Minimally Audible Noise Shaping*

STANLEY P. LIPSHITZ, *AES Fellow,* **JOHN VANDERKOOY,** *AES Fellow,* **AND ROBERT A. WANNAMAKER****

Audio Research Group, University of Waterloo, Waterloo, Ont. N2L 3G1, Canada

Normal quantization or requantization noise is white, but the ear's sensitivity to low-level broad-band noise is not uniform with frequency. By adopting a suitable weighting curve to represent low-level noise audibility, one can design dithered requantizing noise shapers to approximate the inverse of the audibility curve and hence achieve the least audible noise penalty. If Fielder's modified E-weighting curve is adopted as a model of the 15-phon audibility curve, a reduction of 10.9 dB in perceived noise is possible by the use of a simple second-order noise shaper. This result is already within 0.6 dB of the theoretical minimum set by information theory, and almost a 2-bit gain in apparent signal-to-noise ratio. Even greater perceived noise reductions are possible if one adopts an audibility weighting curve which more closely approximates the ear's precipitous high-frequency rolloff, and incorporates a higher order filter into the noise shaper's feedback loop. In fact, a 20-dB apparent reduction in the requantizing noise is then possible with filters of modest order, but the penalty is a significant increase in the total noise power. These questions are explored and some of the available options illustrated. Such noise shaping will soon be advantageous in order to preserve on the 16-bit Compact Disc the lower noise floor of an original 18- or 20-bit master recording.

0 INTRODUCTION

The use of a noise-shaping feedback loop around a quantizer (or, in the case of digital signal processing, around a requantizer which performs word-length reduction) dates from at least the early 1960s, the paper by Spang and Schultheiss [1] being one of the earliest references. In fact, Philips used a combination of simple first-order noise shaping and four-times interpolation ("oversampling") in their first line of Compact Disc (CD) players in the early 1980s in order to eke 16-bit noise from a 14-bit digital-to-analog converter [2]. In this paper we explore some of the possibilities for the design of psychoacoustically optimized noise shapers whose noise spectrum is tailored so as to be least audible. If oversampling is allowed, one can easily make drastic reductions in the perceived baseband requantizing noise [3], [4]. In this paper we address the more difficult question of substantially reducing the apparent noise contributed by a requantization *without* raising the sampling rate.

These questions are relevant for a number of reasons. First, professional master recordings will soon be captured using 18- or even 20-bit analog-to-digital converters. Assuming that these new conversion systems achieve close to their theoretical noise floors, the need will soon arise to be able to perform word-length reductions to the 16-bit CD format without an apparent 12- or 24-dB noise increase in the process. To a lesser extent, the same situation arises when any digital manipulations are performed on already quantized signals, for it is necessary to truncate the longer internal word lengths to the desired lower output word length. Second, any such (re)quantization operation must be appropriately dithered in order to avoid quantization distortion and noise modulation [5]–[7]. This necessitates the use of triangular probability density function (pdf) dither at the input to the quantizer, thus tripling the baseband requantization noise power from $\Delta^2/12$ to $\Delta^2/4$, where Δ represents the least significant bit (LSB)

* Presented at the 88th Convention of the Audio Engineering Society, Montreux, Switzerland, 1990 March 13–16; revised 1991 August 28.

** The author are members of the Guelph-Waterloo Program for Graduate Work in Physics.

step size of the requantization. Even though the added dither noise can be spectrally shaped so as to reduce its audibility somewhat [3], [4] by correlating successive dither samples without modifying its pdf, the white roundoff-noise component remains, and the total audible noise penalty is still considered by some to be a significant disadvantage of properly dithered rounding, even while acknowledging its benefits in all other respects.

The theoretical basis for and limits to the optimization of noise shaping were recently furnished by Gerzon and Craven in a very important paper [8]. These results are applied in this paper, and it will be seen to what extent the theoretical noise limit can be achieved in practice with noise-shaping filters of moderate order.

1 NOISE-SHAPING THEORY

The noise-shaping quantizer configuration under discussion is shown in Fig. 1(a). The input digital words are passed on to a requantizer Q in which they are rounded down to the reduced word length required at the output. In order to guarantee the absence of distortion and noise modulation (and limit-cycle oscillations due to the error feedback) at the rounded output, dither words are added to the signal by injection, as indicated immediately prior to Q [3], [4], [8]. As already mentioned, we shall assume that the dither is white and of triangular pdf of width 2Δ, where Δ is the LSB step size of Q. The added dither noise power over the baseband from 0 Hz up to the Nyquist frequency $f_s/2$ is then $\Delta^2/6$, where f_s denotes the sampling frequency. Q itself contributes a white requantization noise power of $\Delta^2/12$, so that in the absence of the feedback filter $H(z)$ the total requantization noise contributed by the process would be $\Delta^2/4$.

However, the requantization error e is extracted by comparing the output with the input to Q, enclosing the dither within this subtraction loop. This error is fed back through the noise-shaping filter $H(z)$ and subtracted from the input as shown.[1] The error-feedback filter $H(z)$ must incorporate at least a 1-sample delay z^{-1}, and so the most general form of recursive filter structure for $H(z)$ is that shown in Fig. 1(b).[2] The feedforward coefficients are $a_0, a_1, a_2, \ldots, a_n$, while the feedback coefficients are denoted by $b_1, b_2, \ldots, b_n$ for an $(n + 1)$-delay filter structure $(n \geq 0)$. The transfer function $H(z)$ is thus

$$H(z) = z^{-1} \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + \cdots}{1 - b_1 z^{-1} - b_2 z^{-2} - b_3 z^{-3} - \cdots}. \tag{1}$$

For a purely nonrecursive filter realization all the b_i are zero (an FIR filter). It is easily seen [8] that the error e appears at the output modified by the equivalent filter $1 - H(z)$, that is, the output error is $[1 - H(z)]e$, while the signal itself passes through unaffected. Under the dither assumptions that we have made, e is a white noise of baseband power $3\Delta^2/12$, and hence the output noise has a shaped power spectrum of spectral density

$$3 \frac{\Delta^2}{12} \cdot \frac{2}{f_s} \left| 1 - H(\mathrm{e}^{-\mathrm{j}2\pi f/f_s}) \right|^2 \quad \mathrm{Hz}^{-1} .$$

It is also convenient to normalize all powers by the basic quantization noise power of $\Delta^2/12$. We shall do this throughout the sequel. In this case the output noise-power density added to the signal by the noise-shaping requantizer of Fig. 1 is $3N(f)$, where now

$$N(f) = \frac{2}{f_s} \left| 1 - H(\mathrm{e}^{-\mathrm{j}2\pi f/f_s}) \right|^2 \quad \mathrm{Hz}^{-1} . \tag{2}$$

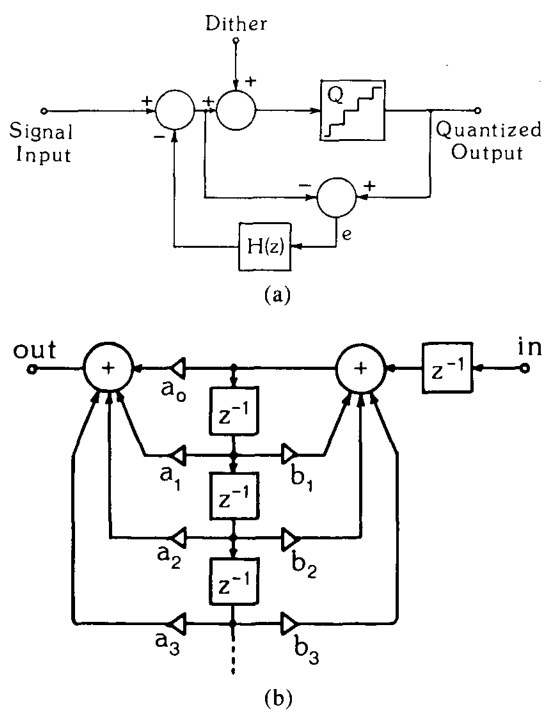


Fig. 1. (a) General noise-shaper configuration. Q—(re)quantizer; $H(z)$—noise-shaping filter, which feeds back (re)quantization error e. Dither is injected as shown. (b) "Direct form 2" realization of most general error-feedback filter $H(z)$ of (a). Only four delays are shown. a_i—feedforward coefficients; b_i—feedback coefficients. Filter is recursive (i.e., IIR) if some of the b_i are nonzero.

[1] The error-feedback loop contains within itself the seeds of a potential instability known as an "overflow instability," which can cause self-sustained full-amplitude oscillations to occur if the quantizer Q is ever driven into clipping. Since this can occur in digital audio systems, a squelch mechanism is essential to prevent such overflow oscillations. This is easily accomplished by ensuring that the input summer in Fig. 1(a) is designed to saturate numerically at the same full-scale values as Q itself. This prevents large error signals e from circulating in the feedback loop. This problem is further discussed and illustrated in [9].

[2] This diagram is intended to be conceptual only. From a practical point of view one would not implement an IIR filter $H(z)$ using this "direct form" structure except in low-order cases, due to internal numerical roundoff and scaling considerations (that is, requantization). A decomposition into cascaded lower order sections would be preferable.

The total (unweighted) baseband power is $3N_u$, where

$$N_u = \int_0^{f_s/2} N(f)\,df\ . \tag{3}$$

In these formulas $N(f)$ and N_u thus represent the normalized powers due solely to the undithered noise shaper, the dithered powers being three times this high. Of course, without the assumption of a properly dithered quantizer Q, the output noise powers and spectra are not represented correctly by these expressions, but it is nevertheless useful to think of $N(f)$ and N_u as the basic quantities from which the actual output powers are derived by appropriate scaling (by a factor of 3). Had no shaping been used [that is, $H(z) \equiv 0$], the output noise power would have been three units of white noise (two from the dither and one from the quantizing operation) spread over the band from 0 to $f_s/2$ Hz. In the general case there is no simple expression for the total power N_u in terms of the filter coefficients a_i and b_i, but for the purely nonrecursive (FIR) case the integral, Eq. (3), evaluates to the following useful expression:

$$N_u = 1 + \sum_{i=0}^{n} a_i^2\ . \tag{4}$$

We now wish to consider how to design the filter $H(z)$ so as to result in the least *audible* output noise spectrum $3N(f)$. This means adopting a suitable psychoacoustic weighting function for low-level broadband noise audibility. Let us denote the power spectral density weighting function of such a noise-weighting filter by $W(f)$. Then as Gerzon and Craven have shown [8], there exists a corresponding noise-shaping filter $1 - H(z)$, and hence an error-feedback filter $H(z)$, which minimizes the total *weighted* output noise power $3N_w$, where

$$N_w = \int_0^{f_s/2} N(f)W(f)\,df\ . \tag{5}$$

This optimal filter $1 - H(z)$ must be minimum phase (that is, have all its zeros and poles within the unit circle in the z plane). Furthermore, it clearly must have a spectrum shape that is the inverse of the weighting function $W(f)$ so that the optimal weighted spectrum becomes perfectly flat. In fact, the optimal filter function $H(z)$ must satisfy the condition [8]

$$\left|1 - H(e^{-j2\pi f/f_s})\right|^2 = \frac{w}{W(f)} \tag{6}$$

where the positive constant w is determined from $W(f)$ according to

$$\log w = \frac{2}{f_s}\int_0^{f_s/2} \log W(f)\,df \tag{7}$$

the logarithms being taken to any desired (but common) base. If base 10 is adopted, then w is expressed in bels. The validity of this theorem requires that the integral on the right-hand side of Eq. (7) be finite, and this will be the case for any rational polynomial weighting function $W(f)$. It should be noted, however, that a weighting function which is zero over any finite portion of the baseband results in an infinitely negative integral, corresponding to $w = 0$. Furthermore, it is assumed that the signal-to-noise ratio remains high over the whole baseband, so that the noise-power density $N(f)$ must be kept moderate at all frequencies if the theory is to apply.

