ANTIDERIVATIVE ANTIALIASING WITH FREQUENCY COMPENSATION FOR STATEFUL SYSTEMS

Pier Paolo La Pastina and Stefano D'Angelo

Orastron Srl
Sessa Cilento, Italy
pierpaolo.lapastina@orastron.com | stefano.dangelo@orastron.com

ABSTRACT

Employing nonlinear functions in audio DSP algorithms requires attention as they generally introduce aliasing. Among others, antiderivative antialiasing proved to be an effective method for static nonlinearities and gave rise to a number of variants, including our AA-IIR method. In this paper we introduce an improvement to AA-IIR that makes it suitable for use in stateful systems. Indeed, employing standard antiderivative antialiasing techniques in such systems alters their frequency response and may cause stability issues. Our method consists in cascading a digital filter after the AA-IIR block in order to fully compensate for unwanted delay and frequency-dependent effects. We study the conditions for such a digital filter to be stable itself and evaluate the method by applying it to the diode clipper circuit.

1. INTRODUCTION

When designing digital signal processing algorithms for audio applications, one of the major challenges consists in dealing effectively with nonlinearities. Indeed, on one hand they have to be computed efficiently if the system is designed for real-time usage, while on the other they generally introduce new harmonics and thus they can create aliasing components.

The traditional way to reduce aliasing is *oversampling* [1], which essentially consists in artificially increasing the sampling rate at which the algorithm operates. This method is easy to implement and effective, yet it normally requires upsampling and downsampling filters and its computational cost increases linearly with the oversampling factor, hence it can prove to be expensive. More recently a new method called *antiderivative antialiasing* (AA) has been proposed by Parker et al. [2]. It virtually works in the following way: first the discrete-time input signal is converted into a continuous-time equivalent by linear interpolation, then the nonlinearity is applied and its output is filtered through a continuous-time lowpass filter, and finally the resulting signal is sampled back to the discrete-time domain. This procedure reduces aliasing because the nonlinearity and the lowpass filter are applied in the continuous-time domain, hence components above the Nyquist limit are attenuated before sampling back.

The original method employs lowpass filters with FIR kernels, such as rectangular and triangular kernel. An extension to higher-order FIR filters has been proposed in [3]. In [4] we developed an analogous method which uses arbitrary lowpass kernels that can be represented by rational functions in the frequency domain.

As these filters are infinite impulse response (IIR), we named this method AA-IIR and we referred to the original one in [2] as AA-FIR in order to distinguish them. This extension allows for a great amount of flexibility as one can obtain the desired trade-off between aliasing reduction and computational cost by simply choosing the most convenient lowpass filter.

Both AA-FIR and AA-IIR are designed for static, single-input single-output (SISO) nonlinearities, and they cannot be generally used for nonlinearities embedded in a dynamical system without further modifications. Indeed, while static nonlinearities introduce no delay and no frequency-dependent effects, the opposite is true of current AA methods, hence they would modify the overall frequency response of the system and potentially lead to stability issues. A solution in very specific cases, where the undesired delay effects of AA can be "merged" with corresponding parts of the stateful system in order to eliminate them, was already suggested in [2] and further explored in [5]. As of today, the only general approach to this problem has been proposed in [6, 7]. Therein, Holters modifies the structure of the system and adjusts its coefficients in order to compensate for the extra delay introduced by AA-FIR. The method has been adapted to wave digital filters in [8]. Although this method works, it can cause significant distortion of the frequency response of the system [6, 7, 9].

In this paper we introduce a different technique for AA in stateful systems. We simply propose to treat any SISO nonlinearity in the system with the regular AA-IIR method and add a digital filter in series to compensate for the delay and frequency-dependent effects introduced. By definition, this filter has to be the inverse of the linearization of the system obtained by applying AA to the nonlinearity: therefore, in order for the compensating filter to be stable, the linearization has to be minimum phase. Notice that this property is not at all guaranteed: for example, it does not hold for classical AA-FIR rectangular and triangular kernels, while it holds for AA-IIR with a single real pole, as we will show later. Moreover, higher order AA-IIR kernels do not always have a minimum phase linearization: we give some examples in the paper, without claiming completeness. The main advantages of this approach are that the changes to the original system are fully local to the nonlinearities and the overall frequency response remains unaffected. On the other hand its effectiveness largely depends on the original system in ways that are yet to be investigated.

The paper is organized as follows. In Section 2 we describe our method, which will be called *AA-IIR method with compensation*, while in Section 3 we discuss the stability of the compensating filter. Sections 4 and 5 analyze other relevant theoretical aspects, namely the influence of linear interpolation on the stability of the compensating filter and the side effects of numerical integration. In Section 6 we consider practical aspects of the proposed method and verify its features in an example application using two

different discretization methods. Finally Section 7 draws conclusions and suggests possible directions for future research, while Appendix A improves the formulation of the AA-IIR method from [4] for multiple poles.

2. THE GENERAL IDEA

We start by recalling the AA-FIR method from [2]. Suppose we have a static nonlinear discrete system of the form $y = f(x)$. We will assume from now on that $f(0) = 0$, without loss of generality: if this is not the case, one can work with the nonlinear function $f^*(x) = f(x) - f(0)$ and then add the offset term at the end. Finally, we will measure time in samples, so that the sampling frequency f_s is unitary.

If one feeds the system with an input x_n, the output may be affected by aliasing because of the nonlinear function f. In [2] Parker et al. construct the AA-FIR method in the following way: first they derive a continuous signal $\tilde{x}(t)$ from x_n by linear interpolation, then they apply a continuous-time causal lowpass filter, and finally they sample the continuous-time output signal back to the discrete-time domain. In formulas,

$$\tilde{x}(t) = x_{n-1} + (t - n + 1)(x_n - x_{n-1}) \tag{1}$$

if $t \in [n-1, n]$ and, if $h(t)$ is the impulse response of the chosen filter,

$$y_n = \int_{-\infty}^{+\infty} f(\tilde{x}(t))h(n-t)\,\mathrm{d}t. \tag{2}$$

By choosing the rectangular kernel one gets

$$\begin{aligned} y_n &= \int_{n-1}^{n} f(x_{n-1} + (t-n+1)(x_n - x_{n-1}))\,\mathrm{d}t \\ &= \int_0^1 f(x_{n-1} + t(x_n - x_{n-1}))\,\mathrm{d}t, \end{aligned} \tag{3}$$

i.e.,

$$y_n = \begin{cases} \dfrac{F(x_n) - F(x_{n-1})}{x_n - x_{n-1}} & x_n \neq x_{n-1} \\ f(\bar{x}) & x_n = x_{n-1} = \bar{x} \end{cases}$$

where F is an antiderivative of f.

