On the Audibility of Midrange Phase Distortion in Audio Systems*

STANLEY P. LIPSHITZ, MARK POCOCK, AND JOHN VANDERKOOY

University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

The current state of our knowledge regarding the audible consequences of phase nonlinearities in the audio chain is surveyed, a series of experiments is described which the authors have conducted using a flexible system of all-pass networks carefully constructed for this purpose, and some conclusions are drawn regarding the audible effects of midrange phase distortions.

It is known that the inner ear possesses nonlinearity (akin to an acoustic half-wave rectifier) in its mechanical-to-electrical transduction, and this would be expected to modify the signal on the acoustic nerve in a manner which depends upon the acoustic signal waveform, and so upon the relative phase relationships of the frequency components of this signal. Some of these effects have been known for over 30 years, and are quite audible on even very simple signals. Simple experiments are outlined to enable the readers to demonstrate these effects for themselves.

Having satisfied ourselves that phase distortions can be audible, the types of phase distortions contributed by the various links in the audio chain are surveyed, and it is concluded that only the loudspeaker contributes significant midrange phase nonlinearities. Confining the investigation to the audibility of such phase nonlinearities in the midrange, circuitry is described which enables such effects to be assessed objectively for their audible consequences. The experiments conducted so far lead to a number of conclusions:

1) Even quite small midrange phase nonlinearities can be audible on suitably chosen signals.

2) Audibility is far greater on headphones than on loudspeakers.

3) Simple acoustic signals generated anechoically display clear phase audibility on headphones.

4) On normal music or speech signals phase distortion appears not to be *generally* audible, although it was heard with 99% confidence on some recorded vocal material.

It is clear that more work needs to be done to ascertain acceptable limits for the phase linearity of audio components—limits which might become more stringent as improved recording/reproduction systems become available. It is stressed that none of these experiments *thus far* has indicated a present requirement for phase linearity in loudspeakers for the reproduction of music and speech.

0 INTRODUCTION

Is phase distortion audible? This innocent-sounding question continues to remain a source of controversy 140 years after Ohm formulated his celebrated acoustic "phase law" [1], [2],[1] according to which only the power spectrum (and not the relative phases of the components) of a sound determines its character. Certainly phase effects are normally sufficiently subtle that it is not surprising that the early investigators, with the relatively insensitive apparatus of the preelectronic age, were for the most part absorbed by the much more blatant audible effects caused by changes in the power spectrum of a sound. Nevertheless, exceptions to the phase law were not entirely unknown, and Helmholtz [4] was cautious to restrict its applicability to the "musical" (that is, continuous or steady-state) portion of a sound, as opposed to its buildup and decay (that is, transient) sections. Indeed, even Rayleigh [5] in 1896 cast serious doubt on the validity of Helmholtz's conclusions regarding phase inaudibility in simple conso-

* Presented at the 67th Convention of the Audio Engineering Society, New York, 1980 October 31–November 3; title revised 1982 June 29.

[1] An excellent survey article on the current state of our knowledge of the human hearing system will be found in Schroeder [3]. This reference also contains a substantial bibliography. We shall cite only a limited selection of the relevant research papers in this publication.

nances—this very early reference is worth consulting.

It is clear that phase distortion of sufficient magnitude *must* be audible. For example, if a transmission channel were sufficiently dispersive to delay one portion of the audible spectrum by, say, 1 second, with respect to the remainder, it is obvious that severe distortion of any normal musical or speech signals would occur. The question thus reduces to one of ascertaining the magnitude of those phase distortions that can be allowed without audible degradation, that is, thresholds for phase distortion must be determined. Stimulated by the obvious need for phase correction of long-distance telephone lines (see, for example, [6]), a considerable amount of careful research has been conducted to obtain audibility thresholds, usually expressed in terms of the allowable group (or envelope) delay as a function of frequency. The group delay distortions introduced by most links in the modern audio chain, including the transducers, are generally regarded as being sufficiently small as to be inaudible, on the basis of the above-determined thresholds.