We note that $\log w$ is the average over the baseband of the logarithm of the power weighting function $W(f)$. The significance of w is easily clarified if we substitute Eq. (6) into Eq. (2) and the latter into Eq. (5), for then we obtain for this optimal filter

$$N_w(\text{opt}) = w \tag{8}$$

that is, w represents the greatest possible improvement in the weighted noise power that can be obtained from any noise-shaping quantizer of the form of Fig. 1. The total weighted noise-power output of the optimal properly dithered noise shaper is therefore $3w$, and this represents a limit which cannot be exceeded by any realizable filter $H(z)$. Thus given a weighting function $W(f)$, we can fairly easily compute w and hence know precisely the ultimate weighted power $3N_w$ ever achievable by this configuration. As Gerzon and Craven show, this limit is equivalent to preserving the Shannon information theory capacity of the channel—a most fascinating result. If w is zero (that is, $-\infty$ decibels), the theorem does not apply. For such weighting functions, which go to zero over part of the band, there is no theoretical limit to the weighted noise-power reduction obtainable from the circuit of Fig. 1. This would be the case if, for example, one assumes that the ear has zero sensitivity between, say, 20 kHz and the Nyquist frequency, and chooses the weighting function to reflect this fact.

It is interesting to examine the total *unweighted* noise power $3N_u$ for the optimal noise shaper. Using Eqs. (2), (3), and (6) we find that

$$N_u(\text{opt}) = \frac{2w}{f_s}\int_0^{f_s/2} \frac{df}{W(f)}\ . \tag{9}$$

The integral on the right-hand side of Eq. (9) does not converge for weighting functions $W(f)$ which roll off with slopes of 6 dB per octave or greater as f tends to zero, and this is certainly the case for any reasonable approximation to the ear's true low-frequency noise sensitivity. Hence the theory predicts that, although an optimal noise-shaping feedback filter $H(z)$ generally exists for any rational weighting function $W(f)$, and can achieve a weighted noise power of $3w$, it produces an unacceptable *infinite* unweighted noise power. Of course, this theoretical conclusion applies to the generally infinite-order filter $H(z)$, which would be needed to approximate in the noise-shaper output spectrum the inverse of $W(f)$. For any practical finite-order realization of such an optimal filter, the unweighted noise power will remain finite, but this prediction warns us to expect it to become very large indeed, as the weighted noise power approaches its optimal lower bound of $3w$. We shall see this in the examples to follow. Moreover, if these results are to be applicable, the noise-power density $N(f)$ must be limited so as to preserve a signal-to-noise ratio substantially greater than unity across the band, as required for the Gerzon–Craven results [8] to be valid. Finally, note the ratio of Eqs. (8) and (9):

$$\frac{N_u(\mathrm{opt})}{N_w(\mathrm{opt})} = \frac{2}{f_s}\int_0^{f_s/2} \frac{\mathrm{d}f}{W(f)} \tag{10}$$

that is, for the *optimal* filter, the ratio of the unweighted to the weighted total noise power in the baseband is given by the (generally infinite) average of $1/W(f)$ over the baseband.

2 CHOICE OF WEIGHTING FUNCTION *W(f)*

In our earlier papers [3], [4] we pointed out that standard A-weighting was not a satisfactory weighting function to use to approximate the ear's low-level noise sensitivity. For this purpose we preferred to choose the *modified* E-weighting curve proposed by Fielder [10]. (This curve is specified in terms of its pole–zero locations in both [3] and [4].) Fielder derived his modified E-weighting curve by approximating somewhat better than does the standard E-weighting function the inverse of the 15-phon ISO equal-loudness curve to which corrections were applied for diffuse field versus free field and noise versus sine-wave sensitivity [11], [12]. The modified E-weighting curve is shown in Fig. 2 together with an alternative curve, which we shall call the *improved* E-weighting curve, which does an even better job of approximating the very low- and high-frequency portions of the corrected 15-phon data. The high-frequency asymptotic rolloff of the modified E-weighting curve is second order. For the improved E-weighting curve, in the absence of good data in the top octave we have extrapolated from the available data by rather arbitrarily setting the curve at −100 dB at 20 kHz. Spline approximations through the data points have been used to generate the improved E-weighting curve shown in Fig. 2. Both curves are normalized so as to have unity power gain over the band 0–20 kHz.[3]

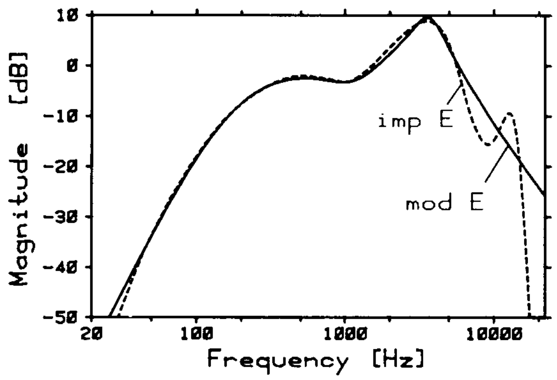


Fig. 2. Solid curve—Fielder's modified E-weighting curve; dashed curve—improved E-weighting curve, which at high and low frequencies better matches the ear's low-level (15-phon) noise sensitivity.

In this paper we shall compare results using both the modified and the improved E-weighting curves as our audibility weighting function $W(f)$. Both curves correctly render the ear's increased sensitivity in the 1–5-kHz region, and this is important if our aim is to shape accurately the requantization noise spectrum to be minimally audible. The improved E-weighting curve goes further by steepening the third-order low-frequency rolloff of the modified E-weighting curve as well as more accurately fitting the high-frequency data around 12 kHz and attempting to better account for the dramatic loss of sensitivity at higher frequencies. We are not specifying this improved E-weighting curve more precisely here because there is some arbitrariness in its derivation. Its purpose at the moment is to show how its better psychoacoustic modeling significantly enhances the performance capabilities of the noise-shaper designs under consideration. We believe that there is a need for more work to standardize a weighting curve to reflect accurately the low-level broad-band noise sensitivity of the average human ear.

Nevertheless, we shall proceed to optimize our noise shapers to the two weighting curves shown in Fig. 2. One can compute the theoretical Gerzon–Craven limit, Eq. (7), to the weighted noise-shaping improvements w possible with these two curves. For modified E-weighting one finds that, for a sampling frequency f_s of 44.1 kHz,

$$w(\text{mod E}) = 0.07184 = -11.44\ \text{dB} \tag{11}$$

while for the improved E-weighting curve,

$$w(\text{imp E}) = 0.002268 = -26.44\ \text{dB}\ . \tag{12}$$

Note that these would represent perceived noise reductions (due to shaping) of approximately 2 and 4.5 bits, respectively, if the inverses of these curves could be realized accurately. These are very substantial numbers. Due to the low-frequency asymptotic slopes of both these curves, the integral in Eqs. (9) and (10) does not converge, and the prediction is that neither curve is realizable—that is, the unweighted noise power will become very large as higher and higher order approximating filters $H(z)$ are tried in an attempt to realize the noise-power spectrum of Eq. (6). The modified E-weighting curve is much less extreme than the improved

[3] The improved E-weighting curve is further refined and explicitly defined (as a new curve, called F-weighting) in [9].

E-weighting curve in this respect, as we shall see. In fact, if we assume that no poles are placed close to 0 Hz in an attempt to mimic the triple zero at 0 Hz in the modified E-weighting curve, the integral, Eq. (9), evaluated from, say, 40 Hz to 22.05 kHz, predicts a total unweighted noise power $3N_u$ on the order of 17.2 units of $\Delta^2/12$, corresponding to a modest increase in the total power of about 12.4 dB including the dither, in return for an audible decrease in the noise of around 11 dB. When compared with a similarly dithered but not noise-shaped requantizer, the unweighted noise power increase is only 7.5 dB in return for this 11-dB perceived reduction. These are reasonable numbers.

Both weighting functions considered so far are continuous functions of frequency. One possible variation on the modified E-weighting curve, which we in fact used in [3], [4], is to "chop" the curve beyond 20 kHz (to $-\infty$ dB) in order to acknowledge that it is not rolling off rapidly enough at high frequencies. Such a truncation results in a value of w of $-\infty$ dB, as mentioned in Sec. 1. The prediction then would be that as much weighted noise reduction as desired is possible by using a high enough filter order to place as much as necessary of the noise in the "dead band" between 20 kHz and the Nyquist frequency. This does, of course, require a very high order filter indeed, and is not desirable in any event because of the very great increase in the unweighted noise which it would entail. We shall therefore not pursue this truncated modified E-weighting curve in this paper.

