

A MODEL OF PARTIAL TRACKS FOR TENSION-MODULATED STEEL-STRING GUITAR TONES

Matthew Hodgkinson, Joseph Timoney and
Victor Lazzarini

Digital Sound and Music Technology Group
National University of Ireland, Maynooth
matthew.hodgkinson@nuim.ie,
jtimoney@cs.nuim.ie,
victor.lazzarini@nuim.ie

ABSTRACT

This paper introduces a spectral model for plucked, steel string tones, based on functional models for time-varying fundamental frequency and inharmonicity coefficient. Techniques to evaluate those analytical values at different time indexes are reviewed and commented. A method to evaluate the unknowns of the fundamental frequency and inharmonicity coefficient functions and match the data of a given tone is presented. Frequency tracks can thereafter be deployed and traced for all values of time. Their accuracy is discussed, and applications for the model are suggested.

1. INTRODUCTION

The idea of spectral modeling of plucked-string tones as introduced in this paper arose when confronting the problem of partial tracking, one of several steps of Spectral Modeling Synthesis (SMS). After identifying the harmonics of a spectrum around a given time index, SMS finds the harmonics of the next analysis frame around the harmonic frequencies previously obtained [1]. Obviously, this method is entirely reliable only if those frequencies remain constant throughout the sound. If this is not the case, a small hop size from one frame to the next is needed so that the partial looked for in the following frame still stands within a reasonably narrow region around the search's center frequency. But even so, strong magnitude beating of a partial may cause the latter to drop below the noise floor for several frames, rendering inconsequent the reduction of the hop size. When it emerges again, the partial may have fallen outside the boundaries of the search, or another one of greater amplitude might have intruded the search region, whose center missed being updated during this time. So instead of basing one's search on adjacent frequency values in time, one might want to rely on a harmonic series expression instead.

For gut or nylon strings, such as those found on Spanish and Classical guitars, it is sufficient to resort to a harmonic series in its simplest form, $f_k = kf_0$, where f_k is the frequency of the k^{th} harmonic, and f_0 , the fundamental frequency of the series. (Also, note that in this case, the fundamental frequency equals the frequency of the first partial.) On the other hand, the harmonic series of strings showing some degree of stiffness (e.g. steel strings) do not exhibit this kind of linearity anymore, and their expression must be replaced with

$$f_k = kf_0 \sqrt{1 + \beta k^2}, \quad (1)$$

where we have introduced β , an inharmonicity coefficient. Added to that, when the string is plucked hard, fundamental frequency and inharmonicity coefficient both become time-dependent. Figure 1 exemplifies this fact, plotting measurements over time of the fundamental frequency and inharmonicity coefficient of the open E3 of an acoustic guitar played *mezzo forte*.

When those two values become time-dependent, the whole series becomes time-dependent too:

$$f_k(t) = kf_0(t) \sqrt{1 + \beta(t)k^2}. \quad (2)$$

(2) suggests that the modeling of $f_0(t)$ and $\beta(t)$ could yield the modeling of a whole spectrum, at least in terms of frequency. As will be shown later, and provided faithful $f_0(t)$ and $\beta(t)$ representations, the model is sufficiently accurate to assist in the tracking of partials, but may also be useful to other applications, such as the synthesis of string tones with subtly evolving spectra.

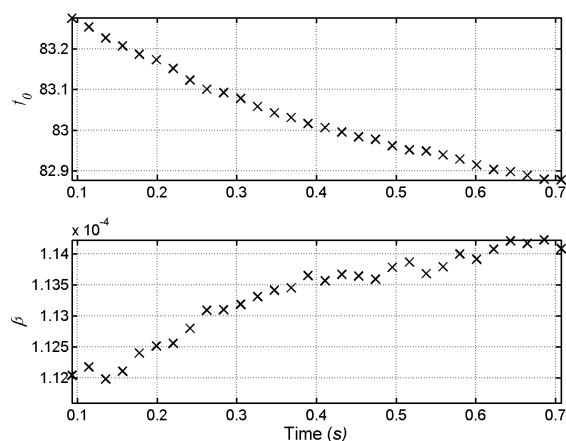


Figure 1: Fundamental frequency and inharmonicity coefficient time measurements performed on mezzo-forte acoustic guitar tone (open E3)

2. CHOICE OF A FUNCTIONAL MODEL

The time-variance of the fundamental frequency and inharmonicity coefficient can be observed from such measurements as those presented in Figure 1, as well as from the study of their analytical expressions. Physically, the fundamental frequency of a string

with length L , tension T and mass density μ (in kilogrammes per metre) is expressed as

$$f_0 = \frac{1}{2L} \sqrt{\frac{T}{\mu}} \quad (3)$$

The inharmonicity coefficient also depends on the string's diameter D and elasticity modulus Q [2], as per

$$\beta = \frac{\pi^3}{64} \frac{QD^4}{TL^2} \quad (4)$$

All those may be considered constant for non-steel strings, or for steel strings when the transverse vibrations are so small the changes in length – and thus in tension and mass density – are negligible. For instance, in spite of showing very high values of inharmonicity, Steinway grand piano tones show no significant trend in either fundamental frequency or inharmonicity changes (e.g. Figure 2).