And yet, there exists a substantial body of evidence that even more subtle phase effects are audible. Part of the difficulty lies in the fact that the group delay is frequently not a physically meaningful measure of the "time delay" of a frequency component of a signal through a system, as pointed out by Heyser [7]. For example, a minimum-phase system can have a group delay which is negative over certain frequency bands, but this cannot be taken to mean that it behaves acausally in these frequency bands, giving an output *before* its signal input. Formulating a useful definition of phase distortion requires consideration of what phase characteristic guarantees the *absence* of phase distortion. If $H(s)$ is the transfer function of a linear system with magnitude $A(\omega)$ and phase $\phi(\omega)$, that is,

$$H(j\omega) = A(\omega)e^{j\phi(\omega)} \tag{1}$$

then for distortionless transmission over a given frequency band it is necessary and sufficient that $A(\omega)$ be constant and $\phi(\omega)$ be linear in ω, that is, $\phi(\omega) = -\tau\omega$. Thus the frequency response must be flat, and the phase response must be a straight line *through the origin*, with slope τ representing the constant time delay to which this system subjects the signal. When $A(\omega)$ and/or $\phi(\omega)$ do not satisfy these conditions, linear distortion is present. (For a detailed discussion, see Preis [8].) The two common measures of phase nonlinearity are the phase delay

$$\tau_p \triangleq -\frac{\phi}{\omega} \tag{2}$$

and the group delay

$$\tau_g \triangleq -\frac{d\phi}{d\omega} \tag{3}$$

both in general being functions of frequency. As already intimated, neither the group delay nor the phase delay by itself is a very meaningful measure of the phase distortion. In seeking a better indicator of the deviation from phase linearity, one is led to define what Leach [9] has called the *differential time delay distortion* $\Delta\tau$ of the system as the difference

$$\Delta\tau \triangleq \tau_g - \tau_p \tag{4}$$

which is a first-order measure of this deviation. The smaller $\Delta\tau$, the smaller the first-order phase distortion, and indeed, $\Delta\tau \equiv 0$ over the relevant bandwidth is a necessary and sufficient condition for phase linearity. The related phase angle

$$\Delta\phi \triangleq \omega\Delta\tau \tag{5}$$

is likewise a measure of the deviation from perfect phase response, and has been called the phase intercept distortion [8] or the differential phase shift distortion [9]. It should be clear that it is not the actual phase shift in a system, as a function of frequency, which matters, but rather the amount by which this phase shift differs from a pure time delay, and $\Delta\tau$ is a good measure of this. It seems not to be generally appreciated how good the phase linearity of most low-pass filters is, and this has often led to ridiculous and unnecessary demands being made on component high-frequency bandwidth in a mistaken belief that such bandwidths are necessary for good time delay performance (see [9]). High-pass filters are another matter. In any event, any purely frequency-domain quantity, whether group delay or differential time delay distortion, must be interpreted with considerable caution when assessing its audible significance, since the human hearing system does not operate purely in one domain but behaves as a mixed time- and frequency-domain processor of daunting complexity.

The last few decades and, as regards transducers, especially the last few years, have seen a considerable improvement in the amplitude flatness of audio components. Some loudspeakers, pickup cartridges, disk-cutting systems and magnetic tape recorders now have impressively flat frequency responses over most of the audio band, and nonlinear distortions which are reasonably low at moderate signal levels. The stage has thus been reached when the only major respect in which technical improvement can be made in their transfer functions is as regards their phase performance. It is for this reason that it is now timely that an assessment be made of whether improvements to their phase responses would provide any audible benefits, even for the most sensitive listeners.

