Audio Engineering Society

Convention Paper 10553

Presented at the 152nd Convention
2022 May, In-Person and Online

This paper was peer-reviewed as a complete manuscript for presentation at this convention. This paper is available in the AES E-Library (http://www.aes.org/e-lib) all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Twang! A physically derived synthesis model for the sound of a vibrating bar

Rod Selfridge[1], Mimmi Andreasson[2], Lina Bengtsson[2], Bjarki Vidar Kristjánsson[2], Emelie Lindborg[2], Matilda Rydén[2], Hazar Emre Tez[3], and Joshua D Reiss[3]

[1]*Edinburgh Napier University*
[2]*KTH Royal Institute of Technology*
[3]*Queen Mary University of London*

Correspondence should be addressed to Rod Selfridge (r.selfridge@napier.ac.uk)

ABSTRACT

A physically derived synthesis model of the sound generated when a ruler is twanged while hanging over the edge of a solid surface is presented. This is a sound effect used in movies, TV, theatre performances and cartoons. The model is derived from the Euler-Bernoulli equation, offering the user a set of physical parameters to control ruler length as well as the material properties. Perceptual evaluation indicates that the model can be perceived as realistic as a recorded ruler twang as well as being able to replicate sounds of similar quality as an alternative synthesis model.

1 Introduction

The sound of a ruler twang has been present in school classrooms as long as rulers have been available. They are used as a classic sound effect employed in movies such as *Raiders of the Lost Arc* [1] and cartoons like *The Magic Roundabout* [2]. This simple sound has a variety of comic inferences but is generated from the combination of two mechanical principles. Figure 1 illustrates a ruler just before the force of being bent over a desk is released to generate a twang sound.

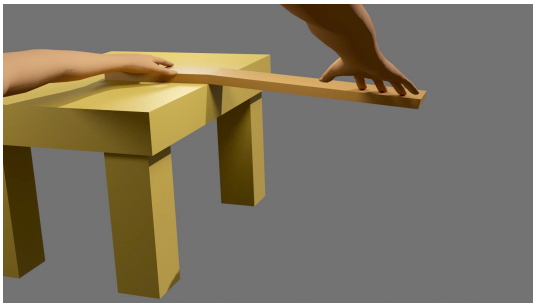

Fig. 1: A ruler being 'twanged' to create the sound of vibrating bar.

This paper examines the mechanical interactions that generate the sound created when a ruler is hung over the edge of a desk, the force of a finger pushes it down and suddenly released causing the ruler to oscillate.

2 Background

The sound generated by the twang of a ruler is based on the acoustics of a vibrating bar. The synthesis of bar instruments was presented in [3] using a banded waveguide technique. Waveguide synthesis models the physics by creating a delay line that propagates the excitation signal, replicating the travelling wave in the object. Banded waveguides are where a separate delay line is implemented for each harmonic.

A boing / twang sound effect was given in [2] using an physically inspired algorithm that the proposed model in this paper is influenced by. Farnell's model uses a mixture of additive synthesis and noise shaping techniques to synthesis both the free bar and clamped bar vibrations.

Physically derived sound synthesis models have been investigated in [4, 5]. Here, equations that were derived in physics are applied to procedural synthesis models for sound effects. The benefit of this approach is that sounds are generated in real-time, and the parameters are highly relevant with respect to how the sounds are generated in nature.

The design of this synthesis model draws from modal synthesis by using values for the harmonics, mathematically predicted from the physics of a vibrating bar. A musical instrument that generates sounds through these principles is the Kalimba or African thumb piano [6]. This differs from our model as it has a bridge support rather than the hard edge of a desk.

This model expands the algorithm in [2] by increasing the number of partials being synthesised, adding in a frequency damping component, an additional slide function and a method of controlling the gain relationship of the partials. User are also provided with a number of practical physical parameters to adjust, enabling them to tailor the sound effect to their requirements.

3 Physics of a Ruler Twang

The ruler twang effect is generated by a combination of different vibration modes of a bar; vibration of a cantilever bar and vibration of a free bar.

3.1 Free Bar

A vibrating free bar is one which vibrates with both of its ends free to move. Figure 2 shows a free bar vibrating at its fundamental frequency.