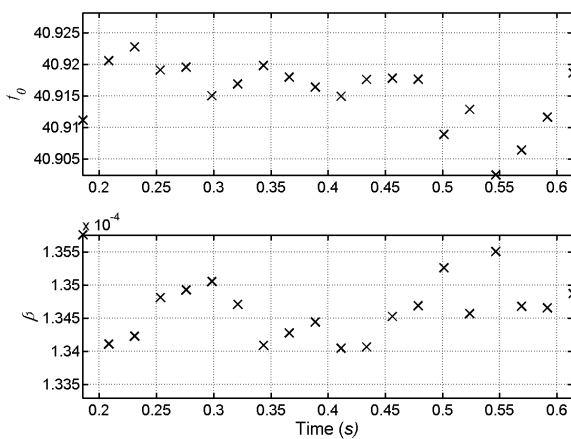


Figure 2: Steinway grand piano fundamental frequency and inharmonicity coefficient measurements on fortissimo E2

Plucked strings, however, can easily show distinctive such trends. The trend observed in the fundamental frequency measurements presented in Figure 1 clearly is that of an exponentially decaying function, not converging towards 0, but toward the “fundamental frequency at rest”. We will therefore try fit a function $f_0(t)$ of the form

$$f_0(t) = (f_{0,0} - f_{0,\infty}) e^{-\frac{t}{\tau_{ff}}} + f_{0,\infty} \quad (5)$$

in our measurements, where $f_{0,0}$ is the function's value at time 0, $f_{0,\infty}$ the value it converges to as its independent variable t approaches infinity, and τ_{ff} the function's decay time (where “ff” stands for “fundamental frequency”).

A functional choice for $\beta(t)$ seems more delicate. Here, further study of the inharmonicity coefficient's analytical expression guided us towards the solution proposed in this paper. Assuming constant string cross-section diameter D for the string, the time-dependent variables in the expression of β come down to the tension and length. Considering the string in two dimensions, where the transverse displacement y is a function of space x and time t ,

the length of the vibrating string can be obtained with the expression

$$L(t) = \int_0^{L_0} \sqrt{1 + \left(\frac{dy}{dx}\right)^2} dx \approx L_0 + \frac{1}{2} \int_0^{L_0} \left(\frac{dy}{dx}\right)^2 dx \quad (6)$$

where L_0 is the length of the string at rest.

(6) shows that the length of a vibrating string can be approximated as the addition of the length of the string at rest and a time-dependent term $L(t)$, i.e. $L(t) = L_0 + L(t)$. A good approximation for L' was developed in [3] and [4] to an infinite sum of exponentials,

$$L'(t) = \frac{\pi^2}{8L_0} \sum_{k=1}^{\infty} A_k^2 k^2 e^{-\frac{2t}{\tau_k}} \quad (7)$$

where k is the index of the k^{th} harmonic, A_k its amplitude at time 0, and τ_k its decay time.

Similarly to $L(t)$, we express the tension $T(t)$ as the sum of the tension at rest T_0 with a time-dependent term $T'(t)$, i.e. $T(t) = T_0 + T'(t)$. In fact, $T'(t)$ is directly related to $L'(t)$, as in

$$T'(t) = \frac{\pi^2 QS}{8L_0^2} \sum_{k=1}^{\infty} A_k^2 k^2 e^{-\frac{2t}{\tau_k}} = \frac{QS}{L_0} L'(t) \quad (8)$$

Substituting those tension and length expressions into (x) yields

$$\beta(t) = \frac{\frac{\pi^3}{64} QD^4}{\frac{QS}{L_0} L'(t)^3 + (T_0 + 2QS)L'(t)^2 + (2T_0 + QS)L_0 L'(t) + T_0 L_0^2} \quad (9)$$

At the denominator of the inharmonicity coefficient, we have the weighted sum of powers of $L'(t)$ plus a constant. It must be remembered that the individual terms of this sum are decaying exponentials, whose individual gradient and curvature are negative and positive, respectively, for all values of time. As their sum, $L'(t)$ inherits those properties, and so do its powers. Therefore, we have at the denominator a value which, like (5), exponentially decays and converges towards a positive value. On that basis, we use an expression for $\beta(t)$ which accommodates a term the form of (5) in its denominator, with a decay time τ_{ic} , and arranged so that the function as a whole equals β_0 for $t = 0$, and converges towards β_∞ as t approaches ∞ :

$$\beta(t) = \frac{\beta_0 \beta_\infty}{(\beta_\infty - \beta_0) e^{-\frac{t}{\tau_{ic}}} + \beta_0} \quad (10)$$

3. PRELIMINARY ANALYSIS

Later in this paper we will discuss the fitting of (5) and (10) in the measurements over time of the fundamental frequency and inharmonicity coefficient, respectively. To do that it is obviously necessary to have the means of evaluating those analytical values as precisely as possible, all the more that order of the range of variation within those two variables is around 1 percent (e.g.

0.5% for the fundamental frequency and 1.8% for the inharmonicity coefficient in Figure 1) – a far greater order of accuracy is therefore necessary to detect a trend within such narrow intervals. In DAFX 2009 was presented an original way of estimating the fundamental frequency and inharmonicity coefficient of a tone from its magnitude spectrum using Median-Adjustive Trajectories (MAT) [5]. Tests supported that the algorithm stood as the most accurate and computationally efficient to date.

