ECE 272/472. Lecture 11: Dynamic Range Control.

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

(If we allow $|X|_{min}$ to be 0, i.e., silence, then the dynamic range is ∞ dB, which doesn't make much sense. In practice, we consider $|X|_{min}$ 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 dB ~ 120 dB ← human ears' range.

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

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

Tchaikovsky's 1812 Overture: pppppp ~ ffff

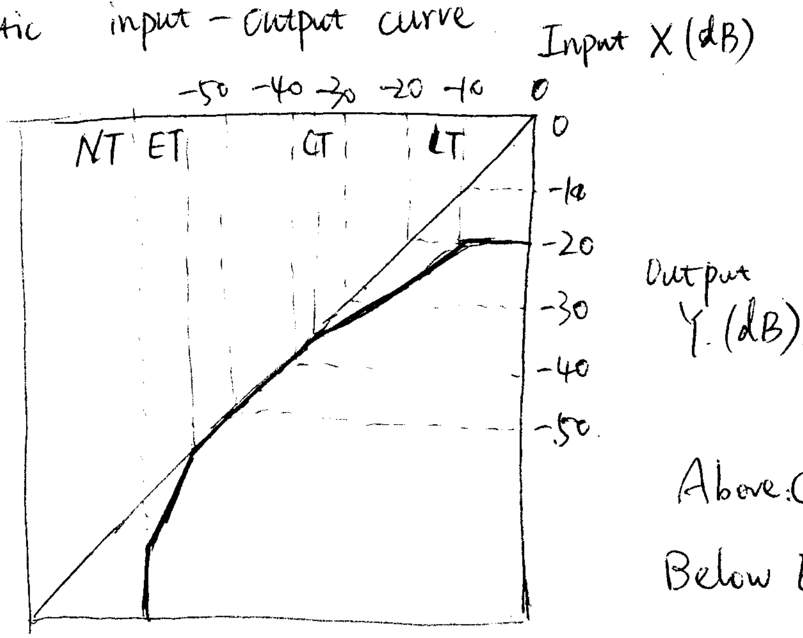
Types of ~~Why~~ dynamic range controls?

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

Why?

- Environment/channel may has small dynamic range, e.g., car, radio broadcast.
- Suppress noise floor
- Storage medium may have small range, e.g., tape

Static input - output curve



NT: noise gate threshold
ET: expander threshold
CT: compressor threshold
LT: limiter - - - -

Above CT: compression
Below ET: expansion

Let's focus ~~of~~ on compression here between CT and LT:

Compression ratio: $R = \dfrac{\Delta P_I}{\Delta P_o}$ ← Input level change
↖ Output - - - - -

$$\therefore R = \frac{X_{dB} - CT}{Y_{dB} - CT}$$

$$Y_{dB} = CT + R^{-1}(X_{dB} - CT), \quad \text{slope: } R^{-1}$$

$R = \infty$: Limiter
$R > 1$: compressor (compression ratio)
$0 < R < 1$: expander (expander ratio)
$R = 0$: noise gate

input signal level in linear scale

go to linear amplitude: $R = \dfrac{\log_{10}\left(\frac{x}{Ct}\right)}{\log_{10}\left(\frac{y}{Ct}\right)}$ ← output - - - - - - - - -

$$\therefore y = 10^{\frac{1}{R}\cdot \log_{10}\frac{x}{Ct}} \cdot Ct = \left(\frac{x}{Ct}\right)^{\frac{1}{R}} \cdot Ct$$

$$\therefore \text{control factor } g = \frac{y}{x} = \left(\frac{x}{Ct}\right)^{\frac{1}{R}-1}$$

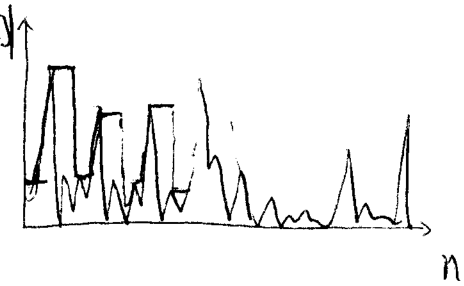
How to measure signal level?