Before examining in more detail some of the phase aberrations exhibited by the various links in the audio chain, an important point must be made. Systems can be classified into two distinct classes which differ in a significant respect. A minimum-phase (lag) system is one in which there exists an intimate relation between the system's logarithmic amplitude *ln* $A(\omega)$ and phase $\phi(\omega)$. In fact, in mathematical terms these two functions are Hilbert transforms of each other (see, for example, [7]), so that they are not independent.[2] Most audio electronics fall into this category, as do many of the remaining components, at least over parts of the frequency range. The reason that a minimum-phase component is so desirable from the technical point of view is that this correspondence between amplitude and phase *guarantees* that equalization (by *any* minimum-phase means) will improve the phase response as the frequency response is flattened, so that in the limit as perfect frequency response is achieved, the corresponding phase response becomes perfectly linear (because of the fact that a perfect system has flat amplitude and linear phase response). To emphasize this point we remark that it follows that even mechanical systems with resonances like loudspeakers or pickup cartridges are guaranteed to be improved by (minimum-phase) equalization, provided only that they themselves are minimum phase to begin with. It is therefore wrong, as is sometimes done, to object to an equalizer on the grounds that it "rings," when this very ringing must *of necessity* occur as a consequence of its frequency response, if it is to accurately correct a complementary frequency response aberration in some other component, and result in a flatter system free from such ringing.

Those systems that are not minimum phase may have their phase response degraded as their frequency response is improved by (minimum-phase) equalization. It is for this reason important to know whether or not a system is of the minimum-phase type before one proceeds to manipulate its frequency response. (This statement is of particular relevance for loudspeaker systems, as we shall see.) Non-minimum-phase systems can be represented as the cascade of a minimum-phase system, a pure time delay, and an all-pass system [that is, one having a flat magnitude $A(\omega)$ = constant, but nonlinear phase $\phi(\omega)$]. The effect of normal minimum-phase equalization on such a system is to make its minimum-phase part perfect, so that we are left with a system whose overall characteristic is that of an all-pass network. A non-minimum-phase equalizer is required to phase linearize such a system. Magnetic tape recorders and loudspeaker systems constitute non-minimum-phase systems in general, and great care must be taken in using such components when trying to assess the audible significance of *additional* phase distortions contributed by some other component. We shall have more to say about this in the sequel. There do, however, exist a few minimum-phase loudspeaker systems with adequately flat frequency response (and hence approximately linear phase response). The amplitude and phase curves of Fig. 1 show that the original Quad full-range electrostatic loudspeaker falls into this category, and so is suitable for conducting meaningful tests of the audibility of phase distortion elsewhere in the chain. For our loudspeaker listening tests we have therefore used the Quad ESL. Still more phase-accurate loudspeakers like the new Quad ESL-63 [10] are now appearing. Similarly, electrostatic headphones are a suitable choice for experimental purposes.

1 THE STATE OF OUR KNOWLEDGE REGARDING PHASE AUDIBILITY

As intimated above, phase effects are controversial. Nevertheless, there are some well-documented results of careful experiments which demonstrate the clear audibility of certain kinds of phase distortion on suitably chosen signals. For example, Mathes and Miller [11] in 1947 and Craig and Jeffress [12] in 1962 showed that a simple two-component tone, consisting of a fundamental and a second harmonic only, changed in timbre as the phase of the second harmonic was varied relative to the fundamental. (See also [5].) This is a very basic and most musical type of signal, and can hardly be called "highly specialized." The effect is so startlingly audible that it is well worth the trouble of demonstrating it for oneself, and this can easily be done. Two sine-wave oscillators of good purity should have their outputs summed and fed in-phase to both ears via a pair of good-quality headphones. The fundamental frequency should be set at 200–300 Hz and at a comfortable sound pressure level. The combined oscillator signal should be viewed on an oscilloscope, and the second oscillator tuned to the second harmonic, and adjusted so that the relative phase between the oscillators slips 360° (that is, 1 cycle) every few seconds. The amplitude of second harmonic should be varied. It will be found that the tone changes timbre cyclically at the slip rate. If the two oscillators are phase locked, or adjusted to

