DESIGN PRINCIPLES FOR LUMPED MODEL DISCRETISATION USING MÖBIUS TRANSFORMS

François G. Germain, Kurt J. Werner

Center for Computer Research in Music and Acoustics (CCRMA)
Stanford University, Stanford, CA, USA
fgermain@stanford.edu, kwerner2@stanford.edu

ABSTRACT

Computational modelling of audio systems commonly involves discretising lumped models. The properties of common discretisation schemes are typically derived through analysis of how the imaginary axis on the Laplace-transform s-plane maps onto the Z-transform z-plane and the implied stability regions. This analysis ignores some important considerations regarding the mapping of individual poles, in particular the case of highly-damped poles. In this paper, we analyse the properties of an extended class of discretisations based on Möbius transforms, both as mappings and discretisation schemes. We analyse and extend the concept of frequency warping, well-known in the context of the bilinear transform, and we characterise the relationship between the damping and frequencies of poles in the s- and z-planes. We present and analyse several design criteria (damping monotonicity, stability) corresponding to desirable properties of the discretised system. Satisfying these criteria involves selecting appropriate transforms based on the pole structure of the system on the s-plane. These theoretical developments are finally illustrated on a diode clipper nonlinear model.

1. INTRODUCTION

Computational modelling of audio systems is a major topic of interest, including emulation of existing electronic and acoustic systems such as vintage audio effects or acoustic instruments. When the dynamics of a physical system are known through its transfer function, a common procedure is to discretise it using the bilinear transform or the forward or backward Euler method.

The general properties of those discretisations are often derived from analysing the mapping of the imaginary axis in the Laplace transform s-plane onto the Z-transform z-plane [1]. Since these mappings fall into the category of conformal mappings, the imaginary axis always maps to a circle or a line [2]. Additional analysis of these transforms can be found in [3], where a method to design digital constant-Q filters is presented using hybrid transforms between the bilinear transform and the backward Euler method. More generally, in the case where a state-space representation of the system is available, all those methods correspond to different numerical schemes [1, 4], which are well-studied methods in the field of numerical analysis [5]. In particular, those methods are well-studied in terms of numerical accuracy and stability as a function of the system dynamics and the chosen sampling rate for linear time-invariant (LTI) systems.

In the audio literature, little attention has been given to the simulation distortion introduced by those transforms due to the distortions of pole properties. However, the analysis of the mapping of the imaginary axis can offer only guarantees regarding the response of LTI systems in a steady-state context. For example, the magnitude response of an LTI system after discretisation with the bilinear transform can be shown to correspond exactly to the magnitude response of the original system up to a contraction and distortion of the imaginary axis [4, 6]. In cases where that frequency distortion is problematic, frequency warping methods can be used with the bilinear transform to compensate for it [4, 7]. However, warping breaks the equivalence between the bilinear transform and its numerical equivalent, the trapezoidal method, adding additional errors terms to the transform to achieve frequency matching. Additionally, warping requires previous knowledge of the system to know the needed level of warping. A method was also proposed in order to derive digital parametric equalizer designs with a non-zero Nyquist-frequency gain and better match the non-warped response of analog equalizers [8]. Those techniques do not consider the distortion introduced in the pole locations, which can produce noticeable differences in the system response for non-steady state conditions (e.g. transients, modulated input). Similar observations have been made in the context of artificial reverberations regarding the noticeable influence of all-pass filters on the short-term audio colouration of those effects despite their flat magnitude response [9]. Additionally, one must remain careful when considering the properties of those methods for discretising nonlinear and/or time-varying systems. Other typical filter designs (e.g. impulse-invariant methods [10], matched Z-transform [11], Prony's method [7]) based on exact pole placement can improve the transient behaviour of LTI systems, but they are typically impractical in the context of nonlinear and/or time-varying systems.

More advanced systematic discretisation schemes (e.g. high-order methods, fractional bilinear transform) are described in the numerical analysis literature [5], or in the audio context, in studies on physical-modelling-based synthesis [12, 13] or filter design [3, 14]. Those methods can often handle systems where the bilinear transform and the Euler methods are limited (e.g. strongly nonlinear systems). However, the bilinear transform and the Euler methods have been historically preferred as they preserve the system order, allowing for compact system representation and efficient computation. Backward Euler and the bilinear transform also have the property of unconditionally preserving stability and minimum-phase properties [1, 13]. This proves useful as digital audio effects work with a prescribed sampling frequency. Another aspect to consider is the limiting of undesirable transient behaviour when several systems are cascaded, more so if those systems are nonlinear and/or time-varying. In this case, the bilinear transform and the Euler methods are generally preferred to filter design methods based on exact pole placement, which can be sensitive to the way the system is implemented, even for LTI systems [1].

One case where the bilinear transform and the Euler methods are limited is systems with highly-damped poles. Numerical analysis literature shows that discretisation of such systems using the trapezoidal rule shows high-frequency oscillations at its output in transient mode [15, 16]. Indeed, the mapping of pole contours with identical damping in the s-plane, while still circles, are not centred around the origin, meaning their damping in the z-plane is not frequency-independent [3]. As a result, the properties of a mapped pole in the z-plane depend on the pole damping in a way that can noticeably impact the dynamics of the discretised system transient response. Methods such as frequency warping cannot compensate efficiently for this despite their use of a priori information of the system pole locations in the s-plane. This information can be available as, for example, we can often know where all the possible poles of a parametrised equalizer, or all the instantaneous poles of a nonlinear analog circuit will lie across all the operating conditions of those systems. However, with such information, we can imagine a generalisation of the bilinear transform and the Euler methods that could enforce a wider range of desirable properties for a discretised system (e.g. stability, pole damping and frequency) under all its known operating conditions.

In Sec. 2, we present a theoretical extension and generalised analysis of typical discretisations (e.g. bilinear transform) in the context of the Möbius transforms. In Sec. 3, we generalise the analysis of s-plane distortion introduced by discretisation and propose criteria in order to design discretisation schemes alleviating typical unwanted behaviour. In Sec. 4, we apply those principles to the discretisation of a typical audio system.

2. DISCRETISATIONS AS MÖBIUS TRANSFORMS

2.1. Lumped models

Audio systems can often be represented as single-input single-output (SISO) LTI lumped models (e.g. analog filters) with input $u(t)$ and output $y(t)$ (with Laplace transforms $U(s)$ and $Y(s)$) can be described by their transfer function in the Laplace domain [1]:

$$H(s) = \frac{Y(s)}{U(s)} = \frac{\sum_{m=0}^{M} b_m s^m}{\sum_{n=0}^{N} a_n s^n}. \tag{1}$$

This generalises to multiple-input multiple-output (MIMO) systems with the state-space representation

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{2}$$

and state variables $\mathbf{x}$. The generalised system transfer function

$$\mathbf{H}(s) = \mathbf{C}\left(s\mathbf{I} - \mathbf{A}\right)^{-1}\mathbf{B} + \mathbf{D} \tag{3}$$

has poles corresponding to the eigenvalues of $\mathbf{A}$ [6].

