

ANALYSIS/SYNTHESIS OF QUASI-HARMONIC SOUNDS BY USE OF THE KARPLUS-STRONG ALGORITHM

J. Laroche, J. Jot

► To cite this version:

J. Laroche, J. Jot. ANALYSIS/SYNTHESIS OF QUASI-HARMONIC SOUNDS BY USE OF THE KARPLUS-STRONG ALGORITHM. Journal de Physique IV Colloque, 1992, 02 (C1), pp.C1-117-C1-120. <10.1051/jp4:1992122>. <jpa-00251187>

HAL Id: jpa-00251187

<https://hal.archives-ouvertes.fr/jpa-00251187>

Submitted on 1 Jan 1992

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ANALYSIS/SYNTHESIS OF QUASI-HARMONIC SOUNDS BY USE OF THE KARPLUS-STRONG ALGORITHM

J. LAROCHE and J.M. JOT

Télécom Paris, 46 rue Barrault, F-75634 Paris cedex 13, France

Abstract

In the Karplus-Strong method, the synthetic sound is obtained by exciting a recursive comb-filter with a burst of noise. An analysis scheme is presented in this paper, which makes it possible to derive the optimal Karplus-Strong synthesis parameters from the recording of a real instrument. In addition, an excitation signal can be calculated and used in place of the usual burst of white or colored noise, in order to obtain more realistic syntheses.

A Introduction

The Karplus-Strong synthesis algorithm [1, 2] offers a number of advantages: it produces realistic harmonic or quasi-harmonic synthetic sounds at a very low computational cost. The algorithm's simplicity makes its real-time implementation quite easy on any digital-signal-processing microprocessor. Finally, as described in [2], the synthesis can be controlled through a small number of parameters. The synthetic sound is obtained by exciting a resonating filter (a recursive comb-filter) with a short noise burst. Although the synthesis technique was thoroughly investigated by Smith and Jaffe, the synthesis parameters and the excitation signal were usually chosen in an ad-hoc way, without reference to the analysis of real instrumental sounds. The purpose of this paper is to present an analysis technique corresponding to the Karplus-Strong synthesis algorithm. Given a real quasi-harmonic sound, we will show that it is possible to estimate the synthesis parameters as well as the excitation signal that best reproduce the original sound. The Karplus/Strong algorithm will be briefly presented in the first part. Then we will turn to the estimation of the analysis parameters and of the excitation signal. Before concluding, we will give an analysis example.

B The Karplus-Strong synthesis model

The Karplus-Strong synthesis model was first presented in [1] then re-worked and improved in [2]. The model involves a simple m -delay feedback-loop in which a filter $H(z)$ is inserted, as shown in fig. 1. The

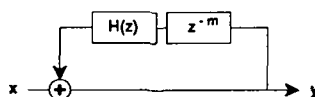


Figure 1: The Karplus Strong synthesis model

condition $|H(z)| < 1$ for any complex z satisfying $|z| = 1$ is not sufficient for the Karplus-Strong model to be stable. However, if filter $H(z)$ is stable and satisfies the preceding condition, then the Karplus-Strong

model can be shown to be stable.

The eigenfrequencies f_k of the model are determined by the time-lag m and by the phase transfer-function of filter $H(z)$. The damping factors α_k of the modes depend on the magnitude transfer-function of filter $H(z)$. When filter $H(z)$ is a frequency-independent real gain, the eigenfrequencies and their damping factors can be calculated explicitly. In the general case however, one must use approximate formulas [3]:

$$f_k = \frac{k}{m} + \frac{\text{Arg}(H(z_k))}{2\pi m} \quad \text{and} \quad \alpha_k = \frac{\log(|H(z_k)|)}{m} + \frac{\text{Arg}(H(z_k))}{2\pi f_k m} \quad \text{where} \quad z_k = e^{2i\pi f_k} \quad (1)$$

A single eigenfrequency produces a damped sinusoidal output of the form:

$$y_n = A_k e^{\alpha_k n} \cos(2\pi f_k n + \varphi_k)$$

where the initial amplitude A_k and phase φ_k depend on the excitation. According to eq. (1), the effect of the phase transfer-function of filter $H(z)$ is to modify the length of the feedback loop, and consequently the values of the eigenfrequencies. The magnitude transfer-function controls the amount of feedback at the various frequencies and in particular the damping factors of the eigenfrequencies.

When filter $H(z)$ is a frequency-independent real gain, the eigenfrequencies are harmonics of a fundamental frequency given by $f_1 = 1/m$. As a consequence, only sounds whose pitch divides the sampling rate can be synthesized that way. In the general case however, one needs to synthesize sounds with arbitrary pitches. To achieve this, one can use eq. (1) to set the phase transfer-function of filter $H(z)$ so that the fundamental frequency becomes arbitrary. Jaffe and Smith showed that a simple first-order all-pass filter was sufficient for that purpose. In addition, the phase behavior of filter $H(z)$ can be used to reproduce the inharmonicity of real sounds.

