Approximation of the short-time Fourier transform

Monika Dörfler
University of Vienna
0000-0001-6139-630X

Martin Ehler
University of Vienna
0000-0002-3247-6279

Karlheinz Gröchenig
University of Vienna
0000-0003-1461-0654

Andreas Klotz
Austrian Academy of Sciences
0009-0008-7212-2724

***Abstract*—We give estimates for the approximation error of the continuous short-time Fourier transform (STFT) by the discrete short-time Fourier transform from a finite number of samples of a signal f.**

***Index Terms*—STFT, FFT, approximation, interpolation, Wiener amalgam space**

I. Introduction

The short-time Fourier transform of a function or signal f with respect to a window g is defined for all $(x, \omega) \in \mathbb{R}^2$ as

$$\mathcal{V}_g f(x, \omega) = \int_{\mathbb{R}} f(u)\overline{g(u-x)}e^{-2\pi \mathrm{i} u\omega}\mathrm{d}u$$

and is an important method in processing time series such as audio signals. The *phase-locked* short-time Fourier transform used almost exclusively in signal processing applications is

$$\begin{aligned}\mathcal{W}_g f(x, \omega) &= \int_{\mathbb{R}} f(u+x)\overline{g(u)}e^{-2\pi \mathrm{i} u\omega}\mathrm{d}u \\ &= e^{2\pi \mathrm{i} x\omega}\mathcal{V}_g f(x, \omega)\,.\end{aligned}$$

The present contribution answers the question, how well $\mathcal{V}_g f(x, \omega)$ and $\mathcal{W}_g f(x, \omega)$ (we focus on the latter from now on) can be approximated from finitely many samples of f.

As $\mathcal{W}_g f(x, \cdot)$ is the Fourier transform of f translated by $-x$ and multiplied with g we can apply our previous results on the approximation of the Fourier transform by the discrete Fourier transform, as obtained in our earlier paper [2].

Using the notation $[n] := \{j \in \mathbb{Z} : -\frac{n}{2} < j \le \frac{n}{2}\}$, the discrete Fourier transform (DFT) $\mathcal{F} : \mathbb{C}^n \to \mathbb{C}^n$ of $y = (y_j)_{j\in[n]} \in \mathbb{C}^n$ is

$$\mathcal{F} y = \left(\frac{1}{\sqrt{n}} \sum_{j\in[n]} y_j e^{-2\pi \mathrm{i}\frac{kj}{n}}\right)_{k\in[n]}.$$

The translate of the function f by $x \in \mathbb{R}$ is $T_x f(u) = f(u-x)$.

AK was supported by the FWF project DISCO (PAT4780023).

The scaled sampling of a function f with step size h and length n is defined as

$$f_{h,n} := \left(\sqrt{h} f(hj)\right)_{j\in[n]} \in \mathbb{C}^n\,.$$

Throughout the text we use the relation $p = hn$ between the interval length p, the step size h and the number of samples n of the discrete Fourier transform used for the discrete STFT.

The discrete version of the short-time Fourier transform that we work with is

$$\begin{aligned}\mathcal{W}_g^{h,n} f(x, k) =& \mathcal{F}\,(T_{-x} f \cdot \overline{g})_{h,n}\,(k) \\ =& \sqrt{\frac{h}{n}} \sum_{j\in[n]} f(hj-x)\overline{g(hj)}e^{-2\pi i\frac{kj}{n}}\end{aligned}$$

for $x \in \mathbb{R}$ and $k \in [n]$. In implementations the time parameter x is often discretized: $x \in Mh\mathbb{Z}$. Here h is the step size as above and $M \in \mathbb{N}$ is a subsampling factor.

Basically we want to approximate the STFT by its discrete version at a point x and the Fourier frequencies k/p for $k \in [n]$ and samples of f at $hj - x$, $j \in [n]$ such that

$$\mathcal{W}_g f(x, \frac{k}{p}) \approx \mathcal{W}_g^{h,n} f(x, k)\,. \tag{1}$$

In signal processing applications one often wishes to impose only weak assumptions on the decay and regularity of the function (class) under analysis, while being able to impose strong conditions on the decay and smoothness of the window. The latter is usually desired to be well-localized in both time and frequency.

We will estimate the error of (1) in terms of the decay of f and the regularity of f, i.e. the decay of $\hat{f}$, *and* in terms of the decay and regularity of g. Therefore we introduce weights and appropriate spaces.