Some nonlinear time-varying systems can also be described in a state-space form as

$$\dot{\mathbf{x}}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t)) \tag{4a}$$

$$\mathbf{y}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t)). \tag{4b}$$

For a given operating point $(t_0, \mathbf{x}_0, \mathbf{u}_0)$, we linearise the system around that point using a small perturbation approximation as the state-space model

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}_0\mathbf{x}(t) + \mathbf{B}_0\mathbf{v}_0(t) \\ \mathbf{y}(t) &= \mathbf{C}_0\mathbf{x}(t) + \mathbf{D}_0\mathbf{v}_0(t) \end{aligned} \tag{5}$$

with $\mathbf{v}_0(t) = \left[(\mathbf{u}(t) - \mathbf{u}_0)^T, \quad (t - t_0), \quad \mathbf{x}_0^T\right]^T$, in which case we interpret the eigenvalues of $\mathbf{A}_0$ as "instantaneous" poles [13].

2.2. Transfer function discretisation and Möbius transforms

Computational simulation of LTI systems typically requires digitising their transfer function $\mathbf{H}(s)$ (Eq. (3)) at a sampling interval T, and several methods have been proposed in the literature [1]. Methods such as the bilinear transform (BT), backward Euler (BE), and forward Euler (FE) have the advantage of being simple, order-preserving, aliasing-free and independent of the transfer function form. Those methods correspond to the substitution of s by $\mathcal{T}(z)$ in the transfer function. They can be interpreted as mappings $z \mapsto s = \mathcal{T}(z)$ (and its inverse $s \mapsto z = \mathcal{T}^{-1}(s)$) between the Laplace transform s-plane and the Z-transform z-plane, with

$$\mathcal{T}_{\mathrm{BT}}(z) = \tfrac{2}{T}\tfrac{1-z^{-1}}{1+z^{-1}}, \mathcal{T}_{\mathrm{BE}}(z) = \tfrac{1-z^{-1}}{T}, \mathcal{T}_{\mathrm{FE}}(z) = \tfrac{1-z^{-1}}{Tz^{-1}}. \tag{6}$$

A point in the s-plane is decomposed as $s = \sigma + j\Omega$ with σ defined as the *damping* and Ω as the *frequency* of that point. A point in the z-plane is decomposed as $z = re^{j\omega}$ ($r \geq 0$ and $\omega \in]-\pi, \pi]$) with $\log(r)/T$ defined as the *damping* and ω/T as the *frequency* of that point.

These mappings belong to the class of *conformal mappings* (i.e. angle-preserving transforms) [2]. They also belong to the subclass of *Möbius transforms* which contains all rational order-preserving mappings. Möbius transforms have the form [17]

$$\mathcal{T}(z) = \frac{a + bz^{-1}}{c + dz^{-1}} \text{ and } \mathcal{T}^{-1}(s) = -\frac{ds - b}{cs - a} \tag{7}$$

with $ad - bc \neq 0$. They map circles (including lines as circles of infinite radius) in the origin plane to circles (including lines) in the target plane [2]. Möbius transform coefficients $(a, b, c, d) \in \mathbb{C}^4$ are defined up to common multiplying factor $\gamma \in \mathbb{C}$ since $(\gamma a, \gamma b, \gamma c, \gamma d)$ corresponds to the same transform. A transform is uniquely defined (up to factor γ) by setting the mapping of 3 separate points.

In digital filter design, filters with real coefficients (or equivalently with a conjugate symmetric Fourier transform with respect to DC) are typically preferred. This corresponds to the use of symmetric mappings with respect to the real axis (or equivalently which maps the real axis onto itself). That property is guaranteed if and only if the transform coefficients (a, b, c, d) are all real up to a common multiplying factor γ (possibly complex).

Also, we can note that, in both the s-plane and the z-plane, the upper part of the plane (i.e. points with positive imaginary part $\Im(\cdot)$) corresponds to points with positive frequency ($\Omega \geq 0$ and $\omega \geq 0$). For this reason, we limit the discussion to transformations mapping the upper plane of the s-plane to the upper plane to the z-plane. For real coefficients (a, b, c, d), we get:

$$\Im(z) = \frac{\Im(s)(ad - bc)}{\Im(s)^2c^2 + (a - \Re(s)c)^2} \tag{8}$$

so that $\Im(s) \geq 0 \Leftrightarrow \Im(z) \geq 0$ is true if and only if $ad - bc > 0$.

Another possible condition of interest is that the origin $s = 0$ maps to a DC point in the z-plane (i.e. with $\omega = 0$); this property holds if and only if a and b have opposite signs. We may also want to consider only mappings such that $|\Omega| \to \infty$ maps to $\omega \to \pm\pi$; this property holds if and only if c and d have same signs.

In the rest of the paper, we consider only Möbius transforms with real coefficients verifying $ad - bc > 0$. By construction, these transforms are all order-preserving and aliasing-free.

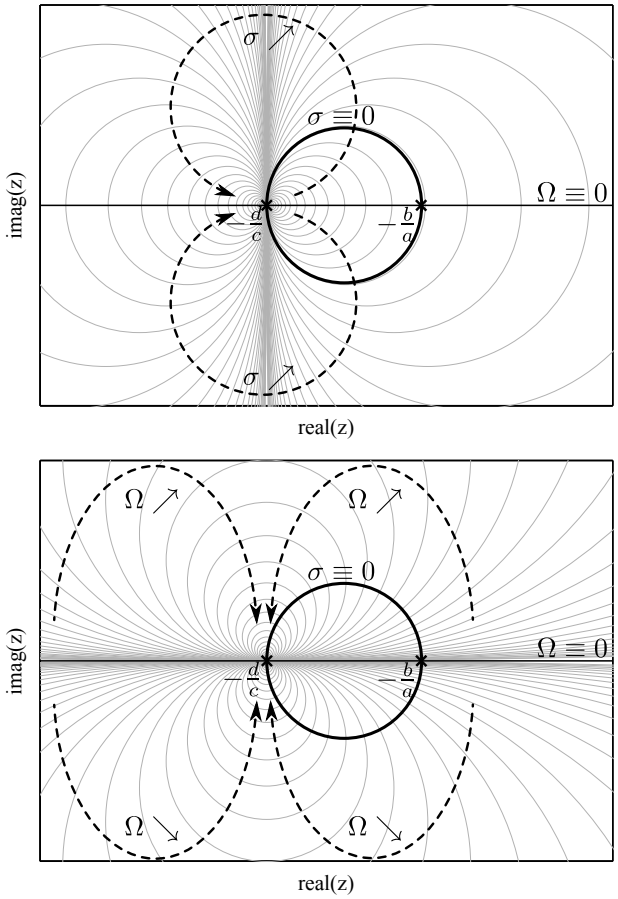


Figure 1: *Circles (in grey) corresponding to constant σ (top) and constant Ω (bottom) in the z-plane. The black horizontal axis corresponds to the real axis in the z-plane, mapped from the real axis $\Omega \equiv 0$. The black circle corresponds to the region mapped from the imaginary axis $\sigma \equiv 0$. In the case of the bilinear transform, the circle correspond to the unit circle, with $d/c = 1$ and $b/a = -1$. When following the dashed arrows, $\nearrow$ indicates increasing quantities and $\searrow$ decreasing ones.*

