REVISITING THE SECOND-ORDER ACCURATE NON-ITERATIVE DISCRETIZATION SCHEME

Martin Holters

Department of Signal Processing and Communication
Helmut Schmidt University – University of the Federal Armed Forces
Hamburg, Germany
martin.holters@hsu-hh.de

ABSTRACT

In the field of virtual analog modeling, a variety of methods have been proposed to systematically derive simulation models from circuit schematics. However, they typically rely on implicit numerical methods to transform the differential equations governing the circuit to difference equations suitable for simulation. For circuits with non-linear elements, this usually means that a non-linear equation has to be solved at run-time at high computational cost.

As an alternative to fully-implicit numerical methods, a family of non-iterative discretization schemes has recently been proposed, allowing a significant reduction of the computational load. However, in the original presentation, several assumptions are made regarding the structure of the ODE, limiting the generality of these schemes. Here, we show that for the second-order accurate variant in particular, the method is applicable to general ODEs. Furthermore, we point out an interesting connection to the implicit midpoint method.

1. INTRODUCTION

In recent years, a variety of methods have been proposed to systematically derive simulation models from circuit schematics for virtual analog modeling, like wave digital filters (e.g. [1, 2, 3]), methods based on state-space systems (e.g. [4, 5, 6]), and port-Hamiltonian systems (e.g. [7, 8]). They have in common that they typically rely on implicit numerical methods, usually trapezoidal rule or the implicit midpoint method, to transform the differential equations governing the circuit to difference equations suitable for simulation. For circuits with non-linear elements, this usually means that a non-linear equation has to be solved at run-time using e.g. Newton-Raphson iteration, which then often dominates the computational requirements.

As an alternative to fully-implicit numerical methods, in [9, 10], Ducceschi et al. have introduced a family of non-iterative discretization schemes for ODEs. These can significantly reduce the computational load, albeit at the cost of reduced stability guarantees and larger (non-asymptotic) error. However, in their presentation, they make several assumptions regarding the structure of the ODE, limiting the generality of these schemes. Here, we show that for the second-order accurate variant in particular, the method is applicable to general ODEs.

In the following, $x(t)$ denotes the state of a system governed by an ODE and $\dot{x}(t)$ its derivative with respect to time. When obvious from the context, we drop the argument t. We seek an approximation $\hat{x}(n) \approx x(nT)$ where T is the sampling interval. For given signals like an input $u(t)$, we likewise denote its samples as e.g. $\hat{u}(n) = u(nT)$.

Copyright: © 2024 Martin Holters. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, adaptation, and reproduction in any medium, provided the original author and source are credited.

2. RECAP OF THE NON-ITERATIVE METHOD

We start with a recap of the main take-aways of [9, 10]. Consider a scalar ODE

$$\dot{x} + f(x) = u \tag{1}$$

where $f(0) = 0$, so that we can rewrite as

$$\dot{x} + g(x) \cdot x = u \tag{2}$$

by letting $g(x) = f(x)/x$ (with continuous extension $g(0) = f'(0)$ at $x = 0$ where $f' = \frac{df}{dx}$). For the zero-input case, the non-iterative scheme is then given by

$$\sigma_P(\hat{x}(n)) \cdot \frac{1}{T}(\hat{x}(n+1) - \hat{x}(n)) + g(\hat{x}(n)) \cdot \frac{1}{2}(\hat{x}(n+1) + \hat{x}(n)) = 0 \tag{3}$$

where

$$\sigma_P(\hat{x}(n)) = \sum_{p=0}^{P} T^p \zeta_p(\hat{x}(n)) \tag{4}$$

and appropriate choice of ζ_p leads to a method of accuracy order $P + 1$. In particular, one finds $\zeta_0(\hat{x}(n)) = 1$ and

$$\zeta_1(\hat{x}(n)) = \frac{1}{2} \cdot \Big(f'(\hat{x}(n)) - g(\hat{x}(n)) \Big). \tag{5}$$

