Red Shift: Doppler distortion in loudspeakers

By Keith Howard • Posted: Nov 21, 2004 (Stereophile)

In the world of digital audio, jitter has been a focus of audiophile attention for well over a decade. It is blamed for many of the sonic ills of which CD and other digital media have been accused. But here's a puzzle: The major source of frequency intermodulation distortion in audio systems—the loudspeaker—has largely escaped such withering inquiry. Why?

Two reasons spring to mind. First, the common origin of these distortions is obscured by the fact that they go by different names and are quantified differently. Whereas frequency intermodulation (FIM) in the digital context is called *jitter* and specified in units of time (because it is considered from the viewpoint of cause rather than of effect), in loudspeakers it is usually called *Doppler distortion* and specified, if at all, as a percentage figure, to align it with other forms of nonlinear distortion.

Never heard of Doppler distortion? Well, that reflects the second reason: Doppler effects in loudspeakers came and went as an issue between the 1960s and early 1980s, the consensus at the end of that period being that FIM distortion in typical hi-fi speakers is inaudible and therefore irrelevant. For many, that closed forever the subject of Doppler distortion. But accepting accepted wisdom can be a hazardous act of faith, particularly when it relies on listening tests conducted years ago and far away.

In what follows, therefore, I describe the results of my own recent efforts to assess the significance of Doppler distortion in loudspeakers, and determine whether the verdict of "irrelevant" is the right one. But before I do that, I must paint in some background: the origins of Doppler distortion in loudspeakers, the details of earlier research, and how Doppler relates to jitter. I also make a short detour into DiAural's "Doppler Decoding," which claims to exploit a loudspeaker's Doppler distortion to cancel that from the recording microphone(s).

Freight train, freight train...

The classic example of the Doppler effect is a locomotive sounding its horn as it passes a nearby listener. As the engine approaches, the wavelength of the sound is compressed and consequently increased in pitch; as it passes the listener and advances into the distance beyond, the wavelength is stretched and the pitch drops. As the locomotive passes, the result is the familiar *Waa-ooo* effect.

Something similar happens when a loudspeaker diaphragm reproduces a sound. While the diaphragm is moving toward the listener, the frequency of the radiated sound increases; as it moves away, the frequency decreases. But the effect is much less obvious here because, even during the large diaphragm excursions required at low frequencies, a loudspeaker cone moves relatively slowly. If a 100Hz tone is reproduced at a cone excursion of 1/2" from peak to peak of the waveform, for instance, the maximum cone velocity (as it passes through the resting position) will only be 4 meters per second—a leisurely 9mph. And as frequency increases, diaphragm velocity decreases for the same sound-pressure level.

With a pure tone, the Doppler effect simply adds harmonic distortion: principally, relatively innocuous second harmonic at practicable diaphragm excursions. But with the complex signals of music, where the diaphragm must reproduce lower and higher frequencies simultaneously, the consequences are more serious. In effect, the higher tones are frequency-modulated by the lower tones, giving rise to intermodulation sidebands. If the 100Hz tone of the previous example were accompanied by another at 1kHz, then FIM sidebands would appear at the sum and difference frequencies (900Hz and 1100Hz, respectively). The amplitudes of these sidebands relative to the higher frequency component depends on the cone velocity generated by the lower tone, and increases with the frequency difference. So if the higher frequency were 3kHz rather than 1kHz, the level of the sidebands would be 3*x* (9.5dB) higher. Of course, the frequency content of the typical music signal is a great deal more complex than this.

So is the pattern of FIM products—frequency modulation generally produces an infinity of sidebands, not the single pair mentioned above. In the case of loudspeakers, the modulation index is usually so low that the higher-order sidebands are at lower amplitude, for which reason it is established practice to consider only the first-order sidebands. But that shouldn't be taken to mean that higher-order components can be ignored. Fig.1 shows a simulated spectrum for the 100Hz/3kHz example above (with a cone excursion of 0.5" peak–peak). Although the first-order sidebands—as predicted by theory, –15.2dB relative to the 3kHz component—have the highest amplitude, the higher-order sidebands are at a level that can hardly be called insignificant.

