WAVELETS THAT WE CAN PLAY: AN UPDATE

GIANPAOLO EVANGELISTA

Laboratoire de communications audiovisuelles
Ecole Polytechnique Fédérale de Lausanne
IN-LCAV Ecublens, CH-1015 Lausanne - Switzerland
gianpaolo.evangelista@epfl.ch

Dedicated to the memory of Aldo Piccialli

ABSTRACT. Musical signals require sophisticated time-frequency techniques for their representation. In the ideal case, each element of the representation is able to capture a distinct feature of the signal and can be attached either a perceptual or an objective meaning. Wavelet transforms constitute a remarkable advance in this field and have several advantages over Gabor expansions or short-time Fourier methods. However, application of conventional wavelet bases on musical signals produces disappointing results for at least two reasons: 1) the frequency resolution of dyadic wavelets is one-octave, too poor for any meaningful acoustic decomposition and 2) pseudoperiodicity or pitch information of voiced sounds is not exploited. Fortunately, the definition of wavelet transform can be extended in several directions, allowing for the design of bases with arbitrary frequency resolution and for adaptation to time-varying pitch characteristic in signals with harmonic or even inharmonic structure of the frequency spectrum. This paper provides an update on refined wavelet methods that are applicable to musical signal analysis and synthesis. Flexible wavelet transforms are obtained by means of frequency warping techniques.

1. INTRODUCTION

In audio signal processing the choice of the representation is crucial for deriving elegant and efficient algorithms for coding, detection and synthesis useful for new and improved applications. Undeniably, the analysis of most features of sound requires mixed time and frequency characterization. The human hearing sense is particularly gifted at classifying acoustic patterns according to both their time of occurrence and brightness of their transitory frequency spectrum. However, our mathematical models, having to cope with the uncertainty principle, cannot achieve indefinitely accurate resolution in both time and frequency domains.

Gabor expansions and the Short-Time Fourier transform (STFT) were among the first techniques to be applied to audio signals in order to analyze their time-varying frequency spectrum characteristic called *spectrogram*. However, our perception of sound is based on a non-linear frequency scale, as reflected by both cochlear functional models and psychoacoustic theories. Musical sounds were indeed classified on a tempered scale, i.e., tones spaced by one octave, corresponding to power of 2 frequency ratios, are given the same name and tones within the same octave progress

Key words and phrases. Wavelet Transform, Musical Signal Processing, Frequency Warping, Pitch-Synchronous Techniques.

as powers of $\sqrt[12]{2}$. This behavior is contrasting with the frequency resolution of the STFT, which is uniform on any portion of the frequency axis.

The introduction of wavelets and multiresolution analysis partly corrected this problem with the construction of integral transforms and expansion bases adjusted on a non-uniform time-frequency scale [1]. Within constant uncertainty product, higher frequencies are represented with higher time resolution and lower frequency with lower time resolution. Indeed, if we are required to detect the pitch of a tone we need to observe at least one period, which has a duration inversely proportional to the frequency. If the integral wavelet transform, in its redundancy, does not introduce limitations on either resolution other than those dictated by the uncertainty principle, the dyadic wavelet bases constrain the frequency resolution to one octave [2][3], with a resolution factor $\frac{1}{2}$.

The reason why dyadic bases are popular lies in the fact that they are easy to generate and fast algorithms for computing the expansion coefficients are available. Wavelet bases based on rational frequency resolution lead to difficult design procedures in order to satisfy regularity constraints [4]. Furthermore, the frequency resolution factor is in practice limited to ratios of small integers. In contrast, frequency warping techniques allow for arbitrary choice of analysis bands. This is an important point if we want to adapt the signal representation either to perceptual, e.g., Bark, scales or to objective bands associated to a single feature. However, computation of the transform based on general warping maps is not efficient unless one is able to derive a structure based on rational transfer functions [5].

In this paper, methods for frequency warping based on the Laguerre transform whose computation is achieved by means of a chain of digital all-pass filters are applied to wavelets. Iterated frequency warping by means of Laguerre maps leads to wavelets with arbitrary frequency resolution, whose bandwidths can be assigned by selecting a set of parameters [6]. Interestingly enough, these wavelets are still based on a dyadic scheme, which is a natural setting for iterated band splitting. The case where the choice of the design parameters leads to orthogonal and complete wavelets with arbitrary resolution factor $0 < a < 1$ is examined in detail. In this context, while single-step frequency warping requires an a-homogeneous selfsimilar map, we show that iterated warping is most conveniently associated with conjugacy and Schröder's equation. This property is also useful for showing convergence of the parameter sequence via an extension of Königs' theorem [7] to parametric maps on the real axis.

Another issue in musical signal analysis is that the frequency spectrum of voiced signals peaks at uniformly or non-uniformly spaced frequencies. While exact periodicity is rarely encountered in real-life signals, exploitation of the pitch information or of pseudoperiodicity is an asset for their representation. The dynamics of most sounds may be captured in at least three phases:

- the attack, where abrupt transients are among the most important factors in our ability to recognize a given instrument,
- the sustain, where the signal is quasi-stationary and quasi-periodic, although small fluctuations in this phase add richness to the dynamics and are important natural factors,
- the release, where the signal fades away, most often with distinct decay rates in different areas of the frequency spectrum.

The Fourier transform is particularly gifted at revealing periodicity, however the STFT represents pseudoperiodic signals as a superposition of partials whose dynamics is captured with uniform time resolution. It follows that transients are blurred if the selected time-resolution is too poor or frequency resolution is at risk when the analysis window is too short.

