ELECTRONOTES
1016 Hanshaw Road
Ithaca, NY 14850

APPLICATION NOTE NO. 428

July 10, 2016

WHY NO RANGE-SWITCH WITH HEARING?

INTRODUCTION:

Many of our studies here involve technical aspects of hearing, as necessary in some application such as audio, speech, and music. It is common to take the "receiver" portion of our considerations (i.e., the ear) as a given. Indeed we will often be told that the details of how the hearing system works is "beyond the scope", or is relegated to many excellent references [1-3]. Often we are asked to approximate something like the way hearing apparently works (like "assume the ear is a Fourier analyzer").

Among the many elements of hearing that tend to be ignored are those that relate to ways in which the hearing mechanism is DIFFERENT from its nearest engineering counterpart. Our perception of loudness is unlike an (pressure) amplitude measurement for a physical vibration. Our perception of pitch is unlike a simple spectrum analyzer. (Often we see laments about how evolution has failed to give us exactly the corresponding mathematical device.) In one important comparison, the ear has a wide range, and reports this wide-ranging parameter without need of a "Range-Switch".

BACKGROUND:

Actually, engineers are very comfortable with the so-called "dual-descriptions" of phenomenon in the "dual-domains" of time and frequency. That is, mathematically we describe something like sound as a mathematical function of time, s(t); or as an alternative, a mathematical function of frequency, S(f). We envision the ear as using either or both – somehow. Engineers jump back and forth between domains as a means of making a problem easier to solve, for insight, or just for entertainment. By in large, they do not even bother to warn their reader/listeners what they are doing.

Perhaps 30 years ago, I gave a colloquium at Cornell on music synthesis. I began by saying that musicians used frequency-domain descriptions – they just wrote their spectrograms in a strange way – simultaneously putting up an overhead slide of a page from an orchestral score. It got the predictable good laugh, and everyone got the point. I went on to insist that this meant that music (indeed, sound in general) was parameterized. We could describe a sound denoted s as a function of time, s(t). We will get further, faster, by saying s(t) is described by parameters that vary (relatively speaking) slowly in time (perhaps once during a “note” of say 1 second duration). Such parameters would involve the loudness, duration, and frequency (more generally, a fuller description in terms of a “spectrum”), and so on. That is, the sound is characterized into an efficient set of quantitative attributes.

Often, alternatively, we get things basically right by saying that the sound has a spectral description S = S(f) [where the upper case is the corresponding frequency-domain variable]. The two descriptions are often said to be related by a “Fourier transform”. Many (likely most) sounds fit better with a spectral description (recall the orchestral score).

In a real sense, s(t) and S(f) are complete, equivalent descriptions. Yet, the signal s(t) is still actually a physical pressure variation. That is, a vibration of air molecules creating regions of over- and under-pressure; bumping about (as longitudinal waves). We could compare the sound wave to changes or atmospheric pressure as on the weather report! However, the variations in barometric values, while tiny (a few percent), are very huge compared to the variations that constitute sound, which may be as small as two ten-billionth of an atmosphere! [Keep in mind that the pressures are extremely small, but the area of the eardrum is very large on this scale, and force is pressure times area.] Barometric variations are also many orders of magnitude slower in time. Further, the total audible dynamic range for pressure variations for the normal ear may be a million to one (threshold of hearing to threshold of pain!). Thus the sound pressures are from about 2×10^{-10} to 2×10^{-4} atmospheres. If you get the impression that the ear is an extremely sensitive instrument of nearly unimaginable capabilities, something an engineer could only dream of, you are right.

The “take-away” at this point is that sound is a vibration that propagates through the air (molecule bumping molecule) but only creates tiny, quite local variations from normal pressure. The ear (human or otherwise) is remarkably well evolved to accommodate pressures from the lowest levels, to the highest that is useful to a given critter (the 1000000:1 ratio). Later we note that pitch perception has a 1000:1 ratio.

RANGE-SWITCHING - LOUDNESS

Builders of instruments for physical measurements most often need to include a “range-switch” either as a manually operated rotary switch, or with pushbuttons; or as an automatic ranging feature. A user of a $15 multimeter is familiar with this. There are several reasons for this. First of all, typically signals are confined in the meter electronics to be safely between random noise on the small side and the available power supply “rails” on the high side. You don’t want to clip, or get lost in the noise.