As already observed in [2], the system (3) has non-flat amplitude response and non-null group delay in linear (small signal) terms, unlike the original system. One quick way to see this is observing that (3) is equivalent to $y_n = \Phi(x_n, x_{n-1})$, where

$$\Phi(u, v) = \int_0^1 f(v + t(u - v))\,\mathrm{d}t.$$

As a consequence, the linearization of (3) around the equilibrium point $(0, 0)$ reads

$$\widetilde{y}_n = \frac{\partial \Phi}{\partial u}(0,0)x_n + \frac{\partial \Phi}{\partial v}(0,0)x_{n-1}.$$

But

$$\begin{aligned} \frac{\partial \Phi}{\partial u}(0,0) &= \int_0^1 f'(0)t\ \ \mathrm{d}t = \frac{f'(0)}{2}, \\ \frac{\partial \Phi}{\partial v}(0,0) &= \int_0^1 f'(0)(1-t)\,\mathrm{d}t = \frac{f'(0)}{2}, \end{aligned}$$

so the linearization is

$$\widetilde{y}_n = f'(0)\frac{x_n + x_{n-1}}{2}. \tag{4}$$

Applying such method to a nonlinearity embedded in a dynamical system would, therefore, affect its overall (small signal) frequency response and potentially undermine its stability. One could think of adding a digital filter after the modified nonlinearity to restore its delay-free static nature, yet the inverse filter of (4) would only be marginally stable. Also, if the triangular kernel was used instead, we would have

$$\widetilde{y}_n = f'(0)\frac{x_n + 4x_{n-1} + x_{n-2}}{6}$$

and it is easily seen that one of the zeros is outside the unit circle, hence its inverse would be unstable.

In the following, we will show that this approach is actually viable if we use our AA-IIR method [4]. We will first recall its construction and then explore conditions to obtain a linearization with stable inverse, i.e., a minimum phase linearization. The AA-IIR method together with the inverse filter will be called *AA-IIR method with compensation*.

3. APPLYING THE AA-IIR METHOD

The AA-IIR method exploits the same idea as Parker's method, but uses IIR continuous-time filters, more precisely filters whose transfer function is a rational function in the Laplace domain. Applying partial fraction decomposition, such a transfer function reads

$$\begin{aligned} H(s) =& A_0 + \sum_{k=1}^{p}\sum_{l=1}^{m_k} \frac{A_{kl}}{(s-\alpha_k)^l} \\ &+ \sum_{k=1}^{q}\sum_{l=1}^{m_k}\left(\frac{B_{kl}}{(s-\beta_k)^l} + \frac{\overline{B}_{kl}}{(s-\overline{\beta}_k)^l}\right) \end{aligned} \tag{5}$$

where $\alpha_1, \ldots, \alpha_p$ are real poles of multiplicities $m_1, \ldots, m_p$ respectively, and $\beta_1, \overline{\beta}_1, \ldots, \beta_q, \overline{\beta}_q$ are complex poles of multiplicities $\mu_1, \ldots, \mu_q$ respectively. We will assume that this filter is stable, i.e., $\alpha_1, \ldots, \alpha_p, \Re(\beta_1), \ldots, \Re(\beta_q) < 0$.

It is clear that one can treat summands of (5) separately and then sum the results. A representation of AA-IIR method with compensation is given in Figure 1, where $H_{\alpha_i}(s)$ and $H_{\beta_i}(s)$ are the summands in (5) relative to poles α_i and to couples of poles β_i, $\overline{\beta}_i$ respectively, and $H_{\text{comp}}(z)$ is the transfer function of the (digital) compensation filter.

Explicit equations for the cases of simple real pole, multiple real pole, couple of simple complex conjugate poles and couple of multiple complex conjugate poles are given in [4]. We will recall them here and address the invertibility issue.

3.1. Simple real pole

Suppose that $H(s) = \frac{A}{s-\alpha}$. In [4] we have shown that the AA-IIR method reads

$$y_n = e^{\alpha} y_{n-1} + A \fint_{x_{n-1}}^{x_n} f(\xi) e^{\alpha \frac{x_n - \xi}{x_n - x_{n-1}}}\,d\xi,$$

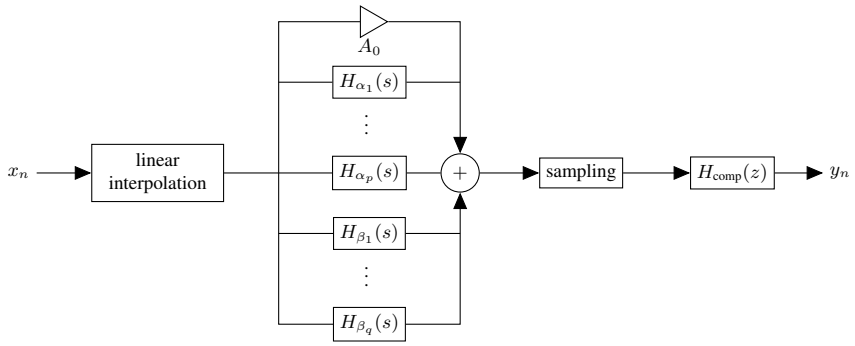

Figure 1: Block diagram of AA-IIR method with compensation.

where the symbol $\fint$ denotes the *mean value* of an integrable function:

$$\fint_a^b g(\xi)d\xi := \frac{1}{b-a}\int g(\xi)d\xi.$$

Similarly, one can find the equivalent formulation:

$$y_n = e^{\alpha}y_{n-1} + A\int_0^1 f(x_{n-1}+t(x_n-x_{n-1}))e^{\alpha(1-t)}\,\mathrm{d}t \quad (6)$$

In the following, we will use the latter as it is more general (it encompasses also the case $x_n = x_{n-1}$) and easier to differentiate.

