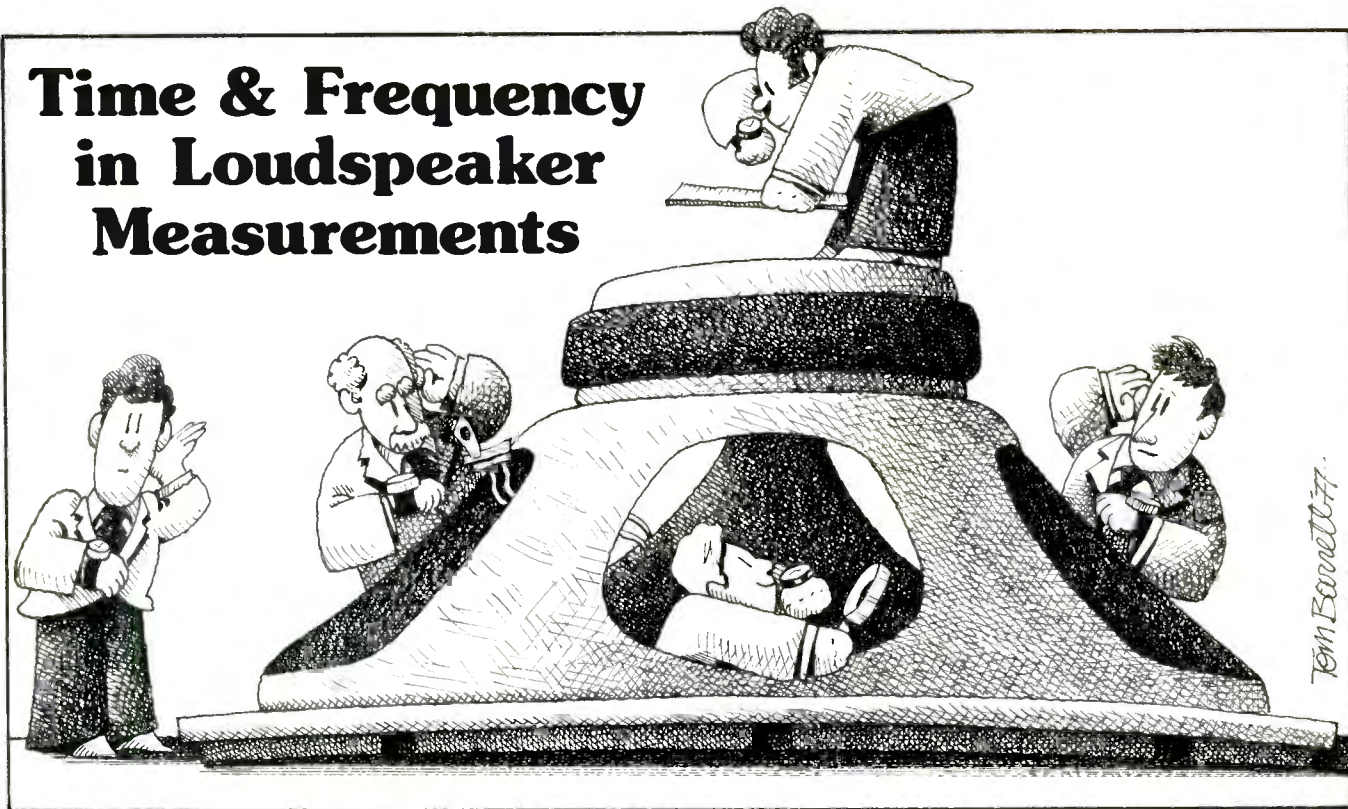


Time & Frequency in Loudspeaker Measurements



80

Richard C. Heyser

I claim that it should be possible to measure audio systems and have those measurements correlate with what we hear out of those systems. We are not doing that now. Our measurements are more precise than ever, but our understanding of what those measurements mean to the way a system "sounds" is still hazy.

