Audio Engineering Society

Convention Paper 8020

Presented at the 128th Convention
2010 May 22–25 London, UK

A Novel Envelope-Based Generic Dynamic Range Compression Model

Adam Weisser[1]

[1] *Oticon A/S, Smørum 2765, Denmark*

Correspondence should be addressed to Adam Weisser (adw@oticon.dk, adam_weisser@fastmail.fm)

ABSTRACT

A mathematical model is presented, which reproduces typical dynamic range compression, when given the nominal input envelope of the signal and the compression constants. The model is derived geometrically in a qualitative approach and the governing differential equation for an arbitrary input and an arbitrary compressor is found. Step responses compare well to commercial compressors tested. The compression effect on speech using the general equation in its discrete version is also demonstrated. This model applicability is especially appealing to hearing aids, where the input-output curve and time constants of the non-linear instrument are frequently consulted and the qualitative theoretical effect of compression may be crucial for speech perception.

1. INTRODUCTION

Audio dynamic range compression is almost always met in today's hearing aids. The positive effect of reducing the perceived dynamic range of signals on speech intelligibility for the hearing-impaired has been well-established (e.g. see [1, 2]). How compressors work is often explained schematically in the audiological literature, using input-output and decay curves of signal envelopes. The underlying implementation of compression may vary significantly between designs, depending on the envelope detection architecture, control of dynamic attenuation and other factors. In spite of its popularity, there have been only a handful of published mathematical compression models in literature [3, 4, 5]. These models, however, have not been used in the context of hearing aid research to-date, as compression effects of hearing aids are often tested directly in clinical settings. Nevertheless, the benefit of having such a model available should not be underestimated, since they can offer both new perspectives on clinical findings, as well as assist with the designs of new instruments. We would like to suggest a new top-level model that would be easier to use along with hearing aids, as well as some other audio applications. The model deals directly with signal envelopes and does not assume a specific topology. It may allow for an analytical solution to be obtained of simple input signals to the system.

2. BACKGROUND

2.1. Dynamic Range Compression

At the heart of the compressor there is an envelope detector that follows the slow changes of the input/output signal level. The compressor has a predefined attenuation for its input signal, which depends on the detected level. The exact attenuation is set according to the compression ratio, or the slope of the input-output (I/O) curve (the "compression-law"). The I/O curve is normally defined logarithmically, so that it is linear in the log representation (but non-linear otherwise). Often the compressor is linear and does not attenuate below a certain input compression threshold (kneepoint), and compressing above it.

The compressor is also characterized by two time constants, which control its speed of changing the gain, between points on the compression law function. The release time is the nominal time it takes the compressor to update its gain, when the input signal level decreases within the compressing region. Conversely, the attack time is the time that it takes to update the compressor gain, when the input level increases.

The effective long-term result of compression is a certain reduction in the dynamic range of the output signal, but otherwise minimal effect on the spectral information of the input. The various constants of the compressor determine the extent of this reduction, in addition to determining the range of the input signal that is actually being compressed (above compression threshold).

The implementation of dynamic range compression can be executed in a number of ways. Such a system has either one of several envelope detection circuit possibilities that follows the output signal (feedback) or the input signal (feedforward). The detected value is then fed-back or fed-forward to an attenuator, which adjusts the gain according to the compression-law [3]. The ample choice of compressor design possibilities results in a wide gamut of behaviors.

2.2. Existing Compression Models

Only a handful of models were found in recent literature that develop the general governing equations of compressors. Unlike particular compressor designs, those models use the compression circuit constants (i.e. ratio, threshold, release and attack times) and an arbitrary input signal to come up with the general functional representations of the output signal, according to specific operational principles.

Oliveira developed basic static and dynamic compression temporal transfer functions based on simple block diagrams, in order to determine the ideal topology of his feedforward compressor design [6]. The equations are voltage based and eventually do not provide a full picture of an arbitrary time-signal that is fed into the system and therefore do not constitute a complete model.

Floru [3] developed accurate models for feedback and feedforward RMS-detection-based linear and logarithmic compression. This detailed analysis is founded on sound electronic basis, which yields accurate results for such designs. Integral transfer functions are derived for the complete time signals. Importantly, he also derives total harmonic distortion levels as a function of frequency.

In another model, Abel and Berners [4] derived differential equations that govern RMS- and peak-detected feedback and feedforward compressors, based on simple electronic concepts. They dwell on the differences between feedback and feedforward topologies and their results agree with those of Floru.

In the last model developed by Simmer et al. [5] an algorithmic approach is introduced using difference equations, which is readily applicable on digital platforms. The authors compared their model to commercial compressors and optimized the fitting. The results are based on the specific way in which the measured compressors were designed. A similar approach was applied to model earlier the compression in the human hearing system by Glasberg and Moore in their loudness model [7].

2.3. Compression in Hearing Aids

Compression is almost ubiquitous in modern digital hearing aids. It is introduced in order to compensate for loudness-recruitment in sensorineural hearing-impairments [8], enable listening comfort and for some other reasons [1]. Its effect has been studied for many decades since its hesitant introduction in the 1930-40's and it appears to be a complicated one that is dependent on many variables (for the early-historical review, see Caraway and Carhart [9]). Multi-band compressors are also very common in hearing aids to offset the changing hearing loss depth over frequency. Compression interacts intimately with the speech – by decreasing the dynamic range of the sentence, words or even syllables themselves and in the case of multi-band compression, by reducing the spectral variation of speech. Those effects are described best by referring directly to the speech envelope [10] and the corresponding reaction of the compressor. Much research has dealt directly with the envelope of the compressed vs. uncompressed or peak-clipped speech and its perception by the hearing-impaired.

For the hearing aid professional (audiologist, ENT physician, speech-language-hearing pathologist, etc.) the exact mechanism by which the hearing aid compression is implemented is unknown and largely uninteresting. It is usually described using its I/O curve, along with the above-mentioned constants. Those are supposed to communicate universally, irrespective of the internal operational principle. For more information, see, for instance, reviews by Souza[2], Moore[11] and Kuk[12].