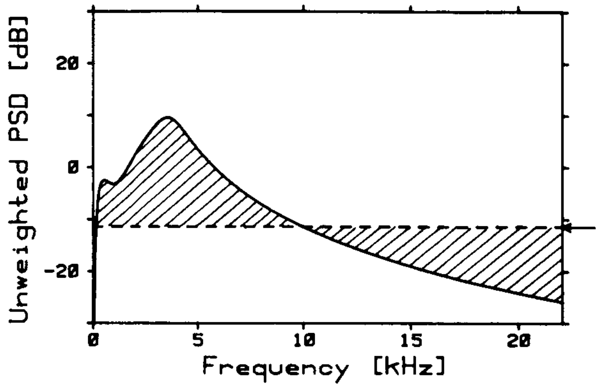
3 PRACTICAL NOISE-SHAPER RESULTS

We shall now investigate the practical realization of low-order noise-shaper designs which attempt to achieve the theoretical noise improvements. We shall throughout assume a sampling rate f_s of 44.1 kHz, which corresponds to the CD standard. In [3], [4] we showed the results of the use of a simple first-order noise shaper with

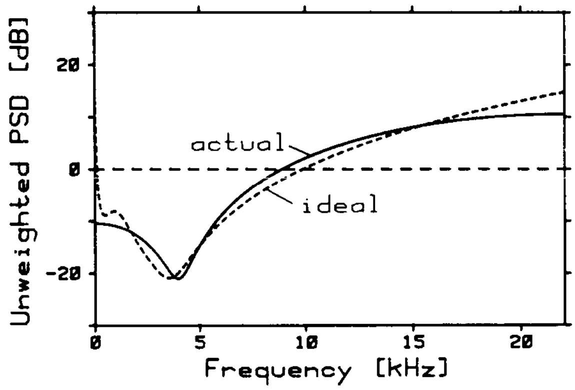
$$H(z) = z^{-1}$$

which corresponds in Fig. 1 to $a_0 = 1$ and all other a_i and b_i being zero. This is the form of noise shaper adopted by Philips in its early CD players (albeit without dither) in conjunction with interpolation by a factor of 4 [2]. In the context of optimized noise shapers without oversampling, the choice of $a_0 = 1$ is not optimal with either modified or improved E-weighting, but does nevertheless result in a lower weighted noise power than plain undithered rounding (see [3], [4]).

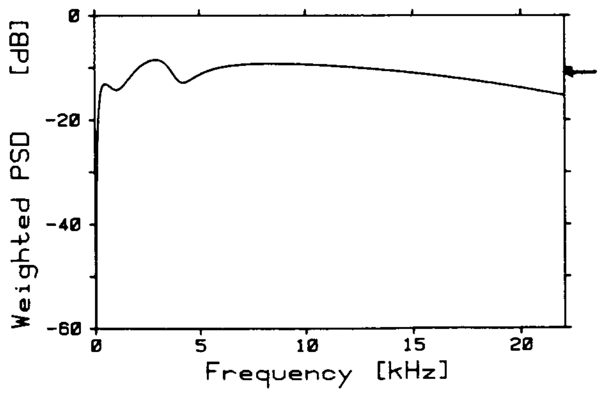
In order to illustrate the general idea it is instructive, before plunging in and investigating the full variety of options, to examine a simple case—say the second-order nonrecursive filter, having a_0 and a_1 nonzero, but all remaining a_i and b_i zero. This is interesting because it does exceptionally well when measured using Fielder's modified E-weighting curve. Fig. 3(a) shows, on a linear frequency scale, the modified E-weighting curve and the Gerzon–Craven limit w of Eq. (11) to illustrate the definition of w, Eq. (7), as the average of the *logarithm* of $W(f)$: the shaded areas are equal above and below the line representing w. Using Eq. (6) we can now construct the optimal noise-shaper power spectrum, namely, $w/W(f)$. This is plotted as the dashed curve in Fig. 3(b) and represents in both shape and



(a)



(b)



(c)

Fig. 3. (a) Modified E-weighting curve showing (shaded) the fact that areas above and below dashed horizontal line are equal. This level, indicated on right by arrow, is the average w of the logarithm of $W(f)$ and equals -11.44 dB [see Eq. (7)]. (b) Dashed curve—ideal noise-shaper power spectrum $w/W(f)$ of Eq. (6); solid curve—spectrum actually realized by optimal second-order nonrecursive noise shaper discussed in text. $a_0 = 1.537$; $a_1 = -0.8367$. (c) Audibility-weighted actual noise-shaper power spectrum. Arrow marks achieved weighted noise level N_w of -10.88 dB. Ideal shaper would produce flat spectrum at -11.44 dB.

vertical positioning the (unachievable) theoretical optimum. It is unachievable (as discussed in Sec. 2) since the total power represented by the area under the curve of $w/W(f)$ is infinite. If one now proceeds to optimize the two shaper coefficients a_0 and a_1 to minimize the weighted noise power of Eq. (5), one finds the following values:

$$a_0 = 1.537 \qquad a_1 = -0.8367 \ .$$

These optimal parameters result in the realized shaper noise-power spectrum shown as the solid curve in Fig. 3(b). Notice how good a fit it is to the desired dashed curve in both shape and height. One easily verifies that the function $1 - H(z)$ thus realized is minimum phase, as predicted in [8]. To appreciate how good this filter is, it is meaningful to plot the modified E-weighted output noise-power spectrum of this design, which we do in Fig. 3(c), omitting the power density factor $2/f_s$ in Eq. (2) so that it may be directly compared with w. The theoretical optimum would produce a totally flat weighted spectrum at height w. The arrow in Fig. 3(c) represents the achieved weighted noise power N_w of 0.08168 units, or -10.88 dB, which is within 0.56 dB of the Gerzon–Craven limit of -11.44 dB. (Achieved values of N_w will likewise be indicated by arrows on all subsequent weighted noise-power spectra.) Using this weighting curve, the further improvement possible with higher order filters $H(z)$ cannot exceed an additional 0.56 dB. The total output noise power $3N_u$, given by Eq. (4), is 12.18 units of $\Delta^2/12$, or about 1 LSB rms, which is very modest, representing an increase of about 6.1 dB over and above that of an unshaped but properly dithered quantizer.

These second-order results are already so good that only minor improvements are possible using higher order noise shapers, as long as we stay with the modified E-weighting curve as our audibility criterion. To see how increasing the order of the filter $H(z)$ modifies the noise-shaper performance, Fig. 4 displays both the unweighted and the weighted noise-power spectra for FIR filters of orders 3, 5, 7, and 9. As in Fig. 3(c), and in all the remaining such curves, the power density factor of $2/f_s$ in Eq. (2) has been omitted to enable easier comparison with the Gerzon–Craven theoretical limit w. Clearly, increasing the order beyond second merely serves to gild the lily, as the approximation can improve only slightly with order. Of course, the unweighted noise power also increases slightly with order. The results for all filter orders from 1 through 9 are summarized in Table 1, which gives the unweighted and weighted noise powers both linearly and in decibels for easier comparison. With increasing order the high-frequency match very soon becomes almost perfect, and the deviation from a flat weighted noise spectrum becomes quite small, except below 250 Hz where the nonrecursive filter topology is unable to compensate for the sixth-order zero in $W(f)$ at 0 Hz. This inability on the filter's part also serves, however, to keep the unweighted total noise power in check and so, as predicted in Sec. 2, $3N_u$ does not rise much above 17 units of power with filters of the order shown, while the weighted noise power N_w reaches within 0.2 dB of the Gerzon–Craven limit for this weighting curve. The higher order FIR filters in this case produce small incremental weighted noise improvements with only small increases in unweighted noise power. The noise-power spectra shown in Fig. 4 are all eminently acceptable, having only moderate low-frequency power densities and a monotonically rising trend toward the Nyquist limit.

Let us now free up the denominator coefficients b_i in Eq. (1) by allowing recursive (IIR)-type filter topologies in the noise shaper. For an n-delay structure ($n \geqslant 1$) this furnishes us with $(2n - 1)$ coefficients $a_0, a_1, \ldots, a_{n-1}, b_1, b_2, \ldots, b_{n-1}$ for minimization of the weighted output noise. The resulting optimal unweighted and weighted noise-power spectra for 3-, 5-, 7-, and 9-coefficient IIR filters are shown in Fig. 5, while Table 2 summarizes the results. For a given number of coefficients, the IIR topology results in a somewhat flatter weighted noise-power curve and an incremental improvement in N_w, but with a slightly larger peak in the weighted noise-power spectrum around 500 Hz. Again, as expected, N_u remains well in check as the order is increased.

We should point out that finding the *global* minimum of a function of many parameters (such as N_w) is a difficult problem, and no optimization routine can guarantee to have found the global rather than a local minimum. In these, and in all the remaining data, we believe that the figures given represent the global minimum subject to the requirement that $1 - H(z)$ be minimum phase, but it is possible that some slightly better minima may exist in some of the higher order cases. It should also be noted that, although the recursive topology of Fig. 1(b) allows poles to be placed at 0 Hz, such marginally stable configurations have been excluded from consideration both for being nonminimum phase and because of the undesirability of substantial infrasonic noise power. Thus even though the weighting curves under consideration have true zeros at 0 Hz, and would require 0-Hz poles for an optimal complementary noise spectrum, we do not allow this to occur.

As we indicated in Sec. 2, the modified E-weighting curve is not as good a match to the ear's noise sensitivity as might be desired at both frequency extremes. A better match is the improved E-weighting curve shown in Fig. 2. In particular, its sharper rolloff at high frequencies results in a substantially reduced Gerzon–Craven limit for w of -26.44 dB, compared with -11.44 dB for the modified E-weighting curve. It is therefore of interest to see how much of this predicted reduction in the perceived noise floor can be achieved with moderate filter orders.

To begin with, we set all the b_i to zero and consider purely nonrecursive implementations. Fig. 6 shows the unweighted and weighted optimal noise-power spectra for FIR filter orders of 3, 5, 7, and 9, while Table 3