The Euler-Bernoulli beam equation, derived from Newton's second law of motion [7] governs the movement of bars, shown in Eqn. 1.

$$\frac{\partial^2 y(x,t)}{\partial t^2} + \frac{EK^2}{\rho}\frac{\partial^4 y(x,t)}{\partial x^4} = 0 \tag{1}$$

where y is the spatial distance in the vertical axis, x is the distance along the length of the bar, t is time, E is the Young's Modulus of the material of the bar, and ρ is the mass density of the bar. The value K is the radius of gyration, for a rectangular cross section of height h, K is given as:

$$K = \frac{h}{\sqrt{12}} \tag{2}$$

By setting:

$$y(x,t) = f(x)g(t) \tag{3}$$

the solution of Eqn. 1 can be written as a standing wave. It can thereafter be separated into spatial and time components giving:

$$f(x)\frac{\partial g^2(t)}{\partial t^2} + \frac{EK^2}{\rho}\frac{\partial f^4(x)}{\partial x^4}g(t) = 0 \tag{4}$$

The time component is given by:

$$\frac{\partial g^2(t)}{\partial t^2} = -\omega^2 g(t) \tag{5}$$

where ω is the angular frequency. This results in $g(t)$ being a multiple of $\sin(\omega t + \phi)$. The spatial component is given by:

$$f(x) = A\sin(\kappa x) + B\cos(\kappa x) + C\sinh(\kappa x) + D\cosh(\kappa x) \tag{6}$$

where:

(a) Maximum point of vibration **(b)** Neutral point of vibration **(c)** Lowest point of vibration

Fig. 2: Plot of a free bar vibrating at its fundamental frequency

$$\kappa^4 = \frac{\omega^2 \rho}{EK^2} \tag{7}$$

The boundary conditions for a bar free at both ends are $x = 0$, $x = l$, $\partial^2 y/\partial x^2 = 0$, $\partial^3 y/\partial x^3 = 0$. This enables us to solve Eqn. 6, shown in [7] giving:

$$\cosh(\kappa l)\cos(\kappa l) = 1 \tag{8}$$

where l is the length of the bar. Setting $\lambda_n = \kappa_n l$, the angular frequency of each node n (ω_n) is given as:

$$\omega_n = \sqrt{\frac{EK^2}{\rho}}\frac{\lambda_n^2}{l^2} \tag{9}$$

Numerical solutions are used to calculate the values of λ_n, given in [7] and reproduced in Table 1. By using Eqn. 9, we can calculate the frequencies of the harmonics using the ratio λ_n^2/λ_1^2. These are shown in Table 1.

3.2 Cantilever Bar

A cantilever bar is one which is free at one end and clamped at the other. Equations 1 to 7 are identical for a cantilever bar, only the boundary conditions are different due to it being clamped at one end. Figure 3 shows a cantilever bar vibrating at its fundamental frequency.

Assume that the end of the bar that is clamped is at $x = 0$ and the free end is at $x = l$. The boundary conditions at $x = 0$ are $y = 0$, $\partial y/\partial x = 0$. At $x = l$ we have $\partial^2 y/\partial x^2 = 0$ and $\partial^3 y/\partial x^3 = 0$.

The boundary conditions give:

$$\cosh(\kappa l)\cos(\kappa l) = -1 \tag{10}$$

The angular frequencies are also given by Eqn. 9 [7]. The values of λ_n for a cantilever bar are given in [7] and reproduced in Table 1. Similar to the free bar, the ratios of the harmonics frequencies can be calculated using λ_n^2/λ_1^2. These are shown in Table 1.

3.3 Damping

In [8] a solution of Eqn. 1 for a vibrating free bar is given as:

$$y(x,t) = \sum_n f_n(x) e^{i\omega_n t} e^{-\alpha_n t} \tag{11}$$

where $e^{-\alpha_n t}$ is a damping factor. The value α_n is given by a quality factor Q_n which is dependent on the wood species and given by [9]:

$$Q_n = \frac{\pi f_n}{\alpha_n} \tag{12}$$

3.4 Interaction

The method of interaction between the two modes of vibration are defined in [2] and what gives the ruler twang its distinctive sound. A brief description is given here as well as additional interaction added in this model.

Initially a ruler is held with approximately half the length hanging over the edge of a desk. One finger is used to press one end onto the desk while another is used to bend the ruler down with a certain amount of force. Once released, the ruler rises up into a cantilever bar as shown in Fig. 4a.