3.1. Median-Adjustive Trajectories

The starting point of the method is the expression, in terms of two partial frequencies f_m and f_n and their partial numbers m and n , of the fundamental frequency,

$$f_0 = \sqrt{\frac{n^4 f_m^2 - m^4 f_n^2}{n^4 m^2 - m^4 n^2}}, \quad (11)$$

and the inharmonicity coefficient,

$$\beta = -\frac{n^2 f_m^2 - m^2 f_n^2}{n^4 f_m^2 - m^4 f_n^2}. \quad (12)$$

If, in reality, the partial frequencies followed (1) exactly, then the measurement of two partials only, as well as their identification (i.e. the attribution to each of a partial number), would be sufficient to get an exact estimate of β and f_0 . However, harmonics are always subject to random frequency fluctuations, as minor as those may be, and frequency measurement ultimately come down to some approximation. It is therefore desirable to perform the calculation upon all possible combinations of two partials, and derive a final estimate statistically. Within a spectrum where K partials were measured, the potential number of β and f_0 estimates by two-partial combinations rises to $(K^2-K)/2$. One might create the vectors

$$\mathbf{B} = -\frac{k^2 f_m^2 - m^2 f_k^2}{k^4 f_m^2 - m^4 f_k^2}, \begin{cases} k = 1, 2, \dots, K-1 \\ m = k+1, k+2, \dots, K \end{cases} \quad (13)$$

and

$$\mathbf{F}_0 = \sqrt{\frac{k^4 f_m^2 - m^4 f_k^2}{k^4 m^2 - m^4 k^2}}, \begin{cases} k = 1, 2, \dots, K-1 \\ m = k+1, k+2, \dots, K \end{cases} \quad (14)$$

From the obtained vectors of fundamental frequency and inharmonicity coefficient estimates, the median values were found to provide the best final estimates.

Identifying the harmonics within a spectrum that is inharmonicity-stretched is, however, a problem that needs to be addressed. Figure 3 shows that the deviation increases in a non-linear, exponential-like way as the partial index increases. Although the stretch is negligible around the first few partials, we see that they may eventually deviate from their harmonic position by more than one index (upper plot), or even, for strongly inharmonic spectra, by more than ten times the fundamental frequency. With such deviation, a harmonic may easily be taken for another, or phantom partials (sustained sinusoids issued from non-linear or longitudinal modes of vibration) may intrude the analysis. Median-Adjustive Trajectories were shown to be an efficient method to overcome this difficulty. Here, the search for

harmonics starts at the bottom of the series, where the inharmonicity effect is negligible, and as the search progresses upwards, all combinations of the previously identified peaks are concatenated into the inharmonicity coefficient and fundamental frequency vectors \mathbf{B} and \mathbf{F}_0 , from which the medians are extracted and substituted in (1) to guess the frequency of the next partial to be found [5].

The accuracy of the inharmonic trajectory this upward progression is going to take depends a great deal on the accuracy of the partial frequency estimates, especially as we reach the upper region of the spectrum, where the level of noisy components starts to rival with that of the pseudo-harmonics we are looking for. As the success of the spectral fit presented below entirely relies on accurate fundamental frequency and inharmonicity coefficient estimates, we deem necessary to allocate a subsection to the problem of frequency component estimates.

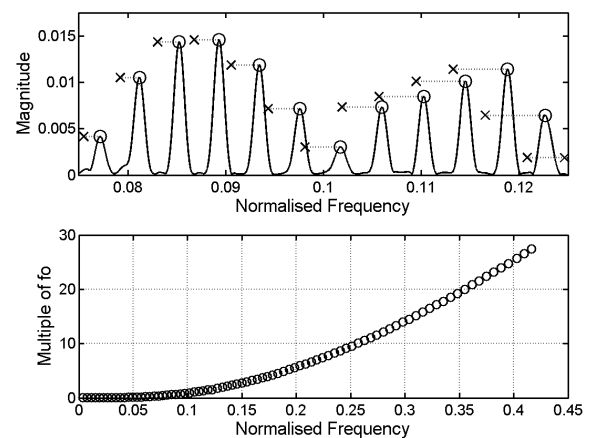


Figure 3: Inharmonicity Stretch (Acoustic Guitar E3, Inharmonicity Coefficient $\approx 1.12 \cdot 10^{-4}$)

3.2. Estimate of Frequency Components

A DFT yields complex values – which can be converted to magnitude and phase data – at discrete, evenly spaced frequency points. A unique, complex-valued frequency component that is Fourier-transformed appears in the transform’s magnitude spectrum as a peak that is even around the frequency where the component actually lies. On that basis, it is common practice to try estimate with precision where the summit of the peak is, and take the corresponding frequency position for the frequency estimate of the underlying partial. To do so, the most popular techniques are zero-padding, and polynomial fitting [5][6].

Zero-padding of the input segment of sound augments the size M (in samples) of the DFT, and therefore reduces the frequency interval between adjacent bins, $\omega_{\text{DFT}} = 2\pi/M$. However, one cannot abuse zero-padding, as it increases the computational cost of the analysis. For further refinement, one may resort to polynomial fitting. Here, the greatest discrete magnitude value as well as its immediate lower and upper neighbours are fit a second-order polynomial. The frequency where the polynomial soars to its maximum, i.e. where its derivative equates to 0, is taken as the final estimate for the frequency of the partial.

