Josh Reiss
Centre for Digital Music
Queen Mary, University of London

www.elec.qmul.ac.uk/digitalmusic/audioengineering/

UNDER THE HOOD OF A DYNAMIC RANGE COMPRESSOR

With thanks to Michael Massberg and Dimitrios Giannoulis

Questions and Challenges

"No two compressors sound alike... each one is inaccurate in its own unique way" – Roey Izhaki, Mixing Audio: Concepts, Theory and Practice

- Given how a compressor is used, how can you build one?
- How can you emulate an analog design?
- What makes compressors sound different?
- Why do attack and release times mean different things for different compressors?
- How can you measure the performance of a compressor?

What we're <u>not</u> going to talk about

- How to use a dynamic range compressor
  - Preferred parameter settings for different tasks
- How not to use a dynamic range compressor (Loudness war)
  - But will show how better design helps prevent misuse
- Hearing aids
  - Still applicable, but we're concerned with compressors for music production, mastering and broadcast
- Sidechaining and parallel compression
- Multiband compression
- Additional parameters
  - Look-ahead, hold...

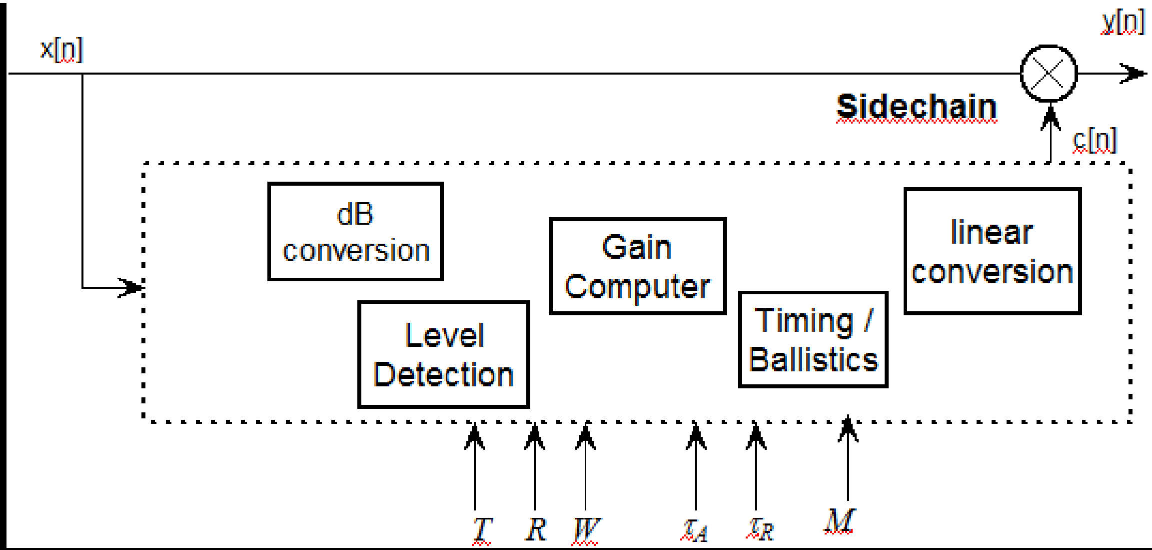
Introducing the dynamic range compressor

- map dynamic range of audio signal to smaller range
  - Reduce signal level of higher peaks while leaving quieter parts untreated
  - Then (optionally) boosting whole signal to make it as loud as, or louder than, original uncompressed signal
- Unlike gain, delay, EQ, panning...
  - Poorly defined
    - Not always agreed what compression parameters do
  - An *adaptive* audio effect
    - Action of the compressor depends on the input signal
  - Nonlinear
    - Generally noninvertible

Compressor Controls

- Threshold- T
  - level above which compression starts
- Ratio- R
  - determines the amount of compression applied.
- Knee Width- W
  - Controls transition in the ratio around the threshold
- Make-up Gain- M
  - Used to match the output signal's overall loudness to the original
- Attack Time- τ_A
  - Time to decrease gain once signal overshoots threshold
- Release Time- τ_R
  - Time to bring gain back to normal once signal falls below threshold

Main components



- Side-chain
  - Level detection
    - Estimate the 'signal level'
  - Decibel conversion
  - Gain computer
    - Apply threshold, ratio, knee
  - Timing (Ballistics, Smoothing)
    - Apply attack and release
- Gain stage
  - Output is input multiplied by control vector produced by side-chain

$x_L[n] \rightarrow y_L[n]$

$y_{dB}[n] = 20\log_{10} x_{dB}[n]$

$x_G[n] \rightarrow y_G[n]$

$x_T[n] \rightarrow y_T[n]$

$y_{dB}[n] = x_{dB}[n] + c_{dB}[n] + M$

Level Detection – the options

- Peak vs RMS
  - Peak detection - Based on absolute value of the signal

$$y_L[n] = |x_L[n]|$$

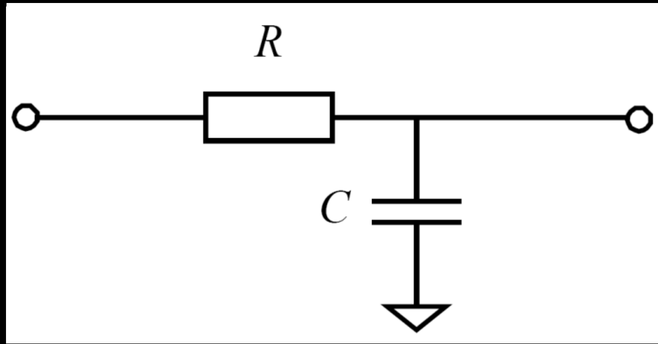
  - RMS - Based on square of the signal
    - Square of the level is average of square of the input