This rough operational understanding of compression in hearing aids is much like many of the popular software plug-in compressors used for audio mixing and mastering, which often display the I/O curve graphically, but do not really reveal the algorithms underneath it.

3. COMPRESSION MODEL

3.1. Model Rationale

Despite the existence of the above-mentioned compression models, a new one is suggested with the primary intent to facilitate the analysis of certain hearing-aid related compression behaviors.

Often, the merit of analytical models is their ability to give insight to the user, simply by looking at the general closed-form results. Current models lack in this ability, especially since the different compression parameters (e.g. ratio, threshold) cannot be always directly used in them. In the hearing aid world it is common to think in envelopes and the I/O curve is the most universal way to characterize compressors. The analytic models available [3, 4] are also not straightforward when it comes to factoring in the compression threshold in a complete signal, i.e. the fact that typical the input signal crosses the compression threshold numerous times and by that changes its subjection to it. We would like to have a model that:

1. does not assume architecture (such as feedback or feedforward)
2. does not assume detection method (RMS/peak/average). This can be computed separately beforehand
3. is formulated in a way that gives us insight with regards to its measurable parameters. Other implicit constants such as reference voltage are to be avoided
4. is translatable to the hearing aid vocabulary and can be used with raw sound pressure levels (SPL). The input and output could be used in final and not intermediate units
5. can be used for analog and digital systems alike

In the presented model, we attempt to trade off electronic and DSP rigor for higher accessibility and versatility to the average user, who either lacks the design specifications of the instrument he/she uses, or is simply uninterested in them. The following model is based on top-level principles of operation. It does not assume any actual methods of implementation, but rather describes the final output of an ideal compressor, which works according to standard definitions in the logarithmic domain. Any implementation of such model is likely going to deviate from it, simply because it does not emanate from any specific electronic or DSP architecture, but adheres to the verbal compression definition, as is often described in the hearing aid and audio production world. In this sense it is qualitative.

3.2. Static Model

The formulae that are developed below are merely models. They do not indicate the internal process inside this or that compressor, but only describe its supposed aggregate behavior. Additionally, they deal explicitly only with broadband compression and time-amplitude phenomena. Multi-band compression should be dealt with separately, although the model here may be harnessed for that analysis simply by adding the appropriate band-pass filters to it.

Any analytic signal, $s(t)$, can be decomposed into an envelope and an instantaneous phase [13]. It can written as:

$$s(t) = x(t) \cos\left[2\pi f_c t + \theta(t)\right] \tag{1}$$

where $x(t)$ is the envelope, f_c is a center frequency and $\theta(t)$ is the instantaneous phase of the signal. Compression works on the slowly varying envelope of an input signal by modifying the gain dynamically accordingly. How the envelope is extracted from the input signal is subject to the particular design. Generally, an envelope detector could be either one of the so-called RMS, peak or averaging detectors [3, 14]. In the sections below, input and output RMS/envelope levels are called for convenience input and output levels, in short.

A general compressor can be represented graphically using its I/O curve (see Figure 1). It is a variation on the gain curve in steady-state operation. When the output is equal to the input, there is no effective gain change and the gain is 0dB. This is represented by a 45° slope of $y = x$, y being the output and x the input. When there is compression in action, there is an effective gain reduction of the output, compared to this zero gain line. The attenuation increases with growing inputs, according to the compression ratio. These two states – linear (no-gain) and compression (level-dependent attenuation) – are both observable in typical compressors.

The following is a functional form of a general compressor, which can have any ratio below or above its single sharp kneepoint, as is often the case with the popular wide-dynamic-range-compression (WDRC) in hearing aids. The time dependence is only implied, as the curve shows the static steady-state of the compressor and not the transitory states:

$$y_s(x(t)) = \begin{cases} \dfrac{x(t)}{C_L} + M & x \leq K \\ \dfrac{x(t) - K}{C_H} + \dfrac{K}{C_L} + M & x \geq K \end{cases} \tag{2}$$

$x(t)$ is the time-dependent input envelope in dB. $y(t)$ is the output envelope in dB. Be it RMS, peak or average, it will have the same envelope interpretation as the input detected envelope type. Occasionally there is a so-called makeup gain, M, that is applied to the signal. It is a level- and frequency-independent gain. In the rest of the discussion we shall set the makeup gain to zero for convenience.

C_L and C_H are the compression ratios below the kneepoint and above it respectively. A value of unity designates linear response (zero gain). Values larger than unity stand for compression and values smaller than unity for expansion.

K is the sharp kneepoint, or compression threshold, in dB (peak or RMS with respect to the input and output units). The threshold need not be sharp ("soft-knee") and its softness varies between different designs. In that case, the two piecewise segments may be asymptotic lines of the continuous compression function. The compression ratio is then the local derivative of the function, rather than a constant factor. Additionally, in comlext designs, there may be multiple kneepoints. However, for the purpose of the discussion here, there is only one kneepoint that is kept sharp, for simplicity.

This piecewise function is continuous between its segments. It includes the most general case of having compression/expansion/limiting or linear response on either side of the kneepoint, without loss of generality. The knee point separates a region compressed by ratio C_L from a compressed region, which is compressed by a ratio C_H. If C_H is very high or infinity, it works like a limiter, where the output signal is absolutely bound. See figure 1 for a complementary graphical description of the various static compression constants.