To give the non-iterative scheme, (3) is solved as

$$\hat{x}(n+1) = \frac{\sigma_P(\hat{x}(n)) - \frac{T}{2} g(\hat{x}(n))}{\sigma_P(\hat{x}(n)) + \frac{T}{2} g(\hat{x}(n))} \cdot \hat{x}(n) = \frac{1 - \kappa(n)}{1 + \kappa(n)} \cdot \hat{x}(n) \tag{6}$$

with

$$\kappa(n) = \frac{T g(\hat{x}(n))}{2 \sigma_P(\hat{x}(n))}. \tag{7}$$

For non-zero input, (3) is extended to

$$\sigma_P(\hat{x}(n)) \cdot \frac{1}{T}(\hat{x}(n+1) - \hat{x}(n)) + g(\hat{x}(n)) \cdot \frac{1}{2}(\hat{x}(n+1) + \hat{x}(n)) = \frac{1}{2}(\hat{u}(n+1) + \hat{u}(n)) \tag{8}$$

which likewise can be solved to

$$\hat{x}(n+1) = \frac{\sigma_P(\hat{x}(n)) - \frac{T}{2} g(\hat{x}(n))}{\sigma_P(\hat{x}(n)) + \frac{T}{2} g(\hat{x}(n))} \hat{x}(n) + \frac{\frac{T}{2}}{\sigma_P(\hat{x}(n)) + \frac{T}{2} g(\hat{x}(n))} (\hat{u}(n+1) + \hat{u}(n)). \tag{9}$$

The ODE (1) is zero-stable (i.e. x converges to zero for $u = 0$) if $\operatorname{sgn}(x) = \operatorname{sgn}(f(x))$, i.e. $g(x) > 0$ for $x \neq 0$. Considering (6), the discretized system is stable if $\kappa(n) \geq 0$, which for $g(x) > 0$ requires $\sigma_P(x) > 0$, which in turn is certainly the case if all $\zeta_p(x)$ are non-negative. For $P = 1$ in particular, a sufficient condition for stability is thus given by $f'(x) \geq g(x) > 0$.

For the vector case, consider an ODE like (1) with

$$f(x) = Bx + Fq(\eta(x)) \tag{10}$$

where $\eta(x) = F^T x + c$ where c is a time-dependent input function and $q(\eta) = \begin{pmatrix} q_1(\eta_1) & \cdots & q_N(\eta_N) \end{pmatrix}^T$ is an element-wise nonlinear function. Then

$$g(x) = B + FD(x)F^T \tag{11}$$

is used in (3) where $D(x) = \operatorname{diag}(q_i(\eta_i)/\eta_i)$ with $\eta(\hat{x}(n)) = F^T\hat{x}(n) + \frac{1}{2}(\hat{c}(n+1) + \hat{c}(n))$ and f' becomes the Jacobian. However, the fact that q is restricted to an element-wise function (i.e. a collection of scalar functions) can be a severe restriction. Generally applicable practical conditions for stability have not been found, but benign behavior has been observed in practice.

3. REVISITING THE SECOND-ORDER CASE

We now only consider an accuracy order of 2, we thus need $P = 1$, i.e.

$$\sigma_1(\hat{x}(n)) = 1 + \frac{T}{2}\Big(f'(\hat{x}(n)) - g(\hat{x}(n))\Big). \tag{12}$$

We again consider an ODE in the form of (1), but vector-valued and assume $g(x)x = f(x)$, where matrix $g(x)$ is not uniquely defined. Focusing on the zero-input case first, we rewrite (3) as

$$\begin{aligned}\left(\sigma_1(\hat{x}(n)) + \frac{T}{2}g(\hat{x}(n))\right)\hat{x}(n+1) \\ = \left(\sigma_1(\hat{x}(n)) - \frac{T}{2}g(\hat{x}(n))\right)\hat{x}(n)\end{aligned} \tag{13}$$

and substitute σ_1 to obtain

$$\begin{aligned}\left(I + \frac{T}{2}f'(\hat{x}(n))\right)\hat{x}(n+1) \\ = \left(I + \frac{T}{2}f'(\hat{x}(n)) - Tg(\hat{x}(n))\right)\hat{x}(n).\end{aligned} \tag{14}$$

We can now exploit $g(x)x = f(x)$ to rewrite to

$$\begin{aligned}\left(I + \frac{T}{2}f'(\hat{x}(n))\right)\hat{x}(n+1) \\ = \left(I + \frac{T}{2}f'(\hat{x}(n))\right)\hat{x}(n) - Tf(\hat{x}(n)).\end{aligned} \tag{15}$$

and solve as

$$\hat{x}(n+1) = \hat{x}(n) - \left(I + \frac{T}{2}f'(\hat{x}(n))\right)^{-1} Tf(\hat{x}(n)). \tag{16}$$

We have thus obtained a non-iterative scheme without explicit appearance of g, which even lets us treat cases where $f(0) \neq 0$. It remains to be verified that the scheme is second-order accurate even then. In particular, we will assume stability for the moment and show that if $\hat{x}(n) = x(t)$ for $t = nT$, then $x(t+T) - \hat{x}(n+1) = O(T^3)$.