An alternative method, Complex Spectral Phase Evolution (CSPE), was introduced in [7]. Here, two DFTs are necessary, the second of which is performed upon a frame of the analysed sound that is delayed by one sample. The angular frequency ω_{CSPE} of the predominant partial in the vicinity of the DFT index

i of a magnitude peak is simply obtained through the product of the complex value produced by the first DFT, $DFT(i)$, with the complex conjugate of the value produced by the delayed DFT, $[DFT'(i)]^*$, as in

$$\omega_{\text{CSPE}}(i) = -\mathcal{L}\left(DFT(i)[DFT'(i)]^*\right). \quad (15)$$

Tests in [5] confronted the accuracy of CSPE with that of second-order polynomial fitting. They were performed upon single-component, real-valued signals, which are known to exhibit two magnitude peaks that are symmetrically opposed around 0Hz and the Nyquist frequency. For signals whose frequency lied halfway between those two limits, in which case the interaction between the symmetrical peaks is negligible, the accuracy of the CSPE was shown to surpass that of polynomial fitting by several orders of magnitudes.

We performed new tests to visualise the effect of the bias in the estimate of a frequency component introduced by the presence of another component, close in frequency and magnitude. It is explained in [1] how, for the analysis to make the distinction between two frequency components of respective angular frequencies ω_1 and ω_2 , the size N of the windowed segment of sound $w[n]x[n]$, $n = 0, 1, \dots, N-1$, should equal at least $2\pi B/|\omega_2 - \omega_1|$, where B is the bandwidth, in frequency bins, of the window $w[n]$ (e.g. 2 for a rectangular window, 4 for a Blackman window). Conversely, for a fixed window size N , ω_2 should differ from ω_1 by at least $\pm 2\pi B/N$. Here, we set $\omega_2 = \omega_1 + \pi B/N$ instead.

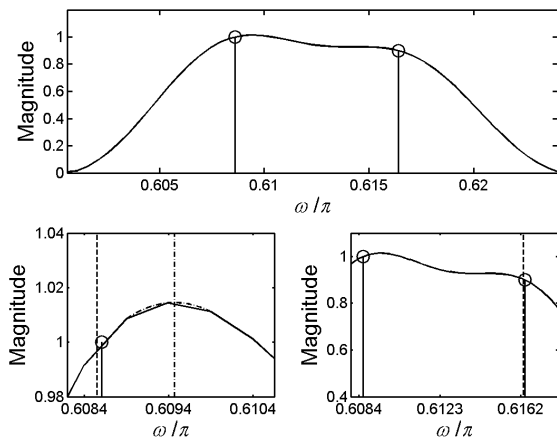


Figure 4: *Magnitude Spectrum of Unresolved Partial (stems), polynomial fit estimate (dash-dotted), CSPE estimates (dashed)*

Figure 4 shows the magnitude spectrum obtained from the DFT of the windowed test signal. The stems reveal the actual components, both in frequency and magnitude. The reader should note that, as opposed to the single complex-valued component case mentioned above, the actual position of the dominant frequency component does not lie underneath the summit of the peak anymore, because of the spectral leakage of the other component. In such case, striving to find the summit with zero-padding and polynomial fitting becomes futile. The spectral leakage affects much less the estimate provided by the CSPE. The bottom-left subplot illustrates the situation.

More importantly perhaps, summit finding would fail completely to estimate the frequency of the component of lesser

magnitude (if that were the component to be estimated) as it simply is too weak to produce a peak of its own. Again, provided that the bin i in $\omega_{\text{CSPE}}(i)$ is sufficiently closer to the weaker partial, the absence of a peak affects little the CSPE approach (bottom-right subplot). Such situation is not uncommon in the frequency estimate of partials that are high up in a sound's spectrum. The fundamental frequency and inharmonicity coefficient estimates returned by the MAT method are more reliable when a maximum number of partials are used, implying the necessity of exploiting those high, noisy spectral regions.

3.3. Sliding MAT

The inharmonicity coefficient and fundamental frequency need to be estimated throughout the sound at different, preferably evenly-spaced time indexes. The procedure here is similar to that for obtaining a spectrogram: a window size N , DFT size M and hop size H need to be defined.

We expressed in the previous subsection the minimum window size N required for the DFT to resolve neighbour partials. Here, as we are looking for near-integer multiples of a fundamental frequency f_0 , N should be equal to or greater than Bf_s/f_0 . As an *a priori* estimate, the constant value to be substituted in place of f_0 can be derived from an equal-temperament scale in function of the note being analysed (e.g. 440 for an A4, $440 \cdot 2^{1/12}$ for an A#4, $440 \cdot 2^{5/12}$ for a D5, etc). Sounds with static spectrograms can be analysed with large DFT sizes. However, acoustic guitar sound spectrograms are only quasi-static, given such fundamental-frequency and inharmonicity coefficient time-variance we witnessed in Figure 1. Still, we favour here rather large window sizes, ranging between $2Bf_s/f_0$ and $4Bf_s/f_0$, for three arguable reasons:

- The statistical nature of the MAT-based FF and IC estimates reduces the effect of errors introduced in the measurement of partial frequencies from the time-variance of the latter.
- All the more sensitive to partial frequency time-variance, the phase data obtained from the DFT is not used.
- At large harmonic indexes, large window sizes facilitate the distinction of pseudo-harmonics from noise components.

The size of the DFT M should be a power of two greater than N . The benefits, limitations and drawbacks of zero-padding were discussed in the previous subsection.

