Reproducing Low-Pitched Signals through Small Loudspeakers*

ERIK LARSEN, *AES Member,* **AND RONALD M. AARTS,** *AES Fellow*

Philips Research Laboratories, 5656AA Eindhoven, The Netherlands

Ever since the invention of the electrodynamic loudspeaker there has been a need for greater acoustical output, especially at low frequencies. From a manufacturer's point of view it has been desirable for a long time to reduce the size of the loudspeaker (and cabinet). These two demands are physically contradictory. Options are being offered to evoke the illusion of a higher low-frequency response of the loudspeaker while the power radiated by the loudspeaker at those low frequencies remains the same, or is even lower. This is feasible by exploiting certain psychoacoustic phenomena. The required nonlinear signal processing is studied for a number of specific implementations. An elaborate analysis of the outcome of a listening test, aimed at assessing the subjective evaluation of the system presented, employing multidimensional scaling and biplots, is also presented.

0 INTRODUCTION

In many sound reproduction applications it is not possible to use large loudspeakers because of size or cost constraints. Typical applications are portable audio, multimedia, TV, and public address systems, to name just a few. Hence the devices are often small in size, and therefore the transducers are inherently small as well. Needless to say, the competitive market just mentioned also dictates the highest possible audio quality of these products. However, probably the most well-known characteristic of small loudspeakers is a poor low-frequency (bass) response. In practice this means that a significant portion of the audio signal may not be reproduced (sufficiently) by the loudspeaker. For loudspeakers used in such applications reproduction below 100 Hz is usually negligible, whereas in some applications this lower limit can easily be as high as several hundred hertz. The bass portion of an audio signal contributes significantly to the sound "impact," and depending on the bass quality, the overall sound quality will shift up or down. Therefore a good low-frequency reproduction is essential.

A traditional and conceptually very simple method to increase the perceived sound level in the lower part of the audible spectrum (below the loudspeaker's resonance frequency, which is usually the lower limit) is to amplify the low-frequency part of the audio spectrum by a fixed or a dynamic amount (depending on the signal amplitude or the reproduction level). A special system that purposefully drives a loudspeaker below resonance is the ELF system, described in Long and Wickersham [1], [2]. From an efficiency point of view these methods are unfavorable, but an even more serious problem is the high cone excursion at low frequencies (quadrupling for every octave down in frequency). For very low frequencies, the mechanical limits of the loudspeaker will limit the stroke the cone can make, leading to distortion and possibly loudspeaker overload. Thus increasing the radiated sound pressure level physically means forcing the loudspeaker to radiate sound in a frequency range for which it is not equipped. It may be better to prevent this completely by using methods to be outlined in this paper. In the process we shall discover several advantages of these methods.

Because the radiation characteristics of (small) loudspeakers are at the core of the topic to be discussed, we shall review these characteristics in terms of the loudspeaker parameters. Because this is fairly well-known material (see, for example, Beranek [3], Olson [4], or Borwick [5] for extensive reviews), this will be discussed in Appendix 1. We will show that in the normal operating range of the loudspeaker (above resonance and below the transition frequency, where the wavelength becomes roughly equal to the cone diameter), the efficiency η is

* Parts of this paper were presented at the 108th Convention of the Audio Engineering Society, Paris, France, 2000 February 19–22, under the title, "Perceiving Low Pitch through Small Loudspeakers." Manuscript received 2001 October 23.

proportional to

$$\eta \propto \frac{S^2}{m_t^2} \tag{1}$$

where S is the cone area and m_t the total moving mass. So a high efficiency requires a large cone area and a small moving mass, which are difficult to realize simultaneously. The resonance frequency ω_0 of the loudspeaker, which is just a simple mass–spring system, equals

$$\omega_0 = \sqrt{\frac{k_t}{m_t}} \tag{2}$$

where k_t is the total spring constant (cone suspension and influence of enclosed air volume), and is in part inversely proportional to the cabinet volume. Any attempt to lower the resonance frequency by increasing m_t will decrease the efficiency, unless the cabinet size or the cone area is increased as well, but this is not possible for a small loudspeaker. More intricate cabinet designs employing (multiple) ports or passive radiators behave somewhat differently in the low-frequency range, but do not overcome the basic problems of good low-frequency reproduction. It is precisely Eqs. (1) and (2) which, for a small loudspeaker, prevent a good and extended low-frequency response—good in the sense of a high radiated sound pressure level and a high efficiency. Obviously all the qualifications of the previous sentence are relative, but they can be seen in light of the applications mentioned and of what is expected by a "demanding" listener.

Now from psychoacoustic theory we know that a pitch perception can occur at a frequency that is not contained in the audio signal. This is possible through nonlinearities in the cochlea (difference tones) or a higher level neural effect in the auditory system (virtual pitch). In Section 1 we review these two effects, which appear to be very suitable effects for our purpose of enhancing bass perception using small loudspeakers. Basically this works by some simple nonlinear (but controlled) processing, replacing very low frequencies in the audio signal by higher frequencies. These will still have the same perceived pitch of the original, using the psychoacoustic effects mentioned previously. Such effects also occurred in transistor radios, where undesired nonlinearities gave rise to a distorted sound. However, the method that we now propose uses nonlinearities in a controlled manner, and restricted to only the lowest frequencies, such that the effect is to our benefit. Without any information about the signal processing employed, we can immediately infer a number of advantages that such a scheme will provide:

1) High radiated sound pressure level because of increased efficiency and decreased cone excursion. Furthermore, at higher frequencies the auditory system is more sensitive, which will also contribute to increased loudness.

2) Less power consumption because of increased efficiency. This can be very important for portable applications.

3) Less disturbance in neighboring areas because of increased absorption of high frequencies in structures.

We propose a signal-processing scheme consisting of a few basic and efficient operations to achieve such a psychoacoustic bass enhancement (see Fig. 1). This method was already introduced by Aarts and coworkers [6]–[8]. The concept of this processing was already mentioned: to prevent the radiation of very low frequencies by radiating higher frequency components instead, leading to the same pitch. Preventing radiation of very low frequencies is done by high-pass filtering the direct signal path through filters HFIL (see Fig. 1). To compensate for this, the input signals are summed and the low-frequency part is obtained through filtering by FIL1. The next step is to create higher harmonics of this bass portion by processing through a nonlinear device (NLD). These harmonics are filtered once again by FIL2 to obtain a suitable spectrum, amplified with a gain of G and added to the left and right output signals. The resulting output signals now do not contain low-frequency elements, but have the low pitch associated with the input (given that the input signals contain low frequencies). Note that the processing as outlined may also be performed for each channel separately, if so desired. As the low-frequency content is generally identical in the left and right signal channels, this is usually not necessary. Furthermore, localization for low frequencies is known to be very poor, so a change in the low-frequency balance should not be detected. In Section 2 we will discuss in more detail the important aspects of the processing scheme proposed in Fig. 1.

As is apparent now, we are focusing on time-domain instead of frequency-domain methods. In the frequency domain we would have the familiar problems such as the connection of consecutive output frames, spectral leakage for frequencies which are not harmonic to the DFT window, and nonstationarity of the input signal in any one frame. By focusing on the time domain we circumvent these problems. Furthermore, in the time domain we can make implementations which are much more efficient computationally, and we can even devise simple analog

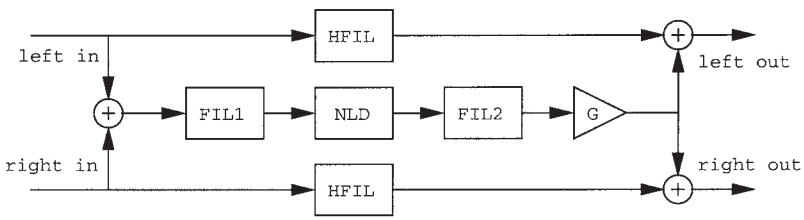


Fig. 1. Signal processing for psychoacoustic bass enhancement. Input signal is summed and filtered to obtain bass portion, then harmonics are created and added to left and right output signals. In direct path a high-pass filter is implemented.

circuitry to do the processing.

About 50 years ago, a circuit based on the same concept was announced in *Radio Electronics* [9]. The circuit was used in a radio receiver and through the nonlinearity of the tubes created odd harmonics of the low-frequency part of the spectrum. The concept has also been discussed by Ben-Tzur and Colloms [10]. However, the focus of that discussion is on achieving equal loudness of the output signal containing the synthetic bass, with respect to the input signal. We will comment on this aspect in Section 1, but in this paper we shall focus more on creating the synthetic bass itself.

Through many informal listening tests and demonstrations of the system it was soon obvious that our method is well appreciated. To show this more formally, and also to be able to compare various implementations in a reliable way, a listening test was conducted. This test will be described and the results presented in Section 3. (Raw data from the listening test are given in Appendix 4.) Part of the analysis of the listening test results will involve the use of biplots. A short introduction to this visualization technique will be presented in Appendix 5. Details of the reproduction device used in the test and the exact design of the filters used in the algorithm are given in Appendix 3.

In Appendix 1 we review the efficiency of an electrodynamic loudspeaker in a closed cabinet. We will treat the last part of our work in Appendix 2. This concerns the key element of the processing scheme shown in Fig. 1—the nonlinear device where the harmonics are created. To analyze what exactly happens to the signal spectrum here is a difficult task. Because of the nonlinear behavior, the operation performed on the signal cannot be represented by a transfer function. Every single frequency component in the output spectrum depends on all frequency components of the input spectrum. However, we have devised a compact formulation for representing the output spectrum in terms of the input spectrum, for all possible signals, in continuous and discrete time.

1 LOW PITCH IN THE ABSENCE OF LOW FREQUENCIES

There are several options to increase the perception of low pitch based on psychoacoustic events. In this section we discuss three such options.

1.1 Frequency Doubling

In previous studies [7], [11]–[13] a frequency doubler was introduced, whereby signals in the low-frequency band also appear one octave higher. It can be considered as one of the options for increased low-pitch perception. The method mentioned, shown in Fig. 2(a), has extreme simplicity as an advantage. By means of a simple nonlinear element, such as a diode, frequencies between, say, $\frac{1}{2}f_0$ and f_0, the frequency region where the loudspeaker does not radiate sufficient power, are shifted to $f_0 - 2f_0$. A drawback is that the pitch has been changed. Furthermore, impulsive sounds with a high low-frequency content are seriously distorted. Nevertheless, informal listening tests have shown that this frequency doubling can be an improvement, and an implementation according to this concept is also included in the listening test described in Section 3.

