REDRESSING WARPED WAVELETS AND OTHER SIMILAR WARPED TIME-SOMETHING REPRESENTATIONS

Gianpaolo Evangelista

Institute for Composition, Electroacoustics and Sound Engineering Education
MDW University of Music and Performing Arts
Vienna, Austria
evangelista@mdw.ac.at

ABSTRACT

Time and frequency warping provide effective methods for fitting signal representations to desired physical or psychoacoustic characteristics. However, warping in one of the variables, e.g. frequency, disrupts the organization of the representation with respect to the conjugate variable, e.g. time. In recent papers we have considered methods to eliminate or mitigate the dispersion introduced by warping in time frequency representations and Gabor frames. To this purpose, we introduced redressing methods consisting in further warping with respect to the transformed variables. These methods proved not only useful for the visualization of the transform but also to simplify the computation of the transform in terms of shifted precomputed warped elements, without the need for warping in the computation of the transform. In other linear representations, such as time-scale, warping generally modifies the transform operators, making visualization less informative and computation more difficult. Sound signal representations almost invariably need time as one of the coordinates in view of the fact that we normally wish to follow the time evolution of features and characteristics. In this paper we devise methods for the redressing of dispersion introduced by warping in wavelet transforms and in other expansions where time-shift plays a role.

1. INTRODUCTION

Linear representations play a central role in sound synthesis, digital audio effects, feature extraction, coding and music information retrieval. In most of these representations, one of the coordinates in the transformed domain is time-shift, as in the case of time-frequency representations [1], e.g., based on the STFT (Short-Time Fourier Transform, also known as phase VOCODER) or as in the case of time-scale representations based on the WT (Wavelet Transform) [2, 3]. Other examples, based on different signal transformations such as the MDCT (Modified Discrete Cosine Transform) [4, 5], windowed Hankel (Fourier-Bessel) [6, 7], to name a few, are available.

In this paper we are mostly concerned with audio signal representations that are exact, i.e., for which an inverse or pseudo-inverse exists such that perfect reconstruction of the original signal is possible from the analysis data. Other analysis-only methods, such as the constant Q transformation [8] may enjoy easier computation but may introduce data loss or have inefficient reconstruction algorithms so they are not directly suitable for the synthesis from analysis.

Generally speaking, mathematical transforms operate on a one-dimentional audio signal and produce a multidimentional representation. For example the STFT yields a 2D signal representation in which one coordinate can be interpreted as time (more precisely time-shift) and another coordinate as frequency, so that the variation of the frequency spectrum in time can be visualized and used for feature detection such as instantaneous frequency or amplitude envelopes of the partials.

Linear signal transformations are computed by taking the scalar product of the signal with a transform nucleus function [9] $K_{\alpha_1,...,\alpha_N}(t)$, where $\alpha_1, ..., \alpha_N$ are the N transform domain coordinates [2]. For example, in the STFT, the nucleus is a modulated window where α_1 is associated to time-shift and α_2 to frequency.

When time-shift is one of the transform coordinates, in a large class of linear 2D representations the transform nucleus is computed by operating on a test function g, e.g., the prototype analysis window or the analysis mother wavelet. In order to formalize the generation of the transform nucleus of this type, one can introduce a time-shift operator $\mathbf{T}_\tau$, whose action on a time signal s is defined as follows:

$$[\mathbf{T}_\tau s](t) = s(t-\tau), \tag{1}$$

and another parametric linear operator $\mathbf{O}_\sigma$. The corresponding transform nucleus can be then written as

$$K_{\tau,\sigma}(t) = [\mathbf{T}_\tau \mathbf{O}_\sigma g](t) \tag{2}$$

and the linear signal transformation computed as follows:

$$S(\tau,\sigma) = \langle \mathbf{T}_\tau \mathbf{O}_\sigma g, s\rangle, \tag{3}$$

where $\langle f, g\rangle$ denotes the scalar product in the signal space[1]. We call "time-something" the generic representation whose representative elements are obtained by cascading time-shift with another operator $\mathbf{O}$ acting on a prototype function g.

In the STFT based representations the operator $\mathbf{O}$ corresponds to modulation

$$[\mathbf{M}_\nu s](t) = e^{j2\pi\nu t} s(t), \tag{4}$$

while in the WT based representations the operator $\mathbf{O}$ corresponds to dilation

$$[\mathbf{D}_a s](t) = \frac{1}{\sqrt{a}} s\left(\frac{t}{a}\right), \tag{5}$$

which we wrote here as a unitary operator.

Notice that in some definitions the time-shift and the $\mathbf{O}$ operators appear in reverse order with respect to that in (3). However, with the introduction of a phase change or of a modification of the transform coordinates, in most useful cases one can redefine the transformation as in (3). For example, in the STFT, we have:

$$[\mathbf{T}_\tau \mathbf{M}_\nu g](t) = e^{j2\pi\nu(t-\tau)} g(t-\tau) = e^{-j2\pi\nu\tau} [\mathbf{M}_\nu \mathbf{T}_\tau g](t). \tag{6}$$

[1] Typically this is the space $\mathbb{L}^2(\mathbb{R})$ of square integrable functions (finite energy signals) on the real line, however, the choice might depend on the signal representation we are considering.

The phase factor $e^{-j2\pi\nu\tau}$ is irrelevant, it cancels out when analysis and synthesis is performed and does not need to be included in the computation. For the wavelet transform, one can show that:

$$\mathbf{D}_a\mathbf{T}_{\tau/\alpha} = \mathbf{T}_\tau\mathbf{D}_a, \tag{7}$$

with obvious reinterpretation of scaled time-shift τ/α.

Most of the mentioned representations can be studied in terms of the time and frequency spreading of the transform nucleus

$$K_{\tau,\sigma}(t) = \left[\mathbf{T}_\tau\mathbf{O}_\sigma g\right](t), \tag{8}$$

as the parameters τ and σ, which are the coordinates of the transformed domain, vary. Directly or indirectly (via the relationship between the coordinate σ and the frequency), this leads to a tiling of the time-frequency plane, in which the specific form of the time-frequency tiles carries a loose meaning in terms of the signal components that get detected or trapped in a certain time-frequency zone.

However, the tiling is often dictated by mathematical simplifications for the definition of the operations used in generating the transform nucleus. Thus, for example, in the STFT based time-frequency representations time and frequency resolutions are both uniform and in the time-scale WT based representations time-shift is uniform in each scale. Also, in the WT, scale and time are interrelated so that the uncertainty product of the representative elements is constant throughout the tiling. In the applications often one would like to achieve higher degree of flexibility in allocating the time-frequency spread of the representative elements in order to produce a tiling prescribed, e.g., by psychoacoustic or physical characteristics.

