EXACT DISCRETE-TIME REALIZATIONS OF THE GAMMATONE FILTER

Elizabeth Ren and Hans-Andrea Loeliger

ETH Zurich, Dept. of Information Technology & Electrical Engineering
{ren, loeliger}@isi.ee.ethz.ch

ABSTRACT

The paper derives an exact discrete-time state space realization of the popular gammatone filter. No such realization appears to be available in the literature. The proposed realization is computationally attractive: a gammatone filter with exponent N requires less than $6N$ multiplications and additions per sample. The integer coefficients of the realization can be computed by a simple recursion. The proposed realization also yields a closed-form expression for the frequency response. The proposed primary realization is not quite in a standard form, but it is easily transformed into another realization whose state transition matrix is in Jordan canonical form.

Index Terms— gammatone filter, auditory filter, discrete-time IIR filter, state space representation

1. INTRODUCTION

Filter banks using gammatone filters have long been popular in applications including mammalian cochlea modeling [1–4], pitch extraction [5], sound source localization [6], speech and audio coding [7–9], speech recognition and non-speech audio classification [10, 11], speech enhancement [12], and unsupervised signal separation [13].

Gammatone filters are traditionally defined as continuous-time bandpass filters with a causal impulse response of the form

$$g(t) = at^{N-1}e^{-\lambda t}\cos(2\pi ft + \phi) \tag{1}$$

for $t \geq 0$ (cf. Fig. 1), which may be viewed as the product of the Gamma distribution from statistics ($t^{N-1}e^{-\lambda t}$) and a cosine tone (up to a scale factor). The main parameters of such a filter are the center frequency f, the bandwidth parameter λ, and the exponent N (in the literature called "order"). The phase ϕ (which is often set to zero) and the scale factor a are of secondary importance. The exponent N is often set to 4 since values in the range $3-5$ yield a good fit to human auditory filtering [14].

A gammatone filter bank consists of multiple gammatone filters with different frequency and bandwidth parameters. Typically, the frequencies are linearly spaced on the equivalent rectangular bandwidth (ERB) scale [5, 15].

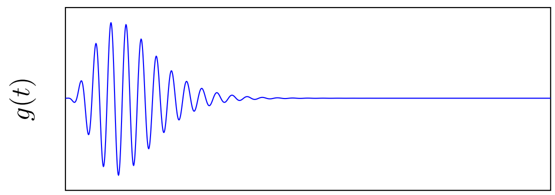

Fig. 1: Impulse response of a gammatone filter with $N = 4$.

For digital signal processing, a discrete-time version of the gammatone filter is needed. Clearly, finite-impulse-response (FIR) filters (e.g., as in [16]) are not very suitable: for a good approximation, the computational load per sample is high and grows linearly with the sampling rate. Therefore, a number of approximate infinite-impulse-response (IIR) realizations of the gammatone filter have been developed. In [5, 17, 18] approximate IIR realizations are obtained by expressing the transfer function as a cascade of lower order filters (with some simplifications in the case of [5, 17]) before applying a suitable transform (e.g., the impulse invariant transform) to discrete time. Another approach begins by sampling a complex version of (1), which is then approximately realized [19, 20]. However, somewhat surprisingly, no exact discrete-time realization of the gammatone filter seems to be available in the literature.

In this paper, we derive an exact realization of a causal discrete-time impulse response of the form

$$g[k] = g(kT) \tag{2}$$
$$= aT^{N-1}k^{N-1}e^{-\lambda Tk}\cos(2\pi fTk + \phi) \tag{3}$$

for $k \geq 0$, where $T > 0$ is the sampling period. The proposed realization is based on a complex state space model of dimension N. The resulting state space filter requires less than $6N$ multiplications and additions per sample, independently of the other filter parameters and of the sampling period T.

The mathematical facts underlying this paper are basically well known, but the apparent absence from the literature of an exact IIR realization of the gammatone filter indicates that the explicit derivations and results presented here may be useful.