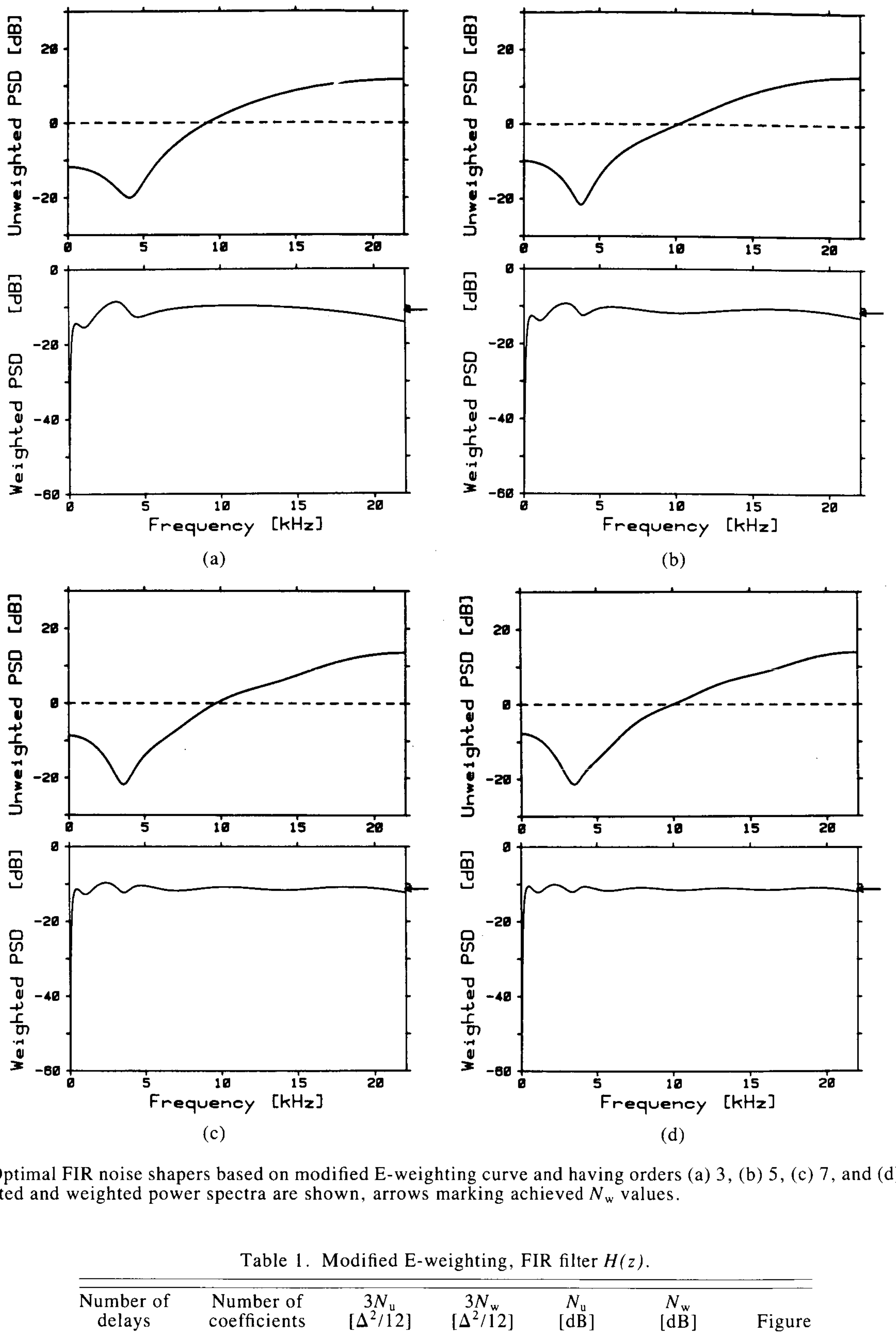


Fig. 4. Optimal FIR noise shapers based on modified E-weighting curve and having orders (a) 3, (b) 5, (c) 7, and (d) 9. Both unweighted and weighted power spectra are shown, arrows marking achieved N_w values.

Table 1. Modified E-weighting, FIR filter $H(z)$.

Number of delays	Number of coefficients	$3N_u$ $[\Delta^2/12]$	$3N_w$ $[\Delta^2/12]$	N_u [dB]	N_w [dB]	Figure
1	1	5.100	0.8169	2.304	−5.650	—
2	2	12.18	0.2451	6.086	−10.88	3
3	3	14.55	0.2404	6.857	−10.96	4(a)
4	4	17.09	0.2302	7.555	−11.15	—
5	5	16.79	0.2300	7.479	−11.15	4(b)
6	6	17.32	0.2273	7.615	−11.21	—
7	7	17.09	0.2269	7.556	−11.21	4(c)
8	8	17.23	0.2257	7.592	−11.24	—
9	9	17.12	0.2254	7.564	−11.24	4(d)
Gerzon–Craven limit		∞	0.2155	∞	−11.44	

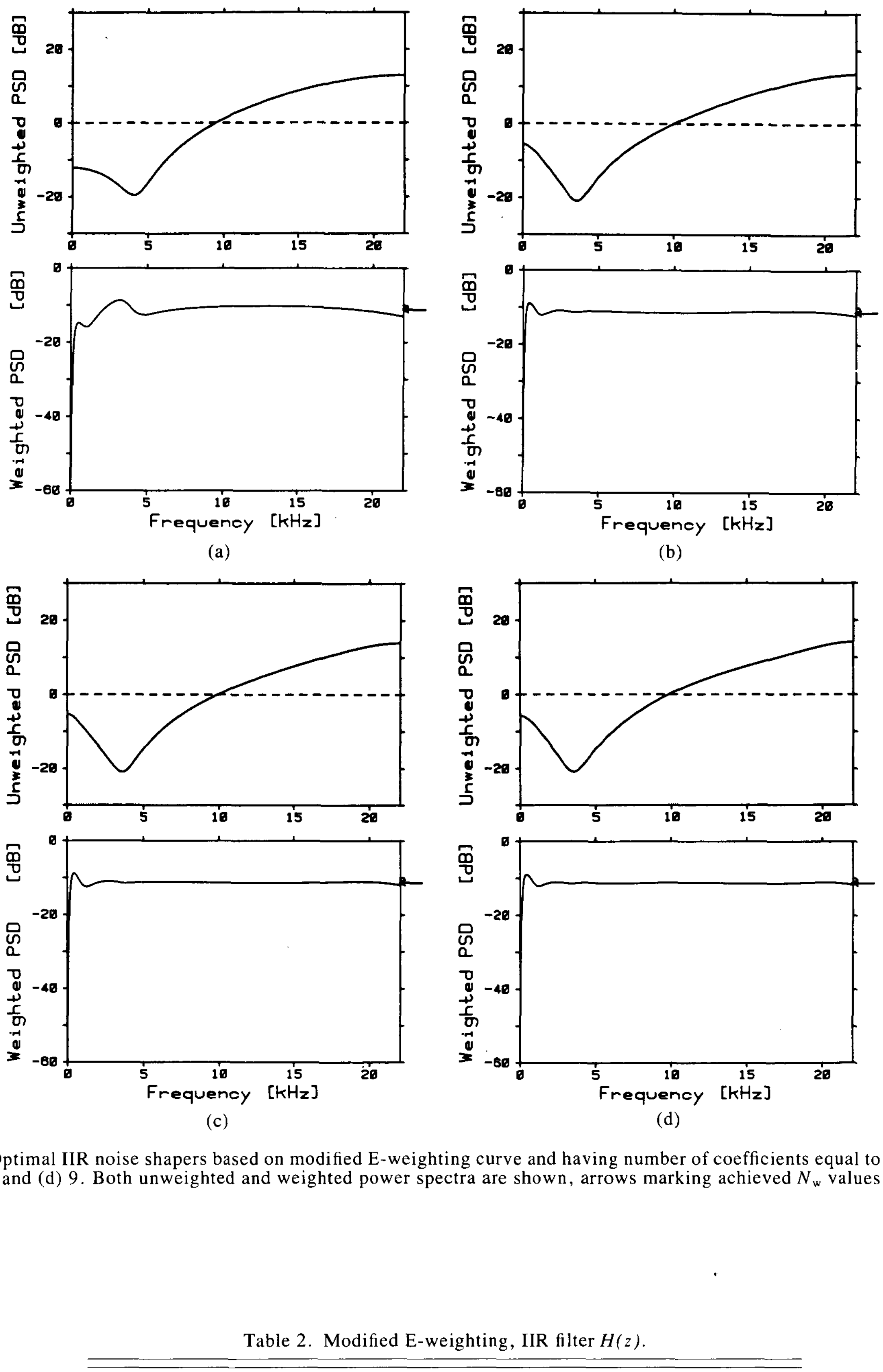


Fig. 5. Optimal IIR noise shapers based on modified E-weighting curve and having number of coefficients equal to (a) 3, (b) 5, (c) 7, and (d) 9. Both unweighted and weighted power spectra are shown, arrows marking achieved N_w values.

Table 2. Modified E-weighting, IIR filter $H(z)$.

Number of delays	Number of coefficients	$3N_u$ $[\Delta^2/12]$	$3N_w$ $[\Delta^2/12]$	N_u [dB]	N_w [dB]	Figure
2	3	16.88	0.2370	7.503	−11.02	5(a)
3	5	16.99	0.2246	7.532	−11.26	5(b)
4	7	17.08	0.2244	7.553	−11.26	5(c)
5	9	17.06	0.2244	7.549	−11.26	5(d)
Gerzon–Craven limit		∞	0.2155	∞	−11.44	

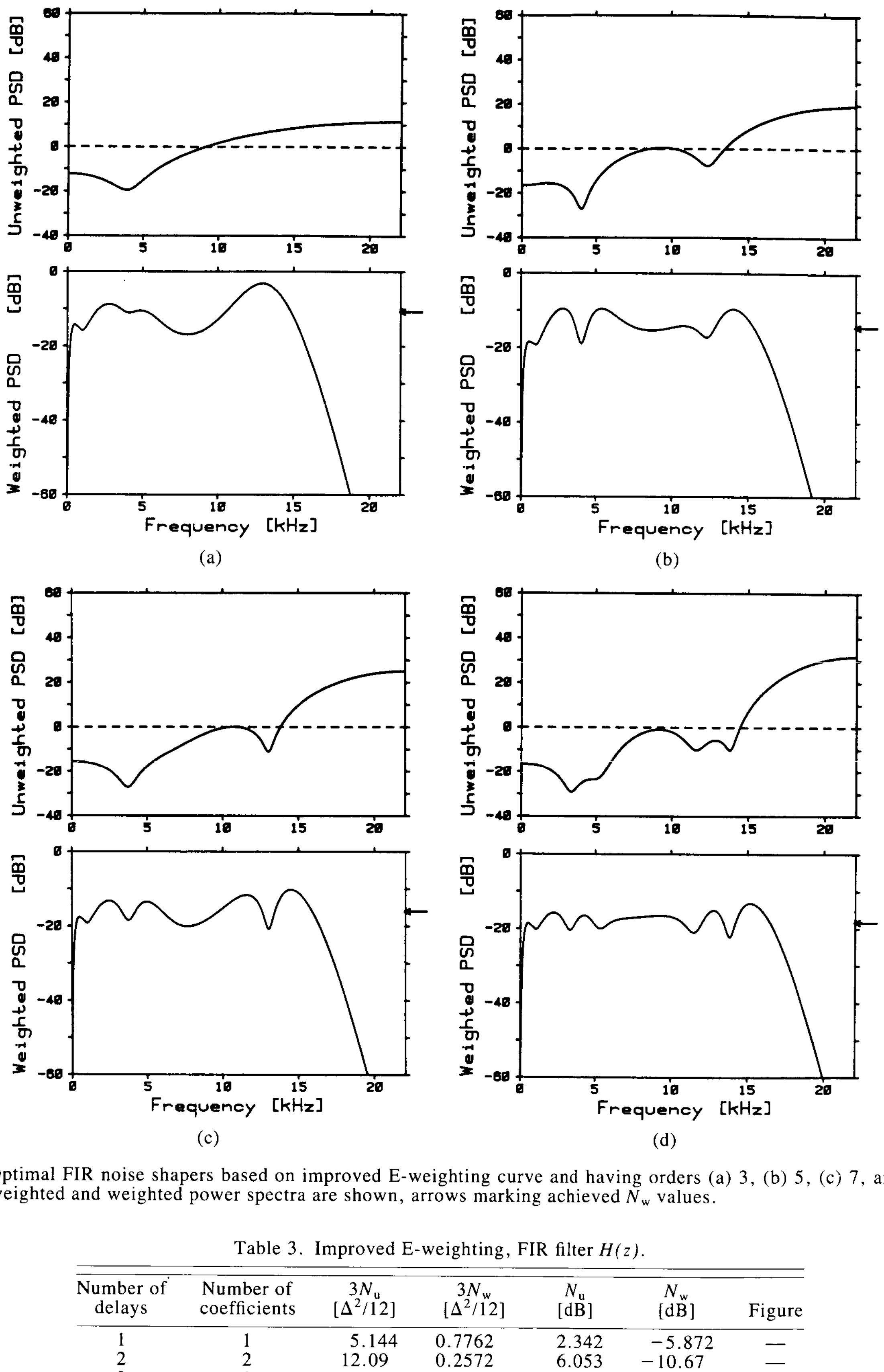


Fig. 6. Optimal FIR noise shapers based on improved E-weighting curve and having orders (a) 3, (b) 5, (c) 7, and (d) 9. Both unweighted and weighted power spectra are shown, arrows marking achieved N_w values.