In previous work, frequency warping has been considered in conjunction with signal representations in order to increase the flexibility and adaptivity of the analysis-synthesis scheme [10, 11, 12, 13, 14]. Since frequency warping introduces dispersion in time-localization, methods for reducing or eliminating this effect are desired. In recent work [15, 16, 17], we proposed techniques, the so called redressing methods, that are directly suitable for the STFT based time-frequency representation once the warping map is defined. These basically consist in the application of suitable operators aimed at linearizing the phase of the dispersive delays resulting from the application of warping to the time-shift operator.

Since, over the continuum, a 2D representation of a 1D signal is generally redundant, sampling can be introduced by selecting a grid of points in the transform domain. Provided that certain conditions on the grid and on the window or wavelet are satisfied, the original signal can be reconstructed from the transform samples evaluated at the grid points only. This leads to a signal expansion in terms of a set of functions constituting a frame or otherwise to a bi-orthogonal / orthogonal and complete set for the signal space.

Sampling introduces some complications in the application of the redressing methods and may impose limitations on the choice of the window in generalized Gabor frames considered in [15, 16, 17]. There, discrete-time frequency warping operators were applied in the sampled time STFT domain in order to partially eliminate dispersion. The results show that for bandlimited windows, the dispersion elimination procedure can be made exact and leads to the same results as ad-hoc methods for the construction of non-stationary Gabor frames [18, 19, 20].

Transform domain discrete-time frequency warping also leads to approximated algorithms for the computation of the generalized warped Gabor frames [21, 22, 23]. On the other hand, computation of discrete-time frequency warping as a digital audio effect can be eased by the use of the STFT [24].

In this paper we extend our methods to other generic time-something representations of which wavelet expansions are an example.

The paper is organized as follows: in Section 2 we discuss the concept of unitary equivalence, applying it to unitary warping in Section 3. In Section 4 we discuss redressing methods for the generic unsampled and sampled time-something representations, with examples to the wavelet expansion. Finally, in Section 5 we draw our conclusions.

Examples and experimental code will be made available at the author's web page:
`http://members.chello.at/~evangelista/`.

2. UNITARY EQUIVALENCE

In order to generalize the linear signal representations and be able to adapt the representing elements to important physical or perceptual characteristics, one can resort to further transformation using invertible operators. This is shown in Figure 1 where the analysis and synthesis blocks allow for perfect reconstruction. However, in the scheme in the figure, the synthesis of the signal s is achieved by further operating with an invertible operator $\mathbf{U}$. In other words, if in the original analysis-synthesis scheme the signal was synthesized in terms of a continuous or discrete set of functions $\psi_\alpha(t)$, in the new scheme the signal is synthesized in terms of the functions $\left[\mathbf{U}\psi_\alpha\right](t)$, where α is either the variable or the index of superposition for the synthesis, e.g.,

$$s(t) = \sum_\alpha C_\alpha \left[\mathbf{U}\psi_\alpha\right](t), \tag{9}$$

in which C_α denote the coefficients of the expansion. Here, in order to simplify the notation, in the variable α we incorporate all of the transform coordinates $\alpha_1, ..., \alpha_N$, such as time-shift and frequency or time-shift and scale.

Clearly, in order to provide the correct analysis algorithm, in the new scheme one needs to operate on the signal with the inverse operator $\mathbf{U}^{-1}$ and obtain the analysis coefficients C_α in (9) from the signal $\left[\mathbf{U}^{-1}s\right](t)$. However, in general, perfect reconstruction is not guaranteed and one needs to prove properties like the norm convergence of the expansion on the right hand side of (9) to the signal on the left hand side.

Perfect reconstruction is guaranteed, with no further proof, if the operator $\mathbf{U}$ is unitary, i.e., if $\mathbf{U}$ is a bounded linear operator satisfying

$$\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}, \tag{10}$$

where $\mathbf{I}$ is the identity operator and $\mathbf{U}^\dagger$ represents the adjoint of $\mathbf{U}$, i.e., the operator satisfying

$$\langle \mathbf{U}f, g\rangle = \langle f, \mathbf{U}^\dagger g\rangle \tag{11}$$

for any pair of functions f and g belonging to the signal space. In this case, the operator $\mathbf{U}^\dagger$ is both the left and right inverse of $\mathbf{U}$ and $\mathbf{U}$ realizes a surjective isometry in the function space. Therefore, properties such as orthogonality, norm and norm convergence are preserved, so one can simply invoke unitary equivalence to verify perfect reconstruction [11].

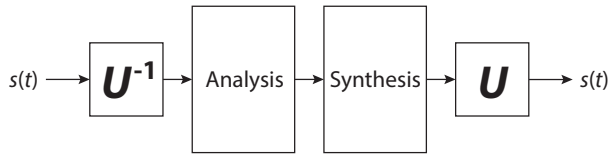

Figure 1: Inserting invertible operators in the analysis-synthesis path of a signal representation

The simple insertion of the invertible operator $\mathbf{U}$ and its inverse in the analysis-synthesis path could, however, lead to dramatic consequences in the interpretation of the transform coordinates. In fact, suppose that the nucleus of the original transform is given as in (8), then the new nucleus is:

$$K_{\tau,\sigma}(t) = [\mathbf{U}\mathbf{T}_\tau\mathbf{O}_\sigma g]\,(t), \tag{12}$$

By inserting neutral pairs $\mathbf{U}^{-1}\mathbf{U}$ in (12), one obtains [11]:

$$K_{\tau,\sigma}(t) = \left[\mathbf{U}\mathbf{T}_\tau\mathbf{U}^{-1}\mathbf{U}\mathbf{O}_\sigma\mathbf{U}^{-1}\mathbf{U}g\right](t), \tag{13}$$

which shows that the new "time" and the new "$\mathbf{O}$" operators are obtained by similarity transformations, i.e., $\mathbf{U}\mathbf{T}_\tau\mathbf{U}^{-1}$ and $\mathbf{U}\mathbf{O}_\sigma\mathbf{U}^{-1}$, respectively, which obviously changes the meaning of the corresponding variables.