$$y_L^2[n] = \frac{1}{M} \sum_{m=-M/2}^{M/2-1} x_L^2[n-m]$$

Or Square of the level is smoothed estimate of square of the input

$$y_L^2[n] = (1-\alpha) y_L^2[n-1] + \alpha x_L^2[n]$$

Exponential Moving Average Filter



$$y[n] = \alpha y[n-1] + (1-\alpha)x[n]$$

- Also known as smoothing filter, one pole low pass filter...
- Step response

$$y[n] = 1-\alpha^n \quad \text{for} \quad x[n]=1, n \geq 1$$

- Time constant - time it takes a system to reach 1-1/e= 63% of its final value

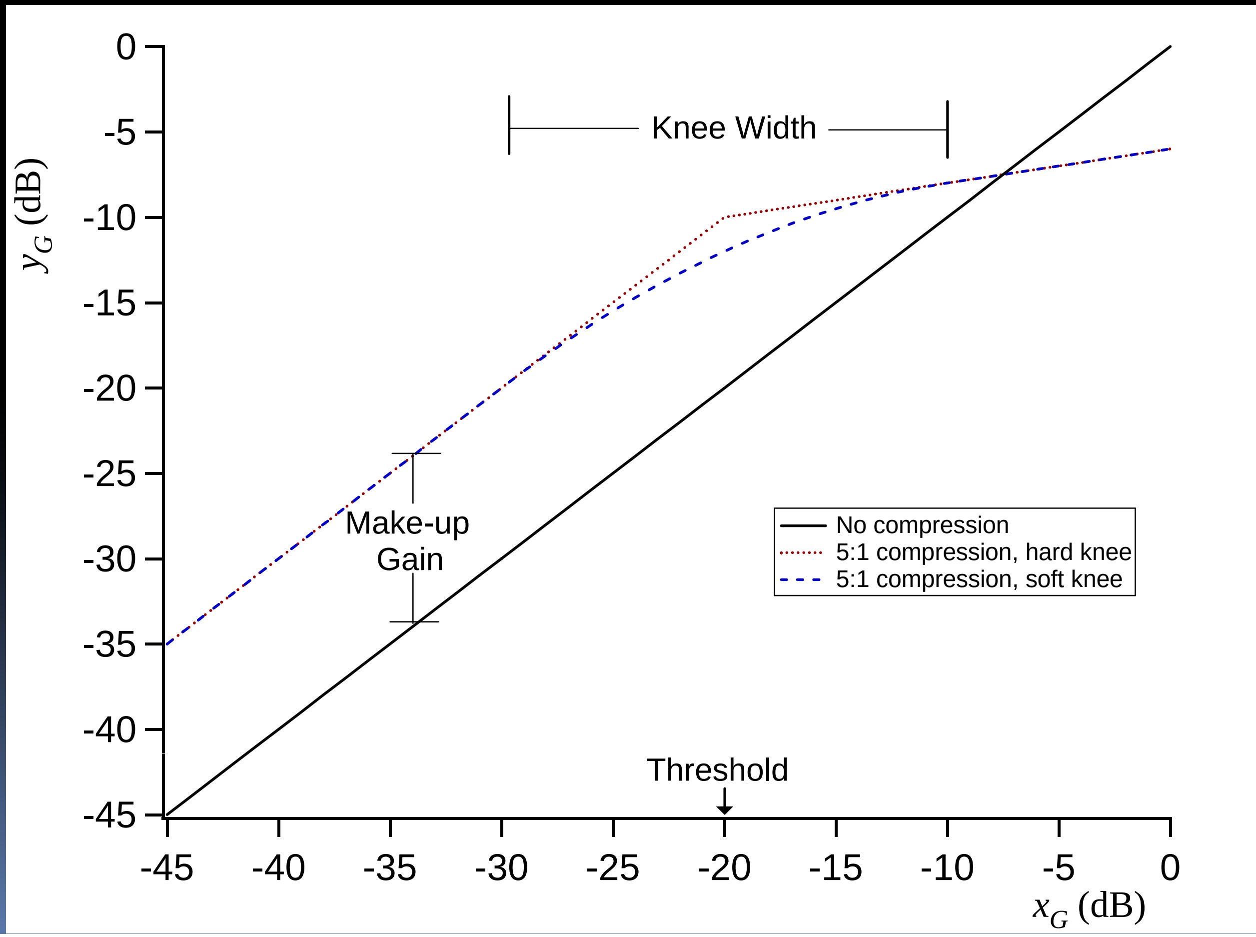
$$y[\tau f_s] = 1-\alpha^{\tau f_s} = 1-e^{-1}$$
$$\rightarrow$$
$$\alpha = e^{-1/\tau f_s}$$

Issues with RMS level detection

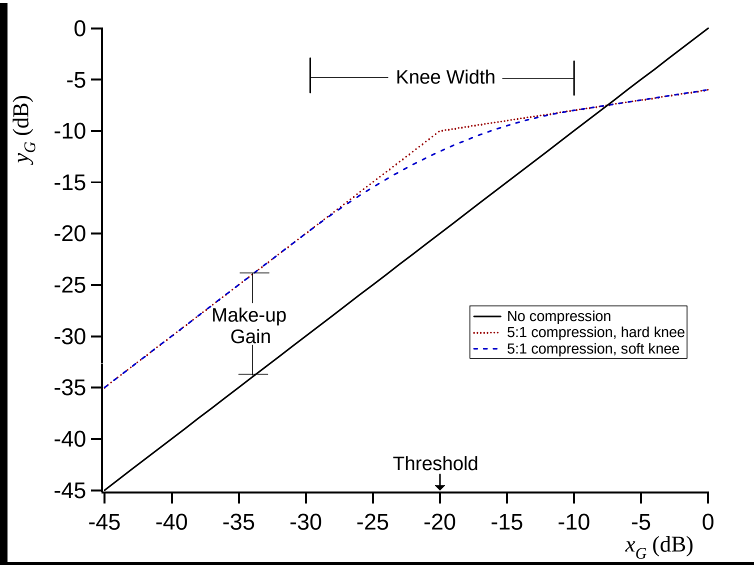
- Introduces significant delay
- Introduces additional parameter, M or τ
- Several studies suggest general behaviour is the same for peak & RMS detection
  - F. Floru, "Attack and Release Time Constants in RMS-Based Feedback Compressors," Journal of the Audio Engineering Society, vol. 47, pp. 788-804, October 1999.
  - J. S. Abel and D. P. Berners, "On peak-detecting and rms feedback and feedforward compressors," in 115th AES Convention, 2003.

The Gain Computer

Static Compression Curve



The Gain Computer



- Compression ratio
  - Hard knee

$$R = \frac{x_G - T}{y_G - T} \quad for \quad x_G > T$$

$$y_G = \begin{cases} x_G & x_G \le T \\ T + (x_G - T)/R & x_G > T \end{cases}$$

  - Soft knee
    - Linear interpolation (bad)

$$y_G = \begin{cases} x_G & 2(x_G - T) < -W \\ T - W/2 + (x_G - T + W/2)(1 + 1/R)/2 & 2|(x_G - T)| \le W \\ T + (x_G - T)/R & 2(x_G - T) > W \end{cases}$$

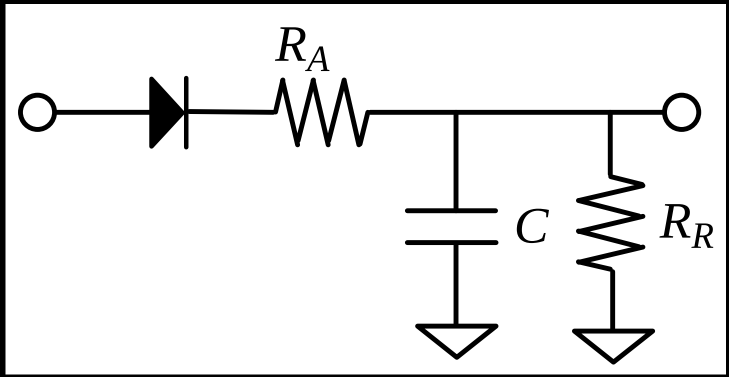
    - Second order interpolation (good)

$$y_G = \begin{cases} x_G & 2(x_G - T) < -W \\ x_G + (1/R - 1)(x_G - T + W/2)^2/(2W) & 2|(x_G - T)| \le W \\ T + (x_G - T)/R & 2(x_G - T) > W \end{cases}$$

Time constants

- From the smoothing filter
- Officially, time it takes a system to reach 1-1/e= 63% of its final value
  - τ_A – level starts at 0, goes to 1. Attack time is the time it takes for level to reach 0.63.
  - τ_R – level starts at 1, goes to 0. Release time is the time it takes for level to reach 1-0.63=0.37.
- But some compressors,
  - Miscalculate it
  - Define it as time to go from 90% of initial value to 10% of final value
  - Define it as time to change level by so many dBs
  - …