Table 3. Improved E-weighting, FIR filter $H(z)$.

Number of delays	Number of coefficients	$3N_u$ $[\Delta^2/12]$	$3N_w$ $[\Delta^2/12]$	N_u [dB]	N_w [dB]	Figure
1	1	5.144	0.7762	2.342	−5.872	—
2	2	12.09	0.2572	6.053	−10.67	—
3	3	14.02	0.2536	6.698	−10.73	6(a)
4	4	22.74	0.1777	8.797	−12.27	—
5	5	49.70	0.1105	12.19	−14.34	6(b)
6	6	105.4	0.07975	15.46	−15.75	—
7	7	155.6	0.07474	17.15	−16.04	6(c)
8	8	324.3	0.05359	20.34	−17.48	—
9	9	612.5	0.04414	23.10	−18.32	6(d)
Gerzon/Craven limit		∞	0.006805	∞	−26.44	

lists the results for all orders 1 through 9. Note the changed scale on the unweighted noise spectra relative to that used when we considered the modified E-weighted curve. The improved E-weighting curve data will need this increased dynamic range for the display of some of the more extreme cases yet to come. It is immediately apparent that the higher order filters improve on the weighted noise powers obtainable with the modified E-weighting curve, but the dramatic rise in total unweighted noise power necessary in order to do so is sobering. For example, the ninth-order FIR filter achieves $N_w = -18.32$ dB (still some way off from the Gerzon–Craven limit of -26.44 dB), but requires a total noise power of 612.5 units of $\Delta^2/12$, or about 7 LSB rms, in order to do so. The corresponding-order FIR filter, based on the modified E-weighting curve, only achieved $N_w = -11.24$ dB, but contributed only 17.12 units of noise (about 1.2 LSB rms). The reasons are apparent from Fig. 6. The rapid rolloff in the weighting curve above 15 kHz allows ample scope for a weighted noise improvement (that is, a flattening of the weighted curve) at the expense of a substantially boosted noise-power output in this region. Note the increase in the high-frequency versus low-frequency boost in the unweighted noise-power curves in Fig. 6(a)–(d) as the order is increased from 3 to 9, and the corresponding flattening, lowering, and extension toward the high frequencies of the weighted noise-power curves.

It is of interest to note how, for FIR filter orders of 5 and greater, advantage is taken of the opportunity to introduce a spectral dip around 12 kHz to complement the audibility peak at this frequency in the improved E-weighting curve. By order 9, double complex zeros are appearing around the weighting function's peaks at 4 and 12 kHz in order to better match their breadth and hence reduce the ripple in the weighted noise curve.

Are noise powers of the amounts seen here acceptable? The answer must be a qualified yes. Although 612.5 units of $\Delta^2/12$ is a significant 23.1-dB increase over and above the 3 units of an unshaped dithered quantizer, it still amounts to only 7 LSB rms of noise. It will not restrict the dynamic range of the digital system, nor will it show up on most digital level meters whose display extends down to 60 dB below clipping. (This noise would be at about -70 dB rms.) Noise powers of this order are quite acceptable. But how much higher can one allow the unweighted noise power to rise and still tolerate it in return for the reduction in *audible* noise? This question arises as soon as we set the full recursive (IIR) topology for $H(z)$ to work on the improved E-weighting curve. For, now there is a benefit in terms of *weighted* noise power to place high-Q poles near the Nyquist frequency. The unweighted noise power is the immediate casualty. Consider, for example, the 5-coefficient (that is, 3-delay) IIR filter. Its "optimum" is represented by the curves of Fig. 7 (note the scale). Compared with the corresponding 5-coefficient FIR case of Fig. 6(b) and Table 3, this new curve is 1.8 dB quieter at $N_w = -16.11$ dB, but its unweighted noise parameters are

$$N_u = 70.30 \text{ dB}$$

$$3N_u = 32\ 145\ 076 \text{ units of } \Delta^2/12$$

$$= 1637 \text{ LSB rms} .$$

This is clearly intolerable. The situation gets worse as the filter order is increased. Our prediction of the unweighted noise power becoming unbounded indeed seems to be coming true.

The most sensible solution seems to be to perform constrained weighted noise-power optimizations, with acceptable limits on N_u. This is what we proceeded to do, with total power limits of N_u = 20, 30, and 40 dB, these corresponding to values of $3N_u$ of 5, 15.81, and 50 LSB rms, respectively. The 20-dB limit is certainly acceptable on all criteria; the 30- and 40-dB limits may represent undesirable amounts of system noise. The results for IIR filters with 3, 5, 7, and 9 coefficients are collected in Table 4 and Figs. 8–10.

The 20-dB power limit IIR filters of up to 9 coefficients do not achieve as low a weighted noise power as did the 9-coefficient FIR filters, and produce decidedly less flat weighted power curves. This is because they still concentrate on placing more power above 15 kHz, not as immoderately as the unconstrained case of Fig. 7 did, but nevertheless to the detriment of the spectral shaping at lower frequencies. With the higher power limits of 30 and 40 dB this trend continues, but

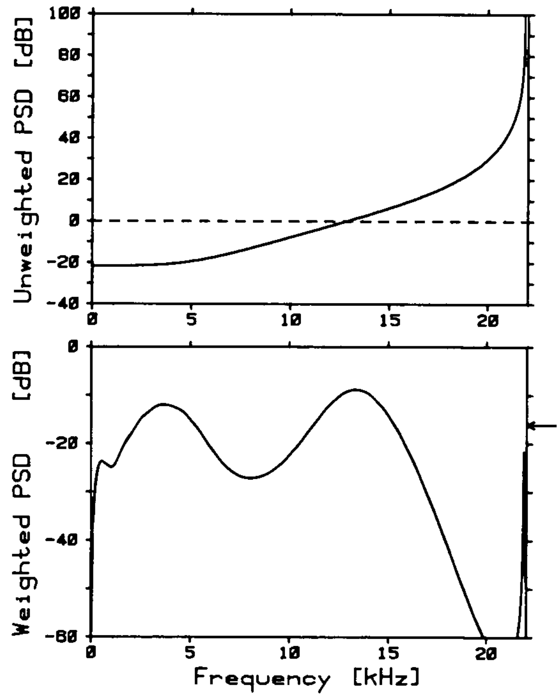


Fig. 7. Unconstrained optimization of 5-coefficient IIR noise shaper based on improved E-weighting curve. Unweighted noise-power spectrum has dynamic range of more than 120 dB and produces high-frequency peak in weighted spectrum while achieving a value of $N_w = -16.11$ dB. Total unweighted noise power is unacceptably large at N_u = 70.30 dB (see text).

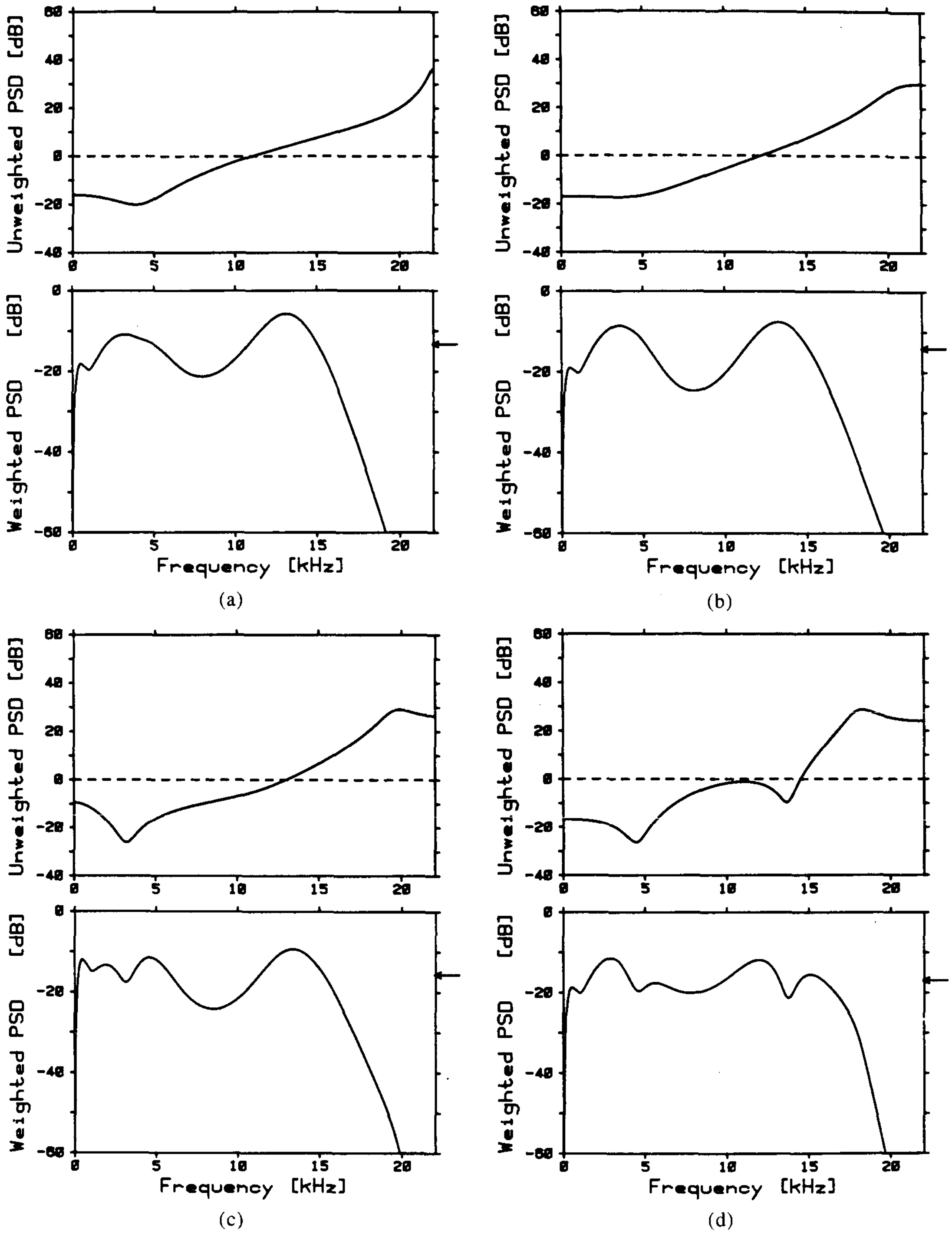


Fig. 8. Optimal IIR noise shapers based on improved E-weighting curve and having a total power constraint of N_u = 20 dB. Curves are for unweighted and weighted power spectral density with number of shaper coefficients equal to (a) 3, (b) 5, (c) 7, and (d) 9.