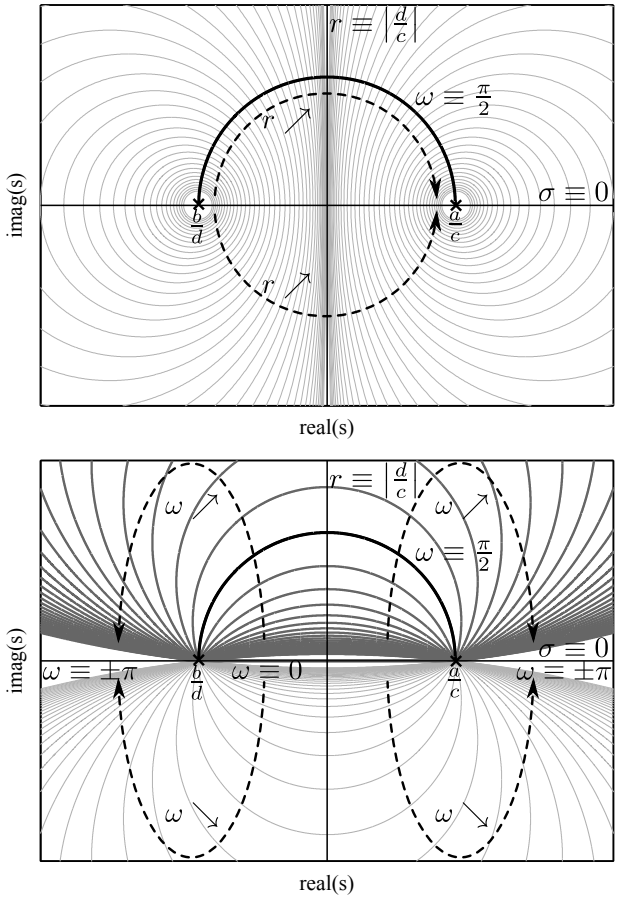


Figure 2: *Circles (in grey) corresponding to constant r (top) and arcs corresponding to constant ω (bottom) in the s-plane. Darker arcs correspond to $\omega \in [0, \pi]$. The black axis corresponds to the real axis in the s-plane, mapped from $\omega \equiv 0$ and $\omega \equiv \pm\pi$. The black arc corresponds to the $\omega \equiv \pi/2$. The black vertical line corresponds to $r \equiv |d/c|$. In the case of the bilinear transform, the arc belongs to the unit circle and the vertical line corresponds to the imaginary axis, with $b/d = -2/T$ and $a/c = 2/T$.*

2.3. Finite difference methods

Computational simulation of a system defined by Eq. (4) is a typical problem in numerical analysis [1, 5]. To digitise the system at a sampling interval T, we can compute the next sample $\mathbf{x}_n$ at $t_n = nT$ by applying common numerical methods. One-step methods correspond to methods which only use quantities evaluated at t_{n-1} and t_n, including $\mathbf{f}_n = \mathbf{f}(t_n, \mathbf{x}_n, \mathbf{u}_n)$ and $\mathbf{f}_{n-1} = \mathbf{f}(t_{n-1}, \mathbf{x}_{n-1}, \mathbf{u}_{n-1})$. Such methods include:

- forward Euler: $\mathbf{x}_n = \mathbf{x}_{n-1} + T\,\mathbf{f}_{n-1}$,
- backward Euler: $\mathbf{x}_n = \mathbf{x}_{n-1} + T\,\mathbf{f}_n$, and
- trapezoidal rule: $\mathbf{x}_n = \mathbf{x}_{n-1} + T\,(\mathbf{f}_n + \mathbf{f}_{n-1})\,/2$.

In the case of LTI systems, these methods are equivalent to order-preserving mappings of the transfer function, with forward and backward Euler being equivalent to the mappings in Eq. (6), and the trapezoidal rule being equivalent to the bilinear transform.

Like the Möbius transforms, these methods can be expressed as

$$a\mathbf{x}_n + b\mathbf{x}_{n-1} = c\,\mathbf{f}_n + d\,\mathbf{f}_{n-1} \tag{9}$$

where the parameters (a, b, c, d) are defined up to a constant factor.

A common way to classify these methods is accuracy, i.e. the order in T of the leading error term of the Taylor series expansion

$$\begin{aligned} a\mathbf{x}_n + b\mathbf{x}_{n-1} - c\,\mathbf{f}_n - d\,\mathbf{f}_{n-1} &= (a+b)\mathbf{x}_{n-1} + aT\dot{\mathbf{x}}_{n-1} \\ &- (c+d)\dot{\mathbf{x}}_{n-1} + (\tfrac{aT^2}{2} - cT)\ddot{\mathbf{x}}_{n-1} + \mathcal{O}(cT^2 + aT^3). \end{aligned} \tag{10}$$

By cancelling some of those terms, we obtain methods with various orders of accuracy. The only 2nd-order method that cancels the terms in $\mathbf{x}$, $\dot{\mathbf{x}}$, and $\ddot{\mathbf{x}}$ corresponds to the trapezoidal rule. 1st-order methods cancel the terms in $\mathbf{x}$ and $\dot{\mathbf{x}}$ with $a = -b$ and $c + d = aT$, and include, among others, backward Euler ($c = aT$ and $d = 0$) and forward Euler ($c = 0$ and $d = aT$). All other methods are 0th-order, even if some of the terms in Eq. (10) can still be cancelled (e.g. the term in $\mathbf{x}$ if we have $a = -b$).

It can seem counter-intuitive to consider 0th-order methods, meaning that error does not unconditionally vanish as we decrease the sampling interval. However, we can observe that:

- in audio, the sampling period is usually fixed so that higher-order error in T does not always translate in less error for the discretised system,
- the true solution for typical conditions are often such that $||\mathbf{x}(t)|| \rightarrow_{t\rightarrow\infty} 0$ and $||\dot{\mathbf{x}}(t)|| \rightarrow_{t\rightarrow\infty} 0$ (e.g. impulse response) so that the numerical solution error vanishes at $t \rightarrow \infty$ for all those methods,
- for other conditions when $||\dot{\mathbf{x}}(t)|| \rightarrow_{t\rightarrow\infty} 0$ (e.g. step response) but $||\mathbf{x}(t)|| \nrightarrow_{t\rightarrow\infty} 0$, the error for methods with $a = -b$ also vanishes at $t \rightarrow \infty$.

It can also seem counter-intuitive to consider 1st-order methods, when a 2nd-order method is available. However, the usage of the backward Euler method shows it is already standard practice to trade numerical accuracy for other types of desirable properties.

Here, it seems logical to restrict to methods such that $a \neq 0$ to ensure that the term in $\mathbf{x}_n$ is present. We can set it to $1/T$ as the coefficients are defined up to a multiplying factor.

