

*This article is an accepted version. Please cite the published version: [10.1007/s11760-020-01701-8](https://doi.org/10.1007/s11760-020-01701-8)*

# Bridging the gap between the Short-Time Fourier Transform (STFT), Wavelets, the Constant-Q Transform and Multi-resolution STFT

(\*) Carlos Mateo<sup>1</sup>, Juan Antonio Talavera<sup>2</sup>

**Abstract - The Short-Time Fourier Transform (STFT) is extensively used to convert signals from the time-domain into the time-frequency domain. However, the standard STFT has the drawback of having a fixed window size. Recently, we proposed a variant of that transform which fixes the window size in the frequency domain (STFT-FD). In this paper, we revisit that formulation, showing its similarity to existing techniques. Firstly, the formulation is revisited from the point of view of the STFT and some improvements are proposed. Secondly, the Continuous Wavelet Transform (CWT) equation is used to formulate the transform in the continuous time using wavelet theory and to discretize it. Thirdly, the constant-Q transform (CQT) is analyzed showing the similarities in the equations of both transforms, and the differences in terms of how the sweep is carried out is discussed. Fourthly, the analogies with multi-resolution STFT are analyzed. Finally, the representations of a period chirp and an electrocardiogram signal in the time-frequency domain and the time-scale domain are obtained and used to compare the different techniques. The analysis in this paper shows that the proposed transform can be expressed as a variant of STFT, and as an alternative discretization of the CWT. It could also be considered a variant of the CQT and a special case of multi-resolution STFT.**

**Index terms - Wavelets, Short-Time Fourier Transform, constant-Q transform, time-frequency, time-scale, electrocardiogram.**

## 1. INTRODUCTION

The Short-Time Fourier Transform (STFT) can be applied to convert a signal from the time-domain into the time-frequency domain. It has been used to process signals in many research areas, for example in image processing [1], speech [2], engineering [3, 4], biology and medicine [5]. The STFT can be used to analyze non-stationary signals, determining how the spectral content of signals changes over time. This transform localizes the signal in time using a window. However, the standard STFT transform has the disadvantage of using a fixed window size. On one hand, long windows have better frequency resolution but poor time resolution. On the other hand, short windows provide better

time resolution but lower frequency resolution [6]. Several alternatives can be used to improve this transform, such as adaptive STFT [7, 8] and multi-resolution STFT [9–11]. Adaptive techniques adjust the window size depending on local signal characteristics. Therefore, with adaptive STFT, different window sizes are used for different time instants. This allows us to use different time-frequency resolutions depending on how the signal evolves over time. In multi-resolution STFT the signal is divided into frequency bands, and each of them is processed with a different window size [9, 10]. This allows us to combine the results of several STFTs to obtain the resulting time-frequency representation (TFR). In this technique different window sizes are used for different frequencies. The transform proposed here is closer to multi-resolution STFT than to adaptive STFT. The similarities and differences of the proposed transform and multi-resolution STFT are discussed in this paper. The multi-scale STFT is a more complex technique that combines both approaches by modifying the window size both in the time and in the frequency domain [12]. Multi-scale STFT allows us to dynamically detect the degree of transience of individual signal components and to automatically process them with adequate time-frequency resolution.

Other alternatives to the Fourier Transform have been proposed. For example, the Fractional Fourier Transform (FRFT) is a generalization of the conventional Fourier analysis which aims to capture the characteristics of functions in the mixed time and frequency domains, called fractional domain [13]. It can be interpreted as a rotation in the time-frequency plane. This domain has applications, for example, in the generalized sampling expansion [14]. The transform is more adequate for processing linear frequency type non-stationary functions. The Short-Time Fractional Fourier Transform is a windowed version of the FRFT [15], which is more adequate for representing strongly non-linear chirp functions [16]. Linear canonical transforms (LCTs) are a generalization of other transforms such as Fourier, fractional Fourier and Laplace. They are a three-parameter family of integral transforms [17], that are a power tool for optics and signal processing, lattice sampling being analyzed in [18]. The offset linear canonical transform has also been proposed, as a time-shifted and frequency modulated version of the LCT with six parameters [19, 20].

<sup>1</sup> Carlos Mateo is with the Institute for Research in Technology, School of Engineering (ICAI), Universidad Pontificia Comillas, Madrid, 28015, Spain. ORCID 0000-0001-9400-8502.

(\*) Carlos Mateo is the corresponding author. Email: [cmateo@comillas.edu](mailto:cmateo@comillas.edu)

<sup>2</sup> Juan A. Talavera is with Esdras A., S.L., Las Rozas de Madrid, Spain.

Wavelets provide an alternative framework that can be applied to various signal processing applications [21]. The Continuous Wavelet Transform (CWT), which is based on wavelet analysis, can be also used as an alternative to STFT in order to obtain a TFR [22, 23]. Wavelets are also used in image processing [24], engineering [25] and medicine [26, 27]. The fidelity factor,  $Q$ , is the inverse of the relative bandwidth [28]. This factor remains constant in wavelets [28]. Wavelets and STFT are linear TFRs, satisfying the superposition or linearity principle [29]. In contrast to the standard STFT which uses a single window size, the Wavelet Transform (WT) uses short windows at high frequencies and long windows at low frequencies [21]. Wavelets rely on the use of a mother wavelet function that can be scaled and shifted, to correlate with the anomalies or events of the signals. In the CWT, the parameters of the wavelets (time and scale) are continuous. When the parameters of the wavelets are discrete, this leads to a Discrete Series Expansion. Finally, the Discrete Wavelet Transform (DWT) can also be used for discrete signals. For non-stationary signals, the conversion into a two-dimensional space allows representing how the properties of the signals evolve over time. The analysis of non-stationary signals was traditionally performed in the time-frequency domain using techniques such as the STFT. Wavelets also allow us to analyze signals in the time-scale domain [30]. Similar to the FRFT, the Fractional Wavelet Transform has also been proposed in wavelet theory [31].