Finally, the hop size H should be chosen in regard of the process used to fit the $f_0(t)$ and $\beta(t)$ models in the data. For reasons that will be explained in the section on curve-fitting, a number W of f_0 and β measurements yields a maximum of

$$C = \binom{W}{3} = \frac{W!}{6(W-3)!} \quad (16)$$

estimates for each three parameters of each $f_0(t)$ and $\beta(t)$ functions. Within the first second of string tones, we typically set W to a value between 20 and 30, so that C ranges between 1,140 and 4,060. H can be determined from W and S , the number of sound samples used in the modeling, via the expression

$$H = \left\lfloor \frac{S-N}{W-1} \right\rfloor. \quad (17)$$

A series of MAT-based inharmonicity coefficient and fundamental frequency estimates, slid across the sound in the manner described above, should provide the user with such data as plotted in Figure 1 and Figure 2.

4. DATA-FITTING OF FUNDAMENTAL FREQUENCY AND INHARMONICITY COEFFICIENT MODELS

The problem now is to find the unknowns the fundamental frequency and inharmonicity coefficient models that will yield the best fit into the data obtained through the preliminary analysis. We opt for the same approach used in the MAT method: solve the models' equations for each of their unknowns, evaluate the unknowns with the greatest number of data point combinations as possible, and take the median of each unknown's vector thus obtained [5].

In (1), there were two unknowns, f_0 and β , which led the solution for each to contain two frequency measurements and corresponding partial indexes, as one can see in (11) and (12). Now, each of the fundamental frequency and inharmonicity coefficient time-dependent models has three unknowns (e.g. $f_{0,0}$, $f_{0,\infty}$ and τ_{ff} for the fundamental frequency), and so each solution will feature three measurements and corresponding time indexes (we have left the frequency domain for the time domain).

Due to the similarity between the fundamental frequency and inharmonicity coefficient time models, the process to solve the unknowns of one is similar to the process to solve the unknowns of the other. We will begin with case of the fundamental frequency, which we will develop fully. One will find the development of the inharmonicity coefficient case more condensed.

4.1. Estimate of Fundamental Frequency Model Parameters

Three unknowns, three equations needed: we take three fundamental frequency estimates $f_{0,p}$, $f_{0,q}$ and $f_{0,r}$ at three different times t_p , t_q and t_r .

First, we introduce $f_{0,p}$, $f_{0,q}$ and solve (5) for $f_{0,0}$,

$$f_{0,0} = \frac{f_{0,p} \left(e^{\frac{t_q}{\tau_{ff}}} - 1 \right) - f_{0,q} \left(e^{\frac{t_p}{\tau_{ff}}} - 1 \right)}{e^{\frac{t_q}{\tau_{ff}}} - e^{\frac{t_p}{\tau_{ff}}}}, \quad (18)$$

and for $f_{0,\infty}$,

$$f_{0,0} = \frac{f_{0,p} e^{\frac{t_q}{\tau_{ff}}} - f_{0,q} e^{\frac{t_p}{\tau_{ff}}}}{e^{\frac{t_q}{\tau_{ff}}} - e^{\frac{t_p}{\tau_{ff}}}} \quad (19)$$

The real difficulty arises when we introduce the third measurement, $f_{0,r}$, into (5), substitute (18) and (19), and try solve for τ_{ff} :

$$e^{\frac{t_p}{\tau_{ff}}} (f_{0,r} - f_{0,q}) + e^{\frac{t_q}{\tau_{ff}}} (f_{0,p} - f_{0,r}) + e^{\frac{t_r}{\tau_{ff}}} (f_{0,q} - f_{0,p}) = 0 \quad (20)$$

An analytical solution for τ_{ff} could not be found. This could have something to do with the constraints of $f_0(t)$. We have seen that $f_0(t)$, as an exponential decaying function, had negative derivative and positive curvature for all t . Therefore, provided that

$t_p < t_q < t_r$, a negative derivative for all t implies for our measurements that

$$f_{0,p} > f_{0,q} > f_{0,r}, \quad (21)$$

and a positive curvature, that

$$(f_{0,r} - f_{0,q}) / (t_r - t_q) > (f_{0,q} - f_{0,p}) / (t_q - t_p). \quad (22)$$

To solve (20), we resort here to a numerical method. We find any combination of three fundamental frequency measurements that satisfy the conditions (21) and (22), and make (20) a function $y(x)$,

$$y(x) = e^{\frac{t_p}{x}} (f_{0,r} - f_{0,q}) + e^{\frac{t_q}{x}} (f_{0,p} - f_{0,r}) + e^{\frac{t_r}{x}} (f_{0,q} - f_{0,p}) \quad (23)$$

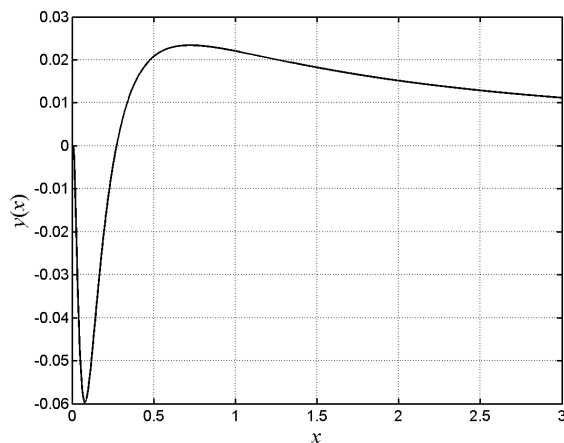


Figure 5: An Example of (23) for Consistent Values of $f_{0,p}$, $f_{0,q}$ and $f_{0,r}$

As illustrated in Figure 5, (23) is a function that tends to 0 as τ_{ff} approaches both 0 and ∞ . Its gradient is positive only within the interval that lies between its global minimum and maximum, which also is the region where it crosses the zero-axis. On that basis, we devised a numerical way of approximating τ_{ff} :

- 1) Evaluate (23) for a reasonably large value x_{\max} (e.g. 3 seconds), check that it is positive, and repeat the operation for a greater x_{\max} if not.
- 2) Evaluate (23), now for a number (e.g. 100) of values of x evenly spaced within the interval $(0, x_{\max}]$.
- 3) Find the index i of the greatest difference in the vector obtained in 2). If $y(x_i)$ is negative/positive, the zero-crossing is further to the right/left. Increment/decrement i until $y(x_i)$ becomes positive/negative, proceed to a linear interpolation between $y(x_{i-1})$ and $y(x_i)$, and find the value of x where the line equates to 0. This value will be read as the solution to (23).