A second reason is more subtle. If you have a voltmeter (for example) reading millivolts, and yet want it to read a power main of perhaps 120 volts, that would be 120,000 millivolts. Well, it would not likely be exactly 120.000 volts on the lines, perhaps 119.634 volts. The meter wouldn't have the precision to reliably report this as 119,634 millivolts, and there are twice as many digits as you likely need or want anyway. Thus we are delighted to provide a device that cheaply and reliably says 120. Either we switched to a 200 volt range, or the meter just did it and reported only 3 digits so as not to tell outrageous lies. Set to other ranges, it might tell us that a battery is 1.42 volts, for example. Plenty good enough. Familiar.

Still it goes further. If we get a sound-level meter, or a frequency-counter intended for audio, we may find it strange that the sound level is measured in decibels. A frequency meter may have a translation to musical notes (like the famous A4=440 Hz). Why logarithms for sound levels, you demand, and how did evolution come to ingrain decibels, and octaves? It didn't. Those are man-made and artificial.

What they are are ratios.

Much as we saw that a voltage measurement made sense against a background of ranges, a critter's hearing seems to be interested in how much a parameter of the hearing differs from a current value or expectations. While we detect signal with pressure variations down to tiny tiny fractions of an atmosphere, if the same pressure variations are added to a larger signal, it is ignored. Why bother! It is like floating point arithmetic. It is worth an example here.

Suppose we have a floating-point format of a three digit mantissa and a one digit exponent. The number 123 would be:

$$a = 0.123 \times 10^3$$

while the number 52 would be:

$$b = 0.520 \times 10^2$$

and the number 0.00521 would be:

$$c = 0.521 \times 10^{-2}$$

We can add a to b to give

$$a + b = 0.175 \times 10^3 = 175$$

but using the exponent 3 when adding a + c gives:

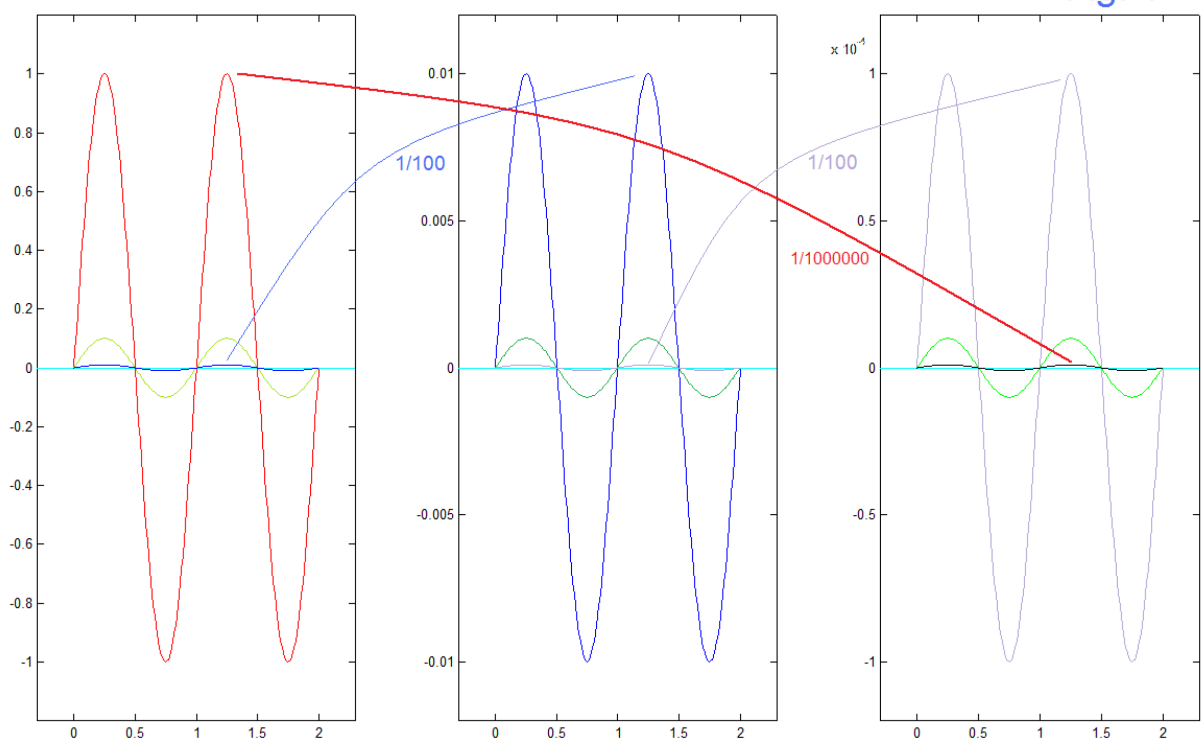
$$a + c = 0.123 \times 10^3 + 0.00000521 \times 10^3 = 0.123 \times 10^3 = 123$$

