

# OPTIMIZATION TECHNIQUES FOR PARAMETRIC MODELING OF ACOUSTIC SYSTEMS AND MATERIALS

PACS: 43.55.Ka

Matti Karjalainen, Tuomas Paatero, and Miikka Tikander  
Helsinki University of Technology  
Laboratory of Acoustics and Audio Signal Processing  
P.O.Box 3000, FIN-02015, HUT  
Espoo  
Finland  
Tel: +358 9 451 2490  
Fax: +358 9 460 224  
E-mail: [matti.karjalainen@hut.fi](mailto:matti.karjalainen@hut.fi)

## ABSTRACT

Linear and time-invariant signal and system models are useful in compact characterization of acoustic transfer functions. In addition to compact representations of responses, such models are efficient in simulating acoustic systems for sound synthesis, artificial reverberation, etc. In this paper we propose parametric modeling techniques for room impulse responses (RIRs), in-situ acoustic material measurements, and musical instrument modeling, based on ARMA models — including Kautz filter models — which require nonlinear optimization of the parameters. Low order models are applied to surface impedance modeling, and high-order models are used for complex responses such as RIRs and musical instruments.

## INTRODUCTION

Physical modeling of acoustic systems is in practice based on numerical simulation, for example by solving partial difference equations. Often the problem is to model the response from a point to another so that the spatial distribution of waves is not of primary interest. This leads to using transfer functions and signal modeling. The signal processing approach has the advantage of being computationally highly efficient, which is important especially when real-time simulation is needed. This is the case for example in model-based sound synthesis of musical instruments, audio effects such as artificial reverberators, auralization of room or concert hall acoustics in acoustics design software, or equalization of loudspeaker-room responses in audio reproduction. In all these cases the system to be modeled can be considered being linear and time-invariant (LTI), and in practice also stable and causal. Further cases where the signal processing methodology is highly useful can be found in acoustic and audio measurements.

For LTI acoustic systems the impulse response or corresponding frequency response can be measured and the task is often to find a compact parametric model to capture the essential features of the target response. Digital filters, in the form of finite impulse response (FIR) and infinite impulse response (IIR) filters, are powerful means of modeling target responses, both from analysis and synthesis points of view. In analysis applications we may obtain essential information about the measured process, such as (eigen)mode parameters of a room or a musical instrument body. In synthesis applications a compact parametric model in the form of a digital filter makes it possible to simulate a given system in real time.

Although LTI models are linear in input-output relationships, the solving for optimal model parameters is not necessarily so. There are a number of parameter estimation techniques for LTI systems [1,2]. MA (moving average) modeling leads to an FIR type filter model with  $a_k$

coefficients being zero in its z-transform, in Eq.(1) below. While MA models are easily estimated, for reverberant and resonating systems they are not compact, computationally efficient, or analytically interesting. AR (autoregressive) models, with only  $b_0$  being non-zero in the numerator of z-transform in Eq.(1), are able to describe recursive feedback in a system, and are also relatively straightforward due to linearity of normal equations used to estimate filter parameters for example in the linear predictive (LP) autocorrelation method [3]. AR models are however, due to their minimum-phase property, limited in modeling capabilities, e.g., in modeling responses exhibiting inherent latencies in their frequency components. ARMA (autoregressive moving average) modeling is the most general form of LTI modeling with both  $a_k$  and  $b_k$  non-zero coefficients in Eq.(1), but there is no way to solve the optimal model parameters in a closed form, and thus iterative techniques of nonlinear optimization are needed, which brings potential problems in convergence, e.g., trapping to a local minimum not close to the global optimum.

$$H(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_N z^{-N}}{1 + a_1 z^{-1} + \dots + a_P z^{-P}} = \frac{\sum_{k=0}^N b_k z^{-k}}{1 + \sum_{k=1}^P a_k z^{-k}} \quad (1)$$

Nonlinear optimization of  $a_k$  and  $b_k$  in a general case is realized in iterative techniques such as Prony's method [4] or Steiglitz-McBride method [1] (both are available in the Matlab Signal Processing Toolbox). In a more general case of (low-order) nonlinear model fitting, iterative methods (such as Matlab function 'curvfit') are available also for constraints in model parameters, as is shown in the surface impedance modeling case below.