4.2. Estimate of Inharmonicity Coefficient Model Parameters

Following the same steps as in the previous subsection, we attain an expression similar to (20) that we would like to solve for τ_{ic} :

$$e^{\frac{t_p}{\tau_{ic}}} \beta_p (\beta_r - \beta_q) + e^{\frac{t_q}{\tau_{ic}}} \beta_q (\beta_p - \beta_r) + e^{\frac{t_r}{\tau_{ic}}} \beta_r (\beta_q - \beta_p) = 0 \quad (24)$$

Likewise, we create a function $z(x)$,

$$z(x) = e^{\frac{t_p}{x}} \beta_p (\beta_r - \beta_q) + e^{\frac{t_q}{x}} \beta_q (\beta_p - \beta_r) + e^{\frac{t_r}{x}} \beta_r (\beta_q - \beta_p), \quad (25)$$

whose root we find numerically.

The first derivative of (10) is positive for all t , so we can set a first condition to the combinations of β_p , β_q and β_r that we are going to substitute in (25), i.e.

$$\beta_p < \beta_q < \beta_r. \quad (26)$$

The second derivative of (10), however, will be negative for positive values of t only if $\beta_\infty > 2\beta_0$. It was mentioned in the introduction to section 3, that the maximal inharmonicity value measured typically was greater than the minimal value by one percent or so. It can therefore reasonably be assumed that $\beta_\infty < 2\beta_0$, and thus we can set the condition of a negative gradient, i.e.

$$(\beta_q - \beta_p)/(t_q - t_p) > (\beta_r - \beta_q)/(t_r - t_q). \quad (27)$$

With those conditions, $z(x)$ becomes a function that tends towards zero as x approaches both zero and infinity, and whose gradient is negative only between its global maximum and minimum, which is also the interval wherein it crosses the zero-axis. We can therefore adapt the numerical method explained in the previous subsection that was used to approximate τ_{ff} to the approximation of τ_{ic} .

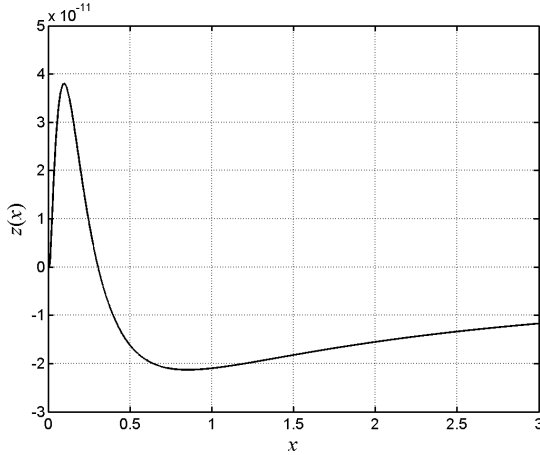


Figure 6: An example of $z(x)$ for values of β_p , β_q and β_r that meet the gradient and curvature conditions.

Once the decay time is approximated, β_0 and β_∞ can be computed as per

$$\beta_0 = \beta_p \beta_q \frac{e^{\frac{t_p}{\tau_{ic}}} - e^{\frac{t_q}{\tau_{ic}}}}{\beta_p \left(e^{\frac{t_p}{\tau_{ic}}} - 1 \right) - \beta_q \left(e^{\frac{t_q}{\tau_{ic}}} - 1 \right)} \quad (28)$$

and

$$\beta_\infty = \beta_p \beta_q \frac{e^{\frac{t_p}{\tau_{ic}}} - e^{\frac{t_q}{\tau_{ic}}}}{\beta_p e^{\frac{t_p}{\tau_{ic}}} - \beta_q e^{\frac{t_q}{\tau_{ic}}}}. \quad (29)$$

The parameters of the time-varying fundamental frequency and inharmonicity coefficient models are thus estimated for all possible combinations of measurements that meet the gradient and curvature conditions listed above. We thus come up with six vectors, $\mathbf{F}_{0,0}$, $\mathbf{F}_{0,\infty}$, \mathbf{T}_{ff} , \mathbf{B}_0 , \mathbf{B}_∞ and \mathbf{T}_{ic} , whose medians provide us with our final $f_{0,0}$, $f_{0,\infty}$, τ_{ff} , β_0 , β_∞ and τ_{ic} estimates.

5. RESULTS AND OBSERVATIONS

The models were fitted in fundamental frequency and inharmonicity coefficient time measurements of acoustic guitar tones. Other instruments, such as the grand piano, electric bass guitar, Chapman stick and Spanish guitars, did not show a distinctive trend in their measurements. The model was not tested on electric guitar tones, however, simply because no such samples could be accessed at the time of the writing of this paper. Yet the similarity of electric and acoustic guitar strings leads us to believe that the technique could be applicable to that instrument as well.

