On the Convergence of Fourier Representations and Schwartz Distributions

Pushpendra Singh*, *Member, IEEE,* Amit Singhal, Binish Fatimah, Anubha Gupta, *Senior Member, IEEE,* and Shiv Dutt Joshi

***Abstract*—Fourier theory is a popular tool for analyzing various signals and interpreting their spectral contents. It finds applications in a wide variety of subject areas. However, some of the popular signals, such as sinusoidal signals, Dirac delta, signum, and unit step, fail to have a convergent Fourier representation (FR) in the conventional sense. Hence, it becomes imperative to utilize the distribution theory to understand and build a suitable representation for these signals. The signal processing and communication engineering literature does not explain these concepts clearly. As a result, many of the concepts of FR for signals that do not conform to the conventional derivations remain obscure to researchers. We attempt to bridge this gap and provide a comprehensive explanation regarding the existence of FR. Further, we have proposed a new linear space of Gauss–Schwartz (GS) functions and corresponding tempered superexponential (TSE) distributions. It is shown that the Fourier transform (FT) is an isomorphism on the GS space of test functions and hence by duality is an isomorphism on TSE distributions. The space GS is smallest in the sense that its dual space, a set of TSE distributions, is the largest linear space over which FT can be defined by duality.**

***Index Terms*—Fourier series, Fourier transform, Schwartz distributions, Gauss functions, tempered superexponential distribution.**

I. Introduction

Fourier theory is one of the most important tools used ubiquitously for understanding the spectral content of a signal, extracting and interpreting information from signals, transmitting and processing, and analyzing the signals and systems. Since its inception in 1807, Fourier theory has been used to solve problems in almost all fields of mathematics, engineering, science and technology to model, design and analyze engineering systems and physical phenomena. Some of the recent applications include decomposition and analysis of signals [1], [2], spatial-audio signal processing [3], analysis of images [4], radiometry [5], radar signal analysis [6], and enhancement of power amplifier [7]. Authors in [4] analyze symmetric-geometry computed tomography with the help of Fourier theory to build an accurate reconstruction method. An imaging technique is developed in [5] by estimating near-field temperatures for synthetic aperture radiometry using Fourier transform (FT). Further, various frequency modulated radar waveforms exhibiting similar time-frequency distributions are differentiated using fractional FT in [6]. On the contrary, an efficiency enhancement scheme for envelope tracking power amplifier is proposed in [7] using short-time FT. Owing to the large application base for non-uniform discrete FT, a faster estimation algorithm [8] is designed recently to achieve better results than the sparse FT.

Although the concepts of Fourier theory are foundational for the areas of signal processing and communication engineering, the popular signal processing literature [9]–[13] does not offer a clear explanation regarding the convergence or the existence of Fourier representations for certain well-known signals. Some clarity regarding convergence of Fourier representation is offered in [14], however, FT for some popular signals cannot be interpreted in the conventional sense. Therefore, it is required to understand the distribution theory that connects delta Dirac functions with the Fourier theory. The distribution theory by Schwartz in 1945 is one of the great revolutions in mathematical function analysis [15]. Despite being a powerful tool for understanding Fourier theory, it is generally ignored in the engineering literature. In this work, we utilize the concepts of this theory to show how some signals that fail to exhibit FT in the conventional sense, can have FT in the distributional sense.

It is shown in the literature that the FT is defined for (i) the functions of at most polynomial growth (i.e., $x(t) = t^n$ for $n \in \{0, 1, 2, 3, \cdots\}$) in the sense of tempered distributions [15]–[17], and (ii) the functions of at most exponential growth (i.e., $x(t) = \exp(at)$ for $a \in \mathbb{C}$) in the sense of distributions corresponding to the space of Gauss functions [18]. These aspects can be easily understood with the help of distribution theory. In this work, we provide a lucid description of the existence and convergence of Fourier representations using the concepts of the distribution theory that may benefit the entire signal-processing, communication and computer engineering community. Suitable examples have been presented to support the text. The main contributions of this work are as follows:

1) We present a comprehensive summary of the convergence of Fourier series (FS) and Fourier transform (FT) as available in the communication engineering, signal processing, and the other literature.
2) We extend the theory of FT by proposing the space of Gauss–Schwartz functions and the corresponding tempered superexponential distributions. Thus, we define the FT for the functions of at most tempered superexponential growth, i.e., $x(t) = \exp\big(\alpha t^2\big)$, where $t \in \mathbb{R}$ and $\alpha \in \mathbb{C}$ such that the real part $\mathrm{Re}(\alpha) < \pi$.

Thus, the theory presented in this work generalizes the FT.

Pushpendra Singh is with the School of Engineering, Jawaharlal Nehru University Delhi, and Department of Electronics and Communication Engineering, National Institute of Technology Hamirpur, Hamirpur (HP) 177005 India (e-mail: pushpendrasingh@iitkalumni.org). *Corresponding author.

Amit Singhal is with the Department of Electronics & Communication Engineering, Netaji Subhas University of Technology, Delhi, India (e-mail: singhalamit2002@gmail.com).

Binish Fatimah is with MaxEye Technologies Private Ltd, Bengaluru, India (e-mail: binish.fatimah@gmail.com).

Anubha Gupta is with the Department of Electronics and Communication Engineering, IIIT-Delhi, Delhi 110020, India (e-mail: anubha@iiitd.ac.in; Lab: http://sbilab.iiitd.edu.in/).

Shiv Dutt Joshi is with the Department of EE, IIT Delhi, Delhi, India (e-mail: sdjoshi@iitd.ac.in).

Moreover, there is no need to restrict FT for functions of at most exponential growth. Any locally integrable function of at most tempered superexponential growth is Fourier transformable. The study is organized as follows: Some prerequisites are presented in Section II. The Section III presents Fourier Transform in $\mathcal{L}_2(\mathbb{R})$. The Schwartz distributions and corresponding Fourier transforms are presented in Section IV. The conditions for convergence of Fourier series and Fourier transform are summarized in Section V. Finally, the conclusion is presented in Section VI.

II. Prerequisites

Definition 1. *The $\mathcal{L}_p([a,b])$ space is defined as*

$$\mathcal{L}_p([a,b]) = \{x : [a,b] \to \mathbb{C} \mid \|x\|_p < \infty\}, \tag{1}$$

where $\|x\|_p$ denotes the $\mathcal{L}_p$-norm of the function x defined as

$$\|x\|_p = \left(\int_a^b |x(t)|^p \, \mathrm{d}t \right)^{1/p} \quad \text{where} \quad 1 \le p < \infty. \tag{2}$$

The space $\mathcal{L}_\infty([a,b])$ is a set of functions bounded almost everywhere.

Definition 2. *Let $x : [a,b] \to \mathbb{C}$ be a function and let $P_n = \{a = t_0 < t_1 < \cdots < t_n = b\}$, $n \in \mathbb{N}$, be a finite partition of $[a,b]$. The total variation of $x(t)$ for $t \in [a,b]$ for all such partitions P_n for any $n \in \mathbb{N}$ is defined as*

$$V(x,[a,b]) = \sup_{P_n} \left\{ \sum_{i=1}^{n} |x(t_i) - x(t_{i-1})| \right\}. \tag{3}$$

A function x is defined to have bounded variation (BV) on $[a,b]$, denoted as $x \in BV([a,b])$, if $V(x,[a,b]) < \infty$.

Similarly, the space $BV(\mathbb{R})$ is defined as $BV(\mathbb{R}) = \{x : \mathbb{R} \to \mathbb{C} \mid x \text{ has bounded variation on } \mathbb{R}\}$.

A. Fourier Series–Representation and Convergence

The FS representation for a periodic signal $\tilde{x}(t)$, with period T_0, is defined as:

$$\text{Synthesis equation:} \qquad \tilde{x}(t) = \sum_{k=-\infty}^{\infty} \widehat{x}_k \exp(jk\omega_0 t), \tag{4}$$

$$\text{Analysis equation:} \qquad \widehat{x}_k = \frac{1}{T_0} \int_{\mathbb{T}} \tilde{x}(t) \exp(-jk\omega_0 t) \, \mathrm{d}t, \tag{5}$$

where $\mathbb{T} = [t_0, t_0 + T_0]$ with $t_0 \in \mathbb{R}$, and $\omega_0 = 2\pi/T_0$. The FS representation (4) converges if $\tilde{x}$ satisfies the Dirichlet conditions [14], i.e., $\tilde{x} \in \mathcal{L}_1(\mathbb{T}) \cap BV(\mathbb{T})$. Further, according to the Carleson–Hunt theorem [19], [20], if $\tilde{x} \in \mathcal{L}_p(\mathbb{T})$ for $p > 1$, then its FS converges at almost all points. The space $\mathcal{L}_1(\mathbb{T})$ is excluded from this theorem because in 1923, Kolmogorov produced an example of a periodic $\mathcal{L}_1(\mathbb{T})$ signal whose FS diverges pointwise almost everywhere [21]. The convergence is understood as how the partial sum [22]

$$S_N(t) = \sum_{k=-N}^{N} \widehat{x}_k \exp(jk\omega_0 t), \tag{6}$$

converges to the original signal $\tilde{x}(t)$ while $N \to \infty$, i.e., whether it converges uniformly, pointwise, or in norm sense. These convergences, uniform, pointwise, and $\mathcal{L}_p$-norm, are defined as

$$\begin{aligned} &S_N \to \tilde{x} \text{ uniformly, if } \lim_{N\to\infty} \left\{ \sup_{t\in\mathbb{T}} |\tilde{x}(t) - S_N(t)| \right\} = 0, \\ &S_N \to \tilde{x} \text{ pointwise, if } \lim_{N\to\infty} S_N(t) = \tilde{x}(t), \quad \forall\, t \in \mathbb{T}, \\ &\text{and } S_N \to \tilde{x} \text{ in } \mathcal{L}_p\text{-norm in } \mathbb{T}, \text{ if } \lim_{N\to\infty} \|\tilde{x}(t) - S_N(t)\|_p = 0. \end{aligned} \tag{7}$$

The uniform convergence is the strongest one that implies both pointwise convergence and $\mathcal{L}_p$-norm convergence, but the converse is not true. The uniform convergence preserves continuity, differentiability and integrability. Further details regarding the convergence of FS, along with some suitable examples, can be found in [14].