now higher-Q poles appear in this frequency band and the weighted noise power drops below −19 dB for the first time. An examination of Figs. 8–10, however, leads us to conclude that for comparable numbers of coefficients (which is the determining criterion in terms of real-time implementation) the optimized FIR filter designs are preferable. With all the designs exhibiting a large noise-power rise above 10 kHz, and especially with the designs of Figs. 9 and 10 with their large ultrasonic noise peaks, there is a real potential for inadvertent tweeter damage to occur if excessive gain is used during replay (for example, if one turns up the gain unduly to "listen to the background noise"). This is one further reason for preferring the lower-Q shaper designs.

The results of all these experimental noise-shaper designs are collected together in Fig. 11 to facilitate comparison. In this diagram we have plotted both unweighted (N_u) and weighted (N_w) noise powers as a function of the number of filter coefficients for both FIR (solid curves) and IIR (dashed curves) types and for both modified E-weighting and improved E-weight-

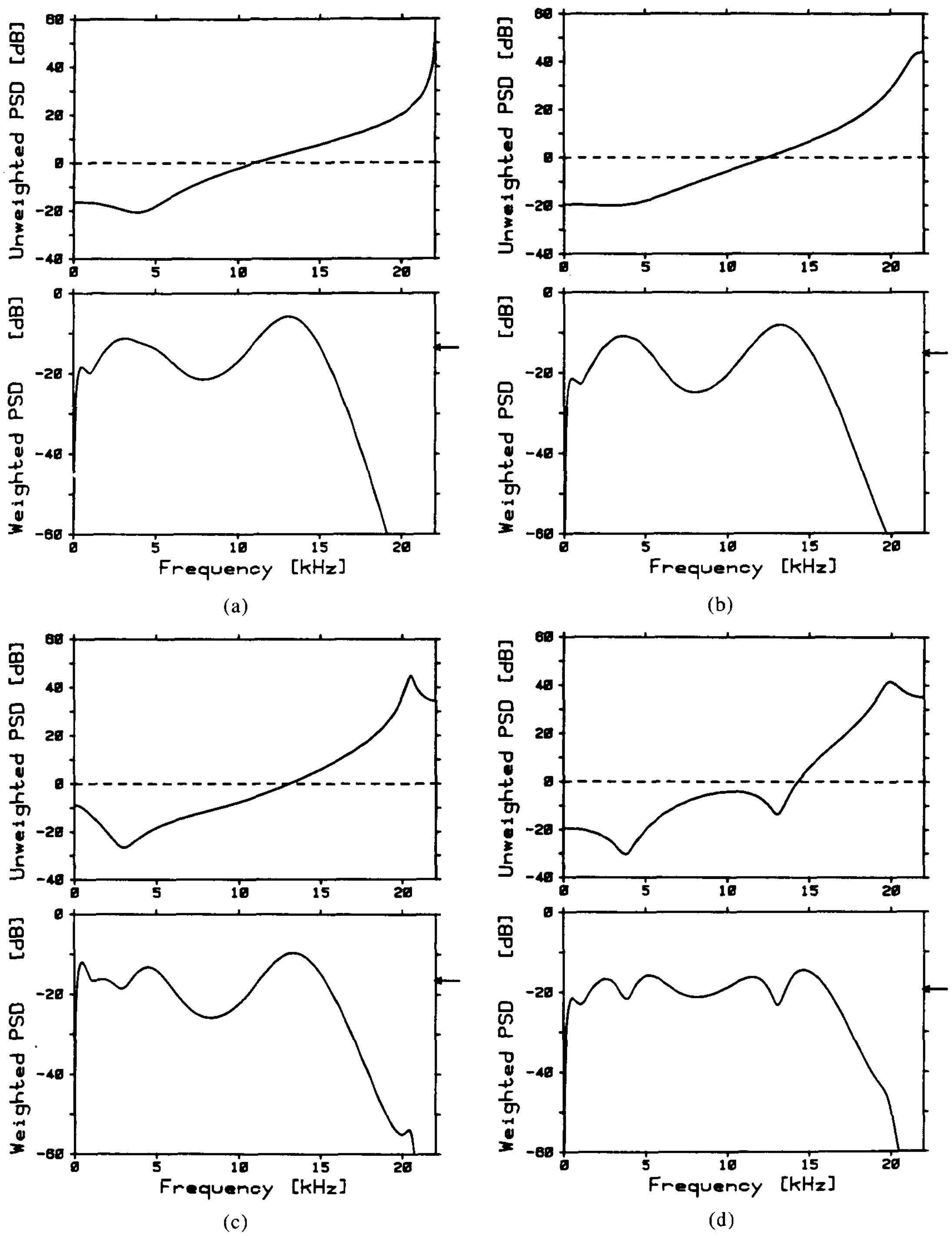


Fig. 9. Same as Fig. 8, but with total power constraint of N_u = 30 dB.

ing audibility curves. It is easily seen how, with modified E-weighting, a saturation soon sets in for both N_u (at around 7.5 dB) and N_w (at around −11 dB) when the number of coefficients exceeds 3. For the improved E-weighting curve, however, the results are somewhat different. For the FIR filter types N_u rises monotonically with order while N_w continues to fall toward the Gerzon–Craven limit of −26.44 dB, and beyond three coefficients these designs improve on those based on the modified E-weighting curve in that they achieve lower N_w figures, but at the expense of higher N_u figures. For IIR-type filters based on the improved E-weighting curve the results shown in Fig. 11 correspond to the N_w values achievable with N_u limited to 20, 30, or 40 dB, as indicated on the diagram. We see that with N_u = 20 dB, the weighted noise powers N_w are worse than the corresponding FIR designs of order 5 or more. The IIR designs with power limits of 30 or 40 dB do indeed improve upon the FIR results, but at the expense of substantially greater N_u values, of course.

Bear in mind that in this investigation we have examined only relatively low-order filters. Our aim has been to see how the theory of [8] applies in practice to two possible audibility-weighting curves. We have

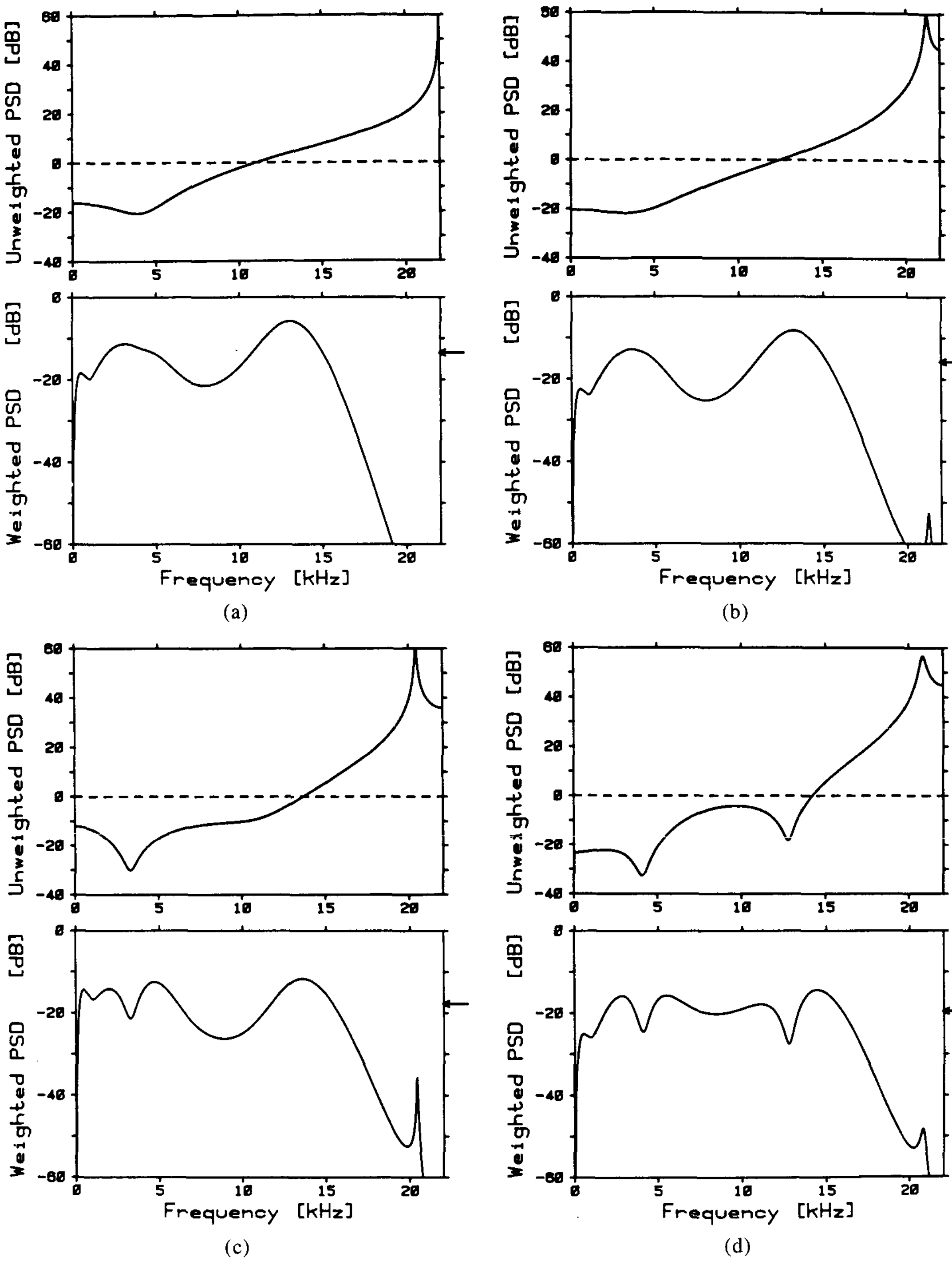


Fig. 10. Same as Fig. 8, but with total power constraint of N_u = 40 dB.

found that, with Fielder's modified E-weighting curve, around 11 dB of audible noise improvement is easily achievable with modest filter orders and only moderate (around 7.5 dB) total noise increases relative to unshaped quantizers. In order to allow for the possibility of greater weighted noise reductions it is necessary to use a weighting curve, like our improved E-weighting curve, which more accurately represents the ear's high-frequency rolloff. Although substantially greater noise reductions are indeed now possible (over 19 dB with 9-coefficient filters), the total unweighted noise power now becomes unacceptably high in order to achieve these large weighted noise reductions. Even with a limit on the total noise power, the IIR filter results are not really satisfactory. On the other hand, the ninth-order FIR filter is perfectly acceptable, producing over 18 dB of noise reduction with only a 23-dB total noise penalty.