We can linearize the system as seen in Section 2. The result is:

$$\widetilde{y}_n = e^{\alpha}\widetilde{y}_{n-1} + f'(0)\frac{A}{\alpha^2}\Big((e^{\alpha}-\alpha-1)x_n + ((\alpha-1)e^{\alpha}+1)x_{n-1}\Big). \quad (7)$$

Notice that this result coincides with the one obtained in [4], Appendix B, for the case $A=-\alpha$, $f'(0)=1$.

Such linearized filter has a single zero

$$\zeta = -\frac{(\alpha-1)e^{\alpha}+1}{e^{\alpha}-\alpha-1}$$

and a direct inspection shows that $-1 < \zeta < 0$ when $\alpha < 0$. Therefore *the system (6) has a minimum phase linearization for every* $\alpha < 0$. This is a remarkable result because the AA-IIR method with a simple real pole shows antialiasing performance comparable to the AA-FIR method with rectangular kernel, provided that the pole is chosen appropriately.

3.2. Multiple real pole

Suppose now that $H(s) = \frac{A}{(s-\alpha)^{r+1}}$, with $r > 0$. From Appendix A we know that the AA-IIR method in this case is described by Eqs. (16)-(17). The linearization of the latter equation takes the form:

$$v_n^{(k)} = e^{\alpha}\sum_{l=0}^{k}\binom{k}{l}v_{n-1}^{(l)} + b_0^{(k)}x_n + b_1^{(k)}x_{n-1}, \quad (8)$$

for $k = 0, \ldots, r$, where

$$b_0^{(k)} = f'(0)\int_0^1 t(1-t)^k e^{\alpha(1-t)}\,\mathrm{d}t,$$
$$b_1^{(k)} = f'(0)\int_0^1 (1-t)^{k+1}e^{\alpha(1-t)}\,\mathrm{d}t.$$

Defining $v_n = (v_n^{(0)}\ v_n^{(1)}\ \ldots\ v_n^{(r)})^t$, $b_0 = (b_0^{(0)}\ b_0^{(1)}\ \ldots\ b_0^{(r)})^t$, $b_1 = (b_1^{(0)}\ b_1^{(1)}\ \ldots\ b_1^{(r)})^t$ and

$$M = \begin{pmatrix} \binom{0}{0} & 0 & 0 & \ldots & 0 \\ \binom{1}{0} & \binom{1}{1} & 0 & \ldots & 0 \\ \binom{2}{0} & \binom{2}{1} & \binom{2}{2} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ \binom{r}{0} & \binom{r}{1} & \binom{r}{2} & \ldots & \binom{r}{r} \end{pmatrix},$$

the system (8) can be concisely expressed as

$$v_n = e^{\alpha}Mv_{n-1} + b_0x_n + b_1x_{n-1}.$$

In the z domain, this translates into

$$V(z) = (I - e^{\alpha}Mz^{-1})^{-1}(b_0 + b_1z^{-1})X(z).$$

In order to determine when the system described by Eqs. (16)-(17) has a minimum phase linearization, one should understand if the zeros of the last element of $(I - e^{\alpha}Az^{-1})^{-1}(b_0 + b_1z^{-1})$ lie inside the unit circle. Now we will show that this cannot be done analytically even for $r = 1$, i.e., for a double pole.

Indeed, for $r = 1$ we have:

$$b_0^{(0)} = f'(0)\frac{e^{\alpha}-\alpha-1}{\alpha^2},$$
$$b_1^{(0)} = f'(0)\frac{(\alpha-1)e^{\alpha}+1}{\alpha^2},$$
$$b_0^{(1)} = f'(0)\frac{(\alpha-2)e^{\alpha}+\alpha+2}{\alpha^3},$$
$$b_1^{(1)} = f'(0)\frac{(\alpha^2-2\alpha+2)e^{\alpha}-2}{\alpha^3},$$
$$M = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

After some computations, one finds that

$$V^{(1)}(z) = f'(0)\frac{\tilde{b}_0 + \tilde{b}_1z^{-1} + \tilde{b}_2z^{-2}}{(1-e^{\alpha}z^{-1})^2}X(z),$$

where

$$\tilde{b}_0 = \frac{(\alpha-2)e^{\alpha}+\alpha+2}{\alpha^3},$$
$$\tilde{b}_1 = \frac{2(e^{2\alpha}-2\alpha e^{\alpha}-1)}{\alpha^3},$$
$$\tilde{b}_2 = \frac{e^{\alpha}((\alpha-2)e^{\alpha}+\alpha+2)}{\alpha^3}.$$

It is clear that the equation $\tilde{b}_0 + \tilde{b}_1 z^{-1} + \tilde{b}_2 z^{-2} = 0$ can be solved analytically and has 2 real solutions ζ_1, ζ_2. However, there is no way of determining analytically when $|\zeta_1|, |\zeta_2| < 1$.

By numerically solving the inequalities one finds that both roots lie inside the unit circle if and only if $\alpha < -3.21$. As the Nyquist frequency corresponds to $\alpha = -\frac{\pi}{2} \approx -1.57$, this limitation is too restrictive for practical applications.

Furthermore, for triple and quadruple poles we found that a necessary condition for the system (16)-(17) to have a minimum phase linearization is $\alpha < -5.65$ and $\alpha < -7.72$, respectively. We expect that if the order increases the upper bound decreases even more, making the filter unsuitable for AA-IIR with compensation.