In the following section, we will show that it is possible to analyze a real, quasi-harmonic sound to derive the Karplus-Strong model that best reproduces the original sound.

C Determining the Karplus-Strong model for a given sound

In this section, we want to reproduce an original quasi-harmonic sound using the Karplus-Strong synthesis algorithm. The parameters to determine are the integer delay m and the filter $H(z)$. When this is done, the excitation signal can be extracted by use of inverse filtering.

1 Determining the delay m and the filter $H(z)$

It is first necessary to estimate the frequencies and the damping factors of the sinusoids composing the original sound. This can be done by use of the technique described in [4]: a cumulated time-varying spectrum is calculated, then a peak extraction step yields estimated values of the frequencies f_k . The damping factors α_k are calculated by linear-regression on a dB scale. If the original sound has a strong perceptual pitch (e.g. guitars, pianos, etc...) the frequencies form a near-harmonic series in which f_k is the k^{th} harmonic. The delay parameter m can be set to:

$$m = \text{floor}\left(\min_k \frac{k}{f_k}\right) \quad \text{with} \quad f_1 < f_2 < \dots < f_L \quad (2)$$

in which $\text{floor}(x)$ represents the closest integer below x , and L is the number of detected harmonics. With a constant filter $H(z)$ and the value of parameter m defined above, a Karplus-Strong model would produce a harmonic sound with fundamental $1/m$. We now use eq. (1) to fine-tune each harmonic frequency: the phase transfer-function of filter $H(z)$ is defined at L different frequencies f_k by the equations:

$$\text{Arg}(H(z_k)) = 2\pi(m \cdot f_k - k) \quad \text{where} \quad z_k = e^{2i\pi f_k} \quad \text{and} \quad k \in [1 L] \quad (3)$$

Note that with our choice of m , $f_k/k < 1/m$ for each k , and $\text{Arg}(H(z_k))$ is negative. This property will be seen to make the estimation of filter $H(z)$ easier. The damping factor of each harmonic is adjusted according to eq. (1):

$$\log |H(z_k)| = m\alpha_k - \frac{\text{Arg}(H(z_k))}{2\pi f_k} = m\alpha_k - \left(m - \frac{k}{f_k}\right) \quad (4)$$

The two equations (3) and (4) define filter $H(z)$ at L different points on the unit circle, in both modulus and phase. Our problem is now to find a stable filter $H(z)$ that satisfies those two equations.

A standard form for filter $H(z)$ is $H(z) = B(z)/A(z)$. Smith, in [5], proposes a number of techniques to determine $B(z)$ and $A(z)$ from the values of $\text{Arg}(H(z_k))$ and $|H(z_k)|$. Two of these techniques will be discussed here: the equation-error method and the Hankel-norm approximation.

1. The equation-error method is based on the resolution of a set of linear equations and therefore is quite easy to implement. Furthermore, if $n_A + n_B + 1 \geq L$ and n_A and n_B are the degrees of polynomials $A(z)$ and $B(z)$, the values at the frequencies f_k of the magnitude and of the phase of filter $H(z)$ are *exactly* those specified by eqs. (3) and (4). The drawback of the equation-error method is that the filter $H(z)$ is not guaranteed to be stable. With our choice of the delay m (eq. 2), the filters obtained in practice are often unstable. To reduce the extent of the problem, it is possible to choose a smaller value of the delay m . This operation does not guarantee that the new filter will be stable, but the "likelihood" of obtaining an unstable filter is reduced [5].
2. The Hankel-norm approximation involves the eigenvalue decomposition of a Hankel matrix and makes use of the Prony algorithm to obtain the coefficients of filters $A(z)$ and $B(z)$. Unlike the equation-error method, the Hankel-norm approximation always yields a stable filter. On the other hand, the values at the frequencies f_k of the magnitude and of the phase of filter $H(z)$ are not necessarily exactly those specified by eqs. (3) and (4). See [5] for more details about the Hankel-norm approximation.

In our problem, we impose another constraint: $|H(z)| < 1 \forall z, |z| = 1$. This constraint cannot be incorporated easily into the calculation of $H(z)$ by either the equation-error method or the Hankel-norm approximation, and we will have to make sure that the filter we obtain is indeed of modulus less than 1 on the unit-circle. Interpolating the magnitude transfer-function between the points specified by the harmonic frequencies with intermediate values bounded by 1 actually increases the likelihood to obtain a filter satisfying our constraint. Moreover, such an interpolation is necessary when the Hankel-norm approximation is used.