Analog Peak Detector



- Assuming ideal diode,

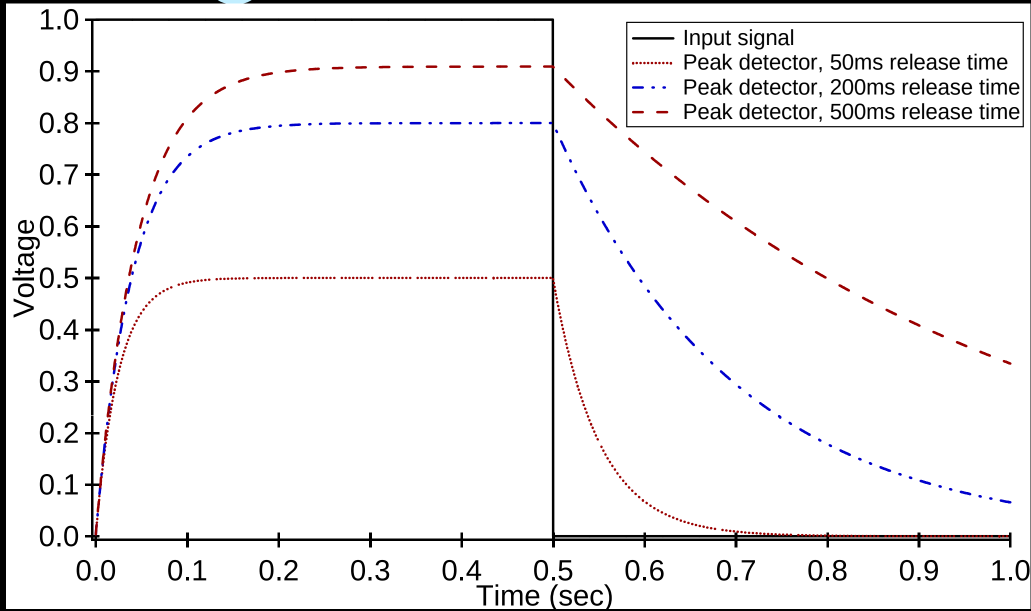
$$y_T[n] = \alpha_R y_T[n-1] + (1-\alpha_A)\max(x_T[n] - y_T[n-1], 0)$$

  - Where $\tau_A = R_A C$ and $\tau_R = R_R C$

$$\alpha_A = e^{-1/\tau_A f_s}, \alpha_R = e^{-1/\tau_R f_s}$$

- Used in
  - D. Berners, "Analysis of Dynamic Range Control (DRC) Devices," Universal Audio WebZine, vol. 4, September 2006.
  - P. Dutilleux and U. Zölzer, "Nonlinear Processing, Chap. 5," in DAFX - Digital Audio Effects, U. Zoelzer, Ed., 1st ed: Wiley, John & Sons, 2002

Analog Peak Detector

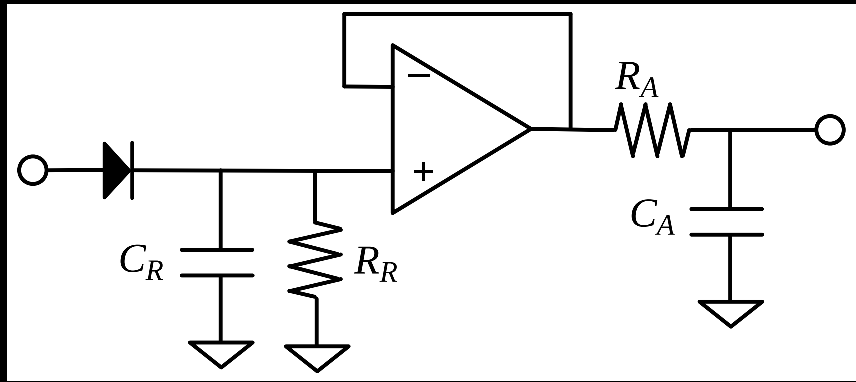


- Step response

$$y[n] = (1-\alpha_A)\sum_{m=0}^{n-1}(\alpha_R+\alpha_A-1)^m \rightarrow \frac{1-\alpha_A}{2-\alpha_R-\alpha_A} \approx \frac{\tau_R}{\tau_R+\tau_A}$$

- correct peak estimate only when release time is considerably longer than attack time
- attack time gets slightly scaled by release time
  - faster attack time than expected when we use a fast release time

Decoupled Peak Detector



$$y_1[n] = \max(x_T[n], \alpha_R y_1[n-1])$$
$$y_T[n] = \alpha_A y_T[n-1] + (1-\alpha_A) y_1[n]$$

- attack envelope is now impressed upon release envelope
- good estimate of actual release time given by adding attack time constant to release time constant

Branching Peak Detector

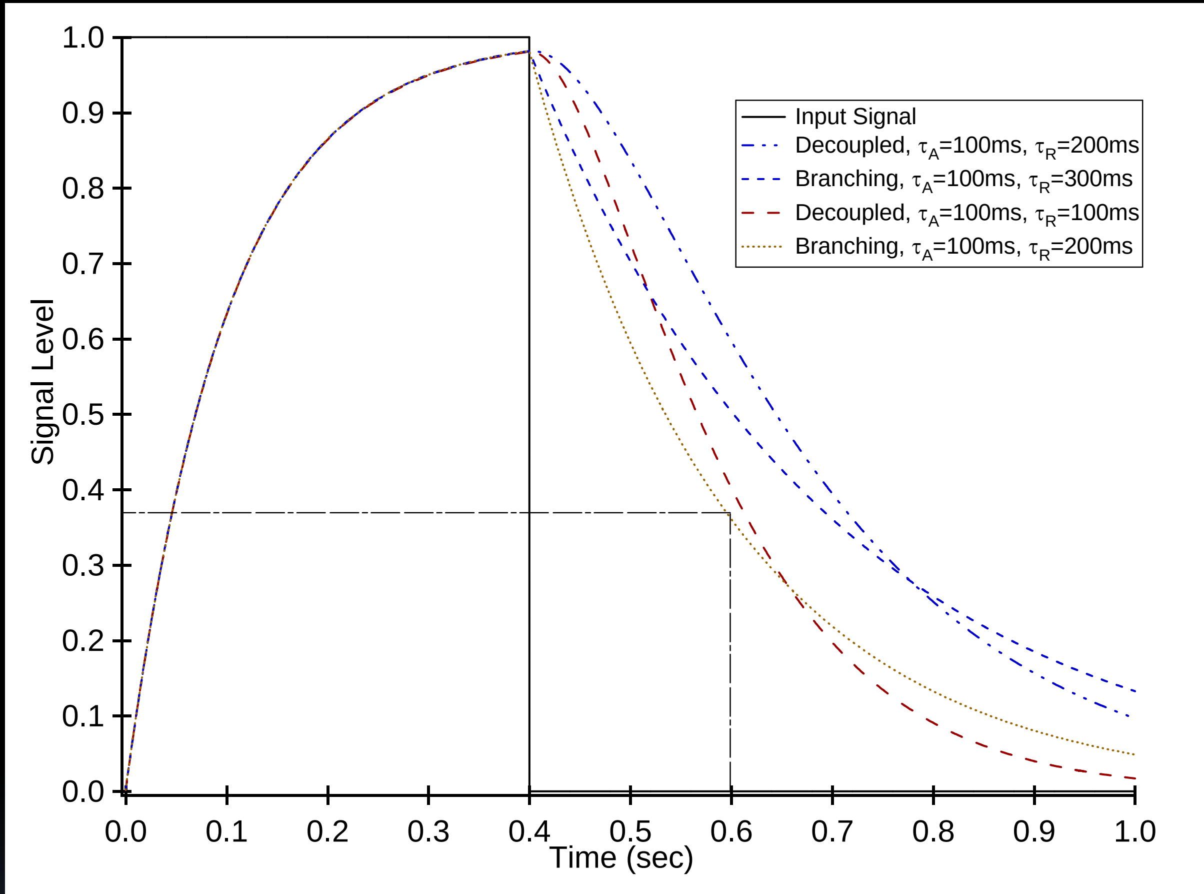
$$y_T[n] = \begin{cases} \alpha_A y_T[n-1] + (1-\alpha_A) x_T[n] & x_T[n] > y_T[n-1] \\ \alpha_R y_T[n-1] & x_T[n] \le y_T[n-1] \end{cases}$$

- Easily accomplished in digital implementation

G. W. McNally, "Dynamic Range Control of Digital Audio Signals," Journal of the Audio Engineering Society, vol. 32, pp. 316-327, May 1984.

