Dynamic FM synthesis using a network of complex resonator filters

Julian Parker
Department of Signal Processing & Acoustics
Aalto University
Espoo, Finland
julian.parker@aalto.fi

Till Bovermann
Department of Media
Aalto University
Helsinki, Finland
till.bovermann@aalto.fi

ABSTRACT

There is a strong analogy between the sinusoidal operator used in FM synthesis, and the resonator filter. When implemented in a direct-form structure, a resonator filter is not suitable for use as a substitute for an FM operator, as it is not stable under centre frequency modulation. Recent, more robust resonator filter structures have made this use a possibility. In this paper we examine the properties of this structure that makes it appropriate for this application, and describe how a network of these filters can be combined to form a dynamic FM synthesis network. We discuss the possible range of sounds that can be produced by this structure, and describe its application to a performance system for improvised electroacoustic music.

1. INTRODUCTION

FM synthesis [1,2] and resonator filters [3–5] are both mature topics in digital audio signal processing. Resonators can in some ways be thought of as a generalisation of a sinusoidal oscillator that can take arbitrary input. Indeed, in the physical world, oscillators are often resonators which are driven in some way. Hence, we have a strong analogy between resonators and the sinusoidal operators at the heart of FM synthesis. This raises the question – what would an FM synthesiser-like structure constructed out of resonators and driven by an audio-signal sound like? Digital filter design has traditionally been dominated by direct-form topologies, which generally have poor time-varying properties. This deficiency seems to have discouraged any development of this idea. This situation is in contrast to the analog synthesis world, where audio-rate modulation of filters has been part of the standard repertoire of techniques since the beginning.

Previous work on the use of time-varying linear filters outside of audio signal processing exists [6–8], and has recently been applied in the analysis of the behaviour of Feedback-AM synthesis [9]. Feedback-AM synthesis can be considered to a be technique based on time-varying filters, and this analogy has been extended to second-order filters, including the direct-form resonator filter [10]. However, the poor time-varying stability of the direct-form filter means that this exploration has been limited to very low modulation depths.

In 2003, the late Max Mathews proposed a better behaved implementation of a resonator based on the idea of complex multiplication [11]. This structure is anecdotally reported to be completely stable under modulation of its parameters. This work has unfortunately seen little attention in the time since, although some analysis of coefficient interpolation schemes has been performed [12]. The resonator design proposed by Mathews is termed the 'phasor filter' or the 'complex resonator', the latter of which is the term used hereforth.

In this work, the idea of an FM synthesiser-like configuration of complex resonators is explored - with arbitrary audio input and the natural decay of the resonators taking the place of envelopes in defining the dynamic behaviour of the sound. In Section 2, we review the complex resonator structure, derive some useful properties of the structure, and examine the output it produces under audio-rate modulation of its centre frequency. In Section 3 we describe how a number of complex resonators can be combined into an FM synthesis network, and qualitatively examine the range of sounds which this structure can produce. In Section 4, we describe how this system has so far been applied in practice to produce musical performance systems. In Section 5, we conclude.

2. THE COMPLEX RESONATOR

The complex resonator is a system first introduced by Mathews and Smith [11]. It arises from the observation that multiplication of a complex number by a complex coefficient is equivalent to rotation around the origin on the complex plane. If we take a complex number $x = re^{i\theta}$ and multiply it by itself, the result is $x^2 = r^2e^{2i\theta}$. If we repeat the multiplication n times, we have $x^n = r^ne^{ni\theta}$. It should be clear that this process represents a continuous rotation around the origin. If $|x| < 1$, this motion is an in-going spiral. If $|x| > 1$, the motion is an out-going spiral. If $|x| = 1$, the motion is a circle around the origin. We can see that this circular motion is analogous to a resonance, with the angular velocity of the motion (defined by θ) being the frequency of the resonance. We can write this process as a pair of difference equations in terms of the real and imaginary parts of the product, and hence derive a system that looks very much like a digital filter:

$$x_{\text{Re}}[n+1] = r\cos(\theta)x_{\text{Re}}[n] - r\sin(\theta)x_{\text{Im}}[n],$$
$$x_{\text{Im}}[n+1] = r\sin(\theta)x_{\text{Re}}[n] + r\cos(\theta)x_{\text{Im}}[n] \quad (1)$$

