J. R. Stuart, R. Hollinshead and M. Capp, "Is High-Frequency Intermodulation Distortion a Significant Factor in High-Resolution Audio?"
J. Audio Eng. Soc., vol. 67, no. 5, pp. 310–318, (2019 May.).
DOI: https://doi.org/10.17743/jaes.2018.0060

Is High-Frequency Intermodulation Distortion a Significant Factor in High-Resolution Audio?

J. ROBERT STUART, *AES Life Fellow*, **RICHARD HOLLINSHEAD,** *AES Member*, **AND**
(jrs@mqa.co.uk)

MICHAEL CAPP, *AES Member*

MQA Ltd., Huntingdon, PE29 6YE, UK

Intermodulation distortion arises when a non-linearity causes two or more signals to interact. We investigated this distortion mechanism by measurement and listening tests using three models of high-quality loudspeaker. Our aim was to discover whether intermodulation distortion of ultrasonic (i.e., above 20 kHz) signal elements could lead to their detectability and confound listening tests or otherwise modify the listening experience. The paper concludes that while such distortion can be found and must be accounted for in some psychoacoustic threshold experiments, it is not pertinent to playback of current high-resolution recordings.

0 INTRODUCTION

High-resolution recording and playback systems should accept signal components above 20 kHz and preferably reproduce above 40 kHz [1, 2].

Considering the frequency domain, it has been questioned whether a listening preference for systems with wider bandwidth could result from the reproduction of signal frequencies above 20 kHz, or alternatively, whether it might arise as a side-effect of filtering in the chain, such as may be encountered when constraining bandwidth to meet a Nyquist criterion within the signal chain. A third, frequency-domain hypothesis, suggests that wider-band signals may cause misbehavior in playback systems.[1]

Amplifiers and loudspeakers tend to exhibit non-linearity giving rise to measurable distortion, e.g., the addition of harmonics to a single applied tone. Alternatively, if the input signal has more than one tone, the output may contain additional frequencies resulting from intermodulation.

This paper examines the general nature of these distortions and considers the circumstances where this mechanism could confound listening tests; we illustrate with measurement and listening experiments on three playback systems.

[1] "The wider we open the window, the more dirt flies in"; attributed to Peter Eckersley, founding chief engineer of the BBC [3, 4].

1 BACKGROUND

Earlier studies noted that intermodulation of higher-frequency signals may give rise to distortion components at lower, more-audible frequencies. Some considered the simple case where all stimuli were contained in the signal [5], while others [6, 7] addressed whether in-band signal components could intermodulate with aliased components arising from digital sampling.

In one experiment, using specific tone complexes above 20 kHz, Ashihara found that intermodulation within a single loudspeaker could produce misleading results that did not arise if the test frequencies were reproduced by separate transducers [8]. This is discussed later in Sec. 7.1.

In listening experiments described by the present authors [9], the addition of specific low-pass filters to a wide-band playback system could be detected, even though the filters attenuated spectral components above the normally accepted range of human sensitivity to pure tones, specifically above 22 kHz. Having shown that the filters could be detected, we decided to investigate whether nonlinear distortion might be a contributing factor to that result.

2 THIS INVESTIGATION

We investigated intermodulation distortion (IMD) arising in combinations of amplifiers and loudspeakers, concentrating particularly on circumstances where very high frequency or ultrasonic signals might induce audible distortions.

Our approach was:

i. To make measurements of such distortions;
ii. To find signal combinations that might yield audible results and measure their levels;
iii. To determine whether such triggers might be present in high-resolution recordings;
iv. To determine whether such distortion exists in a context where it could be detected;
v. To carry out detailed measurement and listening tests on three models of loudspeaker;
vi. To produce a simple model to crosscheck the assumptions.

2.1 Intermodulation Distortion

Non-linear distortion will arise in a playback chain and is unavoidable in analog components such as amplifiers or transducers. Typically, these non-linearities will be greater at the extremes of the working-frequency range and, in well-designed systems, distortion will tend to be higher in transducers (operating without the benefit of negative feedback), rather than in the electronics or a digital channel.

If the principle non-linearity is in a loudspeaker drive-unit, arising for example from non-uniform compliance or magnetic-field effects in the motor, then the system transfer function will be monotonic.[2] Although a simplification (see Sec. 6), such a transfer function can be modeled by a polynomial relating, for example, an amplifier voltage output V_{out} to input V_{in} as follows:

$$V_{out} = K_0 + K_1 V_{in} + K_2 {V_{in}}^2 + K_3 {V_{in}}^3 \ldots \quad (1)$$

When the input signal contains several frequencies, they will interact if there is a non-linearity. If we apply tones of equal amplitude β at frequencies ω_1 and ω_2, such that:

$$V_{in} = \beta \cos(\omega_1 t) + \beta \cos(\omega_2 t) \quad (2)$$

then second-order intermodulation arises from the $K_2 {V_{in}}^2$ term and manifests as outputs at $(\omega_2 \pm \omega_1)$.

Expansion of the 2nd-order term gives an expression for the sum and difference frequencies:

$$Level = K_2 \beta^2 (\cos(\omega_1 - \omega_2)t + \cos(\omega_1 + \omega_2)t) \quad (3)$$

from which we see the general property that the quantity of 2nd-order difference IMD is proportional to the square of the input amplitude of each tone.

Similarly, IMD arising from a third-order nonlinearity, the $K_3 {V_{in}}^3$ term, will give distortion products in magnitude proportional to the cube of the input and at $(2\omega_1 \pm \omega_2)$ and $(2\omega_2 \pm \omega_1)$, some of which may be in the audio band.