1.2 Virtual Pitch

Pitch is a subjective, psychophysical quantity. According to the American Standards Association pitch is "that attribute of an auditory sensation in terms of which sounds may be ordered on a scale extending from low to high." For a pure tone, where the fundamental frequency corresponds to the frequency of the tone, the pitch is unambiguous and —if we neglect the influence of sound level on pitch— one can identify pitch with the frequency of the pure tone. For a complex tone, consisting of more than one frequency, the situation is more complicated. Pitch should then be measured by psychophysical experiments. A pitch that is produced by a set of frequency components [see Fig. 2(b)] rather than by a single sinusoid, is called a *residue*. In Fig. 2(b) the fundamental frequency is missing, yet will still be perceived as a residue pitch, which in this case is also called *virtual pitch*. The psychoacoustic phenomenon responsible for this effect is the "missing fundamental" effect. There is a long history of investigations into pitch perception, also regarding virtual pitch. Famous are the experiments of Seebeck in 1843, and the controversy between him and Ohm (see Plomp [14] for an historical review). There is a vast amount of literature on this topic; just a few interesting references are [15]–[19]. As the frequency of a pure tone decreases to very low values, say, under 100 Hz, the pitch becomes more difficult to determine. This is also true for the missing fundamental

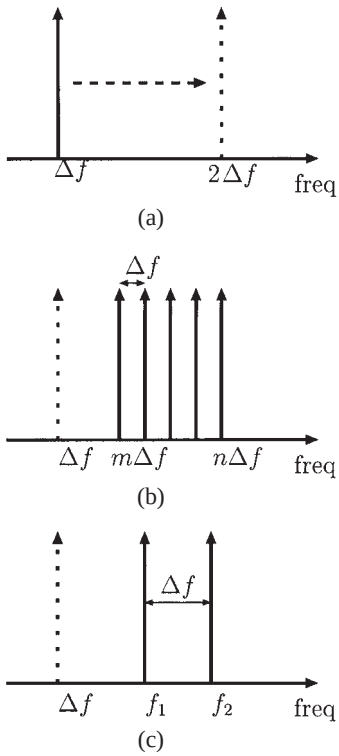


Fig. 2. Possible options for psychoacoustic bass enhancement. Dotted frequency component denotes perceived pitch (but is not necessarily acoustically radiated). (a) Frequency doubling. (b) Residue pitch. (c) Difference tone.

effect, and because our algorithm is aimed at this very low frequency range, we need psychoacoustic data regarding the perception of virtual pitch for this range. Unfortunately only sparse data are available. The work of Ritsma [20], [21] investigates the existence region of the tonal residue for frequencies above 200 Hz.

1.3 Difference Tone

We will illustrate how difference tones are generated by means of an example involving organ pipes. As in any other instrument, the frequency reproduced by an organ pipe has a direct (inverse) relationship with the size of the pipe. Thus for very low notes, very long pipes are required. Now if there is not enough space (in a church) for a pipe long enough to produce such very low notes, one can combine two higher notes to get a similar perceptual effect. The cause is that nonlinearities of the auditory system, within the cochlea, produce the difference tone of these two higher notes. This principle was — according to Helmholtz [22] — discovered in 1745 by Sorge, a German organist; the tones thus achieved are often known as Tartini's tones. Since the end of the sixteenth century many organs include a stop [the 5⅓-ft (1.65-m) fifth] composed of pipes sounding a fifth higher than the pitch of the actual note as played from the musical score. The purpose is to stimulate or reinforce the bass one octave below the pitch of the actual note [that is, to reinforce the 16-ft (4.90-m) sound of the organ]. Of older use, according to Roederer [23], is the use of the 10⅔-ft (3.3-m) fifth in the pedals, which in combination with 16-ft (4.9-m) stops, evokes the 32-ft (9.75-m) bass. Roederer attributes this to residue pitch [Fig. 2(b)]. However, this is doubtful since the effect of residue pitch decreases very fast for low frequencies. Consequently it is probably due to difference tones [see Fig. 2(c)]. As an example this is illustrated in Table 1, showing which frequencies are obtained for the pipes mentioned. The acoustical bass concept for organs has also been considered by Terhardt and Seewann [4]; see also Terhardt [25], which includes a concise overview of various aspects of pitch perception.

1.4 Considerations on Loudness Effects

Loudness perception depends strongly on the frequency of the stimulus presented to the ear(s). This is illustrated in Fig. 3, which shows the equal-loudness-level contours for steady-state pure tones. For the bass frequency range there are two important observations. First the hearing threshold increases sharply for low frequencies. Second the level contours lie close together, implying that small changes in sound level lead to large changes in loudness level.

Relating these properties of the auditory system to the processing scheme of Fig. 1 allows us to make a few a priori observations on the loudness effects we can expect to occur. First, because low frequencies are replaced by higher frequencies, the perceived loudness (of the total signal, but in particular of the bass range) will increase due to the lower hearing threshold. This is beneficial, as it is our goal to maximize the perceived low-frequency reproduction level. Although no formal verification of this assumption has been made, it is quite apparent from informal listening that the loudness does increase, even if the harmonics signal has the same amplitude as the original bass signal. Second, the "dynamic loudness range" of the perceived bass frequencies will be decreased due to the greater separation of the equal-loudness-level contours at higher frequencies. The effect might be opposed, or reversed, by introducing dynamics processing of the harmonics signal in Fig. 1, which Ben-Tzur and Colloms [10] have proposed, based on the contours of Fig. 3.

Another psychoacoustic effect that we need to consider is a timbral change of the bass perception when we replace the original bass by harmonics. Although the pitch should remain identical with our proposed method, the timbre or sound color will be altered after processing. Another cause for a timbral change might be the interference of the harmonics generated with frequency components in the original signal. The phase relationship between these original components and the synthesized harmonics is not clear a priori, and therefore the effect this interference will have on perception is hard to predict.

Table 1. Example of the acoustical bass of an organ.

	Length		Frequency	
f	(ft)	(m)	(Hz)	Key
Δf †	32	9.75	16.35	C_0
f_1	16	4.9	32.70	C_1
f_2	10⅔	3.3	49.05	G_1

†The length of the two organ pipes and their corresponding individual frequencies.

2 GENERATING HARMONICS

2.1 Processing Scheme

As stated previously, Fig. 1 presents the general processing scheme that we propose for psychoacoustic bass enhancement. In this section we will go into more detail regarding the choice of the filters and the nonlinear device.

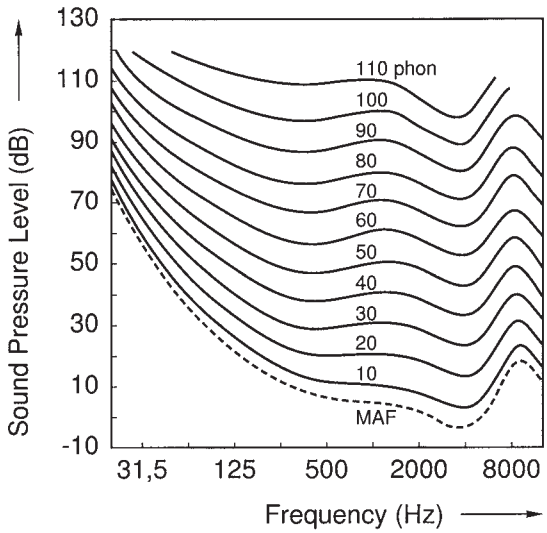


Fig. 3. Equal-loudness-level contours for pure tones (binaural free-field listening, frontal incidence). (From ISO 226-1987(E) [37, Fig. 1].)

As the system is "merely" based on a psychoacoustic model of pitch perception, and uses loudspeaker characteristics in a very general sense (it is only assumed that reproducing lower frequencies is less efficient than reproducing higher frequencies), the method can be employed for any type or size loudspeaker. Therefore emphasis will mostly lie on how to choose the filters in relation to the loudspeaker that is used. In Appendix 3 a complete design for a specific reproduction set is presented.

2.1.1 High-Pass Filter (HFIL)

The high-pass filter in the direct signal path (both left and right for a stereo signal) is used to prevent low-frequency reproduction by the loudspeaker. The primary purpose is to limit the cone excursion, since if this cone excursion is well controlled, we will prevent distortion when large input signals are used or a high output volume is selected. The maximum cone excursion that will be attained depends not only on the input signal or volume level, but also on the gain of the harmonics signal that the system adds, and the frequency range of FIL1. The larger this range, or the higher the gain, the larger the cone excursion will become. Thus the attenuation rate of the high-pass filter must be tuned to suit these parameters. (It is our experience that a fourth-order IIR filter is certainly sufficient.) In less demanding cases the filter might be completely omitted, however. The cutoff frequency of the high-pass filter can be equal to the resonance (or cutoff) frequency of the loudspeaker, because obviously, below this frequency the loudspeaker efficiency decreases drastically. In particular, we should choose this cutoff frequency equal to the cutoff frequency of the low-pass flank of bandpass filter FIL1, as we shall see in Section 2.1.2.

2.1.2 First Filter (FIL1)

This filter extracts frequencies from the left and right input signals, from which harmonics will be created in the nonlinear device. This filter usually is a bandpass filter. For the cutoff frequency of the low-pass flank of FIL1, we choose the cutoff frequency of the loudspeaker. This ensures that all frequencies that are inefficient to reproduce (that is, below the cutoff frequency) will be passed to the nonlinear device or harmonics generator. This cutoff frequency can also be chosen lower if desired — this will in general give a somewhat deeper bass impression of the output signal — at the expense of lower efficiency. The high-pass flank has a cutoff frequency of, say, 20 Hz. This serves merely to prevent dc signals from entering the nonlinear device. This is necessary because the output of the nonlinear device depends on all components of its input. As a dc signal serves no purpose in a reproduction system, we need to ensure that no dc signal enters the nonlinear device. Only if the cutoff frequency of the low-pass flank lies substantially above 150 Hz, leading to a bandwidth of FIL1 of over three octaves, should we choose a higher cutoff frequency of the high-pass flank. At a bandwidth of approximately three octaves or more, the input signal for the nonlinear device contains too wide a frequency range, resulting in too much intermodulation distortion, which would become clearly audible in the reproduced output signal. In such cases the cutoff frequency of the high-pass flank of FIL1 can be tuned to suit the loudspeaker. The low- and high-pass flanks are usually third- or fourth-order IIR.

2.1.3 Nonlinear Device (NLD)

The nonlinear device, or harmonics generator, "shifts" signal components in a low-frequency range to a higher frequency range. The pitch of the input signal is preserved, because the components in the higher frequency range are harmonics of the original components. As discussed in Section 1, the preservation of the original low pitch is due to the virtual pitch (or difference tone effect) of the harmonics signal. Because this element is a nonlinear device, any single output component depends on all input components. Moreover, at the output, frequency components will be generated which are not present at the input. This is a desired effect, since this is how the harmonics are obtained. However, it also leads to sum and difference components, which are not desired because they are not harmonically related to the input signals.

There are many nonlinear functions available to create harmonics, but for our purposes they should be computationally simple (or achievable with analog circuit elements). Another desirable characteristic is amplitude linearity. The operation of the device does not depend on the amplitude of the input signal. For example, an input–output relationship that has a quadratic or higher polynomial element is in principle unsuitable. For such nonlinearities the "amount" of nonlinearity generally increases with increasing magnitude of the input signal. This will lead to weak harmonics for small input signals and strong harmonics for large input signals. For the auditory system, an opposite effect would be more desirable. At low levels we need strong harmonics (strong nonlinearity of the nonlinear device) because the bass frequencies have a very small loudness at low levels. At higher levels the loudness increases sharply for bass frequencies, so we do not need very strong harmonics in such cases. In Sections 2 and 3 some specific examples are presented.