If we add an input $u[n]$ to the real part, and take an output $y[n]$ from the imaginary part we can write the system in state space form:

$$\underline{x}[n+1] = \boldsymbol{A}\underline{x}[n] + \boldsymbol{B}u[n]$$
$$y[n] = \boldsymbol{C}\underline{x}[n] \quad (2)$$

where

$$\boldsymbol{A} = \begin{bmatrix} r\cos(\theta) & -r\sin(\theta) \\ r\sin(\theta) & r\cos(\theta) \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \boldsymbol{C} = \begin{bmatrix} 0 & 1 \end{bmatrix}$$

As we intend in this paper to use the system as a resonator and not as a pure oscillator, it makes sense to parameterise r in a more intuitive way. Instead, we would like to specify a decay time for the response of the filter. Intuitively from understanding the system as a repeated rotation, we can see that the reduction in amplitude at each sample step is given by multiplying by r. This clearly describes an exponential decay. Therefore, we can calculate a desirable value of r from a decay time τ, using the equation $r = e^{-\frac{1}{\tau f_s}}$, where f_s is the sampling frequency. It is also worth noting that we can trivially convert from unity sampling period angular frequency θ to a centre frequency with arbitrary sampling period by the relation $\theta = \frac{f_c}{2\pi f_s}$ where f_c is the centre frequency and f_s is the sampling frequency.

2.1 Normalization

In the form described above, the resonator structure possesses a large gain at its resonant peak. For more predictable use of the resonator in larger signal processing structures, particularly those involving feedback, it is desirable to normalize the filter so that its peak gain is unity. Also, since we are planning on modulating the filter's centre frequency at audio rate, any fluctuations in peak amplitude will introduce additional sidebands due to amplitude modulation. The normalisation should minimise this problem. First, the system is expressed in transfer function form:

$$H_{\text{res}}(z) = \frac{r\sin\theta z^{-2}}{1 - 2r\cos\theta z^{-1} + r^2 z^{-2}} \quad (3)$$

Assuming unity sampling period, and that the peak gain is at the specified centre frequency of the filter (which is correct apart from very close to DC or Nyquist, where the poles interfere with each other), we have:

$$H_{\text{res}}(e^{i\theta}) = \frac{r\sin(\theta)e^{-2i\theta}}{1 - 2r\cos(\theta)e^{-i\theta} + r^2 e^{-2i\theta}} = \frac{r\sin(\theta)}{(1-r)(e^{2i\theta} - r)} \quad (4)$$

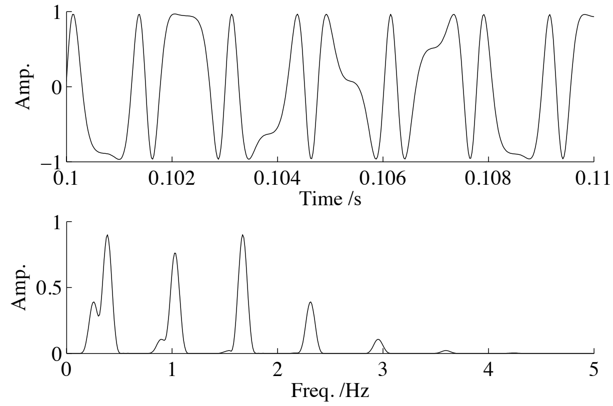

Figure 1. Extract from output of modulated complex resonator, excited with an impulse. Time-domain behaviour at a point near to the start of the decay (top), 4096 sample STFT taken at same point (bottom).

and therefore

$$|H_{\text{res}}(e^{i\theta})| = \left|\frac{r\sin(\theta)}{(1-r)(e^{2i\theta} - r)}\right| = \frac{\sqrt{r^2\sin^2(\theta)}}{(1-r)\sqrt{(\cos(2\theta) - r)^2 + \sin^2(2\theta)}} = \frac{r\sqrt{\frac{1}{2}(1-\cos(2\theta))}}{(1-r)\sqrt{1 + r^2 - 2r\cos(2\theta)}}. \quad (5)$$