With higher orders, the frequency combinations increase but the distortion level for each intermodulation product is proportional to the order of the underlying nonlinearity.

[2] Non-monotonic transfer function effects can arise at the micro level, e.g., stiction, in magnetic domains; in a digital channel (quantization); in electronics as cross-over distortion, thermal envelope or clipping.

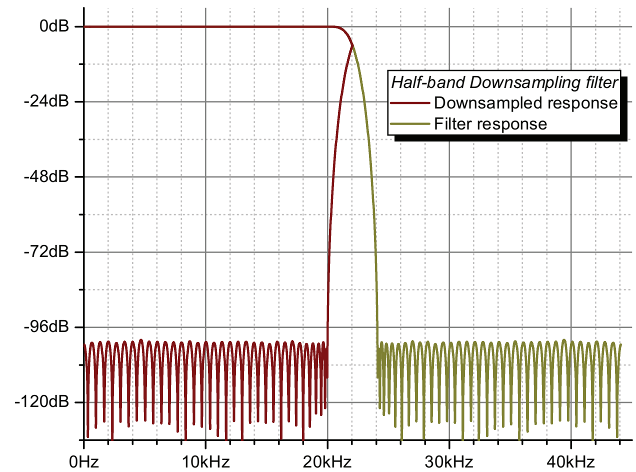

Fig. 1. The olive curve shows the response of a typical half-band filter, commonly found in A/D and D/A converters, in a 44.1 kHz channel. At Nyquist (22.05 kHz), the response is –6 dB (rather than $-\infty$). The wine curve shows how signals above 22.05 kHz alias down into the audio band.

2.2 Selecting Measurement Signals

To investigate IMD, since humans are most sensitive to sounds around 4 kHz and 12 kHz, we first considered combinations of tones above 20 kHz that could produce "downward IMD" (difference products) to these target frequencies—specifically pairs of tones, 24 kHz and above, separated by either 4 kHz or 12 kHz.

In view of the concerns raised by Black [6], we also considered the possibility of intermodulation between high audio band and near ultrasonic frequencies. Guided by typical filters, see Fig. 1, we included stimuli in the range 18–40 kHz; 18 kHz plausibly emulating the lowest significant downward alias of tones above 22 kHz.

2.3 Loudspeaker Systems Tested

Three models of loudspeakers capable of output above 20 kHz were used in this investigation. These were *Meridian* DSP 7200SE (M) loudspeakers (also used in the previous listening tests [9]) and two models of high-quality passive home loudspeakers, one from *Sonus Faber* (S) and the other by *Dali* (D); the latter employed a ribbon tweeter.

All three systems were pre-scanned for suitability in an anechoic chamber. Fig. 2 shows the response of (M). For (S) and (D) the frequency responses were smooth falling by 5 dB and 6 dB respectively at 40 kHz, relative to 1 kHz.

The passive loudspeakers were driven by a Meridian 557 200W low-distortion analog power amplifier. Since the test signals were digitally generated, this amplifier was fed by a Meridian 818 D/A converter configured for variable output. The frequency response of the D/A and power amplifier driving the passive loudspeakers was flat, falling to –0.8 dB at 40 kHz compared to 1 kHz and distortion was very low at –80 dB for the signals used.

Loudspeaker (M) has built-in DSP crossovers that contributed no detectable distortions in the digital domain. These crossovers, at 200 Hz and 2600 Hz, divided the

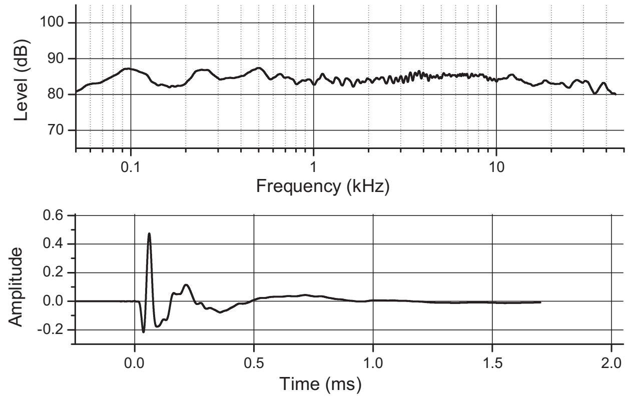

Fig. 2. Frequency (upper panel) and impulse (lower panel) response of the DSP loudspeaker (M) measured at 2 m in semi-anechoic conditions.

incoming signal into three bands, each feeding a D/A converter operating at the signal rate and driving separate 140 W linear analog power amplifiers, directly coupled to the tweeter, midrange, and bass drive units. The power amplifier topology provided high linearity; total harmonic distortion measured below –80 dB up to 100 kHz.

3 TESTS FOR DISTORTION

Each loudspeaker was measured in an anechoic chamber. The test stimuli were generated by an *Audio Precision System 2* that also received the output of a *B&K* microphone preamplifier connected to a *B&K* 4133 microphone placed 1 m in front of the loudspeaker on the tweeter axis. The measuring system was calibrated using a *B&K* 1-kHz 94-dBSPL reference.

In the previous listening tests [9] the gain control of the DSP loudspeaker was set to volume-number 75, giving an acoustic gain of 102 dBSPL in anechoic conditions at 1 m for 0 dBFS input. These experiments used the same gain, confirmed using a –20 dBFS 1-octave band of pink noise centered at 1 kHz. The same acoustic gain was calibrated for loudspeakers (S) and (D) by adjusting the 818 D/A converter output level.

