OPTIMAL INTEGER ORDER APPROXIMATION OF FRACTIONAL ORDER FILTERS

Pier Paolo La Pastina

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

Stefano D'Angelo

Independent researcher
Sessa Cilento, Italy
s@dangelo.audio

ABSTRACT

Fractional order filters have been studied since a long time, along with their applications to many areas of physics and engineering. In particular, several solutions have been proposed in order to approximate their frequency response with that of an ordinary filter. In this paper, we tackle this problem with a new approach: we solve analytically a simplified version of the problem and we find the optimal placement of poles and zeros, giving a mathematical proof and an error estimate. This solution shows improved performance compared to the current state of the art and is suitable for real-time parametric control.

1. INTRODUCTION

Fractional calculus is a classical branch of mathematical analysis that studies possible definitions and properties of derivatives of non-integer order. The origin of fractional derivative dates back to Leibniz [1] and a variety of different approaches have been proposed, including the ones from Abel [2], Riemann-Liouville [3], Riesz [4], Caputo [5] and Caputo-Fabrizio [6]. Apart from its theoretical importance, fractional calculus has been extensively studied in view of its applications in several areas of physics and engineering, such as system identification [7] and filter design [8]; recently it has also been applied to modeling of COVID-19 transmission [9]. An overview of fractional calculus from the signal processing point of view can be found in [10]: see references therein for more examples of applications.

Linear filters exhibit a spectral rolloff that is an integer multiple of -20 dB/decade. Fractional order filters remove this constraint: their steepness can be any real number. In audio DSP, they can be used to transform white noise into pink, blue or any "fractional" noise [11], which in turn can be used for dithering algorithms [12]. They have also been recently applied in physical modeling of brass instruments [13].

Unfortunately, it is not possible to reproduce the exact behaviour of fractional order filters in both ordinary analog circuits and DSP algorithms. Therefore, several attempts to approximate them with ordinary filters have been made so far. One of the most celebrated paper on this subject is [14], where the authors, for the first time to our knowledge, use equally spaced poles and equally spaced zeros on a log scale. The papers by Hélie [15, 16] and Smith and Smith [17] are more specifically designed for audio applications and propose approximations based on different optimality criteria.

Our work goes in the direction of [15]: its aim is approximating the amplitude response of a fractional order low-pass filter with that of an integer order filter in the sense of least squares, on a log-log scale. On the other hand, we use a different approach: after simplifying the problem, we find a closed-form expression for the best solution, give a mathematical proof of optimality and show that the error decays quadratically with the order of the approximating filter. Our formulas for poles and zeros show low computational cost, so they are also manageable in time-varying applications. Finally, they can be easily extended in order to obtain other filters (high-pass, band-pass, etc.) with arbitrary slopes.

The paper is organized as follows. Section 2 defines the precise setting of the problem, including the approximation used. Section 3 is the heart of the paper and contains the optimality proof. In Section 4 we evaluate our result and compare it to the current state of the art approach. Finally, conclusions are drawn in Section 5, where we also suggest ideas for further research.

2. PROBLEM STATEMENT

A *fractional order low-pass filter* is a linear filter whose transfer function is

$$H(s) = \frac{1}{\left(\frac{s}{\omega_c} + 1\right)^\alpha}, \tag{1}$$

where $\alpha > 0$ is the *order* of the filter and $\omega_c > 0$ is the *angular cutoff frequency* corresponding to a *cutoff frequency* $f_c = \frac{\omega_c}{2\pi}$. According to our definition

$$|H(j\omega_c)| = \frac{1}{2^{\alpha/2}}, \tag{2}$$

hence the gain of the filter at the cutoff frequency f_c is

$$|H(j\omega_c)|_{dB} = 20\log_{10}|H(j\omega_c)| \approx -3\alpha \text{ dB}. \tag{3}$$

We aim to develop an approximation of such a filter by another filter of order $\leq N$ whose transfer function is

$$\widehat{H}(s,\zeta,\rho) = \frac{\left(\frac{s}{\zeta_1} - 1\right)\left(\frac{s}{\zeta_2} - 1\right)\dots\left(\frac{s}{\zeta_N} - 1\right)}{\left(\frac{s}{\rho_1} - 1\right)\left(\frac{s}{\rho_2} - 1\right)\dots\left(\frac{s}{\rho_N} - 1\right)}, \tag{4}$$

where $\zeta = (\zeta_1, \dots, \zeta_N)$ are the zeros of $\widehat{H}$ and $\rho = (\rho_1, \dots, \rho_N)$ are the poles. All zeros and poles are non-zero. In our case the gain at dc is 1 and that is coherent with (1).

In this work we also assume that all zeros and poles are real, hence to ensure BIBO stability it must be $\rho_1, \dots, \rho_N < 0$. Finally, the original filter can be seen as the series of two filters, one of integer order and another of order between 0 and 1, so it is not restrictive to assume $\alpha \in (0, 1)$.