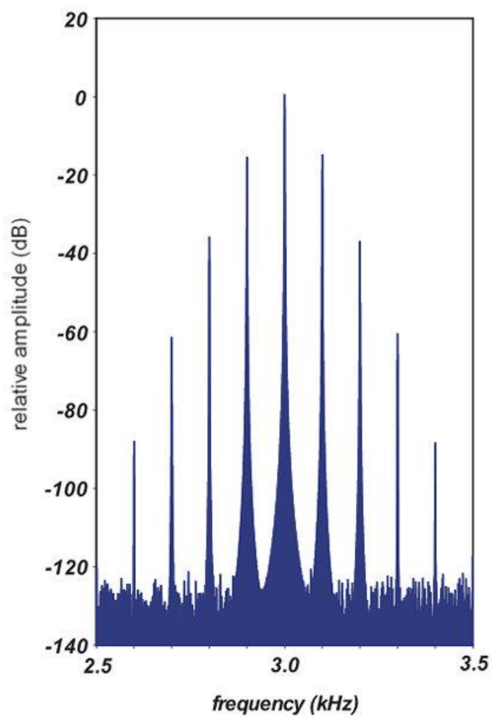


Fig.1 Doppler spectrum for a 3kHz tone reproduced by a loudspeaker diaphragm also radiating a 100Hz tone with a peak–peak excursion of 0.5". For a 200mm (8") cone, this is equivalent to an SPL of 92.5dB at 3m (10'), assuming free-space conditions.

This is an admittedly simplified description of Doppler distortion in loudspeakers. (For a more complete mathematical treatment, I recommend Siegfried Linkwitz's Web page at www.linkwitzlab.com/frontiers.htm#J.) Still, two key messages can be taken from it.

First, the amount of Doppler distortion generated by a loudspeaker is dependent on the amplitude of the signal's low-frequency content, because this is the principal determinant of diaphragm velocity. This means that higher levels of Doppler distortion will be generated by music program with strong bass content.

Second, Doppler distortion worsens as the frequency difference between the modulating and modulated frequencies increases. This inevitably means that, for a given diaphragm size, Doppler distortion will be worst in loudspeakers that use a single, full-range drive-unit. Two-way speakers will suffer less, and three-ways with a lower bass/midrange crossover frequency will suffer less still. For this reason, my simulation models a two-way speaker, because this represents the worst case most of us are likely to encounter in an audiophile system.

Prior art

Interest in Doppler distortion began in 1943, with the publication of a paper by two RCA engineers, G.L. Beers and H. Belar[1], which was the first to identify and quantify this distortion mechanism. Interest in the subjective effects of Doppler distortion intensified in the 1960s with the work of James Moir [2]in the UK and Paul Klipsch [3]in the US, both of whom concluded that it had a significant effect on speaker sound. But these findings relied in major part on comparisons of speakers with different levels of Doppler distortion, the assumption being that this was their only substantive disparity. As other researchers later argued, this was quite possibly wishful thinking. Other factors may have been responsible for the differences.

The first concerted attempt to sidestep this experimental difficulty was made in the UK by Peter Fryer[4], who was then working at Rank Hi-Fi (Wharfedale). In order to be certain that Doppler distortion was the only variable in his listening trials, Fryer built a "Carrier Canceling Doppler Machine"—based on a variable-length, bucket-brigade delay line—to simulate the effect electronically. The conclusion he reached from his listening was that, even when you could hear it, Doppler distortion "is not at all unpleasant up to quite high levels as long as the modulating frequency lies between say 20 and 100Hz," and that "Doppler distortion is unlikely to be of importance in most speaker systems." Detectability, he found, corresponded with a total cone movement (peak–peak) of about 10mm (0.4").

Roy Allison and Ed Villchur[5] later reached broadly similar conclusions using a different experimental technique, although they pointed out that Fryer's detection criterion needed to be relaxed in various ways to reflect real-world loudspeaker use. First, Fryer's work had assumed a

[1] "Frequency-Modulation Distortion in Loudspeakers," reprinted in the *Journal of the Audio Engineering Society*, Vol.29 No.5, May 1981.
[2] "Doppler Distortion in Loudspeakers," *Hi-Fi News*, January 1967. See also Moir's paper of the same title, presented at the AES 46th Convention, September 1973.
[3] "Modulation Distortion in Loudspeakers," *Journal of the Audio Engineering Society*, Vol.17 No.4, April 1969.
[4] "Simulation and Investigation of Doppler Distortion," AES 56th Convention, March 1977.
[5] "The Audibility of Doppler Distortion in Loudspeakers," AES 70th Convention, October 1981.

full-range drive-unit, whereas most high-quality speakers divide the frequency range between two or more drivers. Second, Fryer had taken no account of the fact that the amount of Doppler distortion generated by a moving diaphragm varies with the angle off the forward axis until, at right angles to the direction of diaphragm motion, it falls to zero. As a result of this, the early reflections and reverberant sound in the listening room contain lower levels of Doppler distortion than the direct sound, which, Allison and Villchur claimed, will further raise the threshold of detection and thus make the generation of audible Doppler distortion by most hi-fi loudspeakers even less likely.