In this paper we will explore cases where nonlinear optimization techniques are applied to acoustical modeling problems. We will explore parametric modeling for room impulse responses (RIRs), in-situ acoustic material measurements, and musical instrument modeling, based on pole-zero models, particularly orthonormal Kautz filters. Low-order models are applied to surface impedance case and high-order models are used for complex responses, such as RIRs and musical instruments.

## MODELING OF ROOM IMPULSE RESPONSES

Room impulse responses can be measured easily with modern computerized means, e.g., by using maximum length sequences (MLS). In principle about  $1.5 \times T_{60}$ , where  $T_{60}$  is reverberation time, should be measured to obtain a 90 dB dynamic range, but in practice the achieved dynamic range is often limited to 40-60 dB. Thus a small room response may contain 10000-50000 samples and a concert hall may contain even more than 100000 samples for a sample rate of 48000 Hz. The goal of AR or ARMA modeling of such a response is either to obtain acoustically interesting parametric information or a compact filter model for a practical application.

It turns out that it is not possible to estimate a single ARMA model for the full audio range. This is due to problems in convergence, numerical precision, and related instability of iterative estimation procedures. Even filter orders below 100 ( $P$  and/or  $N$  in Eq.(1)) may turn unstable and above 300 it is seldom possible to get useful results at all. Here we first introduce ARMA modeling based on Kautz filters instead of more traditional pole-zero modeling. Case studies of room responses start with an example of limiting the modeling to low frequencies only. Then we discuss frequency zooming techniques, finally carrying out a demanding concert hall case using frequency-zooming Kautz modeling.

### Kautz Functions and Models

Kautz filters [5,6] are a special class of fixed-pole IIR filters organized structurally to produce orthonormal tap-output impulse responses. The lowest order such filters, corresponding to any given set of desired stable poles, constitute an efficient tapped transversal structure, depicted in Fig.1. A particular Kautz filter is defined by a set of poles  $z_i$  in the unit circle and with a corresponding set of somehow assigned tap-output weights  $w_i$ , Eq.(2). Because of the transversal appearance and tap-output orthonormality, Kautz filters can be seen as a generalizations of the FIR filter. For more details on Kautz filter formulations and aspects of audio applications, see [7].

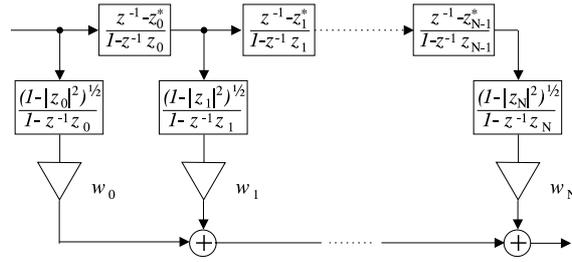


Figure 1. The Kautz filter. For  $z_i = 0$  in Eq.(2) it degenerates to an FIR filter, for  $z_i = a$ ,  $-1 < a < 1$ , it is a Laguerre filter where the tap filters can be replaced by a common pre-filter.

$$H(z) = \sum_{i=1}^N w_i \left( \frac{\sqrt{1 - z_i z_i^*}}{1 - z_i z_i^{-1}} \prod_{j=1}^{i-1} \frac{z^{-1} - z_j^*}{1 - z_j z_j^{-1}} \right) \quad (2)$$

Here we address the approximation of a given target response  $h(n)$  (or  $H(z)$ ) using Kautz filters. The filter parametrization, i.e., choosing of the filter weights  $w_i$ , is done in the least-square sense. For solving optimal values for Kautz poles  $z_i$ , we have adopted a method from [8], called here the BU-method [7]. In the following examples this estimation method is used due to its favorable properties originating from the orthonormality of Eq.(2), although other ARMA techniques can be used also.