2.1.4 Second Filter (BP2)

This filter serves to extract a suitable spectrum from the output of the nonlinear device. The cutoff frequency of the high-pass flank is chosen equal to the cutoff frequency of the low-pass flank of FIL1. The value of the cutoff frequency of the low-pass flank of FIL2 depends somewhat on the choice of the nonlinear device. If its value is too high, the harmonics spectrum is wide, and has a very sharp timbre, which does not sound very pleasant. Thus, usually, this value is chosen roughly one octave above the cutoff frequency of the high-pass flank. Depending on the implementation of the nonlinear device, this value has to be tuned. Orders for the low- and high-pass flanks are again third- or fourth-order IIR.

2.2 Full-Wave Rectifier

We will now focus on the nonlinear device (NLD) of Fig. 1. A very simple nonlinear device is a full-wave rectifier, that is, a device that gives the absolute value of the

input. For a sine-wave input, only even harmonics are generated. So for a pure tone of frequency f_0, we get the output frequency components $2f_0$, $4f_0$, $6f_0$, etc. (and a noninteresting dc component as well), with a decay of 12 dB per octave. Expressions for the output spectrum in terms of the input spectrum for arbitrary periodic inputs are given in Appendix 2. Fig. 4 shows the input (solid line) and output (dashed line) signals for this situation. The output spectrum is indicated in Fig. 5. The fundamental component of the output is $2f_0$, which is not equal to the original pitch. Therefore this nonlinear device cannot be an ideal element for our purpose, but nonetheless, due to its simplicity, it may be an option. Informal listening tests have indicated that in many cases the bass enhancement obtained using a rectifier is an enhancement over the original.

2.3 Full-Wave Integrator

To create a deeper bass impression, corresponding to the pitch of the original fundamental, we must create a harmonics signal containing all (odd and even) harmonics. This is possible by using a full-wave integrator. This nonlinear device integrates the absolute value of the input, and the output is reset to zero when the input has a zero crossing with a positive slope. For a sine-wave input the situation is shown in Fig. 4, the full-wave integrator output being shown as the dash–dotted signal. Fig. 5 shows the input and output spectra — the output contains all the harmonics, which decay relatively slowly.

To explore what happens when the input signal is not a single frequency, but a more realistic (musical) signal, consider that the locations of the zero crossings are very important (as the output signal is reset to zero at every zero crossing with positive slope) and determine the fundamental frequency of the output signal, which should ideally be the same as the fundamental frequency of the input signal. If we consider a weakly stationary time series with spectral distribution $F(\omega)$, we can then write, as shown in Kedem [26],

$$\cos\pi\gamma = \frac{\int_0^{\pi} \cos\omega \, dF(\omega)}{\int_0^{\pi} dF(\omega)} \tag{3}$$

where γ is the expected zero crossing rate of the signal, that is, the expected number of zero crossings in a certain interval (which, for example, is 2 per period for a sine). This is called the *zero-crossing spectral representation*, and it expresses the tendency of $\pi\gamma$ to be attracted to a specific frequency (band) if this frequency (band) is dominant in the signal. This is a well-known empirical fact known as the *dominant frequency principle*. It shows us that we can be confident that if an input signal has a dominant frequency component, this will be reflected in its zero crossings. Consequently this frequency component will be the fundamental frequency of the output of the full-wave integrator. Exact expressions for the output as a function of an arbitrary periodic input signal are given in Appendix 2.

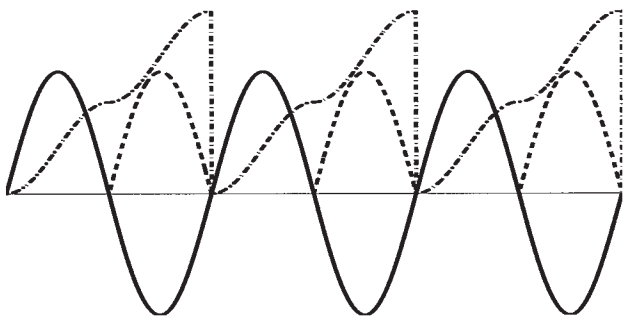


Fig. 4. Input and output signals of NLD. — sine wave input; – – – full-wave rectification; –.– full-wave integration.

3 SUBJECTIVE EVALUATION

In order to gain insight into the subjective appreciation of the psychoacoustic bass enhancement, a formal listening test was conducted. Although subjects may differ in what aspects they consider when determining "best bass quality," this does have the benefit of obtaining an overall assessment, and is therefore deemed to be quite useful.

Besides this obvious reason for performing a listening test, there was another reason. At an early stage of this research it was recognized that while for many repertoires the psychoacoustic bass enhancement gave very pleasant results, in some cases artifacts can occur which could even lead to a degradation of the bass quality. In most cases such artifacts can be classified as bass sounds acquiring an unnatural timbre, or excessively strong or prolonged bass notes. This last case might not actually be an artifact, but merely the unusual effect of a seemingly very good bass reproduction on a small loudspeaker. When the same piece of music was played back over a good subwoofer, the same prolonged bass note could be observed. Two examples of such a repertoire are "Hotel California" by The Eagles (which is one of the tracks used in the listening test), and "That's the Way It Is" by Phil Collins. What must be clear now is that for some repertoires it is questionable whether the applied processing will improve the subjective quality of the music. We call this "difficult" repertoire. In the listening test described next we have mainly used such difficult repertoires.

3.1 Experiment Description

3.1.1 Music Selection

As the system is meant primarily to enhance music reproduction, it is an obvious choice to use some typical music fragments in the listening test. Using musical fragments makes the listening test quite difficult. Instead of testing for one specific auditory attribute such as loudness, pitch, or even speech intelligibility, music appreciation may be seen as a combination of multiple attributes. When testing a bass enhancement system, the result would almost certainly be influenced by whether or not the test subject likes music with a high bass quantity. It is also felt that appreciation of any kind of music depends to some extent on the mood of the person at the moment of listening.

First the material to be used in the test was selected. Obviously this should be music with sufficient deep bass.

The genre of music to be used in the test was rock and pop—it is in these genres that bass enhancement systems find their most frequent application. Three of the four tracks used fall into the category of difficult repertoire, as described. Only track 1 is not difficult repertoire. The following four tracks were used:

1) "Bad," Michael Jackson. This track contains a typical pop bass line.

2) "My Father's Eyes," Eric Clapton. A very deep and strong bass line accompanies the music here, which might soon sound too imposing when enhanced.

3) "Hotel California" (live version), The Eagles. In the first part of this track a bass drum is heard, with some ambient sounds. The fact that besides the bass drum few other instruments are heard could lead to an overemphasis of the bass after processing.

4) "Twist and Shout," Salt n' Peppa. The bass in this track is a tight bass beat, which is difficult to preserve after processing.

3.1.2 Music Preparation

Each track will be processed in various ways, so as to create a number of different versions for each track. We use the following four processing methods:

1) No processing. The unprocessed version will be included as a reference for evaluations afterward. However, the subjects will not be informed that one of the four versions is the original unprocessed fragment. Therefore the subjects have no reference with which to compare the processed versions.

2) Linear bass boost. To compare the psychoacoustic processing with a traditional bass enhancement processing method.

3) Psychoacoustic bass enhancement with rectifier as harmonics generator.

4) Psychoacoustic bass enhancement with integrator as harmonics generator.

The details of how the processing for each of the four versions mentioned was achieved are given in Appendix 3.

3.1.3 Listening Test Procedure

By means of a pilot test it became obvious that the perceptual differences between some of the versions of the same track were very small. This makes a direct ranking of all versions of the same track very difficult. In such a case, paired comparison [27] is often the best method to obtain the subject's preference. All versions of any track are compared to all other versions of the track, and for each pair presented, subjects are to indicate their preference by judging the two versions with respect to bass quality. To ensure a fair and valid result, the experiment is performed randomized double blind, that is, the order of presentation of the pairs is completely randomized (with respect to track order and versions of the track), and neither subject nor experiment leader knows which versions are being presented.

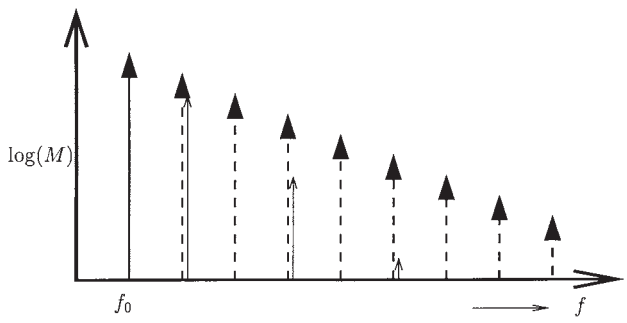


Fig. 5. Output spectrum. Solid arrow at f_0—input spectrum. Arrows at $2f_0$, $4f_0$, etc.—output spectrum for full-wave rectifier; dashed arrows at $2f_0$, $3f_0$, etc.—output spectrum for full-wave integrator.

If T represents the number of tracks and N the number of versions per track, the total number of pairs P presented to a subject in the complete test equals

$$P = T\,\frac{N(N-1)}{2}. \tag{4}$$

Thus in our $T = N = 4$, leading to $P = 24$. The answers of a subject's response in the paired comparisons are stored in a preference matrix $\mathscr{P}$, whose elements p_{ij} can be either 0 or 1, where a 1 indicates that the row element i is preferred over the column element j; a 0 indicates that the column element is preferred over the row element. Each track has an associated $N \times N$ preference matrix. Now for each preference matrix we construct a score vector $\boldsymbol{s}$ by summing the rows of $\mathscr{P}$. The elements of $\boldsymbol{s}$ can now be used as a ranking for the different versions of the track.

Fifteen subjects participated in the listening test, eleven male and four female. Most subjects were in the age group of 25–30 and had normal hearing, with varying degrees of experience in listening tests. Some of the subjects had no experience at all, whereas some may be termed experts. The subjects were also asked to indicate their general music preference (it was thought that this might have some influence on their results). This revealed that pop and rock music was by far the most popular music category of the subjects. They were encouraged to jot down any comments they would have during or after the listening test.

3.2 Listening Test Results

The raw data in the form of score vectors for all subjects and all tracks can be found in Appendix 4. In Fig. 6 we present a box plot of the results obtained from the listening test as described. The distribution of the scores for the four versions is indicated where we have summed the scores over all four tracks. The thick horizontal line indicates the median value for each version, and the bottom and top of the box indicate the 25% and 75% quartiles of the distribution. The whiskers extending from the bottom and top of the box show the remaining spread in the results. Those unfamiliar with the box plot display method are referred to Tukey [28] for a comprehensive treatment. It appears obvious that all three bass enhancement methods are ranked significantly higher than the unprocessed version. Although the median values for the three bass systems do differ, the spread in the rankings per system is quite large. The spread in the results for the two psychoacoustic systems is larger than that for the linear bass boost.

3.2.1 Analysis by Multidimensional Scaling

The overall result of the experiment does not give us insight into how the individual responses lead to this result. It will be interesting to investigate whether there is any structure in the distribution of the subjects' responses. For this we can use multidimensional scaling (MDS).