3.1 Initial Distortion Measurements

Initial tests were performed over a wide range of levels, frequencies, and spacing to discover the degree to which twin-tone IMD could be measured and to be certain that 4- and 12-kHz spacings were a suitable choice.

We confirmed that IMD products could be seen at higher SPL on all models and determined that the principal non-linearity was in the drive units (tweeter).

Fig. 3 shows measurements of the power amplifier when driving loudspeaker (D) with a (24+28) kHz tone pair. Even at the highest level (combined stimulus of –10 dBFS), the 4-kHz IMD component is 105 dB below the total signal. In contrast, from Table 1, showing SPL measurements of the loudspeakers (as described in Sec. 3.2), we can see that the loudspeaker produced an IMD component at a relative

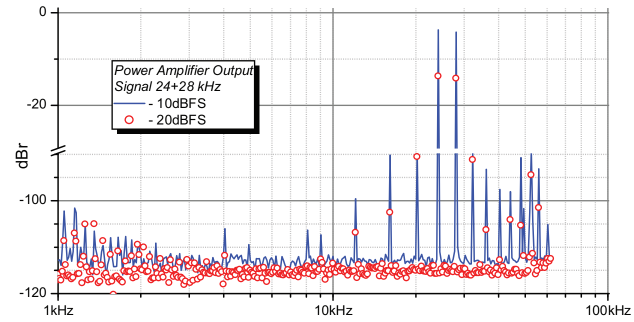

Fig. 3. Showing two measurements of the power amplifier output when driving the 24+28 kHz test signal into loudspeaker (D) with the –10 and –20 dBFS signals.

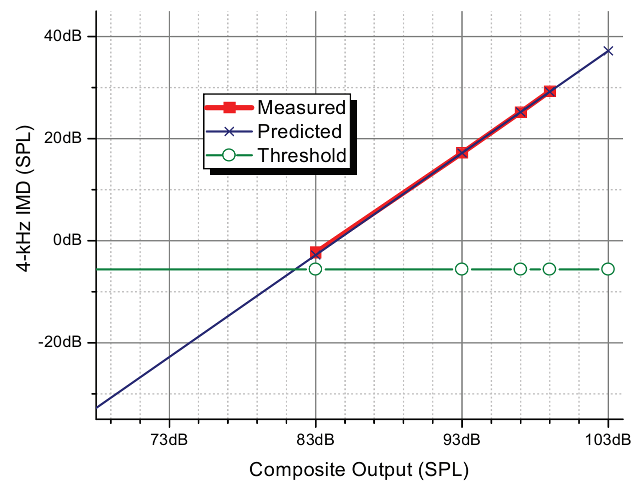

Fig. 4. Showing (red) the measured levels of the 4-kHz 2nd-order difference component for applied (18+22) kHz tones varied with input level when using the DSP loudspeaker. Also shown (blue) is the predicted level for a pure 2nd-order non-linearity. At 1 m, the 4 kHz component falls below the standard ISO 226 threshold (green with circles) with input below –22 dBFS (and the output below 80 dBSPL) [13].

level of –75 dB, confirming that the loudspeaker was the primary source of IMD.

Fig. 4 shows the result of an exploratory test in which 4-kHz IMD can be found in response to an (18+22) kHz tone-pair with loudspeaker (M). It can be seen in the graph that the $(f_2 - f_1)$ intermodulation component (red) falls in level by 2 dB for each 1 dB reduction in the main signal, indicating a 2nd-order non-linearity (see Eq. (3)). Such non-linearity on a single 93 dBSPL tone would give rise to a pure 2nd-harmonic distortion of 0.04%.

3.2 Test 1: Twin-Tone Measurement

We made spectral measurements for all loudspeakers in response to tone pairs spaced by 4 kHz or 12 kHz in the range 18–40 kHz for input levels between –30 and –6 dBFS (resulting output between 72 and 96 dBSPL at 1 m).

The measured output and $(f_2 - f_1)$ IMD products are shown in Table 1.

Table 1. The table entries show the measured SPL at 1 m for two levels of stimulus: –20 dBFS and –10 dBFS (the input level of each tone of the pair being 3 dB below the total level). For each twin-tone test the lower and upper frequencies are shown. Measured results are given for the total level and IMD product, where measurable. Shaded entries are for the 12-kHz difference pairs. Numbers in bold are referenced in the text.

	Input Frequency (kHz)		DSP Loudspeaker M		Loudspeaker S		Loudspeaker D	
Stimulus Level (dBFS)	Lower	Upper	Level (dBSPL)	IMD (dBSPL)	Level (dBSPL)	IMD (dBSPL)	Level (dBSPL)	IMD (dBSPL)
–10	18	22	90	17	93	19	98	24
	20	24	90	16	92	17	94	22
	24	28	90	16	90	9	**96**	**21**
		36	88	26	87	30	94	36
	28	32	**91**	**25**	86	9	93	19
		40	88	29	83	17	91	23
	32	36	89	16	82	14	90	16
	36	40	87	13	75	7	87	< –1
–20	18	22	80	< –1	83	< –1	88	5
	20	24	80	< –1	82	< –1	88	2
	24	28	80	< –1	70	2	86	< –1
		36	**78**	**15**	77	11	84	25
	28	32	**80**	**7**	77	< –1	83	< –1
		40	78	13	74	8	81	19
	32	36	79	8	72	1	80	< –1
	36	40	77	7	66	< –1	77	< –1

