MODELLING THE DECAY OF PIANO SOUNDS

Tian Cheng, Simon Dixon, Matthias Mauch

Centre for Digital Music, Queen Mary University of London, London, United Kingdom

ABSTRACT

We investigate piano acoustics and compare the theoretical temporal decay of individual partials to recordings of real-world piano notes from the RWC Music Database. We first describe the theory behind *double decay* and *beats*, known phenomena caused by the interaction between strings and soundboard. Then we fit the decay of the first 30 partials to a standard linear model and two physically-motivated non-linear models that take into account the coupling of strings and soundboard. We show that the use of non-linear models provides a better fit to the data. We use these estimated decay rates to parameterise the characteristic decay response (decay rates along frequencies) of the piano under investigation. The results also show that dynamics have no significant effect on the decay rate.

Index Terms— double decay and beats, non-linear fitting, piano decay response

1. INTRODUCTION

The energy of a piano note decays after the hammer strikes the string. The decreasing energy causes several problems for piano analysis. One example is the difficulty of determining offset times of notes in real-world performances. We would expect that modelling decay will help to improve performance of piano analysis applications, such as onset-offset detection. The physics of coupled piano strings and factors influencing piano decay have previously been studied [1, 2]. Välimäki et al. first propose to estimate the decay rate of harmonics by linear regression [3], and the non-linear decay caused by coupled vibrations has been tackled in [4, 5, 6, 7]. Ewert and Müller propose a model for estimating note intensities of piano pieces [8]. In this paper, we apply knowledge of piano physical modelling on real-world piano recordings to model the decay of piano sounds.

We track the decay of piano notes from the RWC Music Database to explore how the energy of piano tone partials decays for different frequencies and dynamics. Frequencies of partials for each note are estimated jointly in an NMF framework taking inharmonicity into account [9]. There are several decay patterns of partials due to the coupling between bridge and soundboard and the coupling within the multiple strings of piano notes. Based on the theory of piano string coupling [1], the decay of each partial is classified into three types: linear decay (of log-energy), double decay and curve decay (beats). These decay types are fitted by a linear model, a multi-phase linear model (which is non-linear, despite its name) and a non-linear curve fitting model, respectively. By using the non-linear models, we obtain a better fit to the data. The results show various decay rates along the frequency range with a trend that partials decay faster at higher frequencies. For different dynamics, the estimated decay rates are similar to each other.

Tian Cheng is supported by a China Scholarship Council/ Queen Mary Joint PhD Scholarship. Matthias Mauch is funded by a Royal Academy of Engineering Research Fellowship.

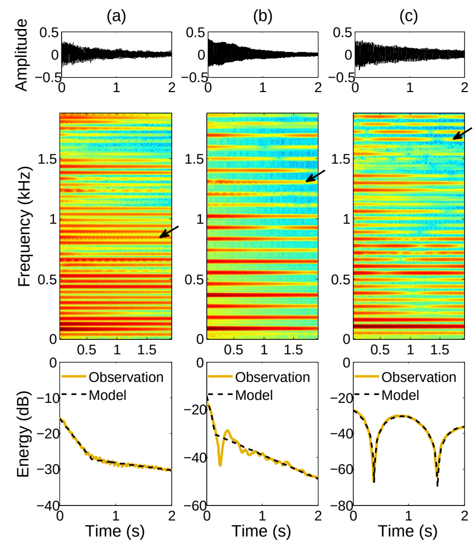

Fig. 1. Different decay patterns of partials from notes (a) F1 (43.7Hz), (b) G♭2 (92.5Hz) and (c) A1 (55Hz). The top and middle panes show the waveforms and spectrograms, respectively. The bottom panes show the decay of selected partials, which are indicated by the arrows on the spectrograms.

In Section 2, a brief introduction to piano acoustics is presented. The methods for finding and modelling decays of partials are introduced in Section 3. The experimental setup and results are described in Sections 4. Section 5 concludes the paper.