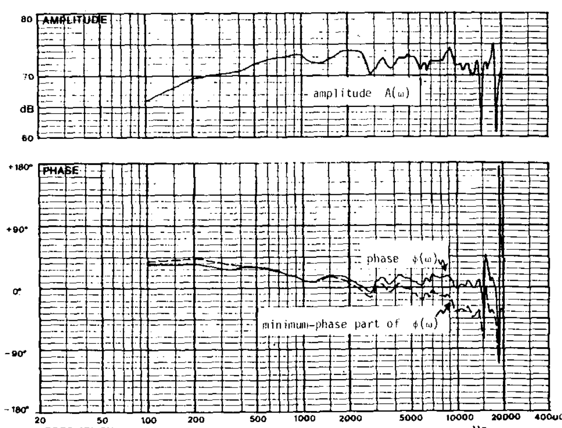

Fig. 1. Free-field amplitude and phase responses of the original Quad electrostatic loudspeaker at 1 m on axis (solid lines), shown together with the computed minimum-phase portion of its phase response (broken line). Note that, apart from a small linear-phase (i.e., time delay) difference in the computed curve, the system is minimum phase to above 15 kHz. The response irregularities are a consequence of the close microphone distance, necessitated by the measuring set-up used. At distances greater than 2 m the loudspeaker displays the flat frequency response for which it is renowned. These curves demonstrate that a loudspeaker can be both minimum phase and phase linear.

[2] The name "minimum-phase" derives from the fact that of all the possible systems with the same amplitude response, the minimum-phase system is the one having the least possible phase lag. Every other system differs from it by an all-pass function, which provides additional phase lag without changing the frequency response.

produce a stationary trace, it can be verified quite easily that the timbre is a function of the relative phase between second harmonic and fundamental. (This is most easily accomplished if one has access to a Fourier synthesizer which permits independent adjustment of harmonic amplitudes and phases.) The waveform seen on the oscilloscope will be asymmetrical positive to negative due to the presence of even (second) harmonics only. This duplicates the results of [11], [12].

But let us proceed further. Adjust the phase difference to produce a stationary pattern which is markedly asymmetrical positive to negative, and compare the sound quality when the acoustic polarity of this signal is reversed, by reversing the polarity of the connections to *both* earpieces simultaneously. This is best accomplished by a switch which allows instantaneous changeover. Again a change will be heard, and it will be found to correspond precisely to the timbre change caused by a 180° shift of the second harmonic, for this is equivalent to a polarity reversal in the case of this composite signal. This experiment can be repeated on different headphones, and also on loudspeakers, and the effect will still be found to be audible, although not as clearly so on loudspeakers unless conducted in an anechoic chamber, due to standing waves in the room. An objection can be raised that any transducer nonlinearity could account for the effect. This can be resolved in a number of ways. First, two transducers could be used, one for each oscillator signal. This would prevent any possible intermodulation, but does not remove the effects of harmonic distortion in each separately. It also introduces path-length difference effects which mean that the listener must not move. Second, one can verify by measurement of the distortion in the transducer's acoustic output that, at the levels and frequencies involved, they are adequately linear. Most available transducers satisfy this requirement easily. Third, if the transducer is of symmetrical planar construction, and so can be auditioned from front and back (such as some electrostatic [10] and planar dynamic designs), it can be absolved *totally* as a contributory factor by the simple expedient of comparing the sound from its front with that from its back when feeding it with an electrical signal of reversed polarity. For, by listening from the back, the acoustic polarity of any transducer asymmetries is reversed (as is the signal); hence simultaneously reversing the electrical signal while listening from the back results in an acoustic signal of the same polarity as originally, but with all transducer asymmetry (that is, nonlinearity) contributions reversed in polarity. If no difference can be heard, it can be deduced that the transducer's asymmetries are below audibility, and so not a contributory factor. We have in fact performed this verification for ourselves on a planar loudspeaker. Indeed, these phase effects have proved audible on all reasonable transducers tried. In this connection it is interesting to note that not all transducers are of the same acoustic polarity, that is, not all produce an acoustic compression in response to a positive going electrical signal. (The original Quad ESL, for example, is polarity-inverting.) Transducer polarity can actually be assessed solely on the basis of the timbre heard when listening to the simple two-tone signal discussed above. This experiment also suggests that an acoustic polarity inversion may be audible on music and speech, and this is indeed true, although the effect is *much* more subtle than those described above.