Because the mantissa was only three decimal places in this example, the small number is lost. The number c is perfectly fine by itself, but can't be added to a much larger number. Thus it is the nature of the floating-point arithmetic to show disinterest when the ratio due to a small alteration is close to 1.

Roughly we think that when it comes to loudness an increase of say 5% to15% (0.4db to 1.2db, or ballpark, 1db) is necessary to be significant. For more discussion of db and indeed of hearing in general, see [4]. [With pitch, something like 3% (half a semitone) is significant.] Indeed, if what is often called a "Just Noticeable Difference" (JND) is around 1db, we understand a full range of 120db as being 120 possible "bins" for loudness which should leave us quite functional in that regard. The logarithmic spacing thus divides a wide range (1,000,000:1 which we need) into a number of divisions (120) that is all we need. Evolution did it.

It may be hard to envision a 1,000,000:1 ratio except by imagining a choice between $1,000,000 and $1. But what does this look like on a graph? Fig. 1 shows that this is hard to graph. We start (left panel) with signal of amplitude 1 (red), two cycles of a sine wave just as an example, and divide it by 10 (yellow-green) and then by 100 (blue). Note that if we divided by 1000 (or more) we would only be overplotting the light blue baseline. So a ratio of 100:1 is about all we can show on an ordinary plot. In order to show more attenuation, we need additional plots, and the middle panel of Fig. 1 shows

Fig. 1



the plots from 0.01 (blue) to 0.001 (green) and to 0.0001 (purple). Again, we reach baseline. Thus it is left to the right panel to show the plots for 0.0001 (purple) to 0.00001 (light green) and finally to 0.000001 (black) to get down by a million. This red to black is the loudness range of the ear. Impressive.

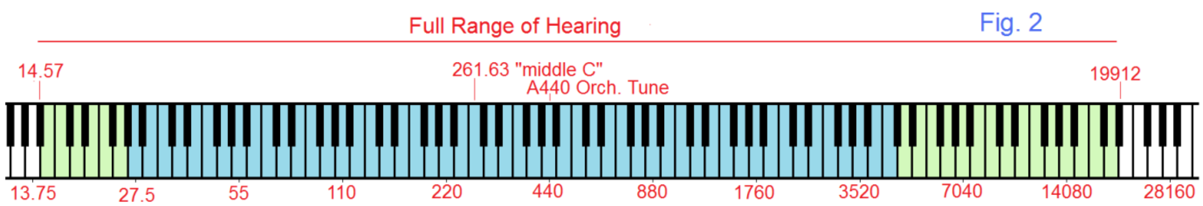
Further, recall that on this scale, we have already started at the 1 = threshold of pain! This was already about 2×10^{-4} atmospheres. Thus we see how tiny the pressure levels are, which we can still hear. A rather astounding number often quoted is that the ear responds to a displacement of the eardrum equivalent to 0.02 the diameter of a hydrogen atom. { Sometimes it is 0.03, and Hartmann allows a full hydrogen atom [2, pg. 61]. } Three questions perhaps come to mind: (1) Is this true? (2) How could it be true? and (3) Who measured this? The answers are (1) seems to be ballpark true, (2) its because of the relatively immense area of the eardrum and (3) it seems to be a calculation, not a measurement.

First, the issue with area. If the news said that an inch of rain fell on a football field in Texas, you might envision a challenging mop-up, but still only 28,050 gallons or about 1/24 of an Olympic swimming pool. If on the other hand, you heard that an inch or rain fell on <u>all</u> of Texas, that would be about 4.46×10^{12} gallons. So area matters a lot. The eardrum has an area of roughly 0.5 cm^2 as compared to the supposed tiny displacement of 0.5×10^{-8} cm.