2. BACKGROUND

When a key of a grand piano is pressed, the hammer strikes from below, setting the string into a vertical motion, which decays quickly. Due to the coupling of bridge and soundboard, the plane of vibration gradually rotates to a horizontal motion which decays more slowly [1]. This is referred to as the typical "double decay" of the piano, as shown in Figure 1(a).

In a piano, most notes have more than one string per note. Usually, only the lowest ten or so notes have one string per note. The subsequent (about 18) notes have two strings, and the rest have three strings. For notes with multiple strings, the decay rate will double or treble if the strings are tuned to exactly the same frequency. To make the sound sustain longer, the strings are tuned to slightly different frequencies. The strings interact via the piano bridge which causes a coupled oscillation. If the mistuning (the frequency difference between strings) is small, the coupled motion will result in a double decay, as shown in Figure 1(b). When the mistuning is large, beats appear. In this case, the decay becomes a periodic decaying

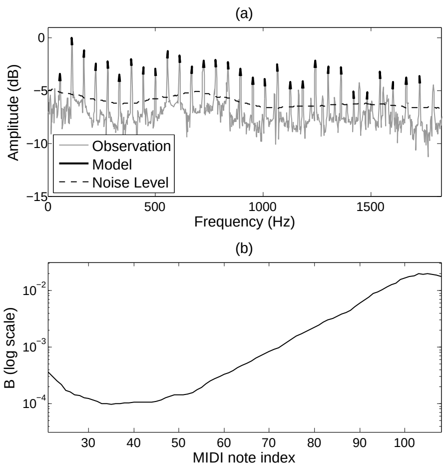

Fig. 2. (a) Partial frequencies of note A1 (55Hz), (b) Inharmonicity coefficient B along the whole compass estimated for the piano of the RWC dataset.

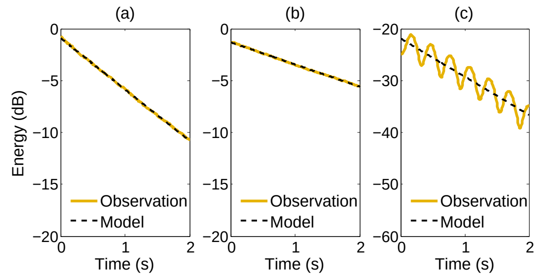

Fig. 3. Linear fitting for: (a) the 3^{rd} partial of note D1 (36.7Hz); (b) the 2^{nd} partial of note A♭1 (51.9Hz), (c) the 30^{th} partial of note D♭2 (69.3Hz).

curve, as shown in Figure 1(c). Note that the beat frequency is not equal to the frequency difference between strings, due to the coupled oscillation of strings [1]. We also observe beats in high partials of single string notes. This is known as "false beats" which is caused by imperfections in string wire or problems at the bridge such as loose bridge pins.

Although the theory of piano acoustics is well understood, measuring it in practice is complicated. In the next section, we track the temporal decay of individual partials of real-world piano recordings based on this theory.

3. METHOD

In order to track the decay of piano notes, we first find the frequencies of partials for each note, taking inharmonicity into account. Then, the decay of the partials is fitted in three typical patterns: linear decay, double decay and beats.

3.1. Finding Partials

Because of string stiffness, partials of piano notes occur at higher frequencies than the harmonics (integer multiples of the fundamental frequency), which is known as inharmonicity. The partial frequencies are given by [10]:

$$f_n = nf_0\sqrt{1 + Bn^2},$$

where f_n is the n^{th} partial of the note with fundamental frequency f_0 and B is the inharmonicity coefficient which varies from note to note. Moreover, during the course of a sounded note partial frequencies can diverge from their ideal inharmonic frequencies due to the coupling between bridge and soundboard. To get the frequencies of partials, we estimate B and f_0 in a non-negative matrix factorization framework proposed in [9]. We provide an implementation of this method for download[1].