A compromise would appear to be the solution to the conflicting requirements of large audible noise reductions coupled with tolerable total noise powers. If one modifies the improved E-weighting curve by shelving its high- and low-frequency rolloffs so that they flatten off below, say, −40 or −50 dB, the

Table 4. Improved E-weighting, IIR filter $H(z)$.

Unweighted power limit	No. of delays	No. of coefficients	$3N_w$ $[\Delta^2/12]$	N_w [dB]	Figure
N_u = 20 dB	2	3	0.1418	−13.26	8(a)
$3N_u$ = 300 $[\Delta^2/12]$	3	5	0.1185	−14.03	8(b)
= 5 LSB rms	4	7	0.08317	−15.57	8(c)
	5	9	0.06012	−16.98	8(d)
N_u3 = 30 dB	2	3	0.1362	−13.43	9(a)
$3N_u$ = 3000 $[\Delta^2/12]$	3	5	0.09169	−15.15	9(b)
= 15.81 LSB rms	4	7	0.06748	−16.48	9(c)
	5	9	0.03458	−19.38	9(d)
N_u3 = 40 dB	2	3	0.1356	−13.45	10(a)
$3N_u$ = 30 000 $[\Delta^2/12]$	3	5	0.07902	−15.79	10(b)
= 50 LSB rms	4	7	0.05345	−17.49	10(c)
	5	9	0.03296	−19.59	10(d)
∞	Gerzon–Craven limit		0.006805	−26.44	

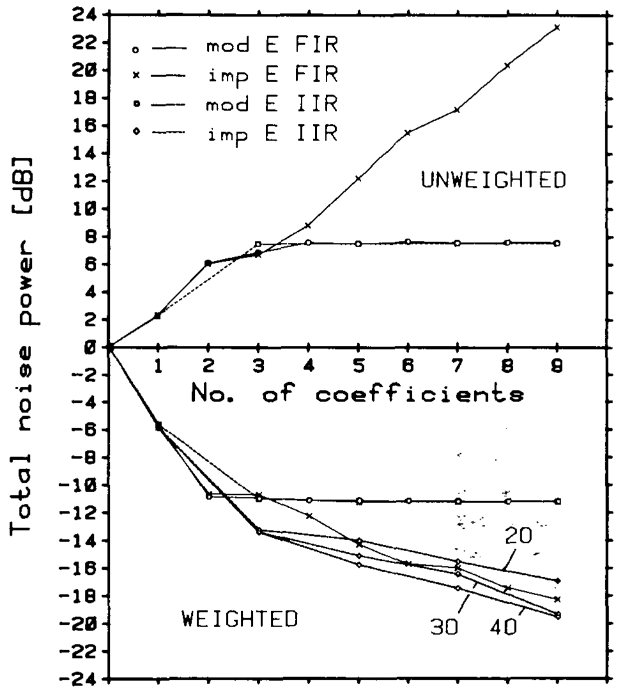


Fig. 11. Total unweighted (N_u) and weighted (N_w) noise powers as a function of number of coefficients for all noise-shaper designs considered in paper. Labels 20, 30, and 40 refer to total unweighted power limitation in decibels for improved E-weighted IIR noise shapers.

Gerzon–Craven theory predicts that not only will a theoretical weighted noise-power reduction of more than 11.44 dB (but less than 26.44 dB) be possible, but it will be achievable with a *finite* total unweighted noise power which should be of an acceptable magnitude. This is explored further in [9], where calculations show that the following should be attainable. With shelving of the weighting curve at −40 dB we predict $N_w \approx -17$ dB with $N_u \approx 18$ dB; with shelving at −50 dB, the predictions are $N_w \approx -19$ dB and $N_u \approx 25$ dB; and with shelving at −60 dB, one should be able to achieve $N_w \approx -21$ dB for a total noise power still acceptable at $N_u \approx 32$ dB. Such an approach yields smoother weighted noise-power spectra than those of Figs. 8–10.

One possible design procedure for the nonrecursive type of noise-shaper filter is to compute the inverse Fourier transform of the minimum-phase spectrum corresponding to the square root of the spectral power density function $w/W(f)$ for such a realizable filter, suitably time-limit it, and use its samples to compute using Eq. (6) the tap weights a_i of an FIR filter $H(z)$. This produces a long, but very good, error-feedback filter. For more details, see [9].

For other than very short filter lengths, and preferably even in such cases, one should beware of the "starting transient" of the noise shaper, during which the noise spectrum is stabilizing, by ensuring that the shaper is kept running continuously when such stop–start operations as digital splice editing are being performed. One cannot simply start the noise shaper "unprimed."

To enable those readers who may wish to experiment with noise-shaper designs of the type discussed in this paper to more conveniently do so, we have listed in the Appendix the noise-shaper coefficient values a_i and b_i for a selection of the filters investigated.

4 CONCLUSIONS

We have shown that noise shapers of orders less than 10 can achieve weighted noise-power reductions of about 2 bits (12 dB) with ease, and over 3 bits (18 dB) with still acceptable total unweighted noise penalties. Of the designs explored in this paper, the nonrecursive error-feedback filter designs generally display flatter weighted noise-power spectra for a given order (that is, number of filter coefficients). The higher weighted noise-power reductions are only possible if one adopts a more realistic audibility weighting function such as the improved E-weighting curve used here, and then one finds that the unweighted noise power rises dramatically if one attempts to achieve any further weighted power reductions. The way out of this dilemma seems to be to tailor the ultimate rolloff of the weighting curve to restrict the unweighted noise power to acceptable levels. This stratagem will enable weighted powers on the order of −20 dB to be achieved with very flat weighted power spectra and acceptable unweighted noise powers, and is explored in a subsequent paper [9].

More work is desirable with the aim of standardizing the most audibly meaningful weighting curve for low-level broad-band noise. Nevertheless, a very significant and worthwhile perceived (re)quantization noise reduction is already possible with the designs covered in this paper. One immediately apparent benefit of these schemes is the ability to hear a considerable apparent increase in low-level signal resolution as a consequence of the reduced aural masking by the (re)quantization noise as a result of its redistribution to regions of the spectrum where it is less audible. In particular, the removal of noise from the region of greatest sensitivity around 3 kHz is mainly responsible for this effect. The ability to reduce wordlengths by 2–3 bits without apparent noise penalty can be of benefit to the audio industry as professional digital recordings of 18 or more bits resolution begin to find their way onto the 16-bit CD medium.

It should be noted, however, that the noise-shaped rounding operation should be the last requantization performed before analog reconversion. For, any subsequent unshaped requantization will add a white requantization noise floor which will "fill in" the dips so carefully engineered in the shaped noise floor around 3 and 12 kHz, greatly degrading the audible signal-to-noise ratio of the channel. This means that such signal-processing devices as equalizers, faders, sampling-standards converters, and even the "oversampling" digital filters associated with interpolating digital-to-analog converters (DACs) need thus to be noise shaped to prevent their numerical truncations from dominating the audible noise spectrum. Unless one intends shaping every such operation (and repeated shaping operations are probably undesirable), it will thus be necessary to carry and store intermediate word lengths of at least 4 bits more than the final word length to which the rounding (or analog reconversion) will be performed. Indeed, all subsequent requantizations need to achieve 3 kHz noise floors corresponding to at least 20 bits in order not to degrade the signal-to-noise ratio. This is a severe requirement and, as far as the DAC is concerned, necessitates the use of a properly noise-shaped converter or a very quiet non-oversampled system if one wishes just to be able to *listen* to the recorded signal without audible change. If the recording is a preemphasized one, one side benefit of the subsequent analog deemphasis operation will be to lower the boosted noise powers above 10 kHz significantly upon playback (and, of course, also to require a separate noise-shaper spectral optimization, in principle).

Another side benefit of noise-shaped rounding schemes is the increased smoothing of DAC nonlinearities contributed by the larger excursions which the increased total noise power provides. Of course, such converter nonlinearities will prevent the achievement of the designed shaped noise spectrum (by "filling in the dips"), and will introduce some measure of noise modulation too. (For further discussion see [9, sec. VI].) So the requirement for accurate DACs remains. We note in passing that care must be exercised in subsequent "butt splice" re-editing of noise-shaped data segments, since the discontinuities so introduced (in the zeroth, first, . . . derivatives of the waveform) disturb the correlation of the shaped noise sequence and can produce audible "clicks." Finally, we should mention that one aspect yet to be examined is the extent to which the complex (approximately chaotic) behavior of the higher order noise-shaper designs can allow the amount of injected dither to be reduced without deleterious audible consequences. Preliminary experiments, however, seem to indicate that full triangular-pdf dither of 2 LSB peak to peak should be used.

5 ACKNOWLEDGMENTS

This work has been supported by the Natural Sciences and Engineering Research Council of Canada. The authors wish to thank Ewan Macpherson for his assistance with the plotting of the graphs.

6 REFERENCES

[1] H. A. Spang and P. M. Schultheiss, "Reduction of Quantizing Noise by Use of Feedback," *IRE Trans. Commun. Sys.*, vol. CS-10, pp. 373–380 (1962 Dec.).

[2] R. J. van de Plassche and E. C. Dijkmans, "A Monotonic 16-Bit D/A Conversion System for Digital Audio," presented at the AES Premiere Conference: Digital Audio, Rye, New York, 1982; in B. Blesser, B. Locanthi, and T. G. Stockham, Jr. (Eds.), *Collected Papers* (Audio Engineering Soc., New York, 1983), pp. 54–60.

[3] J. Vanderkooy and S. P. Lipshitz, "Digital Dither: Signal Processing with Resolution Far Below the Least Significant Bit," *Proc. AES 7th Int. Conf.: Audio in Digital Times* (Toronto, Ont., Canada, 1989).