I further assert that we are locked into that dilemma because we do not truly understand the meaning of those technical concepts which we now use. I don't think I can be more blunt about the matter.

OK. So having shot my mouth off, what am I going to do about it? Well, what I would like to do is present the readers of *Audio* with my personal view of the meaning of some of the more important terms we use in audio. These are some of the results from my own continuing research into the problem of finding out how to bring subjective and objective audio together. What I present here is my own work. I'm laying it out, in a put up or shut up fashion.

But I am not asking you to accept these things blindly. Question it, think about it, because what we really need to do is dig down to these underlying principles, the philosophy of the

problem. In these discussions we go below the equation's mechanical formalism and question what the meaning is behind the equations. Then when we come up to the equations of audio we find that, while there may be no change in form, we often have a completely new perspective on just what they mean, not only to the pedestrian task of measuring components, but to the possible link with subjective perception.

In a previous article, I started out at ground zero and gave my interpretation of the meaning behind a technical term that is commonplace in audio, the term we call frequency. In this article I would like to carry this point further and apply it to the interpretation of certain loudspeaker measurements.

But before I get technical, let me put one thing into perspective. The end product of this whole multi-billion dollar audio industry is the listening experience. It is what we "hear," in the abstract sense of this word, that is important.

It is not the oscilloscope pattern but the listener's perception that is paramount. This does not mean that we should reject technology... quite the contrary. We know that most persons have the same general impressions of the realism and quality of a performance when listening to identical sound reproduction. There is some-

thing that is used by all of us in making our judgment, and that something is tied to the ingredients making up the reproduced sound. If this something is there but not specifically outlined in our present technical measurements, then we need to get even more technical and find out why. We need a Renaissance out of what may prove to be the "middle ages" of audio. The winner, if there is to be a winner, would be the listener, for we would know how to make his enjoyment of sound far better.

In my last article, I pointed out that when we do become very technical and poke around at the precise meaning of terms, a startling fact emerges. Even as fundamental a term as frequency turns out to have a meaning quite different from that which most of us employ in audio.

It is a subtle thing, but sometimes subtle things topple kingdoms. Let me recap. We know that at present there are two major ways of describing an audio signal. There is a time-domain representation and there is a frequency-domain representation. The time-domain representation and the frequency-domain representation are Fourier transforms of each other.

Now what the heck is a Fourier transform? A conventional textbook answer to that question is to write out a certain hairy integral equation and state... "that is a Fourier transform."

Simply writing down some equation, as though it were a Machine of the Gods, doesn't answer anything. Nature does not solve equations, people solve equations. Nature works in spite of us, and at best the equation is some sort of model for the way in which nature works.

In the previous article, therefore, I suggested a different approach. Suppose we have a signal which we agree is a legitimate time-domain representation. And suppose we ask ourselves what form that signal will take if it is observed by a being who uses some other coordinate instead of time. In particular, what would the form of that signal be if it has the same dimensionality but is somehow measured in units that are the reciprocal of the units of time we use?

Remember, we would both be seeing the same signal, but would be using different frames of reference.

Pursuing the point further, we asked what recipe we could use to take our time-domain view and see it within the framework of this other being's coordinate system. We derived the recipe, which turned out to be the Fourier transform. And the coordinate system which this other being uses turned out to be the parameter we call frequency. Exactly the same equation you will find in a textbook, but with a totally new interpretation.

The thing we call time in audio measurements and the thing we call frequency are different coordinates for describing precisely the same signal.

Subjective Descriptions

Oh, yes... ho hum, technicalia. But if we begin to think what this means to audio it gets a bit exciting, because this means that frequency and time are only two out of an infinite number of coordinate systems we can use to characterize a signal. We don't have to go just from time to frequency, we can go from time to some other coordinate. And even more stunning is that since we can have *either* time or frequency, but never both together in a meaningful description, this means that those properties of sound which we perceive and relate to the

words "time" and "frequency" are *not those parameters at all*.