After the ruler reaches the top of its oscillation it starts its return journey down, striking the desk when it passes the horizontal point. The amount of force pressing the ruler on the desk and its position will determine how much of the ruler bends upwards on the desk as the right-hand side of the ruler is oscillating down below the desk level.

A second cantilever is formed by the edge hanging over the desk. This is exaggerated in Fig. 4b to highlight the phenomenon. The ruler then returns on its upward journey. Once the ruler passes the horizontal and separates from the desk it also acts as a vibrating free bar struck in the middle. This is shown in Fig. 4c.

Table 1: Values of λ and harmonic ratios for cantilever and free mode bar vibrations

	Free Bar		Cantilever Bar	
n	λ_n	Harmonic ratio λ_n^2/λ_1^2	λ_n	Harmonic ratio λ_n^2/λ_1^2
1	4.7300	1	1.8751	1
2	7.8532	2.7565	4.6941	6.2669
3	10.9956	5.4039	7.8548	17.5475
4	14.1372	8.9330	10.9955	34.3861
5	17.2788	13.3443	14.1372	56.8426
6	20.4204	18.6379	17.2788	84.9130

(a) maximum point of vibration **(b)** Neutral point of vibration **(c)** Lowest point of vibration

Fig. 3: Plot of a cantilever bar vibrating at its fundamental frequency

4 Implementation

The synthesis model forms part of a series within a browser-based library. It was implemented in JavaScript, making use of the Web Audio API to provide client-side, procedural synthesis [10]. Parameters of vibrato, shape and decay were made available for users to adjust, creating unique sounds in real-time. An image of the interface is given in Fig. 5[1].

The properties of the ruler and its material are also available to the user, including Young's Modulus, density, length and height of the ruler. The DSP design was inspired by [2]. It is shown in Fig. 6 with a description of each property below.

4.1 Synthesis of vibration modes of a bar free at both ends

The fundamental frequency of the clamped mode relating to the length of the whole bar was calculated using Eqn. 9. The values of E, l, ρ and height h are set by the user. The radius of gyration K is calculated using Eqn. 2 and applied to Eqn. 9.

[1]The model can be heard and interacted with at https://nemisindo.com/models/twang.html

For free vibration fundamental frequency, the value of $\lambda_1 = 4.7300$ was used, as given in Table 1. The modes of the free vibration are calculated from the ratios given in Table 1. The fundamental and five harmonics were used for both the clamped and free vibrating part of the synthesis.

The free vibration modes are synthesised using a white noise source and subtractive synthesis. Prior to shaping the noise source is modulated by the clamped vibration modes (see section 4.2) and passed through a single pole low pass filter. The free bar vibration signal is then passed through a series of forth order bandpass filters whose centre frequencies are set at the fundamental frequency along with the 5 harmonics. Each frequency component was then multiplied by a user set gain as per Section 4.3, and decayed at the frequency dependent rate as per section 4.4.

4.2 Synthesis of vibration modes of a bar clamped at one end

The clamped vibration modes were also calculated from Eqn. 9, with the same parameter values as the Free bar vibration modes. For the fundamental clamped mode, $\lambda_1 = 1.8751$ was used, as given in Table 1. This allows the calculation of the next 5 harmonics based on the ratios given in Table 1. Vibrations generated

(a) Vibrating bar at maximum point - full-length **(b)** Vibrating bar at lowest point - part-length cantilever **(c)** Mid-point vibrating bar illustrating free bar vibration

Fig. 4: Plot of a vibrating bar hanging over an edge.

Fig. 5: Interface for the twang sound effect

from the clamped vibration modes were synthesised using phasors ramping at the appropriate frequencies, with their phases aligned when triggered. These then were the input to cosine operators giving an equivalent output to an oscillator.

The output from the phasor ramping at the fundamental frequency of the clamped bar vibration is used to shape the time envelope of the noise source used for the synthesis of the free bar modes (Section 4.1). This is squared first to give it a sweeping rise. Perceptually this can be viewed as having the free bar vibrations while the ruler is at moving above the horizontal plane of the solid surface, which is damped when the ruler moves below the plane of the table.