[4] S. P. Lipshitz and J. Vanderkooy, "High-Pass Dither," presented at the AES 4th Regional Convention, Tokyo, 1989; in *Collected Preprints* (AES Japan Section, Tokyo, 1989), pp. 72–75.

[5] J. Vanderkooy and S. P. Lipshitz, "Resolution Below the Least Significant Bit in Digital Systems with Dither," *J. Audio Eng. Soc.*, vol. 32, pp. 106–113 (1984 Mar.); correction *ibid.*, p. 889 (1984 Nov.).

[6] J. Vanderkooy and S. P. Lipshitz, "Dither in Digital Audio," *J. Audio Eng. Soc.*, vol. 35, pp. 966–975 (1987 Dec.).

[7] S. P. Lipshitz and J. Vanderkooy, "Digital Dither," presented at the 81st Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 34, p. 1030 (1986 Dec.), preprint 2412.

[8] M. A. Gerzon and P. G. Craven, "Optimal Noise Shaping and Dither of Digital Signals," presented at the 87th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 37, p. 1072 (1989 Dec.), preprint 2822.

[9] R. A. Wannamaker, "Psycho-Acoustically Op-

timal Noise-Shaping," presented at the 89th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 38, p. 871 (1990 Nov.), preprint 2965.

[10] L. D. Fielder, private communication, 1988 Mar.

[11] ISO 131-1979(E) "Acoustics—Expression of Physical and Subjective Magnitudes of Sound or Noise in Air," International Standards Organization, Geneva, Switzerland, 1979.

[12] ISO 454-1975(E) "Acoustics—Relation between Sound Pressure Levels of Narrow Bands of Noise in a Diffuse Field and in a Frontally-Incident Free Field for Equal Loudness," International Standards Organization, Geneva, Switzerland, 1975.

[13] L. D. Fielder, private communication, 1991 May.

ADDENDUM: AN ALTERNATIVE MEASURE OF AUDIBLE IMPROVEMENT

In using the integral of an improved E-weighted spectrum to gauge noise audibility, we make the tacit assumption that the noise signal in question is low level but nonetheless audible in all psychoacoustic critical bands. Subject to this condition, N_w represents a valid measure of its audibility, since the total perceived noise power is equal to the sum of the noise powers in all the critical bands.

It has been pointed out by Fielder [13] that an alternative measure of audible improvement assumes that the noise spectrum is at the threshold of hearing, and hence makes use of the 0-phon loudness curve, instead of the 15-phon curve upon which the improved E-weighting curve is based. At this level, detection of noise requires only that the noise power in some particular critical band be above the threshold, and is independent of whatever (inaudible) noise may be present in other critical bands. Thus if we assume that listening levels are adjusted such that the noise is at the threshold of hearing, we cannot use N_w as a measure of audibility. Instead, we adopt the following algorithm to assess the psychoacoustical difference between two noise spectra.

Consider two identical channels carrying noises with the power spectra of interest. Let each noise signal be multiplied by a different analog gain, whose value we will denote by G_1 for the first channel and G_2 for the second. Let G_1 be increased from zero until the associated noise just reaches the threshold of audibility in some critical band, and let G_2 be similarly adjusted with respect to the noise in the other channel (the two critical bands need not be the same). The ratio of the two gains, G_1/G_2 (or, if expressed in decibels, their difference) can be taken as a measure of relative perceptibility for the two noise signals.

Using this technique to gauge the audible improvement of the type of shaped noise spectra discussed in this paper over an unshaped (flat) noise spectrum, typically yields higher improvement figures than those cited in this paper. For instance, this algorithm yields an improvement of about 21 dB for the spectrum of Fig. 6(c), as compared to 16.04 dB (from Table 3) using the weighted integration method. This means that the gain of a system using such a noise shaper can be set 21 dB higher than that of a system which uses no noise shaping, such that in both cases the noise is just on the verge of becoming audible.

It should be noted that if a high-order noise shaper is designed such that the resulting noise spectrum matches the 0-phon curve (as opposed to the improved E-weighting curve) with a higher degree of accuracy, then this new algorithm yields an improvement figure which is simply the maximum amount by which the value of the unshaped spectrum exceeds that of the shaped spectrum (assuming that units of decibels are used for both). Since the 0-phon curve has a minimum at 3.5 kHz, the value of the shaped noise spectrum at this frequency will directly yield the improvement if the unshaped spectrum resides at 0 dB.

Had the noise-shaper designs presented here been produced by optimization of the shaped noise spectra with respect to this alternative measure of audibility, somewhat greater perceptual improvements (~1.5 dB) could have been realized *as measured by this criterion.* This is not a sufficiently compelling reason to depart from the filter design methods used in this paper, especially since using the new measure of audibility renders uncertain the convergence of some filter design algorithms.

APPENDIX

We list here the noise-shaper coefficient values a_i and b_i for a selection of the optimized shaper designs discussed in the text.

1) Fig. 4(a) (mod E, 3-coefficient FIR):
$a_0 = 1.652, \quad a_1 = -1.049, \quad a_2 = 0.1382$

2) Fig. 4(d) (mod E, 9-coefficient FIR):
$a_0 = 1.662, \quad a_1 = -1.263, \quad a_2 = 0.4827$
$a_3 = -0.2913, \quad a_4 = 0.1268, \quad a_5 = -0.1124$
$a_6 = 0.03252, \quad a_7 = -0.01265, \quad a_8 = -0.03524$

3) Fig. 5(a) (mod E, 3-coefficient IIR):
$a_0 = 1.726, \quad a_1 = -0.7678, \quad b_1 = -0.2709$

4) Fig. 5(d) (mod E, 9-coefficient IIR):
$a_0 = 1.655, \quad a_1 = -1.928 \quad a_2 = 0.3396$
$a_3 = 0.09123, \quad a_4 = -0.04640, \quad b_1 = 0.4056$
$b_2 = 0.3921, \quad b_3 = -0.05994, \quad b_4 = 0.03179$

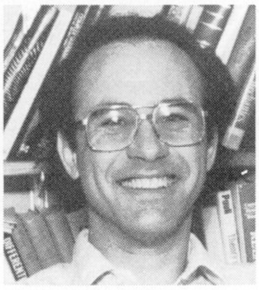
5) Fig. 6(b) (imp E, 5-coefficient FIR):
$a_0 = 2.033, \quad a_1 = -2.165, \quad a_2 = 1.959$
$a_3 = -1.590, \quad a_4 = 0.6149$

6) Fig. 6(d) (imp E, 9-coefficient FIR):
$a_0 = 2.847, \quad a_1 = -4.685, \quad a_2 = 6.214$
$a_3 = -7.184, \quad a_4 = 6.639, \quad a_5 = -5.032$
$a_6 = 3.263, \quad a_7 = -1.632, \quad a_8 = 0.4191$

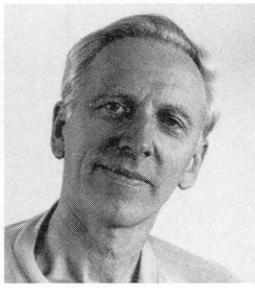
7) Fig. 9(b) (imp E, 30-dB limit, 5-coefficient IIR):
$a_0 = 2.779$, $a_1 = 0.5338$, $a_2 = -0.05967$
$b_1 = -1.814$, $b_2 = -0.8285$

8) Fig. 9(d) (imp E, 30-dB limit, $_4$9-coefficient IIR):
$a_0 = 3.120$, $a_1 = -0.6006$, $a_2 = 1.406$
$a_3 = -1.104$, $a_4 = 0.3365$, $b_1 = -1.643$
$b_2 = -0.7424$, $b_3 = -0.07004$, $b_4 = -0.08775$

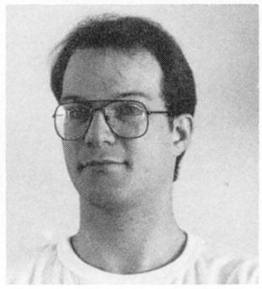
THE AUTHORS



S. P. Lipshitz



J. Vanderkooy



R. Wannamaker

Stanley P. Lipshitz is a professor in both the departments of Applied Mathematics and Physics at the University of Waterloo in Waterloo, Ontario, Canada. There, in addition to the normal teaching associated with a university position, he conducts his own and supervises graduate student research in audio and electroacoustics as a member of the university's Audio Research Group. Prior to joining the faculty of the University of Waterloo in 1970, he studied in South Africa, the country of his birth, where he received his Bachelors (1964), Masters (1966) and Ph.D. (1970) degrees in applied mathematics and physics. He is now a Canadian citizen.

Dr. Lipshitz is a fellow of the AES. He served as a governor of the Society for the period 1984–1986, and as its president for the year 1988–1989. Other society memberships include the IEEE, the Acoustical Society of America, and the Canadian Acoustical Association.

He has presented numerous technical papers at AES conventions, both in North America and overseas, on a wide range of topics including amplifier design, psychoacoustics, loudspeaker crossover design, electroacoustic transducer measurement, acoustics, and digital signal processing for audio. He has participated in many educational workshops and seminars on audio topics including loudspeaker measurement, stereo microphone techniques, and the fundamentals of digital audio. Dr. Lipshitz's publications have appeared frequently in the AES *Journal* and elsewhere. His current research interests include transducer design and measurement, digital signal processing for audio, and the characterization and design of surround-sound systems. He has consulted for a number of companies on audio-related questions.

John Vanderkooy was born in The Netherlands in 1941, but received all of his education in Canada, with a B.Eng. degree in engineering physics in 1963 and a Ph.D. in physics in 1967, both from McMaster University in Hamilton, Ontario. For some years he followed his doctoral interests in low-temperature physics of metals at the University of Waterloo, where he is currently a professor of physics. However, since the late 1970s, his research interests have been mainly in audio and electroacoustics.

A fellow of the AES and a member of the IEEE, Dr. Vanderkooy has contributed a variety of papers at conventions and to the *Journal*. Together with his collegue Stanley Lipshitz and a number of graduate and undergraduate students, they form the Audio Research Group at the University of Waterloo.

Dr. Vanderkooy's current interests are digital audio signal processing, measurement of transfer functions with maximum-length sequences, transducers, diffraction of loudspeaker cabinet edges, and most recently sub-surface analysis techniques using maximum-length sequences.

•

Robert Wannamaker was born in 1967 in Ontario, Canada. In 1988 he received the degree of B.Sc.E. in engineering physics from Queen's University in Kingston, Ontario. He was granted the degree of M.Sc. in physics by the University of Waterloo in 1990 for research on dither and noise-shaping conducted under the aegis of the Audio Research Group. Since graduating, he has continued to work with the ARG as a research assistant. He plays a pretty mean piano and some guitar as well.