Many loudspeaker designers seem to have taken Allison and Villchur's work as the last word on the Doppler issue. If you search the Audio Engineering Society archives at www.aes.org, you will find that, since that presentation in 1981, no paper on the audibility of Doppler distortion has been presented at an AES convention or published in *JAES*. But a few audiophiles keep the Doppler flame alight—notably horn aficionados, who, taking their lead from the late Paul Klipsch, claim low levels of Doppler distortion to be one of the reasons that horn-loaded speakers surpass direct-radiating alternatives.

Doppler distortion may have fallen out of the headlines for much of the past 20 years, but, as alluded to earlier, the same cannot be said of another species of frequency intermodulation distortion: jitter. Jitter in digital systems has been a focus of audio-industry attention for well over a decade, and is widely credited with being one of the reasons that digital audio initially had such a bad reception in the audio press. So let's compare the level of FIM distortion introduced by modern digital source components with that generated by analog sources and loudspeakers. Are they in the same ballpark?

Fig.2 does this by combining data from various sources. The blue line shows the proposed tolerance limit for sinusoidal jitter suggested by W.I. Manson of the BBC in the early 1970s, while the red line represents the most stringent National Association of Broadcasters (NAB) standard for wow and flutter in tape machines running at 15ips. Note that, in both cases, the tolerance limit is dependent on the modulating frequency, with much larger amounts of FM distortion being acceptable when the modulating frequency is low—as it will predominantly be with loudspeakers. The black triangles indicate jitter-detection thresholds recorded by eight people, each of whom listened to three different music excerpts in a recent test conducted by Dolby [6]; the solitary pink diamond shows the threshold of detectability of Doppler distortion identified by Peter Fryer.

[6] Eric Benjamin and Benjamin Gannon, "Theoretical and Audible Effects of Jitter on Digital Audio Quality," AES 105th Convention, September 1998.

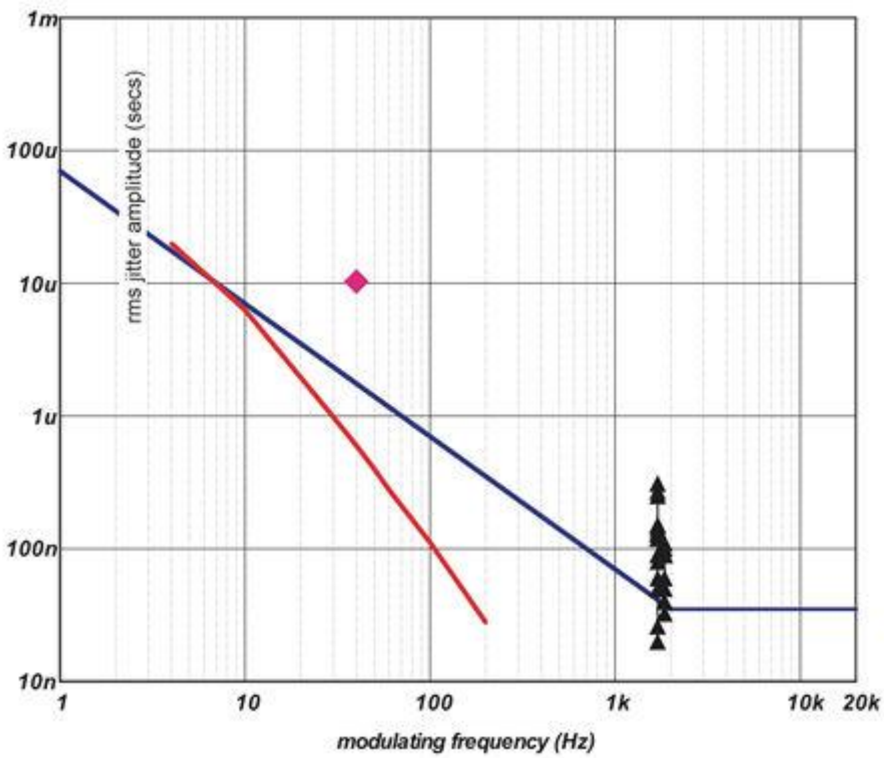


Fig.2 Manson's suggested jitter audibility threshold (blue line), NAB wow and flutter standard (red line), Dolby's jitter listening-test results (black triangles), and Fryer's Doppler audibility threshold (pink diamond).

The fact that all these data agree so well, though they were generated across three decades and in different circumstances, is no more than we might expect, given that the distortion mechanism is fundamentally the same in each case. What will surprise readers who have never seen this information before is how widely these detection thresholds differ from the measured jitter performance of modern CD players, the best of which have total (RMS) jitter levels, measured across a wide frequency range, two orders of magnitude less than anything plotted here.