Figure 7 shows the fitting of the fundamental frequency and inharmonicity coefficient models in the data already presented in Figure 1. Here, $f_{0,0} = 83.4\text{Hz}$, $f_{0,\infty} = 82.8\text{Hz}$, $\tau_{ff} = 0.38\text{s}$, and $\beta_0 = 1.1 \cdot 10^{-4}$, $\beta_\infty = 1.14 \cdot 10^{-4}$, $\tau_{ic} = 0.23\text{s}$.

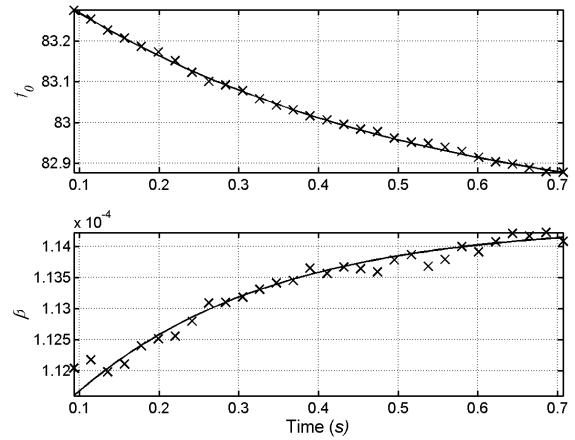


Figure 7: $f_0(t)$ (upper plot) and $\beta(t)$ (lower plot) fits in acoustic guitar mezzo-forte open E3 measurements

Now we shall observe the formulation of our dynamic spectrum, and see how it corresponds to the partial frequency measurements performed in the preliminary sliding spectral analysis. The tracks are drawn substituting (5) and (10) into (2). Figure 8 exemplifies the tracks thus obtained from the fits shown in Figure 7. The partials shown in the upper subplot are partials 40 to 43, and in the lower subplot, 80 to 83.

To give a better global impression of the accuracy of the modeled spectrum, we now formulate the deviation of the actual partial measurements from the modeled tracks. The deviation should be expressed in terms of the frequency difference between consecutive partials. For a static, harmonic spectrum, it can simply be expressed as a fraction of the fundamental frequency, which is the frequency interval that separates one harmonic from

another. Here, however, we have a more complex situation, in that instead of being a constant, f_0 , our partial difference has become a series function of time, $\delta_k(t)$:

$$\delta_k(t) = kf_0(t)\sqrt{1 + \beta(t)k^2} - (k-1)f_0(t)\sqrt{1 + \beta(t)(k-1)^2}. \quad (30)$$

Now we formulate the deviation itself. The frequency of the k^{th} partial measured at time t_n of index n is denoted by $f_{k,n}$, while the frequency of the modeled partial at that same time is denoted by $f_k(t_n)$. The deviation follows as $|f_{k,n} - f_k(t_n)|$, and we express it as a fraction $\varepsilon_k(t_n)$ of the corresponding partial difference $\delta_k(t_n)$, i.e.

$$\varepsilon_k(t_n)\delta_k(t_n) = |f_{k,n} - f_k(t_n)|. \quad (31)$$

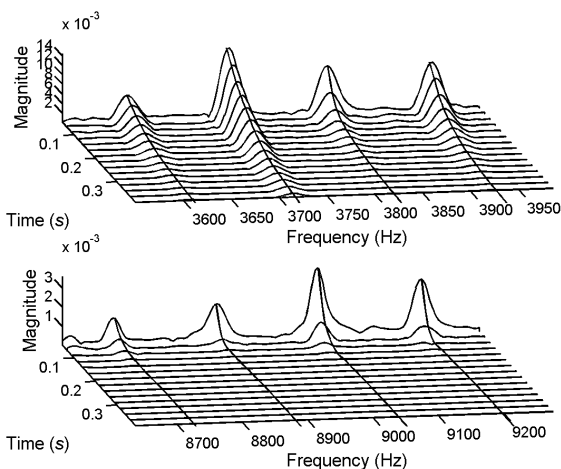


Figure 8: Spectral model tracks on open E3 spectrogram, partials 40 to 43 (upper plot), and 80 to 83 (lower plot).

Still on the same acoustic guitar E3 example, Figure 9 shows the maximal deviation $\varepsilon_k(t_n)$ along the 30 time indexes, for each measured partial. It is encouraging to see that the deviation never exceeds half of the partial difference, which facilitates greatly such a process as partial tracking.

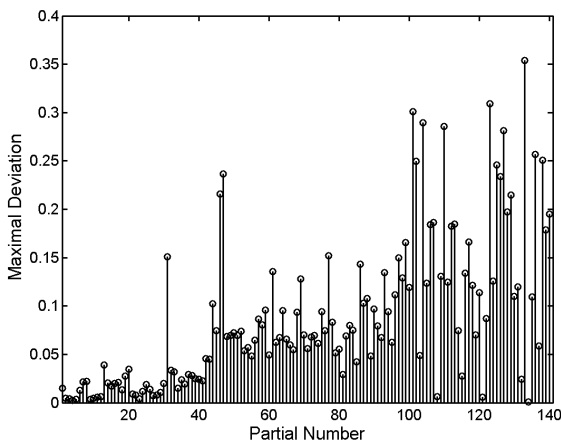


Figure 9: Maximal deviation of measured partials from modeled partial tracks.