For sufficiently small T, we may utilize the geometric series

$$\sum_{i=0}^{\infty}\left(-\frac{T}{2}f'(\hat{x}(n))\right)^i = \left(I + \frac{T}{2}f'(\hat{x}(n))\right)^{-1} \tag{17}$$

so that

$$\begin{aligned}\hat{x}(n+1) &= \hat{x}(n) - \left(\sum_{i=0}^{\infty}\left(-\frac{T}{2}f'(\hat{x}(n))\right)^i\right)Tf(\hat{x}(n)) \\ &= \hat{x}(n) - Tf(\hat{x}(n)) + \frac{T^2}{2}f'(\hat{x}(n))f(\hat{x}(n)) + O(T^3).\end{aligned} \tag{18}$$

Comparing with the Taylor series expansion

$$\begin{aligned}x(t+T) &= x(t) + T\dot{x}(t) + \frac{T^2}{2}\ddot{x}(t) + O(T^3) \\ &= x(t) - Tf(x(t)) + \frac{T^2}{2}f'(x(t))f(x(t)) + O(T^3)\end{aligned} \tag{19}$$

it is obvious that the method is second-order accurate.

Regarding stability, first note that if $f(0) \neq 0$, the ODE itself does not converge to zero. Instead, assume that the ODE converges to some $\bar{x}$, then in the scalar case, $f(\bar{x}) = 0$ and $\operatorname{sgn}(x - \bar{x}) = \operatorname{sgn}(f(x))$ hold and it is easily verified that $0 < \frac{f(x)}{x-\bar{x}} \leq f'(x)$ is a sufficient condition for stability of the discretized system. Unfortunately, a practical stability condition for the vector case remains elusive.

3.1. Including an input signal

Now consider a vector-valued ODE of the form

$$\dot{x} + f(x, u) = 0 \tag{20}$$

where f is arbitrary and includes an input term u. (For simplicity, we only consider a zero right-hand side, as a non-zero right-hand side could easily be subsumed in f.) In the following, let J_x and J_u denote the Jacobians of f with respect to x and u, respectively. We adapt (16) as

$$\hat{x}(n+1) = \hat{x}(n) - \left(I + \frac{T}{2}J_x^n\right)^{-1} Tf^n \tag{21}$$

where

$$f^n = f\Big(\hat{x}(n), \tfrac{1}{2}(\hat{u}(n+1) + \hat{u}(n))\Big) \tag{22}$$

and likewise

$$J_x^n = J_x\Big(\hat{x}(n), \tfrac{1}{2}(\hat{u}(n+1) + \hat{u}(n))\Big). \tag{23}$$

We will show second-order accuracy similarly as before. By following the same reasoning as above, we find

$$\hat{x}(n+1) = \hat{x}(n) - Tf^n + \frac{T^2}{2}J_x^n f^n + O(T^3). \tag{24}$$

The Taylor series of $x(t+T)$ becomes a bit more complicated due to the second (time-dependent) argument to f, which in addition we

have to evaluate at $\frac{1}{2}(u(t+T)+u(t))$. To get there, first observe that

$$\begin{aligned}\dot{x}(t) &= -f(x(t),u(t))\\ &= -f\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\\ &\quad + J_u(x(t),\tfrac{1}{2}(u(t+T)+u(t)))\cdot\tfrac{1}{2}(u(t+T)-u(t))\\ &\quad + O\Big((u(t+T)-u(t))^2\Big)\\ &= -f\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\\ &\quad + \frac{T}{2}J_u(x(t),\tfrac{1}{2}(u(t+T)+u(t)))\cdot\tfrac{1}{2}(\dot{u}(t+T)+\dot{u}(t))\\ &\quad + O(T^2)\end{aligned} \tag{25}$$

as $u(t+T)-u(t)=\frac{T}{2}(\dot{u}(t+T)+\dot{u}(t))+O(T)$. Based on that, we furthermore find

$$\begin{aligned}\ddot{x}(t) &= -\frac{d}{dt}f(x(t),u(t))\\ &= -\frac{d}{dt}f\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)+O(T)\\ &= -J_x\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\dot{x}(t)\\ &\quad - J_u\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\cdot\tfrac{1}{2}(\dot{u}(t+T)+\dot{u}(t))\\ &\quad + O(T)\\ &= J_x\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)f\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\\ &\quad - J_u\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\cdot\tfrac{1}{2}(\dot{u}(t+T)+\dot{u}(t))\\ &\quad + O(T).\end{aligned} \tag{26}$$

Substituting these in the Taylor expansion of $x(t+T)$ then yields

$$\begin{aligned}x(t+T) &= x(t)+T\dot{x}(t)+\frac{T^2}{2}\ddot{x}(t)+O(T^3)\\ &= x(t)-Tf\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\\ &\quad + \frac{T^2}{2}J_x\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\\ &\quad\cdot f\Big(x(t),\tfrac{1}{2}(u(t+T)+u(t))\Big)\\ &\quad + O(T^3).\end{aligned} \tag{27}$$

Comparison with (24) now shows that the method is indeed second-order accurate.

