rms} signals.

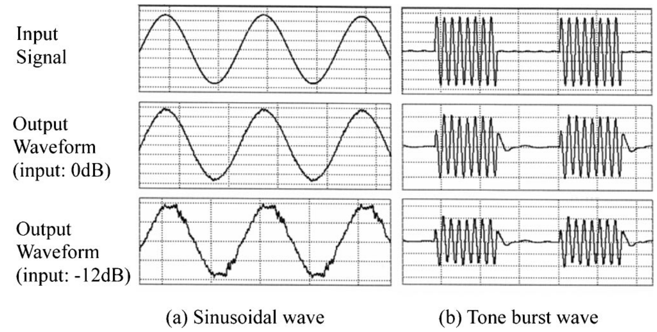

FIG. 7. Output wave forms from our electrodynamic planar digital loudspeaker driven by (a) a pure tone signal (12 bit: 0 dB, and 10 bit: −12 dB) and (b) a tone burst signal (12 bit: 0 dB, and 10 bit: −12 dB) at 100 Hz.

voice coils, the mutual inductance M between each pair of coils (b_{11} and b_{10}) is 0.0287 mH, because the current of the 11 bit pair (I_{11}) is 0.791 A.

The interaction with a double-voice-coil in a multidriver loudspeaker has been modeled in terms of mutual coupling and induced motion emf. Also, our loudspeaker parameters related to the equivalent electric circuit model shown in Fig. 4 were measured: $2R = 9.4\ \Omega$ on average, $2L = 0.068$ mH on average, $M = 0.028$ mH on average, and $2(Bl)^2/Z_m = 0.074\ \Omega$ on average.

IV. RESPONSES TO 12 BIT DIGITAL SIGNALS WITH NOISE SHAPING

A D-A converter of the surface of the panel driven by 11 exciters is used to convert the sampled binary information back in to an analog vibration signal. The conversion is called a zero order hold type where each output sample level is a function of its binary weight value and is held unit the next sample arrives. As a result of the D-A converter step function response it is apparent that large amounts of undesirable high frequency energy are present. To eliminate this, the D-A converter is usually followed by a smoothing filter, having a cutoff frequency no greater than half the sampling frequency. The loudspeaker acts as a low-pass filter with a cutoff frequency of 10 kHz and a slope of about −18 dB/octave. Therefore, the spectrum of the resulting signal is the product of a step function spectrum and the band-limited analog filter spectrum. Furthermore, the radiation produced by the vibration of the surface of the rectangular panel does not have symmetric spherical radiation patterns characteristic of a simple source. However, the radiation produced by the panel can be found by considering the panel to be a group of simple point sources.[1]

A. Experiment

The experiments were carried out in an anechoic room. The on-axis response was measured at a distance of 10 cm from the front of the diaphragm.

The input signal to the loudspeaker was generated by a computer (NEC: LW500J/2), which was conveyed to the PCM driving circuit.[1] The digital audio signal had a sampling frequency of 48 kHz, and was quantized to 12 bits with noise shaping. First, a rectangular tone-burst covering a fairly wide frequency band was used to evaluate the dynamic behavior of the digital speaker. The transient behavior of the loudspeaker was indicated by a change in the envelope of the burst signal. Second, pure tones from 20 Hz to 20 kHz were used every 1/3 octave. The peak level of the input signal was 4096 samples (peak to peak), which corresponds to 0 dB for a 12 bit digital signal. This level was decreased in 6 dB steps and the acoustic responses were measured at each step.

B. Results

The performance was measured using the following indices.

1. Output wave forms

Example output wave forms from our electrodynamic planar digital loudspeaker driven by (a) a pure tone signal (12 bit: 0 dB, and 10 bit: −12 dB) and (b) a tone burst signal (12 bit: 0 dB, and 10 bit: −12 dB) at 100 Hz are shown in Fig. 7. From this figure, we see that our digital loudspeaker can reproduce pure sounds because the timing precision of D-A conversion on the diaphragm is good, and overhanging (ringing after the signal is gone) is present to a small degree.