Example 1. Let us consider the sequence of functions $x_n(t) = t^n$ for $t \in \mathbb{T} = [0,1)$. One can easily show that $x_n(t) \to x(t) = 0$ in $\mathcal{L}_p$-norm and pointwise on $[0,1)$. However, by definition $\lim_{n\to\infty} \{\sup_{t\in\mathbb{T}} |x_n(t) - x(t)|\} = \lim_{n\to\infty} \{\sup_{t\in\mathbb{T}} |t^n - 0|\} = \lim_{n\to\infty} \{1\} = 1 \neq 0$. Thus the sequence of functions t^n does not converge uniformly on the interval $[0,1)$.

B. Fourier Transform–Representation and Convergence

The FT can be obtained from a limiting case of FS (5) with the period $T_0 \to \infty$. Thus, the FT and inverse FT (IFT) can be defined for $x \in \mathcal{L}_1(\mathbb{R})$ and $\widehat{x} \in \mathcal{L}_1(\mathbb{R})$ as

$$\widehat{x}(\omega) = c_1 \int_{-\infty}^{\infty} x(t) \, \exp(-j\omega t) \, \mathrm{d}t, \tag{8}$$

$$\text{and} \quad x(t) = c_2 \int_{-\infty}^{\infty} \widehat{x}(\omega) \, \exp(j\omega t) \, \mathrm{d}\omega, \tag{9}$$

respectively, where $\omega = 2\pi f$ and $c_1 \times c_2 = \frac{1}{2\pi}$. The literature popularly considers $c_1 = 1$ and $c_2 = \frac{1}{2\pi}$ because it corresponds to the following symmetric FT and inverse FT (IFT) representations

$$\widehat{x}(f) = \int_{-\infty}^{\infty} x(t) \, \exp(-j2\pi f t) \, \mathrm{d}t \tag{10}$$

$$\text{and} \quad x(t) = \int_{-\infty}^{\infty} \widehat{x}(f) \, \exp(j2\pi f t) \, \mathrm{d}f, \tag{11}$$

respectively. We denote FT as $\mathcal{F}\{x\} = \widehat{x}$, IFT as $\mathcal{F}^{-1}\{\widehat{x}\} = x$, and FT pair as $x(t) \rightleftharpoons \widehat{x}(f)$. One can obtain the duality property of FT (11) as $\widehat{x}(t) \rightleftharpoons x(-f)$, which is very useful because it can provide FT that may be difficult to compute directly. From (10) and (11), we can write

$$\widehat{x}(0) = \int_{-\infty}^{\infty} x(t) \, \mathrm{d}t \qquad \text{and} \qquad x(0) = \int_{-\infty}^{\infty} \widehat{x}(f) \, \mathrm{d}f. \tag{12}$$

Moreover, we observe that

$$|\widehat{x}(f)| = \left| \int_{-\infty}^{\infty} x(t) \, \exp(-j2\pi f t) \, \mathrm{d}t \right| \le \int_{-\infty}^{\infty} |x(t)| \, \mathrm{d}t = \|x\|_1 \tag{13}$$

$$|x(t)| = \left| \int_{-\infty}^{\infty} \widehat{x}(f) \, \exp(j2\pi f t) \, \mathrm{d}f \right| \le \int_{-\infty}^{\infty} |\widehat{x}(f)| \, \mathrm{d}f = \|\widehat{x}\|_1. \tag{14}$$

Therefore, if $x \in \mathcal{L}_1(\mathbb{R})$, the FT $\widehat{x}$ is uniformly continuous, vanishes at infinity (i.e., $|\widehat{x}(f)| \to 0$ as $|f| \to \infty$), and is bounded by the $\mathcal{L}_1$-norm of the function (Riemann–Lebesgue lemma). Similarly, if $\widehat{x} \in \mathcal{L}_1(\mathbb{R})$, the function x is uniformly continuous, vanishes at infinity, and is bounded by the $\mathcal{L}_1$-norm of the FT. The FT and IFT of a function x are guaranteed if either the Dirichlet condition of $x \in \mathcal{L}_1(\mathbb{R}) \cap BV(\mathbb{R})$ is fulfilled, or $x \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$.

III. Fourier Transform in $\mathcal{L}_2(\mathbb{R})$

The FT of a typical function, $x \in \mathcal{L}_2(\mathbb{R})$ but $x \notin \mathcal{L}_1(\mathbb{R})$ may not converge. Therefore, FT of any function $x \in \mathcal{L}_2(\mathbb{R})$ is defined using an *extension-by-continuity* from the following results (refer [16], [17] for more details and proofs).

Theorem 1. *$\mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$ is dense in $\mathcal{L}_2(\mathbb{R})$.*

This implies that for any $x \in \mathcal{L}_2(\mathbb{R})$, one can find a sequence of functions $\{x_n\}_{n=1}^{\infty}$ in $\mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$, such that $\lim_{n\to\infty} \|x - x_n\|_2^2 \to 0$. In fact, it is easy to find such a sequence of functions $x_n(t) = x(t)\chi_{[-n,n]}(t)$ using indicator functions, where an indicator function is defined as

$$\chi_{[-n,n]}(t) = \begin{cases} 1, & -n \le t \le n \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

for every $n \in \mathbb{N}$. This implies that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{L}_2(\mathbb{R})$, which converges to the function $x \in \mathcal{L}_2(\mathbb{R})$.

Theorem 2. *Let $x \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$. Then $\widehat{x} \in \mathcal{L}_2(\mathbb{R})$. Furthermore, $\|x\|_2^2 = \|\widehat{x}\|_2^2$ (Parseval–Plancherel identity).*

Remark 1. The space $\mathcal{L}_2(\mathbb{R})$ is a Hilbert space, and every Cauchy sequence in $\mathcal{L}_2(\mathbb{R})$ converges to some function in $\mathcal{L}_2(\mathbb{R})$. Thus, for any $x \in \mathcal{L}_2(\mathbb{R})$, one can obtain a sequence of functions $\{x_n\}_{n=1}^{\infty} \subset \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$, such that $\lim_{n\to\infty} \|x - x_n\|_2^2 \to 0$. This implies that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathcal{L}_2(\mathbb{R})$ and hence, for any $m, n \in \mathbb{N}$, $x_m - x_n \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$. According to Theorem 2, $\|x_m - x_n\|_2^2 = \|\widehat{x}_m - \widehat{x}_n\|_2^2$ because $\mathcal{F}\{x_m - x_n\} = \widehat{x}_m - \widehat{x}_n$. This implies that $\{\widehat{x}_n\}_{n=1}^{\infty}$ is also a Cauchy sequence in $\mathcal{L}_2(\mathbb{R})$ and hence, there is a function $\widehat{x} \in \mathcal{L}_2(\mathbb{R})$, such that sequence $\{\widehat{x}_n\}_{n=1}^{\infty}$ converges to $\widehat{x}$ under the norm of $\mathcal{L}_2(\mathbb{R})$. Moreover, for $x, y \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$, if $\widehat{x} = \widehat{y}$, then $x = y$.

Remark 2. It has been shown in the literature [14], [16], [17] that the FT can be defined on

1) $\mathcal{L}_1(\mathbb{R})$ in which $\mathcal{F} : \mathcal{L}_1(\mathbb{R}) \to \mathcal{L}_\infty(\mathbb{R})$ with $\|\widehat{x}\|_\infty \le \|x\|_1$.
2) $\mathcal{L}_2(\mathbb{R})$ in which $\mathcal{F} : \mathcal{L}_2(\mathbb{R}) \to \mathcal{L}_2(\mathbb{R})$ with $\|\widehat{x}\|_2 = \|x\|_2$.
3) $\mathcal{L}_p(\mathbb{R})$ for $1 \le p \le 2$ from the Riesz-Thorin Theorem $\mathcal{F} : \mathcal{L}_p(\mathbb{R}) \to \mathcal{L}_q(\mathbb{R})$, $\frac{1}{p} + \frac{1}{q} = 1$, $1 \le p \le 2$ with $\|\widehat{x}\|_q \le \|x\|_p$ which is Hausdorff–Young inequality.
4) $\mathcal{L}_p([a,b])$ for $p \in [1,\infty]$ from Fourier series $\mathcal{F} : \mathcal{L}_p([a,b]) \to \mathcal{L}_q(\mathbb{R})$, with $\|\widehat{x}\|_q \le \|x\|_p$ where $-\infty < a < b < \infty$, i.e., x has compact support on $[a,b]$ where $x(t) = 0$ for $t \notin [a,b]$.

This is to note that there is equality only for $p = q = 2$, for which FT is an isometry and the FT is invertible. In fact, FT is invertible only for the finite energy signals ($x \in \mathcal{L}_2(\mathbb{R})$) in the normal sense of integration as defined in (11). If a signal is not an energy signal, then the FT integral (11) does not converge in the normal sense. We know that FT of many signals such as $\sin(\omega_0 t)$, $\cos(\omega_0 t)$, $\delta(t)$, $\mathrm{u}(t)$, and $\frac{1}{\pi t}$ are defined, well known, ubiquitous, and are widely used. However, these are not energy signals. We would like to emphasize that the FT of these functions is defined only in the distributional sense [15], which we present in the next section.

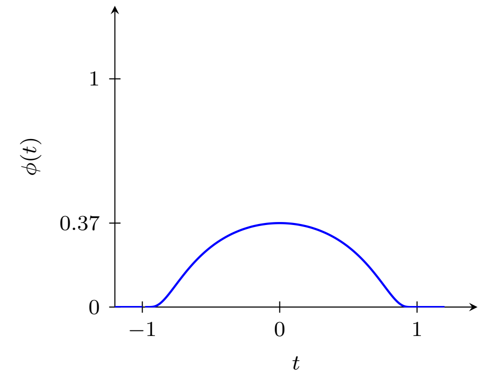

Fig. 1. An example of a test function $\phi(t)$ defined in (16).

IV. Schwartz Distributions and Fourier Transform

In this section, we consider Schwartz distributions, especially the tempered distributions, exponential distributions and their FT in distributional sense. We also propose and define the tempered superexponential distributions and their FT.