4. CONNECTION TO THE IMPLICIT MIDPOINT METHOD

The midpoint method is a commonly used implicit scheme which when applied to (20) and approximating $u((n+\frac{1}{2})T)$ with $\frac{1}{2}(\hat{u}(n+1)+\hat{u}(n))$ [11] gives

$$\hat{x}(n+1)=\hat{x}(n)-Tf\Big(\tfrac{1}{2}(\hat{x}(n+1)+\hat{x}(n)),\tfrac{1}{2}(\hat{u}(n+1)+\hat{u}(n))\Big). \tag{28}$$

Generally, no closed from solution for $\hat{x}(n+1)$ is available and one has to resort to numerical solution methods. In particular, we shall consider Newton-Raphson iteration, which for this case is given by

$$\begin{aligned}&\hat{x}_{i+1} =\\ &\hat{x}_i - \Big(I+\frac{T}{2}J_x\Big(\tfrac{1}{2}(\hat{x}_i+\hat{x}(n)),\tfrac{1}{2}(\hat{u}(n+1)+\hat{u}(n))\Big)\Big)^{-1}\\ &\cdot\Big(\hat{x}_i-\hat{x}(n)+Tf\Big(\tfrac{1}{2}(\hat{x}_i+\hat{x}(n)),\tfrac{1}{2}(\hat{u}(n+1)+\hat{u}(n))\Big)\Big).\end{aligned} \tag{29}$$

Given an appropriate starting point $\hat{x}_0$, $\hat{x}_i$ will converge to the desired solution $\hat{x}(n+1)$. A reasonable choice for the starting point that is often used in practice is the solution from the previous time step, i.e. $\hat{x}_0=\hat{x}(n)$. With this choice, the first iteration becomes

$$\begin{aligned}\hat{x}_1 &= \hat{x}(n)-\Big(I+\frac{T}{2}J_x\Big(\hat{x}(n),\tfrac{1}{2}(\hat{u}(n+1)+\hat{u}(n))\Big)\Big)^{-1}\\ &\qquad\cdot Tf\Big(\hat{x}(n),\tfrac{1}{2}(\hat{u}(n+1)+\hat{u}(n))\Big)\\ &= \hat{x}(n)-\Big(I+\frac{T}{2}J_x^n\Big)^{-1}Tf^n\end{aligned} \tag{30}$$

with the definitions of (22) and (23). We see that if we terminate the Newton-Raphson method after the first iteration and let $\hat{x}(n+1)=\hat{x}_1$, we recover (21). Thus, the non-iterative method of accuracy order 2 can also be viewed as an approximation of the implicit midpoint method with (very) early termination of the numerical solver. Another way of looking at it is to consider (21) the solution of the local linearization of (28) with respect to $\hat{x}(n+1)$ at $\hat{x}(n)$, an approach that has similarly been used with good results (but without theoretical analysis of its properties) for physical modeling in [12]. It is remarkable that even with an exact solution, the implicit midpoint method has accuracy order 2, so this approximation does not decrease the accuracy order (but it does impact the non-asymptotic error as well as stability).

5. EXAMPLES

5.1. Lotka-Volterra equation

As a first example, we consider the Lotka-Volterra equation, since it is one of the simplest ODEs that cannot be addressed with the non-iterative scheme in its original formulation. Setting all coefficients to unity, the Lotka-Volterra equation becomes

$$\dot{x}+f(x)=\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix}+\begin{pmatrix}x_1\cdot(x_2-1)\\ x_2\cdot(1-x_1)\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}. \tag{31}$$

The resulting state trajectories when starting at $x(0)=\begin{pmatrix}2 & 2\end{pmatrix}^T$ are shown in fig. 1. Discretization with the non-iterative method gives convincing results for about $T\leq 1/2$. For larger T, the discretized system suffers from instability. The reason is that the continuous-time system itself becomes unstable for $x_1<0$, and while the true solution can never change sign, the numerical solution can if T is too large, and will then diverge. Note, however, that the same is true for the implicit midpoint method (and likely other schemes), where the threshold on T is not much larger.

