PREPARED PIANO SOUND SYNTHESIS

Stefan Bilbao
University of Edinburgh
United Kingdom
sbilbao@staffmail.ed.ac.uk

John ffitch
Department of Computer Science
University of Bath
United Kingdom
jpff@codemist.co.uk

ABSTRACT

A sound synthesis algorithm which simulates and extends the behaviour of the acoustic prepared piano is presented. The algorithm is based on a finite difference approximation to multiple stiff string vibration, including an excitation method (a hammer) as well as several connected preparation elements, modeled as lumped nonlinearities. Numerical issues and implementation details are discussed, and sound examples are presented.

1. INTRODUCTION

The use of physical modeling principles in sound synthesis leads, eventually, to certain thorny questions. By their very nature, such methods allow very high-fidelity modeling of instrument timbres—and yet, by increasingly close attention to detail in the physics of a musical instrument, one can obtain a result which is, for musical purposes, useless: in the limiting case, an exact reproduction of the sound of an existing musical instrument. Though such sound synthesis methods are undeniably extremely important for purposes of checking the validity of a given musical instrument model, and for commercial sound synthesis applications, the focus here is on the creation of sounds which retain some of the natural texture of those produced by acoustic instruments, but which are, in some sense, new. In this regard, the use of physical models, though admirably suited to the task, is something of a balancing act.

One approach is to begin from a well-defined musical instrument, and to perform variations on its structure. This is essentially what was done early on, in the mechanical setting, in the case of prepared piano (probably best known through the work of John Cage [1], but still in extensive use up to the present day). Objects such as screws, rubber erasers, and washers (among others) are placed in proximity to the strings of a piano, yielding timbres sometimes quite foreign to the familiar piano tone—strongly percussive and bell-like tones may be generated in this way. Such an approach can be easily translated to physical modeling sound synthesis approaches, and leaves open the door to various types of preparation which are unrealizable in an actual piano (at least, not realizable without causing permanent damage to the instrument). In this study, string vibration will be simulated using finite difference schemes [2, 3, 4, 5], and various connecting or exciting elements, such as a standard hammer [2], and abstractions of preparing objects, are dealt with as lumped objects (though this is not strictly necessary). This approach is reminiscent of the massspring network type models used in the past [6], but the use of a fully distributed model of the string is far less unwieldy. Digital waveguides [7, 8] could also potentially be employed, but their efficiency advantage is substantially reduced under stiff string conditions, and when a string is broken up by various elements in contact with it (i.e., with a scattering junction required at each such connecting points, as well as the accompanying fractional delay machinery [2]). Another more important reason for making use of a direct simulation approach such as finite differences is that any results can easily be extended to deal with systems far more complex than simply strings.

In Section 2, a model of the stiff string, suitable for piano sound synthesis is presented, followed in Section 3 by definitions of the hammer interaction and various preparing elements. In Section 4, a finite difference scheme for multiple strings is given, including a discussion of boundary termination and numerical stability. Implementation details, particularly with regard to the csound music programming language [10] follows in Section 5. Numerical results appear in Section 6.

2. A STIFF STRING MODEL

A useful general purposed model of string vibration may be written as [3]:

\[ \dddot{u} = c^2 \dddot{u}'' - \kappa^2 \dddot{u}'''' - \sigma \dddot{u} + b \dddot{u}'' \]  

Here, \( u(x, t) \) represents transverse string displacement in a single polarization, as a function of time \( t \geq 0 \) and \( x \in [0, 1] \). Dots and primes indicate differentiation with respect to time and space, respectively. The first term on the right-hand side is the usual wave equation term, and the parameter \( c \) is related to the fundamental frequency \( f_0 \) of the string by \( f_0 = c/2 \). (Note that in [1], for simplicity, the spatial variable has been non-dimensionalized, so that \( c \) does not have dimensions of velocity.) The second term represents a contribution due to string stiffness, leading to inharmonicity of partials, parameterized by the constant \( \kappa \); in the limit as \( \kappa \) becomes large, Eq. (1) represents the behaviour of a vibrating bar. The final two terms model loss: the first, parameterized by \( \sigma \geq 0 \), describes the global decay rate of the string, and the second, parameterized by \( b \geq 0 \), approximates frequency-dependent loss characteristic of string vibration.

There are various boundary conditions which can be used to terminate the string equation (1) at either end point \( x = 0 \) or \( x = 1 \); two are required at any terminating point. In the case of strings, perhaps most useful are fixed end conditions of the clamped, i.e.,

\[ u = u' = 0 \]  

or simply supported type, i.e.,

\[ u = u'' = 0 \]  

Under low-stiffness conditions, the audible difference between the two types of conditions is negligible. If the equation (1) is intended
to represent the behaviour of a bar, then free boundary conditions may be implemented as well.