3. POLE MAPPING AND WARPING

3.1. Bilinear transform frequency warping interpretation

In digital filter design, the bilinear transform is typically parameterised by a gain factor η as [7]

$$\mathcal{T}(z) = \eta \frac{1 - z^{-1}}{1 + z^{-1}}\,. \tag{11}$$

We interpret η as equal to $2/T'$ with T' a parameter potentially different from the sampling period T such that:

$$\mathcal{T}(z) = \frac{2}{T'} \frac{1 - z^{-1}}{1 + z^{-1}}\,. \tag{12}$$

For the canonical bilinear transform ($T = T'$, see Eq. (6)), we know that the imaginary axis $\sigma \equiv 0$ is mapped to the unit circle $r \equiv 1$ with the frequency distortion [4]

$$\omega = 2\tan^{-1}\left(\Omega T/2\right)/T\,. \tag{13}$$

For any T', those transforms map the imaginary axis to the unit circle, the DC point $s = 0$ to $z = 1$, and $|s| \rightarrow \infty$ to $z = -1$, with different distortion of the frequencies. One additional mapping condition between a point s on the imaginary axis and a point z on the unit circle can be used to uniquely define the desired frequency distortion. The frequency warping method aim at selecting T' so that the point $s_0 = j\Omega_0$ (only for $\Omega_0 \in]-\pi, \pi[$) maps to $z_0 = e^{j\Omega_0}$ (i.e. $\Omega_0 = \omega_0 T$), meaning that the frequencies of s_0 and z_0 match [7]. This is achieved with

$$T' = 2\tan(\Omega_0 T/2)/\Omega_0. \tag{14}$$

Warping can be interpreted in the context of numerical methods by looking at the equation

$$\dot{x}(t) = j\Omega_0 x(t),\ x(0) = 1 \tag{15}$$

and discretising it with the trapezoidal rule

$$x_n = \xi x_{n-1},\ x_0 = 1 \quad \text{with} \quad \xi = \frac{1 + j\Omega_0 T/2}{1 - j\Omega_0 T/2}\,. \tag{16}$$

The solutions to Eq. (15) and Eq. (16) are $x(t) = e^{j\Omega_0 t}$ ($t \geq 0$) and $x_n = \xi^n$ ($n \geq 0$). Since $\xi = e^{j\omega_0 T}$ with

$$\omega_0 = 2\tan^{-1}\left(\Omega_0 T/2\right)/T\,, \tag{17}$$

the solution to the discretised system presents a frequency-dependent phase lag due to eigenvalue distortion introduced by the discretisation [5]. That distortion matches the frequency distortion from the bilinear transform, meaning that the frequency warping can be interpreted in the numerical analysis framework as:

- Modifying Eq. (4a) so that the eigenvalues of the system shift by a multiplicative factor of T'/T. In filter design, that is the case where the filter coefficients are *pre-warped* and the canonical bilinear transform is used, or equivalently
- Changing the time step for the discretisation of Eq. (4a) to T' (q'_n is quantity q evaluated at nT' instead of nT) as

$$\begin{aligned}&(\mathbf{x}_n - \mathbf{x}_{n-1})/T = (\mathbf{f}_n + \mathbf{f}_{n-1})/2\\ &\mapsto (\mathbf{x}'_n - \mathbf{x}'_{n-1})/T = (\mathbf{f}'_n + \mathbf{f}'_{n-1})/2\end{aligned} \tag{18}$$

  with $\mathbf{f}_n = \mathbf{f}(t_n, \mathbf{x}_n, \mathbf{u}_n)$ and $\mathbf{f}'_n = \mathbf{f}(t'_n, \mathbf{x}'_n, \mathbf{u}'_n)$, and modifying the discretisation of Eq. (4b)

$$\mathbf{y}_n = \mathbf{g}(t_n, \mathbf{x}_n, \mathbf{u}_n) \mapsto \mathbf{y}_n = \mathbf{g}(t'_n, \mathbf{x}'_n, \mathbf{u}'_n)\,, \tag{19}$$

  effectively creating a mismatch between the time steps of $\mathbf{y}$ and that of $(t, \mathbf{x}, \mathbf{u})$. In filter design, that corresponds to changing T in T' in the bilinear transform.

Effectively, frequency warping in the bilinear transform is equivalent to *compensating for the phase lag* introduced by the trapezoidal rule in the numerical solution of Eq. (15). As the phase lag is frequency-dependent, it can only be cancelled for a single frequency. This process can be thought to be similar to the issue of minimising numerical dispersion when modelling of the 2D/3D wave equation using finite-difference meshes [12, 18].

3.2. Mapping equations

The analysis presented in the previous section focuses on the mapping of the imaginary axis. However, additional attention can be given to the mapping of all the points of the s-plane in order to further generalise the concept of warping. For any point $s = \sigma + j\Omega$, the transform defined by (a, b, c, d) maps it to $z = re^{j\omega}$ with

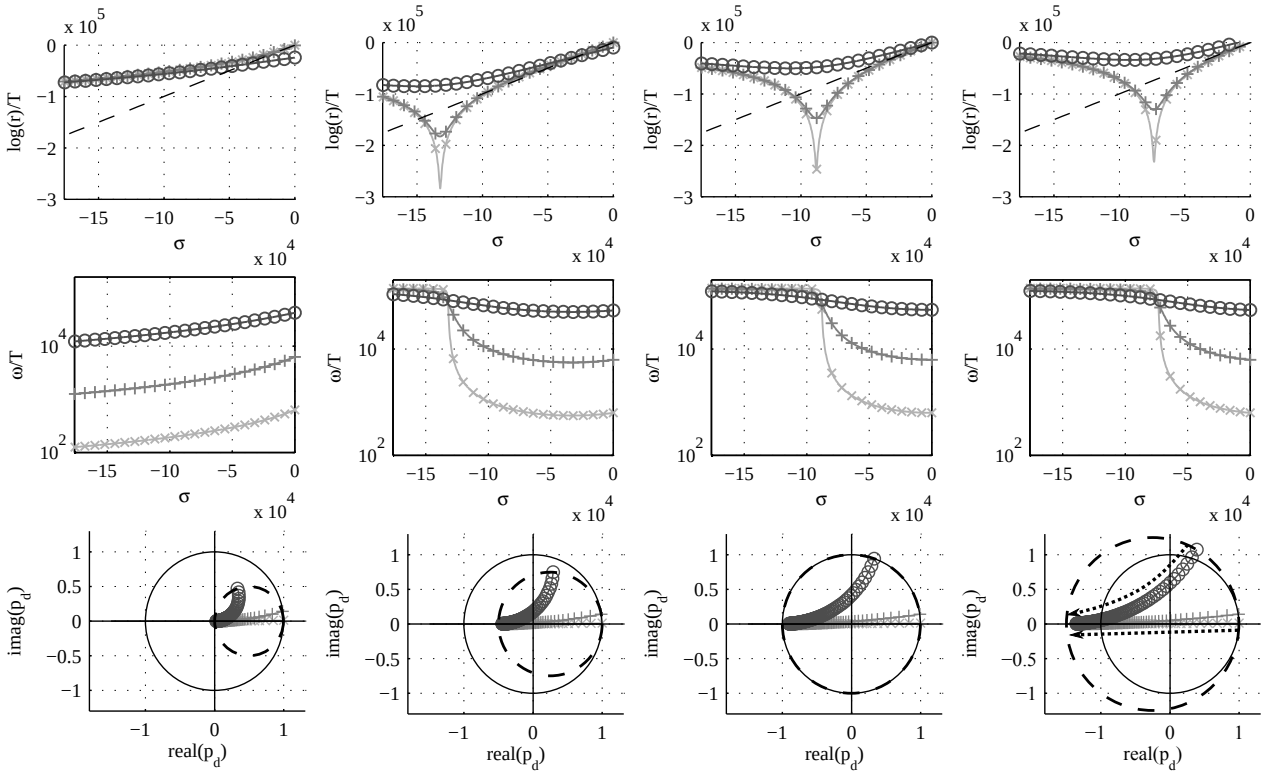
$$\begin{aligned}r &= \sqrt{\frac{\Omega^2 d^2 + (b - \sigma d)^2}{\Omega^2 c^2 + (a - \sigma c)^2}}\\ \tan\frac{\omega}{2} &= \frac{\Omega(ad - bc)}{(a - c\sigma)^2(b - d\sigma)^2 + \Omega^2 cd - r((a - c\sigma)^2 + \Omega^2 c^2)}\end{aligned} \tag{20}$$