II. Weights and spaces

Basically we measure decay of a function f by considering the weighted L^q norm $\|f \cdot v\|_{L^q}$ for a weight

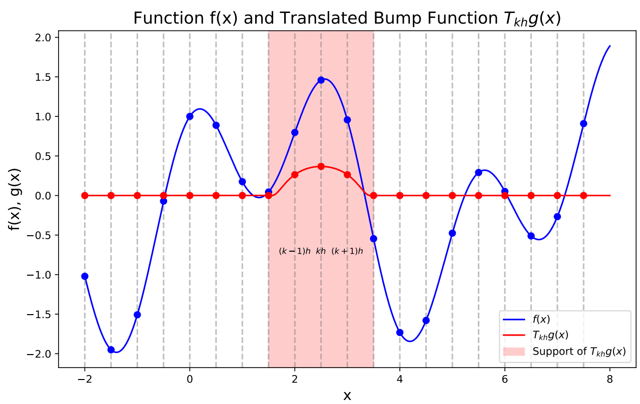

Fig. 1. Sampling pattern for the discrete approximation of the STFT

v. As we need to consider function (spaces) where stable sampling of functions is possible we have to replace L^q by spaces that are somewhat smaller, the Wiener amalgam spaces defined below.

In order to avoid technicalities we restrict the weights under consideration to *standard weights*, i.e., polynomial weights $v_\alpha(x) = (1+|x|)^\alpha$ for $\alpha > \frac{1}{2}$, and subexponential weights $v_{r,\alpha}(x) = e^{r|x|^\alpha}$, for $0 < \alpha \le 1$ and $r > 0$. More general classes of weights can be considered, in particular, all weights that are equivalent to the weights above (in the sense $v \sim w$ if $v/w \asymp 1$) will work as well.

We say $c = (c_l)_{l\in\mathbb{Z}}$ is in the weighted sequence space ℓ^q_v for $1 \le q < \infty$ if the norm $\|c\|_{\ell^q_v} = (\sum_{l\in\mathbb{Z}} |c_l|^q |v(l)|^q)^{1/q}$ is finite (standard modification for $q = \infty$).

We next define a class of Wiener amalgam spaces. Standard references on Wiener amalgams include [1, Sec 2.4], [3], [4].

The Wiener amalgam space $W(C, \ell^q_v)$ consists of all continuous functions f such that

$$\|f\|_{W(C,\ell^q_v)} := \left(\sum_{l\in\mathbb{Z}} \sup_{x\in[0,1]} |f(x+l)|^q \, |v(l)|^q \right)^{1/q} < \infty \,.$$

Note that the embedding $\ell^2_v \subset \ell^1$ is satisfied for all subexponential weights and for the polynomial weights v_α with $\alpha > 1/2$, and it implies the continuous embeddings $W(C, \ell^2_v) \subset W(C, \ell^1) \subset L^1(\mathbb{R})$. Observe that $W(C, \ell^\infty) = C_b$, the space of bounded continuous functions.

If we want to analyze signals that are sinusoidal, e.g., of the form $f(x) = \exp(\mathrm{i}\omega x)$, then we have to deal with Fourier transforms that are measures: Let $\mathcal{M}$ be the space of finite Borel measures on $\mathbb{R}$. The Wiener amalgam space $W(\mathcal{M}, \ell^r_v)$ consists of all Borel measures μ such that

$$\|\mu\|_{W(\mathcal{M},\ell^r_v)} = \left\| \left(\|\mu\chi_{[k,k+1]}\|_{\mathcal{M}} \right)_{k\in\mathbb{Z}} \right\|_{\ell^r_v} < \infty \,,$$

where $\chi_{[k,k+1]}$ denotes the indicator function of the interval $[k, k+1]$ and $\|\mu\|_{\mathcal{M}} = |\mu|(\mathbb{R})$.

III. Results

We use results from [2] about the approximation of the Fourier transform of a function f,

$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi \mathrm{i} x\xi} \mathrm{d}x \,,$$

by the DFT to obtain similar results for the approximation of the STFT. The approximation error of the DFT is defined as

$$E_h^{[n]}(f) = \left(\frac{1}{p} \sum_{k\in[n]} \left| \hat{f}(\tfrac{k}{p}) - h \sum_{j\in[n]} f(hj) e^{-2\pi \mathrm{i} \frac{kj}{n}} \right|^2 \right)^{\frac{1}{2}} .$$