In addition to the clamped vibrations of the full length of the ruler, the same process was carried out to calculate the frequencies for the clamped vibration modes relating to the bar clamped at the edge of the desk. The length of the shorted clamped bar was set at 0.7 of the full length. This is added to the full-length clamped vibration modes with a gain value perceptually set at 0.75 of that of the full-length frequency dependent gain.

4.3 Modal Gains

In his model of a vibrating ruler, Farnell [2] sets the gain of each mode to be half of the previous as an approximation. In [9] the gain of each mode is calculated to match a recording.

In order to provide greater flexibility in this model, we decided to allow the user to shape the gains of each frequency. The harmonic gains were set to values based on a beta probability distribution model, given in [11] as:

$$p(x) = \frac{\Gamma(l+m)}{\Gamma(l)\Gamma(m)} x^{(l-1)}(1-x)^{(m-1)} \qquad (13)$$

where $0 < x < 1$, and $l, m > 0$ are shape parameters. Γ is the Gamma Function given by:

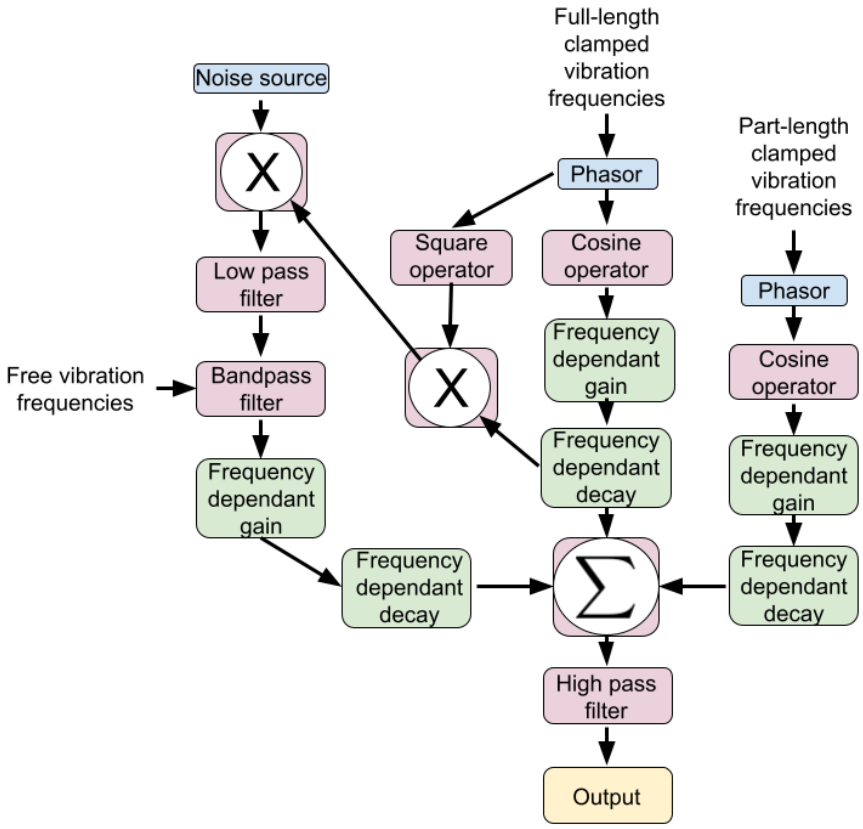

Fig. 6: Signal processing chain used in the creation of synthesis model

$$\Gamma(l) = (l-1)! \tag{14}$$

In this implementation $0.5 < l < 10$ is a parameter available to the user. $m = 4$. The output x is linearly mapped to $0 - f_s/2$. The value of $p(x)_{max}$ was identified and the output normalised,

$$p(x) = \frac{p(x)}{p(x)_{max}} \tag{15}$$

This gave the harmonic gain values as shown in Fig. 7.

4.4 Damping

As outlined in section 3.3, the note decay for a vibrating free bar depends on each frequency and a quality factor. In [8] (Figure 3 in the article), a range of Q values are indicated on a graph. These range from 50 to over 250. In this implementation we gave the user control over the Q value, labelled as 'decay'. The user is able to adjust in real-time the value from 50 up to 800, taking our model beyond what was tested in [8]. These values were also applied to the modes of the cantilever vibration.