3.3. Simple complex conjugate poles

Now suppose that $H(s)$ has only a couple of simple complex conjugate poles, i.e., $H(s) = \frac{B}{s-\beta} + \frac{\overline{B}}{s-\overline{\beta}}$, where $B, \beta \in \mathbb{C}$ and $\Re(\beta) < 0$. Then the AA-IIR method reads

$$\begin{cases} u_n = e^{\beta} u_{n-1} + 2B \int_{x_{n-1}}^{x_n} f(\xi) e^{\beta \frac{x_n - \xi}{x_n - x_{n-1}}} d\xi \\ y_n = \Re(u_n) \end{cases}$$

or, equivalently,

$$\begin{cases} u_n = e^{\beta} u_{n-1} + 2B \int_0^1 f(x_{n-1} + t(x_n - x_{n-1})) e^{\beta(1-t)} \, \mathrm{d}t \\ y_n = \Re(u_n) \end{cases}.$$

The linearized system is:

$$\begin{cases} v_n = e^{\beta} v_{n-1} + 2f'(0) \frac{B}{\beta^2} \Big((e^{\beta} - \beta - 1) x_n \\ \qquad\quad + ((\beta - 1)e^{\beta} + 1) x_{n-1} \Big) \\ \widetilde{y}_n = \Re(v_n) \end{cases}. \tag{9}$$

Therefore, in the z-domain we have:

$$\hat{H}_{\text{lin}}(z) = 2f'(0) \frac{B}{\beta^2} \frac{(e^{\beta} - \beta - 1) + ((\beta - 1)e^{\beta} + 1) z^{-1}}{1 - e^{\beta} z^{-1}};$$
$$H_{\text{lin}}(z) = \frac{1}{2} (\hat{H}_{\text{lin}}(z) + \overline{\hat{H}_{\text{lin}}(\overline{z})}).$$

As one can expect, $H_{\text{lin}}(z)$ is a second-order transfer function with real coefficients, so one can explicitly compute its zeros. However, it is not possible to determine analytically when these zeros lie inside the unit circle, so we performed a numerical evaluation.

In the case of complex conjugate poles, the region where one has to choose β also depends on B. Therefore, in order to get an idea we studied the case when $H(s)$ has unitary DC gain and has no zeros. The latter condition corresponds to $\Re(B) = 0$. In this case, the transfer function becomes

$$H(s) = \frac{-2\Re(B\overline{\beta})}{(s - \beta)(s - \overline{\beta})} \tag{10}$$

and we get

$$B = -\frac{i}{2} \frac{|\beta|^2}{\Im(\beta)},$$

where $\Re(z)$ and $\Im(z)$ denote the real and the imaginary part of a complex number z, respectively.

We chose β so that $-10 \leq \Re(\beta) < 0$, $-20 \leq \Im(\beta) \leq 20$, on an uniform grid with a 0.01 step on both axes, and for each value we found numerically the zeros of $H_{\text{lin}}(z)$. We discovered that the linearized system (9) is not minimum-phase when β belongs to some semi-ellipses with center on the y-axis and with the x-semiaxes slightly longer than the y-semiaxes. The biggest semi-ellipse is centered in the origin, while others appear in pairs, symmetrically w.r.t the x-axis. In the following table we give the measured values:

Center	**x-semiaxis**	**y-semiaxis**
$(0, 0)$	3.21	3.14
$(0, \pm 9.205)$	0.22	0.215
$(0, \pm 15.58)$	0.13	0.12

The x-semiaxis of the first ellipse is 3.21, the parameter we found in Subsection 3.2: this is not a surprise since Eq. (10) tends to $H(s) = \frac{\beta^2}{(s-\beta)^2}$ when the imaginary part of β goes to 0. When we move away from the origin, the dimensions of the semi-ellipses decrease, so we can expect smaller and smaller semi-ellipses to appear if we extend the range of the simulation.

However, please recall that $f_s = 1$, so in order to obtain good aliasing reduction we should choose $|\beta|$ around $\frac{\pi}{2}$. This region is entirely inside the bigger semi-ellipse, so we can conclude that (10 is not a good candidate kernel for the AA-IIR method with compensation.

3.4. Multiple complex conjugate poles

Finally, we briefly observe that similar considerations apply to the case of a couple of multiple complex conjugate poles:

$$H(s) = \frac{B}{(s - \beta)^{r+1}} + \frac{\overline{B}}{(s - \overline{\beta})^{r+1}},$$

with $r > 0$. In this case, the feasible region will depend again on both B and β.

Let us consider the case $r = 1$. If we impose that $H(s)$ has unitary gain at DC and, as before, $\Re(B) = 0$, the transfer function becomes:

$$H(s) = \frac{|\beta|^4}{\Re(\beta)} \frac{s - \Re(\beta)}{(s - \beta)^2 (s - \overline{\beta})^2}.$$

When the imaginary part of β goes to 0, the latter tends to $H(s) = \frac{\beta^3}{(s-\beta)^3}$, that is a filter with no zeros and a triple real pole. From the discussion in Subsection 3.2, it follows that the feasible region for β in this case will intersect the real axis in the half-line $x < -5.65$. As a consequence, we expect a no-go area around the origin analogous to the one found in the last subsection, but even bigger: in this case, we would similarly not use this filter for AA-IIR with compensation. We expect that similar issues appear when $r > 1$.

3.5. The general case

So far, we discussed the cases in which the continuous-time filter only has no zeros and only one real pole or one couple of complex conjugate poles of arbitrary multiplicity. However, in general a filter for AA-IIR has p real poles, q complex pole couples, and up to $p + 2q$ zeros (see Eq. (5)). Therefore, investigating stability conditions in the general case means studying a space that has at least p real and q complex dimensions.

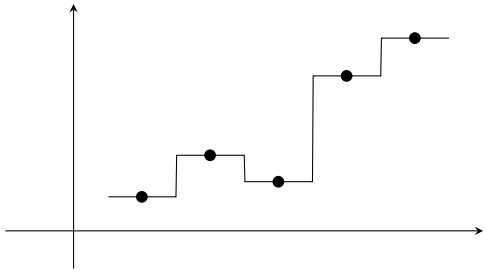

Figure 2: An alternative reconstruction process.

