SIGNAL-MATCHED POWER-COMPLEMENTARY CROSS-FADING AND DRY-WET MIXING

Marco Fink, Martin Holters, Udo Zölzer*

Department of Signal Processing and Communications,
Helmut-Schmidt University Hamburg
Hamburg, Germany
marco.fink@hsu-hh.de

ABSTRACT

The blending of audio signals, called cross-fading, is a very common task in audio signal processing. Therefore, digital audio workstations offer several fading curves to select from. The choice of the fading curve typically depends on the signal characteristics and is supposed to result in a mixed signal featuring power and loudness close to the input signals. This work derives a correlation-based design of the fading curves to achieve exact power consistency to avoid audible fluctuations of the signal's loudness. This principle is extended to the problem of mixing original signals with effect-processed signals using the dry-wet balance. Weighting coefficients for dry and wet signals are derived which realize the desired dry-wet balance but maintain the signal power.

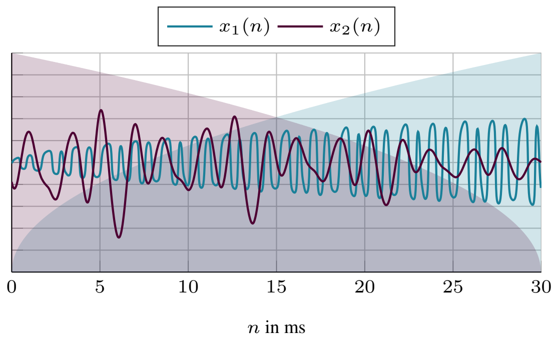


Figure 1: *Exemplary cross-fading of two signals using square-root fading curves*

1. INTRODUCTION

The convenience and affordability of todays audio processing tools allows almost any user to work on audio data, especially digitally. Digital audio workstations (DAWs) offer manifold possibilities for recording, mixing, arranging, adding effects, mastering, etc. One of the essential tools within any DAW is the cross-fader which allows to smoothly blend between separate pieces of audio [1]. A cross-fade is realized by multiplying the signals with fading curves, as illustrated in Fig. 1, and summing the weighted signals. Three main applications for the utilization of cross-fading can be emphasized:

1. Transition between songs:
   Many digital audio players allow to cross-fade between the end of the current song and the beginning of the next song to achieve smooth transitions and continuous sound. Quite long transition times of several seconds are applied in general.
2. Blend between different tracks or takes:
   Oftentimes cross-fading is applied to switch between different takes of a recording to combine the best parts of different performances of the instrumentalist. Typical cross-fade times for this application are in the range of $10-30$ ms [2]. The transition between different recorded tracks or instruments within a song can also be realized using cross-fades, especially in electronic music.
3. Fading between sources in DJ-ing:
   Many DJs artistically recompose elements from recordings to create new songs. This approach requires cross-fading to blend between sources and beats. In other words, turntables and mixers can even be considered as the DJs actual instrument and underlie a continuous evolution [3].

Another application of cross-fading is to blend intact and synthesized concealment audio streams in the context of packet loss concealment [4, 5]. All mentioned applications share the requirement of preserving the signal's loudness while cross-fading. It is well-known that the cross-faded signal power depends on the correlation of the signals to be mixed. Uncorrelated signals are cross-faded using the so-called *Equal Power Crossfade* featuring a -3 dB weighting in the transition center. In contrast, correlated signals are supposed to be linearly cross-faded corresponding to a weighting of -6 dB [2, 3]. However, these rules of thumb are very practically motivated and not perfectly accurate. Therefore, this work derives fading curves analytically based on the exactly measured correlation of the signals to be cross-faded to achieve perfect power preservation. The perfectly power-complementary fading curves are derived in Sect. 2 based on a simple signal model. The power preservation property of the proposed design is validated in Sect. 3. Additionally, Sect. 4 demonstrates that the proposed fading design can also be utilized in the context of dry-wet mixing which shares several properties with the problem of cross-fading. The work is concluded in Sect. 5.