Poles and zeros will be chosen so that $\widehat{H}$ approximates H in terms of magnitude response. In order to evaluate and minimize

the error we choose to measure amplitudes in dB and to consider frequencies on the logaritmic scale. Hence, evaluating the transfer functions on the imaginary axis, i.e. substituting $s = j2\pi f$, and remapping frequencies on the logarithmic scale by $f = 10^x$, our problem can be formally expressed as choosing ρ_i and ζ_i so that the target magnitude response function

$$\begin{aligned}\eta(x) &= 20\log_{10}|H(2\pi j\cdot 10^x)| \\ &= -10\alpha\log_{10}\left(\left(\frac{2\pi}{\omega_c}\cdot 10^x\right)^2+1\right)\end{aligned} \tag{5}$$

is best approximated by

$$\begin{aligned}\widehat{\eta}(x,\zeta,\rho) =&20\log_{10}|\widehat{H}(2\pi j\cdot 10^x,\zeta,\rho)| \\ =&10\sum_{i=1}^{N}\left(\log_{10}\left(\left(\frac{2\pi}{\zeta_i}\cdot 10^x\right)^2+1\right)\right. \\ &\left.-\log_{10}\left(\left(\frac{2\pi}{\rho_i}\cdot 10^x\right)^2+1\right)\right).\end{aligned} \tag{6}$$

Here we notice that, for this purpose, the sign of ζ_i is irrelevant, so we assume $\zeta_i < 0$ to obtain minimum phase approximations. Setting

$$x_0 = \log_{10}\left(\frac{\omega_c}{2\pi}\right),\ z_i = \log_{10}\left(-\frac{\zeta_i}{2\pi}\right),\ p_i = \log_{10}\left(-\frac{\rho_i}{2\pi}\right), \tag{7}$$

the previous equations can be more conveniently expressed as

$$\begin{aligned}\eta(x) =& -10\alpha\log_{10}(10^{2(x-x_0)}+1), \\ \widehat{\eta}(x,z,p) =&10\sum_{i=1}^{N}(\log_{10}(10^{2(x-z_i)}+1) \\ &-\log_{10}(10^{2(x-p_i)}+1)),\end{aligned} \tag{8}$$

and the independent variables for the minimization procedure are now $z = (z_1, ..., z_N)$ and $p = (p_1, ..., p_N)$.

2.1. Error metric

The least squares method can arguably be considered the preferred choice for minimization problems such as ours. In this case the error metric to minimize is

$$E_\eta(z,p) = \int_{x_{\min}}^{x_{\max}}(\eta(x)-\widehat{\eta}(x,z,p))^2dx, \tag{9}$$

for a given range of interest $x \in [x_{\min}, x_{\max}]$.

Unluckily, we have not been able to minimize this quantity analytically so far. Instead, we were successful after further approximating η and $\widehat{\eta}$. Observe that

$$\log_{10}(1+10^{2\beta}) \approx 2\beta u(\beta), \tag{10}$$

where u is the Heaviside function. This is the classical approximation we use to draw Bode diagrams and is already very accurate for $|\beta| > 1$: we will call it *Bode approximation* from now on (see Figure 1). Applying (10) to (8), we obtain the new functions

$$\begin{aligned}\theta(x) &= -20\alpha(x-x_0)u(x-x_0), \\ \widehat{\theta}(x,z,p) &= 20\sum_{i=1}^{N}((x-z_i)u(x-z_i)-(x-p_i)u(x-p_i))\end{aligned} \tag{11}$$

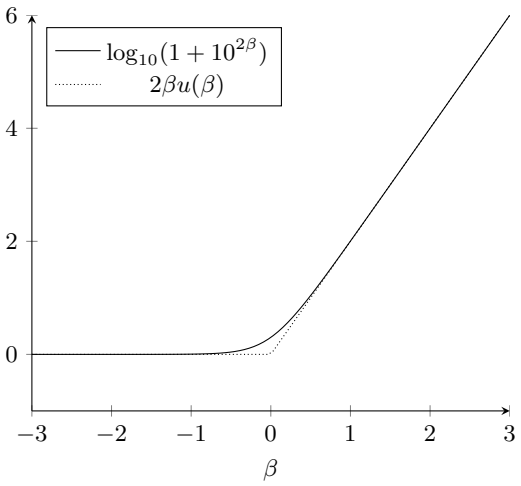

Figure 1: *The Bode approximation.*

and the new error to minimize:

$$E_\theta(z,p) = \int_{x_{\min}}^{x_{\max}}(\theta(x)-\widehat{\theta}(x,z,p))^2dx. \tag{12}$$

For minimization purposes we can safely consider $x_{\min} = x_0$ as the contribution to the error as expressed in (12) would be null for $x \in [x_0, x_{\min}]$, and conversely it would also be irrelevant if $x_0 < x_{\min}$. Clearly, we can suppose that $z_1 \le \cdots \le z_N$ and $p_1 \le \cdots \le p_N$. We also observe that, if $z_i = p_k$ for some $i, k \le N$, the corresponding terms in the previous sum cancel, so we obtain a filter of order lower than N.

Finally, we assume that

$$x_0 \le p_1 \le z_1 \le \cdots \le p_N \le z_N. \tag{13}$$

It is intuitively clear that the optimal curve should satisfy this property. Indeed, the slope of the amplitude response of a filter of order α is -20α dB/decade. If the approximating filter satisfies (13), the slope is alternatively 0 and -20 dB/decade; otherwise, the slope will be less than -20 dB/decade or even positive in some regions.

2.2. Optimal solution

Theorem. The absolute minimum point of (12) in the domain $D \subset \mathbb{R}^{2N}$ defined by

$$x_0 \le p_1 \le z_1 \le \cdots \le p_N \le z_N \le x_{\max} \tag{14}$$

is

$$\begin{aligned}\hat{p}_i &= x_0 + \frac{2i-1-\alpha}{2N+1-\alpha}(x_{\max}-x_0), \\ \hat{z}_i &= x_0 + \frac{2i-1+\alpha}{2N+1-\alpha}(x_{\max}-x_0),\end{aligned} \tag{15}$$

and the minimum value is

$$E_\theta(\hat{z},\hat{p}) = \frac{400}{3}\left(\frac{\alpha(1-\alpha)}{2N+1-\alpha}\right)^2(x_{\max}-x_0)^3. \tag{16}$$

An example of this approximation is shown in Figure 2. Our proposed filters are then obtained by plugging (15) into the original