The Constant- $Q$  transform (CQT) can also be used to transform a discrete time domain signal into the time-frequency domain [32, 33]. This transform was initially proposed for music analysis, transforming against  $\log(\text{frequency})$  to obtain a constant pattern in the frequency domain. This transform can be expressed as a series of logarithmically spaced filters. In the CQT transform, the frequency bins are geometrically spaced, and their  $Q$  factor, which is the number of integer cycles processed, remains constant. Traditionally, one of the reasons for the unpopularity of this transform is the excessive computational complexity [34], which is partially mitigated by the increased computing capacity of modern computers. Besides, working with the data structure of the CQT is more complex than working with the time-frequency matrix of the STFT. This transform has already been applied to the analysis of electroencephalograms [35].

In [36, 37] we presented the Short-Time Fourier Transform with the Window Size Fixed in the Frequency Domain (STFT-FD). In this paper we improve the formulation of that transform and formulate it in continuous time showing how it bridges the gap between the Short-Time Fourier Transform, Wavelets, the Constant- $Q$  Transform and Multi-resolution STFT. Thus in Section 2, we revisit the transform in the context of the STFT, improving the proposed formulation. In Section 3, the STFT-FD is formulated in continuous time using wavelet theory. In Section 4, the similarities and differences with the constant- $Q$  transform are discussed. In Section 5, it is shown that we can also consider the transform as a special case of multi-resolution STFT. Section 6 applies the proposed transform to a period chirp and to an electrocardiogram (ECG) signal, representing them in the time-frequency and time-scale domains, and

comparing the results. Finally, Section 7 summarizes the main conclusions.

## 2. REVISION OF THE FORMULATION AS A STFT

The proposed STFT-FD is defined in [36] following the standard STFT methodology. First, a window is applied around every time instant. Then, the Discrete Fourier Transform (DFT) is computed to obtain each frequency component. The only difference with the standard STFT is that instead of fixing the windows size in the time domain, it is determined by the number of cycles inside the window function, and therefore in this transform, the window size depends on the frequency. In this paper, we propose the following modifications to our initial formulation:

- We normalize the transform to have the same energy at every frequency. In the Fourier Transform or in the standard STFT, each sinusoid is integrated during the same range and therefore all sinusoids have the same energy. However, in STFT-FD the window size is different for each frequency. Therefore, we need to use an energy normalization factor. As justified in Section 3, we propose to divide by the square root of the window size, expressed in number of samples.
- Instead of using a definition of the Discrete Fourier transform with the indexes starting at 1, we change the notation to start at 0.
- In this paper, we present the example of a Hamming Window defined by Eq. (1), but other types of window functions could be used.

$$w\left[\frac{n}{NW}\right] = \begin{cases} 0 & \frac{n}{NW} < 0 \\ a_0 - a_1 \cdot \cos\left(\frac{2\pi \cdot n}{NW}\right) & 0 \leq \frac{n}{NW} < 1 \\ 0 & \frac{n}{NW} \geq 1 \end{cases} \quad (1)$$

Where  $w(n)$  is a window function with a size of  $NW$  samples (which should be even) and  $n$  is a discrete index.

As a result of these changes, the transform results in the following equation:

$$\text{STFT\_FD}\{x[m]\}[m, p] = \frac{1}{\sqrt{p \cdot NC}} \sum_{n=0}^{p \cdot NC - 1} x\left[m - \frac{p \cdot NC}{2} + n\right] \cdot w\left[\frac{n}{p \cdot NC}\right] \cdot e^{-\frac{2\pi n}{p}} \quad (2)$$

$$p \in \left\{2, \dots, \frac{NS}{NC}\right\}$$

Where  $x(m)$  is an input signal with  $NS$  samples and  $m$  is a discrete index. The sweep is linear in the integer variable  $p$ , which represents the number of samples per cycle.  $NC$  is a design parameter that determines the number of cycles inside the window function. As shown in [36], the methodology to obtain this transform is the same as in the standard STFT, with the difference that  $NC$  is fixed and the number of samples of the window function ( $NW$ ) depends on variable  $p$ , according to Eq. (3).

$$NW = p \cdot NC \quad (3)$$

As this transform was proposed following the methodology of the STFT [36], but with a different window size criterion, we can state that it is a variant of STFT. The similarities with wavelets, with the Constant-Q Transform and with multi-resolution STFT are discussed in the following sections.

### 3. FORMULATION AS A WAVELET

The Continuous Wavelet Transform (CWT) of a continuous-time signal  $x(t)$ , at a scale ( $s > 0$ )  $s \in \mathcal{R}^+$ , and translational value  $u \in \mathcal{R}$  is defined as follows.

$$CWT_x(u, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t) \bar{\psi}\left(\frac{t-u}{s}\right) dt \quad (4)$$

Where  $\psi(t)$  is the mother wavelet and  $\frac{1}{\sqrt{s}}$  is an energy normalization factor so that the transformed signal will have the same energy at every scale [38]. We let  $v$  be defined by Eq. (5).

$$v = t - u \quad (5)$$

This leads to the following expression:

$$CWT_x(u, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(v+u) \bar{\psi}\left(\frac{v}{s}\right) dv \quad (6)$$

Where we are directly applying the translation  $u$  to the input signal. In this paper, we propose the following  $\psi(t)$  function to be used in Eq. (6).

$$\psi(t) = w(t) \cdot e^{i2\pi \cdot NC \cdot t} \quad (7)$$

Where  $w(t)$  is a window function with a size of 1s. In this paper, we show as example the case of a Hamming Window defined in the continuous time by Eq. (8), but other types of windows could be used.

$$w(t) = \begin{cases} 0 & t < 0 \\ a_0 - a_1 \cdot \cos(2\pi \cdot t) & 0 \leq t < 1 \\ 0 & t \geq 1 \end{cases} \quad (8)$$

This leads to the following expression.