I have not been able to find a discussion of the question of measurement vs. calculation. Yet it would be quite absurd to suppose we could even measure the position of the eardrum to that precision (roughness, for example), let alone its motion. We do however know the acoustic power needed, and the area of the eardrum. This permits the fair estimation of displacement. That must be what they did. It is indeed tiny.

PITCH IN “BINS” TOO

We saw above that when it came to loudness that the ear had a very large dynamic range (1,000,000:1) and yet our resolution was such that something like100 levels, logarithmically spaced (in decibels), was really enough. A fortuitous display of the analogous situation with regard to pitch is available and familiar as an ordinary piano keyboard. Fig. 2 shows a somewhat extended keyboard. The famous 88-keys are shown in blue. This is just over 7 octaves. The actual range of pitch perception extends beyond the range of the keyboard to just over 10 octaves to include the keys in the green regions on either end. This is 125 “keys” (10.4 octaves) in our range of pitch, a number we compare to 120db. Apparently whomever designed the piano decided that about a hundred notes was enough. More on this below.



Note first here that we again deal with ratios. As soon as we mention "octaves" we are dealing with 2:1 ratios of frequencies. This again puts pitch on a perceptual logarithmic scale. Note well that we say we can normally hear from about 15 Hz to 20,000 Hz, a bit more than 1000:1. (This is less than the range of loudness, but still impressive.) This puts the arithmetic middle of the range of pitch (at about 10,000 Hz) which is about a $D^{\#}$ just over an octave above the highest note on the piano. NOT even on the keyboard! What happened to the famous "middle C"? Well, it's at 261.63 Hz. And the famous "A 440" to which the oboe sounds to tune the orchestra is of course at 440 Hz. Ratios of 2 every 12 notes matter.

So, we understand ratios. In addition to a 2:1 ratio (octave) we know there are other ratios between the octaves that are of interest, such as a "fifth" (3:2) and four or five others. So, why are there 12 tones in an octave in the standard scale? Well, it's complicated but well studied [5]. The selection of 12 tones per octave (as opposed to, say 11 or 13, etc.) is a balance between enough "density" within an octave and (more or less by chance) getting good approximations to low integer ratios. The compromise is a 12-tone "equal-tempered" scale. In this scale, any piano key has a pitch the same as that below it (white or black) multiplied by the "twelfth-root-of-2". This is:

$$2^{(1/12)} = 1.0594631$$

(about 6%) so if middle C is 261.6256 the $C^{\#}$ just above it is 277.1826. The full range can be generated, for example in Matlab, as:

```
r  =  -5:(1/12):+6

f  =  440*2.^(r)
```

which generates from 5 octaves below A=440 or 13.75 Hz to 6 octaves above A=440 or 28,160 Hz (above audible). Note that if you have some frequency in mind, the standard scale will give you a choice within about 3%. The value of 3% is roughly the just noticeable difference for pitch.

LOUDNESS CHANGES WITH PITCH

Having looked at loudness and at pitch, we can now combine the two. It would be quite remarkable if, in stating that we hear from 15 Hz to 20,000 Hz, it were true that we hear all the frequencies equally well, and that outside this region, we hear absolutely nothing. Indeed we expect some notion of a band-pass frequency response where the response tapers at the low and high frequency ends. Further we expect that, regardless of the frequency response, that the actual "detection" or an audible signal depends on the loudness (pressure amplitude basically) of the input sound.

Recognizing this, way back in 1933 Fletcher and Munson [7] ran experiments and obtained a set of curves which pretty much tell us what we need to know today. They, and newer versions, give a set of U-shaped curves (Fig. 3 typical) rolling up below about 100 Hz and above about 6000 Hz, with a modest bump down between 3000 Hz to 4000 Hz (due to a resonance of the ear canal). These are unusual (for an engineer) due to their orientation (U-shape) and the multiplicity of response curves. We live with them, but perhaps forget that we were first confused by them.