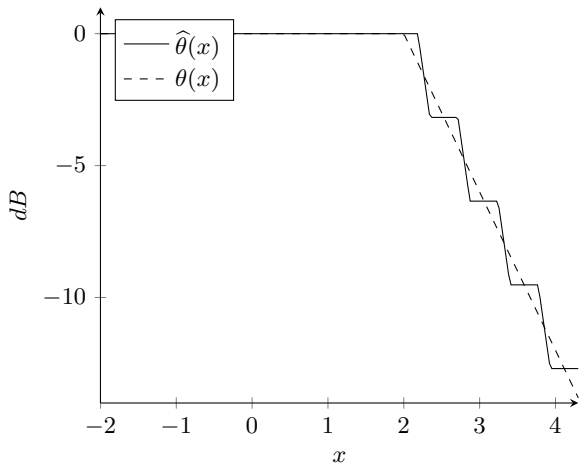

Figure 2: *Optimal solution to the simplified problem ($\alpha = 0.3$).*

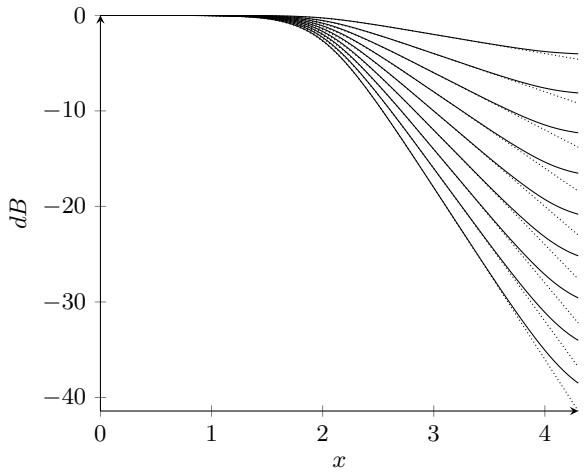

Figure 3: *Exact and approximated fractional order filters ($x_0 = 2$, $\alpha = 0.1, 0.2, \ldots, 0.9$ and $N = 4$).*

approximated magnitude response $\widehat{\eta}(x, z, p)$: Figure 3 shows the result of the approximation for different values of α.

The optimal poles $\hat{\rho}_i$ and zeros $\hat{\zeta}_i$ can be computed from (7) as:

$$\begin{aligned} \hat{\rho}_i &= -2\pi f_c \cdot \left(\frac{f_{\max}}{f_c}\right)^{\frac{2i-1-\alpha}{2N+1-\alpha}}, \\ \hat{\zeta}_i &= -2\pi f_c \cdot \left(\frac{f_{\max}}{f_c}\right)^{\frac{2i-1+\alpha}{2N+1-\alpha}}. \end{aligned} \tag{17}$$

2.3. Other filter types

Once we have built a good approximation of fractional order low-pass filters, it is easy to extend it to other types of fractional order filters.

A *fractional order high-pass filter* is a linear filter whose transfer function is

$$H_{HP}(s) = \left(\frac{\frac{s}{\omega_c}}{\frac{s}{\omega_c} + 1}\right)^{\alpha}. \tag{18}$$

As before, evaluating the transfer function on the imaginary axis and using a log-log scale, we obtain:

$$\begin{aligned} \eta_{HP}(x) &= 20 \log_{10} |H_{HP}(2\pi j \cdot 10^x)| \\ &= 10\alpha \log_{10} \left(\frac{10^{2(x-x_0)}}{10^{2(x-x_0)} + 1}\right) \end{aligned} \tag{19}$$

Now we notice that $\eta_{HP}(x)$ can be obtained from $\eta(x)$ in (8) via the reflection $x \mapsto 2x_0 - x$. In order to get the analogous approximation of the high-pass filter, we simply apply this reflection to (15), obtaining:

$$\begin{aligned} \hat{p}_i &= x_0 - \frac{2i - 1 - \alpha}{2N + 1 - \alpha}(x_0 - x_{\min}), \\ \hat{z}_i &= x_0 - \frac{2i - 1 + \alpha}{2N + 1 - \alpha}(x_0 - x_{\min}). \end{aligned} \tag{20}$$

The corresponding poles and zeroes are therefore:

$$\begin{aligned} \hat{\rho}_i &= -2\pi f_c \cdot \left(\frac{f_{\min}}{f_c}\right)^{\frac{2i-1-\alpha}{2N+1-\alpha}}, \\ \hat{\zeta}_i &= -2\pi f_c \cdot \left(\frac{f_{\min}}{f_c}\right)^{\frac{2i-1+\alpha}{2N+1-\alpha}}. \end{aligned} \tag{21}$$

A *fractional order band-pass filter* can be defined as a series of a fractional order low-pass and a fractional order high-pass filter, possibly of different orders. Accordingly, it can be approximated by the series of two filters of the form (4), where poles and zeroes are given respectively by (17) and (21). Inverse fractional order (low-pass and high-pass) filters are defined again by Equations (1) and (18), with $\alpha < 0$. It is clear that their integer order approximation can be constructed by simply swapping poles and zeroes in (17) and (21).

Finally, other filter types (such as low and high shelving) can be obtained as a combination of the filters above. However, our solution does not contemplate the possibility of placing poles or zeros on the imaginary axis and is bound to chosen frequency ranges which cannot contain dc, therefore it is not suitable for full-range approximations of fractional integrators and differentiators.