U. Zolzer, Digital Audio Signal Processing, 2nd ed.: John Wiley and Sons, Ltd., 2008.

Time constants in decoupled & branching peak detectors



- Output of decoupled and branching peak detector circuits for different release time constants
  - branching peak detector produces intended release time constant
  - decoupled peak detector produces measured time constant ~ $\tau_A + \tau_R$.

Smooth, level corrected peak detector

- Smooth decoupled peak detector

$$y_1[n] = \max(x_T[n], \alpha_R y_1[n-1] + (1-\alpha_R) x_T[n])$$
$$y_T[n] = \alpha_A y_T[n-1] + (1-\alpha_A) y_1[n]$$

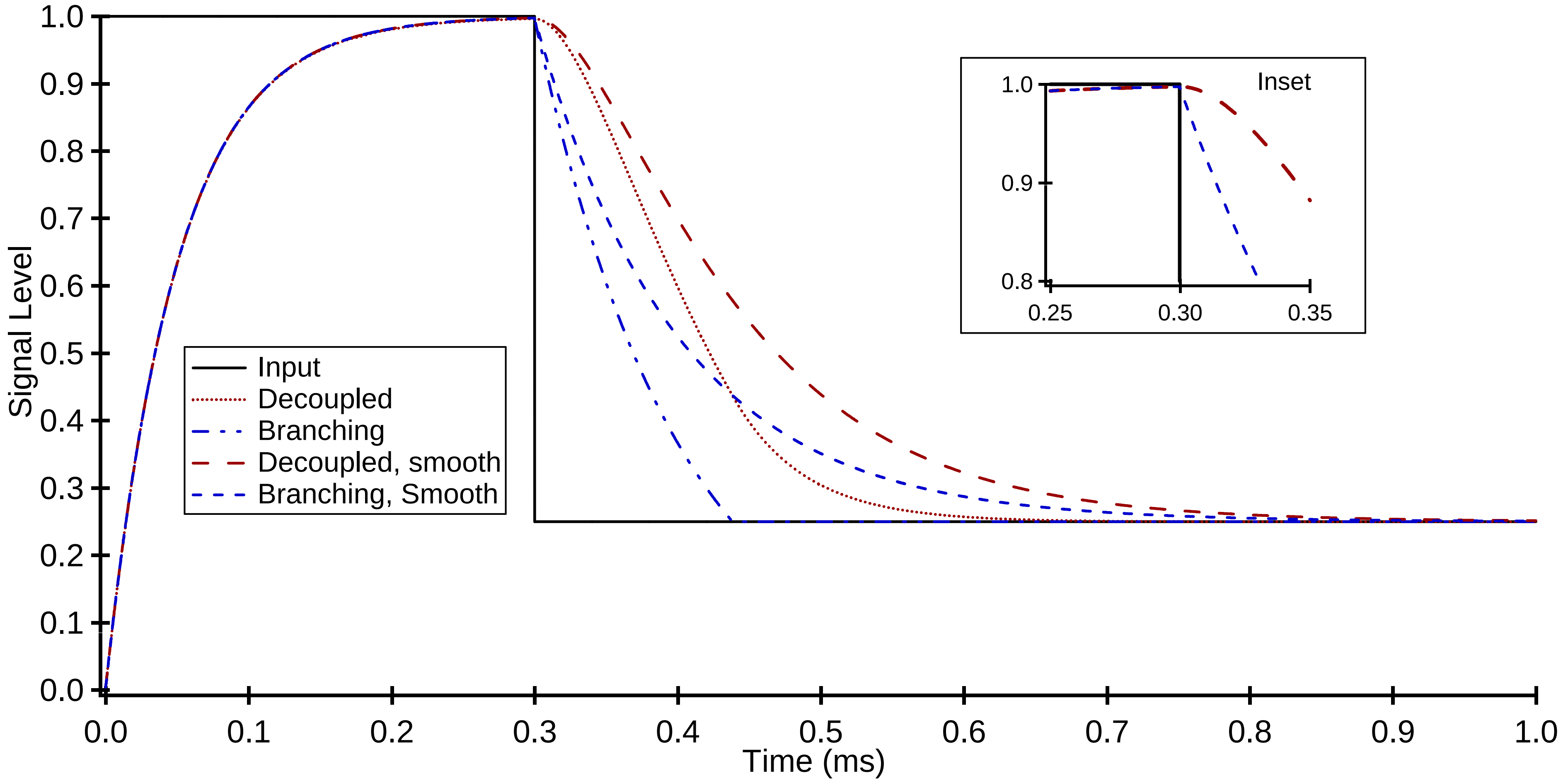
- Smooth branching peak detector

$$y_L[n] = \begin{cases} \alpha_A y_T[n-1] + (1-\alpha_A) x_T[n] & x_T[n] > y_T[n-1] \\ \alpha_R y_T[n-1] + (1-\alpha_R) x_T[n] & x_T[n] \le y_T[n-1] \end{cases}$$

  - Used in
    - P. Dutilleux, et al., "Nonlinear Processing, Chap. 4," in Dafx:Digital Audio Effects, 2nd ed: Wiley, 2011, p. 554.
    - U. Zolzer, Digital Audio Signal Processing, 2nd ed.: John Wiley and Sons, Ltd., 2008.

Comparison of Digital Peak Detectors

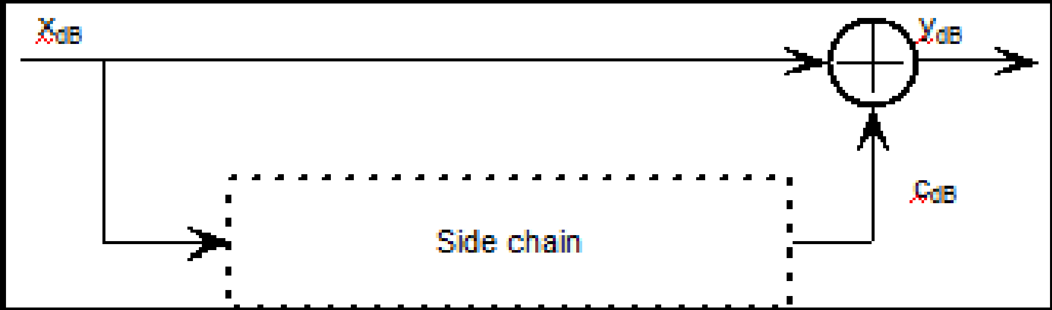
- All envelopes reach maximum peak value and feature similar attack trajectories
- Release envelopes too short for decoupled and branching peak detectors
  - smoothed versions make full use of release time
- Discontinuity for smooth, branching peak detector
  - abrupt switch from attack to release
  - release envelope is continuous for smoothed, decoupled peak detector