In previous papers [8][9] the author suggested methods for exploiting the pitch information in wavelet representations by introducing the Pitch-Synchronous Wavelet Transform (PSWT). The signal is first converted into a sequence of variable-length vectors each containing the samples of one period of the signal, then the sequences of components are analyzed by means of an array of wavelet transforms. This representation is able to capture period-to-period fluctuations of the signal by means of basis elements that are comb-like in the frequency domain. Periodic behavior is trapped in narrow combs adjusted on the harmonic grid and transients at several scales are represented by multiresolution basis elements whose Fourier transforms are organized in ensembles of sidebands of the harmonics. Faster transitions are represented at fine time resolution by larger bands lying far from the harmonics. Vice-versa, slow transitions and quasi-stationary dynamics are represented at coarser time resolution by narrower sidebands lying closer to the harmonics. In particular, this technique allows for separation of perceptually and physically distinct phenomena such as the bow noise and the harmonic resonant component in a violin tone. Importantly enough, this separation is achieved by means of a complete and orthogonal set and the sum of the components exactly reconstructs the original signal.

The PSWT technique is amenable to further generalization, essentially by introducing alternate methods for forming the vector signal, e.g., based on intermediate Gabor-like or STFT representations, which can be formalized in the multiwavelets framework. However these have several drawbacks: they do not generate complete and orthogonal sets when the pitch is variable, they blur the time resolution and the construction of pitch-synchronous bases requires the design of filter banks with a large number of bands.

While the PSWT works efficiently when the frequencies of the partials are arranged on a harmonic grid, the frequency spectrum of many sounds, such as piano tones in the low register, has an inherently inharmonic structure. From a physical model point of view, this phenomenon may be explained by the fact that stiffness of the material is not negligible and dispersive wave propagation occurs within the medium. Furthermore, coupled 2D modes in drums have a peculiar distribution of eigenfrequencies. A partial solution of this problem is provided by frequency warping techniques that allow for redistribution of eigenfrequencies into harmonics. Inharmonic signals are first regularized by means of a frequency map and then analyzed by means of PSWT, leading to the Frequency Warped PSWT [6].

The paper is organized as follows. In section 2 the properties of one-step frequency warping maps of continuous-time signals and wavelets are examined. Methods for frequency warping discrete-time signals and the Laguerre transform are illustrated in section 3. These methods are applied to define discrete-time warped wavelets in section 4. In section 5 an extension of Königs' theorem on iterated maps is presented and applied to the construction of scale a continuous-time warped wavelets.

2. Frequency Warping and Wavelets

2.1. **Frequency Warping.** Frequency warping is obtained by mapping the frequency axis ω by means of a suitable real function $\Omega(\omega)$. Given a continuous-time signal $x(t)$ with Fourier Transform (FT)

$$X(\omega) = \int_{-\infty}^{+\infty} x(t)\; e^{-j\omega t} dt$$

we form the warped signal

$$\widetilde{x}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\Omega(\omega))\; e^{j\omega t} d\omega,$$

i.e., we let

$$\widetilde{X}(\omega) = X(\Omega(\omega)) = \int_{-\infty}^{+\infty} x(t)\; e^{-j\Omega(\omega)t} dt \tag{2.1}$$

Thus, at any frequency $\overline{\omega}$ the FT of the warped signal has the same magnitude as the FT of the signal at frequency $\Omega(\overline{\omega})$, i.e., the frequency spectrum of the warped signal is a deformed version of that of the original signal. Clearly, for the warping operation to be well defined, the map $\Omega(\omega)$ must satisfy a number of requirements. In this paper we are interested in reversible warping operations defined by a strictly increasing, invertible, map Ω of the real axis onto itself. The map is also assumed to have odd parity: $\Omega(-\omega) = -\Omega(\omega)$. In this way frequency ordering is preserved and warping a real signal yields a real signal. Notice that, as defined in (2.1), frequency warping is not energy preserving since narrow frequency bands may be mapped into broader bands with equal peak amplitude. In order to preserve energy in any frequency interval one needs to multiply the right hand side of (2.1) by a suitable amplitude scaling function or, more generally, define a suitable measure of the frequency axis. In the sequel we will assume that the map is almost everywhere differentiable, e.g., that $\Omega(\omega)$ is absolutely continuous, and define a scaled warped operation on the signal $x(t)$ as the one producing the FT

$$\widehat{X}(\omega) \overset{a.e.}{=} \sqrt{\Omega'(\omega)} X(\Omega(\omega)),$$

where the derivative $\Omega'(\omega)$ is positive almost everywhere since the map is assumed to be strictly increasing. In that case one obtains energy preservation in any band. In fact, a frequency band with support $[\omega_0, \omega_1]$ is mapped into a frequency band with support $\left[\Omega^{-1}(\omega_0), \Omega^{-1}(\omega_1)\right]$ and

$$\frac{1}{2\pi} \int_{\Omega^{-1}(\omega_0)}^{\Omega^{-1}(\omega_1)} \left|\widehat{X}(\omega)\right|^2 d\omega = \frac{1}{2\pi} \int_{\omega_0}^{\omega_1} |X(\omega)|^2 \, d\omega.$$

From a perceptual point of view this property is important since the scaled warped signal is properly equalized while in the unscaled version some parts of the frequency spectrum are boosted. From a mathematical point of view scaled warping is associated with an orthogonal operator $\mathbb{W}$:

$$\widehat{x}(t) = \left[\mathbb{W}x\right](t) = \int_{-\infty}^{+\infty} W(t, \tau)\, x(\tau)\, d\tau$$

with real kernel

$$W(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sqrt{\Omega'(\omega)} e^{j(\omega t - \Omega(\omega)\tau)} d\omega$$

and inverse operator $\mathbb{W}^{-1}$ defined by the kernel

$$W^{-1}(t,\tau) = W(\tau, t). \tag{2.2}$$

Finally, we remark that, by exploiting duality of the FT one can define time warping operations in a similar fashion.