$$CWT_x(u, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(v+u) \cdot w\left(\frac{v}{s}\right) \cdot e^{-i2\pi \cdot NC \cdot v} dv \quad (9)$$

For discrete signals, using the number of samples of the window function (NW) for the scale, using  $b$  as the discrete translational value, and  $n$  being a discrete index, the transform can be expressed as follows.

$$X[b, NW] = \frac{1}{\sqrt{NW}} \sum_{n=0}^{NW-1} x[b+n] \cdot w\left[\frac{n}{NW}\right] \cdot e^{-i2\pi \cdot NC \cdot n} \quad (10)$$

For the sum to be centered, we use variable  $m$  as defined by Eq. (11).

$$m = b + \frac{NW}{2} \quad (11)$$

This results in Eq. (12).

$$X[m, NW] = \frac{1}{\sqrt{NW}} \sum_{n=0}^{NW-1} x\left[m - \frac{NW}{2} + n\right] \cdot w\left[\frac{n}{NW}\right] \cdot e^{-i2\pi \cdot NC \cdot n} \quad (12)$$

In the STFT-FD the number of samples of the window is related to  $p$  by Eq. (3). Therefore, Eq. (12) can be rewritten as a function of  $p$  according to Eq. (13).

$$X[m, p] = \frac{1}{\sqrt{p \cdot NC}} \sum_{n=0}^{p \cdot NC - 1} x\left[m - \frac{p \cdot NC}{2} + n\right] \cdot w\left[\frac{n}{p \cdot NC}\right] \cdot e^{-i2\pi \cdot n} \quad (13)$$

This is the same equation as STFT-FD in Eq. (2). Therefore, the transform is an alternative discretization of the CWT. Instead of the more common exponential progression of scales, the proposed transform has a specific linear sweep in the scale variable, where all the dilatation (scale) parameters ( $p$ ) are integers<sup>3</sup>. As shown in Section 6, this kind of sweep is interesting for the analysis of some signals, such as ECGs.

### 4. FORMULATION AS A CONSTANT-Q TRANSFORM

The constant-Q transform proposes to set the window size ( $N_k$ ) according to Eq. (14) [35].

$$N_k = \frac{f_s}{f_k} Q \quad (14)$$

$N_k$  in CQT notation is equivalent to NW in the STFT-FD. The fidelity factor,  $Q$ , in CQT notation is equivalent to the number of cycles inside the window function (NC) in the STFT-FD.  $\frac{f_s}{f_k}$  is the number of samples processed per cycle at a center frequency  $f_k$ , which is equivalent to variable  $p$  in Eq. (3). The CQT transform is defined by Eq. (15).

$$X^{CQ}[m, k] = \sum_{j=m-\lfloor \frac{N_k}{2} \rfloor}^{j=m+\lfloor \frac{N_k}{2} \rfloor} x[j] \cdot a_k \left[ j - m + \frac{N_k}{2} \right] \quad (15)$$

Where

$$a_k[n] = \frac{1}{N_k} w\left[\frac{n}{N_k}\right] e^{-i2\pi \cdot n \cdot \frac{f_k}{f_s}} \quad (16)$$

Let  $n$  be defined by Eq. (17).

$$n = j - m + \frac{N_k}{2} \quad (17)$$

The transform can be reformulated as follows using Eq. (14) to (17),

$$X^{CQ}[m, k] = \frac{1}{N_k} \sum_{n=0}^{2\lfloor \frac{N_k}{2} \rfloor} x\left[m - \frac{N_k}{2} + n\right] w\left[\frac{n}{N_k}\right] e^{-i2\pi \cdot Q \cdot n} \quad (18)$$

It is the same equation as the STFT-FD, except that,

- It divides by the number of samples of the window. Instead, as shown in this paper, in order to have the same energy at every frequency, and according to wavelet theory, we divide by the square root of the number of samples of the window. However, wavelets can also be normalized in terms of amplitude (instead of in terms of energy), and in

<sup>3</sup>  $p$  is the number of samples per cycle in the STFT-FD methodology.

that case the normalization factor would be  $\frac{1}{N_k}$  as in the CQT transform [28].

- It requires rounding the window size.
- The sum is calculated in an odd number of samples. This is not specific for all the implementations of the constant-Q transform. For example in [39] the sum is calculated from 0 to  $N_k-1$ ,

Thus, the equations of both the CQT and STFT-FD transforms are similar. However, there is a major difference between the two transforms. The center frequencies being sampled in the CQT obey Eq. (19) [33].

$$f_k = f_1 \cdot 2^{\frac{k-1}{B}} \quad (19)$$

Where  $f_1$  is the center frequency of the lowest-frequency bin and B is the number of bins per octave. In the CQT, the number of bins per octave (B) is related to the fidelity factor (Q) by Eq. (20) [40].

$$Q = \frac{1}{2^{\frac{1}{B}} - 1} \quad (20)$$

Therefore the number of bins is fixed for a given Q (which is equivalent to NC in the STFT-FD) and cannot be arbitrarily increased, because it also determines the number of cycles inside the window function [36]. In the cases analyzed in this paper, we apply the same Q factor to both the CQT and the STFT-FD

The constant-Q Transform is obtained with a sweep in log(frequency), and it is usually explained as a series of logarithmically spaced filters. This makes it more complicated to work with the data structure of the CQT than with the time-frequency matrix of the STFT. Instead, in the STFT-FD the sweep is linear for the number of samples per cycle ( $p$ ), which is inversely proportional to the frequency. These differences make the constant-Q Transform quicker to compute because it only calculates some frequencies. However, the implementation of that transform requires rounding values, implying therefore a discretization issue in real applications. Instead, in the STFT-FD, the sweep is linear in variable  $p$ , requiring neither this concept of a series of logarithmically spaced filters nor rounding the window size to an integer. In the STFT-FD formulation the window size is always directly an integer. In fact, in the way the STFT-FD is formulated, the natural way of representing is not in the time-frequency domain, but in the time-scale domain, where  $p$  is the scale. However, as explained in [36], Eq. (21) can be used to represent this transform in the time-frequency domain.

$$f = \frac{1}{p \cdot T_s} \quad (21)$$

Where  $f$  is the frequency and  $T_s$  is the sampling interval. Taking into account the similarity of the equations of both transforms and the difference in the sweep, we can state that the STFT-FD is a variant of the CQT that uses a different sweep.