MDS is a technique to assess similarity between objects. It is especially useful if the attributes of the objects are vague or difficult to judge. Through some kind of procedure one determines the proximities between the objects, which can represent either similarity or dissimilarity, and the output will be a spatial map where the objects are represented as locations in an *n*-dimensional space. In case the proximities are dissimilarities, like objects will be positioned close together, and unlike objects far apart. An important parameter of the acquired mapping is the stress value, which indicates how well the distances in the mapping correspond to the true distances (proximities). A comprehensive treatment of MDS can be found in Kruskal and Wish [29], whereas Schiffman et al. [30] present a concise introduction to MDS, including a glossary of commonly used terminology and a treatment of several programs that can be used for MDS calculations. A previous application of MDS to psychoacoustic listening tests was presented in connection with judging crossover filter design in Aarts [31]. Herein the similarities of different designs were compared. We now propose another use of MDS, which is to scale the subject similarities.

Thus in our case the objects to analyze are the subject responses, and as proximities we have taken the dissimilarity between any two subjects. This we have taken to be the euclidean norm (Minkowski's parameter = 2) of the difference between their score vectors. We have pooled the results for the four tracks by summing the four score vectors of each subject. To visualize their configuration in a small number of dimensions, we applied MDS using the program KYST2a [32]. Since we only use the ranking order of the data, the program was used in the nonmetric mode, as opposed to the metric mode, where the data should be in interval or ratio scale. The stress values for one and two dimensions are given as 0.149 and 0.0216, respectively. Using Kruskal's classification [33], we can say that a mapping in one dimension is somewhere between poor and fair, whereas a two-dimensional mapping can be termed excellent. Fig. 7 presents the two-dimensional mapping of the fifteen subjects.

We observe that the fifteen subjects are distributed in a number of subgroups, which have been indicated by numbers in Fig. 7. The largest subgroup, S1, containing seven subjects, may be considered as having the most typical responses in this test. Referring to their results in Appendix 4 we can indeed verify that their score vectors agree reasonably well with the box plot of Fig. 6. The second largest subgroup is S2, and these are subjects C, E, and G, which all have a small preference for the two versions of psychoacoustic bass enhancement. Subgroup S3 includes subjects F and H. Most obvious in their results is the fact that only two subjects of the fifteen have never preferred an unprocessed version over another version. Furthermore, they prefer the two types of psychoacoustic bass enhancement very clearly over a linear bass enhancement, which is not the case for the "average" subgroup S1. From these three subgroups it appears that the position along the horizontal dimension indicates an appreciation of psychoacoustic bass enhancement. Subjects oriented toward the left have a clear preference for linear bass enhancement, subjects on the right clearly prefer psychoacoustic bass enhancement. Those in the middle have no clear preference. Also, since the amount of bass attainable with versions 3 and 4 is greater than with version 2, the horizontal dimension may represent a more general preference for bass. From left to right along the horizontal dimension, the preferences of the subjects appear to shift to those versions which contain a louder bass. Three sub-

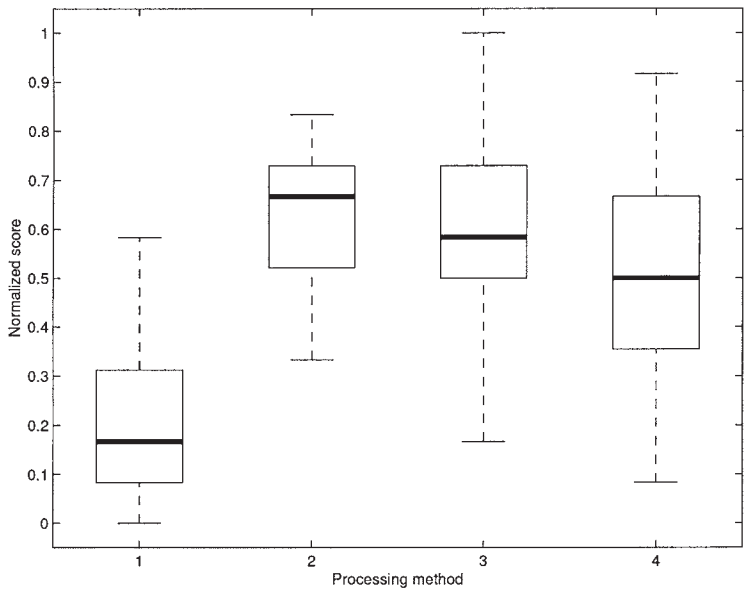


Fig. 6. Distribution of normalized scores per processing method, averaged over all four tracks. Processing methods: 1—unprocessed; 2—linear bass boost; 3—psychoacoustic with rectifier; 4—psychoacoustic with integrator.

jects remain: A, K, and L. For subject L (subgroup S0) we observe that the score for version 3 is much higher than typical, whereas for subjects A and K (subgroup S4) version 4 seems to have a higher preference. So along the vertical dimension it appears that we can observe an indication of which kind of psychoacoustic bass enhancement is preferred (since versions 3 and 4 each use a different nonlinear device to generate higher harmonics). It may be noted that the spread along the vertical dimension is smaller than the spread along the horizontal dimension. From this we might conclude that the associated attribute is less important for the subject responses.

Using MDS we have found that the preferences for the four versions of processing can be described in two dimensions. Furthermore, we have been able to arrive at an interpretation of what these two dimensions indicate. The fifteen subjects appear to be able to be classified into five subgroups, with the preferences in the subgroup being very similar, whereas the preferences between the subgroups are very dissimilar.

3.2.2 Analysis by Biplots

Although multidimensional scaling gives an instant impression of the structure in the subject responses, it does not allow a direct evaluation of how these relate to the four processing systems. A visualization method that yields both is the *biplot*, devised by Gabriel [34]. A biplot is a method to visualize the value of the elements of a rank 2 matrix. The biplot is also applicable to higher rank matrices if the approximation of such a matrix to a rank 2 matrix can be made to a sufficient degree of accuracy. Those not acquainted with the biplot method are referred to Appendix 5 for a short introduction.

Fig. 8 shows the column metric preserving (CMP) biplot for the listening test results. The data over all tracks have been summed, and the four different versions have been labeled 1–4, as well as the fifteen vectors representing the subjects, A – O. The goodness of fit is 0.87, which is not very high, but should not be too low for a useful analysis.

Focusing on the orientation of the vectors, we see a similar set of subgroups as was obtained for the multidimensional scaling. Subjects C, E, and G are again a very clear subgroup; subjects F and H are once again nearly identical in their representation in the biplot, and subjects A and K are together again as well. Only subject L falls right in the middle of the "typical" subgroup, whereas in the multidimensional scaling case this subject falls outside that subgroup. The length of a vector is an indication of the value of the standard deviation of the subject's preferences. Thus from the length of the vector we can see whether a subject has a strong preference for one or more systems or rates all more or less the same. In the last case, the subject's score vector will have a small standard deviation, and its associated vector in the biplot will thus have a small length. This is the case for subjects D and M, as we can see in Fig. 8, and which we can also verify directly from their score vectors given in Appendix 4. Subjects with long vectors, such as subjects E, G, F, and H, apparently have a large standard deviation in their responses, which indicates a strong preference for some systems and little preference for others. Again, referring to their score vectors in Appendix 4, we can verify that this is indeed the case.

Similarity between the different systems can be judged from the distance between their coordinates in the biplot. So we can clearly see that systems 3 and 4 are judged most similar, which are the two psychoacoustic bass enhancement systems. The largest distance is observed between systems 2 and 4, which are the linear bass boost versus the psychoacoustic bass enhancement using a full-wave integrator as harmonics generator. Their (dis)similarity does not differ too much from the (dis)similarity between systems 1 and 2, 1 and 3, and finally 1 and 4.

4 CONCLUSIONS

In this paper we proposed a psychoacoustically based signal-processing system to enhance the perceived bass response of a loudspeaker below its cutoff frequency. The main concept of this system is to replace very low fre-

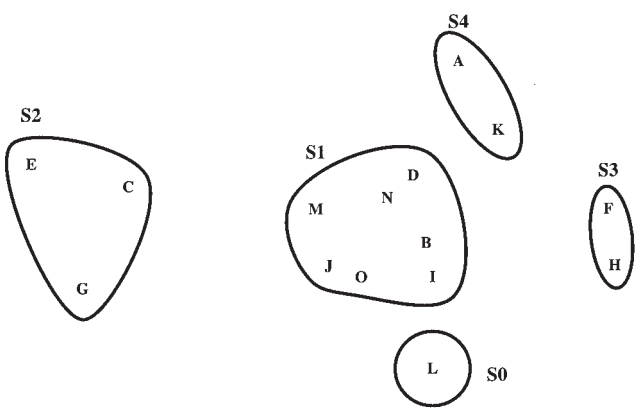


Fig. 7. Two-dimensional scaling result for pooled subject responses. Stress value 0.0216. Note five subgroups S0–S4.

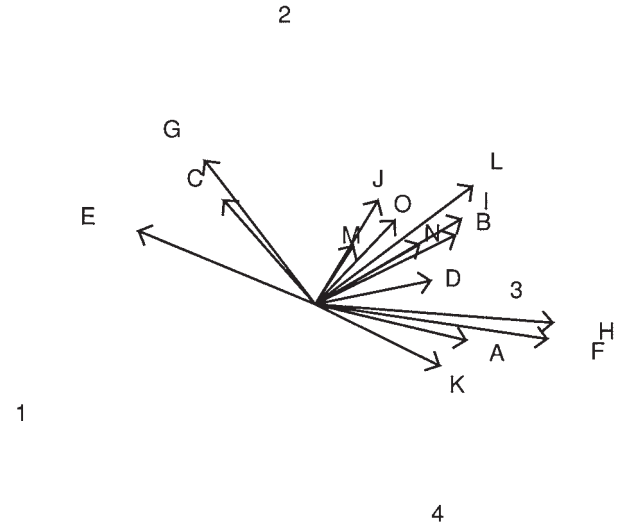


Fig. 8. CMP biplot of listening test results. Goodness of fit 0.87. Subjects are indicated by letters, processing methods by numbers. Processing methods: 1—unprocessed; 2—linear bass boost; 3—psychoacoustic with rectifier; 4—psychoacoustic with integrator.

quency components by their harmonics through controlled nonlinear processing. The resulting harmonics yield the same (virtual) pitch as the original signal, due to the missing fundamental effect. Beneficial characteristics of the system include:

- (Very) low radiation of energy below loudspeaker cutoff frequency.
- Less headroom required compared to a traditional (linear) bass boost for a comparable bass enhancement effect.
- Computationally very efficient, simple circuit in case of analog design.
- Power efficient.
- Tunable to any kind and size of loudspeaker.

Some drawbacks are:

- Intermodulation distortion can lead to audible artifacts. Careful tuning of filters will be beneficial.
- The added harmonics interfere with frequency components in the original audio, which in some cases may alter the timbre of the signal to some extent.

A compact mathematical description of the two nonlinear functions (used to create the harmonics) has been derived and presented.