Fig. 2 shows the filter obtained in the case of a guitar sound: the frequencies and damping factors were estimated by use of the 3-D cumulated spectrum, as described in [4]. Parameter m was derived from formula (2). The filter $H(z)$ was specified at 14 different frequencies by applying eqs. (4) and (3). The equation-error method was used to calculate the filters $B(z)$ and $A(z)$. The two figures present the magnitude transfer-function and the phase transfer-function of the resulting filter $B(z)/A(z)$. It can be seen that the approximation is quite good in the magnitude domain as well as in the phase domain. Moreover, the magnitude of $H(z)$ evaluated on the unit-circle does not exceed 1, and the corresponding Karplus-Strong model is stable.

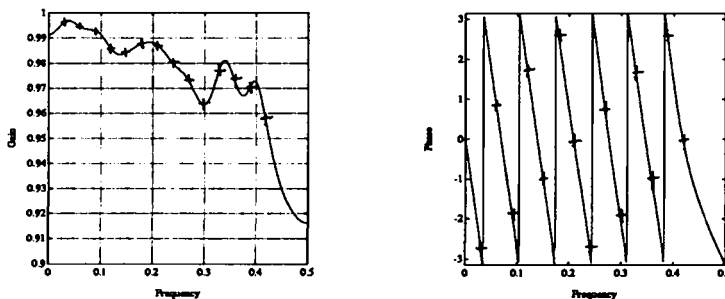


Figure 2: Left: The magnitude transfer-function of filter $H(z)$ is plotted in solid line; the crosses represent the values specified by the analysis of the original guitar sound. Right: Same results for the phase transfer-function of filter $H(z)$.

2 Calculating the excitation signal

The Karplus-Strong synthesis model is usually excited with a burst of noise. Jaffe and Smith in [2] discuss various choices of the excitation noise. We propose to derive the excitation signal from the analysis of a real sound, as was done for the filter parameters. Following [4], once the Karplus-Strong model has been calculated, we can inverse-filter the original sound to obtain the excitation signal. Since filter $H(z)$ is modeled by a rational transfer function $B(z)/A(z)$, the overall Karplus-Strong filter, corresponding to fig. 1 and its inverse filter can be re-written the following way:

$$K(z) = \frac{1}{1 - z^{-m}H(z)} \quad \text{where} \quad H(z) = \frac{A(z)}{B(z)} \quad \text{and} \quad K^{-1}(z) = 1 - z^{-m}H(z) \quad (5)$$

According to equ. (5), the stability of filter $K^{-1}(z)$ stems from the stability of $H(z)$. Consequently, the methods discussed in [4] can be applied to our problem. However, depending on filter $H(z)$, regularizing can be necessary [6] in order to retain only the pertinent information in the excitation signal. If the resonating part of the original sounds is nearly sinusoidal and if the Karplus-Strong model includes all the original frequencies, the excitation signal is generally of short duration. This turns out to be the case for a wide variety of instruments, including the guitar, the piano, the harpsichord... These instruments can therefore be synthesized by using the Karplus-Strong model in conjunction with a proper excitation signal.

D Conclusions

The methods presented in this paper for the determination of the Karplus-Strong parameters make it possible to synthesize realistic instrumental sounds at a small computation/storage cost. Although the calculation of filter $H(z)$ still presents some difficulty, (e.g. the constraint $|H(z)| < 1 \forall z, |z| = 1$) the simulation results are quite satisfactory for a wide variety of sounds. An interesting application could be the synthesis of high-quality piano sounds by means of the Karplus-Strong algorithm, an attractive alternative to the classical sampling/filtering methods.

References

- [1] K. Karplus, A. Strong, "Digital Synthesis of Plucked-String and Drum Timbres," Computer Music Journal, vol. 7 (No. 2), pp. 43-55, Summer 1983.
- [2] D. A. Jaffe, J. O. Smith, "Extensions of the Karplus-Strong Plucked-String Algorithm," Computer Music Journal, vol. 7 (No. 2), pp. 56-69, Summer 1983.
- [3] J.M. Jot, A. Chaigne, "Digital Delay Networks for Designing Artificial Reverberators," Proc. 90th A.E.S Convention, Paris 1991, preprint 3030 (E-2).
- [4] J.L. Meillier, J. Laroche, "Extracting the transient from percussive sounds," these proceedings.
- [5] J. O. Smith, "Techniques for Digital Filter Design and System Identification with Application to the Violin," Ph.D. Dissertation, Stanford University, Stanford, CA, June 1983.
- [6] G. Demoment, "Image Reconstruction and Restoration: Overview of Common Estimation Structures and Problems," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-37 (No. 12), pp. 2024-2036, Dec 1989.