## 5. FORMULATION AS MULTI-RESOLUTION STFT

Multi-resolution STFT proposes to divide the signal into frequency bands, and then process each with a Fast Fourier Transform (FFT) of a different window size [9–11]. In the

STFT-FD the number of cycles inside the window function is fixed. This leads to always taking the NC component of the DFT. The result is then divided by the square root of the number of samples of the window, so that the transformed signal will have the same energy at every frequency, according to Eq. (22).

$$\text{STFT\_FD}\{x[m]\}[m, p] = \frac{1}{\sqrt{p \cdot \text{NC}}} \cdot \text{DFT}^{p \cdot \text{NC}}\{wx[m, n]\}[m, \text{NC}] \quad (22)$$

Where  $wx(m, n)$  is the windowed signal (this is, the signal after applying the window function), and the DFT calculated with a fixed window size (NW) is defined by Eq. (23).

$$\text{DFT}^{\text{NW}}\{wx[m, n]\}[m, k] = \sum_{n=0}^{\text{NW}-1} wx[m, n] e^{-j \frac{2\pi \cdot k \cdot n}{\text{NW}}} \quad (23)$$

In standard STFT the window size is fixed in the time domain, while in the STFT-FD it is fixed in the frequency domain. Therefore, the window size of this transform, expressed in number of samples, depends on  $p$ , according to Eq. (3). Taking the components of the transform from the DFTs of the windowed signal of different window sizes is equivalent to taking the components from FFTs or STFTs of different window sizes. Therefore, the proposed transform could also be considered as multi-resolution STFT. In fact, if we denote  $\text{STFT}^{\text{NW} \cdot T_s}\{x[m]\}[m, f]$  the STFT of a signal  $x[m]$ , calculated with a fixed window size of  $\text{NW} \cdot T_s$  seconds, the STFT-FD that we are proposing in this paper can be formulated by Eq. (24).

$$\text{STFT\_FD}\{x[m]\}[m, p] = \frac{1}{\sqrt{p \cdot \text{NC}}} \cdot \text{STFT}^{p \cdot \text{NC} \cdot T_s}\{x[m]\}\left[m, \frac{1}{p \cdot T_s}\right] \quad (24)$$

This means that for each value of  $p$  we are computing the transform with an STFT of a different window size, similar to multi-resolution STFT. This can also be formulated in the continuous time, using the period (T) instead of  $p$ , where it is clear that we use a window size which is NC times the period that we are analyzing, as explained in [36].

$$\text{STFT\_FD}\{x(t)\}(t, T) = \frac{1}{\sqrt{T \cdot \text{NC}}} \cdot \text{STFT}^{T \cdot \text{NC}}\{x(t)\}\left(t, \frac{1}{T}\right) \quad (25)$$

The major differences between these techniques are that multi-resolution STFT requires a band-pass filter to separate the bands when different window sizes are applied, and in multi-resolution STFT only a few discrete window sizes are typically used [9, 10]. The STFT-FD is a special case of multi-resolution STFT, where every frequency has a different window size. In the present formulation, the result is normalized using the square root of the number of samples of the window.

## 6. TIME-FREQUENCY VS TIME-SCALE

As the proposed transform could be expressed as a variant of the Fourier transform, as well as an alternative discretization of the Continuous Wavelet Transform, it is interesting to explore the results of the algorithm in two domains: time-frequency and time-scale. Compared to wavelets, the selection of a different mother wavelet can lead to different results [25]. Although it can be considered a special case of multi-resolution STFT, that technique is typically computed processing only a couple of different values of window sizes, therefore adapting worse to different frequencies or scales [24]. The results of the STFT-FD were compared against standard STFT, an Adaptive Optimal-

Kernel Time Frequency Representation, multi-resolution STFT and wavelets in [36, 37] using several types of signals. Therefore, the case study of this paper focuses on the comparison with the CQT, and especially in the time-scale domain.

### 6.1. Synthetic signal: Linear Period chirp

Two linear frequency chirps were tested in [36] and the results show that the STFT-FD does not provide good resolution for this kind of signals, especially for high frequencies. This is inherent to the transform, because the sweep in this transform is linear in variable  $p$ , which is proportional to period (and not to frequency).

$$p = \frac{T}{T_s} \quad (26)$$

To demonstrate the potential advantages of this approach, we analyze in this paper a linear period chirp using a sinusoid signal whose period changes linearly with time. We propose a sinusoid signal  $x(t)$ , as shown in Eq. (27).

$$x(t) = \sin(\phi(t)) \quad (27)$$

Where  $\phi(t)$  is the argument of the sinusoid.

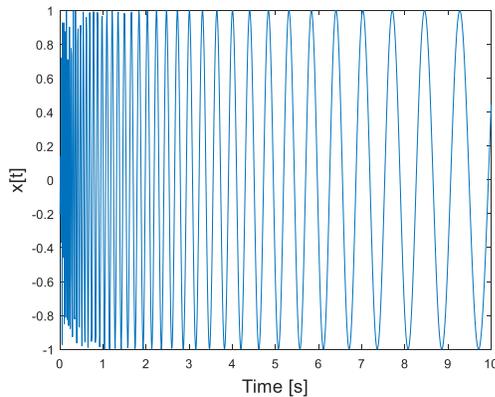
We set the instantaneous period (which is the inverse of the instantaneous frequency [41]) to change linearly with time according to Eq. (28).

$$T(t) = \frac{1}{f(t)} = \frac{1}{\frac{1}{2\pi} \frac{d\phi(t)}{dt}} = T_0 + \kappa \cdot t \quad (28)$$

Where  $T(t)$  is the instantaneous period,  $f(t)$  is the instantaneous frequency,  $T_0$  is the initial period, and  $\kappa$  is the rate of change of the period.