Eq. (1), though arguably suitable for piano sound synthesis, is by no means the most general musically useful string model; various nonlinear models, of varying degrees of complexity, allow rendering of more interesting effects such as pitch glides among others [12], [13], [14]. Any such model could easily be substituted here, with very little effect on the resulting development in this article (particularly with regard to the construction of finite difference schemes, which handle nonlinear distributed systems quite easily).

2.1. Multiple Strings and Excitation

The character of a single piano note is determined by the vibration of several strings, usually very similar in fundamental frequency, and identical in other respects. In the implementation described here, $N_s$ such strings are coupled, each described by an equation of the form

$$
\ddot{u}_{(q)} = c_{(q)} u_{(q)}'' - \kappa u_{(q)}''' - \sigma \dot{u}_{(q)} + b \dot{u}_{(q)} + \delta(x_0) \phi(q) \quad (4)
$$

where here $q$ is an index which runs from $q = 1$ to $q = N_s$; notice that of the defining parameters, only $c(q)$ has any dependence on $q$. Also, a set of excitations $\phi(q)$, acting at $x = x_0$ on each string (through the Dirac delta function $\delta(x_0)$) has been introduced. The excitations represent the behaviour of a coupling element, such as a piano hammer, or one of various types of string preparation, which will be discussed shortly. This description of excitation or coupling is a lumped one—it is by no means necessary to make this assumption, but for many musically meaningful cases, the spatial extent of the excitation is smaller than the shortest audible wavelength in the string, for physically reasonable material sound speeds. See Section 3.5.3 for more discussion of this point.

3. EXCITATION AND PREPARATION

In this section, the hammer interaction, as well as various interactions with preparing elements are described; in each case, the element interacts with the entire set of strings. Though single elements are described here, it is simple to generalize to the case of connections to many of these objects simultaneously.

3.1. Hammers

A simple, but effective model of a hammer strike upon multiple strings from below [2] is easily described in terms of the displacement $u_H(t)$ of the hammer, and that of an individual string $q$ at the strike location $x_0$, $u_q(x_0, t)$, or rather the difference $w_q(t) = u_H(t) - u_q(x_0, t)$:

$$
\phi(q)(w_q) = \begin{cases} 
\mu r^2 w_q^\alpha, & w_q > 0 \\
0, & w_q \leq 0
\end{cases}
$$

In other words, a nonlinear force is exerted when the hammer and string are in contact, with the nonlinearity controlled by the exponent $\alpha$ (a value between 1 and 3 is generally chosen). The parameter $\mu \geq 0$ represents the relative mass density of the hammer relative to the string, and $\gamma$ is a parameter which scales the stiffness of the hammer. The hammer position itself is governed by Newton’s laws, i.e.,

$$
\ddot{u}_H = -\frac{1}{\mu} \sum_{q=1}^{N_s} \phi(q) \quad (5)
$$

The hammer itself must be initialized, typically with an initial position $u_H(0)$ and an initial velocity $\dot{u}_H(0)$.

3.2. Traps

One type of string preparation is to vertically “trap” it, at a given point $x_0$, using nonlinear springs. In this case, the functions $\phi(q)$ are given by

$$
\phi(q)(u_q(x_0, t)) = -\gamma^2 (u_q(x_0, t))^\alpha \quad (6)
$$

where here, $\gamma$ controls the spring stiffness and $\alpha$ is the nonlinearity exponent.

3.3. Rubber Stoppers

In order to model a rubber stopper (such as an eraser) wedged underneath a set of strings, the following model is a useful first approximation. If $u_E(t)$ is the position of the top of the rubber element, then in terms of $w_q(t) = u_E(t) - u_q(x_0, t)$, one has, similarly to the case of the hammer,

$$
\phi(q)(w_q) = \begin{cases} 
\mu r^2 w_q^\alpha, & w_q > 0 \\
0, & w_q \leq 0
\end{cases}
$$

In contrast to the hammer, however, the rubber stopper is anchored, and its dynamics may be written as

$$
\ddot{u}_E = -\sigma E \ddot{u}_E - \gamma^2 \left( u_E + \frac{1}{\mu r^2} \sum_{q=1}^{N_s} \phi(q) \right) \quad (7)
$$

where $\sigma E$ is a damping constant (typically quite large), and $\gamma$, in this case, can be viewed as the fundamental frequency of the rubber element. $\mu$, in $\phi$, is again the mass density ratio of the rubber element to the string.