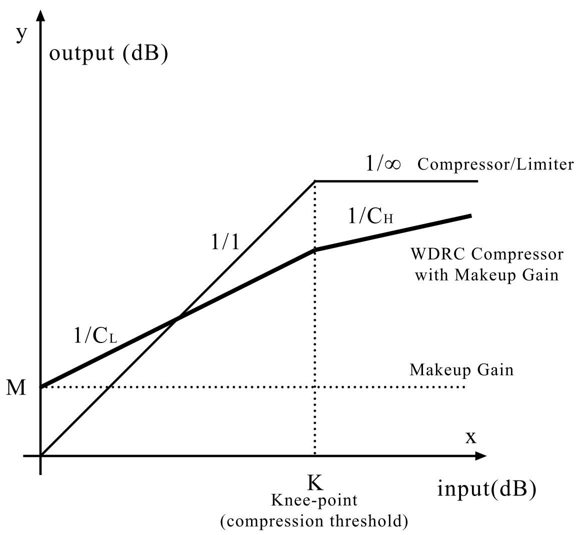

Fig. 1: Basic input output curves of two compressors: the thin line represents a compressor/limiter, which is linear below its kneepoint; The bold line represents a wide-dynamic-range compressor (WDRC) with compression ratios C_L and C_H and makeup gain M. Both have kneepoints at the same input level, but they correspond to different outputs. Various definitions of the kneepoint may actually shift it upwards on the curve, according to the amount of attenuation achieved at that point.

We can translate the static response into static gain, G_s, applied on the original signal, using:

$$G_s(x(t)) = y_s(t) - x(t) \tag{3}$$

which is either zero or negative (in dB) in case of compression. The output signal from the compressor is thus:

$$s_o(t) = s_i(t) + G_s(t) \tag{4}$$

$s_i(t)$ being the complete input signal to the compressor.

3.3. Dynamic Model, General Solution

Let us inspect the time dependency of the compressor. We already know that the direction matters, whether the input level increases or decreases with time, due to the different values of attack and release (hysteresis). This is equivalent to saying that we expect a dependency on the input first time derivative, as well as on the input itself.

In the hearing aid world, we are generally more interested in the release time, since the attack time is often very fast. However, save for the actual constants and sign differences, the attack and release behave in the same way model-wise and the full solution to both is derived commonly.

For simplicity, we shall assume a linear response below the kneepoint, so that $C_L = 1$. Hence, we shall designate $C_H = C$.

What happens exactly when the input level changes? As long as the input level is constant, its corresponding output gain is determined uniquely by the point on either one of the I/O curves in figure 1. When the input level changes, the immediate output after the change has initiated still maintains the same gain as was applied for the previous input level. At this instant, the output is either below or above the steady-state curve, unless the change occurred within the linear segment only. It would be below the curve, when the signal decreases in level. The compressor at first attenuates it too much and has to update its attenuation to be relaxed according to the current lower input level. Conversely, the new point would be above the curves, when the input signal increases in level. Then the compressor does not attenuate it enough and has to update the attenuation to be

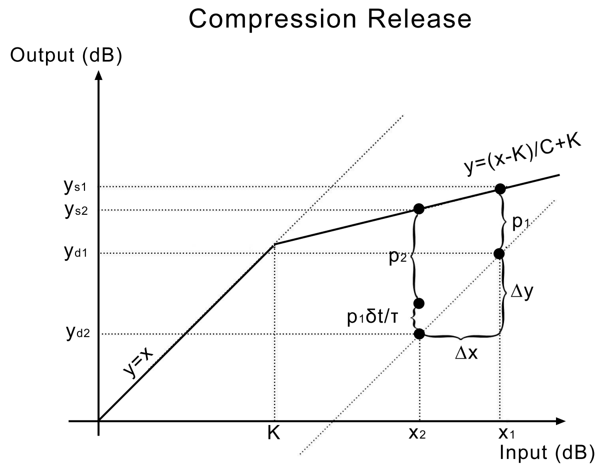

Fig. 2: Calculation of the dynamic compression using the input/output curve. During release the path p is defined as positive length.

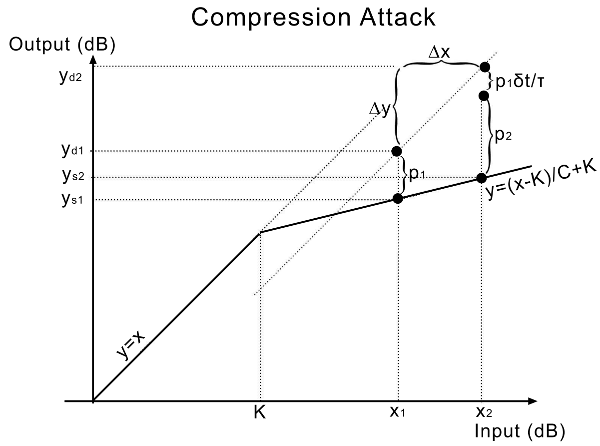

Fig. 3: Calculation of the dynamic compression using the input/output curve. During attack the path p is defined as negative length.

increased. The characteristic time for those updates is defined by the release and attack time constants. The release time determines the reaction time for input level decreases. The attack time does that for input increases.

For both types of actions, we are always interested to know the difference between the two points: the current point outside of the steady-state curve and the one on the curve, which maps the same input level to the prescribed output level. The time constants of the compressor dominate the duration of the attenuation update over that difference.

Let p be defined as the path difference in dB between the target gain and the actual gain, before the compressor reacts. In that case the updated output looks like

$$y_d(x(t), t) = y_s(x(t)) - p\,[y_d(x(t), t)] \qquad (5)$$

y_d is the dynamic output envelope and y_s is the static output as was defined in eq. 2. The difference function p is also a function of y. However, it will become obvious that p depends on the previous condition of the system.

It should be noted that there is no difference between attack and release in terms of the equations followed here. A positive p value corresponds to release and negative value to attack transitions. This is in contradiction to previous models, where the attack are release were modeled independently of one another [4, 15].

In order to find a complete expression for p and later on for y_d, we shall look at an arbitrary I/O point and the path it makes on the input-output plane on the way to a subsequent point in time. Each point has its corresponding distance from the steady-state curve. We calculate geometrically the dependence of p of one point on the second point. The difference function is scalar and relates only to the distance, but not to specific coordinates in the I/O space. This is schematically shown in figures 2 and 3 and is described by the following:

$$p_2 = p_1 + (y_{s2} - y_{s1}) - (y_{d2} - y_{d1}) - p_1\frac{\delta t}{\tau} \qquad (6)$$

where p_2 is the path difference function subsequent to p_1. y_{s2} and y_{s1} are the two static compression points that can be readily calculated as a function of the input. y_{d1} is the initial output of the first point. y_{d2} is the output immediately when the input changes. At this point the compressor still applies the same gain to the input as the previous input and thus it is a brief duration of linearity. The last term represents the actual action of the compressor to diminish the absolute difference from the target gain. It is proportional to the last recorded difference and to the duration of the transition. It is assumed to be first order correction and thus it is indirectly proportional to time constant of the compressor.