Now, think for a moment about those words we often use to characterize the sound of imperfect reproduction. Words such as "grainy" and "forward." These words do not seem to fit in with either an exclusive time description or frequency description. Is it possible that these words belong to some other, as yet unrecognized, coordinate system which is a legitimate mathematical alternative to time and frequency? I claim the answer to this question is yes.

Putting it in blunt language, if we measure the frequency response of a system, and do it correctly, then we know everything about the response of that system. We have all the technical information needed to describe how that system will "sound." But the information we have is not in a system of coordinates that will be recognizable by a subjectively oriented listener. Everything is there, but the language is wrong.

That is the root cause of the continuing fight between subjective and objective audio. It is *not* that *either* is more correct than the other... rather it is *due* to the fact they do not speak the *same language*. And when I say language, I do not mean just the descriptive words, but the very frame of reference upon which these words are based.

Sticking my neck out further, I assert that the reason technical people (and I am one of them) did not recognize the root cause of this problem was due to the fact we did *not* realize there *could* be other meaningful frames of reference besides time and frequency.

And, as a matter of fact, not too many technical people are aware that time and frequency are themselves alternate frames of reference, rather



“And, as a matter of fact, not too many technical people are aware that time and frequency are themselves alternate frames of reference...”

than just two terms to be applied haphazardly to measurement.

There! How's that for tipping over icons?

Loudspeaker Tests

As a reader of *Audio*, you've probably noticed that our loudspeaker reviews have been a bit more technical than is normal industry practice. There's a reason for this. These tests are a first attempt to relate measurement to subjective perception. The various tests we perform did not just happen; each is in some way related to simple mathematical results in the type of geometric structure which we might use in perception. It is a first attempt, and very crude at that. But somebody's got to start the process, so let it be here.

In the remainder of this discussion I would like to explain the technical aspects of spectrum sampling and apodization as they relate to the loudspeaker tests we perform in *Audio*.

Let me begin by recapping a very important concept which I flogged to death in the previous article. That is this mysterious and seemingly sinister thing called the uncertainty principle. There is nothing mysterious about the uncertainty principle at all. It is not something nature does to us, but something we do to ourselves through the definitions we give things.

Here is the point. It makes absolutely no difference whether we start out by defining parameters as being related by the Fourier transform, or somehow discover well along the road that two properties happen to be related through Fourier transformation: when two properties are Fourier transforms of each other, they represent different ways of describing the same thing and hence cannot be thrown together into one common description. The Fourier transform is a map, you see, which converts one coordinate system into another coordinate system.

It is a property of changing from one view to another that each part of one view becomes somehow spread over the entirety of the other view. In particular, the Fourier transformation takes a single coordinate location in

“Nature’s clocks always run forward; at least, the most diligent searching has failed to reveal any experimental results to the contrary.”



one view and makes it into a very special geometric figure in the other view, a figure which we call a wave and which extends over the entire range of coordinates in the other view. If we try to take a restricted range of coordinates in both views, we cannot do so and be precisely accurate. But what we can do is ask what the minimum ranges of coordinates are in *both* views such that “most” of the same information is contained in each. The form this takes for a popular measure of “mostness” is such that the product of these two ranges is greater than or equal to some number. This is called the uncertainty principle.

Let’s see what this means in audio terms. Suppose we are testing a loudspeaker. We kick it with a voltage and the loudspeaker produces some sort of sound. Let’s pick that sound up with a microphone and convert it back to voltage. Now let’s put a switch in the output of the microphone. Suppose the switch is initially open, so that we do not have any sound signal to analyze. Some time after the loudspeaker puts out a pressure wave, we close the switch for one second and then open the switch.

What do we have? In the coordinate of time we have a signal that only has a sound-related value over a period of one second. We have created a one-second chunk of time... a time-domain representation.

Imagine, if you will, how that voltage would appear to some being who does not live in a coordinate called time, but whose frame of reference is something we call frequency.