2. CROSS-FADING

The cross-fading of two signals $x_1(n)$ and $x_2(n)$ of a certain length N can be described as

$$x_{\text{mix}}(n) = w(n)\,x_1(n) + w_i(n)\,x_2(n), \quad n \in [0, \ldots, N-1], \tag{1}$$

* The author is now with Ableton and can be contacted using marco.fink@ableton.com

where $w(n)$ and $w_i(n)$ are the fading and the inverse fading curve. Assuming signals with the same power and hence same signal variance

$$\sigma_x^2 = \mathrm{E}\left[x_1^2\right] = \mathrm{E}\left[x_2^2\right] \tag{2}$$

and identical mean value

$$\mu_x = E\left[x_1\right] = E\left[x_2\right] = 0 \tag{3}$$

allows one to estimate the variance of the faded signal

$$\begin{aligned}
\mathrm{E}\left[x_{\mathrm{mix}}^2\right] &= \mathrm{E}\left[(w\, x_1 + w_i\, x_2)^2\right] && (4)\\
&= \mathrm{E}\left[w^2 x_1^2 + 2\, w\, w_i\, x_1\, x_2 + w_i^2 x_2^2\right] && (5)\\
&= w^2 \mathrm{E}\left[x_1^2\right] + 2\, w\, w_i\, \mathrm{E}\left[x_1 x_2\right] + w_i^2 \mathrm{E}\left[x_2^2\right] && (6)\\
&= w^2 \sigma_x^2 + 2\, w\, w_i\, \mathrm{Cov}(x_1, x_2) + w_i^2 \sigma_x^2. && (7)
\end{aligned}$$

Expressing the covariance as the product of the signal variances and the correlation coefficient $\mathrm{Cov}(x_1, x_2) = r_{x_1,x_2}\, \sigma_x^2$ yields

$$\begin{aligned}
\mathrm{E}\left[x_{\mathrm{mix}}^2\right] &= w^2 \sigma_x^2 + 2\, w\, w_i\, r_{x_1,x_2} \sigma_x^2 + w_i^2 \sigma_x^2 && (8)\\
&= \sigma_x^2 (w^2 + 2\, w\, w_i\, r_{x_1,x_2} + w_i^2). && (9)
\end{aligned}$$

It is typically desirable that the cross-faded signal features the same power and hence variance as the input signals

$$\mathrm{E}\left[x_{\mathrm{mix}}^2\right] = \sigma_x^2 (w^2 + 2\, w\, w_i\, r_{x_1,x_2} + w_i^2) = \sigma_x^2. \tag{10}$$

Canceling the variance holds

$$w^2 + 2\, w\, w_i\, r_{x_1,x_2} + w_i^2 = 1. \tag{11}$$

In the following, the amplitude ratio of w and w_i shall be described with the function depending on the normalized time index $\alpha = \frac{n}{N-1}$

$$f(\alpha) = \frac{w(\alpha)}{w_i(\alpha)}. \tag{12}$$

Rewriting Eq. (12) to $w(\alpha) = f(\alpha)\, w_i(\alpha)$ and inserting it into Eq. (11) holds

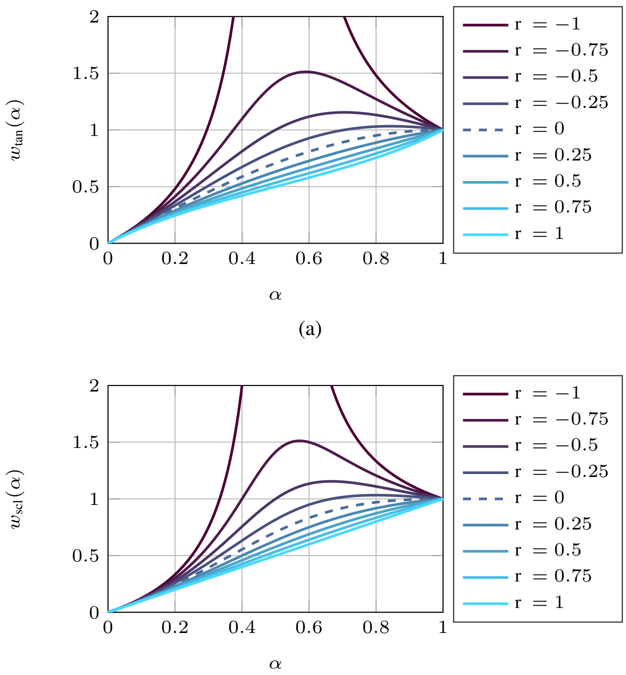
$$w_i^2(\alpha) = \frac{1}{1 + 2\, f(\alpha)\, r_{x_1,x_2} + f^2(\alpha)}. \tag{13}$$

