

SUPPLEMENTARY MATERIAL

A. Proof of Proposition III.1 on temporal integration

At a discrete rate τ^{-1} , the Fourier transform of ϕ_T is

$$\widehat{\phi}_T(\omega) = \mathcal{F}\{\phi_T\}(e^{j\omega}) = \frac{se^{j\omega}}{e^{j\omega} - (1-s)}. \quad (1)$$

The squared magnitude of the denominator simplifies as

$$\begin{aligned} |e^{j\omega} - (1-s)|^2 &= (\cos \omega - (1-s))^2 + (\sin \omega)^2 \\ &= 1 + (1-s)^2 - 2(1-s)\cos \omega \\ &= (1 - (1-s))^2 + 2(1-s)(1 - \cos \omega) \\ &= s^2 + 2(1-s)(1 - \cos \omega). \end{aligned} \quad (2)$$

Let $\omega_0 = \frac{s}{\sqrt{1-s}}$. The equivalent $(1 - \cos \omega) \sim \frac{\omega^2}{2}$ at the limit $\omega \rightarrow 0$ indicates a quadratic decay in the lowest frequencies, i.e. a sidelobe falloff of 10dB per decade. On one hand, the double inequality $-1 \leq \cos \omega \leq 1$ leads to

$$\frac{\omega_0^2}{\omega_0^2 + 2} \leq |\widehat{\phi}_T|^2(\omega) \leq 1. \quad (3)$$

The upper (resp. lower) bound is attained if and only if $\omega = 0$ (resp. $\omega = \pi$). Consequently, ϕ_T has a DC gain of 0dB at 0Hz, and a gain of $10\log_{10} \frac{\omega_0^2}{2+\omega_0^2}$ (in dB) at the Nyquist frequency. On the other hand, the cutoff frequency at 3dB ω_c is defined by $|\widehat{\phi}_T|^2(\omega_c) = \frac{1}{2}$, which can be reformulated as

$$1 - \cos \omega_c = \frac{s^2}{2(1-s)} = \frac{\omega_0^2}{2}. \quad (4)$$

Solving Equation ?? with $\omega_c > 0$ completes the proof. ■

Remark. For large T , we can approximate ω_c by ω_0 , which is polynomial in s . Indeed, applying Taylor's theorem to the cosine function yields the inequality $|1 - \cos \omega - \frac{\omega^2}{2}| \leq \frac{\omega^4}{24}$, of which we deduce

$$\omega_0 \leq \omega_c \leq \frac{\omega_0}{\sqrt{1 - \frac{\omega_c^2}{12}}} \leq \omega_0 + \frac{\omega_0^2}{12} + \frac{\omega_0^3}{12}. \quad (5)$$

B. Proof of Proposition III.2 on adaptive gain control

First, applying Taylor's theorem to $x \mapsto (1+x)^{-\alpha}$ gives $|(\mathbf{M} + \varepsilon)^{-\alpha} - \varepsilon^{-\alpha}| \leq \alpha \frac{\mathbf{M}}{\varepsilon^{1+\alpha}}$, from which we deduce (i) with a relative error of at most $\alpha \frac{\mathbf{M}}{\varepsilon}$. Symmetrically, we have $|(\mathbf{M} + \varepsilon)^{-\alpha} - \mathbf{M}^{-\alpha}| \leq \alpha \frac{\varepsilon}{\mathbf{M}^{1+\alpha}}$ from which we deduce (ii) with a relative error of at most $\alpha \frac{\varepsilon}{\mathbf{M}}$. ■

C. Proof of Proposition III.4 on dynamic range compression

First, applying Taylor's theorem to $x \mapsto (1+x)^r$ gives $|\mathbf{PCEN} - r\delta^{(r-1)}\mathbf{G}| \leq \frac{1}{2}r\delta^{(r-1)}\mathbf{G}^2$, of which we deduce (i) with a relative error of at most $\frac{1-r}{2}\mathbf{G}$. Secondly, we have $\left| \left(1 + \frac{1}{\mathbf{G}^r}\right)^r - \left(r + \frac{r}{\mathbf{G}^r}\right) \right| \leq \frac{r(1-r)}{2\mathbf{G}^{2r}}$, which, after adding the constant $(1-r)$, leads to (ii) with a relative error of at most $\frac{1-r}{\mathbf{G}^r}(1 + \frac{r}{2\mathbf{G}^r})$ by application of the triangular inequality. ■

Remark. In both cases, the relative error is null if and only if $r = 1$, i.e. if there is no root compression.