In fact, if we want to see what he sees, we can convert to his coordinate system by making what we call a spectrum analysis. In order to do this, we have to give up the thing we call time. Time will show in this frequency spectrum, but it will be in the form of the relationship of phase and amplitude of waves in the frequency spectrum.

When we look at the frequency representation, we will see that there is some energy spread over the whole of the frequency coordinate. But the effect of having taken a frequency spectrum from a small chunk of time is that the frequency spectrum will be

very slightly out of focus. The edges will not be sharp, but somehow smeared. The amount of this smear will be on the order of one Hertz, which is the name we give to the unit of measurement in this other being’s coordinate system.

If we had only closed the switch for one-thousandth of a second, and then seen what our frequency-domain friend saw, we would find that the smear was of the order of one thousand Hertz units (I’m only talking in ballpark figures).

That is the manifestation of what is called the uncertainty principle. In performing *Audio’s* loudspeaker tests, I use a 13-millisecond time window to make the three-meter or room test. I want to find out what spectral components are found in that important time period which can establish some measure of timbre or tonal balance of the sound heard from that loudspeaker when placed in a room. This time duration derives from psychoacoustic tests. I cannot legitimately present any frequency measurements focused to an accuracy of better than about 100 Hz, including the range from d.c. to 100 Hz, because of the chunk of time which the data represents. To be safe, therefore, I only give data from about 200 Hz upward.

Apodization

Now there’s this problem called apodization, which literally means “the process of removing feet.”

When we hack off sharp edges, such as closing and opening a switch on a voltage, the equivalent transformed view will be blurred in a most unpleasant manner. There will be foot-like appendages, or sidelobes, which extend outward from each place where there should be a solitary frequency value standing apart from its neighbors.

Again, I must stress this is not due to some caprice of nature, it is due to our definition. If we hack off edges, and if we take a Fourier transform view, then we will find sidelobes. And I don’t give a darn whether we measure the equivalent frequency response with sharp filters or with a computer FFT, our definition requires they be there. The theory determines what we will observe.

In order to minimize (we can never remove) them, it is necessary to do some sort of blurring or defocussing in the hacked-off parameter. The process of removing spectral feet by operating on the original data is called apodization. There are an infinity of apodization processes available, depending upon the type of corresponding blurring we are willing to tolerate in the apodized spectrum. Apodization usually consists of smoothing the sharp edges by using more of what is in the middle of the hacked-off distribution than at the sharp edges. *Audio’s* loudspeaker data is apodized with a nearly raised-cosine weight function when frequency response is plotted, and with a Hamming weight function when time-domain response is plotted.

Time Measurements

Nature’s clocks always run forward; at least, the most diligent searching has failed to reveal any experimental results to the contrary. Where we poor humans get into trouble is when we start out from a frequency measurement and compute the corresponding time-domain response. If we have a *chunk* of frequency response, for example if we have no data above 20 kHz, then the time-domain response will be blurred.

In nature, the sharpest edge of all is at “now.” A computed time-domain response will therefore spread before and after “now.” The computed time-domain response will appear to predict the future... that is not really a prediction, but a blurred edge.

The energy-time loudspeaker measurement we make is a computation from the anechoic frequency response. We band limit from zero frequency to 20 kHz. In order to get the sharpest definition of discrete signal arrivals, such as due to diffraction

from the edge of the enclosure, with the least amount of predictive "feet," we use an apodization function called Hamming weighting. Our measured sidelobes are actually down close to 40 dB below the peak giving rise to them. But you will still see what appears to be a predictive risetime prior to extremely sharp pulses.

As a matter of professionalism, we also check the loudspeaker impulse response by using a raised cosine pulse of voltage that has a 10 micro-second half-width. The loudspeaker impulse response is viewed on an oscilloscope and compared against the computed energy-time response to make sure all is kosher.