This expression could be used to normalize the peak gain of the filter to unity, however it is rather complex and calculating it every time the coefficients are updated would be computationally expensive. With some further analysis, we can construct a simpler approximation to this expression. By observation, we can see that the maxima of the magnitude should occur at $\theta = \frac{\pi}{2}$. Taking a Taylor expansion around this point, we have:

$$|H_{\text{res}}(e^{i\theta})| \approx \frac{r}{1-r^2} - \frac{(1-r^2)}{2(1+r)^3}\left(\theta - \frac{\pi}{2}\right)^2 + O\left[\theta - \frac{\pi}{2}\right]^3 \quad (6)$$

Examining the first two terms of the series, we see that as $r \to 1$, $\left|\frac{r}{1-r^2}\right| \gg \left|\frac{(1-r^2)}{2(1+r)^3}\right|$. The same is true of the higher-order angle-dependent terms. For the purposes of this work, the resonator is generally used with τ on the order of 0.01 seconds and upwards. This corresponds to a resonance of $r = 0.9977$ or higher at a sampling rate of 44.1kHz. Therefore, for these purposes we can make the approximation

$$|H_{\text{res}}(e^{i\theta})| \approx \frac{r}{1-r^2} \quad (7)$$

which can be used to normalize the peak gain of the resonator. The peak gain will still fall to zero when the centre frequency is exactly at DC, but the normalization holds until very close to this point.

It is also possible to exactly normalise the filter with respect to varying centre frequency by using the method described by Smith et al. [3], and inserting two zeros at $z =$

$\pm\sqrt{r}$. This would be achieved by adding an un-attenuated path from the input to output of the system, and by inverting the input to the first state and dividing it by $\sin(\theta)$. However, this multiplication is problematic as it results in a divide-by-zero when the centre frequency is exactly at DC. Given that we may want to allow the centre frequency to take on negative values (to allow large modulation depths), this is unacceptable. This method also produces extremely large signal values at the input to the first state when the centre frequency is close to DC, which will cause numerical issues in fixed-point architectures.

2.2 Time-varying stability

As the filter will be modulated at audio rate, it is wise to examine its stability under time-varying conditions. Intuitively, it seems that it should be stable under arbitrary modulation of r and θ as long as $|r| < 1$, as the rotation that these parameters represent will always be following an in-going spiral. We can express this formally by applying the sufficient (but not necessary) condition for bounded-in bounded out (BIBO) stability of a time-varying filter described by Laroche [13] – which is that $||\boldsymbol{A}(n)||_2 = \sqrt{\lambda_{max}(\boldsymbol{A}^*\boldsymbol{A})} < 1,\ \forall n$ where * denotes the conjugate transpose and λ_{max} is a function which returns the maximum eigenvalue of its argument. Examining $\boldsymbol{A}$ in the case of the complex resonator, we have:

$$\begin{aligned} ||\boldsymbol{A}(n)||_2 &= \sqrt{\lambda_{max}(\boldsymbol{A}^*\boldsymbol{A})} \\ &= \sqrt{\lambda_{max}\begin{pmatrix} r^2\cos^2(\theta) & -r^2\sin^2(\theta) \\ -r^2\sin^2(\theta) & r^2\cos^2(\theta)\end{pmatrix}} \\ &= \sqrt{r^2} \end{aligned} \tag{8}$$

which gives us the condition that $|r(n)| < 1,\ \forall n$ which in this case is simply the normal time-invariant stability condition of keeping the eigenvalues of the state transition matrix within the unit circle.

2.3 Output of a frequency modulated complex resonator

Figure 1 shows a small extract from the signal produced by a resonator when the centre frequency of $f_c = 1028$ Hz is modulated with a 642 Hz sinusoid with a modulation depth of 998Hz. The decay time τ of the filter is 2 seconds. The filter is excited with an impulse. Note that we use absolute modulation depth to denote the amount of modulation, as the usual FM synthesis concept of modulation index is not meaningful in the general case where we do not know the content of the modulating signal.