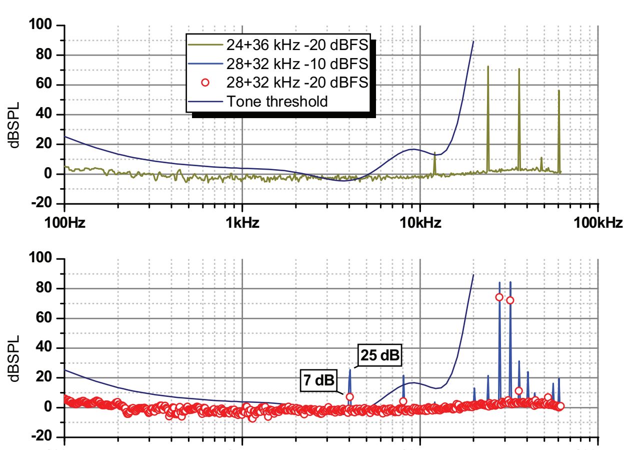

Fig. 5. Measurements of the DSP loudspeaker in response to a twin-tone stimulus. Upper: (24+36) kHz at –20 dBFS, showing the 12-kHz intermodulation component. Lower: (28+32) kHz at two levels. Note that the 4-kHz intermodulation component falls from 25 to 7 dBSPL as the output signals are lowered from 91 to 80 dBSPL. All measurements made at 1 m in anechoic conditions. Both graphs show a normal hearing threshold for pure tones superimposed.

For loudspeaker (M), Fig. 5 (upper) gives an example where a pair of signals at (24+36) kHz, reproduced at 78 dBSPL, introduced a 12-kHz component at 15 dBSPL, which is very close to the nominal threshold of hearing.

For the same loudspeaker, Fig. 5 (lower) shows a (28+32) kHz signal pair reproduced at two levels, 91 and 80 dBSPL. Here 4-kHz IMD products can be seen at 25 and 7 dBSPL respectively, both of which should be audible. In this test we can see other distortion products, notably a similar level at 8 kHz (which results from a 4th-order term).

4 LISTENING TESTS

This section describes aspects of the experimental design common to all listening experiments.

Five male listeners aged between 22 and 69 participated; all were audio engineers by profession and reported normal hearing, although this was not tested formally. Subjects sat in a soundproof listening room approximately 2 m from the loudspeakers (see Figs. 6 and 7).

Signals were presented to a pair of loudspeakers and each test was repeated for the three pairs of loudspeakers described earlier. The playback gain was the same as that in our earlier listening test [9], which had been chosen for comfort, yet high enough for music to be played at a realistic level. Gain was confirmed at 103 dBSPL (for a pair) by measuring at the listening position with the calibration signal described in Sec. 3 fed to both loudspeakers.

An unusual feature of this test was that in many cases, there was nothing to be heard. The goal was to determine the existence of audible components and in very few cases to determine a behavioral threshold. Generally, it was evident

Fig. 6. The listening room.

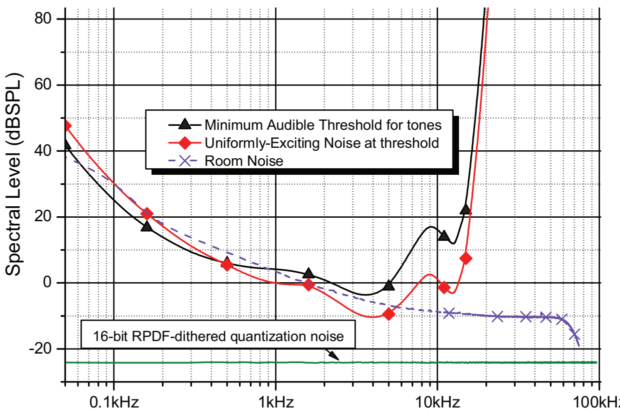

Fig. 7. Showing the measured noise-level in the listening room, along with auditory thresholds, (using 46.88 Hz analysis bandwidth), see [9].

whether an IMD component was audible or not. In cases of doubt we switched to testing single-blind or using level up/down adjustment.

4.1 Test 2: Twin-Tone Listening

Test 2 used the equal-amplitude (24+28) kHz twin-tone signal at various input levels. Since the stimulus was exclusively ultrasonic, listeners should be unable to detect its presence, but with some combinations of loudspeaker and level, intermodulation products might become audible.

Listeners were asked to report detections of *any* sound at the listening position and in some cases were given the opportunity to move closer to the loudspeakers if needed.[3]

[3]Caution was exercised; listeners were not permitted to listen near to the loudspeakers at the highest signal levels. For the highest input (–10 dBFS) we measured 96 dBSPL at 1 m and 118 dBSPL at the tweeter baffle. When the input signal was at –30 dBFS (equivalent to 76 dBSPL at 1 m), although none could hear the signal, three of the listeners were aware of a sensation of "pressure" when close to the loudspeakers.

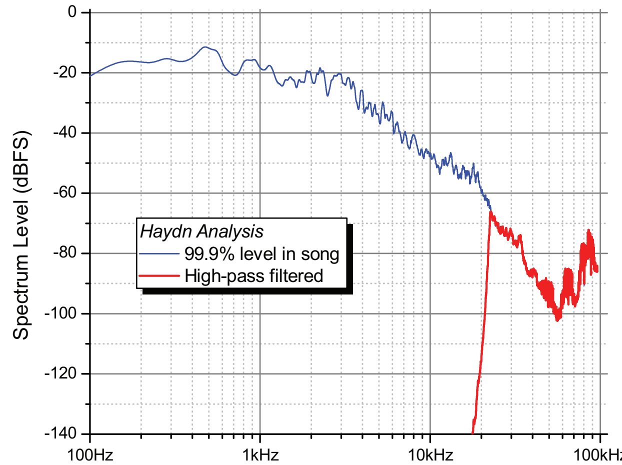

Fig. 8. Showing the 99.9% spectrum level of the Haydn recording before and after high-pass filtering at 23 kHz.