Accordingly we define the approximation error of the STFT *at the point* $x \in \mathbb{R}$ as

$$\begin{aligned} \mathcal{E}_{g,\{x\}}(f) &= E_h^{[n]}(T_{-x} f \cdot \bar{g}) \\ &= \left\| \left(\mathcal{W}_g f(x, \frac{k}{p}) - \mathcal{W}_g^{h,n} f(x,k) \right)_{k\in[n]} \right\| . \end{aligned}$$

This can be generalized to arbitrary finite sets $X \subset \mathbb{R}$ and more general norms as

$$\mathcal{E}_{g,X}(f) = \frac{1}{|X|^{\frac{1}{q}}} \left\| \left(\mathcal{W}_g f(x, \frac{k}{p}) - \mathcal{W}_g^{h,n} f(x,k) \right)_{(x,k)\in X\times[n]} \right\|_{\ell^{q,2}} . \tag{2}$$

For a given standard weight v we need the auxiliary quantity

$$\Phi_v(p) := \left(2 \sum_{m=0}^{\infty} |v(pm + \tfrac{p}{2})|^{-2} \right)^{\frac{1}{2}} .$$

In [2, Lemma 3.4, Lemma 3.7] the following asymptotic bounds on Φ_v are obtained:

$$\begin{aligned} \Phi_{v_\alpha}(p) &\lesssim p^{-\alpha}, \quad \alpha > \frac{1}{2}, \\ \Phi_{v_{r,\alpha}}(p) &\lesssim e^{-r(\frac{p}{2})^\alpha}, \quad 0 < \alpha \le 1, r > 0 \,. \end{aligned}$$

The main technical lemma of [2] can be summarized as follows.

Lemma 1. *Let $f \in W(C, \ell^2_v)$ and $\hat{f} \in W(C, \ell^2_w)$. If v, w are standard weights then, for h small and p large,*

$$E_h^{[n]}(f) \lesssim \Phi_v(p) \|f\|_{W(C,\ell^2_v)} + \Phi_w(h^{-1}) \|\hat{f}\|_{W(C,\ell^2_w)} \,.$$

Applied to the STFT we obtain the error estimate

$$\mathcal{E}_{g,\{x\}}(f) \lesssim \Phi_v(p) \|T_{-x}f \cdot \bar{g}\|_{W(C,\ell^2_v)} + \Phi_w(h^{-1}) \|\widehat{T_{-x}f \cdot \bar{g}}\|_{W(C,\ell^2_w)} . \tag{3}$$

The norms $\|T_{-x}f \cdot \bar{g}\|_{W(C,\ell^2_v)}$ and $\|\widehat{T_{-x}f \cdot \bar{g}}\|_{W(C,\ell^2_w)}$ can be estimated further by an application of Hölder's and Young's inequalities for Wiener amalgam spaces. We specialize to a case important in applications. As mentioned above, the interest is often in weak conditions on the smoothness and decay of the signal f, whereas the conditions on the analysis window g can be stronger.

So let us assume that f satisfies

$$f \in C_b = W(C, \ell^\infty) \text{ and } \hat{f} \in W(\mathcal{M}, \ell^2_w)$$

for a standard weight w (actually the Hausdorff-Young inequality [4]) yields that the assumption $\hat{f} \in W(\mathcal{M}, \ell^2_w)$ already implies $f \in C_b$). The following result estimates the error of approximating the STFT of these functions at one point x for windows of subexponential decay in time and frequency.

Theorem 2. *Let the function f be bounded and continuous, $f \in C_b$, and assume that the window function g and its Fourier transform are of subexponential decay,*

$$g \in W(C, \ell^2_{v_{r,\alpha}}), \qquad \hat{g} \in W(C, \ell^1_{v_{s,\beta}}) .$$

(I) *If $\hat{f} \in W(\mathcal{M}, \ell^2_{v_{t,\gamma}})$ for $\gamma < \beta$ then*

$$\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-r(\frac{p}{2})^\alpha} + e^{-t(2h)^{-\gamma}} .$$

Given h, the asymptotically best choice for p is obtained by balancing both summands. This leads to

$$p_{opt} = 2\left(\frac{t}{r}\right)^{\frac{1}{\alpha}} (2h)^{-\frac{\gamma}{\alpha}} ,$$

and $\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-t(2h)^{-\gamma}}$.

Likewise, if $\gamma > \beta$ then

$$\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-r(\frac{p}{2})^\alpha} + e^{-s(2h)^{-\beta}},$$

$$p_{opt} = 2\left(\frac{s}{r}\right)^{\frac{1}{\alpha}} (2h)^{-\frac{\beta}{\alpha}} .$$

and $\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-s(2h)^{-\beta}}$.