The paper is structured as follows. The basic IIR realization is derived in Section 2. Section 3 addresses the transfer function and the frequency response. Some variations of the basic IIR realization are outlined in Section 4, and Section 5 concludes the paper.

2. A FIRST REALIZATION

2.1. On State Space Realizations

We will use complex state space models with real input $u[k]$ and real output $y[k]$ of the form

$$x[k] = Ax[k-1] + bu[k] \tag{4}$$
$$y[k] = \mathrm{Re}\big(c^{\mathsf{T}}x[k]\big) \tag{5}$$

with $x[k] \in \mathbb{C}^N$, $A \in \mathbb{C}^{N\times N}$, and $b, c \in \mathbb{C}^N$. Clearly, the impulse response of such a model is

$$h[k] = \begin{cases} 0, & \text{if } k < 0 \\ \mathrm{Re}\big(c^{\mathsf{T}}A^k b\big), & \text{if } k \geq 0. \end{cases} \tag{6}$$

The same impulse response (6) can also be realized in the more standard form

$$\xi[k+1] = A\xi[k] + \tilde{b}u[k] \tag{7}$$
$$y[k] = \mathrm{Re}\Big(\tilde{c}^{\mathsf{T}}\xi[k] + \tilde{d}u[k]\Big) \tag{8}$$

with $\tilde{d} = c^{\mathsf{T}}b$ and either with $\tilde{b} = Ab$ and $\tilde{c} = c$ or with $\tilde{b} = b$ and $\tilde{c}^{\mathsf{T}} = c^{\mathsf{T}}A$. However, the derivation below is more easily expressed in terms of (4) and (5).

2.2. The Realization

We begin by noting that the impulse response (3) of the gammatone filter can be written as

$$g[k] = \mathrm{Re}\big(\alpha k^{N-1}\gamma^k\big) \tag{9}$$

(for $k \geq 0$) with complex parameters

$$\alpha = aT^{N-1}e^{i\phi} \tag{10}$$

and

$$\gamma = e^{-\lambda T}e^{i2\pi fT}. \tag{11}$$

In (9), for $N = 1$ and $k = 0$, we define $k^{N-1} = 1$, in agreement with (1).

For the following theorem, we define $\beta_\ell^{(n)}$ (for integers n and ℓ) as

$$\beta_1^{(n)} = \left\{ \begin{array}{ll} 1, & \text{if } n = 0 \\ 0, & \text{if } n > 0, \end{array} \right. \tag{12}$$

$$\beta_\ell^{(n)} = \sum_{\nu=0}^{n-1} \binom{n}{\nu} \beta_{\ell-1}^{(\nu)} \tag{13}$$

for $1 < \ell \leq n+1$, and $\beta_\ell^{(n)} = 0$ otherwise.

ℓ	1	2	3	4	5
$\beta_\ell^{(0)}$	1	0	0	0	0
$\beta_\ell^{(1)}$	0	1	0	0	0
$\beta_\ell^{(2)}$	0	1	2	0	0
$\beta_\ell^{(3)}$	0	1	6	6	0
$\beta_\ell^{(4)}$	0	1	14	36	24

Table 1: Coefficients $\beta_\ell^{(n)}$ according to (12) and (13).

Theorem 1 *The impulse response (6) of the state space model (4), (5) with*

$$A = \gamma \begin{bmatrix} 1 & 1 & 0 & \dots & 0 \\ 0 & 1 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 & 1 \\ 0 & \dots & \dots & 0 & 1 \end{bmatrix}, \tag{14}$$

$c^{\mathsf{T}} = [1, 0, \dots, 0]$*, and* $b = [b_1, \dots, b_N]^{\mathsf{T}}$ *with*

$$b_\ell = \alpha\beta_\ell^{(N-1)} \tag{15}$$

is (9).