There are two ways to regard this large disparity. You could conclude that it shows that much of the fuss about jitter is misplaced, and that most modern CD players, even inexpensive ones, generate levels of jitter that are inaudible. But few audiophiles who have heard the sonic improvement that can be wrought by upgrading the clock circuit in many CD players are likely to accept that view. What seems more likely, in the light of their experience, is that fig.2, for whatever reason, doesn't tell the whole story about the audibility of FIM.

As for Doppler distortion, the situation is too complex to be assessed in any meaningful way from this diagram because, in this case, the audio waveform is both the modulating and the modulated signal, which makes this form of FIM unusually complex. The only reliable way to assess whether typical hi-fi speakers produce audible or inaudible amounts of FIM is to listen to music signals with and without representative levels of Doppler distortion added, and to hear if there is an audible difference.

A new experiment

Peter Fryer's hardware approach to Doppler-distortion simulation has the advantage that while listening, you can literally turn a knob to increase or decrease the level of distortion at will. The disadvantage, as with any hardware solution, is that the Doppler engine must be designed, built, and paid for, and you have to make certain it isn't doing anything else to modify the signal other than generating FIM distortion. A software approach that generates preprocessed files for listening, though much less flexible, costs nothing more than the time required to write the code. It is also pretty straightforward to ensure that it behaves exactly as it should.

To simulate Doppler distortion digitally, you must generate output samples where there are no corresponding input samples; in other words, it's a process of interpolation. There are various ways to do this, the simplest of which is first to upsample the input signal by a large factor to generate a much denser grid of sampling points, and then to use linear interpolation to calculate the output value at the precise point it is required. Linear interpolation is inherently inaccurate, but errors can be reduced to an arbitrary level by making the upsampling ratio sufficiently high.

Fig.3 illustrates the worst-case situation for linear interpolation, where the curvature of the signal waveform is at its greatest. For a sinewave, this occurs at the waveform's positive or negative peak, as illustrated. The black curve is the waveform, the red dots indicate adjacent sampling points symmetrically disposed about the waveform's peak value, and the blue dots indicate the interpolated values. Clearly, the error at the waveform peak is the largest. By specifying (1) the amplitude and frequency of the sinewave, and (2) how small we wish the error to be, it is possible to calculate the maximum sample spacing, and hence the sampling rate required to achieve this performance. If we take a full-scale sinewave at 5kHz as our criterion and require the maximum error to be 0.1 of a least significant bit (*i.e.*, a tenth of the quantization step size) for 16-bit coding, then the required sampling rate turns out to be 6.358MHz. For test material ripped from CD (as I've used in my simulations), the required upsampling ratio is therefore 144x. To simplify the computation, this was applied in four stages, of 2*x*, 3*x*, 4*x*, and 6*x* upsampling, respectively.

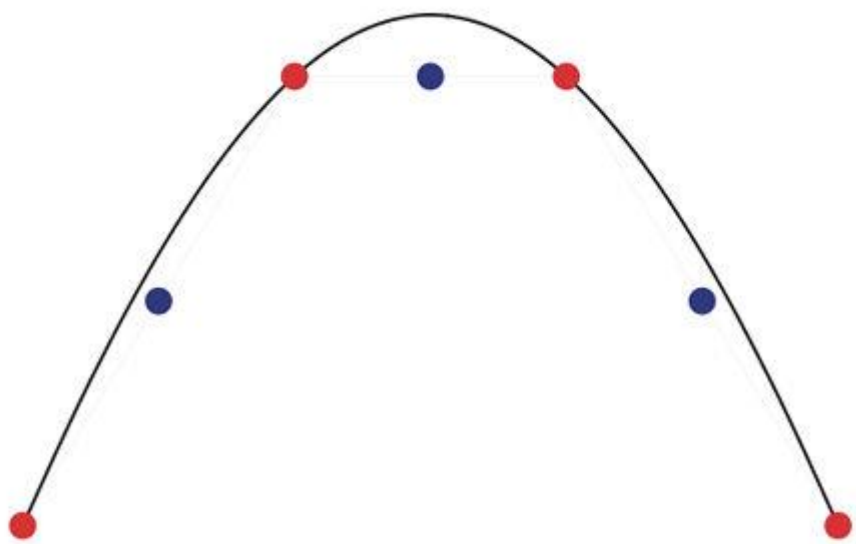


Fig.3 Linear interpolation: the error is largest at the peak of the waveform, where its curvature is highest. Red dots indicate input samples, blue dots interpolated samples.

Each of these upsamplings requires the application of a low-pass digital filter, and these filters should obviously have minimal influence on signal quality. To this end, I designed custom filters with a flat passband to 20kHz and a cube-law rolloff thereafter. (This obviates the passband ripple and excessive time smear of filters designed using equiripple techniques.)