A listening test has been performed, with fifteen subjects participating. The psychoacoustic bass enhancement system in two different versions was compared against a system with a linear bass boost. Unprocessed music was included as a reference. Both psychoacoustic systems as well as the linear bass boost system were appreciated significantly more highly than the unprocessed audio. A large variation in subject responses has been observed. This has been further analyzed by means of multidimensional scaling and biplots. Both visualization methods have indicated a number of groups, out of the total subject pool, which have similar in-group preferences. Personal preference for any kind of processing appears to be highly individual, whereas some subjects prefer no processing at all.

5 ACKNOWLEDGMENT

The authors would like to thank their colleagues Catherine Polisset and Erik van der Tol for their parts in the work presented in this paper. We would also like to acknowledge Guido Janssen for providing great help in deriving the expressions for the nonlinearities and Jan Engel for contributions in the evaluation of the listening test results. We greatly appreciate the participation of all people in the listening test. Finally, we thank the *Journal* reviewers for their constructive criticism.

6 REFERENCES

[1] E. M. Long and R. J. Wickersham, "Method and Apparatus for Operating a Loudspeaker below Resonant Frequency," US patent 4,481,662, filed 1982 (1984).

[2] E. M. Long and R. J. Wickersham, "Guide to ELF Systems. A New Era in Bass Reproduction" (1997), http://www.bagend.com.

[3] L. L. Beranek, *Acoustics* (American Institute of Physics, New York, 1954; reprinted by Acoustical Society of America, 1986).

[4] H. F. Olson, *Acoustical Engineering* (Van Nostrand, Princeton, NJ, 1957).

[5] J. Borwick, Ed., *Loudspeaker and Headphone Handbook* (Butterworths, London, 1988).

[6] R. M. Aarts, *25 jaar AES in Nederland* (Gebotekst, Zoetermeer, The Netherlands, 1999), chap. 17, pp. 158–159.

[7] R. M. Aarts and S. P. Straetemans, "Circuit, Audio System and Method for Processing Signals, and a Harmonics Generator," US patent 6,111,960 (2000 Aug. 29).

[8] G. F. M. De Poortere, C. Polisset, and R. M. Aarts, "Ultra Bass," US patent 6,134,330 (2000 Oct. 17).

[9] "Sonora model RCU-208. A Synthetic Bass Note Circuit," *Radio Electron.*, vol. 37, p. 37 (1948).

[10] D. Ben-Tzur and M. Colloms, "The Effect of MaxxBass Psychoacoustic Bass Enhancement System on Loudspeaker Design," presented at the 106th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 47, p. 517 (1999 June), preprint 4892.

[11] W. Schott, "Low Frequency Audio Doubling and Mixing Circuit," European patent EP0546619 (1991).

[12] M. Oda, "Low Frequency Audio Conversion Circuit," US patent 5,668,885 (1997).

[13] T. Unemura, "Audio Circuit," US patent 5,771,296 (1998).

[14] R. Plomp, "Experiments on Tone Perception," Ph.D. thesis, University of Utrecht, The Nederlands (1966).

[15] F. A. Bilsen and R. J. Ritsma, "Some Parameters Influencing the Perceptibility of Pitch," *J. Acoust. Soc. Am.*, vol. 47, pp. 469–475 (1970).

[16] E. de Boer, "On the 'Residue' in Hearing," Ph.D. thesis, University of Amsterdam, The Netherlands (1956).

[17] A. J. M. Houtsma and J. L. Goldstein, "The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition," *J. Acoust. Soc. Am.*, vol. 51, pp. 520–529 (1972).

[18] J. F. Schouten, "The Perception of Pitch," *Philips Tech. Rev.*, vol. 5, no. 10, p. 286 (1940).

[19] J. F. Schouten, R. J. Ritsma, and B. Lopes Cardozo, "Pitch of the Residue," *J. Acoust. Soc. Am.*, vol. 34, pp. 1418–1424 (1962).

[20] R. J. Ritsma, "Existence Region of the Tonal Residue I," *J. Acoust. Soc. Am.*, vol. 34, pp. 1224–1229 (1962).

[21] R. J. Ritsma, "Existence Region of the Tonal Residue II," *J. Acoust. Soc. Am.*, vol. 35, pp. 1241–1245 (1963).

[22] H. Helmholtz, *Die Lehre von den Tonempfindungen* (1863); (English transl. of 4th German ed. of 1877: Vieweg, Dover, 1954).

[23] J. G. Roederer, *Introduction to Physics and Psychophysics of Music* (Springer, New York, 1975).

[24] E. Terhardt and M. Seewann, "Der 'akustische

Bass' von Orgeln," in *Fortschritte der Akustik DAGA84* (DeGA, Darmstadt, Germany, 1999), pp. 885–888.

[25] E. Terhardt, "Perception of Auditory Pitch," http://www.mmk.e-technik.tumuenchen.de//persons/ter.html (2000).

[26] B. Kedem, *Time Series Analysis by Higher Order Crossings* (IEEE Press, New York, 1994).

[27] H. A. David, *The Method of Paired Comparisons*, 2nd ed. (Griffin, London, 1988).

[28] J. W. Tukey, *Exploratory Data Analysis* (Addison-Wesley, Reading, MA, 1977).

[29] J. B. Kruskal and M. Wish, *Multidimensional Scaling, Quantitative Applications in the Social Sciences* (Sage Publ., Beverly Hills, CA, 1978).

[30] S. S. Schiffman, M. L. Reynolds, and F. W. Young, *Introduction to Multidimensional Scaling* (Academic Press, New York, 1983).

[31] R. M. Aarts, "A New Method for the Design of Crossover Filters," *J. Audio Eng. Soc.*, vol. 37, pp. 445–454 (1989 June).

[32] J. B. Kruskal, F. W. Young, and J. B. Seery, "How to Use KYST, a Very Flexible Program to Do Multidimensional Scaling and Unfolding," Bell Laboratories, Murray Hill, NJ (1973).

[33] J. B. Kruskal, "Multidimensional Scaling by Optimizing Goodness of Fit to a Nonmetric Hypothesis," *Psychometrika*, vol. 29, pp. 1–27 (1964).

[34] K. R. Gabriel, "The Biplot Graphical Display of Matrices with Application to Principal Component Analysis," *Biometrika*, vol. 58, pp. 453–467 (1971).

[35] "Syntrillium," Syntrillium Software Corp., http://www.syntrillium.com (2002).

[36] K. R. Gabriel and C. L. Odoroff, "Biplots in Biomedical Research," *Statist. in Med.*, vol. 9, pp. 469–485 (1990).

[37] ISO 226-1987(E), "Acoustics — Normal Equal-Loudness Level Contours," International Standards Organization, Geneva, Switzerland (1987).

APPENDIX 1
LOUDSPEAKER EFFICIENCY

In this section we argue that, for a small loudspeaker, achieving a low cutoff frequency while maintaining a reasonable efficiency is not possible with a traditional loudspeaker-cabinet construction. Reproducing frequencies below the cutoff frequency is very inefficient and should be avoided as much as possible. This provides enough motivation for investigating the possibility of creating ultralow-pitch perceptions using frequencies above (or at least closer to) the cutoff frequency of the loudspeaker.

In order to show this we explore the relation between various loudspeaker parameters and radiation efficiency. For a more elaborate treatment on the radiation characteristics of various types of loudspeakers the reader is referred to [3]–[5]. We first define the efficiency η at frequency ω as

$$\eta(\omega) = \frac{P_a(\omega)}{P_e(\omega)} \tag{5}$$

where P_a is the time-averaged acoustically radiated power and P_e is the time-averaged electrical power supplied by the generator. We then write

$$P_a(\omega) = \frac{1}{2}\hat{V}^2 \Re\{Z_{rad}(\omega)\} \tag{6}$$

where $\hat{V}$ is the voice-coil velocity amplitude and Z_{rad} the mechanical radiation impedance ($\Re\{\bullet\}$ denotes the real of $\bullet$), and

$$P_e(\omega) = \frac{1}{2}\hat{I}^2 \Re\{Z_{in}(\omega)\} \tag{7}$$

where $\hat{I}$ is the current amplitude delivered by the generator and Z_{in} the total impedance load as seen by the generator.

Using the lumped-element analogon [3], [4], we find the mobility analogous circuit of an electrodynamic loudspeaker in a closed box as shown in Fig. 9. Here we have used the following elements:

E	Generator	
R_E	Resistance of voice coil	[Ω]
L_E	Self-inductance of voice coil	[H]
I	Current delivered by generator	[A]
U	Voltage across voice coi	[V]
V	Velocity of voice coil	[m/s]
F	Force on voice coil	[N]
m_1	Mass of voice coil and cone	[kg]
k_t	Total spring constant (suspension, cabinet)	[N/m]
R_m	Mechanical damping of transducer	[N · s /m]
Z_{rad}	Radiation impedance	[N · s /m]

We can now write for the impedance Z_{in} as seen by the generator,

$$Z_{in}(\omega) = R_E + j\omega L_E + \frac{(Bl)^2}{R_m + R_{rad}(\omega) + j[\omega m_1(\omega) - k_t/\omega]} \tag{8}$$

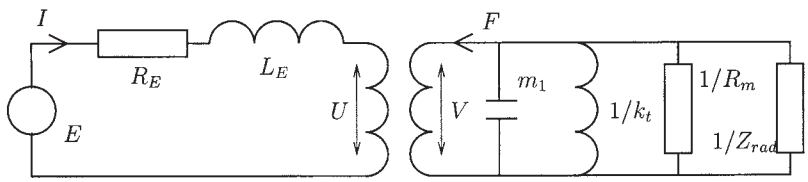


Fig. 9. Mobility analogous circuit for electrodynamic loudspeaker in a box driven by generator *E*. Electrodynamic coupling follows $U = BlV$ and $F = BlI$. Symbols are explained in text.

where B is the magnetic inductance in the air gap in tesla and l is the length of the voice-coil wire in meters. R_{rad} is the real part of the complex radiation impedance, $Z_{rad} = R_{rad} + jX_{rad}$, where X_{rad} has been added to m_1 to form m_t, the total moving mass. Strictly speaking m_t is frequency dependent due to the contribution of X_{rad}, but this contribution is small,[1] and thus we will consider m_t to be frequency independent in the following. Combining Eqs. (7) and (8) we get

$$P_e(\omega) = \frac{1}{2} I^2 \left\{ R_e + \frac{(Bl)^2 |R_m + R_{rad}(\omega)|}{|R_m + R_{rad}(\omega)|^2 + |\omega m_t - (k_t/\omega)^2|} \right\}. \quad (9)$$

Next we shall express the voice-coil velocity amplitude V in I, the current delivered by the generator. We substitute the result in Eq. (6) and then compute the efficiency η by Eq. (5). According to $F = BlI$, the current through the voice coil determines the force F on it. The relation between F and V is determined by the mechanical impedance,

$$V(\omega) = \frac{F(\omega)}{R_m + R_{rad}(\omega) + j(\omega m_t - k_t/\omega)} = \frac{Bl I(\omega)}{R_m + R_{rad}(\omega) + j(\omega m_t - k_t/\omega)}. \quad (10)$$