Multiple assumptions concerning $f(\alpha)$ can be made:

1. $f(0) = 0$ since the fading curve w starts with an amplitude of 0
2. $f(1) = \infty$ since the inverse fading curve w_i ends with an amplitude of 0
3. $f(\alpha) = \frac{1}{f(1-\alpha)}$ since the fading curves w and w_i are symmetric and hence feature the same amplitude in the center of the curve $f(0.5) = 1$.

Several functions fulfill these requirements. In the following, the tangent function $f_{\tan}(\alpha) = \tan(\frac{\pi\,\alpha}{2})$ and the function $f_{\mathrm{scl}}(\alpha) = \frac{\alpha}{1-\alpha}$ are utilized and applied in Eq. (13) to hold the inverse fading curves

$$\begin{aligned}
w_{i,\tan}(\alpha) &= \frac{1}{\sqrt{1 + 2\, \tan(\frac{\pi\,\alpha}{2})\, r_{x_1,x_2} + \tan^2(\frac{\pi\,\alpha}{2})}} && (14)\\
&= \frac{\cos(\frac{\pi\,\alpha}{2})}{\sqrt{1 + 2\, r_{x_1,x_2}\, \sin(\frac{\pi\,\alpha}{2})\, \cos(\frac{\pi\,\alpha}{2})}}
\end{aligned}$$



(a)

(b)

Figure 2: *Power-complementary cross-fading curves for different correlation coefficients*

and

$$\begin{aligned}
w_{i,\mathrm{scl}}(\alpha) &= \frac{1}{\sqrt{1 + 2\, \frac{\alpha}{1-\alpha}\, r_{x_1,x_2} + \frac{\alpha^2}{(1-\alpha)^2}}} && (15)\\
&= \frac{1-\alpha}{\sqrt{1 - 2\,(1 - r_{x_1,x_2})\, \alpha\, (1-\alpha)}}.
\end{aligned}$$

Correspondingly, the fading curves are derived as

$$w_{\tan}(\alpha) = \frac{\sin(\frac{\pi\,\alpha}{2})}{\sqrt{1 + 2\, r_{x_1,x_2}\, \sin(\frac{\pi\,\alpha}{2})\, \cos(\frac{\pi\,\alpha}{2})}} \tag{16}$$

and

$$w_{\mathrm{scl}}(\alpha) = \frac{\alpha}{\sqrt{1 - 2\,(1 - r_{x_1,x_2})\, \alpha\, (1-\alpha)}}. \tag{17}$$

From Eq. (16,17) follows that perfectly power-complementary fading curves can be analytically designed whenever the correlation coefficient

$$r_{x_1,x_2} = \frac{\sum\limits_{n=0}^{N-1} (x_1 - \mu_{x_1})(x_2 - \mu_{x_2})}{\sqrt{\sum\limits_{n=0}^{N-1} (x_1 - \mu_{x_1})^2}\, \sqrt{\sum\limits_{n=0}^{N-1} (x_2 - \mu_{x_2})^2}}. \tag{18}$$

is computed beforehand. The fading curves are plotted in Fig. 2 for different correlation coefficients. Apparently, the $w_{\tan}$ curve evolves from a sine curve to a slightly S-shaped curve for decreasing correlation values. In contrast, the w_{scl} changes from non-symmetric S-shaped curve to the linear curve. It should also be noted that the proposed cross-fading curve design also works flawlessly for negatively correlated signals. Certainly the amplitude