Definition 3. *Let $\mathcal{D} = C_c^\infty(\mathbb{R})$ denotes the space of test functions $\phi : \mathbb{R} \to \mathbb{C}$ that are infinitely differentiable and compactly supported, i.e., there exists a compact (closed and bounded) set $\mathbb{K} \subset \mathbb{R}$ such that $\phi(t) = 0$ for all $t \notin \mathbb{K}$.*

The space $\mathcal{D}$ is endowed with the topology that $\phi_k \to \phi$ in $\mathcal{D}$ as $k \to \infty$, if (i) all the ϕ_k are supported in the same compact set $\mathbb{K}$, i.e., $\phi_k(t) = 0$ if $t \notin \mathbb{K}$ for all $k \in \mathbb{N}$, and (ii) $\frac{\mathrm{d}^m}{\mathrm{d}t^m}\phi_k(t)$ converges to $\frac{\mathrm{d}^m}{\mathrm{d}t^m}\phi(t)$ uniformly for any $m \in \mathbb{N}_0$, where $\mathbb{N}_0 = \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\}$. The notion of convergence makes $\mathcal{D}$ a topological space where continuity can be defined, and these notions are naturally extended to the topologically dual space of distributions.

Example 2. The function $\phi : \mathbb{R} \to \mathbb{R}$ defined by

$$\phi(t) = \begin{cases} \exp\left(\frac{-1}{1-t^2}\right), & t \in (-1, 1) \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

is an example of a test function $\phi \in \mathcal{D}$ which is shown in Figure 1. The support of ϕ, $\mathrm{supp}(\phi) = [-1, 1]$, is the closed interval, as ϕ is non-zero on the open interval $(-1, 1)$ and the closure of this set is $[-1, 1]$.

Definition 4. *Let $\mathcal{S} = \mathcal{S}(\mathbb{R}) = \{\phi \in C^\infty(\mathbb{R}) \,|\, t^k \frac{\mathrm{d}^m}{\mathrm{d}t^m}\phi(t) \to 0$ as $|t| \to \infty \, \forall k, m \in \mathbb{N}_0\}$ denotes the space of Schwartz functions, where ϕ and all its derivatives are of rapid decay.*

For examples (i) $\phi(t) = t^k \exp(-at^2) \in \mathcal{S}$ for any $k \in \mathbb{N}_0$ and $a > 0$, (ii) $\psi(t) = \exp(-|t|) \notin \mathcal{S}$ as it is not differentiable, and (iii) $\psi(t) = \exp(t) \notin \mathcal{S}$ because $\psi(t) \not\to 0$ as $t \to \infty$.

Definition 5. *Let $\mathcal{E} = C^\infty(\mathbb{R})$ denotes the space of smooth functions.*

This is to note that $\mathcal{D} \subset \mathcal{S} \subset \mathcal{E}$. This work considers one-dimensional functions, although the theory is also valid for higher dimensions. Some properties of Schwartz space $\mathcal{S}$ are: (1) $\mathcal{S}$ is a linear vector space; (2) $\mathcal{S}$ is closed under multiplication; (3) $\mathcal{S}$ is closed under multiplication by polynomials; (4) $\mathcal{S}$ is closed under differentiation; (5) $\mathcal{S}$ is closed under convolution; (6) $\mathcal{S}$ is closed under translations and multiplication by complex exponentials; and (7) The functions of $\mathcal{S}$ are integrable: $\int_{\mathbb{R}} |\phi(t)| \,\mathrm{d}t < \infty$ for $\phi \in \mathcal{S}$.

Definition 6. *A **distribution** T is a continuous linear functional on the space of test functions, i.e., $T : \mathcal{D} \to \mathbb{C}$ such that for all $\phi_1, \phi_2 \in \mathcal{D}$ and $c_1, c_2 \in \mathbb{C}$, T is (i) linear: $T(c_1\phi_1 + c_2\phi_2) = c_1T(\phi_1) + c_2T(\phi_2)$, and (ii) continuous: $T(\phi_n) \to T(\phi)$ in $\mathbb{C}$ as $\phi_n \to \phi$ in $\mathcal{D}$.*

The distribution T is a mapping of test function $\phi \in \mathcal{D}$ to $T(\phi) \in \mathbb{C}$ which in general is denoted by $\langle T, \phi\rangle$.

Definition 7. *The **dual space** of $\mathcal{D}$ is denoted as $\mathcal{D}'$ which is a vector space of continuous linear functionals from $\mathcal{D} \to \mathbb{C}$. It is a space of distributions or a set of distributions. One may observe that the space $\mathcal{D}'$ is a linear space because for all $T_1, T_2 \in \mathcal{D}'$, $\phi \in \mathcal{D}$ and $c \in \mathbb{C}$: (i) $\langle T_1 + T_2, \phi\rangle = \langle T_1, \phi\rangle + \langle T_2, \phi\rangle$, and (ii) $\langle cT_1, \phi\rangle = c\langle T_1, \phi\rangle$.*

Definition 8. *A **tempered distribution** is a continuous linear functional on the space of Schwartz functions, i.e., a mapping from $\mathcal{S} \to \mathbb{C}$. The **dual space** of $\mathcal{S}$ is denoted as $\mathcal{S}'$, which is a space of tempered distributions. A **tempered distribution** refers to a **distribution of temperate growth**, meaning thereby a growth that is **at most polynomial**.*

Definition 9. *A **compactly-supported distribution** is a continuous linear functional on the space of smooth functions, i.e., a mapping from $\mathcal{E} \to \mathbb{C}$. The **dual space** of $\mathcal{E}$ is denoted as $\mathcal{E}'$, which is a space of compactly-supported distributions.*

These spaces follow the inclusions as $\mathcal{D} \subset \mathcal{S} \subset \mathcal{E}$, while there is an inclusion-reversing containment of dual spaces: $\mathcal{E}' \subset \mathcal{S}' \subset \mathcal{D}'$. Therefore, a tempered distribution is a kind of distribution, and a compactly-supported distribution is a kind of tempered distribution. Furthermore, the space $\mathcal{S}$ is dense in $\mathcal{L}_2(\mathbb{R})$ [15]. Hence, the chain of containment can be refined as $\mathcal{D} \subset \mathcal{S} \subset \mathcal{L}_2(\mathbb{R}) \subset \mathcal{S}' \subset \mathcal{D}'$, i.e., Compactly supported functions $\subset$ Rapidly decaying functions $\subset$ Energy functions $\subset$ Tempered distributions $\subset$ Distributions.

A function x is locally integrable if, $\int_a^b |x(t)| \,\mathrm{d}t < \infty$, for every finite a and b. For examples, $x(t) = \exp(t)$ and $x(t) = 1/\sqrt{t}$ are locally integrable but $x(t) = 1/t$ is not locally integrable. Every locally integrable function is a distribution, but a distribution is not necessarily a function. A locally integrable function x can be identified with a particular distribution, namely, the distribution $T_x : \mathcal{D} \to \mathbb{C}$, defined as:

$$T_x(\phi) = \langle T_x, \phi\rangle = \int_{-\infty}^{\infty} x(t)\phi(t) \,\mathrm{d}t < \infty, \forall \phi \in \mathcal{D}, \tag{17}$$

where $T_x \in \mathcal{D}'$. In general, a function x determines a distribution T_x by (17). The distributions like T_x that arise from functions in this way are prototypical examples of distributions that are called **regular distributions**. $T \in \mathcal{D}'$ would be referred to as a regular distribution if there exists a locally integrable function x such that $T = T_x$. Since $\langle x, \phi\rangle$ is equal to $\langle T_x, \phi\rangle$ for any test function $\phi \in \mathcal{D}$, T_x is linear with respect to x, i.e., $T_{x+y} = T_x + T_y$ and $T_{cx} = cT_x$. Therefore, $\langle T_x, \phi\rangle$ is usually denoted by $\langle x, \phi\rangle$, which is very useful. With x as a function, a particular t maps to a particular $x(t)$. Similarly, T is a distribution, and for a particular x, there is a distribution T_x as defined in (17). In other words, T_x is a distribution induced by x, characterized by x, or corresponding to x. Let x be a continuous function, then $\langle x, \phi\rangle = 0$ for all $\phi \in \mathcal{D}$ implies that $x = 0$. Moreover, one can observe that $T_x = T_y \Leftrightarrow \langle x, \phi\rangle = \langle y, \phi\rangle \Leftrightarrow \langle x - y, \phi\rangle = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$ almost everywhere, i.e., x and y can differ on a set of Lebesgue measure zero. The functions $\phi \in \mathcal{D}$ are called test functions because one may contemplate that they are being used to test or detect or evaluate the values of a distribution on an open set.

A. The Dirac delta function

There are many distributions that cannot be defined by integration with any function. Examples include the Dirac delta function and distributions defined to act by integration of test functions against certain measures. Delta is not a regular distribution because there is no locally integrable function x (that could be considered as delta) fulfilling $T_x(\phi) = \langle x, \phi\rangle = \phi(0)$ for all $\phi \in \mathcal{D}$. In other words, for a locally integrable function, x, $\lim_{\epsilon\to 0} \int_{-\epsilon}^{\epsilon} x(t) \,\mathrm{d}t = 0$. However, $\delta(\phi) = \langle \delta, \phi\rangle = \phi(0)$, and with the abuse of notation in the literature, it is defined as $\int_{-\epsilon}^{\epsilon} \delta(t) \,\mathrm{d}t = 1$ for $\epsilon > 0$, which is widespread and very convenient for understanding and exploring the properties of δ function.

Definition 10. *Dirac delta is a **singular distribution** defined as $\delta : \mathcal{S} \to \mathbb{C}$ such that $\delta(\phi) = \langle \delta, \phi\rangle = \phi(0)$ for all $\phi \in \mathcal{S}$ and $\delta \in \mathcal{S}'$. Moreover, $\delta_\mu(\phi) = \langle \delta_\mu, \phi\rangle = \phi(\mu)$ where δ_μ is shifted delta function concentrated at a point μ.*

Weak convergence and delta function: The generalized or weak limit [23] of ordinary functions $x_n(t)$ is a distribution or generalized function $x(t)$, and one can say that $x_n \to x$ weakly if $\langle x_n, \phi\rangle = \langle x, \phi\rangle$, i.e.,

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} x_n(t)\phi(t) \,\mathrm{d}t = \int_{-\infty}^{\infty} x(t)\phi(t) \,\mathrm{d}t, \tag{18}$$

for all test functions $\phi(t) \in \mathcal{S}$. Moreover, any weak limit of a distribution is a distribution, and a distribution is the weak limit of a sequence of functions. One can obtain the Taylor series expansion of $\phi(t)$ around $t = \mu$ as

$$\begin{aligned}\phi(t) &= \sum_{m=0}^{\infty} a_m (t-\mu)^m \\ &= \phi(\mu) + \phi^{(1)}(\mu)\,(t-\mu) + \frac{1}{2!}\phi^{(2)}(\mu)\,(t-\mu)^2 + \ldots\end{aligned} \tag{19}$$

where $\phi^{(m)}(t) = \frac{\mathrm{d}^m}{\mathrm{d}t^m}\phi(t)$ and $a_m = \frac{1}{m!}\phi^{(m)}(\mu)$. If the area under $x_n(t)$ is unity for all $n \in \mathbb{N}$, and the support of function $x_n(t)$ approaches to set $\{\mu\}$ which has zero Lebesgue measure as $n \to \infty$. Then $x_n(t)\phi(t)$ can be approximated by using the first term of (19) as $x_n(t)\phi(\mu)$ for $n \to \infty$. Thus (18) can be written as

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} x_n(t)\phi(t)\,\mathrm{d}t = \lim_{n\to\infty} \int_{-\infty}^{\infty} x_n(t)\phi(\mu)\,\mathrm{d}t = \phi(\mu), \tag{20}$$

in the usual sense of integral.