However, in the previous subsections we considered that this is already difficult when $p+q=1$, and it is almost always impossible to do it analytically. Therefore, we do not treat the general case in more detail in this work, but rather suggest case-by-case evaluation.

4. THE ROLE OF THE RECONSTRUCTION PROCESS

Our AA-IIR method was constructed by replacing the rectangular kernel in the AA-FIR method with a lowpass filter with rational transfer function. However, up to now the linear interpolation reconstruction process that virtually converts the input discrete-time signal into a continuous-time one has not undergone any particular investigation. It is perhaps obvious that choosing an interpolation kernel that better approximates the ideal reconstruction filter would yield better anti-aliasing performance, yet here we show that such choice also has consequences on whether the entire process is minimum phase in linear terms, and hence whether it is possible to apply compensation and under which circumstances.

Suppose that x_n is a discrete-time signal and we convert it to continuous-time by (see Figure 2)

$$\widehat{x}(t) = \begin{cases} x_{\lfloor t \rfloor} & \mod(t,1) < \frac{1}{2} \\ x_{\lceil t \rceil} & \mod(t,1) \geq \frac{1}{2} \end{cases}, \tag{11}$$

which corresponds to using a rectangular function of width and height 1 as the interpolation kernel.

We can apply a causal lowpass filter with transfer function $h(t)$ to $\widehat{x}(t)$ and then sample the result, obtaining

$$y_n = \int_{-\infty}^{+\infty} f(\widehat{x}(t)) h(n-t)\, \mathrm{d}t$$

in analogy to (2).

Suppose first that $H(s) = \frac{A}{s-\alpha}$ with $\alpha < 0$, i.e., $h(t) = Ae^{\alpha t}u(t)$, as in Subsection 3.1. A similar procedure to that employed in [4] leads to

$$y_n = e^{\alpha} y_{n-1} + A\frac{e^{\frac{\alpha}{2}}-1}{\alpha} f(x_n) + A\frac{e^{\alpha}-e^{\frac{\alpha}{2}}}{\alpha} f(x_{n-1}).$$

It is straightforward to see that its linearization has a simple zero $\zeta = -e^{\frac{\alpha}{2}}$, so its inverse is always stable as when using linear interpolation.

Now set $H(s) = \frac{A}{(s-\alpha)^2}$. As in Appendix A, one can express the antialiasing method as:

$$\begin{cases} u_n^{(0)} = & e^{\alpha} u_{n-1}^{(0)} + \frac{e^{\frac{\alpha}{2}}-1}{\alpha} f(x_n) + \frac{e^{\alpha}-e^{\frac{\alpha}{2}}}{\alpha} f(x_{n-1}) \\ u_n^{(1)} = & e^{\alpha}(u_{n-1}^{(0)} + u_{n-1}^{(1)}) + \frac{2+(\alpha-2)e^{\frac{\alpha}{2}}}{2\alpha^2} f(x_n) \\ & + \frac{2(\alpha-1)e^{\alpha}-(\alpha-2)e^{\frac{\alpha}{2}}}{2\alpha^2} f(x_{n-1}) \\ y_n = & A u_n^{(1)} \end{cases}.$$

One can linearize this system, find the zeros of its transfer function in analogy to what we did in Subsection 3.2 and study for which values of α they lie inside the unit circle. Surprisingly enough, we discovered that the linearization is minimum phase if and only if $\alpha < -1.69$. In absolute terms, this limit is slightly lower than $\alpha = -\frac{\pi}{2}$ which corresponds to the Nyquist frequency. Hence, in principle, using this reconstruction one could employ the double pole filter for AA-IIR with compensation, while we had to rule out this possibility when using linear interpolation. This suggests that further research could be conducted to find out which reconstruction kernels are better suited to AA-IIR with compensation.

5. NUMERICAL INTEGRATION

The AA-IIR method requires the evaluation of some integrals, such as the one in (6). Depending on the nature of the function f, it may not be possible or otherwise convenient to compute the integral analytically, in which case it is possible to employ numerical integration. For example, if one approximates integrals with the trapezoidal rule, the algorithm (6) for the simple real pole becomes

$$\begin{aligned} y_n =& e^{\alpha} y_{n-1} + \frac{A}{N}\Bigg(\frac{e^{\alpha} f(x_{n-1}) + f(x_n)}{2} \\ &+ \sum_{i=0}^{N} f\Big(x_{n-1} + \frac{i}{N}(x_n - x_{n-1})\Big) e^{\alpha\left(1-\frac{i}{N}\right)}\Bigg), \end{aligned} \tag{12}$$

and its linearization (7) becomes

$$\begin{aligned} \widetilde{y}_n =& e^{\alpha}\widetilde{y}_{n-1} + f'(0)\frac{A}{N}\Bigg(\frac{1}{2} + \sum_{i=0}^{N} \frac{i}{N} e^{\alpha\left(1-\frac{i}{N}\right)}\Bigg) x_n \\ &+ f'(0)\frac{A}{N}\Bigg(\frac{e^{\alpha}}{2} + \sum_{i=0}^{N}\Big(1-\frac{i}{N}\Big) e^{\alpha\left(1-\frac{i}{N}\right)}\Bigg) x_{n-1}. \end{aligned} \tag{13}$$

This approximation has some minor effects on the compensating filter that are worth mentioning. Firstly, the regions where one has to choose poles in order to obtain a minimum-phase linearization will be slightly different from those found in Section 3. Moreover, in general some inaccuracies in DC gain will be introduced, but they will be automatically counterbalanced by the corresponding compensation filter. Similar considerations also hold for multiple real poles and complex poles.

6. PRACTICAL CONSIDERATIONS

The AA-IIR method with compensation aims at reducing aliasing produced by static nonlinearities in a dynamical system without affecting its overall frequency response. This result can be achieved by employing AA-IIR on nonlinearities and then adding a digital filter in series to compensate the unwanted frequency effects of AA-IIR. These two features of the method, namely the preservation of frequency response and the locality of modifications, which

in turn also imply consistency in time-varying behavior and stability in small-signal terms, represent its main advantages over Holters' method [6, 7].