2.2. **Simple Warped Wavelets.** In [10][11] Baraniuk and Jones exploited unitary warping operators in order to define warped wavelets. Their construction may be summarized as follows: the signal is unwarped by means of the inverse warping operator $\mathbb{W}^{-1}$, then the expansion coefficients on a dyadic wavelet basis are computed. Reconstruction is achieved by applying the warping operator $\mathbb{W}$ to the wavelet expansion of the unwarped signal. In other words, given an orthogonal and complete dyadic wavelet set

$$\left\{\psi_{n,m}(t)\right\}_{n,m\in\mathbb{Z}} ,$$

where

$$\psi_{n,m}(t) = 2^{-n/2}\psi_{0,0}(2^{-n}t - m) , \tag{2.3}$$

one computes the expansion coefficients

$$\widehat{x}_{n,m} = \left\langle \mathbb{W}^{-1}x, \psi_{n,m}\right\rangle$$

and reconstructs the signal as follows

$$x(t) = \sum_{n,m\in\mathbb{Z}} \widehat{x}_{n,m}\left[\mathbb{W}\psi_{n,m}\right](t). \tag{2.4}$$

Since the frequency warping operator is orthogonal we have

$$\widehat{x}_{n,m} = \left\langle \mathbb{W}^{-1}x, \psi_{n,m}\right\rangle = \left\langle x, \mathbb{W}\psi_{n,m}\right\rangle$$

and (2.4) may be considered as the expansion on the warped wavelet basis

$$\left\{\left[\mathbb{W}\psi_{n,m}\right](t)\right\}_{n,m\in\mathbb{Z}} .$$

Indeed this basis is orthogonal since

$$\left\langle \mathbb{W}\psi_{n',m'}, \mathbb{W}\psi_{n,m}\right\rangle = \left\langle \psi_{n',m'}, \mathbb{W}^{-1}\mathbb{W}\psi_{n,m}\right\rangle = \delta_{n',n}\delta_{m',m}$$

and complete in view of unitary equivalence.

The FT of the warped wavelets

$$\widehat{\psi}_{n,m}(t) \equiv \left[\mathbb{W}\psi_{n,m}\right](t)$$

is related to the FT of the dyadic wavelets as follows

$$\widehat{\Psi}_{n,m}(\omega) = \sqrt{\Omega'(\omega)}\Psi_{n,m}(\Omega(\omega)) = \sqrt{2^n\Omega'(\omega)}\Psi_{0,0}(2^n\Omega(\omega))e^{-j2^n m\Omega(\omega)} \tag{2.5}$$

from which we see that warped wavelets are not simply generated by dilating and translating a mother wavelet as in dyadic wavelets (2.3). Rather, the "translated wavelets" are generated by generally non-rational all-pass filtering $e^{-j2^n m\Omega(\omega)}$ and scaling depends on the warping map $\Omega(\omega)$ as well. However, frequency warped wavelets have the remarkable property that their frequency resolution, i.e., their essential frequency support, may be arbitrarily assigned by proper choice of the map $\Omega(\omega)$. Indeed, if the cut-off frequencies of dyadic wavelets are fixed at $2^{-n}\pi$, the cut-off frequencies of the warped wavelets are

$$\omega_n = \Omega^{-1}(2^{-n}\pi).$$

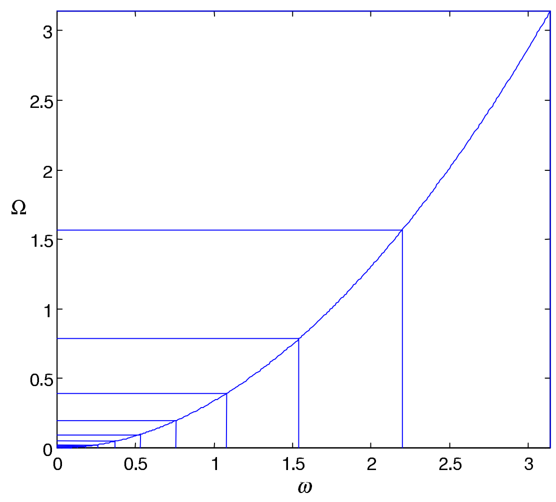

FIGURE 1. Example of warping map for power of a wavelet cut-off choice.

For our purposes we just need to add the remark that genuine scale a wavelets, $0 < a < 1$, may be generated by selecting as warping map an a-homogeneous function

$$\Omega(a\omega) = \tfrac{1}{2}\Omega(\omega) \tag{2.6}$$

such as

$$\Omega(\omega) = \frac{\pi\omega}{|\omega|} 2^{-\log_a \frac{|\omega|}{\pi}},$$

reported in fig.1. Repeated application of (2.6) shows that

$$\Omega\left(a^{-n}\omega\right) = 2^n \Omega(\omega)\ . \tag{2.7}$$

By deriving both sides of (2.7) one obtains

$$\Omega'(a^{-n}\omega) = (2a)^n\, \Omega'(\omega)\ . \tag{2.8}$$

By substituting (2.7) and (2.8) in (2.5) we obtain

$$\widehat{\Psi}_{n,m}(\omega) = a^{-n/2}\widehat{\Psi}_{0,m}(a^{-n}\omega)$$

which shows that warped wavelets generated by an a-homogeneous warping map are obtained by dilating the family of mother wavelets $\widehat{\Psi}_{0,m}(\omega)$, $m \in \mathbb{Z}$, while wavelets at fixed scale level n are all-pass related:

$$\widehat{\Psi}_{n,m}(\omega) = \widehat{\Psi}_{n,0}(\omega)e^{-jm\Omega(a^{-n}\omega)}$$

Without loss of generality we can assume that

$$\Omega(\pi) = \pi$$

so that the cut-off frequencies of the scale a warped wavelets are fixed at

$$\omega_n = a^n\pi\ .$$

If $a > \frac{1}{2}$ then the warped wavelets achieve a finer frequency resolution than the dyadic wavelets.