The sound is like that of a simple struck bell. The response has an exponentially decaying envelope, as would be expected in the case of an unmodulated resonator. Using a different signal as input allows a variety of dynamic behaviours. As the bandwidth of the filters is very narrow, only very little of the input sound is recognisable. The input acts more as a way of controlling the amplitude and spectral balance of the output signal, with the sound output being dominated by distributions of sidebands of the carrier frequency (in this case centre frequency) consistent with those present in standard FM synthesis [1, 2].

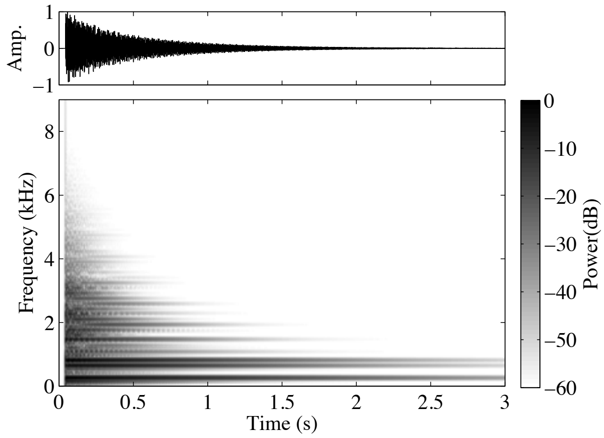

Figure 2. Waveform and spectrogram of cascaded four-resonator system, excited by an impulse.

3. RESONATOR FM NETWORKS

We define a resonator FM network as a vector $\underline{S}$ of resonator systems parameterised by a time-invariant (or slowly changing) centre frequency f_i, a time-varying frequency offset a_i and a decay time τ_i.

$$\underline{S} = \begin{pmatrix} S_1(f_1 + a_1, \tau_1) \\ S_2(f_2 + a_2, \tau_2) \\ \vdots \\ S_i(f_i + a_i, \tau_i) \end{pmatrix} \tag{9}$$

This vector of systems has a state space representation of its own, given by:

$$\begin{aligned} \underline{\chi}[n+1] &= \boldsymbol{\alpha}[n]\underline{\chi}[n] + \boldsymbol{\beta}u[n] \\ \underline{\upsilon}[n] &= \boldsymbol{c}\underline{\chi}[n] \end{aligned} \tag{10}$$

where

$$\underline{\chi} = \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \vdots \\ \underline{x}_i \end{pmatrix}, \underline{\upsilon} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \end{pmatrix}, \boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{A}_1 & 0 & \dots & 0 \\ 0 & \boldsymbol{A}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \boldsymbol{A}_i \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} b_1\boldsymbol{B} & 0 & \dots & 0 \\ 0 & b_2\boldsymbol{B} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_i\boldsymbol{B} \end{pmatrix}, \boldsymbol{c} = \begin{pmatrix} \boldsymbol{C} & 0 & \dots & 0 \\ 0 & \boldsymbol{C} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \boldsymbol{C} \end{pmatrix}. \tag{11}$$

The $\underline{x}_i$ are the individual state vectors of each filter. The $\boldsymbol{A}_i$ are the state update matrices. The y_i are the filter outputs. $\boldsymbol{B}$ and $\boldsymbol{C}$ are as defined in (2), and the b_i are the coefficients representing the gain of the input signal to the input of each filter. We also define an expression for the overall output of the parallel systems:

$$Y[n] = \underline{\kappa} \cdot \underline{\upsilon}[n] \tag{12}$$

Where $\underline{\kappa}$ is the time-invariant (or slowly varying) vector of output mixing coefficients κ_i.

The vector $\underline{a}$ of centre frequency offsets is calculated from the output of the resonators as follows:

$$\underline{a}[n+1] = \mathbf{\Gamma}\underline{\upsilon}[n] \tag{13}$$

We term $\mathbf{\Gamma}$ the FM feedback matrix. Since the outputs $\underline{\upsilon}$ are bounded by the maximum amplitude of the input signal (and likely much lower), the elements of $\mathbf{\Gamma}$ must be quite large – of the order of the depth of frequency modulation in Hz required. In fact, they are not absolute modulation depths but instead place an upper bound on the depth of modulation of a particular resonator by a particular output.