Unfortunately, the ODE does not have a closed-form solution to compare against for a quantitative evaluation. But it is easily verified that [13]

$$V(t)=x_1(t)-\ln(x_1(t))+x_2(t)-\ln(x_2(t))=\text{const} \tag{32}$$

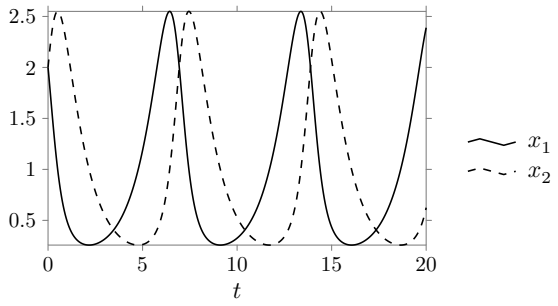


Figure 1: *Solutions to the Lotka-Volterra equation with* $x_1(0) = x_2(0) = 2$

holds, and we can compute the corresponding quantity $\widehat{V}(n)$ for the discretized states $\widehat{x}_1(n)$ and $\widehat{x}_2(n)$ and assess its time-variability. This is done is fig. 2 for different sampling intervals T, where for the same starting point, the difference between $\widehat{V}(n)$ and $V_{\text{ref}} = V(t) \approx 2.614$ is shown. One can nicely verify that the error scales approximately with T^2. The dotted line shows the same quantity when employing the implicit midpoint rule (with a full Newton solver). Obviously it performs quite similar and somewhat surprisingly, even slightly worse. This underlines that the non-iterative method is not just a crude approximation of the implicit midpoint method, but a discretization method in its own right.

5.2. CMOS-based inverting amplifier

The second example we study is a single CMOS-based inverting amplifier stage as found in the Red Llama [14], shown in fig. 3. Assigning the states x_1 and x_2 to the voltages across C_1 and C_2, respectively, and assuming the gate current to be zero, straightforward circuit analysis yields

$$\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix} = -\begin{pmatrix}0\\ \frac{x_2}{RC_2}\end{pmatrix} + \begin{pmatrix}\frac{1}{C_1}\\ \frac{1}{C_2}\end{pmatrix} i_{\text{i}}(u,x) \tag{33}$$

