The **just noticeable difference (jnd)** is the smallest amount of some perceptual phenomenon that we can distinguish. A rough approximation for many (but not all) sensory phenomena is Weber's ([veɪbɚz]) Law, $\Delta S \propto S$, where ΔS is the jnd and S is the original stimulus level. What this means is that as the intensity of the stimulus increases, so does our jnd for that stimulus. In other words, for higher stimulus levels, it becomes harder for us to tell them apart because our jnd is also higher.

One acoustic consequence of Weber's Law is that we can distinguish lower frequency sounds from each other better than we can distinguish higher frequency sounds. For example, it is easier for us to tell the difference between tones at 100 Hz and 105 Hz than tones at 3,000 Hz and 3,005 Hz. The difference of 5 Hz is more noticeable in the lower frequency range than in the upper frequency range.

In music, this explains why we have musical scales set up the way we do. Octaves do not range over the same absolute frequency difference. Instead, octaves range over the same proportional frequency difference. For the Western *do-re-mi* scale, the frequencies in each octave are twice those of the next lower octave. The modern American value for middle A is 440 Hz (which is why it is sometimes referred to as 'A440'), so the A in the next lower octave is at 220 Hz, while the A in the next higher octave is at 880 Hz, not 660 Hz.

Using Weber's Law, we can create a linear scale that corresponds to our perception, rather than to the absolute numerical value of a stimulus. Suppose we want to define such a linear perceptual scale P for musical notes, where middle A has the value 1, and each higher octave is one point higher on this scale. Thus, the next highest A is 2 on our perceptual scale, the next is 3, and so on, while lower As would be 0, −1, −2, etc. We would have:

$$\begin{aligned} P(440\text{ Hz}) &= P(2^0 \cdot 440\text{ Hz}) = 1 \\ P(880\text{ Hz}) &= P(2^1 \cdot 440\text{ Hz}) = 2 \\ P(1760\text{ Hz}) &= P(2^2 \cdot 440\text{ Hz}) = 3 \\ &\vdots \\ P(f) &= P(2^x \cdot 440\text{ Hz}) = x+1 \end{aligned}$$

That is, the value on our perceptual scale for some frequency f (in Hz) would be given by the formula $P(f) = x+1$, where $f = 2^x \cdot 440$. To get $P(f)$ in terms of f, we just need to convert x into a function of f:

$$\begin{aligned} f &= 2^x \cdot 440 \\ f/440 &= 2^x \\ \log_2(f/440) &= x \end{aligned}$$

So our final form of the perceptual function is $P(f) = \log_2(f/440) + 1$. All stimuli subject to Weber's Law can be transformed into a linear scale in this fashion, with the general formula being as follows (note that $\log_k$ (base k) can be converted to common log (base 10) by simple multiplication by a constant):

$$P(S) = a \cdot \log(S+b) + c$$

where a, b, and c are constants that depend on how we wish to define our scale. Thus, for any stimulus described by Weber's Law, our perception of that stimulus will be **logarithmic**. This is reflected in the numerous logarithmic perceptual scales that have been defined for other phenomena besides musical pitch, such as the Richter scale used by seismologists to measure the perceived amplitude of earthquakes and the stellar magnitude scale used by astronomers to measure the perceived brightness of stars.

There are two important acoustic properties that have been given similar scales: frequency and intensity. Besides the musical octave scale, other linear perceptual scales for frequency have also been created. Three of the most widely used scales are the **mel**, **equivalent rectangular bandwidth (ERB)**, and **Bark** scales. They each have different uses and depending on the context, one of the scales may be more appropriate than the others.

The mel scale is defined precisely, with 1000 mel equal to 1000 Hz, and 0 mel equal to 0 Hz. The following formula approximately yields this scale:

$$\text{perceived pitch in mels} \approx 1127 \cdot \ln\left(1 + \frac{f_{\text{Hz}}}{700}\right)$$

The ERB scale is also defined precisely, based upon equal spacing along the basilar membrane in the inner ear, given approximately by the following formula:

$$\text{perceived pitch in ERB} \approx 21.3 \cdot \log\left(1 + \frac{f_{\text{Hz}}}{228.7}\right)$$

The Bark scale is defined empirically on experiments measuring jnd. Interestingly, the results were not purely logarithmic; instead, it was discovered that below about 500 Hz, perception is roughly linear, and logarithmic elsewhere. Various formulas have been proposed for the results of these experiments. Two commonly used formulas are:

$$\begin{aligned}\text{perceived pitch in Bk} &\approx 13 \cdot \arctan(0.00076 \cdot f_{\text{Hz}}) + 3.5 \arctan\left(\frac{f_{\text{Hz}}}{7500}\right)^2 \\ &\approx 7 \cdot \operatorname{arsinh}\left(\frac{f_{\text{Hz}}}{650}\right)\end{aligned}$$

For our perception of loudness, which is dependent on the amplitude of air pressure changes in a sound wave, we use the familiar **decibel** scale, defined such that 0 dB is 20 micropascals (the limit of sensitivity to air pressure in the human ear) and 20 dB is 200 micropascals, yielding the formula:

$$L = 20 \cdot \log(P/20) \quad (P \text{ measured in } \mu\text{Pa})$$

Unfortunately, it turns out that there is more to our auditory perception than just simple logarithmic transformation of physical properties of sound waves (which is why the Bark scale, defined experimentally, is not logarithmic). In addition to the natural logarithmic nature of our perception, the physical structure of our ears causes particular biases in our perception or certain frequencies.

Because the ear is shaped somewhat like a half-closed tube, we can approximate its natural resonance frequencies with the formula $f_n = (2n-1)s/4L$. The adult ear canal is around 2.5 cm long, which means its first resonance frequency is about 3,500 Hz and its second is about 10,500 Hz (the third is at 17,500 Hz, which is very near our limit of about 20,000 Hz, and consequently, not very audible). Thus, frequencies near 3,500 and 10,500 Hz will sound louder than those farther away from these landmarks, with more of an effect 3,500 Hz.

Since these frequencies sound louder, we can tell them apart more easily, so languages will often have more sounds that make use of frequencies near 3,500 Hz than near 7,000 Hz.