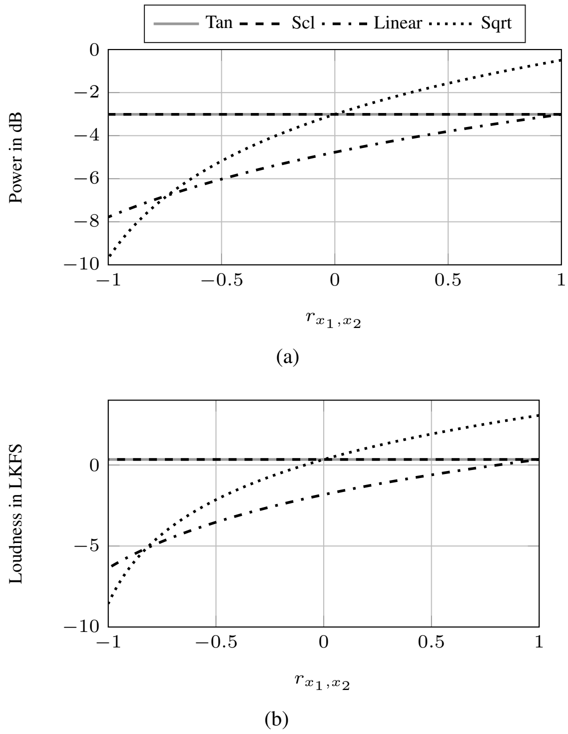


Figure 3: *Power and loudness according ITU BS-1770 of cross-faded sinusoids featuring different correlation coefficients* r_{x_1,x_2}

of the cross-fading curves increases to compensate the power loss caused by destructive interference. Cross-fading completely negatively correlated ($r_{x_1,x_2} = -1$) remains an undesirable scenario since infinite amplification in the transition center is required. However, cross-fading amplitude-inverted signals is improbable in practice anyways. Timbral coloration caused by cancellations of specific signal components remains a typical problem [1].

3. EVALUATION

Two simple experiments shall prove the effectiveness of the proposed fading curves. Two sinusoids of 1 s length featuring the same frequency ω_0 but different phase offsets ϕ

$$x_1(n) = \sin(\omega_0\, n) \tag{19}$$

$$x_2(n) = \sin(\omega_0\, n + \phi), \quad \phi \in [0, \ldots, 2\pi] \tag{20}$$

are cross-faded using the linear curve $w_{\text{lin}}(\alpha) = \alpha$, the square-root curve $w_{\text{sqrt}}(\alpha) = \sqrt{\alpha}$, and the curves from Eq. (16,17). The power of the cross-faded signals is plotted against the correlation coefficient in Fig. 3a). Using the proposed curves yields constant power for all values of r_{x_1,x_2} whereas the linear and square curve produce varying power progressions. The linear and square-root cross-fading is solely power-complementary for fully correlated ($r_{x_1,x_2} = 1$) and uncorrelated signals ($r_{x_1,x_2} = 0$), respectively. In addition to the signal power, the loudness according to ITU BS-1770 [6] was measured and illustrated in Fig. 3b). Basically, the loudness runs the same trend as the power.

In the following, the experiment shall be repeated using a real-world signal. An excerpt from the *Organ Handel* sample of length 1 s from the SQAM database [7] is processed with the mono-to-stereo conversion from [8] resulting in two decorrelated signals. The degree of decorrelation depends on the stereo width parameter and processing band width. The processing band is set from 50 Hz to 16 kHz whereas the stereo width parameter is varied from 0 to 1. Figure 4 illustrates the power and loudness of the cross-faded mono-to-stereo converted signals over the corresponding correlation coefficient. In contrast to the first experiment, the power varies slightly for the proposed fading curves. However, the loudness remains almost constant. Like beforehand, the application of the linear and square-root curves yield correlation-dependent varying results with deviations > 1 dB.