Combining Eqs. (6) and (10), we find

$$P_a = \frac{1}{2} I^2 \frac{(Bl)^2 R_{rad}(\omega)}{|R_m + R_{rad}(\omega)|^2 + (\omega m_t - k_t/\omega)^2}. \quad (11)$$

Now we readily compute the efficiency using Eq. (5),

$$\eta(\omega) = \frac{(Bl)^2 R_{rad}(\omega)}{R_t (\omega_0 m_t)^2 \left\{ \left| \frac{R_m + R_{rad}(\omega)}{\omega_0 m_t} \right|^2 + \frac{R_m + R_{rad}(\omega)}{R_t} \left(\frac{Bl}{\omega_0 m_t} \right)^2 + \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)^2 \right\}} \quad (12)$$

where ω_0 is the resonance frequency of the mass–spring system, given as

$$\omega_0 = \sqrt{\frac{k_t}{m_t}}. \quad (13)$$

To study the behavior of η for various frequencies, we first notice that according to [5, p. 9],

$$R_{rad}(\omega) = \begin{cases} \dfrac{\rho_0 \omega^2 S^2}{2\pi c_0}, & \omega << \dfrac{\sqrt{2}\, c_0}{a} \\ \rho_0 c_0 S, & \omega >> \dfrac{\sqrt{2}\, c_0}{a} \end{cases} \quad (14)$$

where we denote sound velocity by c_0, air density by ρ_0, cone radius by a, and cone area by S. The frequency at which the behavior of R_{rad} changes is the transition frequency ω_t (above which the radiation becomes increasingly directional). We can immediately deduce from Eqs. (12) and (14) that in the normal operating range of the loudspeaker the efficiency is constant and equals

$$\eta(\omega) \approx \frac{(Bl)^2 \rho_0 S^2}{R_t 2\pi c_0 m_t^2}, \qquad \omega_0 << \omega << \omega_t. \quad (15)$$

This clearly presents a problem: a high efficiency requires a large cone area and at the same time a small mass. The precise behavior of η around ω_0 depends on the loudspeaker parameters, but it is not within the scope of this paper to investigate this now. If we focus on what happens for $\omega << \omega_0$, we notice that the last term between the curly braces in the denominator of Eq. (12) is dominant and equal to $(\omega_0/\omega)^2$, and in combination with the ω^2 behavior of R_{rad} we thus have

$$\eta(\omega) \sim \omega^4, \qquad \omega << \omega_0. \quad (16)$$

It is clear that the efficiency decreases rapidly below the resonance frequency of the loudspeaker. Referring to the definition of this cutoff frequency [Eq. (13)], we see that in order to achieve a low cutoff frequency we must have a low k_t and a high m_t. Since the contribution of the cabinet, k_c, to k_t is given by [3, p. 129],

$$k_c = \frac{\rho_0 c_0^2 S^2}{V_0} \quad (17)$$

with V_0 the cabinet volume, it is obvious that for small loudspeaker enclosures, k_t will be relatively high. Furthermore, small loudspeakers obviously have a small mass. The combination of k_t and m_t for a small loudspeaker is precisely opposite to what one needs to obtain a low cutoff frequency. Increasing m_t is not really an option, because as Eq. (15) shows, this will lead to a much lower efficiency.

APPENDIX 2
ANALYSIS OF HARMONICS GENERATION

A2.0 Introduction

In this section we derive analytical descriptions for the full-wave rectifier and the full-wave integrator. Obviously

[1] According to [5, p. 9], for a rigid circular piston in an infinite baffle the mechanical impedance (for frequencies much lower than the transition frequency of the piston, which equals $c_0/\pi a\sqrt{2} \approx 1.5$ kHz using the values given below) $X_{rad} = \omega m_{air} = 8\rho_0 \omega S^2/3\pi^2 a$, with S the cone area, a the cone radius, and ρ_0 the air density. If we use, for example, $\rho_0 = 1.3$ kg/m^3, $S = \pi(0.05)^2$ m^2, and $a = 0.05$ m, then we find $m_{air} \approx 1$ g. Thus from now on we will consider m_t to be frequency independent.

this requires a different approach than the one used for linear time-invariant (LTI) systems, and in general, descriptions of nonlinear systems are extremely complex or pertain to a limited class of signals. However, for our purposes an elegant description has been obtained, which encompasses all practical signals.

- First consider a real periodic signal $f(t)$, having a period of 1, and assume that

$$f(0) = f(t_0) = f(1) = 0\,, \quad f(t) \neq 0\,,\ t \neq 0\,, t_0\,, 1\,. \tag{18}$$

  Thus $f(t)$ has one 0, at t_0, in the interval (0, 1).
- $f(t)$ changes sign at every zero crossing.
- We choose $f'(0) > 0$.
- $f(t)$ is sufficiently smooth in order that its Fourier coefficients a_n [Eq. (19)] decay at a rate of at least $1/n^2$ [which will be satisfied if $f(t)$ is twice continuously differentiable].

We have for $f(t)$ the Fourier series representation

$$f(t) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n t} \tag{19}$$

where $a_n = a_{-n}^*$, because $f(t)$ is real. (x^* represents the complex conjugate of x.) Now we also consider the real periodic function $F(t)$, derived from $f(t)$ by some nonlinear operation. We have for $F(t)$ the Fourier series representation

$$F(t) = \sum_{n=-\infty}^{\infty} b_n e^{2\pi i n t} \tag{20}$$

where $b_n = b_{-n}^*$. In the following we derive expressions for the b_n in terms of the a_n, where each subsection covers a different nonlinear operator (and hence the b_n in the various subsections are not the same). More specifically, it is shown that the b_n obtained by full-wave rectification of $f(t)$ decay as $1/n^2$ and, for large n, can be expressed in the values of $f'(t)$ at the zero crossings of $f(t)$. For the full-wave integrator, the b_n decay as $1/n$ and are mainly determined by the area between $f(t)$ and the horizontal axis.

A2.1 Full-Wave Rectifier

The first nonlinear device that we treat is the full-wave rectifier (see Section 2.2). On the periodicity interval [0, 1) the function $F(t)$ is now given by

$$F(t) = |f(t)|\,, \qquad 0 \leqslant t < 1\,. \tag{21}$$

Now there holds

$$|f(t)| = f(t)h(t;\, 0,\, t_0,\, 1)\,, \qquad 0 \leqslant t < 1 \tag{22}$$

where

$$h(t; 0, t_0, 1) = \begin{cases} 1\,, & 0 \leqslant t < t_0 \\ -1\,, & t_0 < t < 1\,. \end{cases} \tag{23}$$

Letting d_n be the Fourier coefficients of $h(t;\, 0,\, t_0,\, 1)$, so that

$$h(t; 0, t_0, 1) = \sum_{n=-\infty}^{\infty} d_n e^{2\pi i n t} \tag{24}$$

we then obtain from Eqs. (19), (22), and (24) that

$$\begin{aligned} F(t) &= \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n t} \sum_{m=-\infty}^{\infty} d_m e^{2\pi i m t} \\ &= \sum_{k=-\infty}^{\infty} e^{2\pi i k t} \sum_{n=m-k} a_n d_m\,. \end{aligned} \tag{25}$$

Hence from Eq. (20),

$$b_k = \sum_{n=-\infty}^{\infty} a_n d_{k-n}. \tag{26}$$

Due to the piecewise constant nature of $h(t)$ it is easily computed that

$$d_0 = \int_0^1 h(t; 0, t_0, 1)\ dt = 2t_0 - 1 \tag{27}$$

$$\begin{aligned} d_n &= \int_0^1 h(t; 0, t_0, 1)\ e^{-2\pi i n t}\, dt \\ &= \frac{1}{\pi i n}\,(1 - e^{-2\pi i n t_0})\,, \qquad n \neq 0\,. \end{aligned} \tag{28}$$

Therefore, from Eqs. (26)–(28),

$$b_k = (2t_0 - 1)\,a_k + \sum_{n=-k} \frac{a_n}{\pi i\,(k-n)}\,|1 - e^{2\pi i(n-k)t_0}|\,. \tag{29}$$

In principle this solves the problem for the full-wave rectifier. We have expressed the b_k in terms of the a_k and the locations of the zeros of $f(t)$. However, the right-hand side of Eq. (29) shows a decay of the b_k roughly as $1/k$, while we know from the form of $F(t)$ that there should be a decay like $1/k^2$ due to the triangular singularity of $F(t)$ at $t = t_0$. This decay of the b_k can be obtained when we properly use the condition stated in Eq. (18). Accordingly, we have

$$\sum_{n=-\infty}^{\infty} a_n = 0\,, \qquad \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n t_0} = 0\,. \tag{30}$$

Then we get for the series at the far right-hand side of Eq. (29),

$$\begin{aligned} &\sum_{n=-k} \frac{a_n}{\pi i\,(k-n)}\,|1 - e^{2\pi i(n-k)t_0}| \\ &= \sum_{n=-k} \frac{a_n}{\pi i}\left(\frac{1}{k-n} - \frac{1}{k} + \frac{1}{k}\right)|1 - e^{2\pi i(n-k)t_0}| \\ &= \sum_{n=-k} \frac{n a_n}{\pi i k\,(k-n)}\,|1 - e^{2\pi i(n-k)t_0}| \\ &\quad + \frac{1}{\pi i k} \sum_{n=-k} a_n |1 - e^{2\pi i(n-k)t_0}|\,, \qquad k \neq 0\,. \end{aligned} \tag{31}$$

Now it holds that

$$\sum_{n=k} a_n e^{2\pi i(n-k)t_0} = -a_k + \sum_{n=-\infty}^{\infty} a_n e^{2\pi i(n-k)t_0}$$
$$= -a_k + e^{-2\pi i k t_0} \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n t_0}$$
$$= -a_k . \tag{32}$$

In the same way we have that $\Sigma_{n\neq k}\, a_n = -a_k$. Hence, for $k \neq 0$, the second term of Eq. (31) vanishes and we find

$$b_0 = \int_0^1 |f(t)|\, dt \tag{33}$$

$$b_k = (2t_0 - 1)\, a_k - \sum_{n=k} \frac{n a_n}{\pi i k (k-n)} \left| 1 - e^{2\pi i (n-k) t_0} \right| . \tag{34}$$

The right-hand side of Eq. (34) does exhibit the correct $1/k^2$ behavior that we expect from the b_k for large k (since the a_k also decay at a rate of at least $1/k^2$, as explained in the introduction to this appendix). More precisely, assuming that $a_k = 0$ for large k, we have for large k that

$$\sum_{n=k} \frac{n a_n}{\pi i k (k-n)} \left| 1 - e^{2\pi i (n-k) t_0} \right|$$
$$\approx \frac{1}{k^2} \sum_{n=-\infty}^{\infty} \frac{n a_n}{\pi i} \left| 1 - e^{2\pi i (n-k) t_0} \right| . \tag{35}$$