As an approximation to δ using functions in $\mathcal{S}$, one can consider the family of Gaussians $g(t,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-t^2}{2\sigma^2}\right)$, and define that

$$\lim_{\sigma\to 0} \int_{-\infty}^{\infty} g(t,\sigma)\phi(t)\,\mathrm{d}t = \phi(0) = \langle \delta, \phi \rangle, \tag{21}$$

where the limit of the integral makes sense but not the integral of the limit in (21). If we replace $1/\sigma$ by n where $n \in \mathbb{N}$, then we can generate a sequence of functions

$$g_n(t) = \frac{n}{\sqrt{2\pi}} \exp\left(-t^2 n^2/2\right) \text{ where } \lim_{n\to\infty} g_n(t) \to \delta(t) \text{ weakly} \tag{22}$$

as shown in Figure 2. Thus one can define $\lim_{\sigma\to 0} g(t,\sigma) = \delta(t)$ weakly, which is concentrated at $t = 0$; however, the limit does not make sense pointwise. Similarly, one can define δ_μ concentrated at μ as

$$\delta(t-\mu) = \lim_{\sigma\to 0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right), \tag{23}$$

by using the fundamental result of weak limit

$$\lim_{\sigma\to 0} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right)\phi(t)\,\mathrm{d}t = \phi(\mu) = \langle \delta_\mu, \phi \rangle. \tag{24}$$

From the above discussions, one may informally or with the abuse of notation write three properties (i) $\delta(t-\mu) = 0$ for $t \neq \mu$, $\delta(t-\mu) = \infty$ for $t = \mu$, (ii) $\int_{-\infty}^{\infty} \delta(t-\mu)\,\mathrm{d}t = 1$, and (iii) $\int_{-\infty}^{\infty} \delta(t-\mu)\,\phi(t)\,\mathrm{d}t = \phi(\mu)$, which are widely used in the literature to introduce delta function. However, there is no function in the classical sense that has any of these three properties. Next, we consider the following two examples to illustrate the above concepts.

Example 3. Using the indicator function (15), we consider two cases for a sequence of functions $x_n(t) = \frac{\chi_{[-n,n]}(t)}{2n}$ as follows:

1) Let $n \in \mathbb{N}$, and observe that the $\lim_{n\to\infty} x_n \to x$ pointwise where $x(t) = 0$ for all t. In fact, $x_n \to x$ uniformly as $n \to \infty$; however, the integral of limit is not the same as the limit of the integral, i.e., $\lim_{n\to\infty} \left(\int_{-n}^{n} x_n(t)\,\mathrm{d}t\right) = 1 \neq \int_{-n}^{n} \lim_{n\to\infty} x_n(t)\,\mathrm{d}t = 0$.
2) Let $n \in \mathbb{R}$ and consider the limit of integral as $\lim_{n\to 0} \left(\int_{-n}^{n} x_n(t)\phi(t)\,\mathrm{d}t\right) = \lim_{n\to 0} \left(\frac{\Phi(n)-\Phi(-n)}{2n}\right) = \Phi'(0) = \phi(0)$, where $\int \phi(t)\,\mathrm{d}t = \Phi(t) \implies \Phi'(t) = \phi(t)$, and thus $\lim_{n\to 0} x_n \to \delta$ weakly. Here one may observe that the Riemann-integral of limit does not make sense and the Lebesgue-integral of limit is zero because $\lim_{n\to\infty} x_n(t) = 0$ for $t \neq 0$ and $\lim_{n\to\infty} x_n(t) = \infty$ for $t = 0$. We may also construct a sequence of functions as $x_m(t) = \frac{m}{2}\chi_{[-1/m,1/m]}(t)$ for $m \in \mathbb{N}$ such that $x_m \to \delta$ weakly as $m \to \infty$.

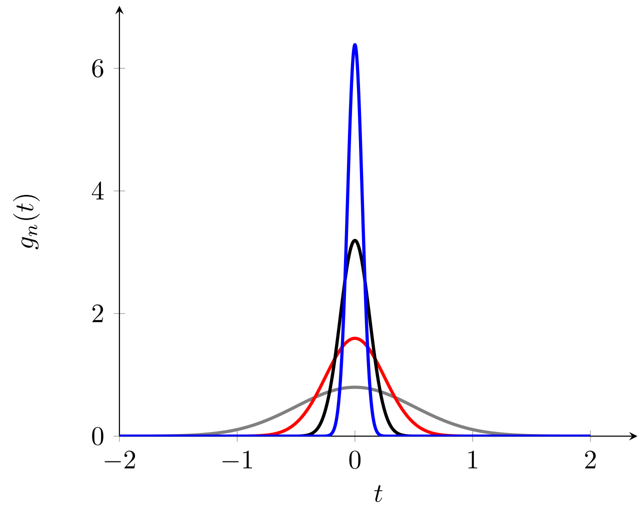

Fig. 2. The sequence of Gaussian functions (22) plotted for $n = 2, 4, 8, 16$, converge to delta function weakly as n approaches to infinity.

Example 4. The sequence of functions $\{x_n\}$ defined by $x_n(t) = \sin(2\pi n f_0 t)$, where $f_0 = 1/T$, converges weakly to the zero function in $\mathcal{L}_2[0,T]$. The integral

$$\int_0^T \sin(2\pi n f_0 t)\, g(t)\,\mathrm{d}t \to 0 \text{ as } n \to \infty \tag{25}$$

for any function $g \in \mathcal{L}_2[0,T]$ by Riemann–Lebesgue lemma, i.e., $\langle x_n, g \rangle = \langle 0, g \rangle = 0$. One can observe that x_n has an increasing number of zeroes in $[0,T]$ as $n \to \infty$; however, it is not equal to the zero function for any value of n. Moreover, x_n does not converge to zero in the $\mathcal{L}_p$-norms for any $p \in [1,\infty]$, and that is why this convergence is weak.

Weak derivative: The derivative of a distribution or the distributional derivative of a function is defined as:

$$\begin{aligned} \langle x', \phi \rangle &= \int_{-\infty}^{\infty} x'(t)\phi(t)\,\mathrm{d}t = x(t)\phi(t)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} x(t)\phi'(t)\,\mathrm{d}t, \\ &= -\int_{-\infty}^{\infty} x(t)\phi'(t)\,\mathrm{d}t = (-1)\langle x, \phi' \rangle, \end{aligned} \tag{26}$$

because $\phi(\pm\infty) = 0$ with $\phi(t)$ being a function of rapid decay. The derivative defined in (26) is a weak derivative, and so x' is a weak derivative of x. The result (26) can be generalized as

$$\langle x^{(n)}, \phi \rangle = (-1)^n \langle x, \phi^{(n)} \rangle \text{ for } \phi \in \mathcal{D} \tag{27}$$

where $\frac{\mathrm{d}}{\mathrm{d}t}x(t) = x'(t)$ and $\frac{\mathrm{d}^n}{\mathrm{d}t^n}\phi(t) = \phi^{(n)}(t)$. Thus, a distributional derivative is defined as

$$\langle T^{(n)}, \phi \rangle = (-1)^n \langle T, \phi^{(n)} \rangle \implies T^{(n)}(\phi) = (-1)^n T(\phi^{(n)}). \tag{28}$$

Therefore, every distribution has a derivative which is another distribution. On the other hand, every function may not have a derivative, but all functions have derivatives which are distributions. The product of two distributions is not defined in general, e.g., δ^2 is not defined. However, one can always multiply an element of $\mathcal{S}'$ with an element of $\mathcal{S}$, e.g., $\delta \in \mathcal{S}'$ and if $x \in \mathcal{S}$, then $\delta(t)x(t) = \delta(t)x(0)$.

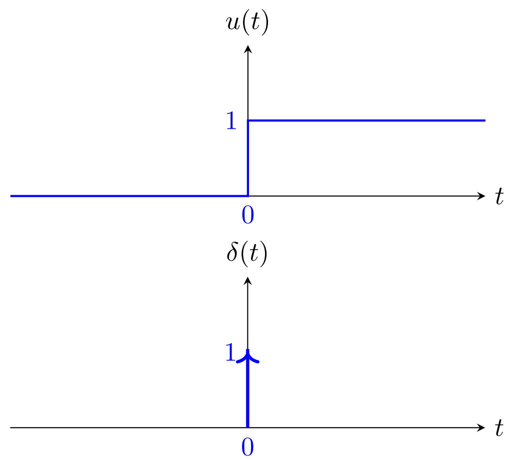

Fig. 3. Unit step function $u(t)$ (29), and its derivative the Dirac delta function $\delta(t)$.

Example 5. Let us consider the derivative of a unit step function defined as

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0, \end{cases} \tag{29}$$

and shown in Figure 3. This function is not differentiable at the point of discontinuity in the normal sense. Therefore, one can obtain the derivative of u in the distributional sense using (27) as $\langle u', \phi \rangle = (-1)\langle u, \phi' \rangle = (-1)\int_{\infty}^{\infty} u(t)\phi'(t)\,\mathrm{d}t = (-1)\int_0^{\infty} \phi'(t)\,\mathrm{d}t = (-1)[\phi(\infty) - \phi(0)] = \phi(0) = \langle \delta, \phi \rangle$ which implies $u' = \delta$. Since u is a tempered distribution, $u' = \delta$ is also a tempered distribution because the derivative of a tempered distribution is always a tempered distribution due to the following lemma [15].