3. PROOF

Now we are ready to prove the theorem in Section 2. According to the discussion therein, $\widehat{\theta}$ has the form:

$$\widehat{\theta}(x,z,p)=\begin{cases}0 & x\le p_1\\ 20(\sum_{k=1}^{i}p_k-\sum_{k=1}^{i-1}z_k-x) & p_i\le x\le z_i\\ 20\sum_{k=1}^{i}(p_k-z_k) & z_i\le x\le p_{i+1}\\ 20\sum_{k=1}^{N}(p_k-z_k) & x\ge z_N\end{cases} \tag{22}$$

and we want to minimize (12) on the domain D. It is clear that the absolute minimum is among the critical points of $E(z,p)$ in the interior of D, or on the boundary of D. We will prove that there is a unique critical point of E in the interior of D and this is the absolute minimum.

Let us assume, by induction, that this is true for $(N-1)$-th order filters. First, we observe that the interior of D is given by points in $\mathbb{R}^{2N}$ that satisfy (14) with strict inequalities, while the boundary is given by points that satisfy (14) with some equality sign. Therefore, the boundary of D is the union of the regions defined by:

$$x_0=p_1\le z_1\le\cdots\le p_N\le z_N\le b, \tag{23}$$

$$x_0\le p_1\le z_1\le\cdots\le p_i=z_i\le\cdots\le p_N\le z_N\le x_{\max} \tag{24}$$

for $i=1,\dots,N$,

$$x_0\le p_1\le z_1\le\cdots\le z_i=p_{i+1}\le\cdots\le p_N\le z_N\le x_{\max} \tag{25}$$

for $i=1,\dots,N-1$ and

$$x_0\le p_1\le z_1\le\cdots\le p_N\le z_N=x_{\max}. \tag{26}$$

We notice that the central equations represent situations in which a zero coincides with a pole, so the order of the filter becomes $N-1$. By induction, we already know the absolute minimum on these regions of the boundary, so we are left with finding the minimum in the regions (23) and (26). We observe that the last region can be identified with the situation in which there are N poles and $N-1$ zeros, because the last zero is $x_{\max}$ and therefore has no influence on the region where we want to approximate the filter.

The regions (23) and (26) can be identified with domains of $\mathbb{R}^{2N-1}$, so we have to understand again if the minimum is in the interior or on the boundary. The boundary is given by points where some inequalities hold with the equality sign. We observe that, if some equality of the form $p_i=z_i$ or $z_i=p_{i+1}$ holds, we are again in a region of the form (24) or (25), where by induction we already know the minimum. If instead one of the extreme inequalities hold, we are in the region

$$x_0=p_1\le z_1\le\cdots\le p_N\le z_N=x_{\max}, \tag{27}$$

where we have to study the problem again.

We will proceed in the following way. First, we study the problem in the interior of (14), we find the critical point and we evaluate the error (12). This error will depend on N, so to show that it is smaller than the error in the regions (24) and (25) it suffices to verify that it decreases with N. Finally, we find the critical points in the regions (23), (26) and (27) and we compute the respective errors: once we have verified that they are greater than the error in the inner critical points, the proof is complete.

In the domain (14), Equation (12) assumes the form:

$$\begin{aligned}E_\theta(z,p)=&\int_{x_{\min}}^{x_{\max}}(\theta(x)-\widehat{\theta}(x,z,p))^2dx\\ =&400\Bigg\{\int_{x_0}^{p_1}\alpha^2(x-x_0)^2dx+\sum_{i=1}^{N}\int_{p_i}^{z_i}\Bigg(-\alpha(x-x_0)\\ &+x+\sum_{k=1}^{i-1}z_k-\sum_{k=1}^{i}p_k\Bigg)^2dx\\ &+\sum_{i=1}^{N-1}\int_{z_i}^{p_{i+1}}\Bigg(-\alpha(x-x_0)+\sum_{k=1}^{i}(z_k-p_k)\Bigg)^2dx\\ &+\int_{z_N}^{x_{\max}}\Bigg(-\alpha(x-x_0)+\sum_{i=1}^{N}(z_k-p_k)\Bigg)^2dx\Bigg\}\\ =&400\Bigg\{\sum_{i=1}^{N}\int_{p_i}^{z_i}\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}z_k-\sum_{k=1}^{i}p_k\Bigg)^2dx\\ &+\sum_{i=0}^{N}\int_{z_i}^{p_{i+1}}\Bigg(-\alpha(x-x_0)+\sum_{k=1}^{i}(z_k-p_k)\Bigg)^2dx\Bigg\}.\end{aligned} \tag{28}$$

For simplicity, we have set $z_0=x_0$ and $p_{N+1}=x_{\max}$.

We notice that E is a C^1 function, even if $\widehat{\theta}$ is not; in fact, it is a cubic polynomial in z and p. In order to compute partial derivatives of E, we use *Leibniz integral rule*:

$$\begin{aligned}\frac{\partial}{\partial x_i}\int_{\lambda(x)}^{\mu(x)}\psi(x,t)dt=&\psi(x,\mu(x))\frac{\partial\mu}{\partial x_i}\\ &-\psi(x,\lambda(x))\frac{\partial\lambda}{\partial x_i}+\int_{\lambda(x)}^{\mu(x)}\frac{\partial\psi}{\partial x_i}(x,t)dt.\end{aligned} \tag{29}$$

We have:

$$\begin{aligned}\frac{\partial E_\theta}{\partial z_l}=&400\Bigg\{\Bigg((1-\alpha)z_l+\alpha x_0+\sum_{k=1}^{l-1}z_k-\sum_{k=1}^{l}p_k\Bigg)^2\\ &+\sum_{i=l+1}^{N}\int_{p_i}^{z_i}2\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}z_k-\sum_{k=1}^{i}p_k\Bigg)dx\\ &-\Bigg(\alpha(z_l-x_0)-\sum_{k=1}^{l}(z_k-p_k)\Bigg)^2\\ &+\sum_{i=l}^{N}\int_{z_i}^{p_{i+1}}2\Bigg(-\alpha(x-x_0)+\sum_{k=1}^{i}(z_k-p_k)\Bigg)dx\Bigg\}\\ =&800\Bigg\{\sum_{i=l+1}^{N}\int_{p_i}^{z_i}\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}z_k-\sum_{k=1}^{i}p_k\Bigg)dx\\ &+\sum_{i=l}^{N}\int_{z_i}^{p_{i+1}}\Bigg(-\alpha(x-x_0)+\sum_{k=1}^{i}(z_k-p_k)\Bigg)dx\Bigg\}.\end{aligned} \tag{30}$$