The authors have demonstrated the two-tone experiment described above to numerous people on different systems. No one has ever failed to hear the timbral change with phase, and discern the polarity reversal on this signal with unvarying accuracy. Indeed, in a double-blind demonstration to eleven members of the SMWTMS audio group [13], the accuracy score was 100% on the summed 200-Hz and 400-Hz tones over loudspeakers, and overall, including musical excerpts, the results on the audibility of the polarity inversion of both loudspeaker channels were 84 correct responses out of 137, this representing confidence of more than 99% in the thesis that acoustic polarity reversal is audible. (See also [14], [15].) Some designers [16] nevertheless still believe this effect to be inaudible.

Let us pursue the polarity reversal question a bit further. The inversion of the polarity of a time signal $[f(t) \rightarrow -f(t)]$ is equivalent to a constant phase shift of π radians in its complex Fourier transform. This is a nonlinear phase distortion (in fact, phase intercept distortion), even though the group delay is zero (that is, no dispersion), for the phase curve is not a straight line through the origin. It leads to severe waveform distortion—in fact, the interchange of positive and negative polarities in the time domain, of course. Now, many musical and speech sounds are markedly asymmetrical [17]. Furthermore, it is now well established that the inner ear behaves largely like a half-wave rectifier (see, for example, [3]) with neural output from the acoustic nerve occurring predominantly during the rarefaction half of the acoustic waveform, for signal frequencies below about 1 kHz. Similar results have been found to apply to other animals, and interestingly enough, a simple creature like the sprat has been found to have hair cells of *two* types, one group responding specifically to compressions, while the other responds to rarefactions [18]. The important point is that there is a well-established mechanism in the inner ear for detecting waveform asymmetries and hence polarity reversal of asymmetrical signals.[3] What is perhaps surprising is how subtle this effect generally appears to be on music and speech. As the above-mentioned experiment [13] indicates, however, it *is* an audible factor, and should be taken into account when performing comparisons of

[3] This is all too often overlooked, and numerous null experiments have been reported on the audibility of phase and polarity effects, based on listening tests with continuous signals like square waves, which are symmetrical positive to negative (that is, which contain only *odd* harmonics), and remain symmetrical irrespective of the type of phase shifting to which they are subjected. Such experiments are not justified in drawing any conclusions about asymmetrical waveforms, and so cannot contribute to a resolution of the issue.

audio components [15], [19], [20]—acoustic polarity should be maintained. In fact, for this very reason, Stodolsky [21] recommended in 1970 that polarity standardization be adopted for all audio components, something which is easy to implement, but has still not been done. There exist polarity standards only for microphones (of necessity, due to potential phase problems in multimiking), and unofficial standards for pickup cartridges (necessitated by CD-4 carrier disks) and loudspeakers. Even these are not uniformly followed. For audio electronics, tape recorders, disk records, and other components no standards have been adopted. We strongly advise the standardization of magnetic (analog and digital) tape recorder and of phonograph disk polarities. A recent proposal [22], [23] for tape recorder polarity unfortunately only exacerbates the situation by recommending precisely the opposite to the standard which Stodolsky recommended 10 years ago.[4]