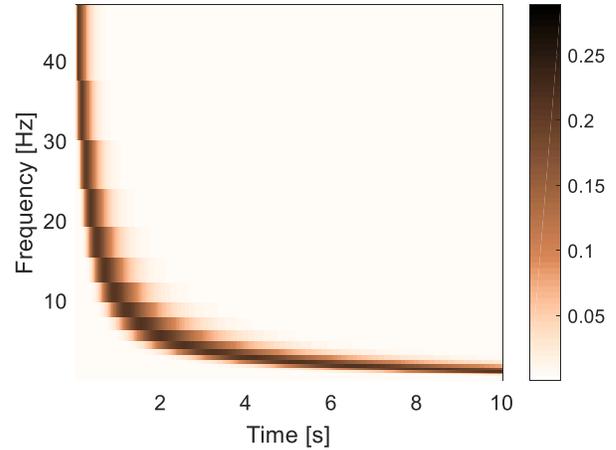
By integrating, we obtain Eq. (29) for the linear period chirp. A synthetic signal has been built with 2000 samples, a sample frequency of 100Hz, an initial period  $T_0 = 20ms$ , and a rate of change of the period  $\kappa = 0.09$ . It is represented in Fig. 1 in the time domain. The aspect of this signal in the time domain is similar to the frequency chirp but the rate of change of the instantaneous frequency is different.

$$x(t) = \sin\left(2\pi \cdot \frac{\ln|T_0 + \kappa \cdot t|}{\kappa}\right) \quad (29)$$



**Fig. 1** Linear period chirp signal represented in the time domain.

This signal is represented using the brute force method of the constant-Q transform [32], both in the time-frequency domain (see Fig. 2) and in the time-scale domain (see Fig. 3).

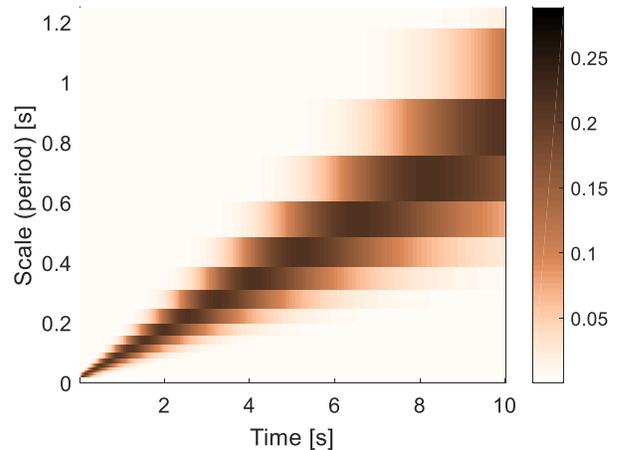


**Fig. 2** Linear period chirp signal represented with the constant-Q transform in the time-frequency domain ( $Q=4$ ).

In the time-scale domain, the period changes linearly with time, as determined by Eq. (28). In the frequency domain, the inverse relation between frequency and period is identified. In the time-scale domain, the width of the signal is also observed to increase with the scale. The reason for this is clear if we re-write CQT Eq. (14) in terms of the period, see Eq. (30), as the window size is increased proportionally to the period. We also observe that the CQT has a very poor resolution in the time-scale domain, especially for high scales.

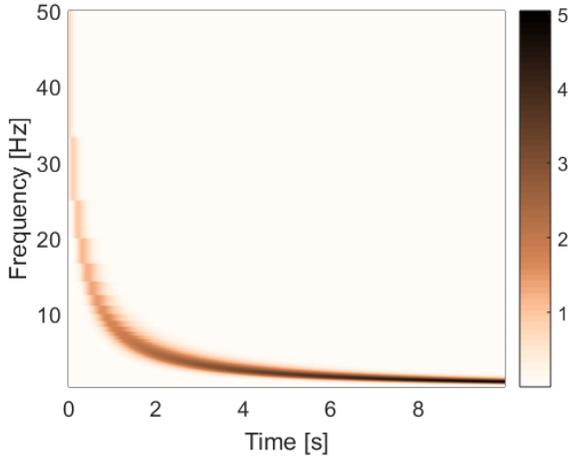
$$N_k = \frac{T_k}{T_s} Q \quad (30)$$

Where  $T_k$  is the period in  $k$ .



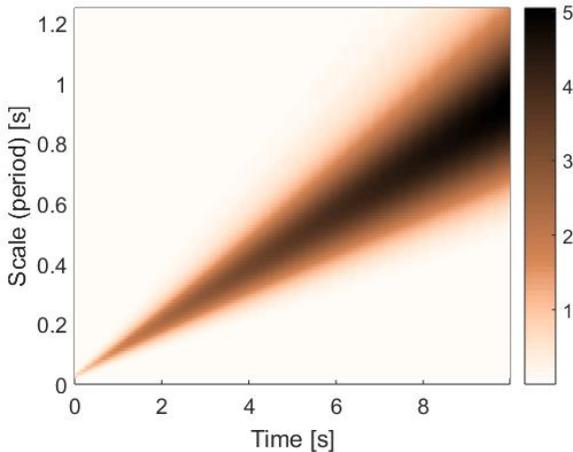
**Fig. 3** Linear period chirp signal represented with the constant-Q transform in the time-scale domain ( $Q=4$ ).

The STFT-FD is represented in Fig. 4 and Fig. 5. The equation is almost the same as in the CQT (except for the normalization factor) and therefore results are very similar, but the way the sweep is carried out in the STFT-FD improves significantly the resolution of this transform.