In [12] it is shown that the natural oscillation requires an adjustment due to damping:

$$\omega_D = \omega_n \sqrt{(1-\varepsilon^2)} \tag{16}$$

where ω_D is the damped angular frequency, ω_N the natural angular frequency and ε is a damping factor. For Q values between 50 and 800, the values of ε vary from 1.25 x 10^{-3} to 62.50 x 10^{-3}. From Eqn. 16 we can see that $\omega_D \approx \omega_N$, hence this effect is negligible and not implemented in this model.

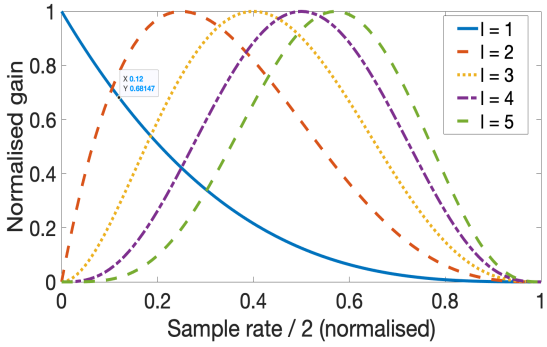

Fig. 7: Normalised gain of the frequency output of synthesis model

4.5 Vibrato

A vibrato option was provided in this implementation. This was set to a depth of 18Hz with the frequency set by the user between 0Hz and 8Hz.

4.6 Slide

For added realism a slide option was supplied to the user. This simulates the ruler being shortened once it has started oscillating by ramping the frequency of the part-length clamped vibrations up by 35%.

5 Perceptual Evaluation

A multistimulus perceptual evaluation, based on MUSHRA [13], was carried out on the sound effect model. It was compared against recordings of a real ruler twang, Farnell's model [2] and an anchor model. Recordings of the real ruler twangs were captured by the authors using a zoom h4n recorder, and Rode shotgun microphone. A similar reverb that corresponds to the space where the real recordings were captured, was added to the synthesised sound effects to minimise spatial differences. An anchor sound effect was created by changing the phasor~ object in Farnell's model to a cos~ to create a perceptually poor model , (see [2] for full model details). This was simply chosen as it maintained the characteristics of a twang sound effect but perceptually was vastly poorer.

Further to the multistimulus test, an A/B pairwise test was carried out to examine if participants were able to identify the real recordings and synthesised ones [14]. Both the Pure Data model of Farnell [2] and this model were compared to the real recordings.

In total, the tests were performed by 21 participants. The distribution between the genders was 7 females and 14 males. Ages ranged from 18 to 59 years with 85% of the participants in ages between 21 and 29 years. 47% of participants had carried out listening tests previously.

The listening test was hosted on PsyToolkit [15]. For the MUSHRA evaluation the participants were presented with a slider ranging from 0 -"Not realistic at all", to 9.99 - "Very Realistic". Two pages of 4 sound clips (obtained from real recordings, Farnell's model, proposed model and an anchor) were presented in a random order and participants invited to rate them. They can also comment on each sound if they choose.

Following this, participants were presented with six pairwise tests. Three of the comparisons were with this model and 3 with Farnell's [2] (again presented in random order). Each of the synthesised sound effects was created to be perceptually close to the recorded sound effects by the authors. After choosing which of the sounds was the real sound, participants were asked to indicate how difficult it was to identify the real recording on a 5-point Likert scale.

6 Results and Analysis

Figure 8 shows the boxplots for the different twang models and the real recordings. As expected, it can clearly be seen that the lowest rated twang sounds were that of both the Anchors with mean rating values of Anchor_1 = 0.61 and Anchor_2 = 0.51. The highest rated sounds were those of the real recordings with Real_1 = 8.10 and Real_2 = 7.07.

Results for sounds generated by Farnell's model [2] show that the sound effect Farnell_1 had a low rating (mean rating of 3.04). The other Pure Data model, Farnell_2, had a mean value of 4.56 and it can be seen in Fig. 8 that there is a greater range in opinion in relation to this sound effect. The sound effect Proposed_1 was given very similar ratings to that of Farnell_2, with a mean of 4.12. The second sound effect from the proposed model has a higher mean of 6.02 and a narrower range of ratings, as seen in Fig. 8.