Even when this degree of care is taken over interpolation accuracy and filtering, there remains the possibility that some subtle signal degradation will occur, not least because of the addition of TPDF (Triangular Probability Density Function) dither, prior to requantizing the processed signal, raises the noise floor. To account for this, I included in the software a “Generate Reference” option, which applies the same processing steps without the addition of frequency modulation, and generates all output samples from points midway between interpolated samples to ensure that the maximum interpolation error is included. It was this reference file, not the original, that was compared with the processed file in the listening test.

A block diagram of the principal processing stages is shown in fig.4. Note that the first step is to divide up the frequency range as it would be in a typical two-way speaker, with only the low-frequency range being subject to Doppler processing. A crossover frequency of 3kHz was chosen, with fourth-order Linkwitz-Riley high-pass and low-pass filter alignments to ensure a flat frequency response when the two portions of the frequency range are reunited. A conventional Linkwitz-Riley crossover introduces phase distortion, but that was obviated here by using a bidirectional IIR (Infinite Impulse Response) filtering technique. Second-order filtering is applied to the signal in both time directions (forward and backward) to achieve the required fourth-order slopes with linear phase.

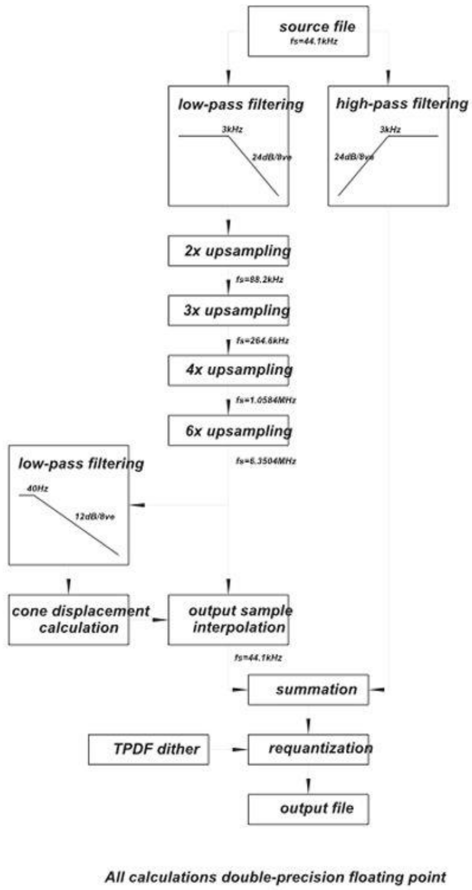


Fig.4 Block diagram of Keith Howard's Doppler-simulation software.

Assuming closed-box bass loading with a total system Q of 0.5 allows the same technique to be used to model the speaker's bass rolloff, although in this case each filter pass is first-order. A bass resonance frequency of 40Hz was chosen for the test, so the response of the modeled speaker would be –6dB at this frequency. (There can't be many two-ways with a 40Hz bass resonance, but, as I said, this *is* a worst-case simulation exercise.) Although this rolloff was used to calculate the cone displacement for the Doppler processing, it was *not* applied to the signal itself.

Eagle eyes will note that no low-pass filtering is applied as part of the decimation process (output sample interpolation). This is because the initial 3kHz rolloff, coupled with the falling spectral content of the test excerpts, makes it unlikely that significant aliasing will be introduced. In any case, if audible aliasing does occur, then this will be (largely) repeated in the comparison file.

Fig.5 shows three spectra, generated during the software's verification, that indicate that any errors are at a very low level. Fig.5a shows the spectrum of the test signal, comprising three sine components, at 100Hz, 1kHz, and 10kHz. Fig.5b shows the spectrum of the reference generated by the simulation software with no Doppler applied (note the raised noise floor caused by the additional dithering), and fig.5c that for the Doppler-processed signal.