The compensation filter typically has no effect on the lowest frequencies and boosts frequencies near the Nyquist limit. Since it is applied at the end of the process, this extra gain indiscriminately affects all components, whether they are "linear terms", thus re-establishing the original frequency response, nonlinear elements, which are incorrectly magnified at highest frequencies, or aliasing components, which should hopefully end up being attenuated at low/mid frequencies but most likely boosted at higher frequencies. Therefore, even if not strictly necessary, it is still advisable to employ mild oversampling over the whole system with the current formulation of the method, as with Holters' method.

The stability of the compensation filter is directly affected by both the reconstruction and continuous-time filter kernels. When using linear interpolation for reconstruction, which is customary to this day for AA methods, only first-order IIR kernels can be serenely chosen for the continuous-time filter. We have however shown in Section 4 how employing other reconstruction kernels can expand the stability range associated with continuous-time filter kernels with steeper rolloff. With a view to conducting further research in this sense, we expect that investigating specific classes of reconstruction filters could yield practically appreciable results. Namely, given the formulations of both AA-FIR and AA-IIR, we suggest studying those kernels $k(t)$ which possess the following characteristics:

1. $k(0) = 1$ and $k(n) = 0$ with $n \in \mathbb{Z}, n \neq 0$, so that the continuous-time signal is equal to the original discrete-time signal at the corresponding discrete-time indices;
2. $\int_{-\infty}^{\infty} k(t)\ \mathrm{d}t = 1$ and $\sum_{i=-\infty}^{\infty} k(i + \Delta t) = 1, \forall \Delta t \in [0, 1)$, so that unitary DC gain is guaranteed overall and with any fixed Δt offset;
3. lead to a causal algorithm, and possibly to both AA and filtering parts being each causal.

In the rest of this Section we evaluate the AA-IIR method with compensation by applying it to the diode clipper circuit, as it was both often studied in previous works [6, 7, 8] and also since its nonlinearity is close to the output, which allows us to clearly examine the effects of the method.

6.1. Example: the diode clipper

The diode clipper is a circuit that prevents the output from exceeding a predefined voltage level. Figure 3 shows a dynamical version of the circuit that includes a first-order lowpass filter. The circuit can be fully described by [10]

$$\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{x - y}{RC} - 2\frac{I_\mathrm{s}}{C}\sinh\left(\frac{y}{V_\mathrm{T}}\right), \tag{14}$$

where x and y are the input and the output voltage, respectively, I_s is the saturation current and V_T is the thermal voltage.

We will construct a digital model of (14) following [11]. First, we discretize the derivative on the left hand side with a general linear one-step method:

$$(Dy)_n = B_0 y_n + B_1 y_{n-1} - A_1 (Dy)_{n-1}.$$

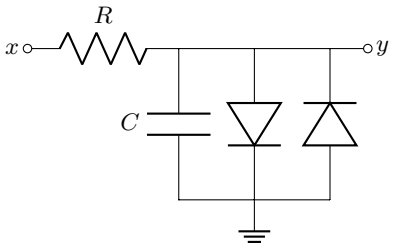

Figure 3: Schematics of the diode clipper circuit.

Then, we approximate $\sinh(x) \approx \frac{1}{2}\mathrm{sign}(x)(e^{|x|} - 1)$. We obtain:

$$B_0 y_n + B_1 y_{n-1} - A_1 (Dy)_{n-1} = \frac{x_n - y_n}{RC} - \frac{I_\mathrm{s}}{C}\mathrm{sign}(y_n)\left(e^{\frac{|y_n|}{V_\mathrm{T}}} - 1\right).$$

This equation can be analytically solved with the help of the Wright Omega function and put in a suitable form for AA-IIR with compensation:

$$\begin{cases} \xi_n = x_n - RC(B_1 y_{n-1} - A_1 (Dy)_{n-1}) \\ y_n = f(\xi_n) \end{cases},$$

where

$$f(x) = \frac{x + I_\mathrm{s} R \mathrm{sign}(x)}{1 + B_0 RC} - V_\mathrm{T}\mathrm{sign}(x)\omega\left(\frac{|x| + I_\mathrm{s} R}{V_\mathrm{T}(1 + B_0 RC)} + \log\left(\frac{I_\mathrm{s} R}{V_\mathrm{T}(1 + B_0 RC)}\right)\right). \tag{15}$$

We chose $R = 1\ \mathrm{k}\Omega$, $C = 33$ nF, $I_\mathrm{s} = 1$ fA, $V_\mathrm{T} = 25$ mV, and $f_\mathrm{s} = 44100$ Hz, and we discretized the derivative using the Euler backwards method, i.e.,

$$B_0 = f_\mathrm{s}, \quad B_1 = -f_\mathrm{s}, \quad A_1 = 0.$$

We then modified this algorithm by applying the proposed method using a single pole filter $H(s) = -\frac{\alpha}{s - \alpha}$, with $\alpha = -\frac{\pi}{4}$. It is clear that the integral in (6) cannot be computed analytically with f as in (15); therefore, we used the numerical algorithm (12) with $N = 5$.

Firstly, we verified that our method does not affect the small-signal frequency response of the overall system by feeding both algorithms with an impulse of amplitude 10^{-6} and then comparing the magnitude spectra of the outputs, thus obtaining practically identical results. Then, we supplied both algorithms with an input sine wave of amplitude 10 and frequency 986.96 Hz to assess the properties of our method. The magnitude spectra of the outputs are shown in Figure 4. As expected, we see that with our method the fundamental has identical amplitude, the harmonics are boosted at increasing frequency, while aliasing components are attenuated at low-mid frequencies and boosted in the highest part of the spectrum.

When using the bilinear transform instead, i.e.,

$$B_0 = 2f_\mathrm{s}, \quad B_1 = -2f_\mathrm{s}, \quad A_1 = 1.$$

the small-signal frequency response is still preserved, but aliasing noise actually increases in the output, as shown in Figure 5. We

conclude that the performance of the method is not always guaranteed and has to be assessed on a case-by-case basis. Further research could be devoted to finding general criteria for the good functioning of the method.