A block diagram of the state space model of Theorem 1 is given in Fig. 2. The coefficients $\beta_\ell^{(N-1)}$ up to $N = 5$ are listed in Table 1. Note that these integer coefficients do not depend on any parameters of the gammatone filter except the exponent N.

2.3. Proof of Theorem 1

Without loss of generality, we assume $\alpha = 1$ and $\gamma \neq 0$.

Lemma *Let the input to the system in Fig. 2 be* $u[k] = \delta[k]$ *(the Kronecker delta). For any fixed* $n \in \mathbb{Z}$ *with* $0 \leq n < N$*, let* $b_\ell = \beta_\ell^{(n)}$ *(implying* $b_\ell = 0$ *for* $n+1 < \ell \leq N$*). Then*

$$x_1[k] = k^n\gamma^k \tag{16}$$

for $k \geq 0$*.*

Note that specializing this lemma to $n = N-1$ yields Theorem 1.

For $n = 0$, $b = [1, 0, \dots, 0]^{\mathsf{T}}$ (cf. Table 1), and $x_1[k] = \gamma^k = k^n\gamma^k$ (for $k \geq 0$) is obvious from Fig. 2. For $n > 0$, we prove the lemma by induction. For the induction step, we assume

$$b = \big[\beta_1^{(\nu)}, \dots, \beta_{\nu+1}^{(\nu)}, 0, \dots, 0\big]^{\mathsf{T}} \ \text{ yields } \ x_1[k] = k^\nu\gamma^k \tag{17}$$

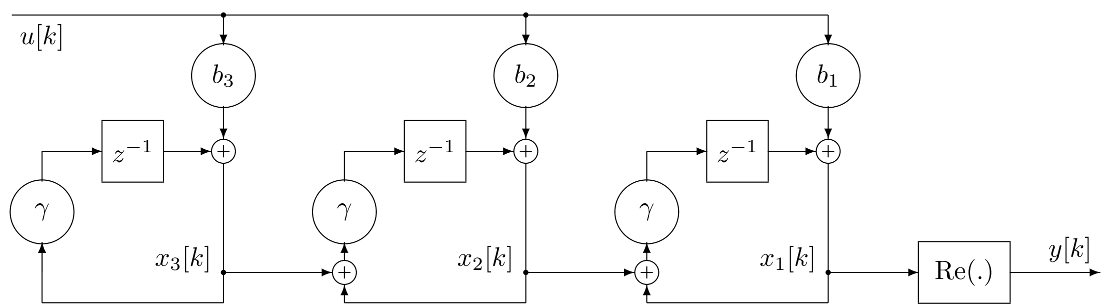

Fig. 2: Block diagram of the complex state space realization according to Theorem 1 for $N = 3$. The boxes labeled z^{-1} are unit-delay cells. (The contents of these delay cells are the state variables in (7) and (8).)

for $k \geq 0$ and $0 \leq \nu < n$. We then show that (17) holds also for $\nu = n$. A key point in the proof is the fact (which is obvious from Fig. 2) that (17) implies

$$b = \big[0, \beta_1^{(\nu)}, \ldots, \beta_{\nu+1}^{(\nu)}, 0, \ldots, 0\big]^{\mathsf{T}} \text{ yields } x_2[k] = k^\nu \gamma^k \tag{18}$$

for $k \geq 0$ and $0 \leq \nu < n$.

Now assume (17) holds for fixed $n > 0$ and all $k \geq 0$ and $0 \leq \nu < n$. We prove the validity of (17) for $\nu = n$ by induction over k. Let

$$b = \big[\beta_1^{(n)}, \ldots, \beta_{n+1}^{(n)}, 0, \ldots, 0\big]^{\mathsf{T}}. \tag{19}$$

For $k = 0$, we then clearly have $x_1[k] = 0$ (cf. Table 1), in agreement with (17). Assuming that (17) holds for some $k \geq 0$, we have

$$x_1[k+1] = \gamma\big(x_1[k] + x_2[k]\big) \tag{20}$$
$$= \gamma k^n \gamma^k + \gamma x_2[k]. \tag{21}$$