Feedforward or feedback design

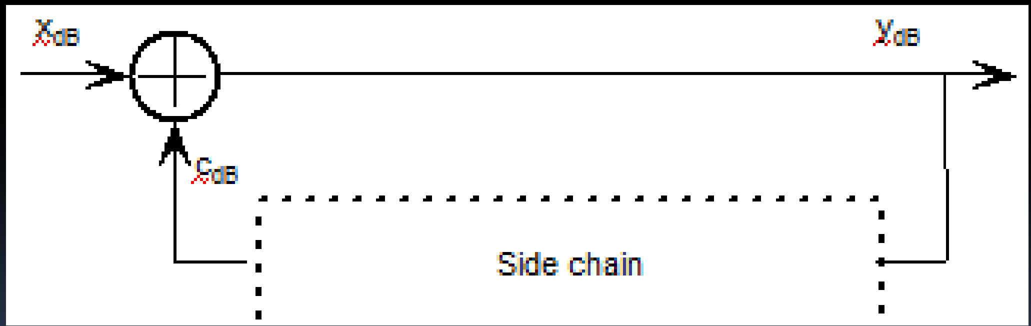
- Feedforward

$$y_{dB} = c_{dB} + x_{dB}$$
$$y_{dB} = T + (x_{dB} - T)/R$$
$$\rightarrow$$
$$c_{dB} = (1/R - 1)(x_{dB} - T)$$



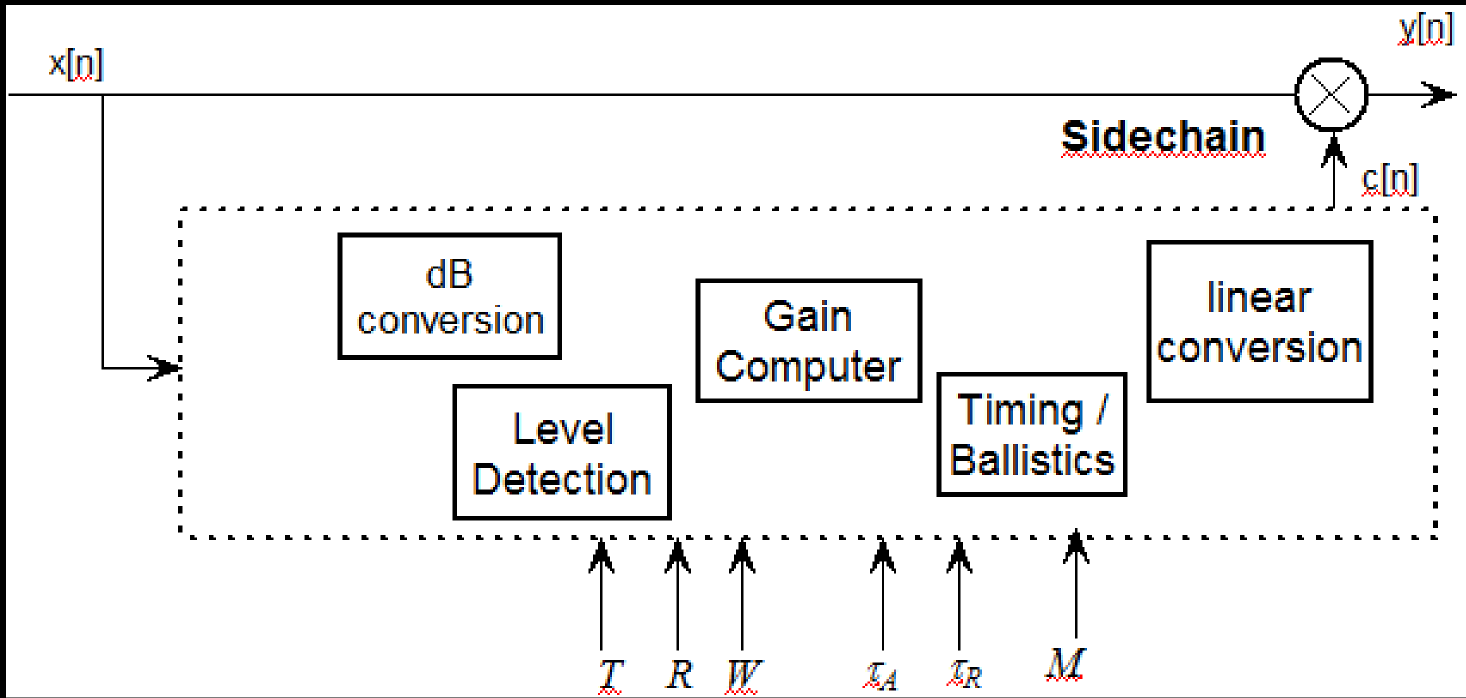
- Feedback

$$y_{dB} = c_{dB} + x_{dB}$$
$$y_{dB} = T + (x_{dB} - T)/R$$
$$\rightarrow$$
$$c_{dB} = (1 - R)(y_{dB} - T)$$



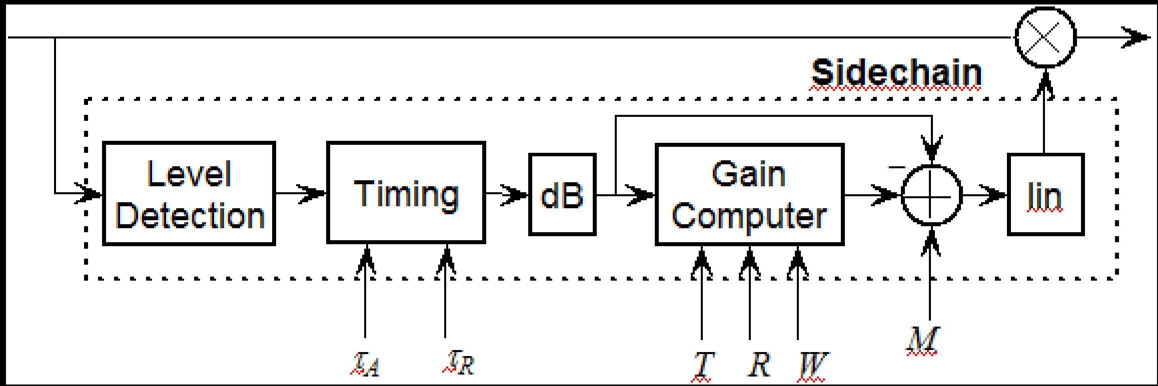
  - Limiter (ratio of ∞: 1) needs infinite negative amplification
  - Not designed for look-ahead

Sidechain configuration



- Level detection $x_L[n] \rightarrow y_L[n]$
- Decibel conversion $y_{dB}[n] = 20\log_{10} x_{dB}[n]$
- Gain computer $x_G[n] \rightarrow y_G[n]$
- Timing (Ballistics, Smoothing) $x_T[n] \rightarrow y_T[n]$