**Fig. 4** Linear period chirp signal represented with the STFT-FD in the time-frequency domain (NC=4).

This improvement is especially noted in the time-scale domain (see Fig. 5 compared to Fig. 3). Please note that the CQT only samples some frequency values, see Eq. (19), and rounds to the nearest point. Instead, the STFT-FD calculates all the terms that can be computed using the equations that define these transforms. This is shown in Eq. (2) with  $p$  starting in 2, which corresponds to Nyquist frequency, and finishing in NS/NC, which is the upper limit, as higher values of  $p$  would require a window size larger than the number of samples of the signal (see Eq. (3)). This is the reason for the improved resolution observed in the case studies of this paper, especially in the time-scale domain. This also has an impact on computing time. The CQT is computed in 0.8s, while the STFT-FD is computed in 3.5s.

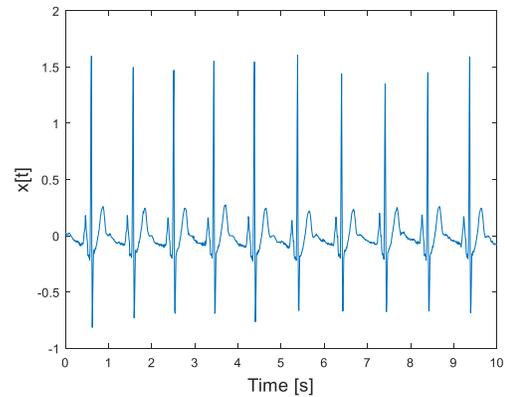


**Fig. 5** Linear period chirp signal represented with the STFT-FD in the time-scale domain (NC=4).

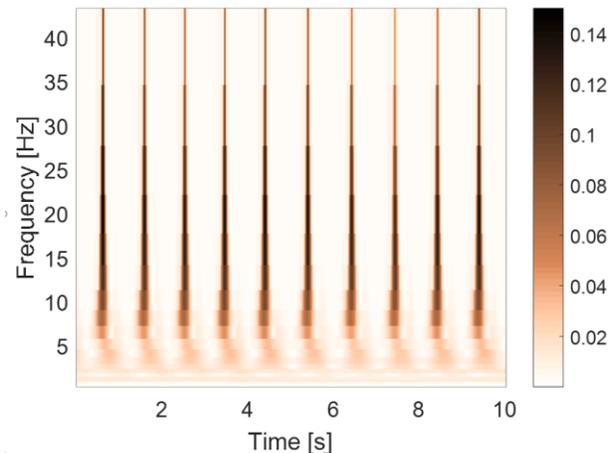
## 6.2. Application to a real signal: ECG

An ECG signal<sup>4</sup>, represented in Fig. 6 in the time domain, is selected as a real case study to compare the time-frequency and time-scale domains. In Fig. 7, the CQT of the ECG signal is represented. In Fig. 8 and Fig. 9, the ECG signal is analyzed

using the STFT-FD in the time-frequency domain and in the time-scale domain. This representation seems to have more resolution than the CQT and uses a different normalization factor. In the time-scale domain we use the period (obtained as  $p \cdot T_s$  according to Eq. (21)) to represent the scale. With the STFT-FD, in both the time-frequency and in the time-scale domain, we can observe the impulses every second, and two low frequency components. The impulses (Area 1) are observed as high frequencies in the time-frequency representation, and low scales in the time-scale representation. The two low frequency components (Areas 2 and 3) have a scale (represented by the period) of about 0.5s and 1s, corresponding to 2Hz and 1Hz respectively. These two frequency components, as well as the impulses, are caused by the cardiac cycle of atrial and ventricular depolarizations and repolarizations. The low frequency components (high scale components) are better observed with the STFT-FD in the time-scale domain.

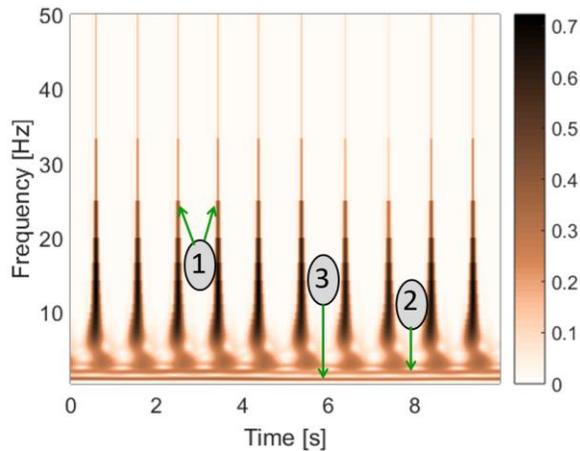


**Fig. 6** ECG signal represented in the time domain

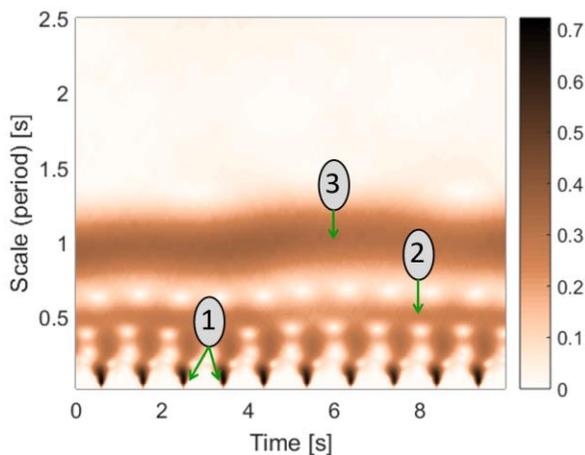


**Fig. 7.** ECG signal represented with the constant-Q transform in the time-frequency domain (Q=4).

<sup>4</sup> The ECG signal corresponds to 10s of a signal downloaded from <http://eleceng.dit.ie/dorran/matlab/ecg.txt>, with a sample frequency of 100Hz (original source: <https://physionet.org>).