The following observations were made for different signal input levels:

- At –30 dBFS nothing could be heard either at the listening position or with an ear right against the tweeter.
- At –20 dBFS:
  - Loudspeakers M and D: All participants heard a 4-kHz tone at a listening position 2 m away.
  - Passive Loudspeaker S: All participants heard a 4-kHz tone 30 cm away but not at 2 m.
- At –15 and –10 dBFS all participants heard a 4-kHz tone at the listening position.
- Using the DSP loudspeakers and listening very close to the tweeters, the threshold of detection of a 4-kHz IMD product occured at an input level of –25 dBFS.

4.2 Test 3: Listening to HF Content

For this experiment the high-frequency components of three music recordings were isolated in *Adobe Audition* using an 18th order Chebyshev high-pass-filter with a 23-kHz cut-off frequency.

Test 3 used a 192 kHz 24-bit PCM recording of Haydn String Quartet Op.76 No.5 in D "Finale, Presto," released by 2L, that we had used in [9]. Fig. 8 shows the 99.9% spectrum attainment level before and after filtering.[4]

Two other recordings ("Harpsichord" and "Music box") were specifically chosen as outliers, both having been made to maximize levels of ultrasonic content.

The harpsichord recording was made with a 100-kHz bandwidth microphone under the instrument lid, close to the plectra [10]. For the music box recording a similar microphone was placed inside a vintage Polyphone music box, adjacent to the plucked reeds [11, 12]. In Fig. 9 we

[4]A signal will be below the attainment level for the percentage of time indicated, e.g., 99.9% in this case.

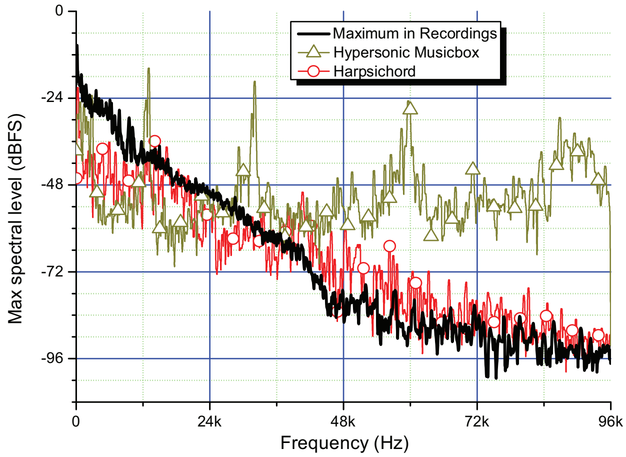

Fig. 9. Spectrum analysis of recordings. In black is the maximum at each frequency of the peak spectra taken over a large corpus of commercial 24-bit recordings. Two deliberately stressful recordings are also included, measured on 99% basis. Shown in red with circles is a harpsichord recording for which a wideband microphone was placed under the instrument lid, near the plectra [10]. The olive curve with triangles shows a recording made with a similar wide-band microphone placed inside a vintage Polyphone music box [11, 12].

compare the spectra of these recordings with those of a large corpus of commercial 24-bit releases.

In this test, the high-pass-filtered extracts were played back at the same acoustic gain and repeated for each of the three pairs of loudspeakers.

1. Haydn: When the high-pass filtered version of the Haydn was played back nothing at all could be heard either in the listening position or with an ear right next to the tweeter.
2. Harpsichord: Nothing could be heard of the harpsichord, except when using Passive Loudspeaker (D). With that loudspeaker, four out of five listeners heard very faint clicks at 30 cm, but these became inaudible with a 6-dB gain reduction even directly listening to the tweeter.
3. Music box: This recording was audible to all participants as a series of very faint clicks. With care, three participants could discern this at the listening position using Passive Loudspeaker (S), and all participants could hear it 30 cm away from all three loudspeaker types.

5 SUMMARY OF RESULTS

The three playback systems showed measurable levels of distortion, originating in the loudspeaker drivers. In the three models tested, the predominantly 2^{nd}-order distortion was equivalent to that producing less than 0.1% total-harmonic distortion at frequencies above 20 kHz.

When a twin-tone (or music signal) is applied, such non-linearity gives rise to measured difference IMD at a level proportional to the square of the signal level.

In an extreme listening test, where the prime stimuli were only ultrasonic, it was possible to detect IMD generated by the system when the test signals comprised a pair of tones spaced by 4 kHz, in the range 22 kHz to 40 kHz, and where the main signal pair was higher than 80 dBSPL at 1 m.

Playing only the ultrasonic part of music recordings at realistic replay gains (see Fig. 10), it was impossible to detect IMD effects at the listening position for either the string quartet or harpsichord music; the latter had a spectrum level above 43 dBSPL up to 50 kHz. With the music box extract, which over the same range was 12 dB higher in level, the IMD products were very faintly heard at the listening position.

It was obvious to all participants that even with the extreme music box test, this very low-level distortion component would be swamped by the baseband noisefloor of the signal, let alone by the (loud) music itself.