Lemma 1. *If $T \in \mathcal{S}'$, then $T^{(n)} \in \mathcal{S}'$ for all $n \in \mathbb{N}_0$.*

B. Tempered distributions and FT

A nice property of tempered distributions $\mathcal{S}'$ is that the FT $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is a linear isomorphism because the FT $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ defined on $\mathcal{S}$ is a linear isomorphism [15]. However, in general, one cannot compute the FT of a regular distribution. If $x, \phi \in \mathcal{S}$, then by Fubini's Theorem:

$$\begin{aligned} \langle \widehat{x}, \phi \rangle &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x(t) \exp(-j2\pi f t)\,\mathrm{d}t \right) \phi(f)\,\mathrm{d}f, \\ &= \int_{-\infty}^{\infty} x(t) \left(\int_{-\infty}^{\infty} \phi(f) \exp(-j2\pi f t)\,\mathrm{d}f \right) \mathrm{d}t, \\ &= \langle x, \widehat{\phi} \rangle. \end{aligned} \tag{30}$$

Thus, the pairing $\langle \widehat{x}, \phi \rangle = \langle x, \widehat{\phi} \rangle$ provides a way to realize the definition of FT in general.

Definition 11. *Let $T \in \mathcal{S}'$, then its FT $\widehat{T} \in \mathcal{S}'$ is defined by $\langle \widehat{T}, \phi \rangle = \langle T, \widehat{\phi} \rangle$ for all $\phi \in \mathcal{S}$. The inverse FT $\check{T} \in \mathcal{S}'$ is defined by $\langle \check{T}, \phi \rangle = \langle T, \check{\phi} \rangle$ for all $\phi \in \mathcal{S}$.*

Theorem 3. *Fourier transform maps the class of tempered distributions onto itself: $T_x \in \mathcal{S}' \Leftrightarrow \widehat{T}_x \in \mathcal{S}'$, $\langle \widehat{T}_x, \phi \rangle = \langle T_x, \widehat{\phi} \rangle$, which implies $\langle \widehat{x}, \phi \rangle = \langle x, \widehat{\phi} \rangle$, where FT: $\mathcal{F}\{T_x\} = \widehat{T}_x = T_{\widehat{x}}$, $\mathcal{F}\{x\} = \widehat{x}$ and $\mathcal{F}\{\phi\} = \widehat{\phi}$.*

This theorem is true because $\phi \in \mathcal{S} \Leftrightarrow \widehat{\phi} \in \mathcal{S}$. The invertibility of the FT on $\mathcal{S}$ implies that $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ is invertible with inverse $\mathcal{F}^{-1} : \mathcal{S}' \to \mathcal{S}'$. Some typical examples of Schwartz function are $\phi(t) = \exp\big(-a\sqrt{1+t^2}\big)$ and $\phi(t) = \exp\big(-at^2\big)$ for all $a > 0$, which can be easily observed to satisfy this theorem. However, the above theorem is not true for $\phi \in \mathcal{D}$ because $\widehat{\phi} \notin \mathcal{D}$ due to the uncertainty principle of FT that states that if a signal is limited in the time-domain, its FT is unlimited in the frequency-domain and vice-versa. One can observe that if $\phi \in \mathcal{D}$, then $\widehat{\phi} \in \mathcal{S}$.

Since the space $\mathcal{S}$ is dense in $\mathcal{L}_2$ [15], the FT $\mathcal{S} \to \mathcal{S}$ extends by *continuity* to a map $\mathcal{F} : \mathcal{L}_2 \to \mathcal{L}_2$. Since $\mathcal{S}$ is mapped to itself by FT, this gives a way to define FT on $\mathcal{S}'$ by *duality* and by extending the Plancherel theorem: $\widehat{T}(\phi) = T(\widehat{\phi})$ for $\phi \in \mathcal{S}$ and $T \in \mathcal{S}'$. The FT on $\mathcal{S}'$ defined via *duality* agrees with the integral definition on $\mathcal{S} \subset \mathcal{S}'$. In other words, $\widehat{T}(\phi) = T(\widehat{\phi})$ for $\phi \in \mathcal{S}$, and $\langle \widehat{x}, \phi \rangle = \langle x, \widehat{\phi} \rangle = \int_{-\infty}^{\infty} x(t)\widehat{\phi}(t)\,\mathrm{d}t < \infty$.

Considering FT $\widehat{\phi}(f) = \int_{-\infty}^{\infty} \phi(t) \exp(-j2\pi f t)\,\mathrm{d}t$, $\widehat{\phi}^{(n)}(f) = \int_{-\infty}^{\infty} (-j2\pi t)^n \phi(t) \exp(-j2\pi f t)\,\mathrm{d}t$, IFT $\phi(t) = \int_{-\infty}^{\infty} \widehat{\phi}(f) \exp(j2\pi f t)\,\mathrm{d}f$ and its n-th derivative, $\phi^{(n)}(t) = \int_{-\infty}^{\infty} (j2\pi f)^n \widehat{\phi}(f) \exp(j2\pi f t)\,\mathrm{d}f$, we can write the FT and IFT of the n-th derivative as

$$(-j2\pi t)^n \phi(t) \rightleftharpoons \widehat{\phi}^{(n)}(f), \tag{31}$$

$$\phi^{(n)}(t) \rightleftharpoons (j2\pi f)^n \widehat{\phi}(f), \tag{32}$$

$$|\widehat{\phi}(f)| = |\mathcal{F}\{\phi^{(n)}(t)\}/(j2\pi f)^n| \leq \|\phi^{(n)}\|_1/(2\pi|f|)^n, \tag{33}$$

and observe that the greater differentiability or smoothness of ϕ leads to a faster decay of the FT.

Example 6. The FT of delta function using (12) can be obtained as: $\langle \widehat{\delta}, \phi \rangle = \langle \delta, \widehat{\phi} \rangle = \widehat{\phi}(0) = \int_{-\infty}^{\infty} \phi(t)\mathrm{d}t = \langle 1, \phi \rangle$ which implies $\widehat{\delta} = 1$ and thus, $\delta(t) \rightleftharpoons 1$. Similarly, The FT of 1 can be obtained as: $\langle \widehat{1}, \phi \rangle = \langle 1, \widehat{\phi} \rangle = \int_{-\infty}^{\infty} \widehat{\phi}(t)\mathrm{d}t = \phi(0) = \langle \delta, \phi \rangle$ which implies $\widehat{1} = \delta$ and thus $1 \rightleftharpoons \delta(f)$. Moreover, FT of $\exp(j2\pi f_0 t)$ can be obtained as follows: $\left\langle \widehat{\exp(j2\pi f_0 t)}, \phi \right\rangle = \left\langle \exp(j2\pi f_0 t), \widehat{\phi} \right\rangle = \int_{-\infty}^{\infty} \exp(j2\pi f_0 t)\widehat{\phi}(t)\mathrm{d}t = \phi(f_0) = \langle \delta_{f_0}, \phi \rangle$ which implies that $\widehat{\exp(j2\pi f_0 t)} = \delta_{f_0}$ and thus, $\exp(j2\pi f_0 t) \rightleftharpoons \delta(f - f_0)$. Therefore, $\cos(2\pi f_0 t) = \frac{1}{2}\big(\exp(j2\pi f_0 t) + \exp(-j2\pi f_0 t)\big) \rightleftharpoons \frac{1}{2}\big(\delta(f - f_0) + \delta(f + f_0)\big)$ and $\sin(2\pi f_0 t) = \frac{1}{2j}\big(\exp(j2\pi f_0 t) - \exp(-j2\pi f_0 t)\big) \rightleftharpoons \frac{1}{2j}\big(\delta(f - f_0) - \delta(f + f_0)\big)$.

Example 7. The Dirichlet function is a pathological function that is defined as

$$\chi_{\mathbb{Q}}(t) = \begin{cases} 1, & t \in \mathbb{Q}, \\ 0, & t \in \mathbb{P}, \end{cases} \tag{34}$$

where $\mathbb{Q}$ is the set of rational numbers, and $\mathbb{P} = \mathbb{R}\backslash\mathbb{Q}$ is the set of irrational numbers. The Dirichlet function is nowhere continuous. This is periodic as $\chi_{\mathbb{Q}}(t) = \chi_{\mathbb{Q}}(t + T)$ for any positive rational number T. It is not Riemann-integrable, whereas it is Lebesgue-integrable, and its integral is zero because it is zero except on the set of rational numbers which are countable and their Lebesgue measure is zero. This function satisfies the Dirichlet conditions [14] for the convergence of FT as there are countably infinite numbers of extrema (maxima and minima) and finite discontinuities, and absolutely integrable. Thus its FT converges to zero function, i.e., $\widehat{\chi_{\mathbb{Q}}} = 0$.

Similarly, the FT of the function

$$\chi_{\mathbb{P}}(t) = \begin{cases} 1, & t \in \mathbb{P}, \\ 0, & t \in \mathbb{Q}, \end{cases} \tag{35}$$

is the delta function, i.e., $\widehat{\chi_{\mathbb{P}}} = \delta$.
The above examples show that the pairing $\langle \widehat{x}, \phi \rangle = \langle x, \widehat{\phi} \rangle$ provides a way to realize the definition of FT in general.

Example 8. *FT of a polynomial in the distributional sense*: The FT of t^n can be obtained using the distributional theory as

$$\begin{aligned} \left\langle \widehat{(t^n)}, \phi \right\rangle &= \left\langle t^n, \widehat{\phi} \right\rangle = \int_{-\infty}^{\infty} t^n \widehat{\phi}(t) \, \mathrm{d}t \\ &= \frac{\phi^{(n)}(0)}{(j2\pi)^n} = \left\langle \left(\frac{j}{2\pi} \right)^n \delta^{(n)}, \phi \right\rangle \\ \implies t^n &\rightleftharpoons \left(\frac{j}{2\pi} \right)^n \delta^{(n)}(f) \quad \text{and} \qquad (36) \\ \delta^{(n)}(t) &\rightleftharpoons (j2\pi f)^n \qquad \text{for all } n \in \mathbb{N}_0. \qquad (37) \end{aligned}$$

Here, we have used $\langle \delta^{(n)}, \phi \rangle = (-1)^n \phi^{(n)}(0)$ and $\phi^{(n)}(0)/(j2\pi)^n = \int_{-\infty}^{\infty} f^n \widehat{\phi}(f) \, \mathrm{d}f$ from (32). Therefore, one can easily obtain the FT of a polynomial: $p_n(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ as $P_n(f) = a_0 \delta(f) + a_1 \left(\frac{j}{2\pi} \right) \delta^{(1)}(f) + a_2 \left(\frac{j}{2\pi} \right)^2 \delta^{(2)}(f) + \cdots + a_n \left(\frac{j}{2\pi} \right)^n \delta^{(n)}(f)$.