GNU Octave implementations of all four algorithms are available on the companion web page for this paper[1].

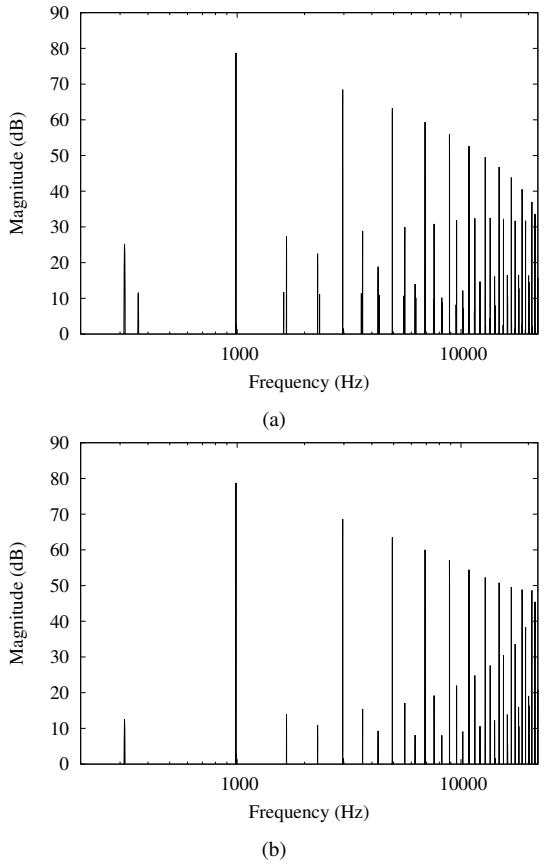

Figure 4: Magnitude spectra of the output of (a) the original and (b) the modified diode clipper simulation algorithms, discretized with Euler backwards, when fed with an input sine wave of amplitude 10 and frequency 986.96 Hz, and running at a sample rate $f_s = 44100$ Hz.

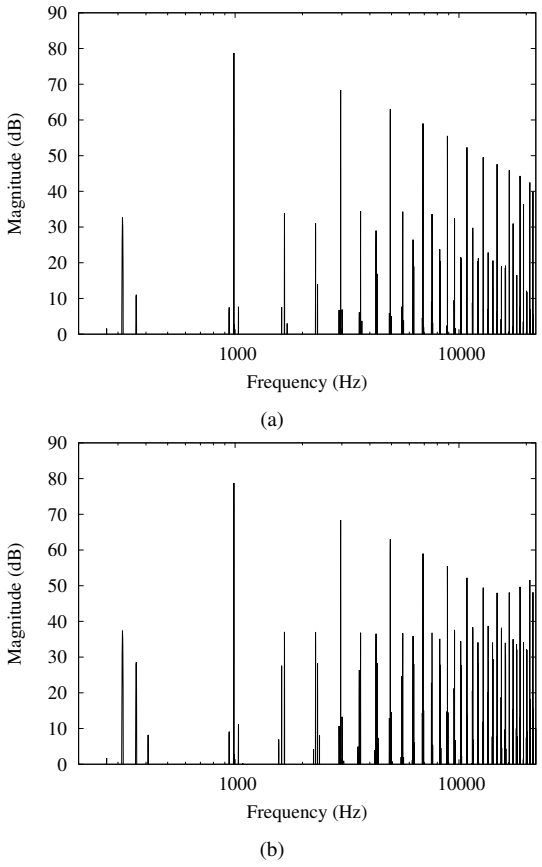

Figure 5: Magnitude spectra of the output of (a) the original and (b) the modified diode clipper simulation algorithms, discretized with the bilinear transform, when fed with an input sine wave of amplitude 10 and frequency 986.96 Hz, and running at a sample rate $f_s = 44100$ Hz.

7. CONCLUSIONS

This paper introduces an improvement to our previous AA-IIR method that is particularly useful for static nonlinearities embedded in dynamical systems: it consists in cascading a digital filter in series with the modified nonlinearities to compensate for unwanted delay and frequency-dependent effects introduced by the AA-IIR method, which in turn also leads to consistency in time-varying behavior and stability in small-signal terms.

The stability of such compensation filter is dependent on the chosen AA-IIR reconstruction and continuous-time filtering kernels. When using linear interpolation for reconstruction, filters with a single real negative pole will always lead to stable compensation filters, while more attention should be paid if one wants to use higher-order kernels. Other choices for the reconstruction phase will lead to different stability criteria, hence further research in this direction would be desirable.

When tested on a virtual analog example, the method in its most obvious fomulation shows very unequal performances when different discretizations are employed. In one instance it reshapes aliasing noise by attenuating it at low/mid frequencies at the expense of boosting nonlinear and aliasing components in the highest part of the digital spectrum, while in another it turns out to be counterproductive as aliasing noise is increased over the whole spectrum. Hence we suggest case-by-case evaluation and possibly coupling it with mild oversampling when it is effective, even though not strictly required. On the other hand we wish that future investigations will shed some light on which systems it is best suited for and possibly propose extensions to expand this class as much as possible.

[1] https://www.dangelo.audio/dafx20in22-aaiir.html

8. REFERENCES

[1] Dimitris G. Manolakis and Vinay K. Ingle, *Applied Digital Signal Processing: Theory and Practice*, Cambridge University Press, 2011.

[2] Julian Parker, Vadim Zavalishin, and Efflam Le Bivic, "Reducing the aliasing of nonlinear waveshaping using continuous-time convolution," in *Proc. 19th Intl. Conf. Digital Audio Effects (DAFx-16)*, Brno, Czech Repulic, 2016, pp. 137–144.

[3] Stefan Bilbao, Fabián Esqueda, Julian D Parker, and Vesa Välimäki, "Antiderivative antialiasing for memoryless nonlinearities," *IEEE Signal Processing Letters*, vol. 24, no. 7, pp. 1049–1053, 2017.