The reason for this belts-and-suspenders approach is due to a fact of apodization that, unfortunately, very few professional people seem to be aware of. Apodization, or a weight kernel, or whatever you choose to call it, has all the properties of the data to which it is applied. This includes the properties of amplitude and phase. In fact, we could take a converse view that the data is actually a weight kernel on the apodizing function.

Now, you know what happens when we take a Fourier transform of a product of two functions in frequency. The result is a time-domain convolution of what would have been the time-domain representation of each by themselves. They get all mixed up.

They get tangled up in phase as well as amplitude. And quite often a messy data signal will "unsmooth" even a good apodizing function. In short, this means that sometimes the computed response is lumpier than we think it should be. But, and computer people take note, unless you have a cross check or precise knowledge of the amplitude and phase of the data being transformed, you don't know it happened.

The geometry of this is too lengthy to go into here, but most apodizing functions used in Fourier transform analysis are non-minimum phase. Mostly they change the amplitude without changing phase. This includes Hamming, Hanning, and the rest. Historically, this is because the interest usually lays in the power spectrum (phase, what's that?). That works swell when the data is minimum phase. But when the data (in our case loudspeaker frequency response) has a maverick phase term, it can unsmooth a good apodizing function. Look at it this way, the effect is as though the loudspeaker response was minimum phase and the excess phase term was thrown into the weight kernel.


I realize that such talk might be highly confusing if you're not in the FFT business, but computer people ought to know what I mean. Other than my own comments in technical journals, I don't believe this fact has been pointed out before.

What it boils down to is that *Audio* makes every effort to be technically accurate, even if we are not terribly popular among some manufacturers when we do so.

Wrap Up

Let me wrap up this little discussion with two observations. First, if we real-

ly want to bring subjective and objective audio together, we need to get down to the fundamentals which can be highly technical. Second, with the editor's permission, I am trying an experiment with these discussions—in using words rather than mathematical symbolism, but I am not watering down the technical level.

Audio's readership covers the full range of involvement in the sound industry, from listener to researcher. Reader survey cards (yes, we do read them) indicate that many of you want more technical articles. And you like straight talk. All right, this was a trial balloon. Want more? 

If you missed the first two issues, it's still not too late.

We've reprinted Volume 1, Number 1 of *The Audio Critic* and have printed enough of Number 2 to keep up with the ever-increasing demand. Our phenomenal growth seems to indicate that noncommercial, consumer-oriented audio journalism has come of age.

One subscriber writes that *The Audio Critic* is as good as the four best-known "underground" audiophile reviews combined. Another calls it the new Bible of high fidelity. Even if you ascribe such reactions to fanaticism, the fact remains that *The Audio Critic* is unlike any other publication or service.

Six times a year, by subscription only, *The Audio Critic* tells you in print what has so far been only spoken of in private among a few enlightened professionals. What's the best. What isn't. What's smart. What's dumb. And why.

Since *The Audio Critic* accepts no commercial advertising, not even by retail stores, and since it has no connections with the audio industry, it has nothing to gain from being kind to unscrupulous or incompetent manufacturers. It names names, calls bluffs, praises and criticizes without restraint.

What's more, *The Audio Critic* doesn't just blithely inform you that the highs are "grainy" or the midrange is "liquid." It makes a serious effort to correlate its listening tests with measurements taken in its own, superbly equipped laboratory.

One year's subscription to *The Audio Critic* (six issues) costs \$28, first-class mail only. For overseas

airmail, add \$5. No single copies are sold for any reason whatsoever, but the unused portion of canceled subscriptions is refundable on request.

The first two issues cover close to thirty preamplifiers, as well as power amps, speakers, subwoofers, tone arms and other subjects. The third issue (May/June) is devoted mainly to speaker systems with state-of-the-art pretensions. We strongly suggest that you begin your subscription with Volume 1, Number 1, in order to own a complete set and be thoroughly familiar with our approach.

Send your \$28 for the first six issues today to *The Audio Critic*, Box 392, Bronxville, New York 10708.