Analogously, we get:

$$\begin{aligned}\frac{\partial E_\theta}{\partial p_l} = &-800\sum_{i=l}^{N}\Bigg\{\int_{p_i}^{z_i}\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}z_k\\ &-\sum_{k=1}^{i}p_k\Bigg)dx+\int_{z_i}^{p_{i+1}}\Bigg(-\alpha(x-x_0)\\ &+\sum_{k=1}^{i}(z_k-p_k)\Bigg)dx\Bigg\}.\end{aligned} \tag{31}$$

Now, if we set

$$\begin{aligned}A_i &= \int_{p_i}^{z_i}\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}z_k-\sum_{k=1}^{i}p_k\Bigg)dx;\\ B_i &= \int_{z_i}^{p_{i+1}}\Bigg(-\alpha(x-x_0)+\sum_{k=1}^{i}(z_k-p_k)\Bigg)dx,\end{aligned} \tag{32}$$

it is clear that $\nabla E_\theta(z,p)=0$ if and only if

$$\begin{aligned}\sum_{i=l+1}^{N}A_i+\sum_{i=l}^{N}B_i=0;\\ \sum_{i=l}^{N}A_i+\sum_{i=l}^{N}B_i=0.\end{aligned} \tag{33}$$

These equations are equivalent to

$$A_1=\cdots=A_N=B_1=\cdots=B_N=0. \tag{34}$$

Computing the integrals (32), we get:

$$\begin{aligned}&\frac{1-\alpha}{2}(z_i^2-p_i^2)+\Bigg(\alpha x_0+\sum_{k=1}^{i-1}z_k-\sum_{k=1}^{i}p_k\Bigg)(z_i-p_i)=0;\\ &-\frac{\alpha}{2}(p_{i+1}^2-z_i^2)+\Bigg(\alpha x_0+\sum_{k=1}^{i}(z_k-p_k)\Bigg)(p_{i+1}-z_i)=0.\end{aligned} \tag{35}$$

We are searching for the critical points in the interior of D, so $p_i<z_i<p_{i+1}$. This means that we can divide the first equation by z_i-p_i and find z_i:

$$z_i=\frac{2}{1-\alpha}\Bigg(\sum_{k=1}^{i}p_k-\sum_{k=1}^{i-1}z_k-\alpha x_0\Bigg)-p_i. \tag{36}$$

Analogously, we divide the second equation by $p_{i+1}-z_i$ and we find p_{i+1}:

$$p_{i+1}=\frac{2}{\alpha}\Bigg(\sum_{k=1}^{i}(z_k-p_k)+\alpha x_0\Bigg)-z_i \tag{37}$$

Now we prove by induction that

$$p_i=x_0+\frac{2i-1-\alpha}{1-\alpha}(p_1-x_0) \tag{38}$$

and

$$z_i=x_0+\frac{2i-1+\alpha}{1-\alpha}(p_1-x_0). \tag{39}$$

For $i=1$, (38) is obvious and (39) follows directly from (36). Suppose then that $1<i\le N+1$ and Equations (38) and (39) hold for every index below i. Then:

$$\begin{aligned}p_i =&\frac{2}{\alpha}\Bigg(\sum_{k=1}^{i-1}(z_k-p_k)+\alpha x_0\Bigg)-z_{i-1}\\ =&\frac{2}{\alpha}\Bigg(\sum_{k=1}^{i-1}\frac{2\alpha}{1-\alpha}(p_1-x_0)+\alpha x_0\Bigg)\\ &-x_0-\frac{2i-3+\alpha}{1-\alpha}(p_1-x_0)\\ =&\frac{4(i-1)}{1-\alpha}(p_1-x_0)+2x_0-x_0-\frac{2i-3+\alpha}{1-\alpha}(p_1-x_0)\\ =&x_0+\frac{2i-1-\alpha}{1-\alpha}(p_1-x_0),\end{aligned} \tag{40}$$

as desired. In the same way you can prove (39) for $1\le i\le N$.

Equation (38) for $i=N+1$ becomes

$$x_{\max}=x_0+\frac{2N+1-\alpha}{1-\alpha}(p_1-x_0), \tag{41}$$

so

$$p_1-x_0=\frac{1-\alpha}{2N+1-\alpha}(b-x_0). \tag{42}$$

Substituting this equation in (38) and (39) we get

$$p_i=x_0+\frac{2i-1-\alpha}{2N+1-\alpha}(b-x_0), \tag{43}$$

$$z_i=x_0+\frac{2i-1+\alpha}{2N+1-\alpha}(b-x_0). \tag{44}$$

Let us denote these values by $(\hat{z},\hat{p})$ and compute the corresponding error. We notice that formula (43) also holds for $p_{N+1}=x_{\max}$, while (44) does not hold for $z_0=x_0$. So, setting $\hat{p}_{N+1}=x_{\max}$, we obtain:

$$\begin{aligned}E_\theta(\hat{z},\hat{p}) =&400\Bigg\{\int_{x_0}^{\hat{p}_1}\Big(-\alpha(x-x_0)\Big)^2dx\\ &+\sum_{i=1}^{N}\int_{\hat{p}_i}^{\hat{z}_i}\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}\hat{z}_k-\sum_{k=1}^{i}\hat{p}_k\Bigg)^2dx\\ &+\sum_{i=1}^{N}\int_{\hat{z}_i}^{\hat{p}_{i+1}}\Big(-\alpha(x-x_0)+\sum_{k=1}^{i}(\hat{z}_k-\hat{p}_k)\Big)^2dx\Bigg\}.\end{aligned} \tag{45}$$