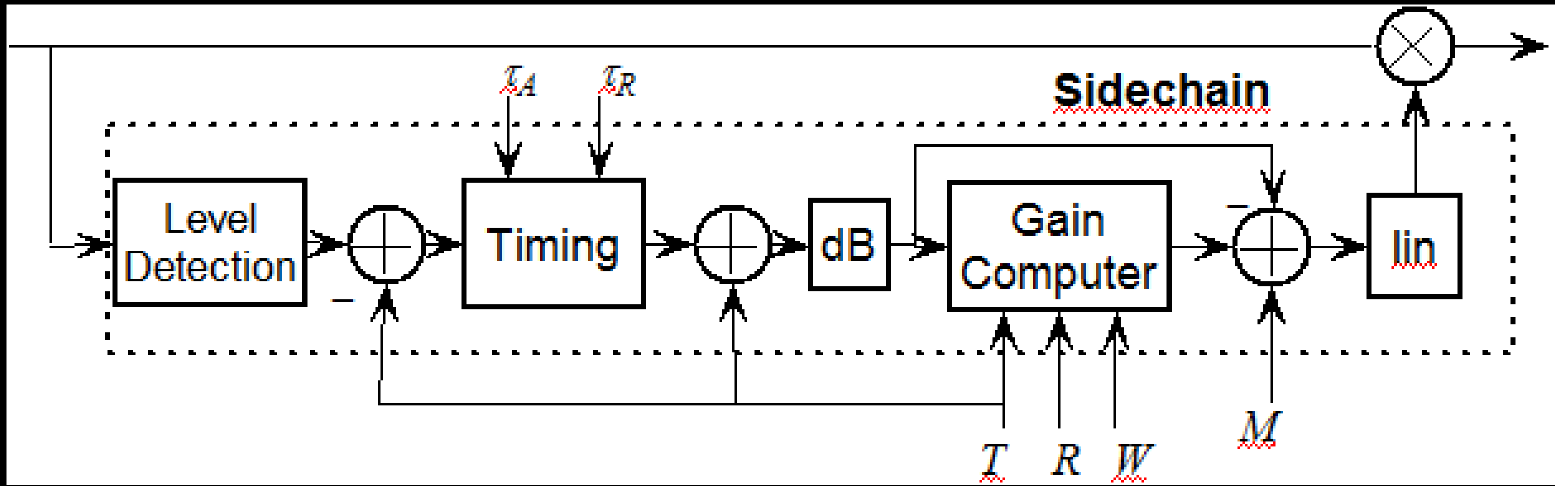
Timing Placement- linear domain



$$x_T[n] = y_L[n]$$
$$x_G[n] = 20\log_{10} y_L[n]$$
$$c_{dB}[n] = y_G[n] - x_G[n] + M$$

- R. J. Cassidy, "Level Detection Tunings and Techniques for the Dynamic Range Compression of Audio Signals," in 117th AES Convention, 2004.
- J. S. Abel and D. P. Berners, "On peak-detecting and rms feedback and feedforward compressors," in 115th AES Convention, 2003.
- U. Zolzer, Digital Audio Signal Processing, 2nd ed.: John Wiley and Sons, Ltd., 2008.
- P. Hämäläinen, "Smoothing of the Control Signal Without Clipped Output in Digital Peak Limiters," in International Conference on Digital Audio Effects (DAFx), Hamburg, Germany, 2002, pp. 195-198.
- S. J. Orfanidis, Introduction to Signal Processing Orfanidis (prev. Prentice Hall), 2010.

Timing Placement- biased linear domain



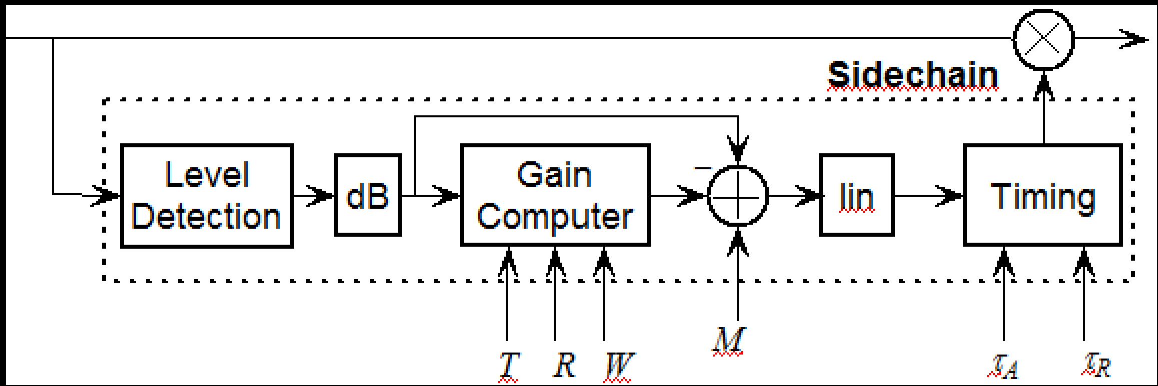
$$x_T[n] = |x_L[n]| - 10^{T/20}$$

$$x_G[n] = 20\log_{10}(y_T[n] + 10^{T/20})$$

$$c_{dB}[n] = y_G[n] - x_G[n] + M$$

- Attack and release now depend on threshold, not zero
- Envelope smoothly fades out once signal falls below threshold

Timing Placement- linear domain, post-gain



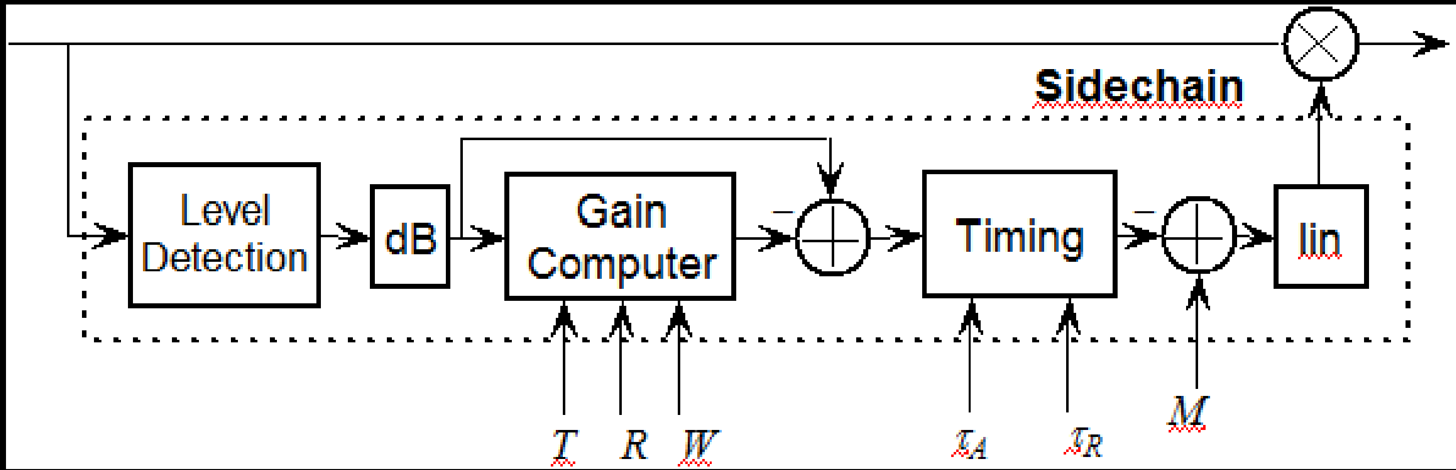
$$x_G[n] = 20\log_{10} x_L[n]$$

$$x_T[n] = 10^{(y_G[n] - x_G[n] + M)/20}$$

$$c[n] = y_T[n]$$

- J. Bitzer and D. Schmidt, "Parameter Estimation of Dynamic Range Compressors: Models, Procedures and Test Signals," presented at the 120th AES Convention, 2006
- L. Lu, "A digital realization of audio dynamic range control," in Fourth International Conference on Signal Processing Proceedings (IEEE ICSP), 1998, pp. 1424 - 1427.