2. Frequency response

Figure 8 shows the frequency responses to 12 bit digital signals (various peak levels) with noise shaping. We found that the sound output [SPL (dB)] from the asymmetric arrangement for a 12 bit digital signal with the noise shaping to input digital signal levels from 0 to −84 dB in 6 dB steps ranged from about 90 to 0 dB in 6 dB steps. We also found that the sound reproduction band was from 40 Hz to 10 kHz.

3. Linearity

For the linearity characteristics from 40 Hz to 10 kHz of the 12 bit digital signal with noise shaping, the relationship between SPL responses is shown in Fig. 8 and input levels from 0 to −84 dB were analyzed with simple regression analysis. The estimated slope coefficients are from 1.006 to 1.188, the intercept coefficients are from 84.0 to 94.0 dB, and the coefficients of determination are from 0.9972 to 0.9991. At the same time, because the linearity range is 84 dB, it can be said that the D-A conversion performed on the diaphragm is equivalent to 16 bit D-A conversion.

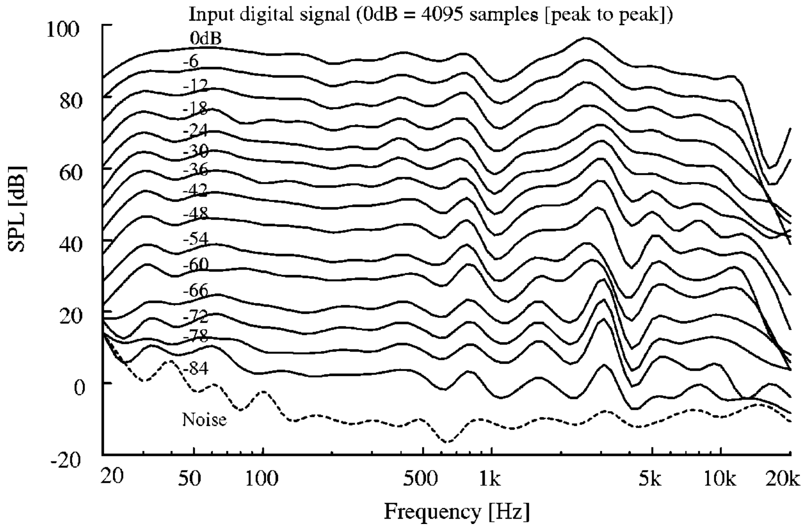

FIG. 8. Frequency response of the digital loudspeaker driven by a 12 bit digital signal with noise shaping at a distance of 10 cm from the center of the diaphragm. The dashed lines show the ambient noise level.

4. Total harmonic distortion

The total harmonic distortion (THD) of the digital loudspeaker using the asymmetric arrangement for a 12 bit digital signal is about 2% at 100 Hz as shown in Fig. 7.

Figure 9 shows the distortion factors when the input signal was changed from 0 to -12 dB in 6 dB steps. The reason why the THD is inferior at lower frequencies is that the voice coils separate from the magnetic circuit. The asymmetric arrangement for the 12 bit digital signal with noise shaping results in an improvement in the characteristics around 4 kHz, which is relatively important for the sense of hearing.

V. CONCLUSIONS

This paper extends our previous work[1] "Acoustic characteristics of an electrodynamic planar digital loudspeaker" by introducing a noise shaping technique into the design to minimize the quantization error resulting from the limited number of bits.

First, the noise shaping filter was designed so that (1) the number of bits used in the dithered quantizer was 12, (2) the sampling frequency was 48 kHz, and (3) the error feedback filter was a 12th order FIR design with coefficients given by: $a(1)=-0.996$, $a(2)=0.196$, $a(3)=0.112$, $a(4)=0.144$, $a(5)=-0.396$, $a(6)=-0.020$, $a(7)=0.208$, $a(8)=-0.032$, $a(9)=-0.140$, $a(10)=-0.048$, $a(11)=0.152$, and $a(12)=-0.100$.