#### Modeling of Low-Frequency Response of a Room

First we apply Kautz models through the BU-method in (low-pass filtered and down-sampled) frequency range of 0-200 Hz to an impulse response, measured in a room of size  $5.5 \times 6.5 \times 2.7 \text{ m}^3$ . Note the relatively long modal decay time (about 1.3 s) at low frequencies. Figure 2 depicts the modeling accuracy both by magnitude response and by temporal envelope. A Kautz order 80 already yields a good fit and with orders  $>100$  the result is perfect for practical purposes.

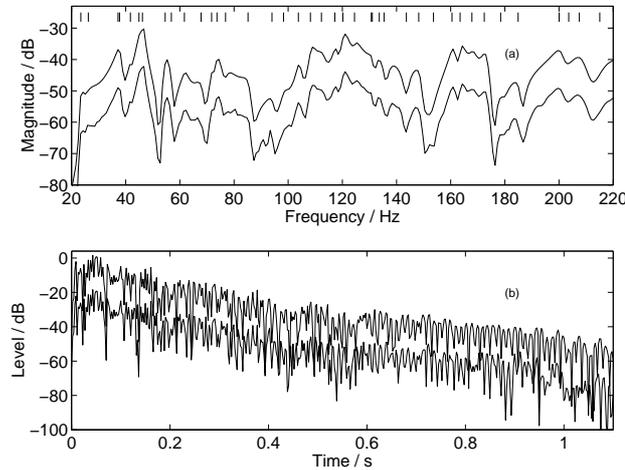


Figure 2. Low-frequency Kautz modeling (0-200 Hz) for a room, (a) magnitude responses of 80th order model (top) and target (bottom), with lines indicating Kautz pole positions, and (b) the decay envelope (top) compared to the target response (bottom).

#### Frequency-Zooming ARMA Modeling

Particularly for large rooms, such as concert halls, the required model order to capture the whole time-frequency range of diffuse modal pattern is not within the capability of direct ARMA modeling. There is, however, a technique to partition a measured response in the frequency domain and do the modeling in subbands as discussed in [9]. We call this *frequency-zooming*

ARMA (FZ-ARMA) modeling. In FZ-ARMA we modulate a subband down in frequency, apply low-pass filtering and decimation, and finally do ARMA modeling in this decimated sample rate domain. This can be done band-by-band, covering the whole audio range. The obtained model can be used as a highly efficient multirate decimated filterbank, or alternatively a composite pole-zero filter for the original sample rate is constructed from the subband filters. The zooming factor  $K_{\text{zoom}}$  for decimation can be selected so that the filter order within each subband remains manageable.

As a challenging case we have applied FZ-ARMA to a 700-seat concert hall, the room response measured by an omnidirectional sound source (100-10000 Hz) at the stage and a microphone in the 7<sup>th</sup> row on the main floor. The audio range was partitioned in frequency bands of 250 Hz, and Kautz models of order around 50-300 (experimentally hand-tuned for each band) were used. The modeling accuracy was tested visually by different time-frequency plots and by informal listening to the original and the modeled response. The required filter order was highest (about 300) around 500-1000 Hz, while above 5-10 kHz much lower orders are enough. Note that from an auditory point of view even lower filter orders should yield perfect late reverberation [10]. Figure 3 shows the waterfall plots of time-frequency behavior of the original and the frequency-zooming Kautz modeled impulse response.