3. Discrete-time frequency warping and Laguerre transform

The continuous-time warped wavelet expansion illustrated in the previous section is applicable to continuous-time signals and requires warping of the signal, a task that is computationally difficult to perform. An important property of dyadic wavelet expansions is that they have a fast algorithm to iteratively compute the expansion coefficients even in the continuous-time case. In this section we consider frequency warping of discrete-time signals and relate this to an orthogonal transform that can be computed by means of a chain of rational all-pass filters.

Since the discrete-time Fourier transform (DTFT) of a sequence $x(n)$ is periodic, reversible frequency warping of discrete-time signals requires an invertible frequency map of the interval $[-\pi, \pi]$ onto itself. As we did in the continuous-time case, we will assume that $\theta(\omega)$ is a.e. differentiable, has odd parity and maps the point π into itself. As discussed in the previous section, we are interested in procedures for unwarping the signal, producing the signal $\widehat{x}(n)$ whose DTFT is

$$\widehat{X}(\omega) = \sqrt{\frac{d\theta^{-1}}{d\omega}} X(\theta^{-1}(\omega)) \ .$$

It follows that the discrete-time unwarping operator $\mathbb{W}^{-1}$ is orthogonal and has kernel

$$\begin{aligned} W^{-1}(n,k) &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \sqrt{\frac{d\theta^{-1}}{d\omega}} e^{j(n\omega - k\theta^{-1}(\omega))} d\omega \\ &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} \sqrt{\theta'(\omega)} e^{j(k\omega - n\theta(\omega))} d\omega. \end{aligned} \tag{3.1}$$

The unwarped signal is

$$\widehat{x}(n) = \left[\mathbb{W}^{-1} x\right](n) = \sum_k W^{-1}(n,k)\, x(k) = \sum_k \lambda_n(k)\, x(k), \tag{3.2}$$

where we defined

$$\lambda_n(k) \equiv W^{-1}(n,k) \ . \tag{3.3}$$

Notice that (3.2) is in the form of a scalar product of the signal with the sequence $\lambda_n(k)$. By the orthogonality of the operator

$$W^{-1}(n,k) = W(k,n)$$

and the set $\{\lambda_n(k)\}_{n\in\mathbb{Z}}$ is orthogonal and complete. In fact:

$$\sum_k \lambda_n(k)\lambda_{n'}(k) = \sum_k W^{-1}(n,k)W(k,n') = \delta_{n,n'}$$

and

$$\sum_n \lambda_n(k)\lambda_n(k') = \sum_k W^{-1}(n,k)W(k',n) = \delta_{k,k'}.$$

Thus, the discrete-time unwarped signal is given by the sequence of expansion coefficients

$$\widehat{x}(n) = \langle x, \lambda_n \rangle \tag{3.4}$$

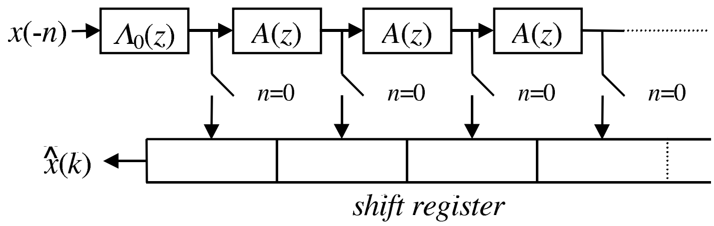

FIGURE 2. Structure for computing discrete-time signal unwarping by means of the Laguerre transform.

of the signal over the orthogonal basis $\{\lambda_n(k)\}_{n\in\mathbb{Z}}$ and the signal $x(n)$ is recovered from $\widehat{x}(n)$ by the expansion series

$$x(n) = \sum_k \widehat{x}(k)\lambda_k(n) \ .$$

In turn, if the signal is causal, i.e., if $x(k) = 0$ for $k < 0$, then the scalar product (3.4) may be computed by convolving the time-reversed signal $y(-k) = x(k)$ by $\lambda_n(k)$ and sampling the result at $n = 0$:

$$\langle x, \lambda_n \rangle = \sum_{k=0}^{\infty} x(k)\lambda_n(k) = \sum_{k=0}^{\infty} y(0-k)\lambda_n(k) \ .$$

Furthermore, from (3.3) and (2.2) it is easy to recognize that the DTFT of $\lambda_n(k)$ satisfies the following recurrence:

$$\Lambda_n(\omega) = \Lambda_{n-1}(\omega)e^{-j\theta(\omega)},$$

with

$$\Lambda_0(\omega) = \sqrt{\theta'(\omega)}.$$

Unwarping may be then computed by means of a cascade of all-pass filters forming a dispersive, i.e., frequency dependent, delay line sampled at $n = 0$. However, for digital implementations one needs to constrain $-\theta(\omega)$ to be the phase of a causal and stable, rational all-pass filter. In this case it is easy to show that $\theta(\omega)$ is one-to-one with odd parity mapping and fixing the point π if and only if $-\theta(\omega)$ is the phase of a first-order real all-pass filter with transfer function

$$A(z) = \frac{z^{-1} - b}{1 - bz^{-1}}, \ -1 < b < 1.$$

The corresponding basis set $\lambda_n(k)$ is recognized to be the discrete Laguerre basis [12][5][13], with

$$\theta(\omega) = \omega + 2 \arctan \frac{b \sin\omega}{1 - b\cos\omega}, \ -1 < b < 1 \tag{3.5}$$

and

$$\Lambda_0(\omega) = \frac{\sqrt{1-b^2}}{1 - be^{-j\omega}}.$$

Laguerre warping is the unique one-to-one warping that can be computed by means of rational filters using the structure of fig. 2. The corresponding family of warping curves (3.5) is shown in fig. 3.