The behaviour of the system is governed by the vector of centre frequencies $\underline{f}$, decay times $\underline{\tau}$ and $\mathbf{\Gamma}$. Since each individual resonator system has unity peak gain and is completely stable under coefficient modulation, the overall system should also remain stable regardless of the values of $\mathbf{\Gamma}$. We can also make some general observations about the relationship between the output of the system and the content of $\mathbf{\Gamma}$. For example, $||\mathbf{\Gamma}||$ gives a measure of the overall depth of modulation, and hence the extent of the FM sidebands and the complexity of the timbre. Feedback loops are generated by entries along the diagonal of $\mathbf{\Gamma}$, and also by symmetrical patterns in the upper and lower triangular parts. When $\mathbf{\Gamma}$ contains only entries in the upper or lower triangular part of the matrix (with the diagonal empty), the system will be in a purely feedforward configuration.

Note that there is no connection between the outputs of the resonators and any of the inputs of the resonators. All of the resonators are excited only by the general input signal, albeit with different weightings. This is a specific choice, and is crucial to the use of the system as a musical instrument as it means that the overall envelope of the output of the system is predominantly a function of the input and the resonator ringing times. A user can therefore interact with the system in a relatively predictable way, in that it only produces sound when some kind of excitation is provided.

3.1 Results

Figure 2 shows the output of a system of four resonators configured in a simple feedforward configuration and excited by an impulse. The centre frequencies are distributed irregularly, and the modulation depths are around 1000Hz. The resulting sound is reminiscent of the idiophones used in Indonesian Gamelan music. Exciting the system with a more complex signal produces strong dynamic behaviour. Low amplitude inputs sound like brushing or blowing on a complex resonant object. Stronger inputs produce extremely dissonant and non-linear behaviour (although the system remains technically linear, just not LTI).

The dynamic behaviour of the sound is completely dependent on the nature of the input signal, and the natural exponentially decaying envelope of the resonators. This imparts a more organic quality to the sound than the precisely defined envelopes used in traditional FM synthesis. Standard FM synthesis strategies can be used when deciding on the topology of connections between resonators and their centre frequencies.

4. APPLICATIONS

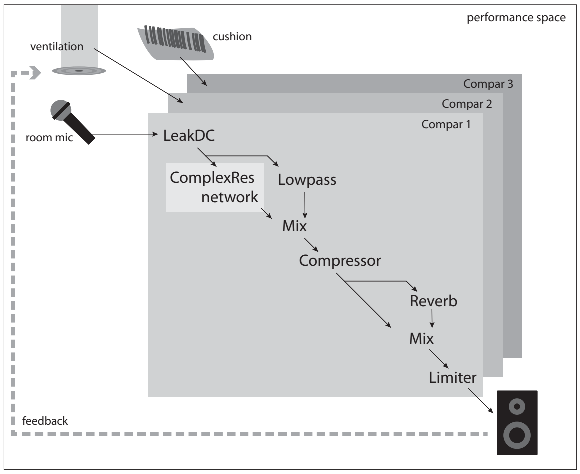

Figure 3. Overview of the performance setup.

The above described resonator network was implemented for practical use, and applied in a performance situation.

4.1 SuperCollider implementation

The complex resonator filter was implemented as a unit generator plugin (*UGen*) for SuperCollider [1] [14]. It has the interface

```
ComplexRes.ar(in, freq, decay)
```

where `in` is the source signal, `freq` the resonance frequency θ and `decay` the decay time τ in seconds.

Both SuperCollider code and sound examples for such networks can be found at the webpage accompanying this publication. [2]

4.2 Performance setup

The SuperCollider ComplexRes implementation was used in two consecutively developed setups: *Compar* is a feedforward resonator network featuring three `ComplexRes` nodes. Its successor *ComparFeedback* implements an FM feedback matrix. The number of nodes can be set when defining the synthesis engine. *ComparFeedback* implements a superset of the *Compar* system.