An even simpler model of such a stopper involves the use of a simple damping term, i.e.,

$$
\phi(q)(u_q(x_0, t)) = -\sigma E \dot{u}_E(x_0, t) \quad (8)
$$

which may be used in conjunction with the trap discussed above.

3.4. Rattling Elements

A rattling element may be modeled, in the simplest case, as a two-point mass “dumbbell,” of length $\epsilon$, on which a force acts only when one of the two masses is in contact with the string. In this case, the coupling may be framed in terms of the displacement $w_q(t) = u_R(t) - u_q(x_0, t)$, where $u_R$ is the vertical midpoint position of the rattling element. The force relationship is given by

$$
\phi(q)(w_q) = \begin{cases} 
\mu r^2 (w_q - \epsilon/2), & w_q > \epsilon/2 \\
0, & -\epsilon/2 \leq w_q \leq \epsilon/2 \\
\mu r^2 (w_q + \epsilon/2), & w_q < -\epsilon/2
\end{cases}
$$

The contact interaction here is piecewise linear, but could easily be made more strongly nonlinear through the introduction of an
exponent, as in the case of the hammer. The rattle position is again governed by Newton’s laws, i.e.,

$$\ddot{u}_R = -\frac{1}{\mu} \sum_{q=1}^{N_s} \phi(q) - g \quad (9)$$

where here, gravitational acceleration $g$ has been introduced, as it will have particular importance in determining the collision dynamics of the rattle. Again, $\gamma$ represents the fundamental frequency of the rattling elements, and $\mu$ in (4) the relative linear mass density of the rattle to the string.

4. A FINITE DIFFERENCE MODEL

For one-dimensional systems such as strings, finite difference schemes are a good choice of numerical method; though not as efficient as digital waveguides [7], they are simple to program, much more general, and allow a complete transparency of the entire state of the simulated system, which is especially useful if one is interested in, say, dynamically changing the position of an excitation or preparing element. As mentioned earlier, they extend easily to many other methods are presented here. In the case of the hammer, (5) may be initialized at zero. It is also clear that any given grid function $u_{n,o},t$ which is defined over a set of grid points indexed by $t = 0, \ldots, N$, where $N = 1/h$, will require access to values outside the problem domain at points near the boundary, according to (10).

For fixed termination, as is the case presented here, one may always set $u_{n,o}, N=0$; in other words, the terminal grid points are permanently set to zero, and need not be computed, satisfying the first of conditions [8] or [9]. To satisfy the second of conditions [8] at, say, the left end of the string, one may set $u_{n,1} = 0$. To satisfy the second of conditions [8], one may set $u_{n,1} = -u_{n}$. Many other methods of setting the boundary conditions are feasible, and details may be found elsewhere; these particular settings are simple, and provably numerically stable.

4.1. Updating for the Connecting Elements

The various connecting elements described in Section[3] are also, for the most part, described by differential equations, and require numerical integration. Though in the case of nonlinear elements, there are obviously many ways of performing this integration, simple methods are presented here. In the case of the hammer, [5] may be updated as

$$u_{n+1}^H = 2u_{n}^H - u_{n-1}^H - \frac{k^2}{\mu} \sum_{q=1}^{N_s} \phi(q)$$

The rattling element, defined by [2] may be dealt with similarly as

$$u_{n+1}^R = 2u_{n}^R - u_{n-1}^R - \frac{k^2}{\mu} \sum_{q=1}^{N_s} \phi(q) - k^2 g$$

For the rubber stopper, one may use

$$u_{n+1}^E = 2u_{n}^E - u_{n-1}^E - \sigma_E k(u_{n}^E - u_{n-1}^E) - \gamma^2 k^2 \left( u_{n}^E + \sum_{q=1}^{N_s} \phi(q) \right)$$

In other words, for a given sample rate and set of material parameters, a minimum spacing between adjacent points must be respected. This spacing relates to the highest audible wavelength in the string. A slight additional complication comes from the fact that the scheme for each string will possess its own separate stability condition, giving rise, potentially, to different grid spacings on each string, which is inconvenient from a programming standpoint; it is useful to make use of a global spacing $h$ over all strings, bounded by

$$h \geq \max_q h_{(q),\text{min}} \quad (11)$$

Use of such a global spacing is advisable only when the fundamental frequencies of the strings are relatively close; otherwise, one may be far from the ideal bound in at least some of the strings, and numerical dispersion [15] will play a significant role, leading to a numerical inharmonicity in partials.