Thus (17) holds for $\nu = n$ if and only if

$$(k+1)^n \gamma^{k+1} = \gamma k^n \gamma^k + \gamma x_2[k], \tag{22}$$

i.e., if and only if

$$x_2[k] = \big((k+1)^n - k^n\big)\gamma^k \tag{23}$$
$$= \sum_{\nu=0}^{n-1} \binom{n}{\nu} k^\nu \gamma^k. \tag{24}$$

But inserting (12) and (13) in (19) yields

$$b = \sum_{\nu=0}^{n-1} \binom{n}{\nu} \big[0, \beta_1^{(\nu)}, \ldots, \beta_{\nu+1}^{(\nu)}, 0, \ldots, 0\big]^{\mathsf{T}}. \tag{25}$$

Using (18), we then see that (24) indeed holds.

2.4. Computational Cost

The required numbers of multiplications and additions per sample are easily seen from Fig. 2: for general $\alpha \in \mathbb{C}$ (and noting that $b_1 = 0$), we have N complex multiplications (by γ), $N - 1$ real-times-complex multiplications, and $2(N-1)$ complex additions, amounting to $6N - 2$ real multiplications and $6N-4$ real additions. In the important special case where α is real or imaginary, the numbers are slightly smaller, cf. Table 2. Further simplifications may be achieved on some processors by noting that the entries in Table 1 are small integers rather than general floating-point numbers.

	multiplications	additions
general ϕ	$6N-2$	$6N-4$
$\phi \in \{0, \pi, \pm\pi/2\}$	$5N-1$	$5N-3$

Table 2: Computations per sample (assuming $N > 1$). In the second row, α is real or purely imaginary.

3. TRANSFER FUNCTION AND FREQUENCY RESPONSE

The transfer function

$$G(z) = \sum_{k=0}^{\infty} g[k] z^{-k} \tag{26}$$

is easily obtained from Theorem 1 as follows. By inspection of Fig. 2, we find the transfer function of the complex part (from $u[k]$ to $x_1[k]$) to be

$$G_{\mathrm{c}}(z) = \sum_{\ell=1}^{N} \frac{b_\ell (\gamma z^{-1})^{\ell-1}}{(1-\gamma z^{-1})^\ell} \tag{27}$$
$$= \frac{\sum_{\ell=1}^{N} b_\ell (\gamma z^{-1})^{\ell-1} (1-\gamma z^{-1})^{N-\ell}}{(1-\gamma z^{-1})^N} \tag{28}$$
$$= \frac{\sum_{k=0}^{N-1} \lambda_k z^{-k}}{(1-\gamma z^{-1})^N} \tag{29}$$

with coefficients

$$\lambda_k = \gamma^k \sum_{\ell=1}^{k+1} b_\ell \binom{N-\ell}{k-\ell+1} (-1)^{k-\ell+1}. \tag{30}$$