We also implied that y_{d1} and y_{d2} are linearly related, as they both follow the same instantaneous gain of the compressor. Hence we can write:

$$y_{d2} - y_{d1} = x_2 - x_1 = \Delta x \tag{7}$$

Setting $p_2 = p(t + \delta t)$, $p_1 = p(t)$, $y_{s2} = y_s(t + \delta t)$ and $y_{s1} = y_s(t)$ yields an explicit time relation of the path difference:

$$p(t + \delta t) = p(t) + [y_s(t + \delta t) - y_s(t)] - \Delta x - p(t)\frac{\delta t}{\tau} \tag{8}$$

Using the definition of the derivative of p:

$$\lim_{\delta t \to 0} \frac{p(t + \delta t) - p(t)}{\delta t} = \frac{dp(t)}{dt} = p'(t) \tag{9}$$

Rearranging the terms, we derive an ordinary differential equation of p:

$$p'(t) = \frac{dy_s}{dt} - \frac{dx}{dt} - \frac{1}{\tau}p(t) \tag{10}$$

Finally, from eq. 2 we know that y_s does not depend directly on t, but rather indirectly through its $x(t)$ dependency. This translates into:

$$\frac{dy_s}{dt} = \frac{\partial y_s}{\partial x} \cdot \frac{dx}{dt} \tag{11}$$

Using eq. 11 in eq. 10 and further rearranging we conclude with:

$$p'(t) + \frac{1}{\tau}p(t) = \frac{dx}{dt}\left(\frac{\partial y_s}{\partial x} - 1\right) \tag{12}$$

This neat ordinary differential equation has an interesting inhomogeneous term on the right. The derivative of the input signal suggests causality. When the input envelope is constant, its derivative is 0 and the compressor would be stay in a steady-state. Then, the input derivative of the static input is simply the compression ratio. In linear regions it is equal to 1 and the compressor need not work. In both cases, the inhomogeneous equation is reduced to a homogeneous one with a trivial solution.

Finally, the solution to eq. 12 can be used in eq. 5 to get the output envelope of the compressor.

We normally differentiate the attack and release decay functions. When the output signal is above the desired target of the steady-state curve, i.e. $y_s(x(t)) - y_d(x(t), t)$ is negative, it is attenuated using the attack time constant. When it is below the target curve ($y_s(x(t)) - y_d(x(t), t)$ is positive), it is de-attenuated using the release time constant. Otherwise, the equations are identical.

The interpretation of all of the above is as follows. There is an input signal $x(t)$, whose level changes with time. The compressor should bring this signal to a desired output level of $y_s(x(t))$. However, it cannot bring it to that output immediately. It constantly measures the gain difference $p(y, t)$ between the current output level and the target output level and works to decrease it. While the input changes its level with time, the compressor is busy correcting the previous output level of the last measured input. So when the output of the new input $x(t + \delta t)$ has been recorded, the compressor has just corrected the previous level $y_d(t)$ by a small amount. This amount is indirectly proportional to the compressor time constant and proportional to the last recorded p. We shall require that for any solution of p at $t \gg \tau_{A,R}$ eq. 5 converges to eq. 2, given that the input level has been constant for a while.

3.4. Dynamic Model, A Particular Solution

The equations above may become quite complicated for arbitrary inputs as the system is non-linear (given the affine steady state function with input-dependent time derivative and the direction dependent time constants). However, we can find a closed solution for a simple input signal. Much like the standard methods of measuring attack and release times in hearing aids [16, 17], we assume a constant input level at $t < 0$ and right after the transition, i.e. a step function of the input level.

$$x(t) = \begin{cases} x_1 & t < 0 \\ x_2 & t > 0 \end{cases} \tag{13}$$

We rewrite it using the standard Heaviside step function $H(t)$, which is defined to be 0 when its argument is negative and unity when it is positive:

$$x(t) = (x_2 - x_1)H(t) + x_1 \tag{14}$$

Next we shall use the Heaviside definition to find the time derivative of the input [18]:

$$\frac{dx}{dt} = (x_2 - x_1)\delta(t) = \Delta x \delta(t) \tag{15}$$

Where $\delta(t)$ is Dirac delta function.

The hard kneepoint in eq. 2 makes y_s non-differentiable in K. However, we can find expressions for inputs below and above it:

$$\frac{\partial y_s}{\partial x} = \begin{cases} 1 & x(t) < K \\ \frac{1}{C} & x(t) > K \end{cases} \tag{16}$$

We know from eq. 12 that our equation becomes trivial for the region below the kneepoint and we shall focus on the region above it.