Figure 2(a) shows the detected frequencies of the first 30 partials of the note A1 (55Hz). The estimated inharmonicity coefficient along the piano compass is given in Figure 2(b).

[1]https://code.soundsoftware.ac.uk/projects/inharmonicityestimation

3.2. Tracking the Decay of Partials

We first represent the spectrogram on a dB scale, then fit the partial decay using three models. We use a 2s sound clip beginning at 0.2s after the onset for each note to discard the transient part. The length is restricted to the duration of the note if it is shorter than 2s. The sampling rate is $f_s = 44100$Hz. Frames are segmented by a 4096-sample Hamming window with a hop-size of 441. A discrete Fourier Transform is performed on each frame with 2-fold zero-padding. The energy is represented on a log scale: $S_{dB} = 20\log_{10}(S)$, where S_{dB} is the log energy spectrogram in dB, and S is the magnitude spectrogram.

3.2.1. *Linear regression*

In most situations, the decay of partials follows a linear function of time, which is modelled by

$$y(t) = at + b,$$

where y is the linear function along time t, a is the decay rate and b is the initial energy. The regression parameters are estimated using ordinary least squares.

Figure 3 shows three kinds of decays which can be fitted in the linear model: (a) is a partial of a single-string note; (b) is a partial of a note with well-tuned strings; and (c) is a partial of a note with large mistuning between strings. When the mistuning is large, there is more than one beat in the decay, but the overall decay rate can be detected correctly using the linear model.

3.2.2. *Multi-phase linear regression*

A multi-phase linear model is employed to model double decay as well as fast decay with noise. Despite the misleading name this is a non-linear regression problem. The decay is modelled by two straight lines, formulated as follows:

$$y(t) = \begin{cases} a_1 t + b_1 & : t_s < t < t_{dp} \\ a_2 t + b_2 & : t_{dp} < t < t_e \end{cases}$$

where y is the estimated function; $a_1,\ a_2$ and $b_1,\ b_2$ are the decay rates and the initial energies of the two lines, respectively; t_{dp} is the demarcation point of the two lines; and t_s and t_e are the starting time and the ending time, respectively. Parameters are estimated using an existing method[2].

Figure 4(a) shows the fit for a partial with two parts, decay and noise. This partial decays quickly, having a low initial amplitude

[2]http://staff.washington.edu/aganse/mpregression/mpregression.html

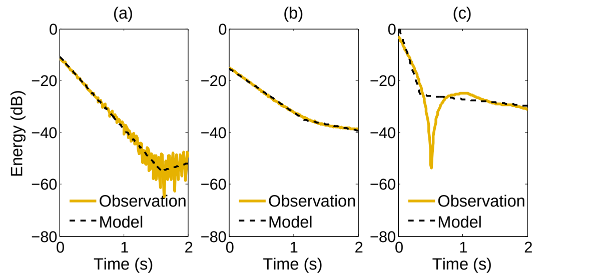

Fig. 4. Multi-phase linear fitting for: (a) the 7^{th} partial of note B♭0 (29.1Hz); (b) the 4^{th} partial of note B♭2 (116.5Hz), (c) the 1^{st} partial of note E5 (659.3Hz).

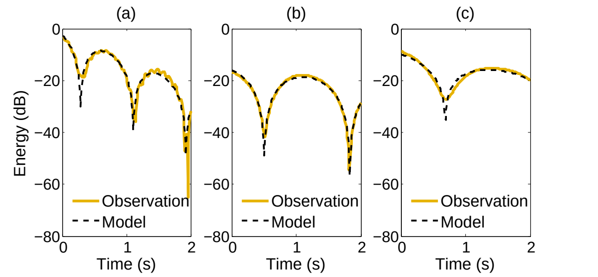

Fig. 5. Non-linear curve fitting for: (a) the 22^{nd} partial of note D♭1 (34.6Hz); (b) the 10^{th} partial of note G1 (49Hz); and (c) the 10^{th} partial of note A1 (55Hz).

due to the hammer impulse position being near a node of the partial's vibration mode, and the late portion is noise which should be discarded. Fitting this decay with the multi-phase model helps to automatically detect the ending time of the partial decay and discard the noisy part. Figures 4(b) and (c) indicate two kinds of double decay. The rate change in (b) is caused by the transmission direction switching from vertical to horizontal, while the reason for the double decay in (c) is a small frequency difference between the strings.