— Use signal instaneous amplitude? Varies too fast, not correlated with perceived loudness.

— Use some smoothed version of signal.

① Signal envelope: sample and hold

discontinuous.



② A smooth peak / envelope follower.

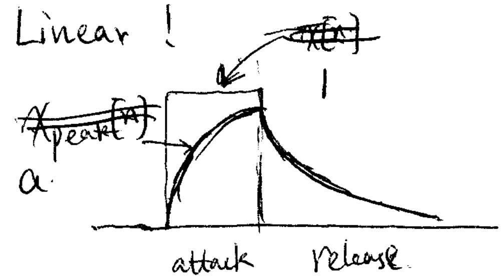
attack
$$X_{peak}[n] = (1-a)\,X_{peak}[n-1] + a\,|x[n]| \qquad 0 \le a \le 1. \quad \text{if } |x[n]| \ge X_{peak}[n-1]$$

release
$$X_{peak}[n] = (1-r)\,X_{peak}[n-1], \qquad \text{if } |x[n]| < X_{peak}[n-1]$$

View the equations as filters from $x[n]$ to $X_{peak}[n]$: Non-linear!

But filter from $|x[n]|$ to $X_{peak}[n]$: Linear!

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



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 $0 \le a, r \le 1$. Low pass filter.

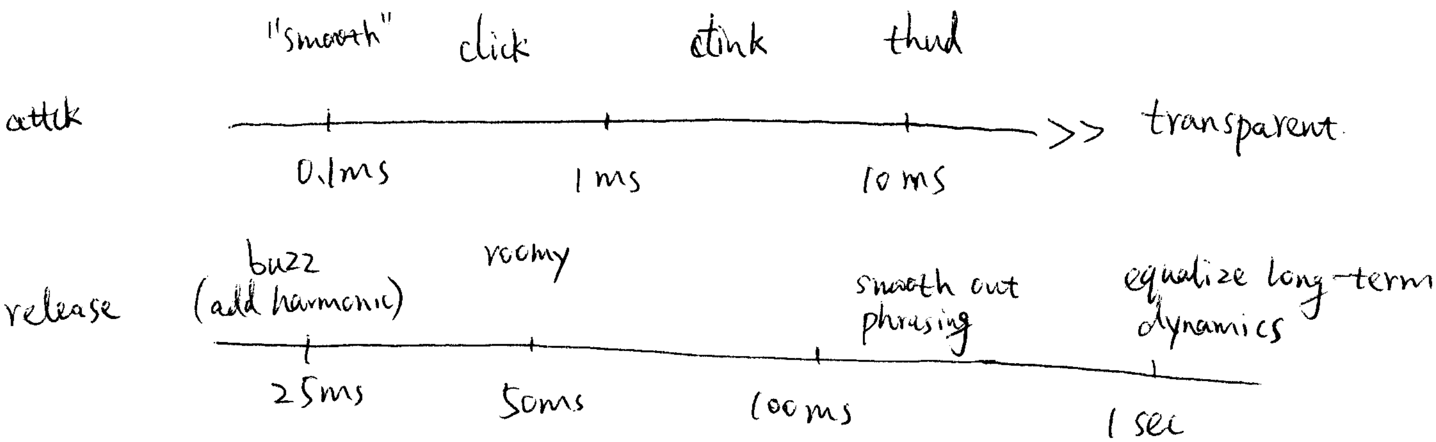
a and r controls how fast the attck and release is.