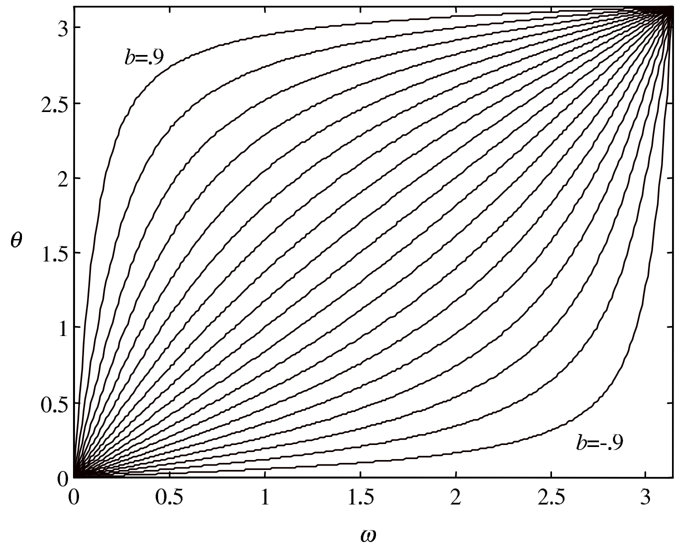

FIGURE 3. Family of warping curves associated to the Laguerre transform.

4. WARPED FILTER BANKS AND DISCRETE-TIME WARPED WAVELETS

4.1. **Two-channel warped filter bank.** In the classical construction of dyadic wavelets [1] a two-channel critically sampled filter bank serves as building block. This is based on two quadrature mirror filters (QMF) $H(z)$, low-pass, and $G(z)$, high-pass. These filters are half-band with cut-off frequency $\omega_c = \frac{\pi}{2}$. In [5] the author, together with S. Cavaliere, considered as building block the structure of fig. 4, where the signal is frequency unwarped by means of Laguerre transform (LT) prior to filtering. Unwarping the signal is equivalent to warping both the filters and the downsampling / upsampling operators [5], as depicted in fig. 5. Thus, on one hand the pass-band of the filters is altered and the cut-off frequency is moved to $\omega_c = \theta^{-1}(\frac{\pi}{2})$, on the other hand the equivalent warped downsampling operator, which embeds unwarping and conventional downsampling, resets the band of the filtered signal to half-band prior to downsampling. By combining the Laguerre transform with an orthogonal filter bank one obtains an orthogonal perfect reconstruction structure whose bands can be arbitrarily allocated by fixing the parameter b.

The structure of fig. 4 can be generalized to include an ordinary discrete wavelet transform structure in place of the two-channel filter bank. This provides a first form of discrete-time warped wavelets. However, the analysis bands cannot be arbitrarily selected since they are determined by transforming the cut-off frequencies $2^{-n}\pi$ according to a single curve in the Laguerre family. As shown in the next section, fully arbitrary bandwidth allocation is achieved by iterating the full warped filter bank structure.

Frequency warping may be successfully employed for transforming inharmonic sounds into harmonic ones or vice-versa. When used in conjunction with PSWT one obtains a PSFWWT version useful for decomposing inharmonic instrument sounds into regular pseudo-periodic components plus fluctuations. The magnitude FT of

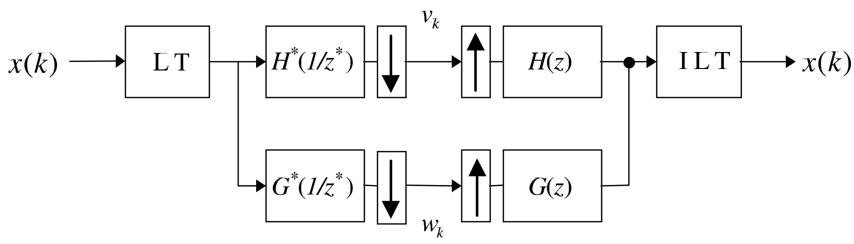

FIGURE 4. Two-channel warped filter bank structure

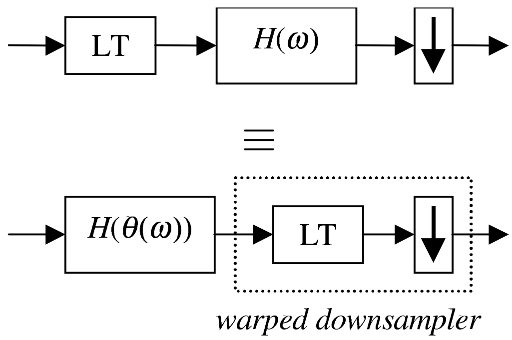

FIGURE 5. Equivalent downsampled warped filtering structures.

the basis elements is reported in fig. 6. The PSFWWT decomposition allows, e.g., to resolve the hammer noise in a piano tone. An account of this method is given in [6][14], where the Laguerre warping family is seen to close match the dispersion curves provided by both experimental data and physical model based on a 4^{th} order PDE.