For double decay, the slope of the line which covers the larger part of the time duration is recorded as the decay rate. For the situation of fast decay with noise, the first decay rate (non-noisy part) is recorded.

3.2.3. *Non-linear curve fitting*

Due to frequency differences between strings of a note, partials decay according to curves resembling the result of amplitude modulation. We use a non-linear curve fitting model to fit these decays with beats. The objective function of the curve fitting is more complex than the first two situations. Based on the theory of coupled strings [1], the formula is simplified as follows:

$$y(t) = at + b + A\log_{10}(|\cos(ft + \varphi)| + \varepsilon).$$

This function illustrates the coupled motion of two strings. It consists of two parts: the decay part is still modelled by a linear function, $at + b$; and the remaining term models the amplitude modulation, where A and f are the magnitude and frequency of the beats, respectively. φ is the initial phase, which is included because the tracking starts at 0.2s after the onset so the phases of the partials are different. $\varepsilon = 0.01$ is added to avoid taking the log of 0.

Figure 5 gives three examples of decay with beats: (a) false beats of a high partial of a single string note; (b) and (c) show beats with different frequencies, which are caused by mistuning of the strings of each note.

The parameters are estimated using a non-linear least squares algorithm [11]. This method requires a good initialisation and ranges for parameters to get a reasonable result. The linear part (a, b) is initialised using the result of the linear model. A is initialised to 20 which is the amplitude of purely resistive coupling in dB. If there is more than one cycle of amplitude modulation (Figure 5(b)), we initialise f according to the time gap between two adjacent troughs. For a curve with less than one cycle (Figure 5(c)), we assume the position of the first trough is half of the period.

The coupling between 3 strings is far more complex than for 2 strings. It is out of the scope of this paper to explore the details of the motion of 3 strings, which we approximate using the models described above.

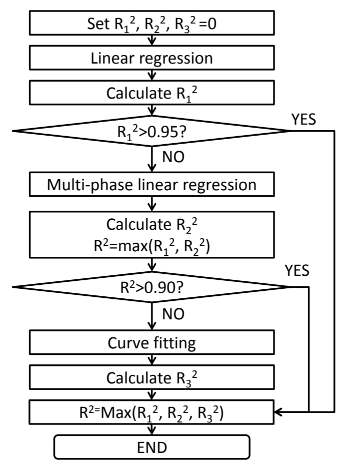

Fig. 6. Flowchart of partial-decay modelling.

4. RESULTS

We use the piano sounds from the RWC Musical Instrument Sound Database [12]. The notes are played at three dynamics: loud (forte, *f*), moderate (mezzo, *m*) and soft (piano, *p*). Each clip consists of 88 isolated notes covering the whole compass of the piano.

The coefficient of determination (R-squared) is used for evaluating the fit between the data and model. It is defined as follows:

$$R^2 = 1 - \frac{\sum_t (y_t - o_t)^2}{\sum_t (o_t - \overline{o})^2},$$

where o_t is the observation of time frame t, y_t is the modelled function and $\overline{o}$ is the mean of the observations o_t. For each partial, if its energy is above the noise level, we fit it to our models according to the algorithm shown in Figure 6.

We illustrate the results of decay tracking. The estimated decay rates are used to parameterise the decay response and to explore the influence of dynamics.

4.1. R-squared

We compare the average coefficient of determination, R^2, between the linear model and all three models described in Section 3 (referred to as the mixed model). The results are presented in Table 1, which indicate that the mixed model has a better fit to the data by around 15%. We also note that the performance is influenced by dynamics.