[4] Pier Paolo La Pastina, Stefano D'Angelo, and Leonardo Gabrielli, "Arbitrary-order IIR antiderivative antialiasing," in *Proc. 24th Intl. Conf. Digital Audio Effects (DAFx20in21)*, Vienna, Austria, 2021, pp. 9–16.

[5] Effrosyni Paschou, Fabián Esqueda, Vesa Välimäki, and John Mourjopoulos, "Modeling and measuring a Moog voltage-controlled filter," in *Asia-Pacific Signal and Inform. Process. Assoc. Annual Summit and Conf. (APSIPA ASC)*, Kuala Lumpur, Malaysia, 2017, pp. 1641–1647.

[6] Martin Holters, "Antiderivative antialiasing for stateful systems," in *Proc. 22nd Intl. Conf. Digital Audio Effects (DAFx-19)*, Birmingham, UK, 2019.

[7] Martin Holters, "Antiderivative antialiasing for stateful systems," *Applied Sciences*, vol. 10, no. 1, 2020.

[8] Davide Albertini, Alberto Bernardini, and Augusto Sarti, "Antiderivative antialiasing in nonlinear wave digital filters," in *Proc. 23rd Intl. Conf. Digital Audio Effects (eDAFx2020)*, Vienna, Austria, 2020, pp. 62–69.

[9] Kurt J. Werner, "An equivalent circuit interpretation of antiderivative antialiasing," in *Proc. 24th Intl. Conf. Digital Audio Effects (DAFx20in21)*, Vienna, Austria, 2021, pp. 17–24.

[10] David T. Yeh, Jonathan Abel, and Julius O. Smith, "Simulation of the diode limiter in guitar distortion circuits by numerical solution of ordinary differential equations," in *Proc. 10th Intl. Conf. Digital Audio Effects (DAFx-07)*, Bordeaux, France, 2007.

[11] Stefano D'Angelo, Leonardo Gabrielli, and Luca Turchet, "Fast approximation of the Lambert W function for virtual analog modelling," in *Proc. 22nd Intl. Conf. Digital Audio Effects (DAFx-19)*, Birmingham, UK, 2019.

A. APPENDIX: A NEW FORMULATION OF AA-IIR FOR MULTIPLE POLES

In this appendix, we derive a formulation of the AA-IIR method for multiple poles different from that proposed in [4], that is clearer and more suitable for our purposes.

A.1. Real case

Suppose that $H(s) = \frac{A}{(s-\alpha)^{r+1}}$, with $A \in \mathbb{R}$, $\alpha < 0$ and $r > 0$. Then the associated impulse response is $h(t) = A\frac{t^r}{r!}e^{\alpha t}u(t)$, where $u(t)$ is the Heaviside function. From Eqs. (1) and (2):

$$\begin{aligned} y_n =& \int_0^n f(\tilde{x}(t))h(n-t)\,\mathrm{d}t \\ =& \frac{A}{r!}\int_0^n f(\tilde{x}(t))(n-t)^r e^{\alpha(n-t)}\,\mathrm{d}t. \end{aligned}$$

If the states $u^{(k)}$ are defined as

$$u_n^{(k)} = \int_0^n f(\tilde{x}(t))(n-t)^k e^{\alpha(n-t)}\,\mathrm{d}t,$$

then

$$y_n = \frac{A}{r!}u_n^{(r)}. \tag{16}$$

From the previous equation:

$$\begin{aligned} u_n^{(k)} =& e^{\alpha}\int_0^{n-1} f(\tilde{x}(t))(n-t)^k e^{\alpha(n-1-t)}\,\mathrm{d}t \\ &+ \int_{n-1}^n f(\tilde{x}(t))(n-t)^k e^{\alpha(n-t)}\,\mathrm{d}t. \end{aligned}$$

Recalling that $a^k = \sum_{l=0}^k \binom{k}{l}(a-1)^l$ and applying the substitution $t \mapsto t+n-1$ to the last integral, we obtain

$$\begin{aligned} u_n^{(k)} =& e^{\alpha}\sum_{l=0}^k \binom{k}{l}\int_0^{n-1} f(\tilde{x}(t))(n-1-t)^l e^{\alpha(n-1-t)}\,\mathrm{d}t \\ &+ \int_0^1 f(x_{n-1}+t(x_n-x_{n-1}))(1-t)^k e^{\alpha(1-t)}\,\mathrm{d}t, \end{aligned}$$

i.e.,

$$\begin{aligned} u_n^{(k)} =& e^{\alpha}\sum_{l=0}^k \binom{k}{l}u_{n-1}^{(l)} \\ &+ \int_0^1 f(x_{n-1}+t(x_n-x_{n-1}))(1-t)^k e^{\alpha(1-t)}\,\mathrm{d}t. \end{aligned} \tag{17}$$

We conclude that the system is described by Eq. (16) and Eq. (17) for $k = 0, \ldots, r$. Notice that the definition of the sequence $u^{(k)}$ does not depend on r, so these states can be used for all summands in Eq. (5) relative to the same pole.

A.2. Complex case

Suppose now that $H(s) = \frac{B}{(s-\beta)^{r+1}} + \frac{\overline{B}}{(s-\overline{\beta})^{r+1}}$, where $B, \beta \in \mathbb{C}$, $\Re(\beta) < 0$ and $r > 0$. Then $h(t) = 2\frac{t^r}{r!}\Re(Be^{\beta t})u(t)$ and the computation goes similarly to the real case. More precisely, if we set

$$u_n^{(k)} = \int_0^n f(\tilde{x}(t))(n-t)^k e^{\beta(n-t)}\,\mathrm{d}t$$

for $k = 0, \ldots, r$, then

$$y_n = \frac{2}{r!}\Re(Bu_n^{(r)}) \tag{18}$$

and the $u^{(k)}$'s are updated according to:

$$\begin{aligned} u_n^{(k)} =& e^{\beta}\sum_{l=0}^k \binom{k}{l}u_{n-1}^{(l)} \\ &+ \int_0^1 f(x_{n-1}+t(x_n-x_{n-1}))(1-t)^k e^{\beta(1-t)}\,\mathrm{d}t. \end{aligned} \tag{19}$$