We can derive the contours resulting from the mapping of regions of interests, as shown in [3] for the bilinear transform, backward Euler and forward Euler. For constant damping σ, s is mapped onto a circle $\mathcal{C}_\sigma(\zeta_\sigma, R_\sigma)$ (Fig. 1) such that

$$\zeta_\sigma = \frac{1}{2}\left(\frac{d\sigma - b}{a - c\sigma} - \frac{d}{c}\right) \quad \text{and} \quad R_\sigma = \frac{ad - bc}{2|c(c\sigma - a)|}\,. \tag{21}$$

For constant frequency Ω, s is mapped onto a circle (Fig. 1)

$$\zeta_\Omega = -\frac{d}{c} + j\frac{ad - bc}{2c^2\Omega} \quad \text{and} \quad R_\Omega = \frac{ad - bc}{2c^2|\Omega|} \tag{22}$$



(a) $\alpha = 0$ (BE) (b) $\alpha = 0.5$ (c) $\alpha = 1$ (BT) (d) $\alpha = 1.5$

Figure 3: *Trajectories of pole p in the z-plane for 3 different constant frequencies Ω as a function of σ (0.1 kHz: light grey $\times$, 1 kHz: medium grey $+$, 10 kHz: dark grey $\circ$). Top: $\log(r)/T$ as a function of σ (The dashed line indicates $\log(r)/T \equiv \sigma$); middle: ω/T as a function of σ; bottom: location of p in the z-plane as a function of σ. The imaginary axis $\sigma \equiv 0$ maps to the dashed circle; p starts on the right part of the dashed circle for $\sigma = 0$ and travels towards the left and the real axis as σ increases, as shown by the dotted arrows.*

such that the centre ζ moves along a vertical line going through the mapped point for infinity, $z = -d/c$.

For both σ and Ω, we see that all the circles contain the mapped point for infinity $-d/c$ and asymptotically shrink towards it, with their centre ζ moving along a line (vertical for constant Ω and horizontal for constant σ). We can also analyse the inverse mapping in order to analyse regions generating digital filters with constant parameters. For constant damping $\log(r)/T$ (or equivalently constant r), z is mapped onto an arc (Fig. 2) such that

$$\zeta_r = \frac{acr^2 - bd}{c^2r^2 - d^2} \quad \text{and} \quad R_r = \frac{r(ad - bc)}{|c^2r^2 - d^2|}\,. \tag{23}$$

For constant frequency ω/T and $\omega \geq 0$, z is mapped onto the part with non-negative imaginary part ($\sigma \geq 0$) of a circle (Fig. 2) such that

$$\zeta_\omega = \frac{ad + bc}{2cd} - j\frac{ad - bc}{2cd\tan\omega} \quad \text{and} \quad R_\omega = \frac{ad - bc}{2|cd\sin\omega|} \tag{24}$$

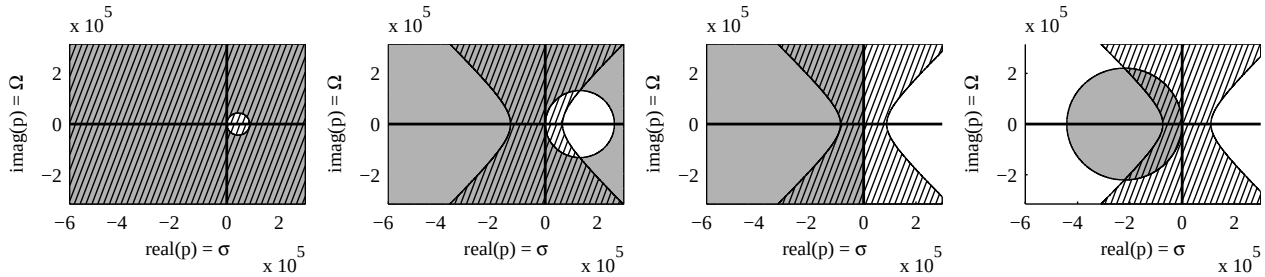
and for $\omega < 0$, z is mapped onto the part with non-positive imaginary part ($\sigma \leq 0$) of that same circle (Fig. 2).

Similarly to the forward mapping, for both r and ω, we see that all the circles have their centre organised around a single line (vertical for constant ω and horizontal for constant r).

3.3. Generalised warping

For the general class of Möbius transforms, three different mapping conditions between s- and z-planes define a unique Möbius transform (up to a multiplicative factor). In typical transforms (BT, BE, FE), we set mapping conditions for the origin ($s = 0$) and infinity ($|s| \to \infty$), so that one mapping degree of freedom remains. However, as mentioned in Sec. 2.2, we wish to limit ourselves to Möbius transforms with real coefficients, which means that not all third mapping conditions can be fulfilled.

While we cannot control the mapping of any additional point, other warping processes can be considered. For example, if a transform with real coefficients and $ad-bc > 0$ maps $s_1 \mapsto z_1$ and $s_2 \mapsto z_2$, then there is a unique circle $\mathcal{C}_s$ centred on the real axis that maps to $\mathcal{C}_z$ such that $\{s1, s2\} \in \mathcal{C}_s$ and $\{z_1, z_2\} \in \mathcal{C}_z$. Those two circles intersect the real axis in two pair of points $(\tilde{s}_1, \tilde{z}_1)$ and $(\tilde{s}_2, \tilde{z}_2)$ mapped to each other. Recall that we can force the mapping of an additional point on the positive imaginary axis $s = j\Omega$ ($\Omega > 0$) onto any point of the upper unit circle $z = e^{j\omega}$ ($\omega \in [0, \pi[$) for the bilinear transform using frequency warping. In the same way, we can show it is possible to find a Möbius transform with real coefficients and $ad - bc > 0$ that will map an additional point p of the upper arc of $\mathcal{C}_s$ to any point p_d on the upper arc of $\mathcal{C}_z$. The warping can be done by altering a single parameter χ such that the warped transform corresponds to $(\chi a, \chi b, c, d)$. This



(a) $\alpha = 0$ (BE) (b) $\alpha = 0.5$ (c) $\alpha = 1$ (BT) (d) $\alpha = 1.5$

Figure 4: *Acceptable pole locations for a given α according to Eq.* (29) *(hatched) and Eq.* (30) *(grey).*

approach can then allow for the design of transforms using similar principles to Sec. 3.1, interpreting warping as changing the phase lag introduced by a given Möbius transform in the numerical solution of $\dot{x}(t) = px(t)$. Phase lag cancellation ($\Omega_p = \omega_{p_d}/T$) can be achieved for only a subset of applicable frequencies Ω_p.