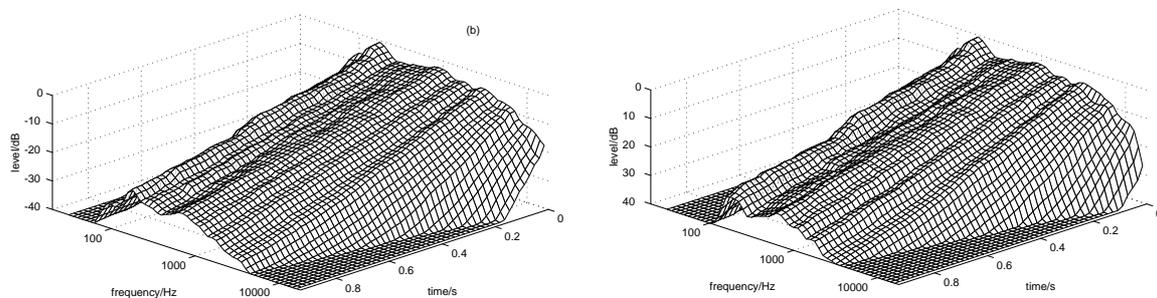


Figure 3: 1/3-octave waterfall time-frequency plots of original (left) and frequency-zooming Kautz modeled concert hall response.

## MODELING OF MUSICAL INSTRUMENTS

As a case of musical instrument sound, the analysis and modeling of bell sounds is explored. A characteristic feature of bell sounds is that they are composed of an inharmonic set of partials [11], such as the one described by magnitude spectrum in Fig.4(b). Each partial is a decaying sinusoid that, in a closer inspection, Fig.4(c), turns out to be a pair or a group of modes very closely located in frequency. This leads to perceptually noticeable beating. In this case the modal group consists primarily of two modes with a frequency difference of about 2.5 Hz.

FZ-ARMA is an excellent method for analyzing the modal groups of bell sounds. Figure 4 shows the envelope match obtained with three different FZ-ARMA orders for the 1310 Hz modal group. The zooming factor  $K_{\text{zoom}}$  is 200 in each case and Steiglitz-McBride iteration was used instead of Kautz modeling. In view (e) the orders are  $N = 0$  and  $P = 4$  cf. Eq.(1). Two pole pairs should in principle be sufficient for a double mode, but this all-pole (AR) case with  $N = 0$  does not allow proper phase matching, and thus the overall match remains poor. For ARMA orders  $N = 40$  and  $P = 4$  in (f) the relatively high number of zeros allows for excellent match with just two pole pairs. The same can be achieved with orders  $N = 2$  to  $P = 6$ , i.e., by adding an extra pole pair and keeping the number of zeros minimal. For all resonances up to 10 kHz for this bell sound, filter orders of  $N = 2...4$  and  $P = 4...6$  are sufficient for good modal decay matching so that a parallel filter, composed of modal group filters with a total order of about  $N = 40$  and  $P = 50$ , can implement efficient and high-quality synthesis for the bell sound at a sampling rate of 22050 Hz.

Other musical instruments (guitar, piano) have been analyzed and modeled using a similar approach, see [9].

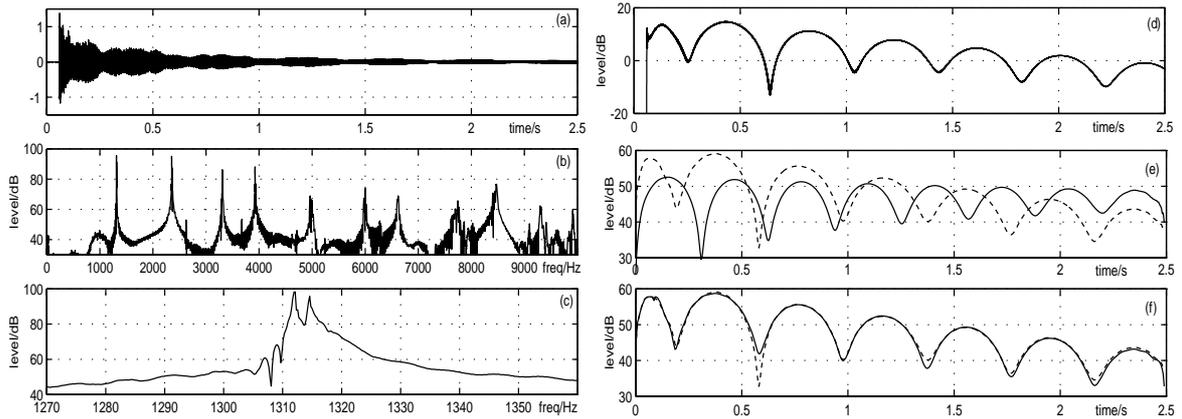


Figure 4. Analysis and modeling of a small bell sound: (a) recorded time-domain signal, (b) magnitude spectrum, (c) magnitude spectrum in the modal region around 1310 Hz, (d) decay envelope of the 1310 Hz modal group, (e) FZ-ARMA(0,4), and (f) FZ-ARMA(40,4).