Using eq. 12, 15 and 16 we shall write the particular equation of our step input:

$$p'(t) + \frac{1}{\tau}p(t) = \Delta x \left(\frac{1}{C} - 1\right) \cdot \delta(t) \tag{17}$$

This equation can be solved using the Laplace transform, by applying it on both sides [18]. We set:

$$\mathfrak{L}\left[p(t)\right] = P(s) \quad s = \frac{1}{t} \tag{18}$$

and eq. 17 becomes:

$$\mathfrak{L}\left[p'(t) + \frac{1}{\tau}p(t)\right] = \mathfrak{L}\left[\Delta x \left(\frac{1}{C} - 1\right) \cdot \delta(t)\right] =$$
$$= sP(s) - p(0) + \frac{1}{\tau}P(s) = \Delta x \left(\frac{1}{C} - 1\right) \tag{19}$$

Rearranging eq. 19 and applying the inverse Laplace transform to find the exact solution to $p(t)$ results in:

$$p(t) = \mathfrak{L}^{-1}\left[P(s)\right] = \mathfrak{L}^{-1}\left[\frac{\Delta x \left(\frac{1}{C} - 1\right) + p(0)}{s + \frac{1}{\tau}}\right] =$$
$$= \left[\left(\frac{1}{C} - 1\right) + p(0)\right] e^{-t/\tau} H(t) \tag{20}$$

Where we used the Heaviside step function $H(t)$ once again. Finally we set the initial conditions, $p(0^-) = 0$, slightly before the step ensues. Determining the relevant input step depends on its relation to the kneepoint and whether it is an attack or release decay at hand. We shall use eq. 20 in eq. 5 with the various possibilities of the step relative to the kneepoint, given in eq. 2:

$$y_d(t < 0) = y_s(x_1)$$
$$y_d(t > 0) =$$
$$= \begin{cases} \left(\frac{x_2 - K}{C} + K\right) - (x_2 - x_1)\left(\frac{1}{C} - 1\right)e^{-t/\tau_R} & K < x_2 < x_1 \\ x_2 - (K - x_1)\left(\frac{1}{C} - 1\right)e^{-t/\tau_R} & x_2 < K < x_1 \\ \left(\frac{x_2 - K}{C} + K\right) - (x_2 - x_1)\left(\frac{1}{C} - 1\right)e^{-t/\tau_A} & K < x_1 < x_2 \\ \left(\frac{x_2 - K}{C} + K\right) - (x_2 - K)\left(\frac{1}{C} - 1\right)e^{-t/\tau_A} & x_1 < K < x_2 \end{cases} \tag{21}$$

We can see from the result that the compressor output target equals to the input and after the attack or release decays have elapsed, the second term would effectively vanish and the dynamic output converges to the static output (first term), as required.

3.5. Hearing Aids Applicability of the Model

The following notes are observations made on real compressors on the market with a special emphasis on hearing aids.

In practice, attack and release time constants may have to be scaled, depending on the exact definition of attack and release that is being used. As the standard hearing aid definitions are inconsistent [16, 17], different scaling factors may be used. Decay functions are normally given time constants according to their relative decay compared to their initial state (for example, in reverberation time or capacitor discharge decay). However, complying with both standards, which require an absolute value of the decay p, such as 2 or $4dB$ below the steady state level, time constants as are expressed in the above equations depend on the kneepoint and on the compression ratio, in addition to the predefined start and end levels.

A comment should be made also about the decay functional form in compressors. Two decay shapes are observed in technical audiological literature about compression: linear and exponential.

Physical decay phenomena are exponential and audio compression-laws are predominantly logarithmic. However, the type of decay seen is inconsistent in the descriptions and is generally not discussed. In the commercial compressors we show here (see 4.1 below) – both standalone and in hearing aids and FM systems for hearing aids – we observed both types of decays. Floru [3] discusses a linear and logarithmic implementation of the detection and gain control circuitry. He derives an expression for the transient decay of a step input, which is non-linear. When converted to decibels, however, the expression becomes of the form $A\log(1+Be^{-t/\tau})$, which is still non-linear. Therefore, a linear decay in compressors measured may suggest a completely controlled decay by a computer program, which works in the logarithmic domain. This range of functionalities affects the measured values of attack and release transient decays and comparison between the two kinds, even if their nominal time constants are identical, is not straightforward. We do not know what is the perceived difference between the two kinds of decay.

4. SIMULATIONS AND MEASUREMENTS

Using a few arbitrary real compressors, we would like to see if our model yields overlapping responses to them, directly employing realistic values to its parameters. In this section the basic equations were tested: The dynamic solution of eq. 21 along with the static compression law approximated by eq. 2 and the governing differential equation, eq. 12 in its discrete form.

4.1. Hearing Aid and FM Compressors

The compressor was tested against two arbitrary hearing aids and an assistive-listening-device (ALD) frequency modulation radio system (FM) for hearing impaired.

The input-output curve was measured in a small anechoic chamber, B&K 4232 Anechoic Test Box. A 1/2" reference condenser microphone, B&K 4192 with a B&K 2669 preamplifier, was placed in front of the chamber loudspeaker. The loudspeaker was driven through a power amplifier, Rotel RB-991, which was fed by the signal generator of an audio analyzer, B&K 2012 that was also connected to the microphone. Using the reference microphone, the loudspeaker was set to give a known SPL at the reference point. Either the FM transmitter microphone or the HA microphone was placed on this point. The output from the hearing aid receiver was connected to the same preamplifier, but via an IEC 711 coupler, B&K 4157 Ear Simulator. Using the calibrated output of the amplifier and coupler microphone, the I/O curves of the hearing aid could be measured accurately by the audio analyzer sine sweep. Sweeps were performed at 1 kHz with 1dB steps.

The FM system I/O curve was measured by supplying its receiver from a battery directly via its DAI male jack ("Euro Pin"), loaded by a standard 4.3kΩ resistor and tapped for its audio output. The output voltage was recorded by the direct channel of the audio analyzer.

The step recording setup was very similar to the one above, except for the replacement of the audio analyzer by a soundcard. Digidesign's Digi Pre, Digi 192 soundcard hardware and Pro Tools HD software were used to generate the sine steps. The output from the coupler was powered by a B&K 5935L Dual Mic Supply and recorded on a mono channel of the software and later analyzed in Matlab.

Figures 4-6 show the release measurements of a 1kHz step between 100 and 60dB SPL. The iterative process was done using the I/O curve of the instruments, from which the kneepoints and compression ratios could be extracted. An example for this curve in conjunction with the release curve of the FM system is given in figure 4. Fitting the I/O accurately allowed, in this case, to reproduce also the dynamic release curve, provided that the time constant is added (iterated). A makeup gain term that was added to the static compression equation includes also the system microphone sensitivity, amplification and conversion from dB SPL input to dBV output.