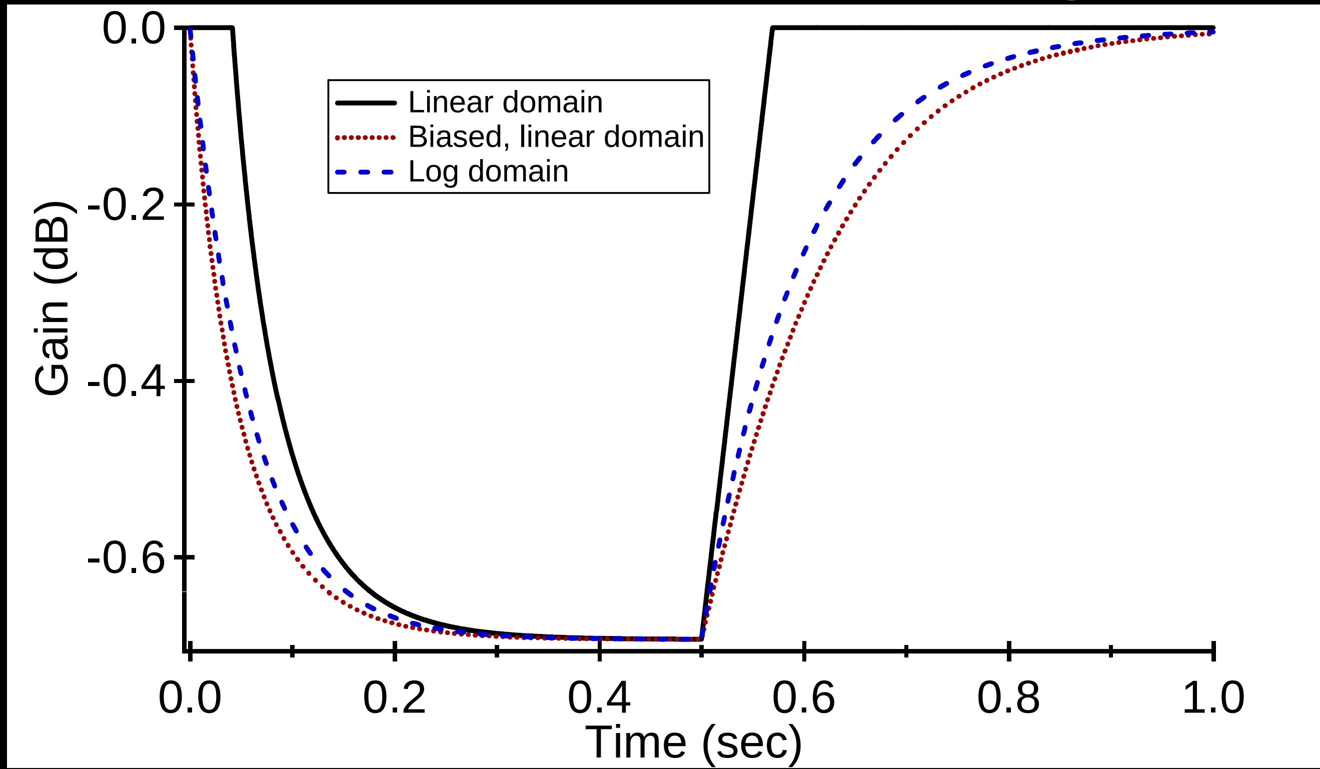
Timing Placement- log (dB) domain



$$x_G[n] = 20\log_{10} x_L[n]$$
$$x_T[n] = x_G[n] - y_G[n]$$
$$c_{dB}[n] = M - y_T[n]$$

- S. Wei and W. Xu, "FPGA implementation of gain calculation using a polynomial expression for audio signal level dynamic compression," J. Acoust. Soc. Jpn. (E), vol. 29, pp. 372-377, 2008.
- S. Wei and K. Shimizu, "Dynamic Range Compression Characteristics Using an Interpolating Polynomial for Digital Audio Systems," IEICE Trans. Fundamentals, vol. E88-A, pp. 586-589, 2005.

Performance – sidechain configuration



gain envelope produced by different placement of ballistics in sidechain configurations

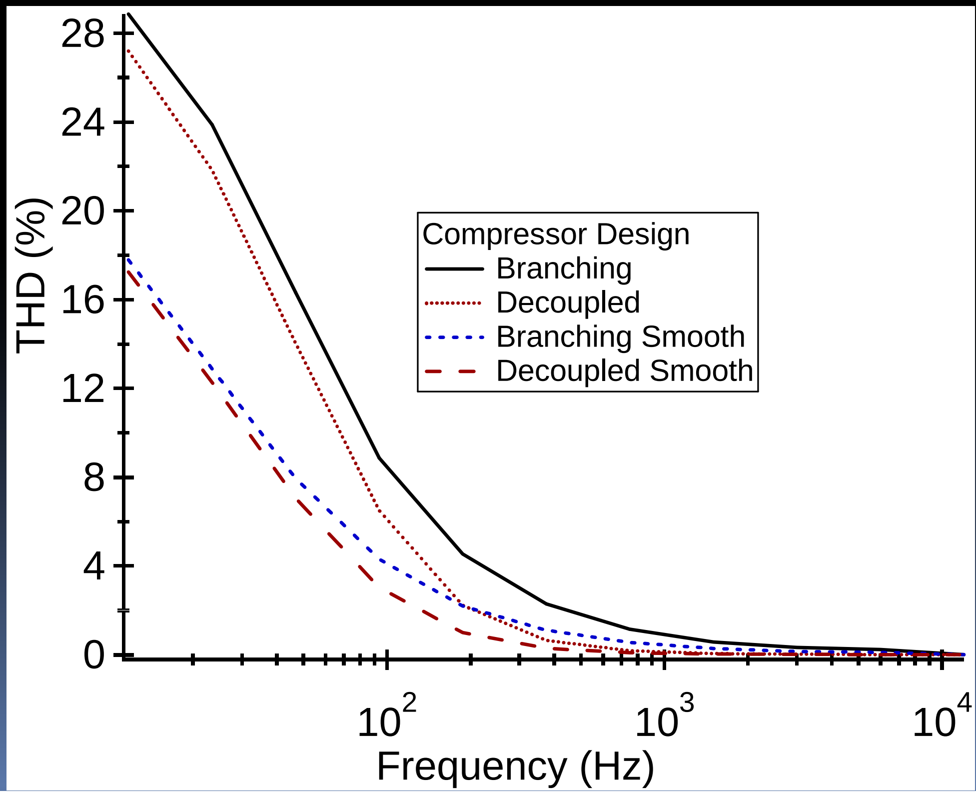
- release discontinuity in linear domain return-to-zero detector
- linear domain return-to-threshold (biased) and log domain detector curves look similar
    - faster attack, slower release for linear domain return-to-threshold

Performance- Artifacts

- How can we measure and characterise performance of a compressor?
- Most artifacts associated with parameter settings, not design
  - Dropouts
    - Long release -> keeps attenuating after transient finished
  - Overshoot
    - Long attack -> misses attenuation of initial transient
  - Lack of clarity
    - Short attack -> squash all transients
  - Pumping
    - Short release -> jump in signal level after transient
  - Breathing
    - Noticeable movement of noise floor

Performance- Noise and Distortion

- Signal to Noise Ratio
  - Not useful
- Total Harmonic Distortion
  - Measures strength of harmonics (introduced by compression) in relation to original frequency