In both designs, the complex resonator network was embedded in a network of other processing structures. To remove unwanted DC offset, the input signal is processed by a high-pass filter (*LeakDC*). After a mixing stage in which the filter output is combined with the (low-passed) raw signal, it is fed through a compressor and processed by

[1] `http://supercollider.sourceforge.net/`
[2] `http://tai-studio.org/index.php/projects/complexres/`

a reverberation unit. A schematic overview of the setup is displayed in Figure 3.

Due to the increased complexity of the *ComparFeedback* unit, it became evident that an extended graphical user interface is needed compared to that for *Compar* (see Figure 5). However *Compar* features, despite its limitations compared to *ComparFeedback*, a unique way of performance which makes it a valuable instrument of its own.

Figure 4. Cushion-shaped musical interface made of conductive yarn.

Both *Compar* and *ComparFeedback* were used as core parts of a feedback-based performance setup, similar to that described by Di Scipio in its general form [15]. These performance systems were designed and implemented as part of the project *Electronic Music Practice for Neurodiverse People* [3]. As shown in Figure 3, three copies of the synthesis engine were used, each with a different input: *Compar 1* connected to a microphone placed in the performance place, *Compar 2* was wired to a contact microphone attached to a ventilation outlet, and *Compar 3* processed the input of a cushion-shaped musical interface made of conductive yarn which, when touched, renders a noisy electrical signal (see Figure 4). The latter served as a source of direct interaction with the sound.

The combination of all three elements created a drone-like soundscape, grounded in the acoustic features of the environment in which it was played. Particularly, the room-modes of the performance space and the resonating frequencies of the (already prominent) ventilation system had a large impact on the resulting sounds. The setup was inspired by works such as Tudor's *Rainforest IV* (1973) and Lucier's *Music on a Long Thin Wire* (1977).

Parameters such as input gain, filter frequencies, modulation depths, decays and reverberation were controlled by the artist during performance. Overall, the implemented systems reacted in a stable manner and were intuitive to play. Sonically, it created organic, FM-like sounds that were highly dependent on the sound colour of the input source: If noisy (e.g. in the case of the cushions), the resulting sounds were noisy, too; the filter network mostly altered the noise colour and added short tonal elements. If the input has a less noisy character the FM becomes much more prominent, adding distinctive sidebands to the output.

[3] http://tai-studio.org/index.php/projects/deind/

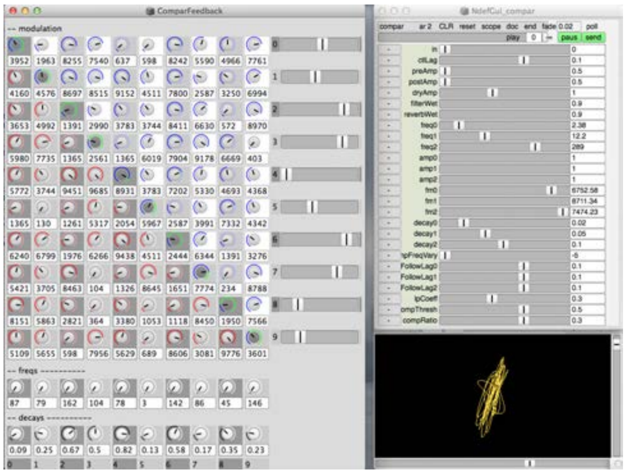

Figure 5. ComparFeedback (left) and Compar GUI (right)

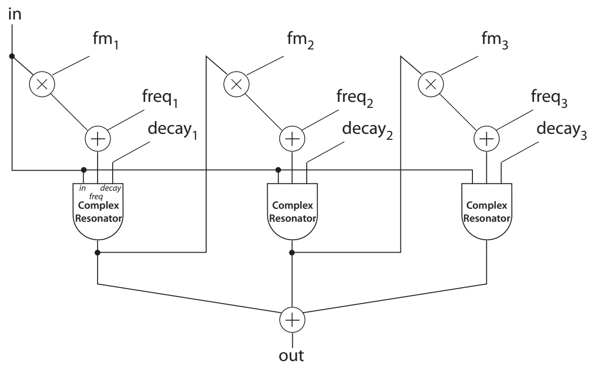

Figure 6. FM network in the Compar synthesis engine.