If $\beta = \gamma$ then

$$\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-r(\frac{p}{2})^\alpha} + e^{-\min(s,t)(2h)^{-\beta}} ,$$

$$p_{opt} = 2\left(\frac{\min(s,t)}{r}\right)^{\frac{1}{\alpha}} (2h)^{-\frac{\beta}{\alpha}} ,$$

and $\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-\min(s,t)(2h)^{-\beta}}$.

(II) *If $\hat{f} \in W(\mathcal{M}, \ell^2_{v_\gamma})$ then*

$$\mathcal{E}_{g,\{x\}}(f) \lesssim e^{-r(\frac{p}{2})^\alpha} + h^\gamma .$$

Given h, the asymptotically best choice for p is

$$p_{opt} = 2(\gamma \log \frac{1}{h} - \log C)^{1/\alpha}$$

for a constant C that depends on the norms of f, $\hat{f}$, g and $\hat{g}$. The approximation error can be estimated as

$$\mathcal{E}_{g,\{x\}}(f) \lesssim h^\gamma .$$

Remark 3. The heuristics is as follows: The approximation error depends on the *decay rate of the window* and on the *minimum* of the regularity of the function and the window (the decay of its Fourier transform)

Remark 4. The asymptotic results of the theorem are satisfied for the more general error measure $\mathcal{E}^q_{g,X}(f)$ (cf. Equation (2)) for all $1 \leq q \leq \infty$, and all fixed finite sets X. Moreover, they still hold for sets X of the form $X = X_{L,M,h} = [-L(p), L(p)] \cap Mh\mathbb{Z}$, where $L(p) \sim \gamma p$ for $0 < \gamma \leq 1$.

Remark 5 (The case of a compactly supported window). If $\operatorname{supp} g \subset [-\frac{p}{2}, \frac{p}{2}]$ then the results of [2][(3.6) together with Theorem 3.3] imply that

$$\mathcal{E}_{g,\{x\}}(f) \lesssim \Phi_w(h^{-1}) ,$$

i.e., the estimate (3) is independent of p. The approximation error depends only on the step size h and on the regularity of f and g. In particular the p dependent estimates for $\mathcal{E}_{g,\{x\}}(f)$ in Theorem 2 (i.e. the first summands) vanish.

This aligns well with common practice in the signal processing community, where the window length $p = nh$ is typically chosen to match the support length of g.

IV. Outlook

In a next step, one can interpolate the values $\mathcal{W}^{h,n}_g f(x_l, k)$, for $k \in [n]$, $x_l \in X$, and obtain an approximation of the full short-time Fourier transform $\mathcal{W}_g f(x, \omega)$. Let Φ be a cardinal interpolation function, i.e., $\Phi(x_{l'} - x_l, \frac{k'-k}{p}) = \delta_{k,k'}\delta_{l,l'}$. Then we consider the interpolation

$$\widetilde{\mathcal{W}}_g f(x,\omega) = \sum_{\substack{x_l \in X\\ k \in [n]}} \mathcal{W}^{h,n}_g f(x_l, k) \Phi(x - x_l, \omega - \tfrac{k}{p}) e^{-2\pi i x_l \omega}$$

for $(x, \omega) \in \mathbb{R}^2$. This function yields an approximation of the short-time Fourier transform $\mathcal{W}_g f$ near the points $(x_l, k/p)$. To obtain an error estimate for $\mathcal{W}_g f - \widetilde{\mathcal{W}}_g f$, one applies known estimates from approximation theory. This task will be carried out in future work.

References

[1] E. Cordero and L. Rodino. *Time-frequency analysis of operators*, volume 75 of *De Gruyter Studies in Mathematics*. De Gruyter, Berlin, [2020] ©2020.

[2] M. Ehler, K. Gröchenig, and A. Klotz. Quantitative estimates: How well does the discrete Fourier transform approximate the Fourier transform on $\mathbb{R}$, 2024.

[3] H. G. Feichtinger. Banach convolution algebras of Wiener type. In *Functions, series, operators, Vol. I, II (Budapest, 1980)*, volume 35 of *Colloq. Math. Soc. János Bolyai*, pages 509–524. North-Holland, Amsterdam, 1983.

[4] H. G. Feichtinger. Wiener amalgams over Euclidean spaces and some of their applications. In *Function spaces (Edwardsville, IL, 1990)*, volume 136 of *Lecture Notes in Pure and Appl. Math.*, pages 123–137. Dekker, New York, 1992.