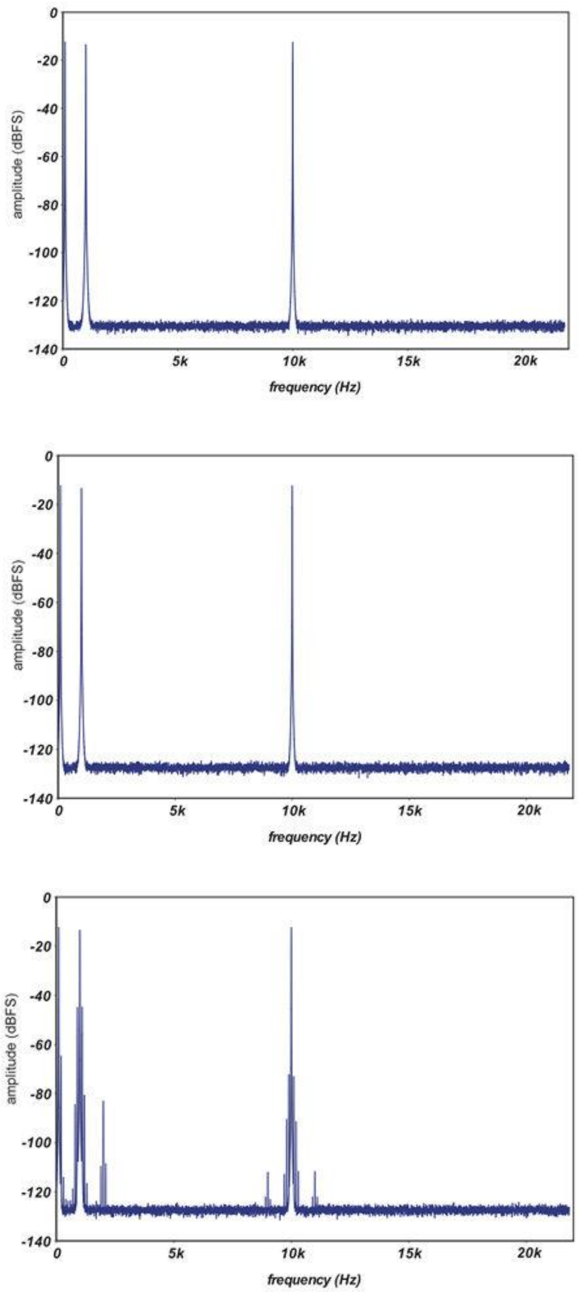


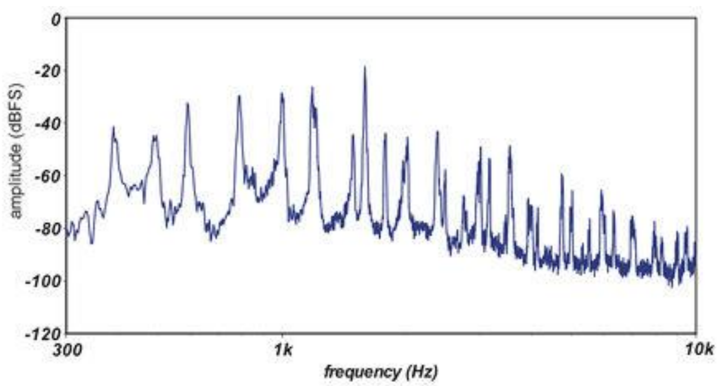
Fig.5 Spectra of a three-tone test signal used to verify the software: (a) input signal, (b) processed comparison file, (c) the full Doppler output file.

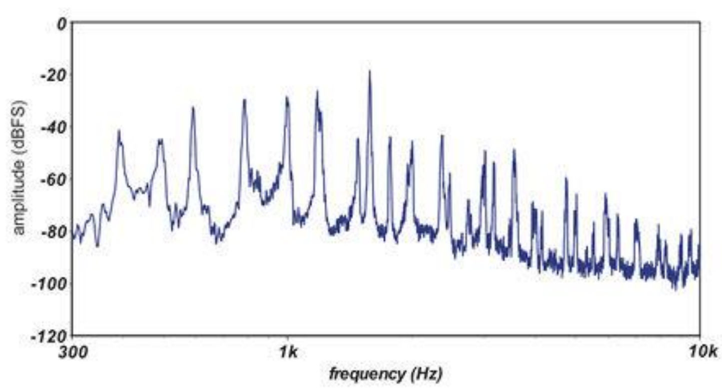
Listening results

Choosing appropriate source material is key to any listening test for Doppler distortion. Ideally, you're looking for sustained high-level bass content coupled with a clearly defined but sparse harmonic structure at higher frequencies, so that there will be minimal masking of the FIM sidebands. You can either sift available recordings for suitable examples or, more practically, concoct your own, which is what I did.

Instruments of the flute family are ideal for providing the delineated high-frequency harmonic structure. I used the single-instrument flute recording (track 12) from the European Broadcasting Union's CD of Sound Quality Assessment Material (SQAM). A short excerpt was taken from this—a few seconds of the right material are quite enough for detecting whether Doppler sidebands are audible—and mixed with a 100Hz sinewave at –10dBFS to ensure constant high-level low-frequency excitation. This file was then processed to simulate peak diaphragm excursions of 1mm, 3.16mm, and 10mm (*ie*, in 10dB steps), following which the 100Hz component was (largely) removed from the output files by mix-pasting it a second time, but in antiphase.

All the files, including the processed reference, were then burned to CD for replay via a low-jitter CD player (the Meridian 508.24). High-quality headphones (Precide Ergo AMT) were used for the listening rather than loudspeakers, to ensure the minimum of residual Doppler and other nonlinear distortions. Spectra for the left channels of the original, reference, and processed files are shown in fig.6 for comparison.