To complete the overview of the model, Figure 10 shows a match of the partial frequency measurements $f_{k,n}$ (crosses) onto

the modeled partial tracks $f_k(t_n)$ (solid lines), for $k = [20 \ 21]$ (lower plot), and $k = [80 \ 81]$ (upper plot). The dashed lines are representative of the maximal deviation ε_{max} (about 0.35 in Figure 9) found across all partials and all indexes of time: $\varepsilon_{k,\text{max}} = \max[\varepsilon_{k,n}]$ for $n = 1, 2, \dots, W$, and $\varepsilon_{\text{max}} = \max[\varepsilon_{k,\text{max}}]$. No partial deviates beyond those lines. Figure 10 also brings about an important point, regarding the trajectories of the partials. At low partial indexes, the effect of an exponentially decaying fundamental frequency predominates over the effect of the time-varying inharmonicity, and the partial trajectories themselves look as if they simply are exponentials decaying towards a constant. However, this simplification does not hold at high partial indexes anymore, where the changing inharmonicity may actually cause the partials to rise.

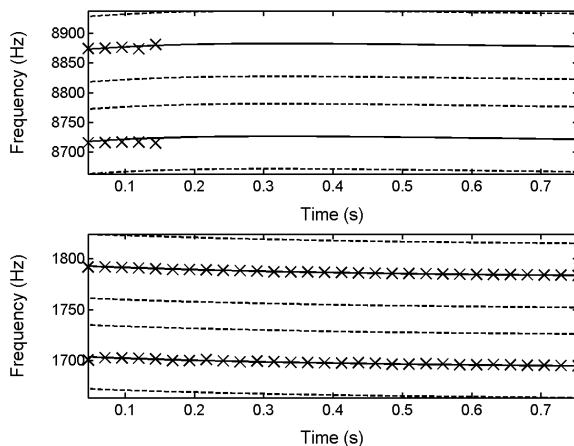


Figure 10: Partial tracks (solid) with deviation bands (dashed), for partials 20 and 21 (lower plot), and partials 80 and 81 (upper plot).

6. CONCLUSION AND FUTURE WORK

The E3 subject tone used in the results presented above was issued from the *Ovation Piezo Guitar* sound bank of the Yellow Tools Independence sampler. More measurements should be performed on acoustic and electric guitar tones for which the string number and fret position is known so as to map the parameters of (5) and (10) across the neck of the instrument – the fret position matters in that the inharmonicity coefficient is inversely proportional to the square of the length of the vibrating string segment, as shown in (4). A trend in the parameters of (5) and (10) depending on those criteria could emerge and be used in a synthetic model. For the latter to be user-friendly, arranging the model's parameters to a reduced, more intuitive set of parameters could be desirable.

More excitingly, the upper plot in Figure 10 illustrates the fact that the model traces the frequency evolution of partials long after those have died away. In fact, those tracks could be carried until the moment in time the fundamental frequency and inharmonicity coefficient have stabilised, and beyond. In Spectral Model Synthesis, this could be the ground for a re-synthesis effect where sustain of the partials is arbitrarily extended – to be distinguished with the time-stretch effect, where the frequency tracks themselves are compressed or lengthened.

Finally, and this is how it was inspired, the model can readily be used for partial tracking. As robust as they are to find and put numbers on pseudo-harmonics, Median-Adjustive Trajectories on the one hand require long window sizes, which compromises

the phase estimates of the DFT analyses – all-important in certain applications – and on the other hand, their computational cost is significantly larger than the cost of Fast-Fourier Transforms alone. Now relatively few MAT analyses can be performed to construct the model, and a tighter Short-Time Fourier Transform can thereafter be run, where the modelled tracks are used as initial guesses in the search and matching of frequency component peaks.

REFERENCES

- [1] J.O. Smith III and X. Serra, “PARSHL, An Analysis/Synthesis Program for Non-Harmonic Sounds Based on a Sinusoidal Representation”, *Proceedings of the International Computer Music Conference (ICMC-87)*, Tokyo, 1987.
- [2] H. Fletcher, E.D. Blackham and R. Stratton, “Quality of Piano Tones”, *Journal of the Acoustical Society of America*, vol. 36, no. 6, pp. 749-761, 1962.
- [3] K.A. Legge and N.H. Fletcher, “Nonlinear generation of missing modes on a vibrating string”, *Journal of the Acoustical Society of America*, vol. 76, no. 1, pp. 5-12, July 1984.
- [4] B. Bank, “Energy-Based Synthesis of Tension Modulation in Strings”, *Proceedings of the 12th Int. Conference on Digital Audio Effects (DAFx-09)*, Como, Italy, September 1-4, 2009.
- [5] M. Hodgkinson, J. Wang, J. Timoney and V. Lazzarini, “Handling Inharmonic Series with Median-Adjustive Trajectories”, *Proceedings of the 12th Int. Conference on Digital Audio Effects (DAFx-09)*, Como, Italy, September 1-4, 2009.
- [6] J. Rauhala, H.M. Lehtonen and V. Välimäki, “Fast automatic inharmonicity estimation algorithm”, *Journal of the Acoustical Society of America*, vol. 121, no. 5, pp. EL184-EL189, 2007.
- [7] K. M. Short and R. A. Garcia, “Signal Analysis using the Complex Spectral Phase Evolution (CSPE) Method”, *Audio Engineering Society 120th Convention*, May 2006, Paris, France. Paper Number: 6645.