4.2. **Discrete-time warped wavelets with arbitrary band allocation.** Discrete time warped wavelet decomposition may be obtained by iterating the two-channel warped filter bank. Similarly to the ordinary wavelet transform, identical warped filter bank structures are nested in the low-pass branch. Each warping stage is characterized by a different value b_k of the Laguerre parameter. One can show [6] that the DTFT of the level n warped wavelet has the form

$$\Psi_{n,m}(\omega) = \Lambda_{k,0}(2\Omega_{k-1}(\omega))G(\Omega_k(\omega))\Phi_{n-1,0}(\omega)e^{-j2m\Omega_n(\omega)}$$

where

$$\Phi_{n,0}(\omega) = \prod_{k=1}^{n} \{\Lambda_{k,0}(2\Omega_{k-1}(\omega))H(\Omega_k(\omega))\} \tag{4.1}$$

is the DTFT of the warped scaling sequence,

$$\Lambda_{k,0}(\omega) = \sqrt{\theta_k'(\omega)}\frac{\sqrt{1-b_k^2}}{1-b_k e^{-j\omega}}$$

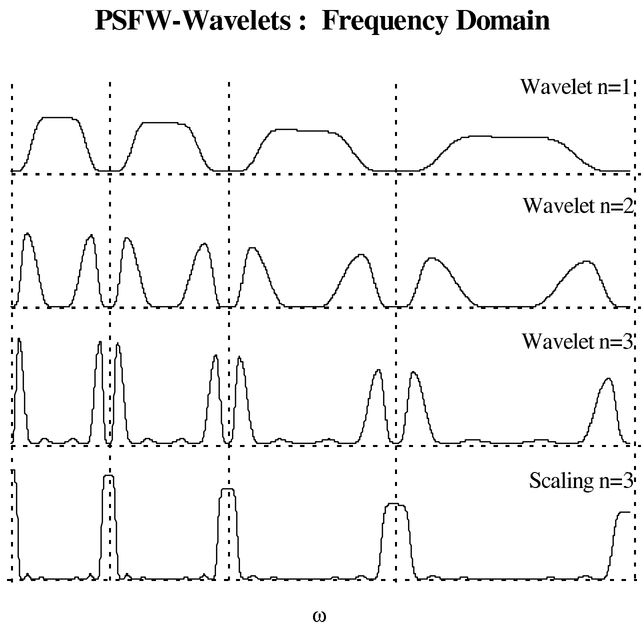

FIGURE 6. Frequency domain characteristics of pitch-synchronous frequency warped wavelets.

and

$$\Omega_k(\omega) = \frac{1}{2}\vartheta_k \circ \vartheta_{k-1} \circ \cdots \circ \vartheta_1$$

with

$$\vartheta_k(\omega) = 2\theta_k(\omega) \ .$$

The cut-off frequency of the level n warped scaling sequence is given by

$$\omega_n = \Omega_n^{-1}(\frac{\pi}{2}) \ ,$$

which is the cut-off frequency of the narrowest band filter $H(\Omega_n(\omega))$ in the product (4.1).

Given a choice of the cut-off frequencies $\omega_n < \omega_{n-1} < \cdots < \omega_1$, there exists a unique solution for the Laguerre parameters b_k, $k = 1, ..., n$, obtained by the following recurrence

$$\begin{aligned} b_1 &= \tan\left(\frac{\pi - 2\omega_1}{4}\right) \\ b_k &= \tan\left(\frac{\pi}{4} - \Omega_{k-1}(\omega_k)\right) \end{aligned} \tag{4.2}$$

Consequently, to any choice of the cut-off frequencies is uniquely associated a corresponding set of parameters. The iterated warped filter bank is a perfect reconstruction orthogonal structure equivalently described by a discrete-time warped wavelet basis. Cut-off frequencies may be arranged on a perceptual scale, such as Bark scale, leading to perceptually organized orthogonal transforms [15][16].

5. Continuous-time warped wavelets

Expansion coefficients over ordinary continuous-time wavelet bases may be iteratively computed by means of an infinite filter bank structure. In this section we start from the definition of discrete-time warped wavelets given in the previous section to construct continuous-time warped wavelets such that the expansion coefficients can be computed by means of an iterated warped filter bank structure. Two problems arise concerning the convergence of the warping map and the construction of suitable scaling functions. As we will show next, conjugacy rather than homogeneity is a key requirement for obtaining scale a wavelets.

5.1. **Schröder's equation and extension of Königs' theorem.** Our theory on iterated warping maps is based on Königs' models for solving Schröder's equation. The latter is an eigenfunction equation for the composition operator. The classical theory provides results for identical maps of the inner unit disk into itself [7]. In order to apply the results to iterated warping maps we need to extend the theory to allow for parametric maps of the real axis. We were able to prove the following

Theorem 1. *Let $f(\omega;\eta)$ be a 1-parameter family of maps of the real line into itself, fixing the point $\omega = 0$ for any value of the parameter η.*

Let $f_n(\omega) \equiv f(\omega;\eta_n)$, where η_n, $n \in \mathbb{N}$, is a sequence of values of the parameter η such that

- *$\forall n \in \mathbb{N}$, f_n has bounded second derivative and $|f_n'(0)| > 0$*
- *$|f_n(\omega)| < |\omega|$, for all but finitely many integers n*

then the sequence $\dfrac{F_n(\omega)}{F_n'(0)}$, with $F_n \equiv f_1 \circ f_2 \circ \cdots \circ f_n$ and $F_n'(0) = \prod_{k=1}^{n} f_k'(0)$, converges uniformly on any compact subset of $\mathbb{R}$.