Example 9. *FT of $\frac{1}{t^{n+1}}$ for $n \geq 0$ in the distributional sense*: A signum function is defined as

$$\operatorname{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1 & t < 0 \end{cases} \tag{38}$$

and its derivative can be obtained as $\langle \operatorname{sgn}', \phi \rangle = (-1) \langle \operatorname{sgn}, \phi' \rangle = (-1) \int_{-\infty}^{\infty} \operatorname{sgn}(t) \phi'(t) \, \mathrm{d}t = \int_{-\infty}^{0} \phi'(t) \, \mathrm{d}t - \int_{0}^{\infty} \phi'(t) \, \mathrm{d}t = (\phi(0) - \phi(\infty)) - (\phi(\infty) - \phi(0)) = 2\phi(0) = \langle 2\delta, \phi \rangle$ which implies $\operatorname{sgn}' = 2\delta$. Using the FT of derivative of a function (32) and delta function, we can write $\mathcal{F}\{\operatorname{sgn}'\} = \mathcal{F}\{2\delta\} \implies (j2\pi f) \widehat{\operatorname{sgn}}(f) = 2$. Thus, FT of sgn can be written as

$$\operatorname{sgn}(t) \rightleftharpoons \widehat{\operatorname{sgn}}(f) = \frac{1}{j\pi f}. \tag{39}$$

Using the duality of FT, we can write $\frac{1}{\pi t} \rightleftharpoons -j \operatorname{sgn}(f)$ and using the differentiation property of FT from (32), we can write $\frac{(-1)^n}{\pi} \frac{n!}{t^{n+1}} \rightleftharpoons -j \operatorname{sgn}(f) (j2\pi f)^n$. This implies that

$$\frac{1}{t^{n+1}} \rightleftharpoons \frac{\pi}{n!} (-j)^{n+1} \operatorname{sgn}(f) (2\pi f)^n, \quad n = 0, 1, 2, \ldots, \tag{40}$$

$$\frac{\pi}{n!} j^{n+1} \operatorname{sgn}(t) (2\pi t)^n \rightleftharpoons \frac{1}{f^{n+1}}, \quad n = 0, 1, 2, \ldots \tag{41}$$

Since $\operatorname{sgn}(t) \in \mathcal{S}'$, it implies that $\widehat{\operatorname{sgn}}(f) = \frac{1}{j\pi f} \in \mathcal{S}'$ and hence, (40) and (41) are the FT pairs in the sense of tempered distributions. Unit step function and its FT can be obtained as $u(t) = \frac{1}{2}\big(1 + \operatorname{sgn}(t)\big) \implies \widehat{u}(f) = \frac{1}{2}\left(\delta(f) + \frac{1}{j\pi f}\right)$. From these results, one may observe that (i) the FT of an even function is a real-valued function, (ii) the FT of an odd function is an imaginary function, and (iii) the FT of neither an even nor an odd function is a complex-valued function.

Example 10. Let us consider signals $x(t)$ and $y(t)$ that are differentiable even and differentiable odd functions, respectively.

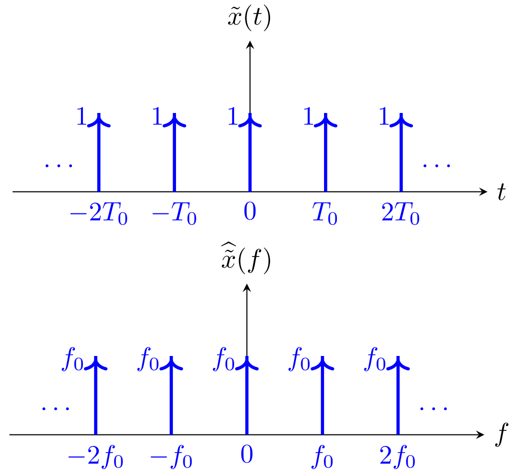

Fig. 4. A train of delta function $\tilde{x}(t)$ (42) and its FT $\widehat{\tilde{x}}(f)$.

This implies that $x'(t)$ and $y'(t)$ are odd and even functions, respectively. Mathematically, it is easy to show that $x(-t) = x(t) \implies x'(t) = -x'(-t)$, and $y(-t) = -y(t) \implies y'(t) = y'(-t)$. However, if the differentiation of $x(t)$ is an even function, then it does not imply that $x(t)$ function is an odd function. For example, if $x(t) = 1 + t$, then $x'(t) = 1$, which is an even function, yet $x(t)$ contains both even and odd part functions. This can also be observed from Example 5 and Example 9 as $u' = \delta$ and $\operatorname{sgn}'(t)/2 = \delta$, where delta is an even function, signum is an odd function, and unit step function is neither an even nor an odd function. Further, $\int \delta(t) \, \mathrm{d}t = \operatorname{sgn}(t)/2 + c = x(t)$. Now, if $x(0) = 0 \implies c = 0$ and $x(t) = \operatorname{sgn}(t)/2$, and if $x(0) = 1/2 \implies c = 1/2$ and $x(t) = \operatorname{sgn}(t)/2 + 1/2 = u(t)$. That is why FT of unit step is computed from $u(t) = \frac{1}{2}(1 + \operatorname{sgn}(t))$ and generally not from $u' = \delta$. Because differentiation kills the DC information present in a function, we have to add $c\delta(f)$ in the result corresponding to FT of DC component while computing the FT of x from x', i.e., $x(t) + c \rightleftharpoons \widehat{x}(f) + c\delta(f)$ and $x'(t) \rightleftharpoons (j2\pi f)\widehat{x}(f) + c(j2\pi f)\delta(f)$, where $c(j2\pi f)\delta(f) = 0$ providing $x'(t) \rightleftharpoons (j2\pi f)\widehat{x}(f)$. Therefore, FT of u from $u' = \delta$ can be obtained as follows: $(j2\pi f)\widehat{u}(f) + c(j2\pi f)\delta(f) = 1 \implies \widehat{u}(f) = \frac{1}{j2\pi f} - c\delta(f)$. Now, in order to obtain the value $-c$ we use $u(0) = \int_{-\infty}^{\infty} \widehat{u}(f) \, \mathrm{d}f = -c$, and by taking $u(0) = 1/2 \implies -c = 1/2$. This provides us the final result $\widehat{u}(f) = \frac{1}{2}\big(\frac{1}{j\pi f} + \delta(f)\big)$.

Example 11. Here, we consider the example of a train of delta functions, shown in Figure 4, which is well-known and is widely used in signal processing applications:

$$\tilde{x}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_0). \tag{42}$$

Let us compute both the FS and FT representations in the sense of distribution. Using (5), one can obtain Fourier coefficients $\widehat{x}_k = 1/T_0 = f_0$. Hence, its FS representation and the corresponding FT can be written as

$$\tilde{x}(t) = \frac{1}{T_0} \sum_{k=-\infty}^{\infty} \exp(jk\omega_0 t) \rightleftharpoons \frac{1}{T_0} \sum_{k=-\infty}^{\infty} \delta(f - kf_0), \tag{43}$$

where $\omega_0 = 2\pi f_0$ and $f_0 = 1/T_0$. One can also compute the FT of (42) as

$$\sum_{n=-\infty}^{\infty} \delta(t - nT_0) \rightleftharpoons \sum_{n=-\infty}^{\infty} \exp(-j2\pi f n T_0). \tag{44}$$

Thus, using the theory of distributions, we obtain

$$\sum_{n=-\infty}^{\infty} \exp(-j2\pi f n T_0) = \frac{1}{T_0} \sum_{k=-\infty}^{\infty} \delta(f - kf_0). \tag{45}$$

Example 12. The next logical question would be to explore the way to obtain the FT of $\exp(t)$ in the distributional sense using (36) and (37) because

$$\exp(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{j}{2\pi}\right)^n \delta^{(n)}(f), \tag{46}$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{j}{2\pi}\right)^n \delta^{(n)}(t) \rightleftharpoons \exp(-f) = \sum_{n=0}^{\infty} \frac{(-f)^n}{n!}, \tag{47}$$

$$\exp(-t) = \sum_{n=0}^{\infty} \frac{(-1)^n t^n}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-j}{2\pi}\right)^n \delta^{(n)}(f), \tag{48}$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-j}{2\pi}\right)^n \delta^{(n)}(t) \rightleftharpoons \exp(f) = \sum_{n=0}^{\infty} \frac{f^n}{n!}. \tag{49}$$

Since the growth of $\exp(t)$ is more than a polynomial, these FTs are not valid in the sense of tempered distributions. Corresponding to the function $x(t) = \exp(t)$, T_x is not a tempered distribution because $\exists\, \phi \in \mathcal{S}$ such that $\int_{-\infty}^{\infty} x(t)\phi(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \exp(t)\phi(t)\,\mathrm{d}t = \infty$. For example, $\phi(t) = \exp\bigl(-\sqrt{1+t^2}\bigr)$. One can observe that $\exp(t)\exp\bigl(-\sqrt{1+t^2}\bigr) \not\to 0$ as $t \to \infty$. In fact, $\exp(t)\exp\bigl(-\sqrt{1+t^2}\bigr) \to 0$ as $t \to -\infty$, and $\exp(t)\exp\bigl(-\sqrt{1+t^2}\bigr) \to 1$ as $t \to \infty$.

Now, the main questions are: Can we make FTs (46)–(49) valid in some sense? Can we define the bigger space than tempered distributions where FT is valid in the distributional sense? We answer these questions by proposing the space of Gauss functions in the next section.

C. The space of Gauss functions, exponential distributions and FT

The FT of a Gaussian function is again a Gaussian function: $\exp\bigl(-\alpha t^2\bigr) \rightleftharpoons \sqrt{\frac{\pi}{\alpha}}\exp\bigl(-\pi^2 f^2/\alpha\bigr)$ or $\frac{1}{\sqrt{2\pi\sigma^2}}\exp\bigl(-t^2/2\sigma^2\bigr) \rightleftharpoons \exp\bigl(-2\sigma^2\pi^2 f^2\bigr)$ for $\alpha > 0$, and $\exp\bigl(-\pi t^2\bigr) \rightleftharpoons \exp\bigl(-\pi f^2\bigr)$ when $2\sigma^2 = 1/\pi$. From this observation, we can define the space of Gauss functions as follows.