The first integral is trivial:

$$\int_{x_0}^{\hat{p}_1}\Big(-\alpha(x-x_0)\Big)^2dx=\frac{\alpha^2}{3}\Bigg(\frac{(1-\alpha)(x_{\max}-x_0)}{2N+1-\alpha}\Bigg)^3. \tag{46}$$

We solve the second integral in (45) by the substitution

$$\xi=(1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}\hat{z}_k-\sum_{k=1}^{i}\hat{p}_k, \tag{47}$$

and we get:

$$\begin{aligned}&\int_{\hat{p}_i}^{\hat{z}_i}\Bigg((1-\alpha)x+\alpha x_0+\sum_{k=1}^{i-1}\hat{z}_k-\sum_{k=1}^{i}\hat{p}_k\Bigg)^2dx\\ &\qquad=\frac{2}{3}(1-\alpha)^2\Bigg(\frac{\alpha(x_{\max}-x_0)}{2N+1-\alpha}\Bigg)^3.\end{aligned} \tag{48}$$

The third integral in (45) can be computed in the same way:

$$\int_{\hat{z}_i}^{\hat{p}_{i+1}} \left(-\alpha(x-x_0) + \sum_{k=1}^{i} (\hat{z}_k - \hat{p}_k) \right)^2 dx = \frac{2}{3}\alpha^2 \left(\frac{(1-\alpha)(x_{\max}-x_0)}{2N+1-\alpha} \right)^3. \tag{49}$$

Finally,

$$\begin{aligned} E_\theta(\hat{z},\hat{p}) =& 400 \left\{ \frac{2N}{3}(1-\alpha)^2 \left(\frac{\alpha(x_{\max}-x_0)}{2N+1-\alpha} \right)^3 \right. \\ &\left. + \frac{2N+1}{3}\alpha^2 \left(\frac{(1-\alpha)(x_{\max}-x_0)}{2N+1-\alpha} \right)^3 \right\} \\ =& \frac{400}{3} \left(\frac{\alpha(1-\alpha)}{2N+1-\alpha} \right)^2 (x_{\max}-x_0)^3. \end{aligned} \tag{50}$$

We observe that E_θ decreases with $\frac{1}{N^2}$ as desired, in particular it goes to 0 for $N \to \infty$.

Now we find the critical points in the regions (23), (26) and (27) and the corresponding errors. The computation is formally analogous to the one we performed for the interior of D, so we only give the results. In the interior of (23), the only critical point is

$$\begin{aligned} p_i &= x_0 + \frac{2(i-1)-\alpha}{2N-\alpha}(x_{\max}-x_0) \quad i = 2,\dots,N, \\ z_i &= x_0 + \frac{2(i-1)+\alpha}{2N-\alpha}(x_{\max}-x_0) \quad i = 1,\dots,N \end{aligned} \tag{51}$$

and the corresponding error is

$$\frac{400}{3} \left(\frac{\alpha(1-\alpha)}{2N-\alpha} \right)^2 (x_{\max}-x_0)^3. \tag{52}$$

In the interior of (26) the only critical point is

$$\begin{aligned} p_i &= x_0 + \frac{2i-1-\alpha}{2N-1+\alpha}(x_{\max}-x_0) \quad i = 1,\dots,N, \\ z_i &= x_0 + \frac{2i-1+\alpha}{2N-1+\alpha}(x_{\max}-x_0) \quad i = 1,\dots,N-1 \end{aligned} \tag{53}$$

and the error is

$$\frac{400}{3} \left(\frac{\alpha(1-\alpha)}{2N-1+\alpha} \right)^2 (x_{\max}-x_0)^3. \tag{54}$$

Finally, in the interior of (27) the only critical point is

$$\begin{aligned} p_i &= x_0 + \frac{2(i-1)-\alpha}{2(N-1)+\alpha}(x_{\max}-x_0) \quad i = 2,\dots,N, \\ z_i &= x_0 + \frac{2(i-1)+\alpha}{2(N-1)+\alpha}(x_{\max}-x_0) \quad i = 1,\dots,N-1 \end{aligned} \tag{55}$$

and the error is

$$\frac{400}{3} \left(\frac{\alpha(1-\alpha)}{2(N-1)+\alpha} \right)^2 (x_{\max}-x_0)^3. \tag{56}$$

But $2N-\alpha$, $2N-1+\alpha$ and $2(N-1)+\alpha$ are all smaller than $2N+1-\alpha$, so we conclude that (50) is the absolute minimum of $E_\theta(z,p)$.

4. EVALUATION

Equation (17) gives explicit expressions for pole and zero placement which are shown to minimize the approximated error (12). Such expressions are relatively easy to compute, hence suitable for real-time usage, and fully parametric: they depend on the original fractional filter order α, the chosen approximating filter order N, the maximum frequency in the range of interest $f_{\max}$, and the cutoff frequency f_c, which can also be considered as the minimum frequency in the range of interest as pointed out in Subsection 2.1.