Figure 5 shows an example for a hearing aid with an exponential decay. In this case, the measurement was dirty and the release undershoot is modulated by noise. However, the trend is visible. The model parameters were also extracted from the I/O curves and the time constant iterated. Here, the system has compression ratio below the kneepoint (WDRC) and so the static equation used is the most general one.

Figure 6 shows a limitation of the model, when compared to a linear release decay. The exponential de-

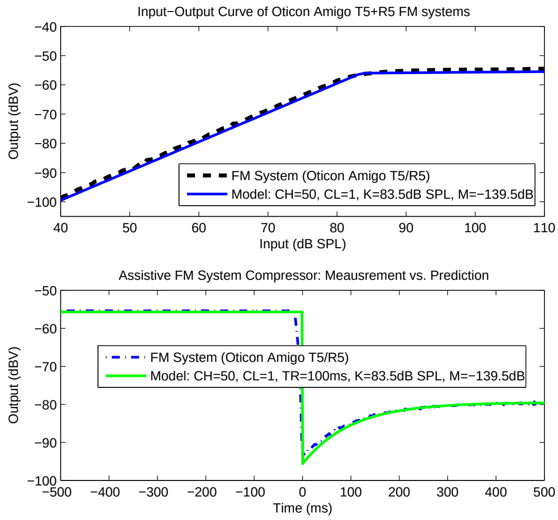

Fig. 4: Static curve and model of Amigo T5-R5 FM system and release decay measured with a 100 to 60 dB SPL step, according to the IEC hearing aid standard[17].

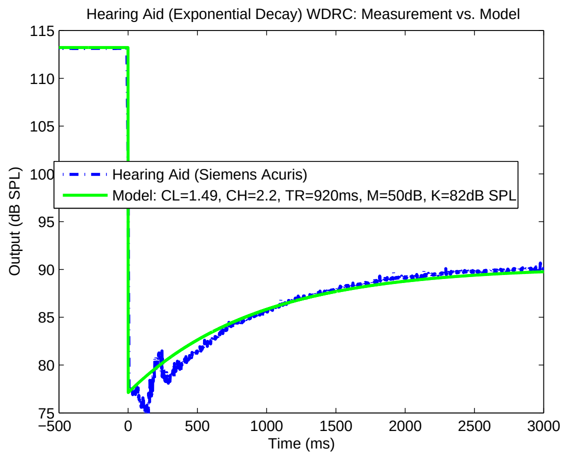

Fig. 5: Siemens Acuris hearing aid release decay measured with a 100 to 60 dB SPL step, according to IEC[17]. Although measurement appears noisy, the trend is clear

cay obviously does not compare too well to it and the usage of the model is restricted to qualitative or boundary effects.

4.2. Comparison to a Plug-In Compressor

In order to examine further the model behavior, it is also compared with a commercial plug-in hi-fidelity compressor (Waves C1), which was run internally on a Pro Tools LE virtual studio environment software (by Digidesign). This kind of plug-in may be used extensively by, for instance, sound engineers for the production and mixing of music and films.

The setup included an internal sine signal generator, which was shaped according to the measurement specifications. This signal was the input to the compressor. The commercial compressor used included a parametric setting for attack, release, ratio, kneepoint and makeup gain. The makeup gain was kept on zero. The knee is soft on this compressor and therefore not so well-defined and its exact level had to be reiterated for best fit. Other nominal parametric values had to be scaled a bit to obtain a better match. Nevertheless, fitting was done manually and only crudely, since the purpose of the comparison is

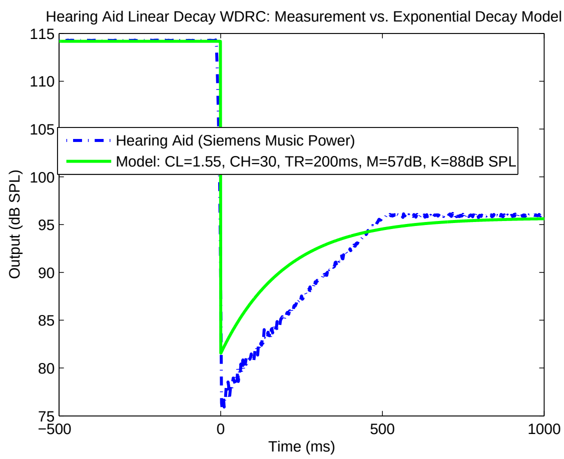

Fig. 6: Siemens Music Power hearing aid release decay measured with a 100 to 60 dB SPL step, according to IEC[17]. The decay in this hearing aid is linear and not exponential as before and thus matching the model to it is incompatible, in terms of the recovery speed and the undershoot of the model vs. the aid

largely qualitative and does not pertain to this or that specific compressor algorithm.

The output was recorded internally on a virtual track. The 44.1 kHz, 24-bit WAV output recordings were exported to Matlab, where their RMS levels were calculated and visually compared.

4.2.1. Step Response

The step response was set to vary between -5dBFS and -35dBFS (peak values). We recall that the model is ignorant to the type of signal detected (peak or RMS), so in order to obtain a correct result both model and recording have to be the same. Since the step is a simple sine, we know that the difference here would be a -3.01dB correction of the peak-to-RMS value (the crest factor of a pure sine).

In figures 7-10 all attack and release time constants were scaled to be 80% of their nominal value. Decibels on the ordinate represent peak values. The nominal kneepoint values are in fact 0.8-1.7 dB compression points, already on the compression curves. Those values may also reflect peak detection (rather than RMS), but we cannot state that with certainty yet. So the actual mathematical knee was iteratively extracted, using eq. 2 and 3.

All eight compressors show a very good fit, given the above tweaks. The very short 10ms release shows an increased undershoot compared to the measurement. The release time prediction is a little slower than the measured one. However, it appears that the very short release time is still modeled well, opposite to results obtain with another model [5]. Additionally, in the particular compressor tested here, the attack decay indeed follows the same exponential rule, contrary to a linear decay (in the log domain) suggested earlier [4]. That is mention not in order to show that one model is more correct than the other, but rather to underline the many possibilities of implementing compression.