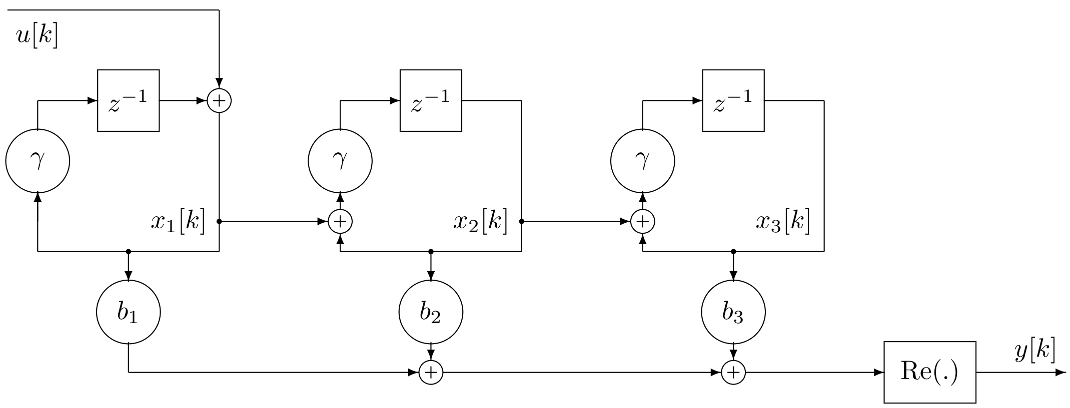

Fig. 3: The transpose realization according to Section 4.1 for $N = 3$.

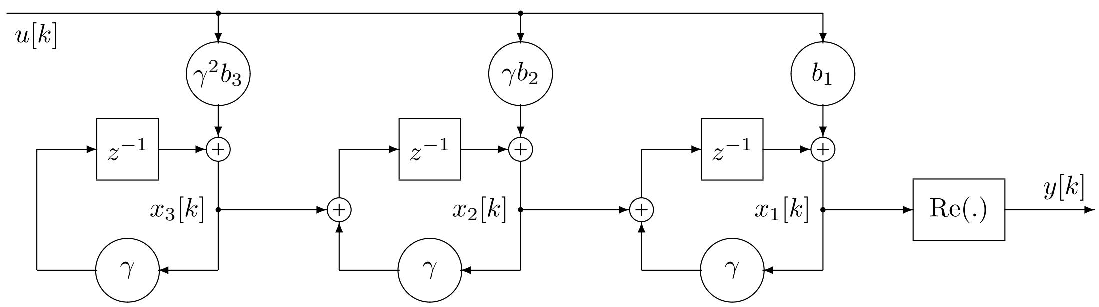

Fig. 4: A realization (4), (5) with state transition matrix A in Jordan form. The contents of the delay cells are the state variables in (7) and (8).

The overall transfer function (26) is thus

$$G(z) = \frac{1}{2} \sum_{k=0}^{\infty} \left(g_c[k] + \overline{g_c[k]} \right) z^{-k}, \tag{31}$$

from which we obtain the frequency response

$$G(e^{i\Omega}) = \frac{1}{2} \left(G_c(e^{i\Omega}) + \overline{G_c(e^{-i\Omega})} \right). \tag{32}$$

4. ALTERNATIVE REALIZATIONS

4.1. Transpose Realization

Since

$$c^\mathsf{T} A^k b = b^\mathsf{T} (A^\mathsf{T})^k c, \tag{33}$$

the impulse response (6) remains unchanged if we exchange c and b, and change A into A^T. Fig. 2 is thus changed into Fig. 3. The computational effort (as in Table 2) turns out to remain the same.

4.2. Realizations in Jordan Canonical Form

The matrix (14) is almost in Jordan canonical form [21], but not quite. However, the realization according to Theorem 1 is easily modified so that the state transition matrix is indeed in the Jordan form

$$A = \begin{bmatrix} \gamma & 1 & 0 & \dots & 0 \\ 0 & \gamma & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \gamma & 1 \\ 0 & \dots & \dots & 0 & \gamma \end{bmatrix}. \tag{34}$$

As illustrated in Fig. 4, the modification consists of moving the multiplications by γ in Fig. 2 behind the adders, which is compensated by extra powers of γ in the input vector b. (By contrast, the vector b in Fig. 2 does not depend on γ.)

Of course, the realization of Fig. 4 can also be transposed as in Section 4.1.

5. CONCLUSION

An exact discrete-time realization of the gammatone filter appears to be missing in the literature. We derived such a realization and outlined some variations of it. The proposed realization (Fig. 2) appears to be new and attractive.

6. REFERENCES

[1] D. Wang and G. J. Brown, *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications*, chapter 1, Wiley/IEEE Press, 2006.