6 MODELING DISTORTION

To crosscheck IMD generation in the music extracts and to examine distortion products throughout the audio band, a plug-in was built for *Reaper* that processed audio in accordance with the first seven terms in Eq. (1). By matching the signal levels and adjusting k_n coefficients, a set was found that gave predicted results similar to measurements made on the loudspeakers for the narrow set of signals under consideration. For example, loudspeaker (M) was modeled by:

$$[k_1 \ldots k_7] = [1, -0.008, -0.005, 0.04, -0.45, 0.25, 2.0]$$

In general, a loudspeaker system requires more complex modeling to take account of movement, drift, and temperature shifts; specifically, a transducer's response will include a "memory" of prior signals. In this paper however, we aren't seeking a precise global model of the loudspeaker but an adequate description of intermodulation products on specific signals.

7 DISCUSSION

7.1 Comparison with Other Work

In an earlier set of experiments, Ashihara [8] illustrated the potential for audible IMD when using a baseband signal comprising a 60 dBSPL 2 kHz tone with equal amplitudes of odd harmonics at 6, 10, 14, 18 kHz, in combination with an ultrasonic probe combination at 22, 26, 30, 34, and 38 kHz, together forming a comb for which IMD would result in interstitial harmonic components based on 4 kHz. In those tests, using a single loudspeaker to carry just the lower-frequency set, no distortion components could be detected, which seems to be consistent with our findings at that SPL.

However, when a single loudspeaker carried both signal sets, Ashihara found a detectable difference when the level of the ultrasonic group was adjusted so that each tone was higher than 66 dBSPL at the listening position. In a