3. ALTERNATIVE UNITARY EQUIVALENT REPRESENTATIONS OBTAINED BY WARPING

Appealing signal transformations that can be cascaded to a given analysis-synthesis scheme consist of time or frequency warping, which provide new signal representations that are unitary equivalent to well-known ones, as discussed in the previous section [10, 11, 12, 13, 14].

In this section we illustrate the fundamental notions of warping in connection with time-something representations. As these apply with same formalism to both continuous and discrete (sampled) representations, we will review these concepts without specifying the nature of the representation.

In order to make the concepts more concrete, in Section 3.2 we provide an example of warping map for changing the scale factor in complete wavelet sets obtained by sampling in time-scale.

In Section 4 we provide a solution for the redressing of the warped time-shift operator, in order to eliminate or reduce dispersion resulting from frequency warping. Redressing proves easy in the case of continuous (with respect to time-shift) representations. It requires a bit more ingenuity when sampling is introduced in the transformed space.

By duality arguments, the principles illustrated in this paper also apply to the time dispersion of frequency localization when time warping is employed.

3.1. Warping Operators

Warping operators perform function composition in order to remap the abscissa. Thus, a time warping operator $\mathbf{W}_\gamma$ remaps the time axis by means of the following action:

$$s_{tw} = \mathbf{W}_\gamma s = \mathbf{C}_\gamma s = s \circ \gamma, \tag{14}$$

where s_{tw} is the time-warped version of the signal s, γ is the time warping map, $\mathbf{C}_\gamma$ is the composition by γ operator and $\circ$ denotes function composition.

One can show that the composition operator $\mathbf{C}_\gamma$ is bounded on $\mathbb{L}_2(\mathbb{R})$ if the map is a monotonic function and $1/\gamma'$ is essentially bounded, where γ' is the first derivative of the map, i.e., γ' must be essentially bounded from below. This, together with more general and necessary and sufficient conditions for the continuity of the composition operator, which require the absolute continuity of the measure introduced by the map and the essential boundedness of the Radon-Nikodym derivative, can be found in [25, 26].

Similarly, a frequency warping operator $\mathbf{W}_{\tilde{\theta}}$ is completely characterized by a function composition operator $\mathbf{W}_\theta$ in the frequency domain:

$$\hat{s}_{fw} = \widehat{\mathbf{W}_{\tilde{\theta}}s} = \hat{\mathbf{W}}_{\tilde{\theta}}\hat{s} = \mathbf{C}_\theta\hat{s} = \hat{s} \circ \theta, \tag{15}$$

where θ is the frequency warping map, which transforms the Fourier transform $\hat{s} = \mathcal{F}s$ of a signal s into the Fourier transform $\hat{s}_{fw} = \mathcal{F}s_{fw}$ of another signal s_{fw}, the frequency warped version of the signal s, where $\mathcal{F}$ is the Fourier transform operator and the hat over a symbol denotes the Fourier transformed quantity (signal or operator). We affix the ˜ symbol over the map θ as a reminder that the map operates in the frequency domain. Accordingly, we have $\mathbf{W}_{\tilde{\theta}} = \mathcal{F}^{-1}\hat{\mathbf{W}}_{\tilde{\theta}}\mathcal{F} = \mathcal{F}^{-1}\mathbf{C}_\theta\mathcal{F}$.

For the case of frequency warping, in order to transform real signals to real signals, one has to constrain the map θ to have the odd parity

$$\theta(-\nu) = -\theta(\nu). \tag{16}$$

This way, positive (negative) frequencies are mapped to positive (negative) frequencies. In general, this is not a big constraint as the frequency band allocation of the representation is usually symmetric (same bandwidth in positive and negative frequencies).

In this paper, we consider warping operators that are invertible. One can show [27] that the composition operator (14) is invertible if and only if the map γ is invertible and the composition operator $\mathbf{C}_{\gamma^{-1}}$ based on the inverse map γ^{-1} is bounded on $\mathbb{L}_2(\mathbb{R})$. If the map γ is a monotonic function then also its inverse γ^{-1} is monotonic. Then, a sufficient condition for the invertibility of the composition operator is that also the first derivative of γ^{-1} is essentially bounded from below. Theorem 3 in [27] also shows that in the monotonic case the operator $\mathbf{C}_\gamma$ is invertible if and only if γ' is essentially bounded and is absolutely continuous on the finite intervals in $\mathbb{R}$.

If the warping map is almost everywhere strictly increasing, one-to-one and differentiable then a unitary form of the warping operator can be defined by amplitude scaling, as given by the square root of the magnitude derivative of the map (dilation function). For example, a unitary frequency warping operator $\mathbf{U}_{\tilde{\theta}}$ has frequency domain action

$$\hat{s}_{fw}(\nu) = \left[\widehat{\mathbf{U}_{\tilde{\theta}}s}\right](\nu) = \sqrt{\left|\frac{d\theta}{d\nu}\right|}\hat{s}(\theta(\nu)), \tag{17}$$

where ν denotes frequency. We assume henceforth that all warping maps are almost everywhere increasing so that the magnitude sign can be dropped from the derivative under the square root.

The unitary warping operator (17) is a special case of weighted composition operator [28]. It turns out that the weight function – the derivative of the map in our case – helps releasing some constraints on the maps that guarantee the continuity of the weighted composition operator. A classical example [28] is the map $\gamma(x) = x^2$ in the space $\mathbb{L}^2([0,1])$. The map does not define a bounded composition operator (γ' is not bounded from below).

However, with the weight $\sqrt{\gamma'(x)} = \sqrt{2x}$ one obtains a bounded (unitary) weighted composition operator. More general conditions for the definition of unitary weighted composition operators can be found in [29].

Incidentally, note that the dilation operator introduced in (5) is a particular case of unitary time warping operator $\mathbf{U}_\gamma$ where the warping map $\gamma(t) = t/a$ is linear. By the Fourier scaling theorem, this is also equivalent to a frequency warping unitary operator $\mathbf{U}_{\tilde{\theta}}$ with map $\theta(\nu) = a\nu$.

When a unitary frequency warping operator $\mathbf{U}_{\tilde{\theta}}$ acts as modifier of a time-shift based signal representation, the time-shift operator is affected by a similarity transformation as in (13). By a calculation similar to the one conducted in the Appendix one can conclude that

$$\left[\mathbf{U}_{\tilde{\theta}}\mathbf{T}_\tau\mathbf{U}_{\tilde{\theta}}^{-1}s\right](t) = \int_{-\infty}^{+\infty} d\nu e^{j2\pi\nu t} e^{-j2\pi\theta(\nu)\tau}\hat{s}(\nu), \quad (18)$$

which shows how, by the effect of warping, the delay τ converts into a dispersive delay represented, in the frequency domain, by the factor $e^{-j2\pi\theta(\nu)\tau}$.