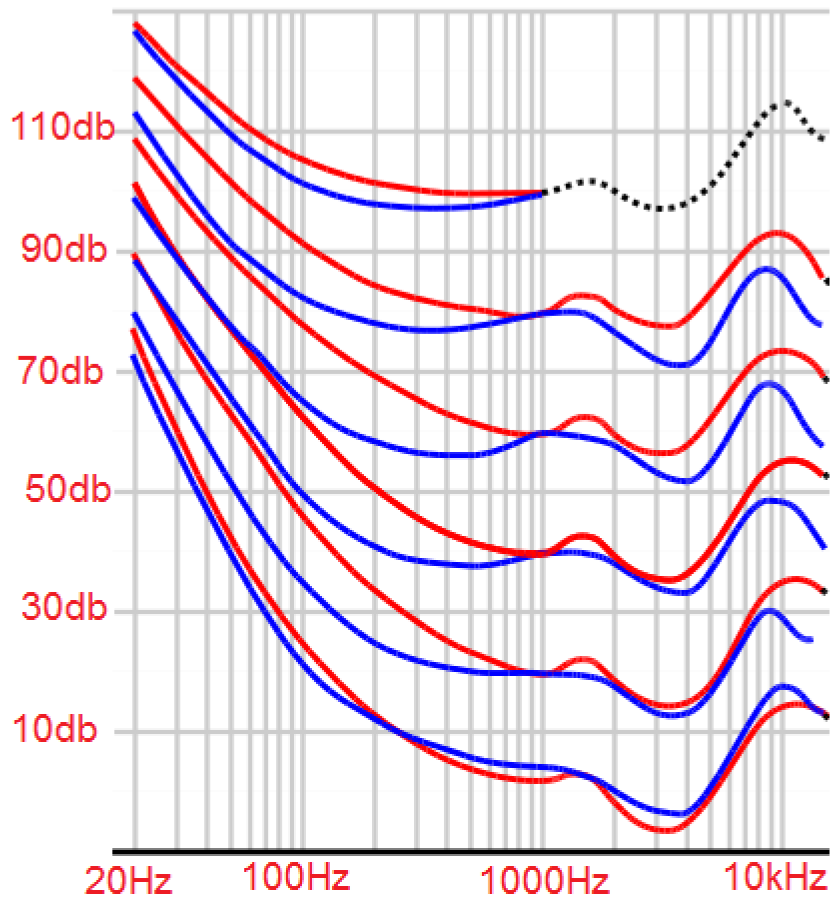
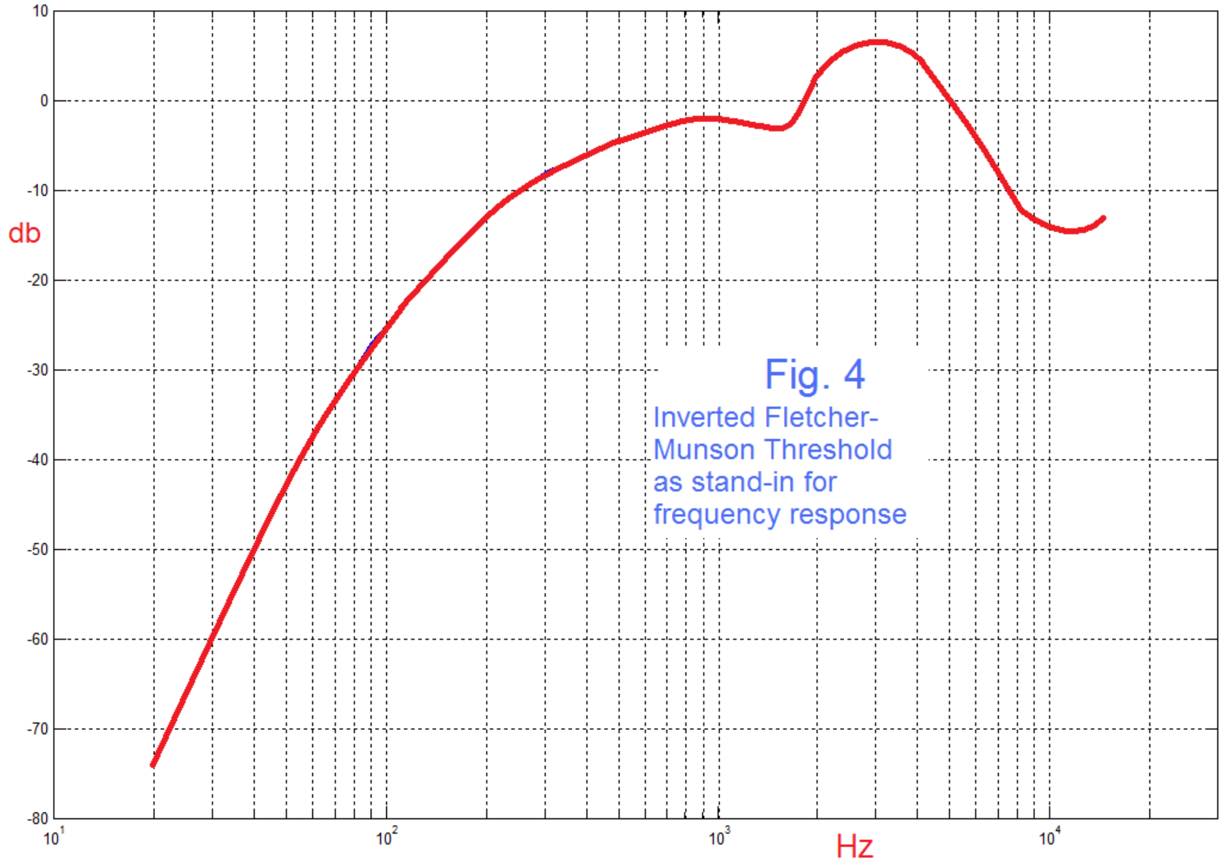