Eq. (10) is implemented as a two-step recursion; it may be initialized as

$$h_{(q),0}^n = 0$$

where here, $h_{(q),n}$ approximates $\phi(q)$ at $t = nk$, and at grid location $i = i_0$ corresponding to $x_n = i_0 h_{(q)}$ (hence the use of the Kronecker delta). $h_{(q)}$ is the spacing between adjacent grid points, and $k$ is the time step (1/k is the sample rate, generally set a priori). The scheme coefficients $\alpha_{(q),i}$ and $\beta_{(q),i}$ are given by

$$\alpha_{(q),0} = \frac{2 - \frac{c^2(q) h^2}{h^2_{(q)}} - \frac{a(n^2 - 1)}{h^2_{(q)}} - \frac{2bk}{h^2_{(q)}}}{1 + \sigma k/2}$$

$$\alpha_{(q),1} = \frac{\frac{c^2(q) k^2}{h^2_{(q)}} + \frac{a(n^2 - 1)}{h^2_{(q)}} + \frac{bk}{h^2_{(q)}}}{1 + \sigma k/2}$$

$$\alpha_{(q),2} = \frac{-\frac{k^2}{h^2_{(q)}}}{1 + \sigma k/2}$$

$$\beta_{(q),0} = \frac{-\frac{bk}{h^2_{(q)}}}{1 + \sigma k/2}$$

$$\beta_{(q),1} = \frac{\frac{k^2}{h^2_{(q)}}}{1 + \sigma k/2}$$

A necessary stability condition, which follows from von Neumann (Fourier) analysis of scheme (10) under zero input conditions [15], is

$$h_{(q)} \geq h_{(q),\text{min}}$$

with

$$h_{(q),\text{min}} = \sqrt{\frac{c^2(q) k^2 + 2bk + \sqrt{(c^2(q) k^2 + 2bk)^2 + 16n^2 k^2}}{2}}$$

DAFX-79
The trapping element requires no additional state, and to approximate one may simply use
\[ \phi(q)(u(q),i_0) = -\gamma^2(u(q),i_0) \]

If loss is introduced, one may use, as an approximation to \( q \),
\[ \phi(q)(u(q),i_0) = -\sigma E \left( u^{n+1}(q),i_0 - u^n(q),i_0 \right) \]

5. IMPLEMENTATION DETAILS

5.1. Operation Count

The bulk of the calculation in this prepared piano simulator is carried out in performing the update of the string state; hammers and various connections are of negligible computational cost. For a single string, of fundamental frequency \( f_0 \), and under low inharmonicity conditions, at sample rate \( f_s \), the number of multiplies per second is approximately \( 2 f_s / f_0 \); the computational cost obviously scales inversely with the fundamental frequency.

5.2. csound Implementation

This algorithm was prototyped in the Matlab programming language, but the aim is to make the sound available to composers. Ideally such an implementation should be playable in real time, or at least close to real time, and from the point of view of practicality there needs to be some attempt to minimise the memory use. For composers, the flexibility of the operation is also significant; csound would appear to be a good choice of programming environment.

The internal operations naturally rely on three arrays to represent the values \( u^{n-1}(q),i \), \( u^n(q),i \), and \( u^{n+1}(q),i \) for each string. At each time step these are cycled by pointer assignment, taking care to correct for the boundary conditions. Similar arrangements are taken to maintain the state of the preparations.

In Csound, as in all MusicV descended systems, the operations are split between an initialisation method and then a perform method that creates a short section of audio. Much of the speed of Csound comes from this batching of operations, with its attendant good cache behaviour and the less frequent adjustment of parameters. In the case of the prepared piano there are twelve fundamental parameters covering basics like fundamental frequency and the number of strings, as well as stiffness, loss parameters and the characteristics of the hammer, most of which cannot reasonably be adjusted after the initialisation (although there is a future action to investigate this more closely). The harder problem is how to specify the various preparations. Our implementation uses two \( f \)-tables, one to define the characteristics of rubber preparations and one to do likewise for the rattles. Either of these can of course be omitted.

The other design question is how to handle multiple hammer strokes. Our implementation does this by having one stroke for each call, but it is possible to skip the initialisation phase, and so keep the previous state of the system. There is also an issue on control of the amplitude of the output. As in many physical models the parameters determine the inputs rather than the resulting amplitude. The code performs some scaling based on experience.

5.3. Performance

The csound opcode has been tested in situ in two simple cases, with one rattle and one rubber stopper, and with no preparations. Clearly the processor speed is important when we are concerned with how close to real time the code runs. The tests were run on two widely differing systems both running Csound5.01; a 1.4GHz Centrino laptop with Linux with floating-point samples, and an AMD Athlon64 3400+ (2.4GHz) with Linux, and double-precision samples. The actual test program is shown in Figure 1 where a very low frequency string \( f_0 = 60 \text{ Hz} \) has been chosen, as a "worst case."