Another option would be to remove the condition either for the origin or infinity, in which case we can exactly specify the mapping of another point s. In this case, the transform coefficients are still ill-defined. An additional condition is needed, for example by cancelling one or several of the terms in Eq. (10), if it does not conflict with the mapping conditions. Otherwise, we can also find sets of transforms that would satisfy an alternative set of criteria such as the one presented in the next two sections.

3.4. Damping monotonicity conditions

In the rest of the paper, we consider only transforms of the form $(a, b, c, d) = (\frac{1+\alpha}{T}, -\frac{1+\alpha}{T}, 1, \alpha)$, even though all the discussed properties extend readily to all previous Möbius transforms. This subclass of transforms includes the bilinear transform ($\alpha = 1$), the forward ($\alpha \to \infty$), and the backward ($\alpha = 0$) Euler methods. As numerical schemes, they are all 1st- or 2nd-order schemes and can be viewed as forward Euler with an added dissipative term [16]:

$$\begin{aligned}\mathbf{x}_n &= \mathbf{x}_{n-1} + T(\mathbf{f}_n + \alpha\mathbf{f}_{n-1})/(1+\alpha) \\ &= \underbrace{\mathbf{x}_{n-1} + T\mathbf{f}_{n-1}}_{\text{FE}} + \underbrace{T^2\ddot{\mathbf{x}}_{n-1}/(1+\alpha)}_{\text{dissipative term}} + \mathcal{O}(T^3)\,. \end{aligned} \quad (25)$$

We also assume we have knowledge of a region in the s-plane enclosing all the possible poles of the studied system (or all the possible instantaneous poles for systems modelled as Eq. (4)).

Through such transform, a pole p is mapped to p_d as

$$p = \sigma + j\Omega \mapsto p_d = re^{j\omega} = \frac{1+\alpha+\alpha Tp}{1+\alpha-Tp}\,. \quad (26)$$

The mapping also generates poles at $z = -\alpha$ from the mapping of the zeroes of the continuous systems, but these poles are all cancelled by the zeroes at $z = -\alpha$ from the mapping of the poles if we only consider proper systems [7]. By adapting Eq. (20) to the class of transforms considered here, the mapping between quantities (σ, Ω) and r simplifies to:

$$r^2 = \frac{(1+\alpha+\alpha T\sigma)^2 + (\alpha T\Omega)^2}{(1+\alpha-T\sigma)^2 + (T\Omega)^2} \quad (27)$$

From Eqs. (20) and (27), we can observe in more details the behaviour of the trajectories (Fig. 3) of the poles for different conditions such as $\log(r)/T$ and ω/T at constant Ω as σ decreases from zero into the region of stable poles $\sigma < 0$.

In the plots of $\log(r)/T$ at a constant σ we observe that while the continuous-time and discrete-time damping have a monotonic relationship for lower Ω, we ultimately reach an inflection point for which an increase in the continuous-time damping results in a decrease in discrete-time damping. Passed the inflection point, the discrete-time frequency moves quickly towards $\pm\pi$, so that p_d becomes a resonant (and possibly unstable) pole at the Nyquist frequency. For a given Ω, it is possible to obtain an analytical expression of that inflection point σ_m by solving the equation $\frac{\partial r}{\partial \sigma} = \frac{1}{2r}\frac{\partial r^2}{\partial \sigma} = 0$. From Eq. (27), we get two solutions:

$$\sigma_\pm = \frac{(\alpha^2-1) \pm \sqrt{(\alpha+1)^4 + (2\alpha T\Omega)^2}}{2\alpha T}\,. \quad (28)$$

We can see that $\sigma_+ > 0$ and $\sigma_- < 0$. Also, σ and r have a monotonic relationship for any $\sigma \in [\sigma_-, \sigma_+]$. In the s-plane, this monotonicity condition is verified inside a rectangular hyperbola of semi major axis $\frac{(\alpha+1)^2}{2\alpha T}$ and centre $\left(\frac{\alpha^2-1}{2\alpha T}, 0\right)$:

$$\left(\sigma - \frac{\alpha^2-1}{2\alpha T}\right)^2 - \Omega^2 \le \left(\frac{(\alpha+1)^2}{2\alpha T}\right)^2 \text{ for } \alpha > 0\,, \quad (29)$$

which becomes the half-plane $\sigma > -\frac{1}{T}$ for $\alpha \to \infty$ (i.e. forward Euler) and $\sigma \in \mathbb{R}$ for $\alpha = 0$ (i.e. backward Euler).

If we have prior knowledge of the pole possible locations (e.g. from the physics of the system), we can select α so that all poles lie in that region, and consequently ensure the monotonic relationship between σ and r at constant Ω. From Eq. (25), we can interpret this process as adding enough dissipation to the discretised system in order to move its poles away from the Nyquist frequency.

3.5. Stability conditions

Another condition to verify, in particular if we also want to allow for values $\alpha > 1$ (i.e. discretisation schemes that can produce unstable discrete-time poles from stable continuous-time poles such as forward Euler), is that all the potential poles will be stable, i.e. that $r < 1$. While complete stability analysis of general time-varying system cannot be guaranteed through the sole analysis of its instantaneous poles [19], a preliminary analysis is generally done by verifying the condition $r < 1$ for those poles [13]. In

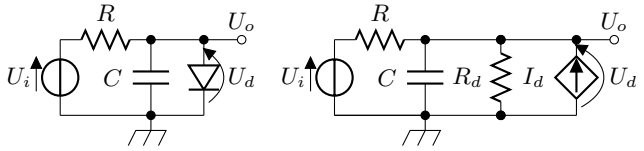


Figure 5: *Diode clipper circuit (left: original; right: linearised).*

the s-plane, the condition is verified over the region described by:

$$\begin{aligned} \left(\sigma - \tfrac{1}{T}\tfrac{1+\alpha}{1-\alpha}\right)^2 + \Omega^2 &> \left(\tfrac{1}{T}\tfrac{1+\alpha}{1-\alpha}\right)^2 \quad \text{for } \alpha \in [0,1[\\ \left(\sigma - \tfrac{1}{T}\tfrac{1+\alpha}{1-\alpha}\right)^2 + \Omega^2 &< \left(\tfrac{1}{T}\tfrac{\alpha+1}{\alpha-1}\right)^2 \quad \text{for } \alpha > 1\,, \end{aligned} \tag{30}$$

which degenerates in the half-plane $\sigma < 0$ for $\alpha = 1$.