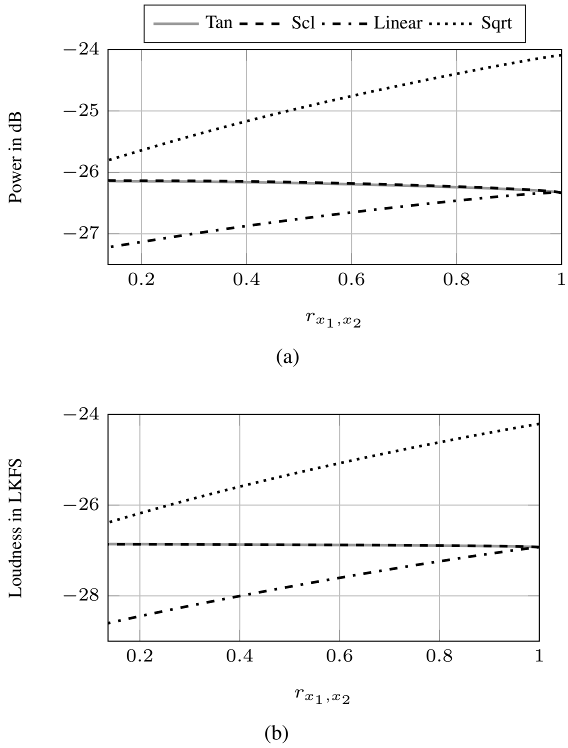


Figure 4: *Power and loudness according to ITU BS-1770 of cross-faded mono-to-stereo converted signals featuring different correlation coefficients* r_{x_1,x_2}

Both experiments validate the effectiveness of this novel approach of power-complementary cross-fading. Fortunately, it is also useful in the context of dry-wet mixing as shown in the following.

4. DRY-WET MIXING

Involving audio effects to shape the sound of a single or multiple instruments is a key task of musicians, audio engineers and producers. Many effect units offer a so-called dry-wet balance knob defining the proportion of unprocessed (dry) and processed (wet) signal in the output. Hence, the resulting mix can be described as a weighted sum

$$x_{\text{mix}}(n) = g_d\, x_d(n) + g_w\, x_w(n), \tag{21}$$

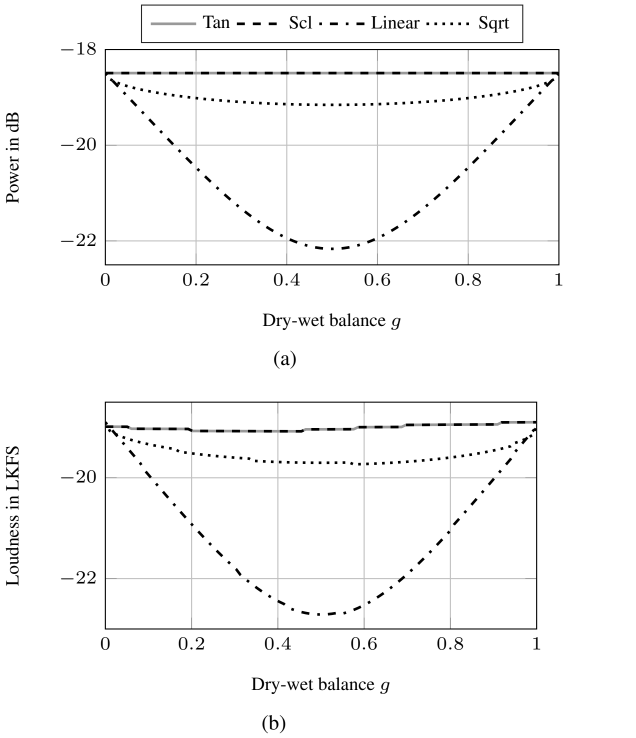


Figure 5: *Power and loudness according to ITU BS-1770 of reverberated and original signal mix using different dry-wet balances*

of dry signal $x_d(n)$ and wet signal $x_w(n)$. The relationship of the weighting coefficients g_d and g_w defines the dry-wet balance g. Typically, the dry-wet balance covers the value range from 0 (full dry) to 1 (full wet). Since the amount of decorrelation added to the dry signal through the effect processing is typically unknown, one has to expect power and resulting loudness variations in the mixed signal for different dry-wet balances similar to the cross-fading problem of the previous section. The design formulas Eq. (14-17) can directly be utilized to obtain the weighting coefficients by inserting the dry-wet balance g instead of the normalized time index α as proven by the following experiment.