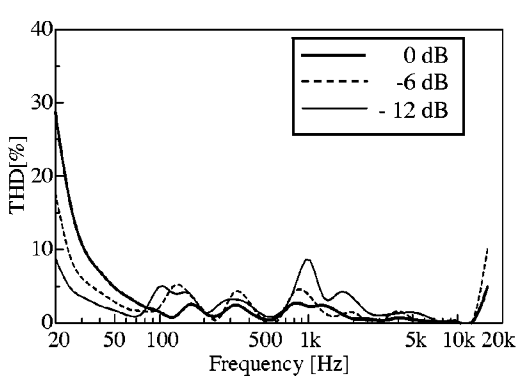

FIG. 9. Total harmonic distortion (THD) vs frequency when the input signal was changed from 0 to -12 dB in 6 dB steps.

Second, the noise shaping filter's interaction with double-voice-coils in our multidriver loudspeaker was modeled in terms of the mutual coupling and the induced motion emf. Our multidriver digital loudspeaker parameters related to the equivalent electric circuit model were measured as follows. (4) $2R=9.4\ \Omega$ on average, $2L=0.068$ mH on average, $M=0.028$ mH on average, and $2(Bl)^2/Z_m=0.074\ \Omega$ on average. (5) The efficiency (the ratio of the motional impedance to the free impedance) was about 0.87%.

Finally, to improve the linear range of our multidriver digital loudspeaker, noise shaping technology was applied to solve the problem that 12 bit PCM cannot resolve details smaller than the LSB. The responses to 12 bit digital signals with noise shaping were measured as follows. (6) Our digital loudspeaker can reproduce pure sounds because the timing precision of D-A conversion on the diaphragm is good, and overhanging (ringing after the signal is gone) is present to a small degree. (7) As for the linearity characteristics from 40 Hz to 10 kHz, the estimated slope coefficients were from 1.006 to 1.188, the intercept coefficients were from 84.0 to 94.0 dB, and the coefficients of determination were from 0.9972 to 0.9991. (8) Because the linearity range was 84 dB, it can be said that the D-A conversion performed on the diaphragm is equivalent to 16 bit D-A conversion.

To summarize, we have constructed and demonstrated a 12 bit digital loudspeaker using noise shaping. This loudspeaker has good performance, a wide bandwidth, and a linear dynamic range of 84 dB.

[1] K. Furihata, A. Hayama, D. K. Asano, and T. Yanagisawa, "Acoustic characteristics of an electrodynamic planar digital loudspeaker," J. Acoust. Soc. Am. **114**, 174–184 (2003).

[2] M. A. Gerzon and P. G. Craven, "Optimal noise shaping and dither of digital signals," Audio Engineering Society 87th Convention, New York, 1989, preprint, p. 2822.

[3] M. A. Gerzon, P. G. Craven, J. R. Stuart, and R. J. Wilson, "Psychoacoustic noise shaped improvements in CD and other linear digital media," Audio Engineering Society 94th Convention, Berlin, 1993, reprint, p. 3501.

[4] J. R. Stuart and R. J. Wilson, "A search for efficient dither for DSP applications," On the possibility and problems of the PCM digital loud-

speaker (in Japanese), Audio Engineering Society 92nd Convention, Vienna, 1992, preprint, p. 3334.
[5] M. Akune, R. M. Heddle, and K. Akagiri, "Super Bit Mapping: Psychoacoustically optimized digital recording," Audio Engineering Society 93rd Convention, San Francisco, 1992, preprint, p. 3371.
[6] W. Verhelst and D. D. Koning, "Least squares theory and design of optimal noise shaping filters," Audio Engineering Society 22nd International Conference on Virtual, Synthetic and Entertainment Audio, Espoo, Finland, 2002, preprint, pp. 216–222.
[7] S. P. Lipshitz, J. Vanderkooy, and R. A. Wannamaker, "Minimally audible noise shaping," J. Audio Eng. Soc. **39**, 836–852 (1991).
[8] J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech* (Springer, New York, 1976).
[9] Y. Huang, S. C. Busbridge, and P. A. Fryer, "Interactions in a multiple-voice-coil digital loudspeaker," J. Audio Eng. Soc. **48**, 545–552 (2000).
[10] M. R. Bal and T. Huang, "Development of panel loudspeaker system: Design, evaluation and enhancement," J. Acoust. Soc. Am. **109**, 2751–2761 (2001).