## MODELING OF ACOUSTIC SURFACE IMPEDANCE

As the last example, a modeling problem of different scope is discussed. Acoustic surface impedance can be measured by different techniques. Particularly in *in-situ* measurements the accuracy tends to remain poor at low frequencies and when absorption is low. By low-order model fitting the data can be smoothed or the material can be characterized with few parameters.

Figure 5 (left) shows an *in-situ* measurement setup [11] to obtain reflection impulse response with a microphone close to a material surface, and by time windowing of the response. The same absorption material is also measured in impedance tube and reverberation room. Fig. 5 (right) shows a case by absorption coefficient data, along with *in-situ* measurement using free-field and hard surface calibration references. It can be found that there is relatively much variance in data, and *in-situ* results look unreliable at low frequencies (negative absorption!).

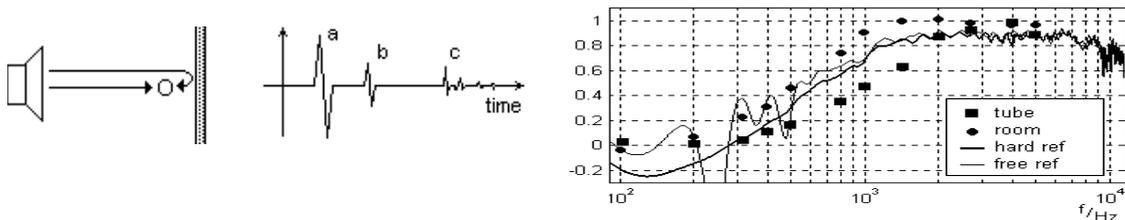


Figure 5: Left: *in-situ* system setup with (a) direct sound, (b) reflection (c) parasitic reflections to be removed. Right: data from different absorption coefficient measurements, including *in-situ* results with free-field and hard-wall references, for mineral wool (20 mm thick, 40 kg/m<sup>3</sup>).

The same material data was modeled by fitting a first-order pole-zero filter  $H_r(z)$  to the measured reflection magnitude response, which is constrained to value 1.0 (no loss) at zero frequency:

$$H_r(z) = \frac{1 + 1 - ab + abz - 1}{1 + az - 1}$$

Nonlinear optimization of parameters  $a$  and  $b$  was computed by Matlab function 'polyfit' to yield a least squares match. Figure 6 (left) shows the corresponding absorption coefficient, compared to raw data from free-field calibrated *in-situ* measurement, and a curve by model filter fitting that is obviously more useful at low frequencies than the raw data. For simple homogeneous absorbers this seems to work well, although there is no clear physical basis for such modeling.

Figure 6 (right) depicts results of more physical model fitting in a case where there is an air gap between 20 mm of the same absorbent as above and a hard wall. The flow resistance vs. acoustic impedance models by Delany-Bazley [12] and Mechel [13] were least-squares fitted to measured impedance tube data. As can be seen, generally the models are in fairly good agreement with measurements except at high frequencies.