[2] R. D. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C. Zhang, and M. Allerhand, "Complex sounds and auditory images," in *Auditory Physiology and Perception, Proc. 9th Int. Symp. Hearing*, 1992, pp. 429–446.

[3] P. I. M. Johannesma, "The pre-response stimulus ensemble of neurons in the cochlear nucleus," in *Proc. Symp. Hearing Theory*, 1972, pp. 58–69.

[4] E. de Boer, "Synthetic whole-nerve action potentials for the cat," *J. Acoust. Soc. Am.*, vol. 58, no. 5, pp. 1030–1045, 1975.

[5] R.D. Patterson, I. Nimmo-Smith, J. Holdsworth, and P. Rice, "Svos final report: The auditory filterbank," Tech. Rep., MRC Applied Psychology Unit, Cambridge, 1988.

[6] N. Ma, T. May, and G. J. Brown, "Exploiting deep neural networks and head movements for robust binaural localisation of multiple sources in reverberant environments," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 12, pp. 2444–2453, 2017.

[7] G. Kubin and W. B. Kleijn, "On speech coding in a perceptual domain," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, 1999, vol. 1, pp. 205–208.

[8] E. Ambikairajah, J. Epps, and L. Lin, "Wideband speech and audio coding using gammatone filter banks," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, 2001, vol. 2, pp. 773–776.

[9] S. Tabibi, A. Kegel, W. K. Lai, and N. Dillier, "Investigating the use of a gammatone filterbank for a cochlear implant coding strategy," *J. Neuroscience Methods*, vol. 277, pp. 63–74, 2017.

[10] R. Schlüter, I. Bezrukov, H. Wagner, and H. Ney, "Gammatone features and feature combination for large vocabulary speech recognition," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, 2007, vol. 4, pp. 645–649.

[11] X. Valero and F. Alias, "Gammatone cepstral coefficients: Biologically inspired features for non-speech audio classification," *IEEE Trans. Multimedia*, vol. 14, no. 6, pp. 1684–1689, 2012.

[12] T. S. Gunawan, E. Ambikairajah, and J. Epps, "Perceptual speech enhancement exploiting temporal masking properties of human auditory system," *Speech Communication*, vol. 52, no. 5, pp. 381–393, 2010.

[13] B. Gao, W. L. Woo, and S. S. Dlay, "Unsupervised single-channel separation of nonstationary signals using gammatone filterbank and itakura-saito nonnegative matrix two-dimensional factorizations," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 60, no. 3, pp. 662–675, 2013.

[14] R. D. Patterson and B. C. J. Moore, "Auditory filters and excitation patterns as representations of frequency resolution," in *Frequency Selectivity in Hearing*, pp. 123–177. Academic Press Ltd., London, 1986.

[15] B. R. Glasberg and B. C. J. Moore, "Derivation of auditory filter shapes from notched-noise data," *Hearing Research*, vol. 47, pp. 103–138, 1990.

[16] L. Van Immerseel and S. Peeters, "Digital implementation of linear gammatone filters: Comparison of design methods," *Acoustic Research Letters Online*, vol. 4, no. 3, 2003.

[17] M. Slaney, "An efficient implementation of the patterson-holdsworth auditory filter bank," Tech. Rep., Apple Computer, Inc., 1993.

[18] A. M. Darling, "Properties and implementation of the gammatone filter: a tutorial," *Speech Hearing and Language, work in progress*, vol. 5, pp. 43–61, 1991.

[19] V. Hohmann, "Frequency analysis and synthesis using a gammatone filterbank," *Acta Acust.*, vol. 88, no. 3, pp. 433–442, 2002.

[20] T. Herzke and V. Hohmann, "Improved numerical methods for gammatone filterbank analysis and synthesis," *Acta Acust.*, vol. 93, no. 3, pp. 498–500, 2007.

[21] R. A. Horn and C. R. Johnson, *Matrix Analysis*, chapter 3, Cambridge University Press, 1985.