Furthermore, if $f_\infty(\omega) \equiv \lim_n f_n(\omega)$ exists finite and differentiable at $\omega = 0$ then Schröder's equation

$$F \circ f_\infty = f_\infty'(0) F, \tag{5.1}$$

is satisfied by $F \equiv \lim_n \dfrac{F_n}{F_n'(0)}$.

This theorem can be applied to the iterated inverse map

$$\Omega_n^{-1}(\omega) = \vartheta_1^{-1} \circ \vartheta_2^{-1} \circ \cdots \circ \vartheta_n^{-1}(2\omega)$$

where

$$\vartheta_n^{-1}(\omega) = \vartheta^{-1}(\omega; b_n) = \frac{\omega}{2} - 2 \arctan \frac{b_n \sin \frac{\omega}{2}}{1 + b_n \cos \frac{\omega}{2}}.$$

In the exponential cut-off assumption

$$\Omega_n^{-1}\left(\frac{\pi}{2}\right) = a^n \pi \tag{5.2}$$

the family $\vartheta^{-1}(\omega; b)$ satisfies the hypotheses of Theorem 1. Since

$$\frac{d}{d\omega}\, \Omega_n^{-1}\left(\frac{\omega}{2}\right)\Big|_{\omega=0} = \prod_{k=1}^{n} \frac{\nu_k}{2} = \gamma_n \,,$$

where

$$\nu_k = \frac{1 - b_k}{1 + b_k},$$

the sequence

$$\frac{1}{\gamma_n}\Omega_n^{-1}\left(\frac{\omega}{2}\right)$$

converges, uniformly on any compact subset of $\mathbb{R}$ to an increasing and differentiable map $\widetilde{\Omega}^{-1}(\omega)$. Convergence together with the cut-off condition (5.2) imply that for $\omega = \pi$ we have

$$\widetilde{\Omega}^{-1}(\pi) = \lim_n \frac{a^n \pi}{\gamma_n} = \pi \prod_{k=1}^{\infty} \frac{2a}{\nu_k}.$$

Hence

$$\lim_n \nu_n = 2a \tag{5.3}$$

with exponential convergence rate:

$$\nu_n = 2a + O(a^{2n-1}).$$

The limit (5.3) is an important result since it guarantees that the sequence of parameters ν_n, whose dependency on a is extremely complex in view of the iterated map in (4.2), rapidly converges to a linear function of a.

By Theorem 1, the map $\widetilde{\Omega}^{-1}(\omega)$ is an eigenfunction of the composition-by-ϑ_∞^{-1} operator with eigenvalue a:

$$\widetilde{\Omega}^{-1} \circ \vartheta_\infty^{-1} = a\widetilde{\Omega}^{-1}, \tag{5.4}$$

where ϑ_∞^{-1} has parameter $\nu_\infty = 2a$. Since any solution of Schröder's equation (5.4) is unique up to a multiplicative constant, one can normalize $\widetilde{\Omega}^{-1}$ by defining

$$\Omega^{-1}(\omega) = \frac{\pi\widetilde{\Omega}^{-1}(\omega)}{\widetilde{\Omega}^{-1}(\pi)} = \lim_n \frac{\Omega_n^{-1}(\omega)}{a^n},$$

which is a solution of Schröder's equation (5.4) satisfying $\Omega^{-1}(\pi) = \pi$.

The direct map $\Omega(\omega)$ satisfies the equation

$$\vartheta_\infty \circ \Omega(\omega) = \Omega\left(\frac{\omega}{a}\right).$$

Since

$$\begin{aligned} \vartheta_\infty(\omega) &= \Omega\left(a^{-1}\Omega^{-1}(\omega)\right) \\ \vartheta_\infty^{-1}(\omega) &= \Omega\left(a\Omega^{-1}(\omega)\right), \end{aligned} \tag{5.5}$$

the composition of several ϑ_∞ or ϑ_∞^{-1} maps is conjugated, respectively, to negative or positive powers of the eigenvalue a via the eigenfunction and its inverse. This conjugacy relationship can be exploited to define continuous-time scale a wavelets from the discrete-time warped wavelets. This leads to the following form for the FT of the "mother" scaling function:

$$\widehat{\Phi}_{0,0}(\omega) = \sqrt{\Omega'(\omega)} \prod_{k=1}^{\infty} \frac{H\left(\frac{1}{2}\Omega(a^{k-1}\omega)\right)}{\sqrt{2}} \tag{5.6}$$

with "translation" given by

$$\widehat{\Phi}_{0,m}(\omega) = \widehat{\Phi}_{0,0}(\omega)e^{-jm\Omega(\omega)}.$$

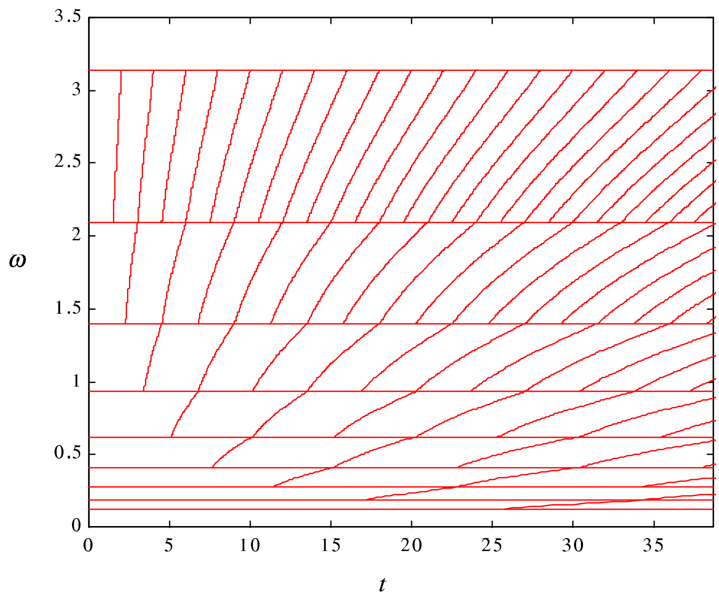

FIGURE 7. Tiling the time-frequency plane by means of frequency warped wavelets.