$$y = u - x_1 - x_2 \tag{34}$$

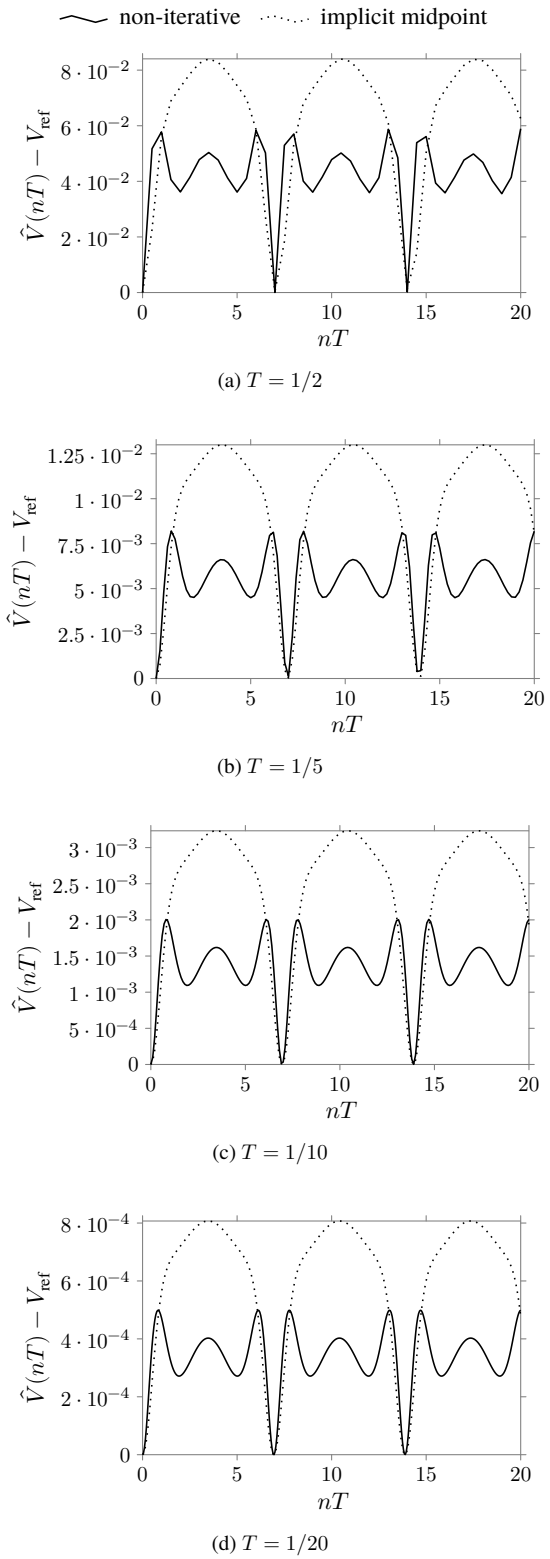
where

$$\begin{aligned} i_{\text{i}}(u,x) = {} & i_{\text{D}}(u - x_1, u - x_1 - x_2) \\ & - i_{\text{D}}(V_{\text{dd}} - u + x_1, V_{\text{dd}} - u + x_1 + x_2) \end{aligned} \tag{35}$$

is the CMOS inverter output current, i.e. the combined current through both MOSFET transistors, each of which is individually modelled by

$$i_{\text{D}}(v_{\text{GS}}, v_{\text{DS}}) = \begin{cases} 0 & \text{if } v_{\text{GS}} \le V_{\text{T}} \\ \alpha \cdot (v_{\text{GS}} - V_{\text{T}} - \frac{v_{\text{DS}}}{2}) v_{\text{DS}} & \text{if } v_{\text{DS}} \le v_{\text{GS}} - V_{\text{T}} \\ \frac{\alpha}{2} \cdot (v_{\text{GS}} - V_{\text{T}})^2 & \text{otherwise.} \end{cases} \tag{36}$$

A more complicated model with variable α and V_{T} like in [14] could also be employed. But even with this simple model, the non-linearity cannot be decomposed into a linear combination of univariate functions, so the non-iterative method as originally presented is not applicable, while applying (21) is straight-forward after working out the required Jacobian.



(a) $T = 1/2$

(b) $T = 1/5$

(c) $T = 1/10$

(d) $T = 1/20$

Figure 2: *Deviation of* $\widehat{V}(n)$ *from the correct* $V_{ref} \approx 2.614$ *for different sampling intervals* T *as obtained with the non-iterative method (solid) and the implicit midpoint method (dotted)*

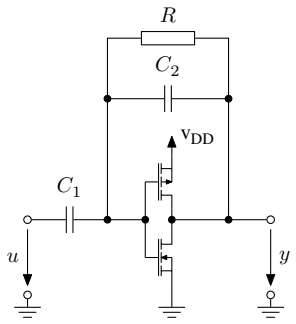
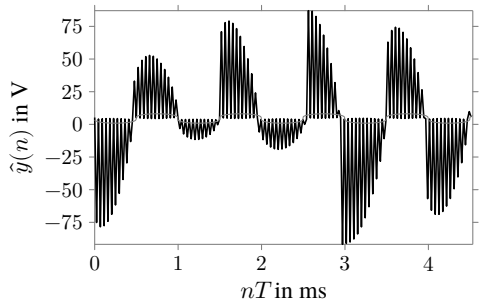


Figure 3: *CMOS-based inverting amplifier stage*

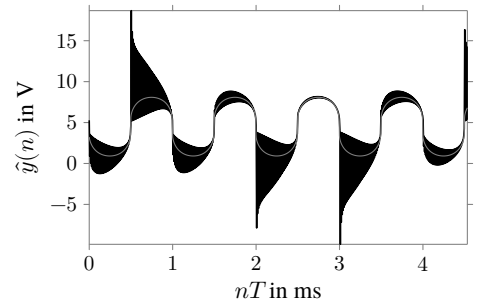
Setting $C_1 = 33\,\text{nF}, C_2 = 100\,\text{pF}, R = 1\,\text{M}\Omega, \alpha = 1\,\text{mA/V}^2$ and $V_\text{T} = 0.7\,\text{V}$ as taken from the Red Llama gives a small-signal gain of approximately $45\,\text{dB}$ at $1\,\text{kHz}$ and a pole above $36\,\text{kHz}$, so the simulation may be expected to present numerical challenges. For evaluation, the system is excited by a sinusoidal input u with an amplitude of $1\,\text{V}$ and a frequency of $1\,\text{kHz}$. And indeed, the output produced by the non-iterative method shown in fig. 4 for different oversampling factors M relative to the sampling rate $44.1\,\text{kHz}$ confirms that relatively high oversampling is required to obtain reasonable output. But note that especially for $M = 4$ and $M = 8$, the observed artifacts follow a pattern: There are instants with high errors followed by a phase of decaying oscillation around the proper solution. In the light of section 4, this is caused by the first Newton-Raphson iteration giving a bad result when the combination of old state and current input indicate operation in the linear regime, where the circuit has high gain, but the true result already being in the clipping region. The solution obtained by just a single Newton-Raphson iteration then overshoots the correct value, and the subsequent samples then converge towards the true solution. Of course, any explicit scheme will suffer such overshoots whenever the non-linear part of the ODE undergoes sudden changes. However, any subsequent lowpass filtering for downsampling (not considered here) would reduce those overshoots and also attenuate the Nyquist-rate oscillation following them.