3.2. Example: Warping Map for Change of Scale in Sets of Wavelet

Sets of wavelets for signal expansions are obtained by choosing a grid of points for sampling the WT in time-scale space of the type $(a^n m, a^n)$, where m and n are integers and a is a constant scale factor. It is well known that the dyadic scheme, which fixes $a = 2$ and octave band resolution, is the easiest way to generate orthogonal and complete sets of wavelets. For sound and music processing scopes, it is highly interesting to improve the frequency resolution, for example to half-tones. One can achieve this by means of a warping map θ that allows for a change of the scale factor from a to $b < a$. In other words, we are requiring the map to satisfy:

$$a\theta(\nu) = \theta(b\nu), \quad (19)$$

so that, by repeated applications of (19) one can show that

$$\sqrt{\frac{d\theta}{d\nu}}\sqrt{a^n}\hat{\psi}\left(a^n\theta(\nu)\right) = \sqrt{b^n}\sqrt{\frac{d\theta}{d\nu}}\hat{\psi}\left(\theta(b^n\nu)\right) \quad (20)$$

where $\hat{\psi}$ is the Fourier transform of a scale-a mother wavelet. This shows that the frequency warped wavelet $\tilde{\psi}$, whose Fourier transform is

$$\hat{\tilde{\psi}}(\nu) = \sqrt{\frac{d\theta}{d\nu}}\hat{\psi}\left(\theta(\nu)\right) \quad (21)$$

properly scales through powers of b to the warped versions of the a-scaled versions of the original mother wavelet ψ:

$$\sqrt{b^n}\hat{\tilde{\psi}}(b^n\nu) = \sqrt{\frac{d\theta}{d\nu}}\sqrt{a^n}\hat{\psi}\left(a^n\theta(\nu)\right). \quad (22)$$

A particular map that satisfies the generalized homogeneity condition (19) is given by the exp-log function:

$$\theta(\nu) = C\,\mathrm{sgn}\,(\nu)a^{\log_b|\nu|}, \quad (23)$$

where $\mathrm{sgn}\,(\nu)$ is the signum function and C is an arbitrary constant that can be set, for example, by constraining the warping map to fix a specific frequency. This capability is important in sampled

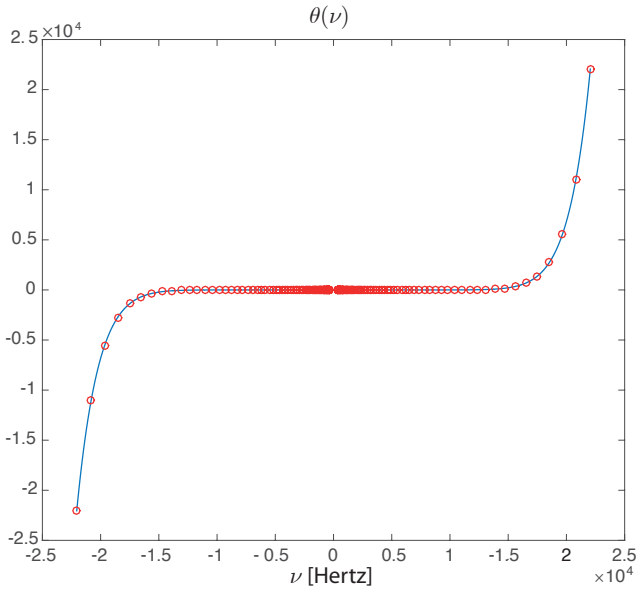

Figure 2: Frequency warping map to convert octave-band to halftone-band wavelets. The circles in the plot denote the halftone mapping points.

systems in order to preserve the sampling rate, in which case we set:

$$C = \nu_N a^{-\log_b \nu_N}, \quad (24)$$

where ν_N is the Nyquist frequency, so that $\theta(\nu_N) = \nu_N$ and the overall bandwidth is preserved.

Notice that for the map in (23) we have:

$$\frac{d\theta}{d\nu} = C\log_b(a)\frac{a^{\log_b|\nu|}}{|\nu|}, \quad (25)$$

so that, for the case of interest where $a > b > 1$, which corresponds to higher frequency resolution warped wavelets, we have $\lim_{\nu\to 0}\theta'(\nu) = 0$, so that θ' is not essentially bounded from below and therefore the corresponding composition operator may not be bounded. However, in representations for sound and music signals, we are not interested in extreme accuracy around zero frequency. In order to obtain a bounded composition operator, one can therefore replace the warping map (23) with one that is identical to it except that, on a small interval around zero frequency, the original map is replaced by a linear segment continuously going from 0 to the value of the original map at the extremes of the interval. Warping with this map will still yield an exact representation, with the desired frequency band allocation.

A map suitable for the conversion of octave-band wavelets into halftone-band wavelets is shown in Figure 2. On a log-log scale the map would appear as linear. The Fourier transforms of the corresponding frequency warped wavelets based on an octave-band pruned tree-structured FIR Daubechies filter bank [3] of order 31 are shown in Figure 3.

The dispersion pattern of 1/3-octave resolution frequency warped wavelets is shown in Figure 4, where each area delimited by the band edge limits (horizontal lines) and two adjacent group delay curves roughly represents the time-frequency resolution or, more precisely, the uncertainty zone of the corresponding

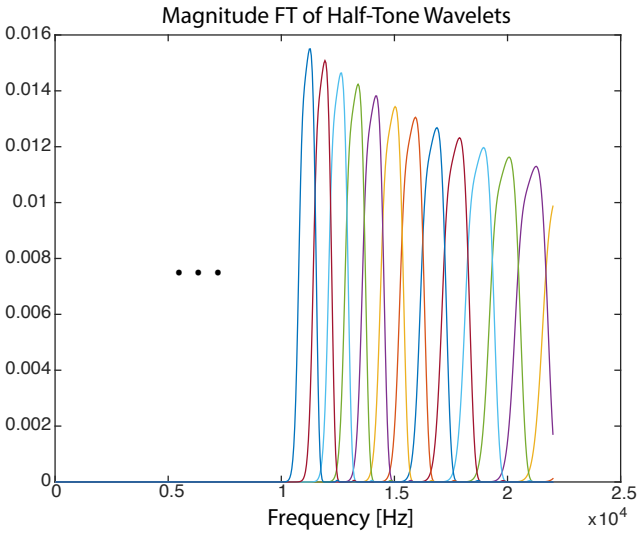

Figure 3: Magnitude Fourier transform of frequency warped wavelets with halftone frequency resolution.