This construction is compatible with powers of a scaling

$$\widehat{\Phi}_{n,m}(\omega) = a^{-n/2}\widehat{\Phi}_{0,m}\left(a^{-n}\omega\right)$$

in that at each fixed level n scaling functions are orthogonal and their span forms an orthogonal subspace V_n complementary in V_{n-1} to the space spanned by the level n wavelets. With the usual requirements on the QMF filter H [1] one can show that the set of wavelets is orthogonal and complete in $L^2(\mathbb{R})$. By applying conjugacy relationships (5.5) one obtains the following warped two-scale equation

$$\frac{1}{\sqrt{a}}\widehat{\Phi}_{0,0}\left(\frac{\omega}{a}\right) = \bar{\Lambda}_{\infty,0}\left(\Omega(\omega)\right) H\left(\tfrac{1}{2}\Omega\left(\frac{\omega}{a}\right)\right)\widehat{\Phi}_{0,0}(\omega)\,.$$

This result allows us to iteratively compute the expansion coefficients over the warped wavelet basis by means of the iterated warped filter bank structure with identical parameters $b_\infty = \frac{1-2a}{1+2a}$.

Warped wavelet decomposition leads to an unconventional tiling of the time-frequency plane shown in fig. 7. Due to frequency dependent delay of each basis element, the uncertainty zones are characterized by curved boundaries.

6. Conclusions

In this paper we explored the panorama of wavelet transforms obtained by means of frequency warping techniques, with particular attention to their realizability in computationally efficient structures based on rational discrete-time all-pass filters and two-channel filter banks. The construction of scale a warped wavelets is based on an original extension of Königs' theorem on iterated maps.

The flexibility of the transform is exploited for the representation of music signals in perceptually or objectively meaningful components, allowing for adaptation to perceptual scales or for orthogonal separation of transients and noise in harmonic or inharmonic sounds via pitch-synchronous methods.

Frequency warping emerges as an appealing technique in musical signal processing for the creation of new effects and for the analysis and synthesis of features of sounds by means of adapted representations.

References

[1] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Reg. Conf. Series in Appl. Math. SIAM, Philadelphia, PA., 1992.

[2] G. Evangelista, "Wavelet transforms that we can play", in *Representations of Musical Signals*, A. Piccialli G. De Poli and C. Roads, Eds. 1991, pp. 119–136, MIT Press, Cambridge, MA.

[3] G. Evangelista, "Wavelet representations of musical signals", in *Musical Signal Processing*, C. Roads A. Piccialli, G. De Poli and S. T. Pope, Eds. 1997, pp. 127–153, Swets & Zeitlinger, Amsterdam, Holland.

[4] T. Blu, "A new design algorithm for two-band orthonormal filter banks and orthonormal rational wavelets", *IEEE Trans. on Signal Processing*, vol. 46, no. 6, pp. 1494–1504, June 1998.

[5] G. Evangelista and S. Cavaliere, "Frequency warped filter banks and wavelet transforms: A discrete-time approach via laguerre expansion", *IEEE Trans. on Signal Processing*, vol. 46, no. 10, pp. 2638–2650, October 1998.

[6] G. Evangelista and S. Cavaliere, "Discrete frequency warped wavelets: Theory and applications", *IEEE Trans. on Signal Processing*, vol. 46, no. 4, pp. 874–885, April 1998, Special issue on Theory and Applications of Filter Banks and Wavelets.

[7] J. H. Shapiro, "Composition operators and schroeder's functional equation", *Trans. of AMS*, vol. Contemporary Mathematics, no. 213, pp. 213–228, 1998.

[8] G. Evangelista, "Pitch synchronous wavelet representations of speech and music signals", *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3313–3330, December 1993, special issue on Wavelets and Signal Processing.

[9] G. Evangelista, "Comb and multiplexed wavelet transforms and their applications to signal processing", *IEEE Trans. on Signal Processing*, vol. 42, no. 2, pp. 292–303, February 1994.

[10] R. G. Baraniuk and D. L. Jones, "Warped wavelets bases: Unitary equivalence and signal processing", *Proc. ICASSP'93, IEEE*, vol. III, pp. 320–323, 1993.

[11] R. G. Baraniuk and D. L. Jones, "Unitary equivalence : A new twist on signal processing", *IEEE Trans. Signal Processing*, vol. 43, no. 10, pp. 2269–2282, October 1995.

[12] P. W. Broome, "Discrete orthonormal sequences", *J. Assoc. Comput. Machinery*, vol. 12, no. 2, pp. 151–168, 1965.

[13] D. H. Johnson A. V. Oppenheim and K. Steiglitz, "Computation of spectra with unequal resolution using the fast fourier transform", *Proc. IEEE*, vol. 59, pp. 299–301, February 1971.

[14] G. Evangelista and S. Cavaliere, "Analysis and regularization of inharmonic sounds via pitch-synchronous frequency warped wavelets", *Proc. of ICMC '97*, pp. 51–54, September 1997.

[15] G. Evangelista and S. Cavaliere, "Arbitrary bandwidth wavelet sets", *Proc. of ICASSP98*, vol. III, pp. 1801–1804, May 1998.

[16] G. Evangelista and S. Cavaliere, "Auditory modeling via frequency warped wavelet transform", *Proc. of EUSIPCO98*, vol. I, pp. 117–120, September 1998.