4.2.2. Speech Response

A final qualitative comparison and test was made using a speech sample as input to the discrete version of the governing differential equation of the compressor, which becomes a difference equation with a

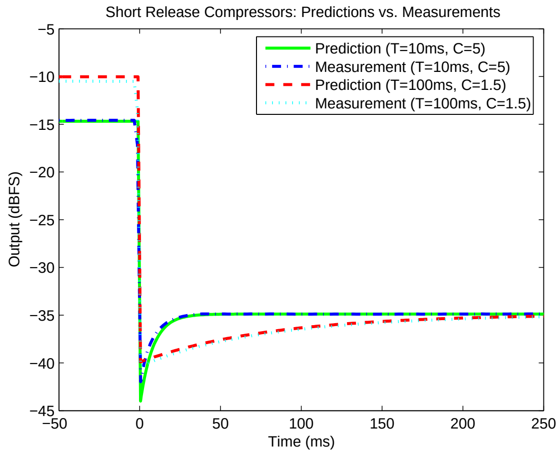

Fig. 7: Dynamic behavior of a single compressor: measurement on a commercial plug-in vs. prediction according to the model

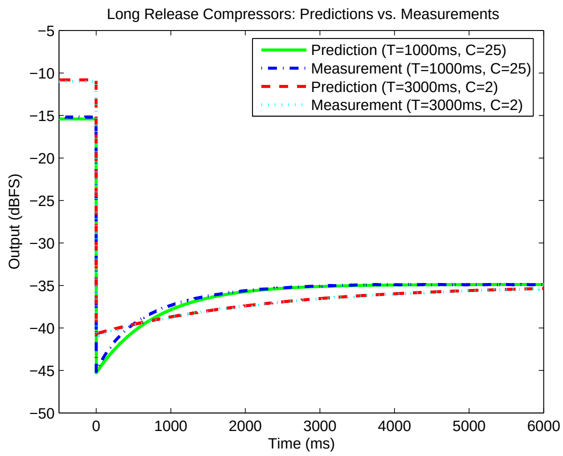

Fig. 8: Dynamic behavior of a single compressor: measurement on a commercial plug-in vs. prediction according to the model

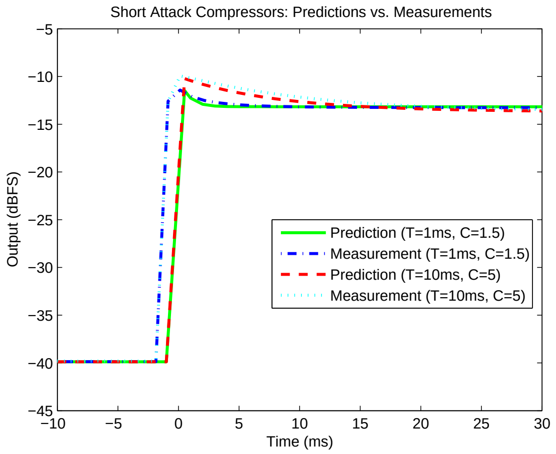

Fig. 9: Dynamic behavior of a single compressor: measurement on a commercial plug-in vs. prediction according to the model

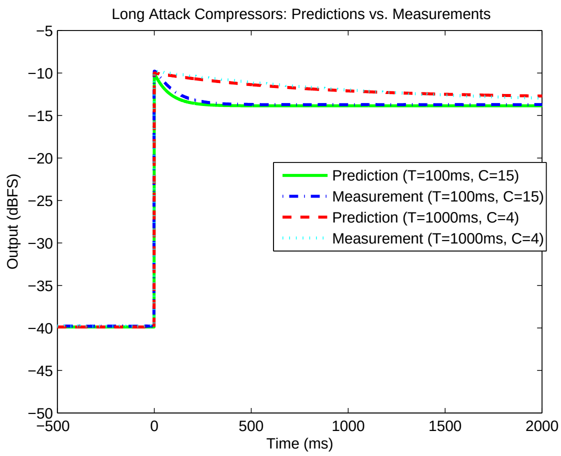

Fig. 10: Dynamic behavior of a single compressor: measurement on a commercial plug-in vs. prediction according to the model

hard-knee compression law:

$$p(n) = \frac{p(n-1) + [x(n) - x(n-1)]\left(\frac{1}{C(n)} - 1\right)}{1 + \frac{1}{\tau_n}} \tag{22}$$

where n designates a sample, $C(n)$ is the instantaneous compression ratio and τ_n is the equivalent attack/release time in samples. In figure 11 the compression effect is shown side by side the same Waves commercial soft-knee compressor. Signal amplitudes are compared as well as the instantaneous gains of the compressors, according to different envelope detection techniques. Preliminary simulations showed the best fit with the peak detection (no dynamic correction is needed for the nominal kneepoint). The RMS and averaging detectors needed a correction for the nominal kneepoint by -7dB. As -20dB appeared to be more accurate for peaks, 7dB is the approximate peak-to-RMS value (crest factor) of this speech sample. None of the predictions is completely similar to the commercial compressor that was potentially combined with a smoothing (or decimation) algorithm and an RMS evaluator. Additionally, the different knee implementation – soft vs. hard – also contributes to the variation.

Previous comparisons [5] showed perhaps a better fit, but did not include a direct gain comparison and they included smoothing algorithms, which are not implemented here. However, we get a reasonable fit to the commercial compression, which is enough to indicate that the expression we got is indeed able to correspond to versatile dynamic inputs and is not limited to step functions.

5. CONCLUSIONS

A model for compression is introduced, which was derived using geometric considerations and not according to a specific compressor topology. The model uses envelopes as input and output variables and thus skips the level detector element, which may assume an arbitrary detection method. Spectral effects of compression are not factored in directly. An emphasis was given to have all model parameters reflect those that are displayed in common specification sheets and parametric user interfaces of various commercial compressors.