An excerpt of the SQAMs *Guitar Sarasate* sample of length 10 s is convolved with the *M7 - 3 Rooms 08 Music Room* impulse response of the Samplicity's Impulse Response Library [9] to realize a reverberation effect. The power and loudness of $x_{\text{mix}}(n)$ is measured for different dry-wet balances g and plotted in Fig. 5. The dry and wet signal feature a correlation coefficient $r_{x_d,x_w} = -0.1417$. Hence, destructive interferences occur and the power of the mixed signal is significantly reduced for the linear and square-root relation of the weighting coefficients. In contrast, obtaining weighting coefficients with the proposed design results in constant power and almost constant loudness for all dry-wet balances. However, applying this approach is solely possible for off-line processing since the correlation coefficient has to be computed for the whole signal length before the dry-wet mix is performed and consequently the property of power preservation is solely achieved in temporal average. Nevertheless, the proposed approach can be utilized in a real-time application when the cross-correlation coefficient is sample-wise estimated as shown in [10].

5. CONCLUSION

Cross-fading is a key tool in audio editing of whatever form. The selection of fading curves by audio professionals is mainly based on experience and hence, is more or less imprecise. This work analytically derives perfectly power-complementary fading curves based on the exactly measured correlation of the signals to be cross-faded. The derivation is based on a simple mixing signal model and some basic assumptions which are met by most audio signals. Experiments on synthetic and real-world signals validate the effectiveness of the fading curve design. The proposed fading curve design can also be utilized in the context of dry-wet mixing due to similar boundary conditions. An additional experiment exposes the power-preservation for mixing dry and reverberated signals for any dry-wet balance using the proposed design.

6. REFERENCES

[1] S. Langford, *Digital Audio Editing: Correcting and Enhancing Audio in Pro Tools, Logic Pro, Cubase, and Studio One*, Digital Audio Editing: Correcting and Enhancing Audio in Pro Tools, Logic Pro, Cubase, and Studio One. Taylor & Francis, 2013.

[2] T. Görne, *Tontechnik*, Hanser, 2008.

[3] Jeffs Rick, "Evolution of the DJ Crossfader," Tech. Rep., Rane Corporation, 1999.

[4] Marco Fink, Martin Holters, and Udo Zölzer, "Comparison of various predictors for audio extrapolation," in *Proceedings of the 16th International Conference on Digital Audio Effects, DAFx-13, Maynooth, Ireland, September 2-5, 2013*, Erlangen, 2013.

[5] Marco Fink and Udo Zölzer, "Low-delay error concealment with low computational overhead for audio over IP applications," in *Proceedings of the 17th International Conference on Digital Audio Effects, DAFx-14, Erlangen, Germany, September 1-5, 2014*, Erlangen, 2014.

[6] "ITU Recommendation ITU-R BS.1770-3, Algorithms to measure audio programme loudness and true-peak audio level," Tech. Rep., International Telecommunications Union, August 2012.

[7] "EBU - TECH2353, Sound Quality Assessment Material recordings for subjective tests," Tech. Rep., European Broadcasting Union, September 2008.

[8] Marco Fink, Sebastian Kraft, and Udo Zölzer, "Downmix-compatibe conversion from mono to stereo in time-and frequency-domain," in *Proceedings of the 18th International Conference on Digital Audio Effects, DAFx-15, Trondheim, Norway, Nov 30-Dec 3, 2015*, Trondheim, 2015.

[9] "Samplicity's Bricasti M7 Impulse Response Library v1.1," `http://www.samplicity.com/bricasti-m7-impulse-responses/`.

[10] Ronald M. Aarts, Roy Irwan, and Augustus J.E.M. Janssen, "Efficient tracking of the cross-correlation coefficient," *Speech and Audio Processing, IEEE Transactions on*, vol. 10, no. 6, pp. 391–402, 2002.