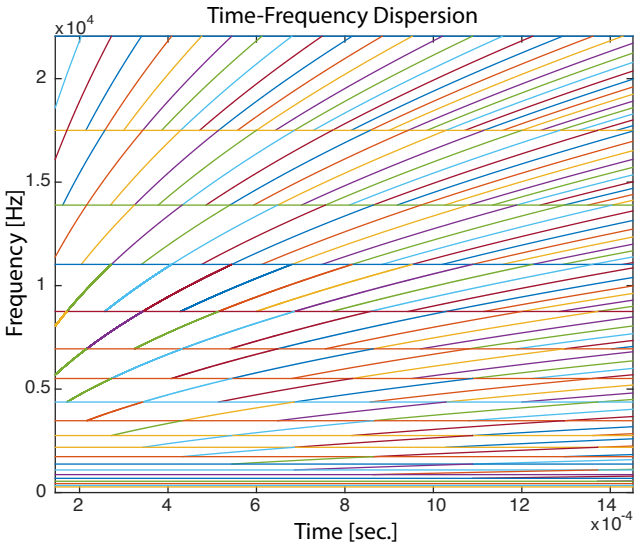

Figure 4: Uncertainty zones delimited by the group delay dispersion profiles of 1/3-octave warped wavelets without redressing.

wavelet. The fact that the time boundaries are curved is not good news for the time localization of signal components or features. Fortunately, redressing methods, which we will illustrate in the remainder of the paper, are available for the elimination or reduction of dispersion.

4. REDRESSING

As shown in Section 3.1, frequency warping results in a similarity transformation of the time-shift operator, which introduces frequency dependency, i.e., the dispersive time-shift in (18). Once established, the result proven in the Appendix that

$$\mathbf{U}_{\tilde{\theta}}^{(\tau)}\mathbf{U}_{\tilde{\theta}}\mathbf{T}_{\tau} = \mathbf{T}_{\tau}\mathbf{W}_{\tilde{\theta}}, \tag{26}$$

provides a simple way to counteract dispersion.

Here the operator $\mathbf{U}_{\tilde{\theta}}^{(\tau)}$ is a genuine warping operator as discussed in Section 3.1, however, it operates in the transform domain, i.e., with respect to the time-shift coordinate instead of the signal's time. Basically, by performing a supplementary frequency warping operation, denoted by $\mathbf{U}_{\tilde{\theta}}^{(\tau)}$, with respect to the time-shift coordinate in the transform domain, one obtains a commutation rule of the time-shift operator with the frequency warping operator. The latter occurs in both its unitary $\mathbf{U}_{\tilde{\theta}}$ and non-unitary $\mathbf{W}_{\tilde{\theta}}$ versions in the commutation rule (26).

Equation (26) shows that the pure time-shifting of the warped nucleus of the transform leads to a perfect reconstruction scheme, which is equivalent to operating in both signal and transform domain with unitary operators. The insertion of signal frequency warping operators $\mathbf{U}_{\tilde{\theta}}$ and $\mathbf{U}_{\tilde{\theta}}^{\dagger}$ for frequency remapping and frequency warping operators $\mathbf{U}_{\tilde{\theta}}^{(\tau)}$ and $\mathbf{U}_{\tilde{\theta}}^{(\tau)\,\dagger}$ in the transform domain for the computation of a time-something representation is shown in Figure 5.

We remark that the scheme illustrated in the picture is a conceptual one, as redressing may provide a great simplification, which makes the online computation of the warping operators unnecessary. In fact, once we apply the redressing operator $\mathbf{U}_{\tilde{\theta}}^{(\tau)}$,

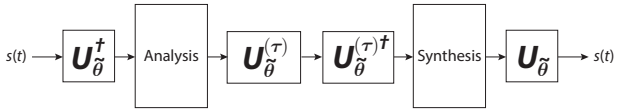

Figure 5: Inserting frequency warping operators for both frequency remapping and redressing in the analysis-synthesis path of a signal representation

the frequency warped nucleus of the transform may become shift-invariant or approximately so. Thus, in the computation of the transform, one can compute the scalar products of the signal with time-shifted versions of the frequency warped modulated prototype window or dilated mother wavelet.

4.1. Redressing Sampled Time-Something Representations

In most time-something representations the time-shift, as well as other variables, can be sampled, while preserving the capability of perfectly reconstructing the signal. Sampling in time-frequency the STFT leads to the concept of Gabor frames [30], while sampling the WT leads to orthogonal / biorthogonal wavelets or to wavelet frames [2, 3]. However, since frequency warping does not commute with the time sampling operator, the redressing procedure does not carry over to the sampled transforms in an exact way.

Once time sampling is performed in the transform domain, the time-shift operator $\mathbf{T}_{\tau}$ becomes $\mathbf{T}_{nT}$. Thus, we cannot apply the operator $\mathbf{U}_{\tilde{\theta}}^{(\tau)}$ with respect to the variable τ in order to eliminate the dispersive delays as we did in (26). Instead, we can try to apply a discrete version of frequency warping operator acting on sampled time, i.e. on the index n.

A discrete-time frequency warping operator can be built from an almost everywhere differentiable warping map ϑ that is one-to-one and onto $[-\frac{1}{2}, +\frac{1}{2}[$. In this case, one can form an orthonormal

basis of $\ell^2(\mathbb{Z})$ as follows:

$$\mu_m(n) = \int_{-\frac{1}{2}}^{+\frac{1}{2}} \sqrt{\tfrac{d\vartheta}{d\nu}} e^{j2\pi(n\vartheta(\nu)-m\nu)} d\nu, \tag{27}$$

where $n, m \in \mathbb{Z}$ (see [31, 32, 14, 12, 13]).

The map ϑ can be extended over the entire real axis as congruent modulo 1 to a 1-periodic function. One can always write the map in terms of a 1-periodic function plus an integer staircase function, so that

$$\vartheta(\nu + k) = \vartheta(\nu) + k \tag{28}$$

for any $k \in \mathbb{Z}$.