Even when α approaches the limit values 0 and 1, we get the correct result. For $\alpha = 0$, the filter (1) reduces to an identity filter (the output coincides with the input). If α goes to 0 in (15), we get $\hat{p}_i = \hat{z}_i$ for every i, so every pole coincides with the following zero and we get the identity filter. For $\alpha = 1$, the filter (1) becomes a first-order filter. If α goes to 1 in (15), we get $\hat{z}_i = \hat{p}_{i+1}$ for $i = 1,\dots,N-1$, so all these zeros and poles cancel. Only the first pole $x_0 + \frac{x_{\max}-x_0}{2N-1}$ and the last zero $x_{\max}$ do not get canceled, and the latter has no effect in the range $[x_{\min}, x_{\max}]$. The first pole does not coincide with the pole x_0 of the original filter, but it goes to x_0 as N goes to infinity.

4.1. Analytic and numerical solution

Equation (15) gives an analytic expression for the minimum of (12), which approximates (9), but this result can be further improved by means of numerical optimization. In this subsection, we choose two instances of the problem and we compare, for each of them, the frequency response of (1) with the ones obtained by plugging both (15) and its numerical improvement in (4). In both cases we run the optimization on the interval $[x_{\min}, x_{\max}] = [\log_{10} 2 - 2, \log_{10} 2 + 4]$, which corresponds to $[f_{\min}, f_{\max}] = [0.02\text{ Hz}, 20000\text{ Hz}]$. This range is suitable if the cutoff frequency varies in the whole spectrum 20 Hz-20 kHz of audible frequencies: indeed, the passband is at least as wide as the stopband, so the numerical result is reasonably accurate in both parts of the spectrum. We stress that this choice is arbitrary, the most suitable range depends on the application at hand. A GNU/Octave implementation using the `leasqr` function in the `optim` Octave Forge package for the numerical optimization can be found at `http://www.dangelo.audio/dafx2021-fracfilt.html`.

The first instance we consider is $x_0 = 2$, $\alpha = 0.3$, $N = 4$. In Table 1 we report the analytic and the numerical values for the z_i's and the p_i's, while in Figure 4 we plot the error $\widehat{\eta}(x) - \eta(x)$ in the two cases.

Table 1: *Values obtained for* $x_0 = 2$, $\alpha = 0.3$, $N = 4$.

	Analytic	Numerical
z_1	2.343832068317607	2.279616275567339
z_2	2.872804481113924	2.881335053060714
z_3	3.401776893910242	3.526431869946629
z_4	3.930749306706559	4.244046225353371
p_1	2.185140344478711	2.137606600592802
p_2	2.714112757275029	2.691901860686028
p_3	3.243085170071347	3.331052735238103
p_4	3.772057582867664	4.004548535923683

We do the same for the second instance: $x_0 = 3$, $\alpha = 0.8$, $N = 5$. Results are given in Table 2 and Figure 5.

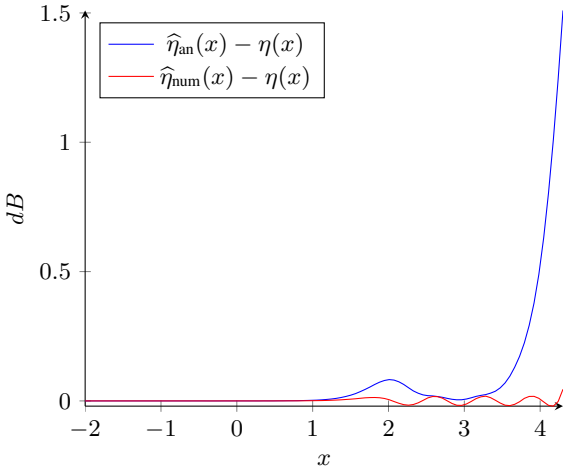

Figure 4: *Frequency responses for* $x_0 = 2$, $\alpha = 0.3$, $N = 4$.

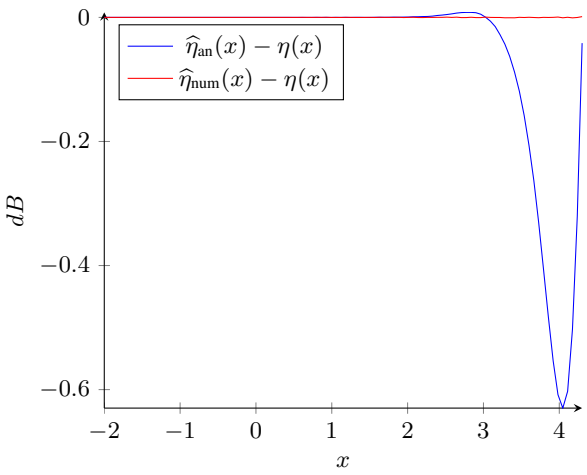

Figure 5: *Frequency responses for* $x_0 = 3$, $\alpha = 0.8$, $N = 5$.

Table 2: *Values obtained for* $x_0 = 3$, $\alpha = 0.8$, $N = 5$.

	Analytic	Numerical
z_1	3.229593528646585	3.148247223022702
z_2	3.484697449365012	3.439993019764529
z_3	3.739801370083440	3.784192323454005
z_4	3.994905290801868	4.176617234105366
z_5	4.250009211520296	4.904210559694439
p_1	3.025510392071843	3.011234625407794
p_2	3.280614312790270	3.199230918296403
p_3	3.535718233508698	3.506038645221784
p_4	3.790822154227126	3.857245939142739
p_5	4.045926074945553	4.269913348230429