To examine the different statistical significance's between the different model ratings, a pairwise

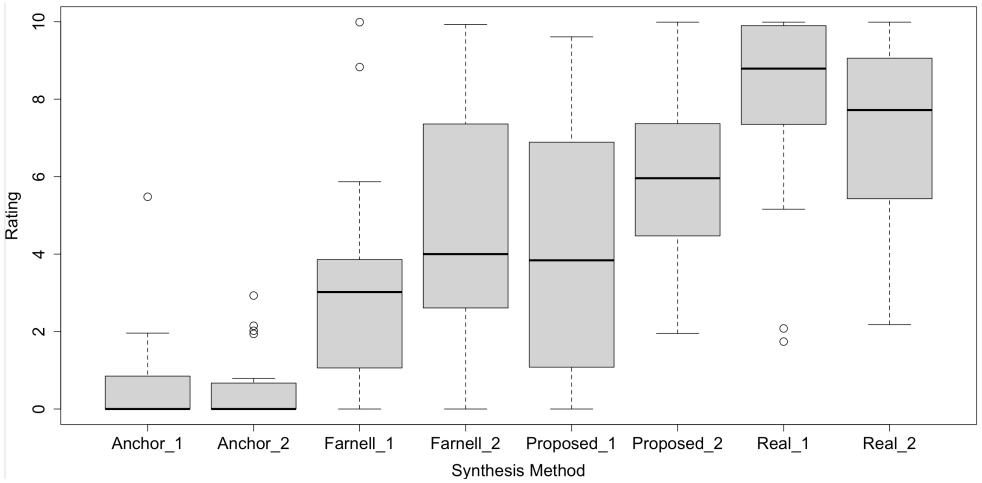

Fig. 8: Boxplot of realistic ratings from participants.

Wilcoxon's rank-sum test of each sound effect rating was carried out and P-values obtained [16]. The results of this test are shown in Table 2. We can see from Table 2 that one of the sound effects generated using this model (Proposed_2) has significant difference to one of the Pure Data models (Farnell_1). While the other sound effect generated using this model (Proposed_1) has no significant difference in ratings to both the sounds obtained from Farnell's model. This would indicate that the proposed model presented in this paper has potential to be able to replicate the sounds generated using Farnell's model in terms of realism but can also perform better.

Further examination of Table 2 shows that there was no significant difference in realism rating for one of the sounds generated using this model (Proposed_2), and one of the real recordings of a ruler twang (Real_2). This indicates that the model presented here is able to replicate a sound viewed as realistic as a real ruler twang.

Comments from participants during this test in relation to the proposed model included *"Sounds like a real sound, but not like a ruler on a table but more like a guitar string", "Sounds fully realistic" and "Kind of sounds like a twang but does sound like it synthesised"*.

The difference between the real ruler twangs sound effects and the synthesised ones was more apparent following the pairwise A / B test. When asked to identify the real ruler twangs compared to those generated using Farnell's model [2], participants were able to identify the real recordings 89% of the time. Similarly, when asked to identify the real recordings when compared to the proposed model, they were able to identify the real recordings 90% of the time.

Participants were asked to rate how easy it was to identify the recorded sound effects compared to the synthesised ones on a 5-point Likert scale. The median value for both the real recordings compared to those generated from the Pure Data as well as those created using this model is 4. This relates to "Easy" on the Likert scale and indicates that although the proposed model can produce sounds as realistic as recorded ones it is still easy for participants to identify which sound had been synthesised.

Comments when describing the reasons for identifying the real sound when compared to the one created using this model included *"The ring on the synth was to long", "the fake one "vibrated" too long" and "Both realistic, shorter the safer answer"*. It appeared that the length of the synthesised sound was a influential component for identifying the recorded sound effect. Other comments included *"the second sound (proposed model) sounds digitized" and "The second one (proposed model) had a more realistic difference in the start and the ringing out"*.

7 Discussion

It is clear from the results that the synthesis model presented can produce sounds that are realistic but can gen-

Table 2: Table showing P-values indicating significance differences between the different Twang models and real recordings. (ns = P>0.05, * = P<0.05, ** = P<0.01, *** = P<0.001, **** = P<0.0001).

	Anchor_1	**Anchor_2**	**Farnell_1**	**Farnell_2**	**Proposed_1**	**Proposed_2**	**Real_1**
Anchor_2	ns						
Farnell_1	***	***					
Farnell_2	****	****	ns				
Proposed_1	***	***	ns	ns			
Proposed_2	****	****	***	ns	ns		
Real_1	****	****	****	***	***	**	
Real_2	****	****	***	*	**	ns	ns

erally be identified when directly compared to recorded sound effects. It is also seen that Farnell's model does not produce sounds as realistic as real recordings, implying that the proposed model has progressed the state of the art but improvements are still required to create sounds as realistic as recorded ones.