Given that an oversampling rate of $M = 16$ appears to be necessary to obtain acceptable results, one may wonder whether the implicit midpoint rule with a full Newton solver at no or lower oversampling would not be more efficient overall. The answer is a potentially unsatisfactory "it depends." To gain some insight, the non-iterative method and the implicit midpoint method are compared for different oversampling rates M in terms of root-mean-square error (RMSE) and, for the implicit midpoint method, the required number of Newton iterations in table 1. The same sinusoidal input is used as before and the RMSE is computed with respect to a reference signal obtained with very high oversampling (as included in fig. 4). The Newton iterations have been terminated when the norm of the residual was below 10^{-3}, a value that is sufficiently low to ensure that errors due to the discretization scheme dominate in the output while not causing unduly many iterations.

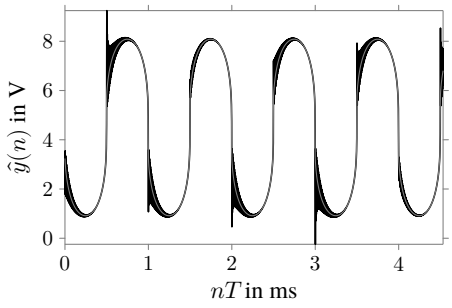
It can be observed that the RMSE at low M is also relatively high for the implicit midpoint point method. This is due to the fact that, for this numerically challenging ODE, the implicit midpoint method also tends to introduce high-frequency oscillations. To compare the computational cost, we assume it is dominated by computing the Jacobian and solving the linear system. Consequently, multiplying M with the average number of required New-



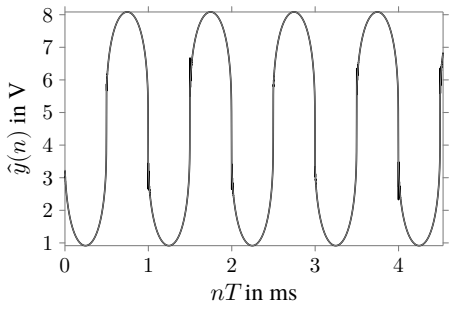
(a) $M = 1$



(b) $M = 4$



(c) $M = 8$



(d) $M = 16$

Figure 4: *Simulation results obtained for sinusoidal input at various oversampling factors M (relative to $44.1\,\text{kHz}$). Reference obtained with trapezoidal rule and high oversampling overlaid in gray.*

Table 1: *Comparison of non-iterative method and implicit midpoint method in terms of root-mean-square error (RMSE) and (for the latter) number of iterations of the Newton solver*

	non-iterative	implicit midpoint		
M	RMSE	RMSE	max. # iters	avg. # iters
1	35.507	1.218	12	4.013
4	2.143	0.534	11	2.991
8	0.346	0.109	10	1.829
12	0.080	0.036	9	1.470
16	0.044	0.018	9	1.283

ton iterations for the implicit midpoint method provides an estimate of the oversampling factor at which the non-iterative method could be operated at similar cost. For instance, the implicit midpoint method for $M = 8$ has a computational cost comparable to the non-iterative scheme at $M = 14$ or $M = 15$, as $8 \cdot 1.829 = 14.632$, but its RMSE is even higher than the non-iterative scheme at $M = 12$. Similarly, the implicit midpoint method at $M = 12$ has computational cost comparable to the non-iterative method at $M = 17$ and has a slightly lower RMSE than the latter at $M = 16$. Overall, for this ODE and this input signal, the two approaches exhibit similar efficiency. The non-iterative method, however, has the added benefit of a more even distribution of computational load over time and a slightly simpler implementation.