Modeling was divided into static and dynamic solutions. The static solution is a repetition of known

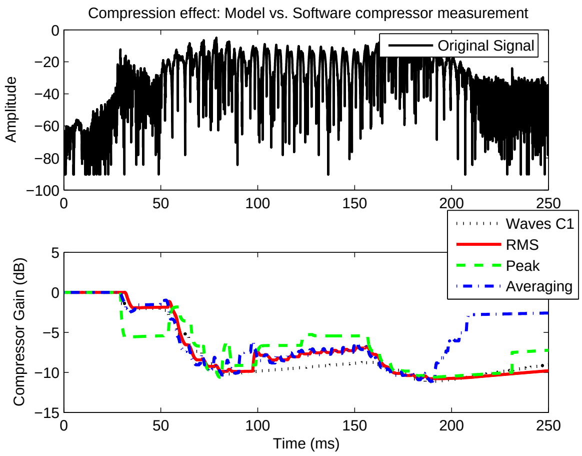

Fig. 11: Dynamic behavior of a single compressor: measurement on a commercial plug-in vs. predictions according to the discrete version of the differential equation of the model and various primitive envelope detectors. The parameters were: C=3, $\tau_R = 500ms$, $\tau_A = 1ms$, K=-20dBFS. The sample contains English male saying: "Come to". RMS detection was done with a 2-pole, 50Hz, Butterworth lowpass filter. Peak detection was done with no decimation using Peakdet[19], set to a tolerance of 0.15. Averaging detection was done using a moving average of 200 samples.

compressor properties. The dynamic treatment converges into an ordinary differential equation that is both visibly more intuitive and mathematically easier to deal with analytically. The particular response for a step input is then calculated accurately, as an example, which results in a typical exponential decay function.

The general and particular model results are all tested against real compressors of hearing aids and an FM system for hearing aids and a plug-in effect used for audio mixing. The model with the nominal compression parameters shows a satisfactory convergence to the measurements with only minor tweaking.

The model is limited, though, to exponential decays and has to be used with the adequate signal levels, representative of the omitted signal detector section. Moreover, the model does not account directly for multi-band compression, which is common in hearing aids. Effects resulting from the band-limited signals should be computed separately using the adequate filters in combination with the otherwise broadband compression model.

6. ACKNOWLEDGMENTS

The author would like to thank Peter Anker Nielsen in Oticon A/S for supporting this research. Special thanks to Karsten Bo Rasmussen for his comments.

7. REFERENCES

[1] H. Dillon, *Hearing Aids*, Thieme, New-York Stuttgart, Boomerang Press, Sydney, 2001

[2] P. E. Souza, "Effects of Compression on Speech Acoustics, Intelligibility and Sound Quality", *Trends in Amplification*, vol. 6, no. 4, pp. 131–165, 2002

[3] F. Floru, "Attack and Release Time Constants in RMS-Based Feedback Compressors", *J. Audio Eng. Soc.*, vol. 47, no. 10, pp. 788–804, Oct 1999

[4] J. S. Abel and D. P. Berners, "On Peak-Detecting and RMS Feedback and Feedforward Compressors", *presented at the AES 115th Convention, New York, NY, USA*, Paper 5914, 2003

[5] U. Simmer, D. Schmidt and J. Bitzer, “Parameter Estimation of Dynamic Range Compressors: Models, Procedures and Test Signals”, *presented at the AES 120th Convention, Paris, France* Convention Paper 6849, 2006

[6] A. J. Oliveira, “Digital Dynamic Range Compressor for Audio”, *J. Audio Eng. Soc.*, vol. 37, no. 4, pp. 226–240, April, 1989

[7] B. R. Glasberg and B. C. J. Moore, “A Model of Loudness Applicable to Time-Varying Sounds”, *J. Audio Eng. Soc.*, vol. 50, 331–342, May, 2002

[8] E. P. Fowler, “A Method for Early Detection of Oto-Sclerosis”, *Archives of Otolaryngology*, vol. 24, no. 6, pp. 731–741, 1937

[9] B. J. Caraway and R. Carhart, ”Influence of Compressor Action on Speech Intelligibility”, *J. Acous. Soc. Am.*, vol. 41, no. 6, pp. 1424–1433, 1967

[10] S. Rosen, “Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects”, Philos. Trans. R. Soc. Lond., B, Biol. Sci., vol. 336 no. 1278, pp. 367-373, 1992.

[11] B. C. J. Moore, “The Choice of Compression Speed in Hearing Aids – Theoretical and Practical Considerations and the Role of Individual Differences”, *Trends in Amplification*, vol. 12, no.2, pp. 103–112, Jun, 2008

[12] F. K. Kuk, “Theoretical and Practical Considerations in Compression Hearing Aids”, *Trends in Amplification*, vol. 1, pp. 5–39, 1996.

[13] J. Proakis and D. Manolakis, *Digital Signal Processing: Principles, Algorithms and Applications*, 3rd edition, 1996, Pearson Education Inc., Singapore

[14] R. B. Randall, *Frequency Analysis*, pp. 87–100, Brĕl & Kjær, 3rd edition, Nærum, Denmark, 1987

[15] E. F. Stikvoort, “Digital Dynamic Range Compressor for Audio”, *J. Audio Eng. Soc.* vol. 34, no. 1/2, pp. 3–9, Jan-Feb. 1986

[16] American National Standards Institute (ANSI), “Specification of Hearing Aid Characteristics ANSI 3.22-2003”, *American National Standard*, 2003

[17] International Electrotechnical Commission (IEC), “International Standard - Hearing Aids with Automatic Gain Control Circuits”, *International Electrotechnical Commission 118-2*, 2nd Edition, 1983

[18] S. Mauch, *Introduction to Methods of Applied Mathematics or Advanced Mathematical Methods for Scientists and Engineers*, *http://www.cacr.caltech.edu/∼sean/applied_math.pdf*, pp. 904, 2001

[19] Billauer E., ”Peakdet: Peak Detection Using MATLAB”, *http://www.billauer.co.il/peakdet.html*, 2008