1 Definitions

The CQT transform X^{CQ} of a discrete time-domain signal $x(n)$ is defined by

$$X^{CQ}(k,n) = \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} x(j) a_k^*(j - n + N_k/2) \tag{1}$$

where $k = 1, 2, ..., K$ indexes the frequency bins of the CQT, $\lfloor \cdot \rfloor$ denotes rounding towards negative infinity and $a_k^*(n)$ denotes the complex conjugate of $a_k(n)$. The basis functions $a_k(n)$ are complex-valued waveforms, here also called time-frequency atoms, and are defined by

$$a_k(n) = \frac{1}{N_k} w(\frac{n}{N_k}) \exp\left[-i2\pi n \frac{f_k}{f_s}\right] \tag{2}$$

where f_k is the center frequency of bin k, f_s denotes the sampling rate, and $w(t)$, is a continuous window function (for example Hann or Blackman window), sampled at points determined by t. The window function is zero outside the range $t \in [0, 1]$.

2 Proof of the inequality

Let prove that

$$\sup_{k,n\in\mathbb{Z}} |X^{CQ}(k,n)| = \|X^{CQ}\|_\infty \lessapprox \|x\|_\infty \cdot \|w\|_1 = \sup_{n\in\mathbb{Z}} |x(n)| \cdot \left(\int_{t\in\mathbb{R}} |w(t)| \, \mathrm{d}t \right) . \tag{3}$$

First, let prove that we have $\forall k, n \in \mathbb{Z}$

$$|X^{CQ}(k,n)| \leq \|x\|_\infty \cdot \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |a_k^*(j - n + N_k/2)| . \tag{4}$$

Indeed,

$$
\begin{aligned}
|X^{CQ}(k,n)| &= \left| \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} x(j) a_k^*(j-n+N_k/2) \right| \\
&\leq \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |x(j) a_k^*(j-n+N_k/2)| \\
&= \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |x(j)| \cdot |a_k^*(j-n+N_k/2)| \\
&\leq \sup_{j\in\mathbb{Z}} |x(j)| \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |a_k^*(j-n+N_k/2)| \\
&= \|x(j)\|_\infty \cdot \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |a_k^*(j-n+N_k/2)| \ .
\end{aligned}
$$

Let now see that

$$
\sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |a_k^*(j-n+N_k/2)| \approx \int_{t\in\mathbb{R}} |w(t)| \, \mathrm{d}t \, . \tag{5}
$$

In one hand,

$$\begin{aligned}
\sum_{j=n-\lfloor N_k/2\rfloor}^{n+\lfloor N_k/2\rfloor} |a_k^*(j-n+N_k/2)| &\approx \sum_{j=n-\lfloor N_k/2\rfloor}^{n+\lfloor N_k/2\rfloor} |a_k^*(j-n+\lfloor N_k/2\rfloor)| \\
&= \sum_{j=n-\lfloor N_k/2\rfloor}^{n+\lfloor N_k/2\rfloor} |a_k(j-n+\lfloor N_k/2\rfloor)| \\
&\overset{l=j-(n-\lfloor N_k/2\rfloor)}{=} \sum_{l=0}^{n+\lfloor N_k/2\rfloor-(n-\lfloor N_k/2\rfloor)} |a_k(l)| \\
&= \sum_{l=0}^{2\lfloor N_k/2\rfloor} |a_k(l)| \leq \sum_{l=0}^{N_k} |a_k(l)| \\
&= \sum_{l=0}^{N_k} \left| \frac{1}{N_k} w(\frac{l}{N_k}) \exp\left[-i2\pi j \frac{f_k}{f_s}\right]\right| \\
&= \sum_{l=0}^{N_k} \left| \frac{1}{N_k} w(\frac{l}{N_k})\right| \left|\exp\left[-i2\pi j \frac{f_k}{f_s}\right]\right| \\
&= \sum_{l=0}^{N_k} \left| \frac{1}{N_k} w(\frac{l}{N_k})\right| \\
&= \frac{1}{N_k} \sum_{l=0}^{N_k} \left| w(\frac{l}{N_k})\right|
\end{aligned}$$

In the other hand,

$$\begin{aligned}
\int_{t\in\mathbb{R}} |w(t)|\,\mathrm{d}t &\overset{\mathrm{supp}(w)=[0,1]}{=} \int_0^1 |w(t)|\,\mathrm{d}t \\
&\overset{t=\frac{u}{N_k}\Rightarrow \mathrm{d}t=\frac{1}{N_k}\,\mathrm{d}u}{=} \int_0^{N_k} \frac{1}{N_k}|w(\frac{u}{N_k})|\,\mathrm{d}u \\
&= \frac{1}{N_k}\int_0^{N_k} |w(\frac{u}{N_k})|\,\mathrm{d}u \\
&\overset{u=l,\mathrm{d}u=1}{\approx} \frac{1}{N_k}\sum_{l=1}^{N_k} |w(\frac{l}{N_k})|
\end{aligned}$$

So we get (5).

Substituting in (4), we get $\forall k, n \in \mathbb{Z}$

$$|X^{CQ}(k,n)| \leq \|x\|_\infty \cdot \sum_{j=n-\lfloor N_k/2 \rfloor}^{n+\lfloor N_k/2 \rfloor} |a_k^*(j-n+N_k/2)| \approx \|x\|_\infty \int_{t\in\mathbb{R}} |w(t)| = \|x\|_\infty \|w\|_1$$

$$\Rightarrow |X^{CQ}(k,n)| \lessapprox \|x\|_\infty \|w\|_1 \,.$$

Which proves that

$$\|X^{CQ}(k,n)\|_\infty \lessapprox \|x\|_\infty \|w\|_1 \,.$$

3 Application to signal processing

If we asumes that $\forall t \in \mathbb{R}$, $x(t) \in [-1, 1]$ (which is the case when dealing with audio signals in floating point format) then we have $\|x\|_\infty \leq 1$.

If we also impose that the window is normalized in the $L^1(\mathbb{R})$ space, (*i.e.*: $\|w\|_1 = 1$) then we have

$$\|X^{CQ}(k,n)\|_\infty \lessapprox \|x\|_\infty \|w\|_1 \leq 1 \cdot \|w\|_1 = 1 \cdot 1 = 1 \,.$$

That means that all the magnitudes of the CQT coefficients are less or equal than 1, which in dB scale means less or equal than 0.