D. Proof of Proposition IV.1 on the link between T and s

Let $\omega_0^2 = 2(1 - \cos \frac{2\pi\tau}{T})$. Proposition III.1 yields $\omega_0^2 = \frac{s^2}{(1-s)}$, and thus the quadratic equation $s^2 + \omega_0^2 s - \omega_0^2 = 0$. Its discriminant is $\Delta = \omega_0^4 + 4 \times \omega_0^2 = \omega_0^2(\omega_0^2 + 4) > 0$, and its two real-valued roots are

$$s = \frac{-\omega_0^2 \pm \omega_0 \sqrt{\omega_0^2 + 4}}{2} = \frac{\omega_0}{2} \left(\pm \sqrt{\omega_0^2 + 4} - \omega_0 \right) \quad (6)$$

We retain the positive root $s > 0$. Replacing ω_0 by its definition in the above equation completes the proof. ■

E. Proof of Proposition III.3 on invariance to impedance curve

Proposition. Let $\mathbf{s}(t)$ a realization of AWGN with null mean and unit variance. Let $\mathbf{a}(t) > 0$ a deterministic amplitude envelope $\mathbf{h}(t)$ a filter. Let $\mathbf{E}_x(t, f)$ the mel-frequency spectrogram associated to the source-filter model $\mathbf{x}(t) = \mathbf{a}(t) \times (\mathbf{s} * \mathbf{h})(t)$. If

- 1) $\forall t_0, \frac{d \log \mathbf{a}}{dt}(t_0) \ll \frac{1}{T}$,
- 2) $\forall t_0, \int_{t_0}^{t_0+\tau} \mathbf{h}(t) dt \ll \frac{1}{\tau}$, and
- 3) $\forall f_0, \frac{d \log |\mathbf{h}|}{df}(f_0) \ll \frac{1}{\Delta f_0}$,

where Δf_0 is the frequency interval between f_0 and its adjacent subbands on the mel scale, then $\mathbf{PCEN}_x(t, f) \approx \mathbf{PCEN}_s(t, f)$.

Proof. We adapt a result from [25]. The mel-frequency spectrogram $\mathbf{E}_x(t, f)$ can be defined as the complex modulus of the convolution between the signal $\mathbf{x}(t)$ and a filterbank $\psi_f(t)$ of N wavelets whose center frequencies f are tuned to the mel scale, ranging between f_{\min} and f_{\max} . For a given f , the first hypothesis allows to factorize the amplitude envelope $\mathbf{a}(t)$ out of the convolution $(\mathbf{s} * \mathbf{h} * \psi_f)(t)$. Furthermore, one has:

$$\begin{aligned} \mathbf{E}_{(\mathbf{s} * \mathbf{h})}(t, f) &= |\mathbf{s} * \mathbf{h} * \psi_f|(t) \\ &= \frac{1}{2\pi} \left| \int_{\mathbb{R}} \widehat{\mathbf{s}}(\omega) \widehat{\mathbf{h}}(\omega) \widehat{\psi}_f(\omega) \exp(2\pi i \omega t) dt \right|. \end{aligned} \quad (7)$$

The second and third hypotheses allow to approximate $\widehat{\mathbf{h}}(\omega)$ by $\widehat{\mathbf{h}}(f)$ in the equation above. This approximation leads to $\mathbf{E}_{(\mathbf{s} * \mathbf{h})}(t, f) \approx |\widehat{\mathbf{h}}(f)| \mathbf{E}_s(t, f)$. Combining the factorization of the amplitude term $\mathbf{a}(t)$ with the factorization of the filter $|\widehat{\mathbf{h}}(f)|$ leads to an approximation of the form:

$$|\mathbf{E}_x(t, f) - \mathbf{a}(t)| \widehat{\mathbf{h}}(f) \mathbf{E}_s(t, f)| = \eta(t, f). \quad (8)$$

An upper bound on the expectation of the stochastic residual η is given in [25, Appendix]. Likewise, the first hypothesis allows to approximate $\mathbf{M}_x(t, f)$ by $\mathbf{a}(t)|\widehat{\mathbf{h}}(f)|\mathbf{M}_s(t, f)$ with a stochastic residual term $v(t, f)$ of bounded expectation. In the active regime ($\mathbf{M} \gg \varepsilon$), AGC cancels $\mathbf{a}(t)$ and $|\widehat{\mathbf{h}}(f)|$ in Equation 3. Thus, Proposition III.2 leads to the inequality

$$|\mathbf{G}_x - \mathbf{G}_s|(t, f) \leq \frac{\eta(t, f)}{(\varepsilon + \mathbf{M}_s(t, f))^\alpha} + \frac{\alpha v(t, f)}{(\varepsilon + \mathbf{M}_s(t, f))^{1+\alpha}}. \quad (9)$$

Because $\delta > 1$ and $r \leq 1$, DRC is nonexpansive (see Proposition III.4). Therefore, it maps each member of the approximate equality $\mathbf{G}_x(t, f) \approx \mathbf{G}_s(t, f)$ to $\mathbf{PCEN}_x(t, f) \approx \mathbf{PCEN}_s(t, f)$ while reducing the expectation of the residual (right-hand side of Equation ??), which completes the proof. ■