1 Signal-to-Noise Ratio

The ear is a logarithmic measurement instrument. We can hear a difference in loudness, for example, if the sound pressure has doubled, regardless of the volume. For this reason and since computations with amplifiers and noise become simple additions and subtractions, the loudness of acoustic signals and the amplitude of their electrical representations is also expressed in a logarithmic measure.

The decibel is the most common unit of measurement. It is not an absolute measurement, rather it measures the ratio between the power or amplitude of some reference signal and the signal of interest.

For a ratio between two power value p_1 and p_2, the ratio in decibels is defined as

$$r = 10 \log_1 0 p_1/p_2$$

For example, a power ratio of 2 is equivalent to 3 dB, a ratio of 50 to about 17 dB. (The name of the unit is derived from Alexander Graham Bell, scaled by a factor of 10.) The human hearing spans 120 dB between the softest sound and the threshold of pain. The range between the softest and loudest part of speech or music is called *dynamic range*.

Since decibel are so convenient, they are also used to measure absolute power values, even though the dB is a ratio measurement. To measure an absolute power value, telecommunications engineers simply assume a reference value of 1 mW and abbreviate this measure as 1 dBm. For example, 20 dBm would be 100 mW. (The analog telephone system produces about 1 mW of power at the receiver.)

We designate the input and output sequence of a coding system by $x(n)$ and $y(n)$, for $n \in [0, \ldots]$. Thus, $x(0)$ is the first input sample being compressed by the codec. The *reconstruction error* $r(n)$ is defined as the difference between the input and output signal:

$$r(n) = x(n) - y(n)$$

For simple codecs, the reconstruction error is often also referred to as *quantization noise*, as it sounds like background noise.

We denote the variance of these signals by σ_x^2, σ_y^2 and σ_r^2. The variance is equivalent to the power contained in the signal. For any sequence u with length M samples and zero mean, this variance or power can be measured as the mean square:

$$\sigma_u^2 = \frac{1}{M} \sum_{n=1}^{M} u^2(n)$$

(For most audio signals, we can safely assume that the signal averages to zero.)

One simple measure of the quality of the codec is to measure the ratio of the power of the signal and that of the noise introduced by the coding.

The signal-to-noise ratio (SNR), measured in decibel, is defined as

$$\text{SNR (dB)} = 10 \log_1 0(\sigma_x^2/\sigma_r^2),$$

where σ_x^2 is the power (variance) in the original input signal and σ_r^2 is the power of the reconstruction error.

If we are given the statistics of a signal rather than its sample values, the power for that signal is computed as

$$\sigma_x^2 = \int_{-V}^{+V} (x - \bar{x})^2 p(x) dx$$

where the signal ranges in amplitude from $-V$ to $+V$ and has $p(x)$ as its probability density function (pdf). For a zero-mean signal ($\bar{x} = 0$) with uniform distribution, the pdf $p(x)$ is a constant with the value $1/2V$.

To compute the SNR for a linear quantizer with B bits, it suffices to compute the power of the signal and that for the quantization step, $V/2^B$.

Using this equation for any B bit linear quantizer, one can compute the power of the signal as $\sigma_x^2 = ...$ and the power of the quantization noise as $\sigma_r^2 = ...$, yielding an SNR of .... For example, a linear 8-bit quantizer has an SNR of XX dB, while audio CD with their 16-bit quantizer have an SNR of XX dB.

This computation doesn't quite hold for non-linear quantization. The 8-bits-per-sample logarithmic quantizer used in the telephone system, called μ-law for the American and A-law for the European version, has an SNR of 38 dB. The lower SNR is compensated for by a larger dynamic range.