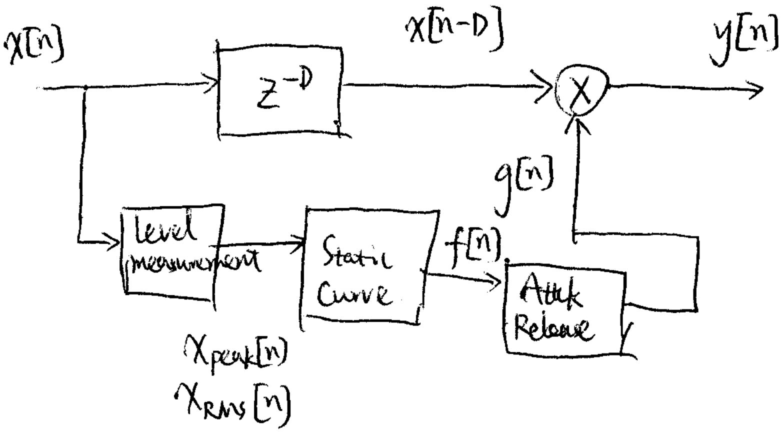
Set $a = 1 - e^{-1/\tau_a f_s}$, $r = 1 - e^{-1/\tau_r f_s}$

τ_a: attck time $\qquad$ τ_r: release time

Usually set $T_a \sim 10\,ms$. $T_r \sim 100\,ms$.

Perceptual range of attck and release time





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

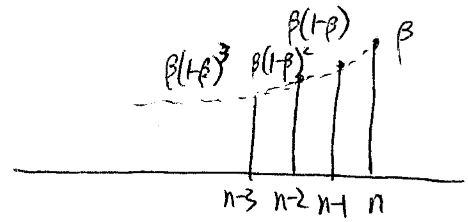
— Common formular: $x_{RMS}[n] = \sqrt{\frac{1}{N}\sum_{i=0}^{N-1} x^2[n-i]}$

Problem: computationally expensive: time or memory.

— An auto regressive formular:

$$x^2_{RMS}[n] = (1-\beta)\, x^2_{RMS}[n-1] + \beta\, x^2[n]$$

weights on perious samples



$x^2[n] \longrightarrow x^2_{RMS}[n]$

Linear filter

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

when $z=1$, $H(z)=1$, so DC component got unit gain

$\therefore \sum h[n] = 1$.

To make gain factor even smoother, we apply attck/release time on gain factor as well!

$$g[n] = (1-k)\, g[n-1] + k f[n]$$

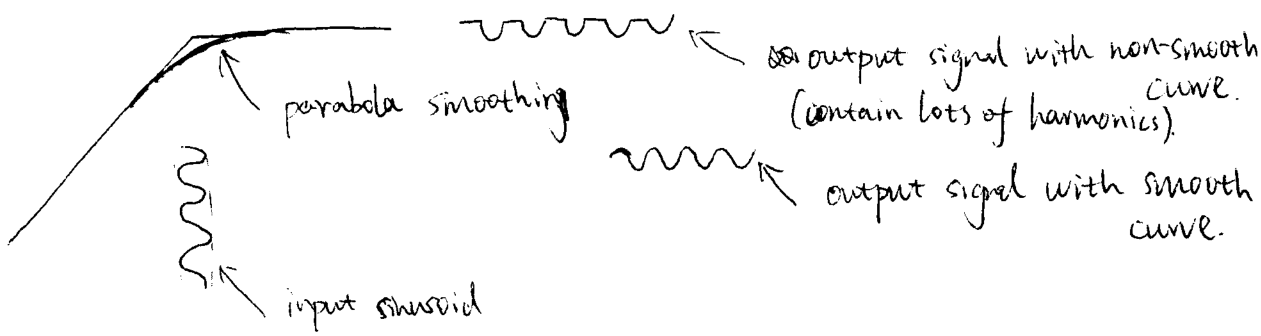
$f[n]$: gain factor from the static curve.

k: a or r, depending the state of gain factor $\begin{cases} \text{attck: } f[n] \geq g[n-1] \\ \text{release: } f[n] < g[n-1] \end{cases}$

Some implementation details:

1. Downsample $X_{peak}[n]$ or $X_{RMS}[n]$ by 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 ~~and~~ with ~~from~~ both channels jointly. Apply same gain factor on both channels.