Again, with prior knowledge of the pole possible locations, it is possible to ensure that all the poles will lie in those two regions by selecting α appropriately, so that the system stability is guaranteed for all potential poles of the system. The intersection of this condition with the monotonic damping condition in the s-plane can be seen in Fig. 4 for different values of α.

4. SIMULATIONS

We study the diode clipper in Fig. 5 [20] to illustrate those concepts. Following the Shockley diode law, the current I through a diode is modelled as a function of the voltage U across a diode as

$$I = f(U) = I_s \left(e^{U/V_t} - 1\right). \tag{31}$$

If we measure the output voltage U_o around the diode as a function of the driving voltage U_i, the system is described through the state-space representation with state variable U_d:

$$\begin{aligned} \dot{U}_d &= (U_i - U_d)/(RC) - f(U_d)/C \\ U_o &= U_d\,. \end{aligned} \tag{32}$$

At a given time instant t_0, the diode nonlinear characteristic can be linearised into a small perturbation model around U_{d0} with a resistor R_d in parallel with a constant current source I_d as

$$\begin{aligned} f(U_d) &\approx U_d/R_d - I_d \\ R_d &= U_t e^{-U_{d0}/V_t}/I_s \\ I_d &= I_s\left(1 + (U_{d0}/V_t - 1)e^{U_{d0}/V_t}\right) \end{aligned} \tag{33}$$

and the linearised state-space representation becomes

$$\begin{aligned} \dot{U}_d &= -\frac{U_d}{C(R||R_d)} + \begin{bmatrix}\frac{1}{RC}, & \frac{1}{C}\end{bmatrix}\begin{bmatrix}U_i \\ I_d\end{bmatrix} \\ U_o &= U_d\,. \end{aligned} \tag{34}$$

The linearised system has one real pole at

$$p = -\frac{1}{C}\left(\frac{1}{R} + \frac{1}{R_d}\right) = -\frac{1}{C}\left(\frac{1}{R} + \frac{I_s}{V_t}e^{U_{d0}/V_t}\right). \tag{35}$$

In general, it is not possible to know analytically the range of output voltages of this system, but we can estimate it by looking at

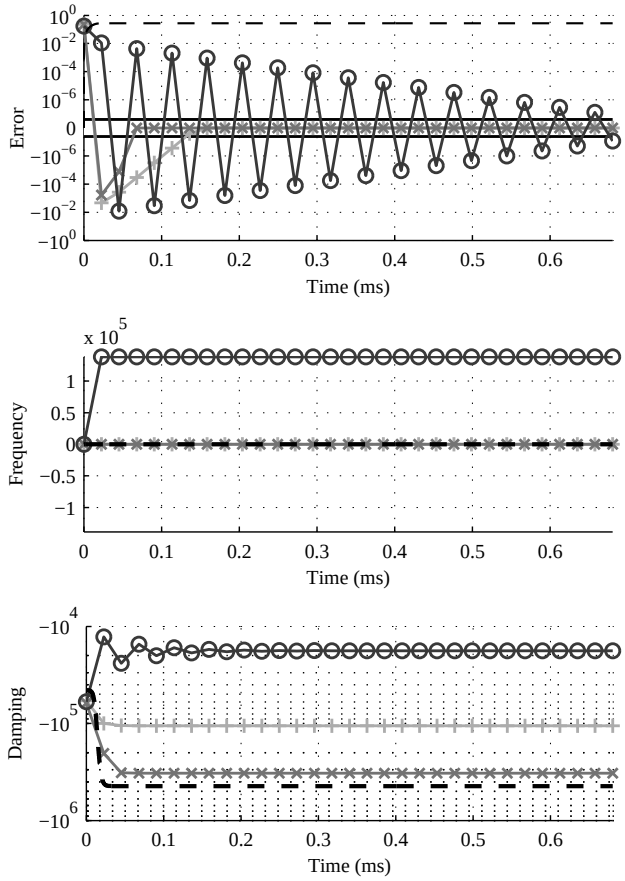


Figure 6: *System response to a $U_i(t) = 0.5u(t)$ for $\alpha = 0$ (light grey +), 0.11 (medium grey ×) and 1 (dark grey ∘). Top: reference U_o^{ref} (dashed) and signed error $\hat{U}_o - U_o^{\text{ref}}$ for each α in log-amplitude; middle: instantaneous pole frequencies Ω^{ref} (dashed) and $\hat{\omega}/T$ for each α; bottom: instantaneous pole dampings σ^{ref} (dashed) and $\log(\hat{r})/T$ for each α.*

the steady-state response of the system at maximum and minimum input voltage U_i as the solution of

$$0 = (U_i - U_d)/(RC) - f(U_d)/C \tag{36}$$

that can be found empirically. For a N914 switching diode ($I_s = 2.52\,\text{nA}$, $V_t = 25.85\text{mV}$), $R = 2.2\,\text{k}\Omega$ and $C = 0.01\,\mu\text{F}$, having an input voltage in $[-0.5\,\text{V}, 0.5\,\text{V}]$ produces steady-state output voltages in $[-0.5\,\text{V}, 0.275\,\text{V}]$. As a result the pole p has a continuous-time damping σ in $[-4.42 \times 10^5, -4.55 \times 10^4]$.

To ensure that we pick a Möbius transform of the form $(a, b, c, d) = (\frac{1+\alpha}{T}, -\frac{1+\alpha}{T}, 1, \alpha)$, with α such that all possible continuous-time pole locations $p = \sigma$ ($\sigma < 0$) fall inside the hyperbola described in Eq. (29), we have the condition:

$$\begin{aligned} \forall\sigma,\ -(\alpha+1)/(\alpha T) \le \sigma &\Leftrightarrow -(\alpha+1)/(\alpha T) \le \sigma_{\min} \\ &\Leftrightarrow \alpha(T\sigma_{\min} + 1) \ge -1\,. \end{aligned} \tag{37}$$

For $\sigma \le -1/T$, this condition is always satisfied. Hence, one way to avoid issues with the mapping of the poles is to reduce the sampling interval T (e.g. oversample). For $\sigma < -1/T$, the condition is $\alpha \le -1/(1+T\sigma_{\min})$. For our system sampled at 44.1 kHz