**Fig. 8.** ECG signal represented with the STFT-FD in the time-frequency domain (NC=4).



**Fig. 9.** ECG signal represented with the STFT-FD in the time-scale domain (NC=4).

## 7. CONCLUSIONS

In this paper we have revisited the STFT-FD, improving its formulation according to wavelet theory, in order to normalize the energy for all the frequencies. In this paper the transform has been analyzed and compared with four existing techniques. Following [36], it is shown that the transform can be considered a variant of the standard STFT. Besides, this paper also shows that the transform can be formulated as an alternative discretization of the Continuous Wavelet Transform. The similarities and differences with the constant-Q transform have also been discussed, showing that the main difference is in the way the sweep is carried out. Finally, it is shown that the proposed transform can also be expressed as a special case of multi-resolution STFT, without requiring the band-pass filters of that approach.

Instead of the more common exponential progression of scales, the proposed transform has a specific linear sweep in the scale variable, where all the dilatation (scale) parameters ( $p$ ) are integers. In some signals, like ECGs, the proposed sweep can enhance the quality of the representation.

To sum up, the paper demonstrates how the proposed transform can be expressed as a variant of discrete Fourier transforms, as

well as an alternative discretization of wavelet transforms. Besides, it is a variant of the CQT and a special case of the multi-resolution STFT. Therefore, the transform can serve as a common framework for these four techniques. The major drawback of the present transform (as in the CQT) is the computational complexity, this being an area of further research [36, 37]. However, the fact of combining different perspectives within a single transform, enhances the value of the proposed approach allowing it to obtain good representations of signals, as shown in this paper, and analyzed in more detail in [36, 37]. Finally, as the sweep is inversely proportional to the frequency, and proportional to the period, the transform is more adequate for analyzing low frequency components or signals whose components vary proportionally to the period. Thus, in the paper it is shown that the low frequency components of ECG signals are better represented in the time-scale or time-period domain using the STFT-FD than in the time-frequency domain using the CQT.

Given these advantages, it is expected that the transform will be useful in fields where currently wavelets and the STFT are now applied, such as engineering, speech, biology and medicine. The areas of engineering and medicine may be especially interesting for using the time-scale representation of this transform. In the same way that the frequency domain contributed to new theories in engineering, the application of this kind of signal techniques, not only in research, but also in the day-to-day of medicine could also enhance the understanding of the information extracted from biomedical signals.

## 8. REFERENCES

- Xing, F., Chen, H., Xie, S., Yao, J.: Ultrafast Three-Dimensional Surface Imaging Based on Short-Time Fourier Transform. *IEEE Photonics Teh. Lett.* 27, 2264–2267 (2015)
- Zhang, H., Hua, G., Yu, L., Cai, Y., Bi, G.: Underdetermined blind separation of overlapped speech mixtures in time-frequency domain with estimated number of sources. *Speech Commun.* 89, 1–16 (2017). <https://doi.org/10.1016/j.specom.2017.02.003>
- Liu, H., Li, L., Ma, J.: Rolling Bearing Fault Diagnosis Based on STFT-Deep Learning and Sound Signals. *Shock Vib.* 2016, (2016)
- Zhang, W.Y., Hao, T., Chang, Y., Zhao, Y.H.: Time-frequency analysis of enhanced GPR detection of RF tagged buried plastic pipes. *NDT E Int.* 92, 88–96 (2017). <https://doi.org/10.1016/j.ndteint.2017.07.013>
- Kara, S., Içer, S., Erdogan, N.: Spectral broadening of lower extremity venous Doppler signals using STFT and AR modeling. *Digit. Signal Process.* 18, 669–676 (2008)
- Gabor, D.: Theory of Communication. *J. Inst. Electr. Eng. London.* 93, 429–457 (1946). <https://doi.org/10.1049/ji-3-2.1946.0074>
- Xie, H., Lin, J., Lei, Y., Liao, Y.: Fast-varying AM-FM components extraction based on an adaptive STFT. *Digit. Signal Process.* 22, 664–670 (2012)
- Pei, S.C., Huang, S.G.: STFT with adaptive window width based on the chirp rate. *IEEE Trans. Signal Process.* 60, 4065–4080 (2012)
- Gnann, V., Becker, J.: Signal reconstruction from multiresolution STFT magnitudes with mutual initialization. In: 45th International Conference: Applications of Time-Frequency Processing in Audio. pp. 2–3 (2012)
- Pihlajamäki, T.: Multi-resolution Short-time Fourier Transform Implementation of Directional Audio Coding, (2009)
- Bonada, J.: Automatic Technique in Frequency Domain for Near-Lossless Time-Scale Modification of Audio. *Int. Comput. Music Conf.* 396–399 (2000)
- Juillerat, N., Müller Arisona, S., Schubiger-banz, S.: Enhancing the quality of audio transformations using the multi-scale Short Time