Another very simple experiment can be used to demonstrate conclusively that the inner ear responds asymmetrically. It is common experience that reversing the polarity of only *one* channel of a pair of headphones produces a markedly oppressive and very audible effect on both monaural and coincident stereophonic material. The effect is predominantly one affecting frequency components below 1 kHz, as can be verified by listening to filtered music or noise signals restricted to frequency bands below and above 1 kHz. Since the reversal of the polarity does not introduce any time-delay or dispersive effects into the acoustic signal, but merely changes compressions into rarefactions and vice versa, the audible effects are due solely to the constant 180° phase shift which polarity reversal entails. Since no interaural cross correlations occur before the olivary complexes to which the acoustic nerve bundles connect, it *must* follow that the acoustic nerve output from the cochlea is *changed* by the polarity reversal. This change is due to two factors: the cochlea is now responding to the opposite polarity half of the waveform, and this waveform has a shifted time relationship relative to the signal heard by the other ear. This is merely a very simple confirmation of the asymmetry (that is, half-wave rectifying nature) of the inner ear. This experiment does *not* demonstrate that polarity reversal of *both* channels (or monaural polarity reversal if the signal is applied to one ear only) is audible. What it *does* show is that the neural output signal from the ear is phase sensitive, and this suggests that, when further processed by the brain, it would be surprising if no monaural, that is, single ear, phase effects were audible. The earlier demonstrations mentioned above serve to confirm this audibility, although the effect is more subtle than one would anticipate. The mechanism responsible for this pronounced transduction asymmetry appears to be the hair cells, which respond unilaterally to motion of the cilia induced by fluid and basilar membrane vibrations in the cochlea. This is well documented [3]. Individual neurons in the acoustic nerve can, for brief periods, fire at up to 1000 times per second, although for continuous stimulation, rates of a few hundred pulses per second are the norm. Since these firings occur predominantly during acoustic rarefactions, it follows that the ear has, at frequencies below 1 kHz, the ability to "follow" the negative half of the waveform of an acoustic stimulus by modulating the neuron firing rate in sympathy with the signal waveform. This feature can explain many, if not most, of the phase and polarity effects discussed above.

[4] We have therefore suggested [24] that Stodolsky's standard be adopted, and have prepared a polarity test tape [25] which permits determining the polarity of a record/reproduce system relative to Stodolsky's convention. We will gladly supply a short segment of this test tape on request.

To conclude this section, we would like to give a few specific references to the literature to guide the further reading of those who may be interested in pursuing the subject in greater depth. The paper by Hansen and Madsen [26] may fairly be classed as one of the most influential in recent years. The authors discuss the effect on single sinusoids of an adjustable dc offset and starting phase, and find pronounced phase effects. In [27], Berkovitz and Edvardsen survey the field and comment further on the results reached in [26]. The research of Plomp and Steeneken [28] enables a comparison of phase sensitivity effects with those due to amplitude (that is, frequency response) changes, and is most interesting. In [29] Cabot *et al.* use as their test signal a 400-Hz fundamental with third harmonic added, this producing only symmetrical waveforms, and consequently leading to much less pronounced phase effects than those displayed by asymmetrical signals. Blauert and Laws [30] use specially constructed all-pass filters in their phase distortion tests, but are forced to bandlimit their test signals rather severely in order to stay within the flat response region of their all-pass filters. Consequently their pulse test signals were already quite badly dispersed before being subjected to further phase shifting. The paper is nevertheless interesting and reflects many of the ideas which we adopted in the experiments to be outlined below. Most recently, Suzuki *et al.* [31] have performed careful phase audibility experiments reinforcing many of our conclusions in the sequel.

An argument frequently put forward to justify why phase distortion cannot be significant for material recorded and/or reproduced in reverberant surroundings is that the reflections cause gross phase distortions themselves, which are very position sensitive. This is true, but in both cases the first-arrival *direct* sound is not subject to these distortions, and very important directional and other analyses are conducted during the first few milliseconds after its arrival, and before the predominant reverberation arrives. We do not accept that the presence of reverberation renders phase linearity irrelevant, and for confirmatory evidence refer to a recent paper by Bridges [32]. Some fascinating new work on the overriding importance of phase preservation in at least the long-term spectrum of acoustic signals is given in [33], [34], and leads one to speculate about possible experiments which could be conducted to re-

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.