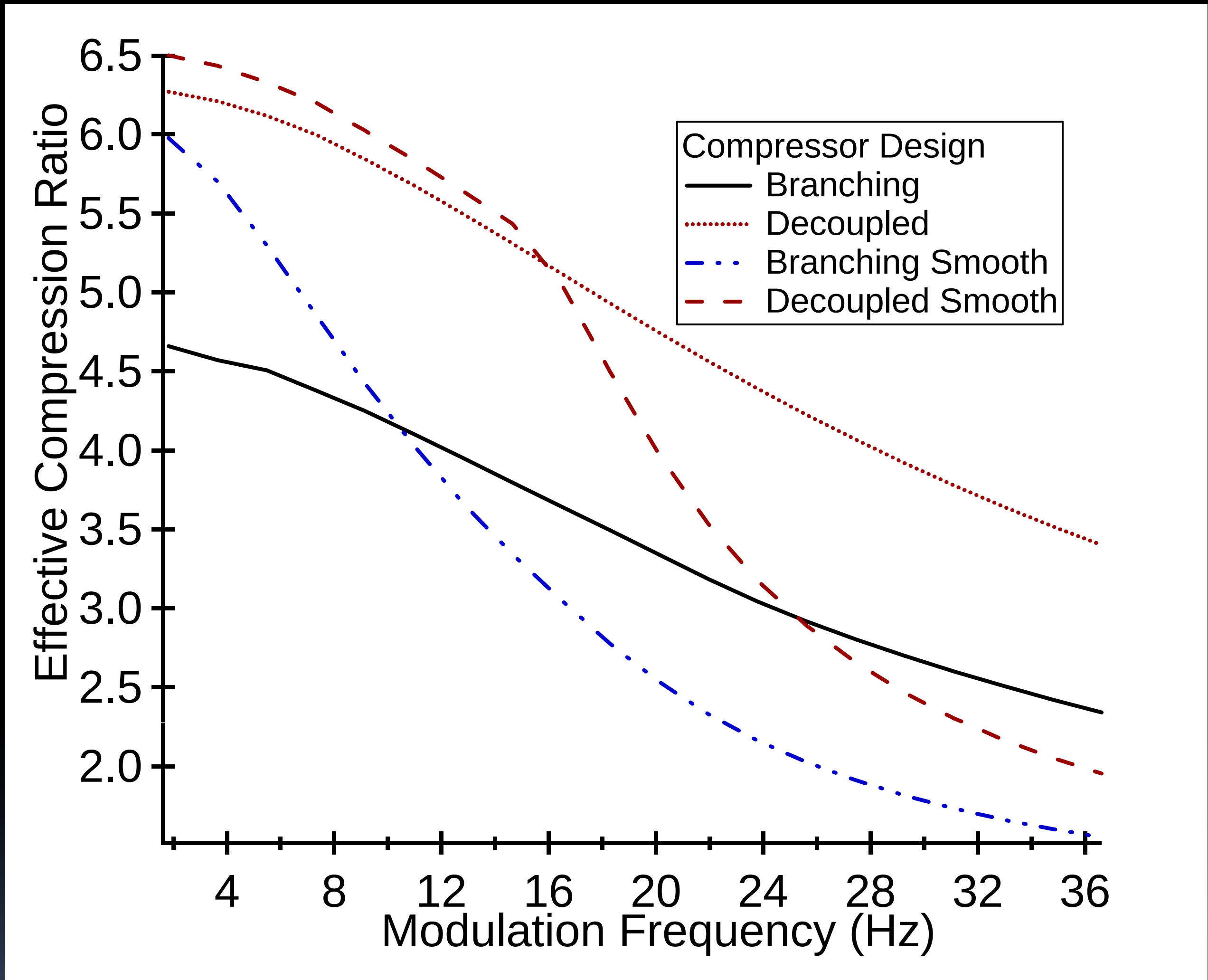
Performance – frequency dependence

- Effective compression ratio
  - used in assessing compressors in cochlear implants and hearing aids
  - amplitude modulated sine wave applied at input

$$x[n] = (1 + m\cos(2\pi f_m n))\cos(2\pi f_c n)$$

  - spectrum consists of a carrier and two side bands
  - ΔS_i difference between amplitude of side bands and amplitude of carrier of input signal
  - ΔS_o difference between amplitude of side bands and amplitude of carrier of compressed signal is found,.
  - effective compression ratio is then given by $\Delta S_i / \Delta S_o$.

Performance – effective compression ratio



- Interpretation?

Performance- Ballistics and sidechain configuration

Fidelity of the envelope shape

Detector Placement	Detector Type	Fidelity of Envelope Shape (FES)			
		Guitar	Bass	Drums	Vocals
Log domain	Branching	0.884	0.945	0.766	0.952
	Decoupled	0.899	0.932	0.755	0.941
	Branching Smooth	0.852	0.927	0.640	0.941
	Decoupled Smooth	0.859	0.911	0.648	0.936
Linear domain	Branching	0.836	0.879	0.517	0.932
	Decoupled	0.790	0.831	0.537	0.930
	Branching Smooth	0.825	0.856	0.456	0.932
	Decoupled Smooth	0.775	0.805	0.461	0.934

- Measures correlation between envelopes of compressed and uncompressed signals
- Nonsmooth designs do well since they often apply no compression
- Placing ballistics (attack and release) in log domain outperforms placing them in linear domain

Pseudocode

```
% Smoothed branching digital dynamic range compressor.
% Parameters
%   x: input signal
%   fs: sample rate (kHz)
%   tauAttack: attack time constant (ms)
%   tauRelease: release time constant (ms)
%   T: logarithmic threshold
%   R: compression ratio
%   W: decibel width of the knee transistion
%   M: decibel Make-up Gain
% Returns
%   y: output (compressed) signal
%LEVEL DETECTION
y_L = max(abs(x), 1e-6);
```

Pseudocode

```
%DECIBEL CONVERSION
x_dB= y_L;
y_dB = 20*log10(x_dB);
x_G=y_dB;
%GAIN COMPUTER
slope = 1 / Ratio - 1;% Feed-forward topology
overshoot = x_G - T;
if   (overshoot <= -W/2)
    y_G =x_G;
if ((overshoot > -W/2) & (overshoot < W/2))
    y_G =x_G+slope* (overshoot + W / 2) .^ 2/ (2*W)  ;
if   (overshoot >= -W/2)
    y_G =x_G+slope* overshoot ;

x_T=y_G-x_G;
```

Pseudocode

```
%BALLISTICS
alphaAtt = exp(-1 / (tauAttack * fs));
alphaRel = exp(-1 / (tauRelease * fs));
if x_T(0) > 0
   y_T(0) = (1-alphaAtt)*x_T(0);
else
   y_T(0) = (1-alphaRel)*x_T(0);
end
for i=2:length(x_T)
   if x_T(i) > y_T(i-1)
      y_T(i)= alphaAtt * y_T(i-1) + (1-alphaAtt)*x_T(i);
   else
      y_T(i)= alphaRel * y_T(i-1) + (1-alphaRel)*x_T(i);
   end
end
c_dB = M-y_T; %CONTROL VECTOR, WITH MAKE-UP GAIN
gain = 10.^(c_dB./20); %CONVERT TO LINEAR
y = x .* gain; %GAIN STAGE
```

The big unknown – psychoacoustics

- Compressor designs
  - in the literature: >20
  - Commercial designs: >40
- Studies of parameter settings
  - for hearing aids: >10
  - for music production and broadcast: 0 ?
    - Anecdotal recommendations : 10?
- Studies of compressor designs
  - for hearing aids: <5?
  - for music production and broadcast: 0 !

Recommendations

- Feedforward compressors are preferred
  - more stable and predictable than feedback designs
  - Benefits of feedback designs are only relevant in analog
- Peak level detection suggested over RMS
  - RMS introduces additional parameter
    - Often not user controlled
  - Smoothing done anyway by attack and release
  - May introduce unnecessary delay
- Soft knee with high order interpolation
  - Linear interpolation gives two hard knees

Recommendations

- Ballistics in the log domain after the gain computer
  - Maintain envelope shape
  - no attack lag
  - easy implementation of variable knee width
- Smooth, peak detector for attack and release
  - Decoupled
    - Low harmonic distortion
    - Prevent discontinuities
  - Branching
    - detailed knowledge of the effect of the time constants
    - Minor discontinuities in the slope of the gain curve
- Make-up gain at end of side-chain
  - Agrees with expected use
  - Not intended or needed to be smoothed

Thanks

- Michael Massberg
  - See his talk this afternoon, at 2:30pm
  - “Digital Low-Pass Filter Design with Analog Matched Magnitude Response”
- Dimitrios Giannoulis