Definition 12. *Let $\mathcal{G} = \mathcal{G}(\mathbb{R}) = \{\phi \in C^\infty(\mathbb{R}) \mid \bigl(c_1 t^k + c_2 \exp(at)\bigr)\phi^{(m)}(t) \to 0$ as $|t| \to \infty\, \forall k, m \in \mathbb{N}_0,\ c_1, c_2, a \in \mathbb{C}\}$. This is the space of Gauss functions, where $\phi(t) = c\,\exp\bigl(-\sigma(t-t_0)^2\bigr)$ with $t_0 \in \mathbb{R}$, $c \in \mathbb{C}$, $\sigma > 0$, and all its derivatives have Gaussian type decay.*

One can easily observe that the set of Gaussian-type decay functions is a subset of the set of rapidly decaying functions. This $\mathcal{G}$ space is obtained by excluding many functions from the space $\mathcal{S}$ such as (i) all functions having a lower decay than Gaussian, e.g., $\phi(t) = \exp\bigl(-\sqrt{1+t^2}\bigr) \in \mathcal{S}$, but $\exp\bigl(-\sqrt{1+t^2}\bigr) \notin \mathcal{G}$, and (ii) all other functions which have a higher decay than Gaussian type decay, e.g., $\exp\bigl(-t^4\bigr)$, and compactly supported functions. Thus, the space $\mathcal{G}$ has a linear combination of set of functions $\{t^k \exp(at)\exp\bigl(-\sigma t^2\bigr)\}$ for $k \in \mathbb{N}_0$, $a \in \mathbb{C}$ and $\sigma \in (0, \infty)$. We can also consider $\sigma \in \mathbb{C}$ such that its real part $\mathrm{Re}(\sigma) > 0$.

Thus, the space of test functions $\mathcal{G}$, with $\phi(t) = c\,\exp\bigl(-\sigma t^2\bigr)$ for $c \in \mathbb{C}$ and $\sigma > 0$, is a linear subspace of $\mathcal{S}$ such that the FT can be defined for its dual space $\mathcal{G}'$. Since $\mathcal{G} \subset \mathcal{S}$, it implies that $\mathcal{S}' \subset \mathcal{G}'$. From a test function, $\phi(t) = \exp\bigl(-\sigma t^2\bigr)$, we can generate infinitely many test functions by (i) shifting $\phi(t - t_0)$, $\forall\, t_0 \in \mathbb{R}$, (ii) amplitude scaling $c_i\,\phi(t)$, $\forall\, c_i \in \mathbb{C}$, (iii) time scaling $\phi(\lambda t)$ for $\lambda \in \mathbb{R} \setminus \{0\}$, (iv) forming linear combinations $\sum_i c_i\phi_i(t)$, $\forall\, c_i \in \mathbb{C}$, and (v) considering products $\phi(t_1, \cdots, t_m) = \phi(t_1)\cdots\phi(t_m)$ to obtain examples in higher dimensions.

Clearly, $\mathcal{G}$ is a linear vector space because (i) $\phi_1 + \phi_2 \in \mathcal{G}$ for all $\phi_1, \phi_2 \in \mathcal{G}$, and (ii) $c\,\phi \in \mathcal{G}$ for all $c \in \mathbb{C}$ and $\phi \in \mathcal{G}$ and satisfies the properties of a vector space. The properties of space $\mathcal{G}$ are summarized as: (1) $\mathcal{G}$ is a linear vector space; (2) $\mathcal{G}$ is closed under multiplication; (3) $\mathcal{G}$ is closed under multiplication by polynomials and *exponentials*; (4) $\mathcal{G}$ is closed under differentiation; (5) $\mathcal{G}$ is closed under convolution; (6) $\mathcal{G}$ is closed under translations and multiplication by complex exponentials; and (7) The functions of $\mathcal{G}$ are integrable, i.e., $\int_{\mathbb{R}} |\phi(t)|\,\mathrm{d}t < \infty$ for $\phi \in \mathcal{G}$.

Definition 13. *An* ***exponential distribution*** *is a continuous linear functional on the space of Gauss functions, i.e., a mapping from $\mathcal{G} \to \mathbb{C}$. The* ***dual space*** *of $\mathcal{G}$ is denoted as $\mathcal{G}'$, which is a space of exponential distributions.*

An **exponential distribution** refers to a **distribution of at most exponential growth**, meaning thereby a growth that is **at most** $\exp(at)$ with $a \in \mathbb{R}$. This is to note that $\mathcal{G} \subset \mathcal{S} \implies \mathcal{S}' \subset \mathcal{G}'$. Since a test function $\phi \in \mathcal{G}$ and its FT $\mathcal{F}\{\phi\} = \widehat{\phi} \in \mathcal{G}$, $\langle \widehat{T_x}, \phi\rangle = \langle T_x, \widehat{\phi}\rangle \implies \langle \widehat{x}, \phi\rangle = \langle x, \widehat{\phi}\rangle$ for all $T_x \in \mathcal{G}'$, where $\phi(t) = c\,\exp\bigl(-\sigma t^2\bigr)$ for $c \in \mathbb{C}$ and $\sigma > 0$. A nice property of exponential distributions, $\mathcal{G}'$, is that the FT is a linear isomorphism for the dual space $\mathcal{G}'$, i.e., $\mathcal{F} : \mathcal{G}' \to \mathcal{G}'$ because FT is a linear isomorphism for the space of Gauss functions, i.e., $\mathcal{F} : \mathcal{G} \to \mathcal{G}$. Thus, we can obtain FT of $\exp(at)$ with $a \in \mathbb{R}$ in the distributional sense as

$$\exp(at) = \sum_{n=0}^{\infty} \frac{a^n t^n}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{a^n}{n!} \left(\frac{j}{2\pi}\right)^n \delta^{(n)}(f). \tag{50}$$

Example 13. Similarly, we can obtain FT of $\sin(1/t)$, $\cos(1/t)$

and $\exp(j/t)$ using (40) as

$$\sin\left(\frac{1}{t}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\frac{1}{t^{2n+1}} \rightleftharpoons \sum_{n=0}^{\infty} \frac{j(-1)^n}{(2n+1)!}\frac{\pi}{(2n)!}\,\mathrm{sgn}(f)(2\pi f)^{2n}, \tag{51}$$

$$\cos\left(\frac{1}{t}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\frac{1}{t^{2n}} \rightleftharpoons \delta(f) + \sum_{n=1}^{\infty} \frac{1}{(2n)!}\frac{\pi}{(2n-1)!}\,\mathrm{sgn}(f)(2\pi f)^{2n-1}, \tag{52}$$

$$\exp\left(\frac{j}{t}\right) = \sum_{n=0}^{\infty} \frac{j^n}{n!}\frac{1}{t^n} \rightleftharpoons \delta(f) + \sum_{n=1}^{\infty} \frac{1}{n!}\frac{\pi}{(n-1)!}\,\mathrm{sgn}(f)(2\pi f)^{n-1}. \tag{53}$$

The next logical question is: Can we further expand the scope of the FT for a larger class of signals? Can we define a larger space of distributions such that the FT is valid in that space? The answers to these questions are explored in the next section.

D. The proposed tempered superexponential distributions and FT

Let us consider a test function $\phi(t) = \exp\big(-\sigma t^2\big)$ for $\sigma > 0$, and $x(t) = \exp\big(\alpha t^2\big)$ such that $\exp\big(\alpha t^2\big)\phi^{(m)}(t) \to 0$ as $|t| \to \infty$ for all $\alpha < \sigma$, and $m \in \{0, 1, 2, \cdots\}$. We observe that

$$\langle \widehat{x}, \phi\rangle = \langle x, \widehat{\phi}\rangle = \sqrt{\frac{\pi}{\sigma}} \int_{-\infty}^{\infty} \exp\big(\alpha t^2\big)\exp\big(-\pi^2 t^2/\sigma\big)\mathrm{d}t = \sqrt{\frac{\pi}{\sigma}} \int_{-\infty}^{\infty} \exp\big(-t^2((\pi^2/\sigma)-\alpha)\big)\,\mathrm{d}t < \infty, \tag{54}$$

which is finite for only $\alpha < (\pi^2/\sigma)$ and $\alpha < \sigma \implies \alpha < \min(\sigma, \pi^2/\sigma)$. This is to note that $\alpha < \pi$ for $\sigma = \pi$ is the optimum value of σ. Based on these observations, we propose and define a linear space of **Gauss–Schwartz (GS)** functions as presented next.

Definition 14. *Let* $\mathcal{G}_s = \mathcal{G}_s(\mathbb{R}) = \{\phi \in C^\infty(\mathbb{R}) \mid \big(c_1 t^k + c_2\exp(at) + c_3\exp\big(\alpha t^2\big)\big)\,\phi^{(m)}(t) \to 0$ *as* $|t| \to \infty\, \forall k, m \in \mathbb{N}_0,\ c_1, c_2, c_3, a \in \mathbb{C},\ \alpha < \pi\}$. *This is the space of GS functions, where* $\phi(t) = c\exp\big(-\pi(t-t_0)^2\big)$ *with* $t_0 \in \mathbb{R}$, $c \in \mathbb{C}$, *and all its derivatives have Gaussian type decay.*

Thus, the space $\mathcal{G}_s$ has a linear combination of the set of functions $\{t^k\exp(at)\exp\big(-\sigma t^2\big)\}$ for $k \in \mathbb{N}_0$, $a \in \mathbb{C}$ and $\sigma = \pi$. We can also consider $\sigma \in \mathbb{C}$ such that its real part $\mathrm{Re}(\sigma) = \pi$. Here, a function of maximum decay is $\exp\big(-\pi t^2\big)$. It is well-known that $\mathcal{S}$ is the *largest* subspace of $\mathcal{L}_1(\mathbb{R})$ closed under multiplication by polynomials on which the FT and IFT exist. Similarly, the space of test functions $\mathcal{G}_s$ is the smallest linear subspace of $\mathcal{S}$, where FT can be defined for its dual space $\mathcal{G}_s'$. Since $\mathcal{G}_s \subset \mathcal{G}$, it implies that $\mathcal{G}' \subset \mathcal{G}_s'$. The space $\mathcal{G}_s$ is smallest in the sense that its dual space $\mathcal{G}_s'$ is the largest linear space over which FT can be defined by duality. The proposed space $\mathcal{G}_s(\mathbb{R})$ is dense in $\mathcal{L}_2(\mathbb{R})$ as the linear combinations of Gaussians with a single variance are dense in $\mathcal{L}_2(\mathbb{R})$ [24].