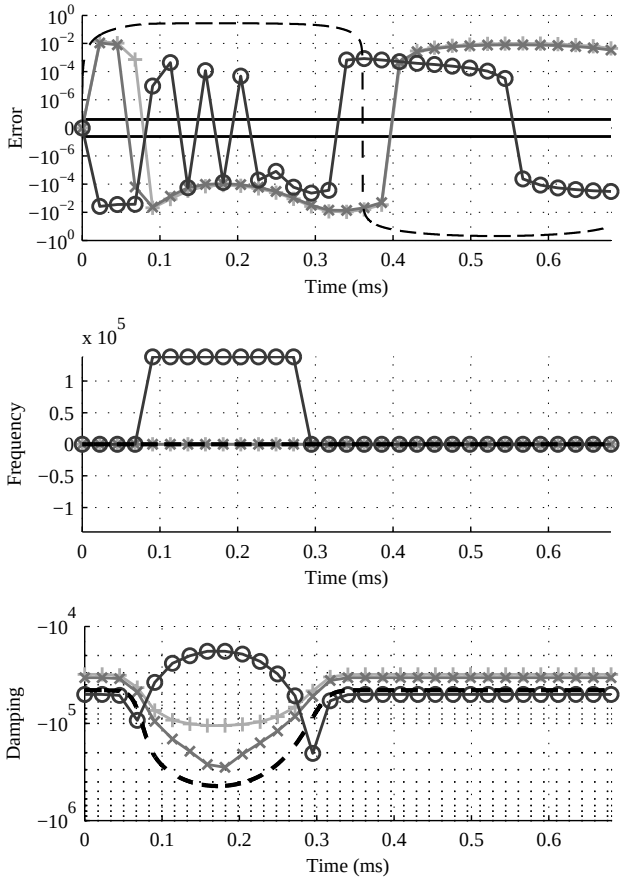


Figure 7: *System response to* $U_i(t) = 0.5\sin(2940\pi t)u(t)$ *(first cycle) for* $\alpha = 0$ *(light grey +),* 0.11 *(medium grey ×) and* 1 *(dark grey ∘). Top: reference* U_o^{ref} *(dashed) and signed error* $\hat{U}_o - U_o^{\mathrm{ref}}$ *for each* α *in log-amplitude; middle: instantaneous pole frequencies* Ω^{ref} *(dashed) and* $\hat{\omega}/T$ *for each* α*; bottom: instantaneous pole dampings* σ^{ref} *(dashed) and* $\log(\hat{r})/T$ *for each* α*.*

($T \approx 22.7\mu$s), the condition to have pole damping monotonicity is $\alpha \lesssim 0.11$. The stability condition translates into a less tight bound $\alpha < (T\sigma_{\min} + 2)/(T\sigma_{\min} - 2)$, i.e. $\alpha \lesssim 1.5$.

We study the impact of instantaneous pole locations by simulating the system response to a step function $U_i(t) = 0.5u(t)$ (Fig. 6) and the onset of a sinusoid $U_i(t) = 0.5\sin(2940\pi t)u(t)$ (Fig. 7), with unit step function $u(t) = \mathbf{1}\{t \geq 0\}$. As reference U_o^{ref} for the system response, we compute the solution to Eq. (32) using the variable-step Runge–Kutta solver `ode45` from MATLAB with absolute and relative accuracies set to machine precision and a maximum time step of $T/8$ (resampled using piecewise cubic interpolation). Then, we compute the response $\hat{U}_o$ of Eq. (32) after discretising it using $\alpha = 0$ (BE), $\alpha = 1$ (BT) and $\alpha = 0.11$ (nearly critical damping monotonicity condition). We solve all implicit update equations using Newton's method [1] to machine precision. For each simulation, we compute the (signed) error $\hat{U}_o - U_o^{\mathrm{ref}}$, and the coordinates (damping and frequency) of the reference instantaneous poles in the s-plane (σ^{ref} and Ω^{ref}, computed using U_o^{ref} as approximation to $U_o(t)$), and of the discretised systems in the z-plane ($\log(\hat{r})/T$ and $\hat{\omega}/T$). The simulations show how the bilinear transform instantaneous pole frequency shifts to π/T when the voltage across the diode becomes high, which results in spurious high-frequency oscillations. On the other hand, backward Euler and $\alpha = 0.11$ instantaneous pole frequencies are always 0 so that no oscillations are triggered.

5. CONCLUSION

In this paper, we presented an generalisation of common methods for discretising transfer functions using Möbius transforms. We study the properties of these transforms when used as discretisation methods in the context of mapping functions from the s-plane to the z-plane and in the parallel context of single-step numerical methods for state-space representations of systems. We introduce general considerations regarding the distortion of pole properties (damping and frequency) as a function of the transform coefficients when mapping from one plane to the other. Finally, we present some criteria based on desirable properties of the pole locations. This allows us to design Möbius transforms that yield discretised systems with those properties as a function of the location of all possible continuous-time system poles. These concepts are illustrated through the simulation of a typical diode clipper circuit, using small-perturbation analysis to predict the transient behaviour of the system as a function of the pole location.

6. REFERENCES

[1] J. O. Smith, *Physical Audio Signal Processing*, W3K Pub., 2010.
[2] Z. Nehari, *Conformal mapping*, Dover Publications, 2011.
[3] T. Stilson, *Efficiently-Variable Non-Oversampled Algorithms in Virtual-Analog Music Synthesis*, Ph.D. thesis, Stanford Univ., 2006.
[4] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall, 3rd ed., 2009.
[5] P. Moin, *Fundamentals of engineering numerical analysis*, Cambridge University Press, 2010.
[6] G. F. Franklin, D. J. Powell, and M. L. Workman, *Digital Control of Dynamic Systems*, Prentice Hall, 3rd ed., 1997.
[7] J. O. Smith, *Introduction to Digital Filters with Audio Applications*, W3K Pub., 2007.
[8] S. J. Orfanidis, "Digital parametric equalizer design with prescribed nyquist-frequency gain," *J. Audio Eng. Soc*, vol. 45, no. 6, 1997.
[9] J. A. Moorer, "About this reverberation business," *Comput. Music J.*, vol. 3, pp. 13–28, 1979.
[10] T. W. Parks and C. S. Burrus, *Digital filter design*, Wiley, 1987.
[11] R. M. Golden, "Digital filter synthesis by sampled-data transformation," *IEEE Trans. Audio Electroacoust.*, vol. 16, no. 3, 1968.
[12] S. Bilbao, *Numerical Sound Synthesis: Finite Difference Schemes and Simulation in Musical Acoustics*, Wiley, 1st ed., 2009.
[13] D. T. Yeh, J. S. Abel, A. Vladimirescu, and J. O. Smith, "Numerical methods for simulation of guitar distortion circuits," *Comput. Music J.*, vol. 32, no. 2, pp. 23–42, 2008.
[14] S.-C. Pei and H.-J. Hsu, "Fractional bilinear transform for analog-to-digital conversion," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 2122–2127, 2008.
[15] B. Lindberg, "On smoothing and extrapolation for the trapezoidal rule," *BIT Numerical Mathematics*, vol. 11, no. 1, pp. 29–52, 1971.
[16] W. Gao, E. Solodovnik, R. Dougal, G. J. Cokkinides, and A. P. S. Meliopoulos, "Elimination of numerical oscillations in power system dynamic simulation," in *Proc. 18th IEEE Appl. Power Electron. Conf. and Expo.*, 2003, vol. 2, pp. 790–4.
[17] T. Needham, *Visual complex analysis*, Oxford Univ. Press, 1998.
[18] B. Hamilton and A. Torin, "Finite difference schemes on hexagonal grids for thin linear plates with finite volume boundaries," in *Proc. Int. Conf. Digital Audio Effects*, 2014, pp. 592–599.
[19] J. Laroche, "On the stability of time-varying recursive filters," *J. Audio Eng. Soc*, vol. 55, no. 6, pp. 460–471, 2007.
[20] J. Macak and J. Schimmel, "Nonlinear circuit simulation using time-variant filter," in *Proc. Int. Conf. Digital Audio Effects*, 2009.