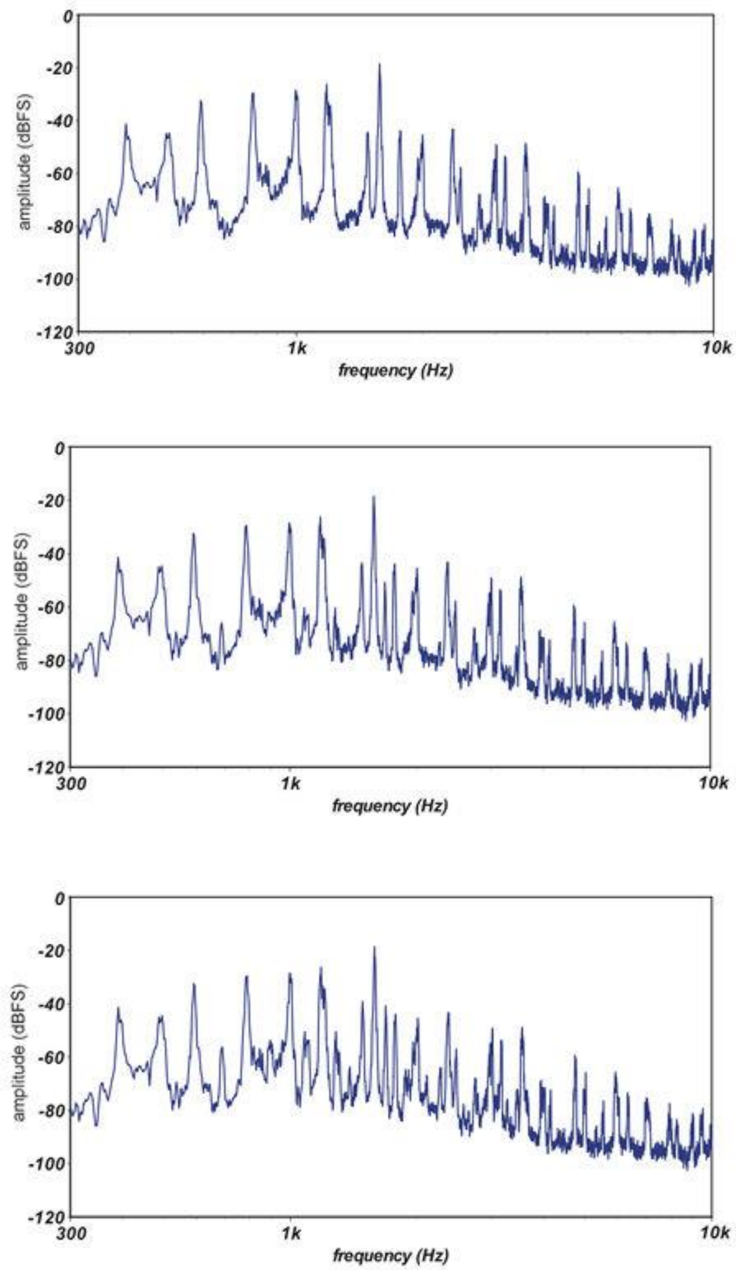


Fig.6 Spectra (300Hz–10kHz) of the flute excerpt used for the listening: (a) original, (b) processed reference, (c) Doppler 1mm, (d) Doppler 3.16mm, (e) Doppler 10mm.

The results were intriguing. Distortion of the flute was gross at 10mm peak diaphragm displacement and not in the least bit euphonic. On the contrary, Doppler made the sound as harsh as you might expect of a distortion mechanism that introduces intermodulation products. At 3.16mm peak displacement (below Fryer’s suggested detectability threshold) the distortion level was obviously lower but still clearly audible; **and even at 1mm it could still be heard affecting the flute’s timbre and adding “edge.”**

Everyone who uses a two-way speaker (me included) can take heart from the fact that most music signals are less revealing of Doppler distortion than this special brew. But these findings undermine the view, widely accepted in the last two decades, that Doppler distortion in loudspeakers is not something we should trouble about. **Having done the listening, I side with Moir and Klipsch more than with Fryer, Allison, and Villchur on this issue**—something that may come as no surprise to anyone who has heard the effects of low-level jitter and sees where the Fryer criterion appears in fig. 2.

It has often been claimed that, with a two-way speaker, there are audible benefits to using a crossover frequency below the typical 3kHz, the usual explanation being that this removes the crossover from the ear's area of greatest sensitivity. But I wonder. Perhaps this not-uncommon experience actually has much more to do with the "D" word. A three-way solution is potentially even better. Three-way speakers bring new design challenges, of course, in particular the need to achieve another perceptually seamless handover between drivers. But from the Doppler perspective, having a crossover for the bass driver at 400Hz or 500Hz is, unquestionably, better.