Similarly to the continuous counterpart, one can define the action of the discrete-time unitary warping operator $\mathbf{V}_{\tilde{\vartheta}}$ on any sequence $\{x(n)\}$ in $\ell^2(\mathbb{Z})$ as follows:

$$\left[\mathbf{V}_{\tilde{\vartheta}} x\right](m) = \langle \mu_m, x \rangle_{\ell^2(\mathbb{Z})}. \tag{29}$$

Given $\tilde{x}(m) = \left[\mathbf{V}_{\tilde{\vartheta}} x\right](m)$, in the frequency domain one obtains

$$\hat{\tilde{x}}(\nu) = \sqrt{\tfrac{d\vartheta}{d\nu}} \hat{x}(\vartheta(\nu)), \tag{30}$$

where here the $\hat{}$ symbol denotes discrete-time Fourier transform. The sequences $\overline{\eta_m(n)}$, where the overbar symbol $\overline{x}$ denotes complex conugation, define the nucleus of the inverse unitary frequency warping $\ell^2(\mathbb{Z})$ operator $\mathbf{V}_{\tilde{\vartheta}^{-1}} = \mathbf{V}_{\tilde{\vartheta}}^{\dagger}$. where $\eta_m(n) = \mu_n(m)$.

Discrete-time frequency warping operators were applied in the sampled time STFT domain in order to partially eliminate dispersion in [15, 16, 17]. We showed that complete elimination of dispersion, i.e., full linearization of the non-linear phase resulting form the application of frequency warping to the time-shift operator is possible in the case of bandlimited analysis window. In the more general case of non-bandlimited window we showed that one can resort to high-performance approximations [23, 21].

In order to extend the redressing methods to other time-something representations, we assume that also the other transform domain coordinate(s) are being sampled. For example, in the case of dyadic wavelets, the basis elements are obtained by time-shifting a scaled version of the mother wavelet $\psi_r(t) = \sqrt{2^{-r}}\psi\left(\frac{t}{2^r}\right)$, in which the scale variable has been exponentially sampled to powers of 2.

In order to simplify our notation, we assume that the redressing method is to be applied to basis or frame elements denoted by ψ_r, which are indexed just by one index r like in the wavelet example. Moreover, we assume for sake of generality that the time-shift factor T_r depends on the index r. Thus, redressing can be achieved by means of a collection of discrete-time frequency warping operators $\mathbf{V}_{\tilde{\vartheta}_r}^{(m)}$, each acting on the frequency warped time-shifted representation element of index r:

$$\tilde{\psi}_{r,m} = \mathbf{V}_{\tilde{\vartheta}_r}^{(m)} \mathbf{U}_{\tilde{\theta}} \mathbf{T}_{mT_r} \psi_r, \tag{31}$$

where the superscript (m) in the discrete-time warping operator is a reminder that it operates with respect to the time-shift index m.

Exploiting the odd parity (16) of the $\theta(\nu)$ map and (28), one can show that (31) becomes:

$$\tilde{\psi}_{r,m}(t) = \int_{-\infty}^{+\infty} d\nu A_r(\nu) e^{j2\pi\left(\theta^{-1}(\varphi_r(\nu))t - m\nu\right)} \hat{\psi}_r\left(\varphi_r(\nu)\right), \tag{32}$$

where

$$\begin{aligned} \varphi_r(\nu) &= \frac{\vartheta_r(\nu)}{T_r} \\ A_r(\nu) &= \sqrt{\frac{d\theta^{-1}\left(\varphi_r(\nu)\right)}{T_r d\nu}}. \end{aligned} \tag{33}$$

From the phase in (32) one can conclude that if, for some constant $\tilde{T}_r > 0$, we had

$$\theta^{-1}(\varphi_r(\nu)) = \frac{\nu}{\tilde{T}_r} \tag{34}$$

then, since by the invertibility of the map θ the last equation is equivalent to

$$\varphi_r\left(\nu\tilde{T}_r\right) = \theta(\nu), \tag{35}$$

we would have

$$\tilde{\psi}_{r,m}(t) = \sqrt{\frac{\tilde{T}_r}{T_r}} \int_{-\infty}^{+\infty} d\nu e^{j2\pi\nu\left(t - m\tilde{T}_r\right)} \hat{\psi}_r\left(\theta(\nu)\right). \tag{36}$$

This shows that, for each r, the redressed warped representation elements could be obtained from the non-unitarily warped original elements just by altered time-shift $\mathbf{T}_{m\tilde{T}_r}$. This would be the discrete counterpart of the unsampled result described in the previous section. However, as we previously remarked, since the maps ϑ_r are constrained to be congruent modulo 1 to a 1-periodic function with odd parity, their shape can only be arbitrarily assigned on a band of width 1/2.

Collecting (33) and (35) together we have the following condition on the discrete-time warping map ϑ:

$$\vartheta\left(\nu\tilde{T}_r\right) = T_r\theta(\nu). \tag{37}$$

Generally, condition (37) cannot be satisfied everywhere but only in a small bandwidth of size $1/2\tilde{T}_r$. However, $\tilde{T}_r$ is a design parameter, which can be selected at will, with the trade-off that it also affects the final sampling rate of the transform. If the warped wavelet, whose Fourier transform is $\hat{\psi}_r\left(\theta(\nu)\right)$, is strictly bandlimited to a sufficiently small band, then it is only necessary to satisfy (37) in its bandwidth. In the general case, one can satisfy (37) within the essential bandwidth of the warped wavelet. We remark that if the complete chain of warping operators defining the warped transform and the redressing method is computed, the partially redressed warped transform has perfect reconstruction.

As a design example, in the dyadic wavelet case choose $T_r = 2^r T$, where $1/T$ is a reference frequency given by the upper cutoff frequency of the mother wavelet. In designing warped wavelet transforms one generally desires to improve the octave band frequency resolution of the dyadic wavelets. Changing to smaller bandwidths for the warped wavelets, and suitably choosing $\tilde{T}_r$, results in a rescaling of the frequency interval of the ϑ_r map for the linearization of the phase within the essential bandwidths of the wavelets, i.e., where the wavelets have essentially non-zero magnitude Fourier transform. This is shown in Figure 6 for a warping map carrying dyadic octave-band wavelets to 1/3 octave resolution. In this example, with reference to the map in Section 3.2, we chose $a = 2$ and $b = 2^{1/3}$, which provides warped wavelets with bandwidths $(b-1)/2Tb^r$, with cutoff frequencies $1/2Tb^r$, $r = 1, 2, \ldots$. Thus, one can select $\tilde{T}_r = Tb^r/(b-1)$ in order to define the branch of the map ϑ_r according to (37), shown, for the case $r = 2$, in the highlighted (red) portion of the map in Figure 6. For any r, this choice defines ϑ_r on a width $1/2$ frequency band