Since

$$f'(t) = \sum_{n=-\infty}^{\infty} 2\pi i n a_n e^{2\pi i n t} \tag{36}$$

we then obtain

$$\sum_{n=k} \frac{n a_n}{\pi i k (k-n)} \left| 1 - e^{2\pi i (n-k) t_0} \right|$$
$$\approx \frac{1}{k^2} \frac{1}{\pi i} \frac{1}{2\pi i} \sum_{n=-\infty}^{\infty} 2\pi i n a_n \left| 1 - e^{2\pi i (n-k) t_0} \right|$$
$$= -\frac{1}{2\pi^2 k^2} \left| f'(0) - f'(t_0)\, e^{-2\pi i k t_0} \right| . \tag{37}$$

Thus if $a_k = 0$ for large k, then for large k,

$$b_k \sim \frac{1}{2\pi^2 k^2} \left| f'(0) - f'(t_0)\, e^{-2\pi i k t_0} \right| . \tag{38}$$

A2.2 Full-Wave Integrator

Now we consider the full-wave integrator (see Section 3) on the periodicity interval [0, 1). We get

$$F(t) = \int_0^t |f(s)|\, ds , \qquad 0 \le t < 1 . \tag{39}$$

The jumps of $F(t)$ at the resetting moments are given by

$$-\alpha_0 = -\int_0^1 |f(s)|\, ds \tag{40}$$

at time $t = k$, where k is an integer. For $t \neq 0$, t_0, 1, we have that $F'(t) = |f(t)|$. Thus,

$$F'(t) = |f(t)| - \alpha_0 \sum_{n=-\infty}^{\infty} \delta(t-n) . \tag{41}$$

Denoting the Fourier coefficients of $|f(t)|$ by c_k, so that

$$|f(t)| = \sum_{k=-\infty}^{\infty} c_k e^{2\pi i k t} \tag{42}$$

and using

$$\sum_{n=-\infty}^{\infty} \delta(t-n) = \sum_{k=-\infty}^{\infty} e^{2\pi i k t}$$

we can write Eq. (41) as

$$\sum_{k=-\infty}^{\infty} 2\pi i k b_k e^{2\pi i k t} = \sum_{k=-\infty}^{\infty} (c_k - \alpha_0) e^{2\pi i k t} \tag{43}$$

so that

$$b_k = \frac{c_k - \alpha_0}{2\pi i k}, \qquad k \neq 0 . \tag{44}$$

The b_k show a decay of roughly $1/k$, which is what we expect due to the discontinuity of $F(t)$ at $t = 1$. The c_k can be found using the expressions of the previous subsection. For $k = 0$, and using partial integration, we obtain

$$b_0 = \int_0^1 F(t)\, dt$$
$$= [tF(t)]_0^1 - \int_0^1 t \left\{ |f(t)| - \alpha_0 \sum_{n=-\infty}^{\infty} \delta(t-n) \right\} dt$$
$$= \int_0^1 (1-t)\, |f(t)|\, dt . \tag{45}$$

A2.3 Generalization to Arbitrary Periodic Signals

We can generalize the preceding results for arbitrary periodic signals that may have more than one zero crossing in the fundamental period. The derivations follow exactly the same procedure as given before, with the following considerations:

- The start and end points of a period are now t_{-1} and t_z, respectively. In between there are z zeros at $t_{0,1,\cdots,z-1}$. Due to the periodicity requirements and the fact that $f(t)$ changes sign at every zero, z must be odd.
- The fundamental frequency of the signal is $\nu_0 = 1/(t_z - t_{-1})$.
- For the full-wave integrator $F(t)$ is reset to zero at $t_{1,3,\cdots,z-1}$. The corresponding jumps are given by $\alpha_{0,1,\cdots,(z-1)/2}$.

An example signal with the t_m indicated is shown in Fig. 10 for the case $z = 7$. The result for the full-wave rectifier is as follows:

$$b_0 = \nu_0 \int_{t_{-1}}^{t_z} |f(t)|\, dt \tag{46}$$

$$b_k = \left[1 + 2\nu_0 \sum_{m=0}^{z} (-1)^m t_m\right] a_k - \sum_{n=k} \frac{n a_n}{\pi i k (k-n)} \sum_{m=0}^{z} (-1)^m e^{2\pi i \nu_0 (n-k) t_m} \tag{47}$$

and if $a_k = 0$ for large k, we then get for large k

$$b_k \sim \frac{1}{2\pi^2 \nu_0 k^2} \sum_{m=0}^{z} (-1)^m f'(t_m) e^{-2\pi i \nu_0 k t_m}. \tag{48}$$

For the full-wave integrator (and denoting the Fourier coefficients of the full-wave rectified signal by c_k), we get

$$b_k = \frac{c_k - \sum_{m=0}^{(z-1)/2} \alpha_m e^{-2\pi i \nu_0 k t_{2m+1}}}{2\pi i \nu_0 k}, \quad k \neq 0 \tag{49}$$

$$b_0 = -\int_{t_0}^{t_z} t\,|f(t)|\,dt + \sum_{m=0}^{(z-1)/2} \alpha_m t_{2m+1}. \tag{50}$$

A2.4 Discrete Time

We can derive expressions for sampled signals as well (amplitude continuous). We use square brackets to indicate the time-discrete property of the signals, such as $f[n]$. The number of samples per period is N, the sample time Δt, and the n_m ($m = -1, 0, \ldots, z$) are the zero crossings of $f[n]$, analogously as in the continuous-time case. We define a zero crossing in the discrete-time sequence to occur at n_m if $f[n]$ changes sign at $n = n_m$ (or if $f[n_m] = 0$).

For the full-wave rectified signal we get

$$b[k] = \left[1 + \frac{2}{N} \sum_{m=0}^{z} (-1)^m n_m\right] a[k] + \frac{1}{N} \sum_{n=k} a[n] \sum_{m=0}^{z} (-1)^m \times e^{\pi i (n_m + n_{m-1})(k-n)/N} \frac{\sin \pi (n_m - n_{m-1})(k-n)/N}{\sin \pi (k-n)/N}. \tag{51}$$

For the full-wave integrator (using $c[k]$ for the full-wave rectified signal) the $b[k]$ become

$$b[k] = \Delta t \frac{c[k] - \sum_{m=0}^{(z-1)/2} \alpha_m e^{-2\pi i k n_{2m+1}/N}}{1 - e^{-2\pi i k/N}}, \quad k \neq 0 \tag{52}$$

$$b[0] = -\Delta t \sum_{k=0}^{n-1} k\,|f[k]| + \Delta t \sum_{m=0}^{(z-1)/2} \alpha_m n_{2m+1}. \tag{53}$$

In the limit that $\Delta t \downarrow 0$, the discrete-time and continuous-time expressions are expected to become identical. By substituting $N = 1/(\Delta t \nu_0)$ and $\Delta t n_m = t_m$ in Eqs. (51)–(53) and taking the limit $\Delta t \downarrow 0$ it can be seen that this is indeed the case.

APPENDIX 3
PROCESSING DETAILS

Here we present the details on the listening test: music preparation, algorithm design, and reproduction set.

Because the processed tracks will have a larger amplitude than the unprocessed tracks, we first need to provide some headroom. This was achieved by normalizing all tracks to full scale and subsequent attenuation by 10 dB. To achieve approximate equal loudness for all tracks, track 1 had to be amplified by a factor of 2, and track 4 by a factor of 1.5. At this time the volume setting of the playback system was also fixed. Fig. 11 shows a measurement of the frequency response of the reproduction set, which was a small portable audio set.

The bass boost was achieved by a graphical equalizer, using a PC and Cool Edit Pro 1.2 [35]. The equalizer settings were determined by trying to achieve a maximum amount of bass boost, which was chosen such as to optimize the sound quality for all tracks simultaneously. The values of the amplification per one-half-octave band are given in Table 2. The processing for the psychoacoustic bass enhancement was the same for the two different kinds of nonlinear devices. Both filters were chosen to be IIR filters in order to achieve a reproduction quality as close as possible to an eventual analog implementation. The actual processing was done on a Silicon Graphics Indy workstation, and was implemented in C programming language. This implementation enables a real-time demonstration,

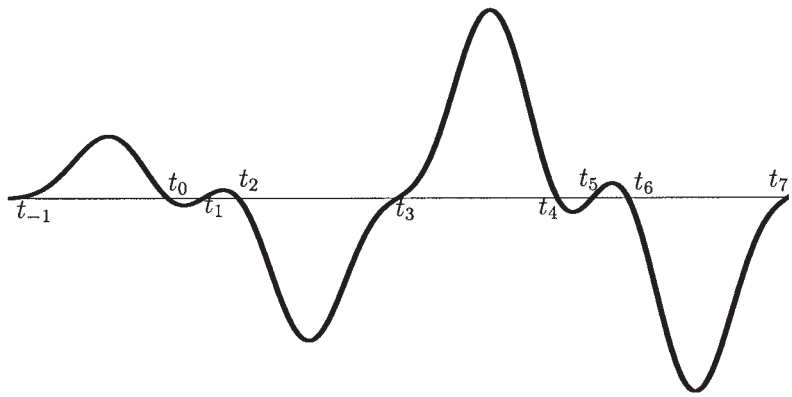


Fig. 10. Arbitrary periodic signal with zeros indicated in notation of Appendix 2. Here $z = 7$.

but in this case the processing was done on a file basis. The processing specifications are as follows.

1) The first filter, selecting the low-frequency portion of the audio signal, was implemented as a Chebyshev type I filter with 1.0 dB of ripple in the passband. The filter order was 2 (thus having two 12-dB per octave flanks). For the first filter the object is to minimize the order, so as to minimize computational complexity as much as possible. The passband was 20–70 Hz. In this case the high-frequency cutoff frequency is not quite up to the cutoff frequency of the loudspeaker.

2) The nonlinear device was either a full-wave rectifier or a full-wave integrator.

3) The second filter was an elliptic filter with 3.0 dB of ripple in the passband and 30 dB of ripple in the stopband. The order was 3 (two 18-dB per octave flanks). For this second filter the order is very important. If the low-pass flank order is too low, the timbre of the bass will be very sharp; if it is too high, the bass enhancement effect is reduced. The high-pass flank is less crucial, but choosing the order too low will give a stronger frequency content at lower frequencies, increasing the chance that distortion will occur. The bandwidth of this second filter is 70–140 Hz. Usually choosing the bandwidth of this second filter equal to an octave is a good choice. A higher bandwidth will start to give sharp timbre to the bass effect, whereas too small a bandwidth will reduce the psychoacoustic effect (because, in effect, the number of harmonics will become very small).

The gain applied to the harmonics signal now obtained was again chosen so as to maximize the bass effect without creating audible distortion. Again, for all tracks this gain value was chosen equal, the value being 15 dB. Note that we can achieve a much higher gain value than for the linear bass boost, which is maximally 9 dB.