Dynamics	NP	linear model	mixed model
f	1512	0.727	0.878
m	1423	0.686	0.831
p	1226	0.643	0.800

Table 1. Average R^2 of the linear and mixed models. NP is the number of partials above the noise level for each dynamic level.

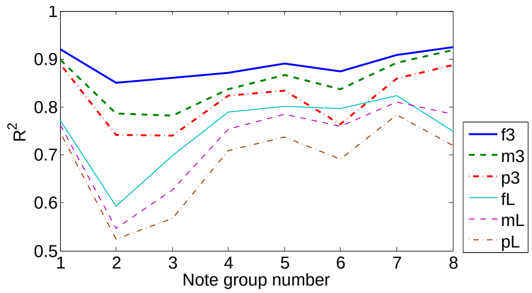

Fig. 7. Average R^2 of different note groups. *f, m, p* stand for dynamics, while L and 3 indicate the linear and mixed models, respectively. The order of labels in the legend corresponds to the order of lines from top to bottom.

If the note starts at a lower energy, it will decay to the noise level more quickly, resulting in there being fewer data available for modelling, not only fewer partials above the noise level, but also shorter duration of notes. This reduction in data makes parameter estimation more difficult, resulting in worse performance for lower dynamics.

In order to provide a more detailed investigation of the results for different notes, we group every adjacent 11 notes together, giving 8 groups. The average R^2 of each group is shown in Figure 7. We find that the mixed model improves the performance of all note groups, with the biggest improvement of around 0.2 occurring at Group 2 for all dynamics, corresponding to the observation of clear beats in the decay of these notes. Beats appear extensively in notes from Groups 2 and 3, hence the linear model performs poorly on these notes and the largest improvements are attained by the mixed model. Although we don't explicitly model the details of motion in three strings, the results show that the decays of notes in the high frequency range are approximated quite well by the mixed model.

4.2. Decay Response

Figure 8 shows the decay rates of all partials along the whole compass of the piano for notes played forte. The figure illustrates the well-known fact that high frequency partials decay faster. The spread of observed decay rates is large, and increases with frequency. Note that some frequencies in the low range, around 80 Hz (MIDI index 41) and 150 Hz (MIDI index 50), exhibit particularly fast decay rates. As partials from different notes may have the same frequency, different decay rates of these partials could be used as a clue to decide which note the partial belongs to. However, in music performances, overlapped partials increase the difficulty of tracking partial decay, which is a topic needing further investigation.

Figure 9 shows the decay rates of the first five partials of notes played at different dynamics. We observe that dynamics have no significant absolute effect on the decay rate. In the low frequency range, the decay rates of different dynamics are almost identical,

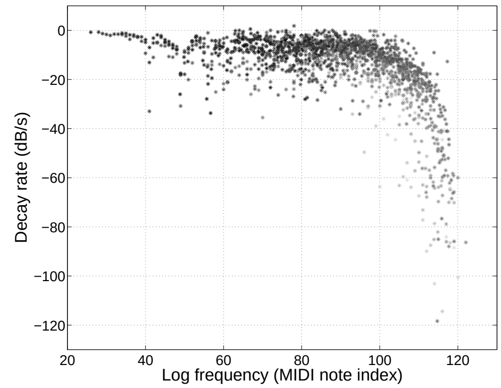

Fig. 8. Decay response: decay rates against frequency. Lower values mean faster decay. The greyscale is used to indicate partials of the same note, with darker colours corresponding to lower pitches.