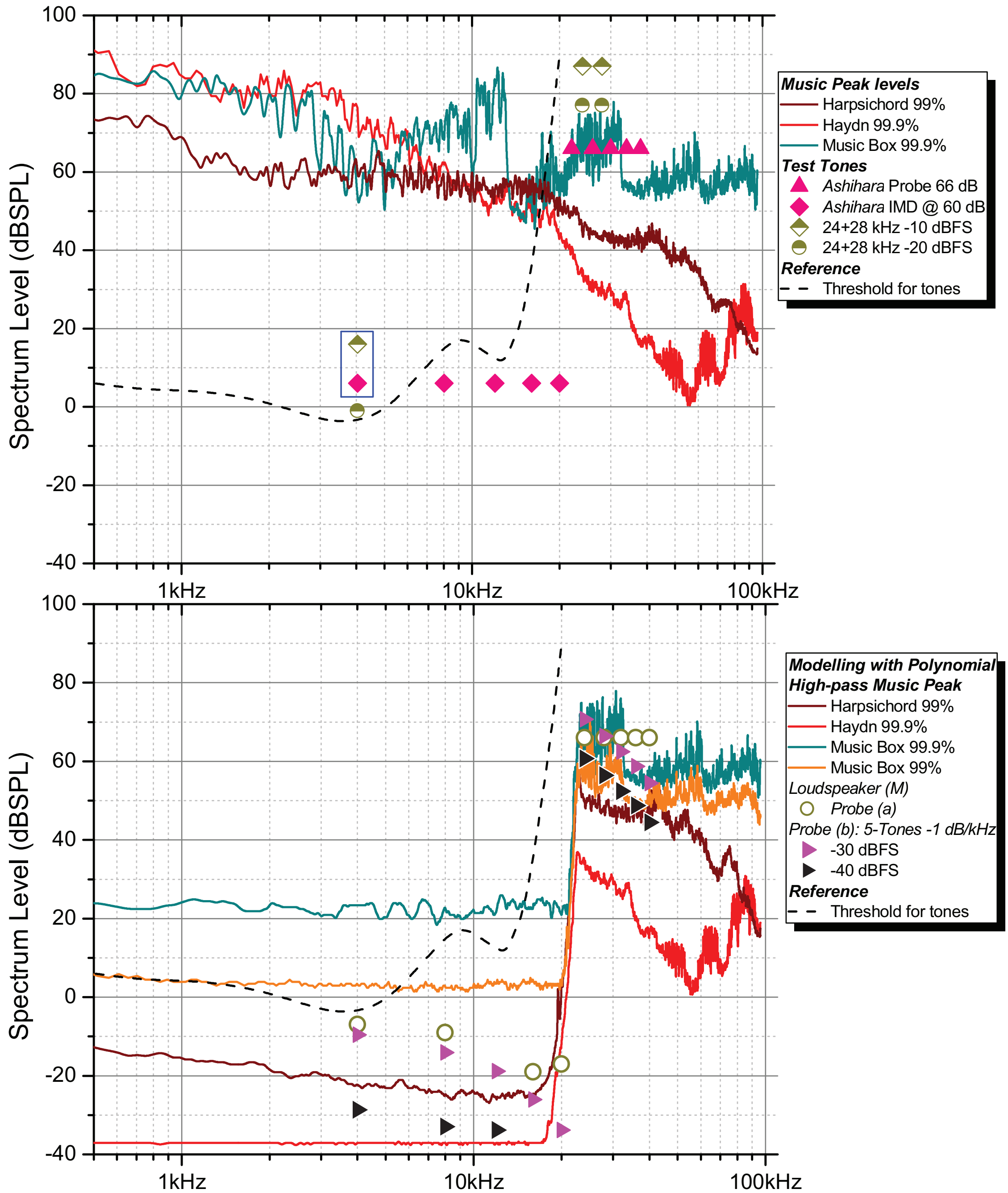

Fig. 10. Summarizing both measurements and simulation results at the listening gain of 103 dBSPL. (upper) (i) Spectrum analysis showing 99% levels for the Harpsichord (wine) and 99.9% levels for the Haydn (red) and Music box (dark cyan). (ii) Measurements with (24+28) kHz twin-tone at –10 and –20 dBFS showing output signals and IMD for loudspeaker (M). (iii) From Ashihara [8]: the 5-tone probe at reported threshold of 66 dBSPL per component (pink triangles); and (transcribed from their Fig. 5) IMD when the 5-tones were at 60 dBFS (pink diamonds). (lower) (iv) Showing the spectra resulting from passing signals through a polynomial modeling loudspeaker (M): (a) 5 tones at 24, 28, 32, 36, and 40 kHz (olive circles); (b) for which the 5-tones followed a –1 dB/kHz trend, with overall levels of –30 dBFS (magenta) and–40 dBFS (black). (v) The spectra resulting from passing the high-pass-filtered music through a polynomial modeling loudspeaker (M), showing 99.9% levels for Haydn (red) and Music box (dark cyan), as well as the 99% level for the Harpsichord (wine) and Music box (orange). Note that in each case the predicted IMD below 23 kHz has a raised but level spectrum. It would not be expected that a tweeter could radiate much below 2 kHz. (vi) Both graphs include the nominal hearing threshold for tones (black dashed line).

crosscheck, it was shown that the presence of these ultrasonic components could not be detected if each of the five tones were reproduced individually from its own loudspeaker, thereby eliminating IMD effects in the signal, the listener, or in air.

With a 2nd-order nonlinearity, five equally spaced harmonics should produce an IMD component 12 dB higher than that produced with just two tones. This can be observed in Fig. 10, where the signal used by Ashihara is indicated as pink triangles (five tones, each at 66 dBSPL), as is the distortion he measured when that signal was 6 dB lower (pink diamonds)[5] along with simulated IMD for loudspeaker (M) using a similar signal set at 24, 28, 32, 36, and 40 kHz (olive circles).

We can infer from Fig. 10 that the IMD component in Ashihara's experiment should be audible; whereas for loudspeaker (M) under the same conditions, IMD should be inaudible. This was confirmed in a follow-up listening test.

7.2. Context

In our listening experiment, when the audio band up to 20 kHz was silent, the (24+28) kHz twin tone signal allowed 4-kHz IMD to become audible when the tone-pair exceeded about 82 dBSPL. The 5-tone set used by Ashihara drives higher levels of downward IMD and in his tests the threshold for IMD audibility occurred at a composite level of about 75 dBSPL. These two results, despite being of different loudspeakers show very good corroboration. However, when put in the context of a loud but not unpleasant playback level, these ultrasonic SPL levels are about 20 dB higher than would be encountered from commercially available music recordings according to a corpus analysis as summarized in Fig. 9.

As a final check, a –40dBFS test signal was generated containing 24, 28, 32, 36, and 40 kHz tones, but in which each successive component was 4 dB lower than the previous, matching the corpus spectral trend of –1 dB/kHz in that frequency range seen in Fig. 9. This set is included (black triangles) in Fig. 10, as is the predicted IMD for loudspeaker (M), which should be inaudible at this level or even 10 dB higher (magenta triangles). This was confirmed in a listening test.

The music signals showed no cases where signal above 23 kHz was not accompanied by higher-level components below 20 kHz, which would therefore be expected to mask distortion products at the scale we measured. This characteristic is to be expected since very-high-frequency elements in music tend to be harmonics of co-temporal audible sounds. Such harmonics also tend to have lower amplitudes than the fundamentals and, unlike our extreme tests, do not exist in isolation above 20 kHz with silence in the audio band. More importantly perhaps, when we look at the instances of highest ultrasonic content such as the music box, the total level of the inband audio is at least 40 dB higher, at a corresponding frequency, than the worst-case measured distortions in our experiments.

Therefore, the fact that intermodulation products could not be heard in Test 3, where the inband frequency component had been removed, strongly suggests a wide safety margin for these playback systems. IMD cannot therefore be a contributing factor to false-positive detection of components above 23 kHz when listening to normal music.

The measurement of intermodulation in Tests 1 and 2 also confirmed a wide safety margin.

[5]We can safely predict that at the higher 66-dBSPL-per-component level, the IMD should increase by 6–12 dB; highlighted in the blue box in Fig. 10 (upper), where the pink diamond would move close to the olive diamond.

8 CONCLUSIONS

Using selected ultrasonic signals higher than 70 dBSPL, IMD could be measured in all three loudspeaker systems tested and it was just audible in the absence of any signals below 20 kHz. Care should therefore be exercised when designing experiments to determine hearing thresholds for combinations of high-frequency tones in the absence of correlated sounds below 20 kHz.

For a large corpus of commercially available recordings, where the peak playback signal spectrum level above 23 kHz will be lower than 40 dBSPL, there is a very wide safety margin between any resulting IMD and the high masking level of the baseband music itself.

The three tests described here also suggest that, for the music and system used in the listening tests [9], intermodulation distortion would not have been a confounding factor—answering one of the authors' original questions.

Although more equipment characterization would be needed if considering headphone listening, the spectral levels found in actual recordings suggest that for amplifiers and loudspeakers intended for high-quality applications and, specifically where electronic non-linearity is below –60 dB, IMD is very unlikely to be detectable.

9 AKNOWLEDGMENTS

This work was stimulated by questions raised by Brian Moore when reviewing work continuing from [9]. Originally intended as an appendix thereto, this work was cast into a stand-alone paper and the experimental work extended to include passive loudspeakers. The authors are grateful for early informal reviewing comments provided by Joshua Reiss. The authors are grateful to Cosmin Frateanu, Haydon Cardew, and Daniel Wolff from MQA.