APPENDIX 4
RAW DATA FROM LISTENING TEST

Table 3 gives the score vectors of all fifteen subjects, for each track individually as well as summed for all tracks. Using these data all other results presented in the text can be derived. The version numbers are as follows:

1 Unprocessed
2 Linear bass boost
3 Psychoacoustic bass enhancement using full-wave rectifier
4 Psychoacoustic bass enhancement using full-wave integrator

Also, the number of circular triads (CT) is indicated for each subject. A circular triad occurs in the case where the subject's response is inconsistent. This will occur if $1 > 2 > 3 > 1$, where $A > B$ indicates a preference of system A over system B. A high number of circular triads is most

Table 2. Amplification per one-half octave band for linear bass boost.

f (Hz)	Gain (dB)
44.0	6
62.5	9
88.0	9
125.0	6
PS1	15
PS2	15

PS1, PS2—psychoacoustic algorithms 1, 2. Single gain value is gain applied after second bandpass filter (see Fig. 1).

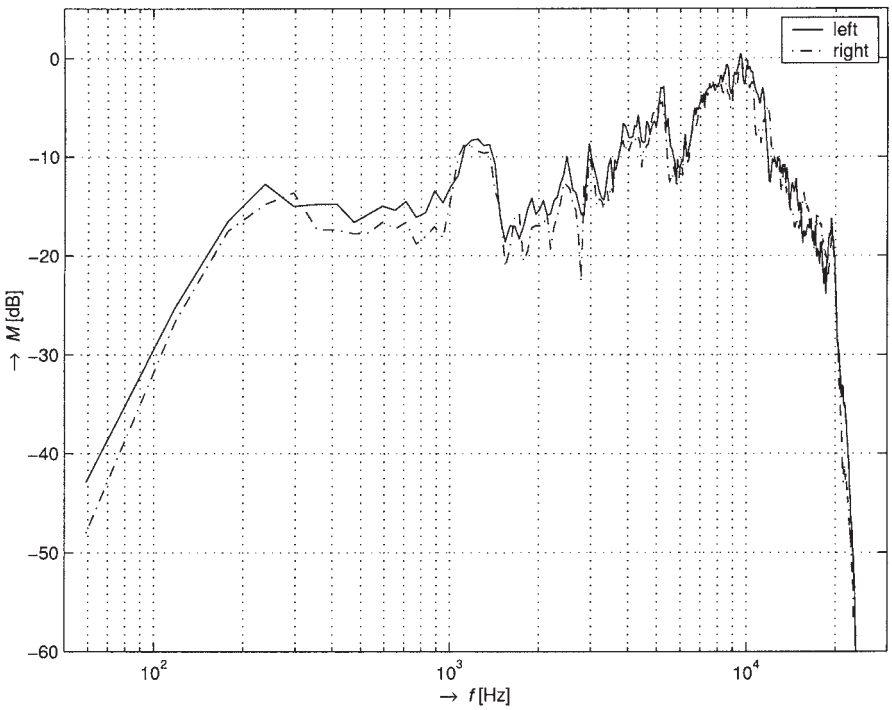


Fig. 11. Magnitude of frequency response of reproduction set.

probably an indication that the pairwise comparison process is difficult. This would necessitate some caution when evaluating the results. The number of circular triads occurring in our data is for most subjects 0 or 1, which is normal.

APPENDIX 5
BIPLOTS

In the following we present an introduction to the biplot method, largely based on Gabriel [35]. Consider an $n \times m$ rank 2 matrix $\mathcal{Y}$, which we factorize as

$$\mathcal{Y} = \mathcal{G}\mathcal{H}' \tag{54}$$

into an $n \times 2$ matrix $\mathcal{G}$ and an $m \times 2$ matrix $\mathcal{H}$, both of rank 2. This factorization is not unique. One way of factorizing $\mathcal{Y}$ is to choose the two columns of $\mathcal{G}$ as an orthonormal basis of the column space of $\mathcal{Y}$, and to compute $\mathcal{H}$ as $\mathcal{Y}'\mathcal{G}$. We can write Eq. (54) as

$$y_{ij} = \boldsymbol{g}_i^t \boldsymbol{h}_j \tag{55}$$

for each i and j, where y_{ij} is the element in the ith row and the jth column of $\mathcal{Y}$, $\boldsymbol{g}_i^t$ is the transpose of the ith row of $\mathcal{G}$, and $\boldsymbol{h}_j$ is the jth row of $\mathcal{H}$. As the $\boldsymbol{g}_i$ and $\boldsymbol{h}_j$ represent the rows and columns of $\mathcal{Y}$, they are termed *row effects* and *column effects*, respectively. These $n + m$ vectors can now be plotted in a plane, giving a representation of the nm elements of $\mathcal{Y}$ by means of the inner products of the corresponding row effect and column effect vectors. This plot is called the *biplot* since it allows the row effects and column effects to be plotted jointly.

In case the original matrix $\mathcal{Y}$ is of rank higher than 2, say r, we can approximate $\mathcal{Y}$ by use of the singular value decomposition according to

$$\mathcal{Y} = \sum_{\alpha=1}^{r} \lambda_\alpha \boldsymbol{p}_\alpha \boldsymbol{q}_\alpha^t \tag{56}$$

with $\alpha = 1, \ldots, r$, the λ_α being the singular values, and $\boldsymbol{p}_\alpha$ and $\boldsymbol{q}_\alpha$ orthonormal sets of vectors. The best fit to $\mathcal{Y}$ in the least-squares sense is then given by the rank 2 matrix $\mathcal{Y}_{(2)}$ as

$$\mathcal{Y}_{(2)} = \sum_{\alpha=1}^{2} \lambda_\alpha \boldsymbol{p}_\alpha \boldsymbol{q}_\alpha^t \ . \tag{57}$$

The goodness of fit can be defined as

$$\rho = \frac{\sum_{\alpha=1}^{2} \lambda_\alpha^2}{\sum_{\alpha=1}^{r} \lambda_\alpha^2} \ . \tag{58}$$

The approximate biplot of $\mathcal{Y}$ is then the exact biplot of

$$\mathcal{Y}_{(2)} = [\boldsymbol{p}_1, \boldsymbol{p}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{q}_1^t \\ \boldsymbol{q}_2^t \end{bmatrix} \tag{59}$$

with a goodness of fit as given by Eq. (58). Now there are a number of useful factorizations of Eq. (59) to arrive at a form as in Eq. (54), depending on what we desire to visualize in the biplot; refer to Gabriel [34].

Instead of biplotting the original data matrix $\mathcal{Y}$, we can first substract our the column averages (the mean of each variable). If we then choose for the factorization of Eq. (59)

$$\begin{aligned} \mathcal{G} &= \sqrt{n}\,(\boldsymbol{p}_1, \boldsymbol{p}_2) \\ \mathcal{H} &= \frac{1}{\sqrt{n}}(\lambda_1 \boldsymbol{q}_1, \lambda_2 \boldsymbol{q}_2) \end{aligned} \tag{60}$$

with n the number of observations, we can construct a so-

Table 3. Score vectors obtained in listening test.

Track	Version Track	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Bad	1	0	0	0	0	3	0	2	0	0	1	0	0	1	0	0
	2	2	2	3	1	2	1	3	1	2	2	2	2	3	1	1
	3	1	2	2	2	1	3	0	3	2	2	2	3	0	3	3
	4	3	2	1	3	0	2	1	2	2	1	2	1	2	2	2
Eyes	1	1	0	3	1	3	0	3	0	0	0	0	0	2	1	2
	2	1	2	1	2	1	1	2	1	3	2	1	2	2	2	3
	3	1	2	0	1	0	2	1	3	1	2	2	2	0	0	1
	4	3	2	2	2	2	3	0	2	2	2	3	2	2	3	0
Hotel	1	0	1	2	1	2	0	1	0	1	0	2	1	1	1	1
	2	1	2	3	3	3	1	2	1	2	2	0	2	2	3	3
	3	2	3	1	2	1	3	3	3	3	2	3	3	3	2	1
	4	3	0	0	0	0	2	0	2	0	2	1	0	0	0	1
Twist	1	0	0	2	0	2	0	2	0	0	2	1	0	0	0	0
	2	2	2	3	1	3	1	3	1	1	3	1	2	1	2	1
	3	2	1	0	2	0	3	1	3	3	1	1	3	3	2	3
	4	2	3	1	3	1	2	0	2	2	0	3	1	2	2	2
Total	1	1	1	7	2	10	0	8	0	1	3	3	1	4	2	3
	2	6	8	10	7	9	4	10	4	8	9	4	8	8	8	8
	3	6	8	3	7	2	11	5	12	9	7	8	11	6	7	8
	4	11	7	4	8	3	9	1	8	6	5	9	4	6	7	5
CT		2	2	0	1	0	0	0	0	1	3	2	1	1	1	1

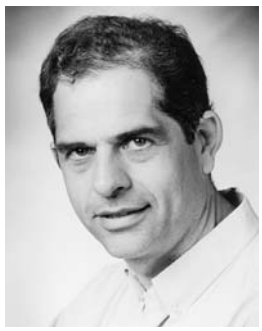
* A–O — fifteen subjects; CT — number of circular triads.

called column metric preserving (CMP) biplot. The vectors $\boldsymbol{h}_j$ now have the attractive properties that the norms of the $\boldsymbol{h}_j$ approximate the standard deviations of the variables, and the cosines of the angle between the vectors $\boldsymbol{h}_j$ approximates the correlations between the variables, as shown in Gabriel [34]. Moreover, the standardized distance between the observations is approximated by the distance between the vectors $\boldsymbol{g}_i$.

Many illustrative examples on the use of biplots may be found in Gabriel and Odoroff [36]

THE AUTHORS

E. Larsen



R. M. Aarts

Erik Larsen was born in Leidschendam, The Netherlands, in 1975. In 1998 he received the M.Sc. degree in applied physics from the Delft University of Technology. His graduation project was done in the Acoustics group, concerning variable hall acoustics, using the wavefield synthesis system developed in the group.

In 1998 he joined the Digital Signal Processing Group of Philips Research in Eindhoven, where he is working on signal-processing algorithms with application to sound reproduction.

Mr. Larsen is a member of the AES, the Acoustical Society of America, the Dutch Acoustical Society, IEEE, and the Royal Institute for Engineers. He served as papers vice chair at the 110th Convention of the AES in Amsterdam, 2001 May, and is currently a member of the board of its Dutch section. E-mail: erik.larsen@philips. com

●

Ronald Aarts was born in Amsterdam, The Netherlands, in 1956. He received the B.Sc. degree in electrical engineering in 1977 and the Ph.D. degree from Delft University of Technology in 1994.

In 1977 he joined the Optics Group of Philips Research Laboratories, Eindhoven, The Netherlands, where he was engaged in research into servos and signal processing for use in both video long-play players and Compact Disc players. In 1984 he joined the Acoustics Group of the Philips Research Laboratories and was engaged in the development of CAD tools and signal processing for loudspeaker systems. In 1994 he became a member of the Digital Signal Processing Group of the Philips Research Laboratories. There he has been engaged in the improvement of sound reproduction by exploiting DSP and psychoacoustical phenomena.

Dr. Aarts has published a number of technical papers and reports and is the holder of several patents in his fields. He was a member of the organizing committee and chair for various conventions. He is a senior member of the IEEE, a fellow of the AES, the Dutch Acoustical Society, and the Acoustical Society of America. He is past chairman of the Dutch section of AES. E-mail: ronald.m .aarts@philips.com