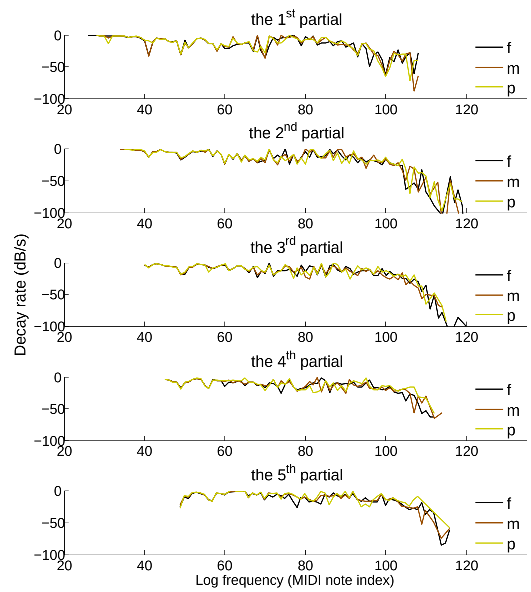

Fig. 9. Decay rates for the first five partials for different dynamics

while in the high frequency range this is less clear, possibly due to higher measurement error.

5. CONCLUSIONS

In this paper we model the decay of piano notes based on piano acoustics theory. Two non-linear models (multi-phase linear model and non-linear curve fitting model) are used to fit double decay and beats of piano tone partials, respectively. The results show that the use of non-linear models provides a better fit to the data, especially for notes in the low register. The decay response of the piano shows various decay rates along the frequency range. The results also indicate that dynamics have no significant effect on the decay rate.

In the future, we would like to investigate using the decay information for onset-offset detection, auditory scene analysis and other music analysis applications.

6. REFERENCES

[1] G. Weinreich, "Coupled piano strings," *The Journal of the Acoustical Society of America*, vol. 62, no. 6, pp. 1474–1484, 1977.

[2] T. C. Hundley, H. Benioff, and D. W. Martin, "Factors contributing to the multiple rate of piano tone decay," *The Journal of the Acoustical Society of America*, vol. 64, no. 5, pp. 1303–1309, 1978.

[3] V. Välimäki, J. Huopaniemi, M. Karjalainen, and Z. Jánosy, "Physical modeling of plucked string instruments with application to real-time sound synthesis," *Journal of the Audio Engineering Society*, vol. 44, no. 5, pp. 331–353, 1996.

[4] M. Aramaki, J. Bensa, L. Daudet, P. Guillemain, and R. Kronland-Martinet, "Resynthesis of coupled piano string vibrations based on physical modeling," *Journal of New Music Research*, vol. 30, no. 3, pp. 213–226, 2001.

[5] B. Bank, "Accurate and efficient modeling of beating and two-stage decay for string instrument synthesis parallel waveguides resonator bank," in *Proc. MOSART Workshop on Current Research Directions in Computer Music*, Barcelona, Spain, 2001, pp. 134–137.

[6] M. Karjalainen, P. Antsalo, A. Mäkivirta, T. Peltonen, and V. Välimäki, "Estimation of modal decay parameters from noisy response measurements," *Journal of the Audio Engineering Society*, vol. 50, no. 11, pp. 867–878, 2002.

[7] N. Lee, J. O. Smith, and V. Välimäki, "Analysis and synthesis of coupled vibrating strings using a hybrid modal-waveguide synthesis model," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 18, no. 4, pp. 833–842, 2010.

[8] S. Ewert and M. Müller, "Estimating note intensities in music recordings," in *proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2011, pp. 385–388.

[9] F. Rigaud, B. David, and L. Daudet, "A parametric model and estimation techniques for the inharmonicity and tuning of the piano," *The Journal of the Acoustical Society of America*, vol. 133, no. 5, pp. 3107–3118, 2013.

[10] H. Fletcher, E. D. Blackham, and R. Stratton, "Quality of piano tones," *The Journal of the Acoustical Society of America*, vol. 34, no. 6, pp. 749–761, 1962.

[11] J. Fox and S. Weisberg, "Nonlinear regression and nonlinear least squares in R," in *Appendix to An R Companion to Applied Regression, second edition*, 2010, pp. 1–20.

[12] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC Music Database: Music Genre Database and Musical Instrument Sound Database," in *Proc. the 4th International Conference on Music Information Retrieval (ISMIR)*, 2003, pp. 229–230.