6. CONCLUSION AND OUTLOOK

As shown in this paper, the non-iterative method of [9, 10] in the second-order accurate variant can be generalized to be applicable to arbitrary vector-valued ODEs. This makes it useable in many more contexts. One aspect that needs further investigation is the question of stability. Initial results indicate that, depending on the ODE, the discretized system can be stable for arbitrary T or there can be an upper limit on T for stability (assuming stability of the original ODE). However, practical criteria for stability are still sought.

Another limitation is the restriction to explicit ODEs. If the evaluation of $f(x)$ itself requires numerical root-solving – which is often the case in virtual analog modeling – the benefits of the non-iterative method vanish, only possibly allowing to reduce the dimensionality of the nonlinear problem. This is especially true since implicit schemes like the implicit midpoint method or the trapezoidal rule can be combined with implicit ODEs in such a way that only a single implicit equation results. (This happens implicitly in methods that discretize individual elements, like wave digital filters or the nodal DK method.) However, the second contribution of this paper may show a path forward, namely the relationship between the non-iterative scheme and the midpoint method with early termination of the Newton-Raphson iteration. It may well be possible that a similar approach is possible for the resulting equation when the midpoint rule is applied to an implicit ODE. However, the details are more involved and exploration of this idea is left to future research.

7. ACKNOWLEDGEMENTS

The author would like to thank Maarten van Walstijn and Lasse Köper for stimulating discussions that led to the idea for this paper.

8. REFERENCES

[1] A. Sarti and G. De Poli, "Toward nonlinear wave digital filters," *IEEE Transactions on Signal Processing*, vol. 47, no. 6, pp. 1654–1668, 1999.

[2] K.J. Werner, V. Nangia, A. Bernardini, J.O. Smith, and A. Sarti, "An improved and generalized diode clipper model for wave digital filters," in *Audio Engineering Society Convention 139*. Audio Engineering Society, 2015.

[3] A. Bernardini, *Special Topics in Information Technology*, chapter Advances in Wave Digital Modeling of Linear and Nonlinear Systems: A Summary, pp. 3–15, Springer International Publishing, Cham, 2020.

[4] K. Dempwolf, M. Holters, and U. Zölzer, "Discretization of parametric analog circuits for real-time simulations," in *Proceedings of the 13th International Conference on Digital Audio Effects (DAFx'10)*, 2010.

[5] D.T. Yeh, J.S. Abel, and J.O. Smith, "Automated physical modeling of nonlinear audio circuits for real-time audio effects—part I: Theoretical development," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 4, pp. 728–737, 2010.

[6] M. Holters and U. Zölzer, "A generalized method for the derivation of non-linear state-space models from circuit schematics," in *Proceedings of the 23rd European Signal Processing Conference (EUSIPCO)*, 2015, pp. 1073–1077.

[7] A. Falaize and T. Hélie, "Guaranteed-passive simulation of an electro-mechanical piano: A port-Hamiltonian approach," in *Proceedings of the 18th International Conference on Digital Audio Effects (DAFx-15)*, 2015.

[8] R. Muller and T. Hélie, "Power-balanced modelling of circuits as skew gradient systems," in *Proceedings of the 21st International Conference on Digital Audio Effects (DAFx-18)*, 2018.

[9] M. Ducceschi, S. Bilbao, and C.J. Webb, "Non-iterative schemes for the simulation of nonlinear audio circuits," in *Proceedings of the 24th International Conference on Digital Audio Effects (DAFx-21)*, 2021, pp. 25–32.

[10] M. Ducceschi and S. Bilbao, "Non-iterative simulation methods for virtual analog modelling," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 30, pp. 3189–3198, 2022.

[11] F.G. Germain, "Fixed-rate modeling of audio lumped systems: A comparison between trapezoidal and implicit midpoint methods," in *Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17)*, 2017, pp. 168–175.

[12] V. Chatziioannou, A. Hofmann, and S. Schmutzhard, "A real-time physical model to simulate player control in woodwind instruments," in *Proceedings of the International Symposium on Music Acoustics (ISMA)*, 2019, pp. 279–283.

[13] Y. Nutku, "Hamiltonian structure of the lotka-volterra equations," *Physics Letters A*, vol. 145, no. 1, pp. 27–28, 1990.

[14] L. Köper and M. Holters, "Taming the Red Llama–modeling a CMOS-based overdrive circuit," in *Proceedings oft he 23rd International Conference on Digital Audio Effects (DAFx-20)*, 2020, pp. 54–61.