<table>
<thead>
<tr>
<th>Machine</th>
<th>Time</th>
<th>% real</th>
<th>Time</th>
<th>% real</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laptop</td>
<td>8.143</td>
<td>47%</td>
<td>8.239</td>
<td>46%</td>
</tr>
<tr>
<td>Workstation</td>
<td>0.591</td>
<td>643%</td>
<td>0.595</td>
<td>639%</td>
</tr>
</tbody>
</table>

The preliminary timings suggest that the cost of the preparations is small. The workstation is clearly showing usable performance while on the laptop the speed is too slow for real time, but not too bad for some uses. These results are preliminary, and there may be opportunities for further optimisation, such as moving some calculations from the sample-rate loop to the control rate.

6. NUMERICAL RESULTS

In this section, several illustrative plots of numerical output from the string model, under various types of preparation, are presented; all were created at the audio sample rate of 44.1 kHz. The effect of brightening with strike velocity for a nonlinear hammer of the type presented here has been discussed elsewhere. Considering the case of a rattling element, in Figure 2 are shown output waveforms generated, as well as spectrograms, in the case of a struck string at fundamental frequency 600 Hz, both with and without a rattling element present. Though it is difficult to show the precise time-dependent effects of the rattling element, at the crudest level, one sees a considerable disruption of the output waveform, as well as a substantial presence of high frequencies in the case of the rattle.
Figure 2: Top: output waveform and spectrogram, for a struck piano tone at 600 Hz. Bottom: output waveform and spectrogram for a struck piano tone at 600 Hz, including the effects of a rattling element placed 0.3 of the way along the string from the left end.

Consider the case of a trapping element, in the case of a string of fundamental frequency 200 Hz, a spectrogram of the output waveform is given in Figure 3. The string is struck successively with blows of initial velocity 5, 10, 50 and 100 m/s. Notice in particular the increase in high frequency energy with strike velocity (due to the effect of both the nonlinear hammer interaction, and trapping element), as well as the rather complex effects of modulation visible in the upper harmonics of the tone under high amplitude striking velocity.

Figure 3: Top: output spectrogram, for a struck piano tone at 200 Hz, with a trap placed 0.3 of the way along the string from the left end, under strikes of increasing amplitude.

Under the same conditions as the example above, but with the addition of a rubber stopper (in this case at the same location as the trapping element), the output sounds become, predictably, shorter and more percussive (see Figure 4) with the additional feature that, under high amplitude conditions, a new, higher fundamental frequency emerges. The effect of a prepared piano sounding at a pitch different from that of the note struck is, of course, a well-known effect.

It is rather difficult to give the full flavor of the sound examples produced by presenting plots of waveforms and spectrograms; the reader is directed to a series of sound examples, created in Matlab, which are available on the internet [16].

7. CONCLUSION

In this short article, a brute force numerical simulation of a prepared piano has been presented. As noted in Section 5, though more computationally demanding than techniques such as digital waveguides, these methods are already capable of real time performance.

One significant problem with the algorithm as presented here has to do with numerical stability concerns, with regard to the nonlinear interactions with the hammer and various preparing elements. The condition [17], which is derived using Fourier principles applied to the string scheme under linear conditions is merely necessary. One may develop energy-conserving or dissipating algorithms to deal with the various nonlinearities (indeed, in all cases, the model, including preparing elements, may be shown to be strictly dissipative) [17], but such algorithms may lead to implicit schemes, necessarily requiring the use of iterative methods, without a guarantee of convergence. Another approach would be to make use of so-called symplectic methods [18], but this approach is not strictly applicable to problems involving dissipation. The question is open as to how to construct robust (i.e., provably numerically stable) numerical methods for complex nonlinear systems such as that presented here.

It is worth noting that although the algorithm can be used to simulate the behaviour of a piano-like instrument, the model equation, as defined by [1] is quite general, and can be used, without modification to simulate the behaviour of various other percussion instruments, including bar-based instruments such as xylophones or marimbas, again under prepared conditions. One element which is missing here is a physical model of a soundboard, which can be easily modeled through the use of a finite difference plate model, connected to the string models mentioned here. Preparation of such a soundboard itself is, of course, a desirable option.

8. ACKNOWLEDGEMENTS

This work was supported in part by the Engineering and Physical Sciences Research Council UK, under grant number C007328/1.
9. REFERENCES