Sidebar: Doppler and DiAural

Although the initial fuss that accompanied Ray Kimber's 1999 launch of DiAural—the proprietary speaker-crossover technology developed by Eric Alexander—has died down, the fact remains that it was and still is promoted as a means of canceling the Doppler distortion introduced by microphones: "Doppler Decoding," in DiAuralspeak. To my knowledge, however, this claim has never been challenged in the audio press. Although DiAural's two patents (US 6,115,475 and 6,310,959 B1, copies of which you can download from the US Patent Office make no mention of Doppler distortion, it has been a central plank of the company's marketing message—as reported, for instance, in Barry Willis's Web news reports on February 22, 1999 and April 4, 1999), and in Ken Kessler's writeup in the June 1999 *Hi-Fi News*.

How credible is the claim that a loudspeaker, whatever its crossover configuration, can cancel microphone Doppler distortion?

The first problem with this notion is that moving-coil loudspeaker diaphragms are mass-controlled through their working range, whereas high-quality (*ie*, capacitor) microphones are stiffness-controlled. This results in quite different behavior in respect to diaphragm displacement *vs* frequency. Whereas, assuming constant sound-pressure level, the loudspeaker diaphragm's displacement decreases with the square of frequency (at least, until it starts to become significantly directional), the microphone diaphragm's displacement is constant. It follows that, as a result, their Doppler distortions vary quite differently with frequency. To further complicate matters, dynamic microphones are resistance-controlled and ribbon microphones are mass-controlled, so diaphragm displacement *vs* frequency behavior is different for different microphone types.

Second, for Doppler Decoding to work even notionally, it would have to rely on the recorded and reproduced SPLs being matched, and on absolute polarity being maintained throughout the recording/reproduction chain. Obviously, these conditions are unlikely to be met, particularly in the case of multimiked or multitracked recordings.

Third, for a microphone to introduce significant levels of Doppler distortion, its diaphragm would have to undergo large excursions—of the same order as the loudspeaker diaphragm, if cancellation is to be feasible. In fact, microphone diaphragm excursions are minuscule. I asked Stephan Peus, president of development at Georg Neumann GmbH, to provide me with some representative figures. He e-mailed back a document, "Some Amazing Facts with Condenser Microphone Capsules," which quotes the diaphragm excursion for Neumann's KM 184 microphone (a miniature cardioid) as being just 10 nanometers—that's four-tenths of a millionth of an inch, or about a 40th the wavelength of blue light—at an SPL of 94dB. Compare this with the 7.5mm peak excursion required to generate this SPL at 100Hz in free space and at 3m (10') listening distance, using a drive-unit of 200mm (8") effective diaphragm diameter. The two figures differ by a factor of 750,000! Depending on their mechanical characteristics, the diaphragms of other capacitor microphones may undergo larger excursions, but their displacements will still be orders of magnitude smaller than a loudspeaker cone's.—**Keith Howard**

Keith Howard returned to DiAural in January 2005 (Vol.28 No.1):

When I wrote the "Doppler and DiAural" sidebar for my recent feature about Doppler distortion in loudspeakers ("Red Shift," November 2004, p.67), I fully expected a retort from DiAural founder Ray Kimber, of Kimber Kable. In the event, both he and Eric Alexander, the originator of DiAural's crossover design concept, took up their pens to point out to me that the Doppler decoding explanation with which DiAural was launched had been withdrawn before my article was published.

I wish, of course, that I had known this before writing "Doppler and DiAural," although it would not have dissuaded me from including the sidebar or much altered its content. A thorough explanation of why a loudspeaker cannot cancel out the Doppler distortion generated by a microphone (and vice versa) is just as valid now as it was before DiAural's retraction.

In my correspondence with Ray Kimber, I pointed out that one reason I (and, surely, countless others) believed the Doppler decoding explanation was still current was the continued citation—and, in one case, linking on DiAural's website—of press articles repeating the Doppler claim, without any cautionary note to the effect that it had since been rescinded. Promptly and, I thought, in a very gentlemanly way, Kimber conceded this point and immediately set in train changes to the website that will prevent it from, in the future, promulgating this discredited explanation of DiAural's sonic effect. (Which, as I won't need to remind *Stereophile* readers of long standing, earned the launch of DiAural enthusiastic column-inches in much of the audio media.)

Eric Alexander has assured me that he has never developed a crossover topology that was intended to eliminate Doppler artifacts, which reinforces my observation in "Doppler and DiAural" that DiAural's two patents, which cite Alexander as the inventor, make no mention of Doppler distortion in their texts. Doppler decoding, it seems, was merely an attempt to explain the sonic benefits of DiAural, and so came after the event, not before.—**Keith Howard**