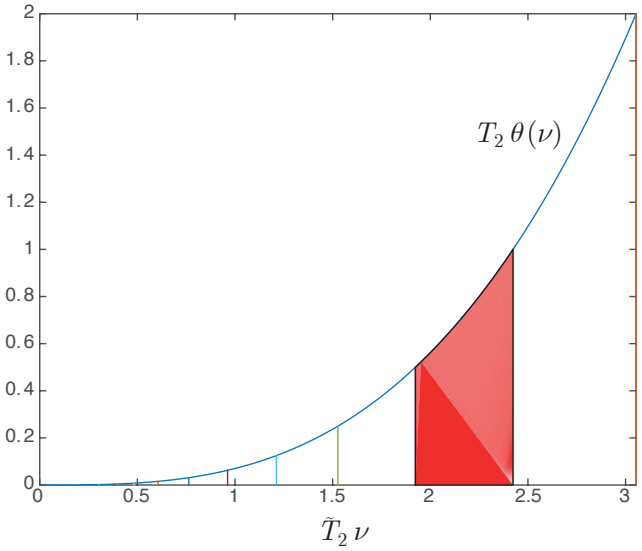

Figure 6: Segment of the map ϑ_2 for the redressing of 1/3-octave warped wavelets obtained from dyadic ones.

with total variation 1/2. The remaining part of each map ϑ_r is then extended by odd symmetry about 0 frequency and modulo 1 periodicity. Actually, in this specific example, in view of property (19), the maps are the same for any r.

Of course, in general, the detailed approximation margins depend on the particular time-something representation at hand. However, the flexibility in the choice of the sampling of the time-shift variable in the transform domain can lead to good results as in the case of generalized Gabor frames [23, 21].

5. CONCLUSIONS

In this paper we have revisited recently introduced redressing methods for mitigating dispersion in warped signal representations, with the objective of generalizing the procedure to generic time-something representations like wavelet expansions.

We showed that redressing methods can be extended to generalized settings in both unsampled and sampled transform domains, introducing notation to simplify their application to signal analysis-synthesis scheme.

The phase linearization techniques lead to approximated schemes for the computation of the warped and redressed signal representations, where one can pre-warp the analysis window or mother wavelet and compute the transform without the need for further warping.

6. REFERENCES

[1] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Applied and Numerical Harmonic Analysis. Birkhäuser Boston, 2001.

[2] Stephane Mallat, *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*, Academic Press, 3rd edition, 2008.

[3] Ingrid Daubechies, *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.

[4] J. P. Princen, A. W. Johnson, and A. B. Bradley, "Subband transform coding using filter bank designs based on time domain aliasing cancellation," in *IEEE Proc. Intl. Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1987, pp. 2161–2164.

[5] H. S. Malvar, "Modulated QMF filter banks with perfect reconstruction," *Electronics Letters*, vol. 26, no. 13, pp. 906–907, June 1990.

[6] S. Ghobber and S. Omri, "Time-frequency concentration of the windowed Hankel transform," *Integral Transforms and Special Functions*, vol. 25, no. 6, pp. 481–496, 2014.

[7] C. Baccar, N.B. Hamadi, and H. Herch, "Time-frequency analysis of localization operators associated to the windowed Hankel transform," *Integral Transforms and Special Functions*, vol. 27, no. 3, pp. 245–258, 2016.

[8] J. Brown, "Calculation of a constant Q spectral transform," *Journal of the Acoustical Society of America*, vol. 89, no. 1, pp. 425–434, January 1991.

[9] Y.A. Abramovich and C.D. Aliprantis, *An Invitation to Operator Theory*, vol. 50 of *Graduate studies in mathematics*, American Mathematical Society, 2002.

[10] A. V. Oppenheim, D. H. Johnson, and K. Steiglitz, "Computation of spectra with unequal resolution using the Fast Fourier Transform," *Proc. of the IEEE*, vol. 59, pp. 299–301, Feb. 1971.

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

[12] G. Evangelista and S. Cavaliere, "Frequency Warped Filter Banks and Wavelet Transform: A Discrete-Time Approach Via Laguerre Expansions," *IEEE Transactions on Signal Processing*, vol. 46, no. 10, pp. 2638–2650, Oct. 1998.

[13] G. Evangelista and S. Cavaliere, "Discrete Frequency Warped Wavelets: Theory and Applications," *IEEE Transactions on Signal Processing*, vol. 46, no. 4, pp. 874–885, Apr. 1998, special issue on Theory and Applications of Filter Banks and Wavelets.

[14] G. Evangelista, "Dyadic Warped Wavelets," *Advances in Imaging and Electron Physics*, vol. 117, pp. 73–171, Apr. 2001.

[15] G. Evangelista, M. Dörfler, and E. Matusiak, "Phase vocoders with arbitrary frequency band selection," in *Proceedings of the 9th Sound and Music Computing Conference*, Copenhagen, Denmark, 2012, pp. 442–449.

[16] G. Evangelista, M. Dörfler, and E. Matusiak, "Arbitrary phase vocoders by means of warping," *Music/Technology*, vol. 7, no. 0, 2013.

[17] G. Evangelista, "Warped Frames: dispersive vs. non-dispersive sampling," in *Proceedings of the Sound and Music Computing Conference (SMC-SMAC-2013)*, Stockholm, Sweden, 2013, pp. 553–560.

[18] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. A. Velasco, "Theory, implementation and applications of nonstationary Gabor Frames," *Journal of Computational and Applied Mathematics*, vol. 236, no. 6, pp. 1481–1496, 2011.

[19] G. A. Velasco, N. Holighaus, M. Dörfler, and T. Grill, "Constructing an invertible constant-Q transform with non-stationary Gabor frames," in *Proceedings of the Digital Audio Effects Conference (DAFx-11)*, Paris, France, 2011, pp. 93–99.

[20] N. Holighaus and C. Wiesmeyr, "Construction of warped time-frequency representations on nonuniform frequency scales, Part I: Frames," *ArXiv e-prints*, Sept. 2014.

[21] T. Mejstrik and G. Evangelista, "Estimates of the Reconstruction Error in Partially Redressed Warped Frames Expansions," in *Proceedings of the Digital Audio Effects Conference (DAFx-16)*, Brno, Czech Republic, 2016, pp. 9–16.