- Fourier Transform. Proc. 10th IASTED Int. Conf. Signal Image Process. 623, 379–387 (2008)
13. Miao, H., Zhang, F., Tao, R.: Fractional Fourier Analysis Using the Möbius Inversion Formula. *IEEE Trans. Signal Process.* 67, 3181–3196 (2019). <https://doi.org/10.1109/TSP.2019.2912878>
  14. Wei, D., Li, Y.M.: Generalized sampling expansions with multiple sampling rates for lowpass and bandpass signals in the fractional Fourier transform domain. *IEEE Trans. Signal Process.* 64, 4861–4874 (2016). <https://doi.org/10.1109/TSP.2016.2560148>
  15. Tao, R., Li, Y.L., Wang, Y.: Short-time fractional fourier transform and its applications. *IEEE Trans. Signal Process.* 58, 2568–2580 (2010). <https://doi.org/10.1109/TSP.2009.2028095>
  16. Capus, C., Brown, K.: Short-time fractional Fourier methods for the time-frequency representation of chirp signals. *J. Acoust. Soc. Am.* 113, 3253 (2003). <https://doi.org/10.1121/1.1570434>
  17. Oktem, F., Ozaktas, H.: Exact Relation Between Continuous and Discrete Linear Canonical Transforms. *IEEE Signal Process. Lett.* 16, 727–730 (2009). <https://doi.org/10.1109/LSP.2009.2023940>
  18. Wei, D., Yang, W., Li, Y.M.: Lattices sampling and sampling rate conversion of multidimensional bandlimited signals in the linear canonical transform domain. *J. Franklin Inst.* 356, 7571–7607 (2019). <https://doi.org/10.1016/j.jfranklin.2019.06.031>
  19. Wei, D., Li, Y.M.: Convolution and Multichannel Sampling for the Offset Linear Canonical Transform and Their Applications. *IEEE Trans. Signal Process.* 67, 6009–6024 (2019). <https://doi.org/10.1109/TSP.2019.2951191>
  20. Bao, Y.P., Zhang, Y.N., Song, Y.E., Li, B.Z., Dang, P.: Nonuniform sampling theorems for random signals in the offset linear canonical transform domain. Proc. - 9th Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA ASC 2017. 2018-Febru, 94–99 (2018). <https://doi.org/10.1109/APSIPA.2017.8282008>
  21. Rioul, O., Vetterli, M.: Wavelets and signal processing. *IEEE Signal Process. Mag.* 8, 14–38 (1991). <https://doi.org/10.1109/79.91217>
  22. Mallat, S.: *A Wavelet Tour of Signal Processing*. Academic Press (2009)
  23. Gargour, C., Gabrea, M., Ramachandran, V., Lina, J.M.: A Short Introduction to Wavelets And Their Applications. *IEEE Circuits Syst. Mag.* 9, 57–68 (2009). <https://doi.org/10.1109/MCAS.2009.932556>
  24. Fanaee, F., Yazdi, M., Faghihi, M.: Face image super-resolution via sparse representation and wavelet transform. *Signal, Image Video Process.* 13, 79–86 (2019). <https://doi.org/10.1007/s11760-018-1330-9>
  25. Chen, J., Phung, T., Blackburn, T., Ambikairajah, E., Zhang, D.: Detection of high impedance faults using current transformers for sensing and identification based on features extracted using wavelet transform. *IET Gener. Transm. Distrib.* 10, 2990–2998 (2016). <https://doi.org/10.1049/iet-gtd.2016.0021>
  26. Tantawi, M.M., Revett, K., Salem, A.B., Tolba, M.F.: A wavelet feature extraction method for electrocardiogram (ECG)-based biometric recognition. *Signal, Image Video Process.* 9, 1271–1280 (2015). <https://doi.org/10.1007/s11760-013-0568-5>
  27. Alajlan, N., Bazi, Y., Melgani, F., Malek, S., Bencherif, M.A.: Detection of premature ventricular contraction arrhythmias in electrocardiogram signals with kernel methods. *Signal, Image Video Process.* 8, 931–942 (2014). <https://doi.org/10.1007/s11760-012-0339-8>
  28. Sheng, Y.: Wavelet Transform. In: *The Transforms and Applications Handbook: Second Edition* (2000)
  29. Hlawatsch, F., Boudreaux-Bartels, G.F.: Linear and Quadratic Time-Frequency Signal Representations. *IEEE Signal Process. Mag.* 9, 21–67 (1992)
  30. Gu, Y.H., Bollen, M.H.J.: Time-frequency and time-scale domain analysis of voltage disturbances. *IEEE Trans. Power Deliv.* 15, 1279–1284 (2000). <https://doi.org/10.1109/61.891515>
  31. Wang, J., Ding, Y., Ren, S., Wang, W.: Sampling and reconstruction of multiband signals in multiresolution subspaces associated with the fractional wavelet transform. *IEEE Signal Process. Lett.* 26, 174–178 (2019). <https://doi.org/10.1109/LSP.2018.2883832>
  32. Brown, J.C.: Calculation of a constant Q spectral transform. *J. Acoust. Soc. Am.* 89, 425–434 (1991). <https://doi.org/10.1121/1.400476>
  33. Bachhav, P.B., Todisco, M., Mossi, M., Beaugeant, C., Evans, N.: Artificial bandwidth extension using the constant Q transform. *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc.* 5550–5554 (2017). <https://doi.org/10.1109/ICASSP.2017.7953218>
  34. Dorran, D.: Audio Time-Scale Modification. <http://arrow.dit.ie/engdoc>, (2005)
  35. Bo, H., Li, H., Ma, L., Yu, B.: A Constant Q Transform based approach for robust EEG spectral analysis. In: *2014 International Conference on Audio, Language and Image Processing*. pp. 58–63. IEEE (2014)
  36. Mateo, C., Talavera, J.A.: Short Time Fourier Transform with the Window Size Fixed in the Frequency Domain. *Digit. Signal Process.* 77, 13–21 (2018)
  37. Mateo, C., Talavera, J.A.: Short Time Fourier Transform with the Window Size Fixed in the Frequency Domain: Implementation. *SoftwareX.* 8, 5–8 (2018)
  38. Bravo, J.P., Roque, S., Estrela, R., Leão, I.C., De Medeiros, J.R.: Wavelets: a powerful tool for studying rotation, activity, and pulsation in Kepler and CoRoT stellar light curves. *Astron. Astrophys.* 568, A34 (2014). <https://doi.org/10.1051/0004-6361/201323032>
  39. Brown, J.C., Puckette, M.S.: An efficient algorithm for the calculation of a constant Q transform. *J. Acoust. Soc. Am.* 92, 2698–2701 (1992). <https://doi.org/10.1121/1.404385>
  40. Dobre, R.A., Negrescu, C.: Automatic music transcription software based on constant Q transform. Proc. 8th Int. Conf. Electron. Comput. Artif. Intell. ECAI 2016. (2017). <https://doi.org/10.1109/ECAI.2016.7861193>
  41. Chang, C.C., Hsu, H.Y., Hsiao, T.C.: The interpretation of very high frequency band of instantaneous pulse rate variability during paced respiration. *Biomed. Eng. Online.* 13, (2014). <https://doi.org/10.1186/1475-925X-13-46>