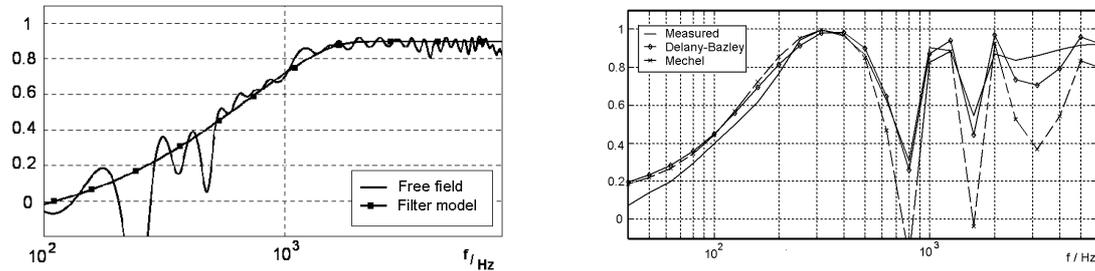


Figure 6: Left: low-pass reflection filter model fitted to in-situ measurement data (same as in Fig. 5) plotted by absorption coefficient as a function of frequency. Right: Absorption by the Delany-Bazley [12] and Mechel [13] impedance models based on flow resistance parameter, compared to impedance tube measurement data (same mineral wool as in Fig. 5 with 200 mm air gap).

## SUMMARY

Nonlinear optimization techniques have been applied in this study to parametric modeling of various acoustic systems. This is primarily based on finding discrete-time pole-zero transfer function models using ARMA methods, except in the last case, acoustic impedance modeling based on flow resistance parameter. Frequency-zooming ARMA modeling is shown to be a powerful technique in dealing with highly complex room responses and fairly complex musical instrument sounds. The estimated models can be applied also in real-time simulation and synthesis purposes, for example in artificial reverberation or spatialization/auralization systems for virtual acoustic reality.

## ACKNOWLEDGMENTS

Tuomas Paatero's work has been part of the Academy of Finland project "Technology for Audio and Speech Processing". Miikka Tikander's work has been part of the Tekes project "TAKU".

## REFERENCES

- [1] M.H. Hayes, *Statistical Digital Signal Processing and Modeling*. Prentice-Hall, 1996.
- [2] S.M. Kay, *Fundamentals of Statistical Signal Processing: Vol. I: Estimation Theory*. Prentice-Hall, 1993.
- [3] J.D. Markel and J.H.G. Gray, *Linear Prediction of Speech Signals*. Springer-Verlag, Berlin, 1976.
- [4] T.W. Parks and C.S. Burrus, *Digital Filter Design*. New York: John Wiley & Sons, 1987.
- [5] W.H. Kautz, "Transient Synthesis in the Time Domain," *IRE Trans. Circuit Theory*, vol. CT1, pp. 29-39, 1954.
- [6] P.W. Broome, "Discrete Orthonormal Sequences," *J. Assoc. Comp. Mach.*, vol. 12, (2), pp. 151-168, 1965.
- [7] T. Paatero and M. Karjalainen, "Kautz Filters and Generalized Frequency Resolution -- Theory and Audio Applications," Preprint 5378 *AES 110th Convention*, Amsterdam, 2001.
- [8] H. Brandenstein and R. Unbehauen, "Least-Squares Approximation of FIR by IIR Digital Filters," *IEEE Trans. Sig. Proc.*, vol. 46, (1), pp. 21-30, 1998.
- [9] Karjalainen et al., "AR/ARMA Analysis and Modeling of modes in resonant and reverberant systems," Preprint in *AES 110th Convention*, Munich, 2002.
- [10] Karjalainen M. and Järveläinen H., "More About this Reverberation Science: Perceptually Good Late Reverberation," Preprint 5415 *AES 110th Convention*, New York, 2001.
- [11] Karjalainen, M., and Tikander, M., "Reducing Artefacts of In-Situ Surface Impedance Measurements," in *Proceedings of the 17th Int. Congr. on Acoust. (ICA)*, vol. 2, p. 393, Rome, Italy, September 2-7, 2001.
- [12] Delaney, M.E. and Bazley, E.N., "Acoustical properties of fibrous absorbent materials", *Appl Acoust.* 3(1970), pp. 105-116
- [13] Mechel, F.P and Ver, I., "Sound absorbing materials and sound absorbers", in the book: Beranek L, Ver I (ed.), *Noise and Vibration Control Engineering*, John Wiley & Sons, New York, 1992.