[22] T. Mejstrik, "Real Time Computation of Redressed Frequency Warped Gabor Expansion," Master thesis, University of Music and Performing Arts Vienna, tommsch.com/science.php, 2015.

[23] G. Evangelista, "Approximations for Online Computation of Redressed Frequency Warped Vocoders," in *Proceedings of the Digital Audio Effects Conference (DAFx-14)*, Erlangen, Germany, 2014, pp. 1–7.

[24] G. Evangelista and S. Cavaliere, "Real-time and efficient algorithms for frequency warping based on local approximations of warping operators," in *Proceedings of the Digital Audio Effects Conference (DAFx-07)*, Bordeaux, France, Sept. 2007, pp. 269–276.

[25] R.K. Singh and Ashok Kumar, "Characterizations of invertible, unitary, and normal composition operators," *Bulletin of the Australian Mathematical Society*, vol. 19, pp. 81–95, 1978.

[26] R.K. Singh and J.S. Manhas, *Composition Operators on Function Spaces*, vol. 179, North Holland, 1993.

[27] R.K. Singh, "Invertible composition operators on $\mathbb{L}^2(\lambda)$," *Proceedings of the American Mathematical Society*, vol. 56, no. 1, pp. 127–129, April 1976.

[28] James T. Campbell and James E. Jamison, "On some classes of weighted composition operators," *Glasgow Mathematical Journal*, vol. 32, no. 1, pp. 87–94, 1990.

[29] V.K. Pandey, *Weighted Composition Operators and Representations of Transformation Groups*, Ph.D. thesis, University of Lucknow, India, 2015.

[30] O. Christensen, "Frames and pseudo-inverses," *Journal of Mathematical Analysis and Applications*, vol. 195, no. 2, pp. 401–414, 1995.

[31] P. W. Broome, "Discrete orthonormal sequences," *Journal of the ACM*, vol. 12, no. 2, pp. 151–168, Apr. 1965.

[32] L. Knockaert, "On Orthonormal Muntz-Laguerre Filters," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, pp. 790 –793, apr 2001.

7. APPENDIX: PROOF THAT $\mathbf{U}_{\tilde{\theta}}^{(\tau)}\mathbf{U}_{\tilde{\theta}}\mathbf{T}_\tau = \mathbf{T}_\tau\mathbf{W}_{\tilde{\theta}}$

In this Appendix we prove an important result on which the redressing methods for the warped time-shift operators are based. We show that the application of the same warping operator $\mathbf{U}_{\tilde{\theta}}$ but with respect to time-shift τ instead of time t, which we denote by $\mathbf{U}_{\tilde{\theta}}^{(\tau)}$, results in the commutation of the warping operator $\mathbf{U}_{\tilde{\theta}}$ with the time-shift operator $\mathbf{T}_\tau$ in its non-unitary form.

We are going to prove our result under the assumption that the map θ is almost everywhere increasing and has odd parity:

$$\theta(-\nu) = -\theta(\nu). \tag{38}$$

In this case, the first derivative

$$\theta'(\nu) = \frac{d\theta}{d\nu} \tag{39}$$

is almost everywhere positive and has even parity:

$$\theta'(-\nu) = \theta'(\nu). \tag{40}$$

From (17) we have:

$$\left[\mathbf{U}_{\tilde{\theta}}s\right](t) = \int_{-\infty}^{+\infty} d\alpha K_\theta(t,\alpha)s(\alpha) \tag{41}$$

where

$$\begin{aligned} K_\theta(t,\alpha) &= \int_{-\infty}^{+\infty} d\nu \sqrt{\frac{d\theta}{d\nu}} e^{j2\pi(\nu t - \theta(\nu)\alpha)} \\ &= \int_{-\infty}^{+\infty} d\nu \sqrt{\frac{d\theta^{-1}}{d\nu}} e^{j2\pi(\theta^{-1}(\nu)t - \nu\alpha)} \end{aligned} \tag{42}$$

where we have used the fact that the map θ is invertible almost everywhere, so that

$$\theta^{-1}(\theta(\nu)) = \nu \tag{43}$$

far almost all $\nu \in \mathbb{R}$, from which it follows that

$$1 = \frac{d\left[\theta^{-1}(\theta(f))\right]}{df} = \left.\frac{d\theta^{-1}}{d\alpha}\right|_{\alpha=\theta(f)} \frac{d\theta}{df} \tag{44}$$

almost everywhere.

Thus, given any finite energy signal $s(t)$, we have

$$\begin{aligned} &\left[\mathbf{U}_{\tilde{\theta}}^{(\tau)}\mathbf{U}_{\tilde{\theta}}\mathbf{T}_\tau s\right](t) = \\ &= \int_{-\infty}^{+\infty} d\alpha K_\theta(\tau,\alpha) \int_{-\infty}^{+\infty} d\beta K_\theta(t,\beta) s(\beta-\alpha) \\ &= \int_{-\infty}^{+\infty} d\alpha K_\theta(\tau,\alpha) \int_{-\infty}^{+\infty} d\eta K_\theta(t,\eta+\alpha) s(\eta) \end{aligned} \tag{45}$$

Using (42), by direct calculation, it is easy to show that

$$\int_{-\infty}^{+\infty} d\alpha K_\theta(\tau,\alpha) K_\theta(t,\eta+\alpha) = L_\theta(t-\tau,\eta), \tag{46}$$

where $L_\theta(t,\alpha)$ is the nucleus of the non-unitary warping operator $\mathbf{W}_{\tilde{\theta}}$, i.e.,

$$\begin{aligned} L_\theta(t,\alpha) &= \int_{-\infty}^{+\infty} d\nu e^{j2\pi(\nu t - \theta(\nu)\alpha)} \\ &= \int_{-\infty}^{+\infty} d\nu \frac{d\theta^{-1}}{d\nu} e^{j2\pi(\theta^{-1}(\nu)t - \nu\alpha)}. \end{aligned} \tag{47}$$

Thus, (45) becomes

$$\begin{aligned} \left[\mathbf{U}_{\tilde{\theta}}^{(\tau)}\mathbf{U}_{\tilde{\theta}}\mathbf{T}_\tau s\right](t) &= \int_{-\infty}^{+\infty} d\eta L_\theta(t-\tau,\eta) s(\eta) \\ &= \left[\mathbf{T}_\tau\mathbf{W}_{\tilde{\theta}}s\right](t), \end{aligned} \tag{48}$$

which is what we needed to prove.