Dynamic range of signal: difference in dB between highest and lowest levels of signal: 
$$20 \log_{10} \left( \frac{|x|_{\text{max}}}{|x|_{\text{min}}} \right)$$

(If we allow $|x|_{\text{min}}$ to be 0, i.e., silence, then the dynamic range is $\infty$ dB, which doesn't make much sense. In practice, we consider $|x|_{\text{min}}$ as the lowest level of the signal that still convey useful information.)

For audio with 16 bits quantization levels, maximally possible range: 96 dB

For most audio files: $40 \text{ dB} - 120 \text{ dB}$

Music notation: $p$ (piano), $f$ (forte), $mp$ (mezzo piano)

$$\cdots \quad pp < p < mp < mf < f < ff \cdots$$

Tchaikovsky's 1812 Overture: PPPPPP ~ ffff

**Types of Dynamic Range Controls**

- Compression: reduce dynamic range
- Expansion: increase
- Compression then Expansion: Compander

Why?

- Environment/medium may have small dynamic range, e.g., car, radio broadcast
- Suppress noise floor
- Storage medium may have small range, e.g., tape
Let's focus on compression here between CT and LT:

Compression ratio: $R = \frac{\Delta P_i}{\Delta P_o} \Leftarrow \text{Input level change}$

\[ R = \frac{X_{\text{db}} - CT}{Y_{\text{db}} - CT} \]

\[ Y_{\text{db}} = CT + R^{-1} (X_{\text{db}} - CT), \quad \text{slope: } R^{-1} \]

- $R = \infty$: Limiter
- $R > 1$: Compressor (compression ratio)
- $0 < R < 1$: Expander (expander ratio)
- $R = 0$: Noise gate

Go to linear amplitude: $R = \frac{\log_{10} \left( \frac{X}{CT} \right)}{\log_{10} \left( \frac{Y}{CT} \right)}$ 

\[ Y = 10^{\frac{R}{R-1}} \log_{10} \frac{X}{CT} , \quad CT = \left( \frac{X}{CT} \right)^{\frac{1}{R}} \cdot CT \]

Control factor $g = \frac{Y}{X} = \left( \frac{X}{CT} \right)^{\frac{1}{R-1}}$
How to measure signal level?

- Use signal instantaneous amplitude? Varies too fast, not correlated with perceived loudness.
- Use some smoothed version of signal.

1. Signal envelope: sample and hold $|x[n]|$
   discontinuous.

2. A smooth peak/envelope follower:

   \[
   \begin{align*}
   X_{\text{peak}}[n] &= (1 - a) X_{\text{peak}}[n-1] + a |x[n]|, & 0 \leq a \leq 1, \\
   & \quad \text{if } |x[n]| \geq X_{\text{peak}}[n-1] \\
   X_{\text{peak}}[n] &= (1 - r) X_{\text{peak}}[n-1], & \quad \text{if } |x[n]| < X_{\text{peak}}[n-1]
   \end{align*}
   \]

   View the equations as filters from $|x[n]|$ to $X_{\text{peak}}[n]$: Non-linear!
   But filter from $|x[n]|$ to $X_{\text{peak}}[n]$: Linear!

   Transfer function from $|x[n]|$ to $X_{\text{peak}}[n]$

   Attack: $H(z) = \frac{a}{1 - (1-a)z^{-1}}$

   Release: $H(z) = \frac{1}{1 - (1-r)z^{-1}}$

   Single pole at $z = 1-a$ or $1-r$. Since $0 \leq a, r \leq 1$, low pass filter.

   $a$ and $r$ controls how fast the attack and release is.

   Set $a = 1 - e^{-\frac{1}{2T_a}}$, $r = 1 - e^{-\frac{1}{2T_r}}$

   $T_a$: attack time $T_r$: release time
Usually set $T_a \sim 10 \text{ ms}$, $T_r \sim 100 \text{ ms}$.

Perceptual range of attack and release time

"smooth" click  click  thud

\[
\begin{align*}
\text{attack} & \quad 0.1 \text{ ms} & 1 \text{ ms} & 10 \text{ ms} & \Rightarrow \text{transparent} \\
\text{release} & \quad \text{buzz (add harmonics)} & \text{roomy} & \text{smooth out phasing} & \text{equalize long-term dynamics}
\end{align*}
\]

\[
\begin{align*}
0.2 \text{ ms} & \quad 5 \text{ ms} & 100 \text{ ms} & 1 \text{ sec}
\end{align*}
\]

\[
x[n] \rightarrow Z^{-D} \rightarrow x[n-D] \rightarrow y[n]
\]

\[
\begin{align*}
\text{level measurement} & \quad (X_{\text{peak}}[n]) \\
\text{X-ray} & \quad (X_{\text{ray}}[n])
\end{align*}
\]

\[
\begin{align*}
\text{Static Curve} & \quad f[n] \quad \text{Attack/Release}
\end{align*}
\]

\[
g[n]
\]

Spectrum of $y[n]$ is convolution of spectrum of $X[n]$ and spectrum of $g[n]$.

If $g[n]$ varies fast $\Rightarrow G(f)$ is broad band $\Rightarrow$ smears $X(f)$

This is the reason why we use smoothed way to estimate signal level.
Signal Level: RMS value.

- Common formula: \[ X_{\text{RMS}}[n] = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} X^2[n-i]} \]

  Problem: computationally expensive; time or memory.

- An auto regressive formula:
  \[ X_{\text{RMS}}^2[n] = (1 - \beta) X_{\text{RMS}}^2[n-1] + \beta X^2[n] \]

  Weights on previous samples

\[ X[n] \rightarrow X_{\text{RMS}}[n] \]

Linear filter

\[ H(z) = \frac{\beta}{1 - (1-\beta)z^{-1}} \]

When \( z=1 \), \( H(z) = 1 \), so DC component got unit gain

\[ \sum h[n] = 1 \]

To make gain factor even smoother, we apply attack/release time on gain factor as well:

\[ g[n] = (1 - \frac{\beta}{k}) g[n-1] + k f[n] \]

\( f[n] \): gain factor from the static curve.

\( k \): \( \alpha \) or \( \gamma \), depending the state of gain factor

| attack: \( f[n] > g[n-1] \)
| release: \( f[n] < g[n-1] \)
Some implementation details:

1. Downsample $X_{peak}[n]$ or $X_{rms}[n]$ by a factor of 2 or 4 to reduce computation. They are smooth signal anyway.

2. Use smoother curve to reduce bandwidth of gain factor and reduce harmonic and inharmonic distortion.

---

3. Stereo processing: estimate signal level with both channels jointly.

Apply same gain factor on both channels.