10 REFERENCES

[1] ADA, "Proposal of Desirable Requirements for the Next Generation's Digital Audio," *Advanced Digital Audio Conference*, Japan Audio Society (1996 Apr.).

[2] JEITA, "On the Designation of High-Res Audio," 25 JEITA-CP No. 42 (2014 Mar.).

[3] C. McKenzie, quoting Eckersley, P. P., *The Gramophone*, vol. 50, p. 558 (1972 Sep.).

[4] P. P. Eckersley, *The Power Behind the Microphone* (London: Jonathan Cape, 1941).

[5] K. Ashihara and K. Shugo, "Influence of Expanded Frequency Band of Signals on Non-Linear Characteristics of Loudspeakers," *ASJ*, vol. 56, no. 8, pp. 549–555 (2000 Aug.).

[6] R. Black, "Anti-Alias Filters—The Invisible Distortion Mechanism in Digital Audio?" presented at the *106th Convention of the Audio Engineering Society* (1999 May), convention paper 4966. http://www.aes.org/e-lib/browse.cfm?elib=8214

[7] J. Dunn, "Anti-Alias and Anti-Image Filtering: The Benefits of 96 kHz Sampling Rate Formats for Those Who Cannot Hear above 20 kHz," presented at the *104th Convention of the Audio Engineering Society* (1998 May), convention paper 4734. http://www.aes.org/e-lib/browse.cfm?elib=8446

[8] K. Ashihara, "Audibility of Complex Tones above 20 kHz," *29th Int. Cong. and Exhibition on Noise Control Engineering (InterNoise)* (2000).

[9] H. M. Jackson, M. D. Capp, and J. R. Stuart, "The Audibility of Typical Digital Audio Filters in a High-Fidelity Playback System," presented at the *137th Convention of the Audio Engineering Society* (2014 Oct.), convention paper 9174. http://www.aes.org/e-lib/browse.cfm?elib=17497

[10] W. Woszczyk and G. Massenburg, McGill University, Montreal, Canada. Private communication (2016 May).

[11] http://www.e-onkyo.com/music/album/hsh001/

[12] T. Harada, S. Ito, et al., "Effect of High-Resolution Audio Music Box Sound on EEG," *International Music J.*, vol. 23, no. 1, pp. 1–3 (2016 Apr.). https://www.researchgate.net/publication/299410459_Effect_of_High-Resolution_Audio_Music_Box_Sound_on_EEG

[13] ISO, "Acoustics–Normal Equal-Loudness Contours," ISO 226 (1987).

THE AUTHORS

J. Robert (Bob) Stuart

Richard Hollinshead

Michael Capp

J. Robert (Bob) Stuart was born in 1948. He studied electronic engineering and acoustics at the University of Birmingham and took an M.Sc. in operations research at Imperial College, London. While at Birmingham he studied psychoacoustics under Professor Jack Allison, which began a lifelong fascination with the subject. In 1977 he co-founded Meridian Audio and served as CTO until early 2015. In 2014 he founded MQA Ltd. where he is currently full time as Chairman and CTO. At the request of Hiro Negishi and Raymond Cooke, Bob chaired the advocacy group Acoustic Renaissance for Audio between 1994 and 2002. In the 1990s he worked with Michael Gerzon and Peter Craven on lossless compression and was instrumental in its adoption for optical discs. Bob has contributed to DVD-Audio and BluRay standards and has served on the technical committees of the National Sound Archive, JAS and the ADA (Japan). Bob's professional interests are the furthering of analog and digital audio and developing understanding of human auditory perception mechanisms relevant to live and recorded music. His specialties include the auditory sciences and the design of analog and digital electronics, loudspeakers, audio coding, and signal processing. Bob joined AES in 1971, has been a fellow since 1992, and is a member of ASA, IEEE, and the Hearing Group at Cambridge. Bob has a deep interest in music and spends a good deal of time listening to live and recorded material.

•

Richard Hollinshead is a consultant currently contracted to MQA Ltd. Having studied natural sciences at Cambridge University he later completed a Postgraduate Diploma in software development at the Open University. From 2007 he was Director of Engineering and then CTO at Meridian Audio Ltd., leaving in 2017 to become an independent consultant. He led the development of numerous products including DSP and active loudspeakers and digital to analog converters. From 2014 he was also responsible for the development of Meridian's audio technology available in McLaren and Jaguar Land Rover vehicles. Richard holds patents for digital audio encoding and group delay correction. His current research interests include spatial audio, loudspeaker crossover algorithms, and high performance analog circuity. He has been an AES member since 1985.

•

Michael Capp is currently Research Manager at MQA Ltd. He received a B.Sc. in maths and physics from the University of East Anglia and completed a Ph.D. in image and audio signal processing with Leicester University, while based at the University of Northampton under the supervision of Prof. Phil Picton. He spent a year at the Open University in Milton Keynes as a research fellow, investigating the application of Ambisonic sound in virtual environments, before joining Meridian Audio Ltd. in 2001. In 2015 he moved to MQA Ltd. to continue in the development of the MQA technology. Michael has worked on numerous research projects, including MLP and Dolby TrueHD, as well as the algorithmic development of Room Correction. He also co-invented the Meridian Enhanced Bass Alignment, Enhanced Boundary Control, and Centre Elevation algorithms, and has acted as industrial supervisor for students from the Universities of York, Surrey, Southampton, and Queen Mary, London. Dr. Capp is a member of the AES, the IEEE, and the IOP.