One of the major advantages of the proposed model is the exposure of the parameters which directly relate to the equations governing the twang process. This allows the user to change the material of the ruler as well as its dimensions. The Q values relating to the decay also gives the user a flexibility not offered by the other synthesis model.

Allowing the user to adjust the gain relationship between the harmonics is another dimension which would create more unrealistic or cartoon sound effects. This may contribute to added difficulty when trying to create specific realistic sounds as it is know that the harmonic relationships have a strong influence over the timbre of the sounds.

7.1 Future Developments

One development of the model which could improve its flexibility and usability would be to map x of the gamma function (Eqn. 13) to log of $\frac{f_s}{2}$. This would change the relationship of the harmonics to one closer to human hearing for better perceptual control.

This model does not investigate the material of the desk which, like a mallet for a bar-based musical instrument, has an influence on the overall sound. For example, a xylophone struck by a hard wooden mallet will sound different when struck by a softer rubber one. Adding the effects of the material of the desk will expand the model.

The addition of a bridge to this model could be investigated. The model could then be developed to synthesis the sound of the Kalimba or African thumb piano.

8 Conclusion

The synthesis model presented in this paper was inspired by the work presented in [2] but exposes considerably more physical attributes of the mechanical system that produces the sound effect to the user. The ability to manipulate these physical parameters has been shown to allow this model to produce a broader range of sounds in relation to realism, ranging from ones that lack realism to sounds that were perceived equivalent to real recordings.

References

[1] Holman, T., *Sound for film and television*, Taylor & Francis, 2012.

[2] Farnell, A., *Designing sound*, MIT Press Cambridge, 2010.

[3] Essl, G., Serafin, S., Cook, P. R., and Smith, J. O., "Musical applications of banded waveguides," *Computer Music Journal*, 28(1), pp. 51–63, 2004.

[4] Selfridge, R., Moffat, D., and Reiss, J., "Sound synthesis of objects swinging through air using physical models," *Applied Sciences*, 7(11), p. 1177, 2017.

[5] Selfridge, R., Moffat, D., Avital, E. J., and Reiss, J. D., "Creating real-time aeroacoustic sound effects using physically informed models," *Journal of the Audio Engineering Society*, 2018.

[6] Chapman, D. M., "The tones of the kalimba (African thumb piano)," *The Journal of the Acoustical Society of America*, 131(1), pp. 945–950, 2012.

[7] Benson, D., *Music: A mathematical offering*, Cambridge University Press, 2006.

[8] Aramaki, M., Baillères, H., Brancheriau, L., Kronland-Martinet, R., and Ystad, S., "Sound quality assessment of wood for xylophone bars," *The Journal of the Acoustical Society of America*, 121(4), pp. 2407–2420, 2007.

[9] Aramaki, M., Bailleres, H., Brancheriau, L., Kronland-Martinet, R., and Ystad, S., "Relationship between sound classification of xylophone-like bars and wood species properties," in *Proceedings of the 13th International Congress on Sound and Vibration*, 2006.

[10] Reiss, J. D., Tez, H. T., and Selfridge, R., "A comparative perceptual evaluation of thunder synthesis techniques," in *Audio Engineering Society International Conference on Audio for Virtual and Augmented Reality*, 2020.

[11] Johnson, N. L., Kotz, S., and Balakrishnan, N., *Continuous univariate distributions, volume 2*, volume 289, John wiley & sons, 1995.

[12] Fahy, F. and Thompson, D., *Fundamentals of sound and vibration*, CRC Press, 2015.

[13] Series, B., "Method for the subjective assessment of intermediate quality level of audio systems," *International Telecommunication Union Radiocommunication Assembly*, 2014.

[14] De Man, B. and Reiss, J. D., "A pairwise and multiple stimuli approach to perceptual evaluation of microphone types," in *Audio Engineering Society Convention 134*, 2013.

[15] Stoet, G., "PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments," *Teaching of Psychology*, 44(1), pp. 24–31, 2017.

[16] Field, A., Miles, J., and Field, Z., *Discovering statistics using R*, Sage publications, 2012.