Fig. 3

Equal Loudness Curves [6]



Consider first what Fletcher and Munson (F-M) were trying to measure: Loudness (subjective) as a function of input intensity and frequency (objective). Measuring objective parameters was not difficult with equipment available in 1933. Then, as today, however, subjective loudness is a challenge. It would be hard to take a sound at a given level of intensity at a given level of frequency, and say, for example, that at a different frequency the result was subjectively "half as loud". Instead, F-M allowed the subjects of the experiment to move to a different frequency and adjust the sound pressure level so that subjectively the sound was of equal loudness to the first frequency. Setting a set of intensity levels at midrange, it was possible to find and plot how much more pressure was necessary to reach subjective equality. Hence the curves bend upward at the ends of the range.

Contrast this to the way an engineer measures the frequency response of an electrical filter. We set a convenient input amplitude, sweep frequency, and objectively (like with an AC voltmeter) the output. This takes out the subjective element. Further, our filter is probably near nominal when it comes to linearity, so we only need one curve. If we have a bandpass-like response (it roll down at the extremes) we have smaller numbers on the end, and an inverted U. With subjective equal-loudness, we had larger numbers (the amount of upward adjustment) which was larger on the ends. The response of the ear is supposed to roll-off (down) on the ends – what gives!

So it is not really correct to say that F-M and more modern versions are measuring the frequency response, but it is still basically the general idea, but applied to a more unwieldy item (subjective loudness). Oh – and we may well flip the curve over. Fig. 4 shows a single level (threshold of hearing) for one such set of curves, flipped over, so that it more closely resembles a band-pass. This I find very helpful, in that the response rolls-off on the ends. (I don't know why these curves tip up ever so slightly just before disappearing on the high side – they are quite hard to hear at all up there regardless of loudness.) Note that Fig. 4 resembles the "A-Weight" curve of a loudness meter [4].

REFERENCES

[1] Roederer, J.G., ***The Physics and Psychophysics of Music***, 3ed, Springer (1995)

[2] Hartmann, W.M., ***Signals, Sound, and Sensation***, AIP Press (1997)

[3] Heller, E.J. ***Why You Hear What You Hear***, Princeton U. Press (2013)

[4] B. Hutchins, "Considerations for a Noise Ordinance," Electronotes Application Note AN-357, November 2003 http://electronotes.netfirms.com/AN357.pdf

[5] B. Hutchins, "Revisiting: Musical Scale Mathematics", ***Electronotes***, Vol. 22, No. 213, January 2013 http://electronotes.netfirms.com/EN213.pdf

[6] Wickipedia https://commons.wikimedia.org/w/index.php?curid=16477782

[7] Fletcher, H. and Munson, W.A., "Loudness, its definition, measurement and calculation", ***Journal of the Acoustic Society of America,*** Vol. 5, pp 82-108 (1933)