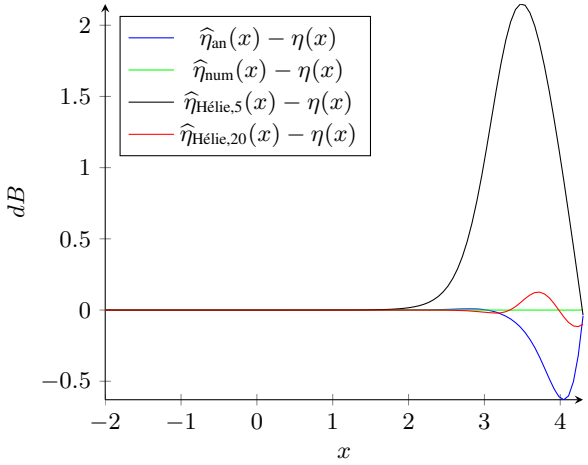

Figure 6: *Comparison with [15].*

From the plots we observe that the analytic solution is rather accurate, except for the highest frequencies: in particular, in the first example the error $|\widehat{\eta}_{\text{an}}(x) - \eta(x)|$ goes up to 1.5 dB. Indeed, from Table 1 and 2 one can notice that the last zeros and poles are sensibly higher in the numerical solution. To address this issue, one can use Equation (15) with a higher value of $x_{\max}$: this will give a less accurate solution in the range $[x_0, x_{\max}]$, but will reduce the absolute error near $x_{\max}$.

4.2. Current state of the art

Now we compare our solution (15) with that proposed in [15]. Firstly, we notice that our formulas are remarkably simpler and easier to compute, which is especially desirable in real-time and time-varying applications. Then, Hélie's approximating filter (7) in [15] does not exhibit unitary dc gain, so for this comparison we added dc compensation. We compare Hélie's filter of order 5 and 20 with our analytic and numerical solutions of order 5, for $x_0 = 3$ and $\alpha = 0.8$ (second example of Subsection 4.1) in Figure 6.

We notice that our approach leads to sensibly reduced error when using 5 poles, while the solution in [15] needs 20 poles to produce comparable results. This is probably due to the fact that in [15] pole frequencies are chosen beforehand, while in our case pole and zero frequencies are subject to optimization.

5. CONCLUSIONS AND FUTURE WORK

This paper finds the best way to approximate a fractional order low-pass filter with an ordinary filter, after a natural simplification of the problem. Compared to the current state of the art, our approach results in both remarkably simpler formulas, which facilitates real-time control in the time-varying case, and reduced error at the same filter order, or equivalently a sensible reduction of filter complexity to reach a given performance target. Applying numerical optimization to the pole and zero frequencies given by our method further improves the accuracy of the approximating filter.

Our results could be further improved by allowing the usage of complex zeros and poles in the approximating filter or by studying the problem without the Bode approximation (10). Finally, the most general formulation of the problem should also take into account approximation of the phase response.

6. ACKNOWLEDGMENTS

We owe thanks to Julius O. Smith for fruitful discussions and for suggesting us some applications and future directions.

7. REFERENCES

[1] G. W. Leibniz, "Letter from Hanover, Germany, to GFA L'Hopital, September 30; 1695," *Mathematische Schriften*, vol. 2, pp. 301–302, 1849.

[2] N. H. Abel, "Oplösning af et par opgaver ved hjelp af bestemte integraler," *Magazin for naturvidenskaberne*, vol. 2, no. 55, pp. 2, 1823.

[3] B. Riemann, "Versuch einer allgemeinen Auffassung der Integration und Differentiation," *Gesammelte Werke*, vol. 62, no. 1876, 1876.

[4] M. Riesz, "L'intégrale de Riemann-Liouville et le problème de Cauchy," *Acta mathematica*, vol. 81, pp. 1–222, 1949.

[5] M. Caputo, "Linear models of dissipation whose Q is almost frequency independent—II," *Geophysical Journal International*, vol. 13, no. 5, pp. 529–539, 1967.

[6] M. Caputo and M. Fabrizio, "A new definition of fractional derivative without singular kernel," *Progr. Fract. Differ. Appl*, vol. 1, no. 2, pp. 1–13, 2015.

[7] S. Das, *Functional fractional calculus*, Springer Science & Business Media, 2011.

[8] A. G. Radwan, A. M. Soliman, and A. S. Elwakil, "First-order filters generalized to the fractional domain," *Journal of Circuits, Systems, and Computers*, vol. 17, no. 01, pp. 55–66, 2008.

[9] S. Ahmad, A. Ullah, Q. M. Al-Mdallal, H. Khan, K. Shah, and A. Khan, "Fractional order mathematical modeling of COVID-19 transmission," *Chaos, Solitons & Fractals*, vol. 139, pp. 110256, 2020.

[10] R. Magin, M. D. Ortigueira, I. Podlubny, and J. Trujillo, "On the fractional signals and systems," *Signal Processing*, vol. 91, no. 3, pp. 350–371, 2011.

[11] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian motions, fractional noises and applications," *SIAM review*, vol. 10, no. 4, pp. 422–437, 1968.

[12] R. A. Ulichney, "Dithering with blue noise," *Proceedings of the IEEE*, vol. 76, no. 1, pp. 56–79, 1988.

[13] R. L. Harrison, S. Bilbao, and J. Perry, "An algorithm for a valved brass instrument synthesis environment using finite-difference time-domain methods with performance optimisation," in *Proceedings of the 18th International Conference on Digital Audio Effects, Trondheim, Norway*, 2015.

[14] A. Oustaloup, F. Levron, B. Mathieu, and F. M. Nanot, "Frequency-band complex noninteger differentiator: characterization and synthesis," *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 47, no. 1, pp. 25–39, 2000.

[15] T. Hélie, "Real-time simulation of a family of fractional-order low-pass filters," in *Audio Engineering Society Convention 135*. Audio Engineering Society, 2013.

[16] T. Hélie, "Simulation of fractional-order low-pass filters," *IEEE/ACM Transactions on audio, speech, and language processing*, vol. 22, no. 11, pp. 1636–1647, 2014.

[17] J. O. Smith and H. F. Smith, "Closed form fractional integration and differentiation via real exponentially spaced pole-zero pairs," *arXiv preprint arXiv:1606.06154*, 2016.