5. CONCLUSIONS

In this work the complex resonator filter has been examined, and results derived which are useful when applying it in a time-varying context. These results are an approximate normalisation of the filter to unity peak gain, and a derivation of the stability condition for the filter under parameter modulation. A new structure for dynamic FM synthesis was proposed, based on an arbitrary number of these filters configured to frequency modulate each other, and a formal description of this system given. The system is able to produce sounds within a very wide timbral space, and possesses a unique organic quality due to the natural exponential decay of the resonators and the use of audio input as excitation. The resonators were implemented as a SuperCollider UGen, and networks constructed from them within this environment. The resulting tools were applied within a performance system where the resonator network is excited by both the environmental sound of the space of the performance, and the signals generated by a cushion made of conductive thread.

Acknowledgments

This research was part of the DEIND project on designing electronic instruments for neurodiverse people, funded by the Aalto Media Factory of Aalto University, Helsinki.

6. REFERENCES

[1] J. M. Chowning, "The synthesis of complex audio spectra by means of frequency modulation," *J. Audio Eng. Soc.*, vol. 21, no. 7, pp. 526–534, 1973.

[2] M. Le Brun, "A derivation of the spectrum of fm with a complex modulating wave," *Computer Music J.*, vol. 1, no. 4, pp. 51–52, 1977.

[3] J. O. Smith and J. B. Angell, "A constant-gain digital resonator tuned by a single coefficient," *Computer Music J.*, vol. 6, no. 4, pp. 36–40, 1982.

[4] K. Steiglitz, "A note on constant-gain digital resonators," *Computer Music J.*, vol. 18, no. 4, pp. 8–10, 1994.

[5] J. Dattorro, "Effect design, part 1: Reverberator and other filters," *J. Audio Eng. Soc.*, vol. 45, no. 9, pp. 660–684, 1997.

[6] S. C. Scoular, I. B. Rogozkin, and M. S. Cherniakov, "Review of soviet research on linear time-variant discrete systems," *Signal Processing*, vol. 30, no. 1, pp. 85–101, 1993.

[7] M. Cherniakov, V. Sizov, and L. Donskoi, "Synthesis of a periodically time-varying digital filter," *IEE Proc. Vision, Image and Sig. Processing*, vol. 147, no. 5, pp. 393–399, 2000.

[8] M. Cherniakov, *An introduction to parametric digital filters and oscillators.* Wiley, 2003.

[9] J. Kleimola, V. Lazzarini, V. Välimäki, and J. Timoney, "Feedback amplitude modulation synthesis," *EURASIP J. Advances Sig. Processing*, vol. 2011, Dec. 2011.

[10] V. Lazzarini, J. Kleimola, and V. Välimäki, "Aspects of second-order feedback AM synthesis," in *Proc. of the Int. Computer Music Conf. (ICMC)*, Huddersfield, 2011.

[11] M. Mathews and J. O. Smith, "Methods for synthesizing very high Q parametrically well behaved two pole filters," in *Proceedings of the Stockholm Musical Acoustics Conference (SMAC 2003)*, Royal Swedish Academy of Music, August 2003.

[12] D. Massie, "Coefficient interpolation for the Max Mathews phasor filter," in *133rd Convention of the Audio Eng. Soc.*, 2012.

[13] J. Laroche, "On the stability of time-varying recursive filters," *J. Audio Eng. Soc.*, vol. 55, no. 6, pp. 460–471, 2007.

[14] J. McCartney, "Rethinking the computer music language: SuperCollider," *Computer Music J.*, vol. 26, no. 4, pp. 61–68, 2002.

[15] A. Di Scipio, "Listening to yourself through the otherself: On background noise study and other works," *Organised Sound*, vol. 16, no. 02, pp. 97–108, 2011.