From a test function $\phi(t) = c\exp\big(-\pi t^2\big)$ for $c \in \mathbb{C}$, we can generate infinitely many test functions (i) by shifting $\phi(t - t_0)$, $\forall\, t_0 \in \mathbb{R}$, (ii) by amplitude scaling $c_i\,\phi(t)$, $\forall\, c_i \in \mathbb{C}$, (iii) by forming linear combinations $\sum_i c_i\phi_i(t)$, $\forall\, c_i \in \mathbb{C}$, and (iv) by considering products $\phi(t_1, \cdots, t_m) = \phi(t_1)\cdots\phi(t_m)$ to obtain examples in higher dimensions.

Clearly, $\mathcal{G}_s$ is a linear space because (i) $\phi_1 + \phi_2 \in \mathcal{G}_s$ for all $\phi_1, \phi_2 \in \mathcal{G}_s$ and (ii) $c\,\phi \in \mathcal{G}_s$ for all $c \in \mathbb{C}$ and $\phi \in \mathcal{G}_s$. Some properties of the GS space $\mathcal{G}_s$ are: (1) $\mathcal{G}_s$ is a vector space and is closed under linear combinations; (2) $\mathcal{G}_s$ is not closed under multiplication. For example, $c_1\exp\big(-\pi t^2\big) \in \mathcal{G}_s$ and $c_2\exp\big(-\pi t^2\big) \in \mathcal{G}_s$. However, $c_1 c_2\exp\big(-2\pi t^2\big) \notin \mathcal{G}_s$; (3) $\mathcal{G}_s$ is closed under multiplication by polynomials and exponentials; (4) $\mathcal{G}_s$ is closed under differentiation; (5) $\mathcal{G}_s$ is closed under translations and multiplication by complex exponentials; and (6) The functions of $\mathcal{G}_s$ are integrable, i.e., $\int_{\mathbb{R}}|\phi(t)|\,\mathrm{d}t < \infty$ for $\phi \in \mathcal{G}_s$.

Example 14. Let us consider a function $\phi(t) = \exp\big(-\pi t^2\big) \in \mathcal{G}_s$. The time shifting of $\phi(t)$ by t_0 corresponds to amplitude scaling and multiplication by an exponential as $\phi(t - t_0) = \exp\big(-\pi t_0^2\big)\exp(2\pi t_0 t)\exp\big(-\pi t^2\big)$. Similarly, the time shifting along with n-times differentiation ($\phi^{(n)}(t - t_0)$) corresponds to amplitude scaling, multiplication by an exponential, and an nth degree polynomial. The FT of these functions are (i) $\phi(t - t_0) \rightleftharpoons \exp(-j2\pi f t_0)\,\widehat{\phi}(f)$, (ii) $\phi^{(n)}(t - t_0) \rightleftharpoons (j2\pi f)^n\exp(-j2\pi f t_0)\,\widehat{\phi}(f)$, where $\phi(t) \rightleftharpoons \widehat{\phi}(f) = \exp\big(-\pi f^2\big)$. Therefore, the proposed GS space is closed under multiplication by only functions of polynomial and exponential growth and decay. Thus, the product of functions in $\mathcal{G}_s$ is not in $\mathcal{G}_s$ because the decay of the resulting function after multiplication will be faster than that defined for functions in $\mathcal{G}_s$.

Definition 15. *A* ***tempered superexponential distribution*** *is a continuous linear functional on the space of GS functions, i.e., it is a mapping from* $\mathcal{G}_s \to \mathbb{C}$. *The* ***dual space*** *of* $\mathcal{G}_s$ *is denoted as* $\mathcal{G}_s'$, *which is a space of tempered superexponential distributions.*

A **tempered superexponential distribution** (TSE) refers to a **distribution of temperate superexponential growth**, meaning thereby a growth that is **at most** $\exp\big(\alpha t^2\big)$ with $\alpha < \pi$. One can observe that the linear spaces follow the following containment: $\mathcal{D} \subset \mathcal{S} \subset \mathcal{L}_2 \subset \mathcal{S}' \subset \mathcal{G}' \subset \mathcal{G}_s' \subset \mathcal{D}'$ and $\mathcal{G}_s \subset \mathcal{G} \subset \mathcal{S} \subset \mathcal{L}_2 \subset \mathcal{S}' \subset \mathcal{G}' \subset \mathcal{G}_s' \subset \mathcal{D}'$ as shown in Figure 5. Moreover, $\mathcal{D} \not\subset \mathcal{G}_s$ and $\mathcal{D} \not\subset \mathcal{G}$. Since a test function $\phi \in \mathcal{G}_s$ and its FT $\mathcal{F}\{\phi\} = \widehat{\phi} \in \mathcal{G}_s$, $\langle \widehat{T_x}, \phi\rangle = \langle T_x, \widehat{\phi}\rangle \implies \langle \widehat{x}, \phi\rangle = \langle x, \widehat{\phi}\rangle$ for all $T_x \in \mathcal{G}_s'$, where $\phi(t) = c\exp\big(-\pi t^2\big)$ for $c \in \mathbb{C}$. A nice property of TSE distributions belonging to $\mathcal{G}_s'$ is that the FT is a linear isomorphism, i.e., $\mathcal{F} : \mathcal{G}_s' \to \mathcal{G}_s'$ because FT is a linear isomorphism for space of GS functions, i.e., $\mathcal{F} : \mathcal{G}_s \to \mathcal{G}_s$. Thus, we can obtain FT of $\exp\big(\alpha t^2\big)$ with $\alpha < \pi$ in the distributional sense as

$$\exp\big(\alpha t^2\big) = \sum_{n=0}^{\infty} \frac{\alpha^n t^{2n}}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{\alpha^n}{n!}\left(\frac{j}{2\pi}\right)^{2n}\delta^{(2n)}(f), \tag{55}$$

$$\sum_{n=0}^{\infty} \frac{\alpha^n}{n!}\left(\frac{j}{2\pi}\right)^{2n}\delta^{(2n)}(t) \rightleftharpoons \exp\big(\alpha f^2\big) = \sum_{n=0}^{\infty} \frac{\alpha^n f^{2n}}{n!}. \tag{56}$$

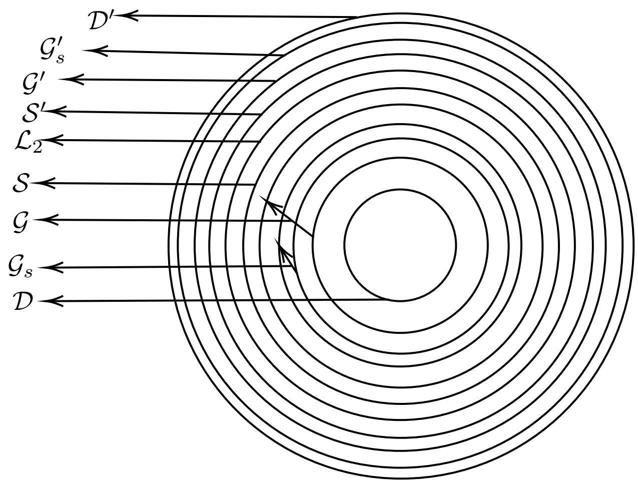

Fig. 5. Venn diagram representation of the linear spaces wherein $\mathcal{D} \subset \mathcal{S} \subset \mathcal{L}_2 \subset \mathcal{S}' \subset \mathcal{G}' \subset \mathcal{G}'_s \subset \mathcal{D}'$ and $\mathcal{G}_s \subset \mathcal{G} \subset \mathcal{S} \subset \mathcal{L}_2 \subset \mathcal{S}' \subset \mathcal{G}' \subset \mathcal{G}'_s \subset \mathcal{D}'$. Moreover, $\mathcal{D}$ is neither a subset of $\mathcal{G}_s$ nor a subset of $\mathcal{G}$.

V. Summary

Convergence of Fourier Series

FS exists if any one or more of the following conditions are fulfilled:

1) $\tilde{x} \in \mathcal{L}_1(\mathbb{T}) \cap BV(\mathbb{T})$.
2) $\tilde{x} \in BV(\mathbb{T})$.
3) $\tilde{x} \in \mathcal{L}_p(\mathbb{T})$ for $p > 1$.

Convergence of Fourier Transform

FT exists if one or more of the following conditions are satisfied:

1) $x \in \mathcal{L}_1(\mathbb{R})$.
2) $x \in \mathcal{L}_1(\mathbb{R}) \cap BV(\mathbb{R})$.
3) $x \in BV(\mathbb{R})$.
4) $x \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$.
5) $x \in \mathcal{L}_2(\mathbb{R})$.
6) $x \in \mathcal{L}_p([a, b])$ for $p \geq 1$, where $x(t) = 0,\ t \notin [a, b]$.
7) $x \in \mathcal{S}'(\mathbb{R})$ that has at most polynomial growth, FT exists in the sense of tempered distribution.
8) $x \in \mathcal{G}'(\mathbb{R})$ that has at most exponential growth, FT exists in the sense of exponential distribution.
9) $x \in \mathcal{G}'_s(\mathbb{R})$ that has at most tempered superexponential growth (i.e., $x(t) = \exp\big(\alpha t^2\big)$ for $\alpha < \pi$), FT exists in the sense of tempered superexponential distribution.

Thus in a nutshell: FS exists if $\tilde{x}$ is an $\mathcal{L}_p(\mathbb{T})$ function for $p > 1$; and FT exists if x is a tempered superexponential distribution $\mathcal{G}'_s(\mathbb{R})$.

VI. Conclusion

In this work, we have investigated and provided a detailed description of the various conditions that guarantee the existence of Fourier representation for a given signal, along with some suitable examples. The space of Gauss–Schwartz functions and corresponding distributions have been proposed. The distribution theory has been leveraged to show that FT can be defined for distributions of at most tempered superexponential growth. We have also elaborated on the interpretation of FT for some popular signals because these clarifications are lacking in the popular signal processing literature. The findings from this discussion can help in building a clear and complete understanding of Fourier theory for all researchers and practitioners working in the related areas.

References

[1] P. Singh, S. D. Joshi, R. K. Patney, and K. Saha, "The Fourier decomposition method for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 473: 20160871, pp. 1–27, 2017.

[2] P. Singh, A. Singhal, B. Fatimah, and A. Gupta, "A novel PRFB decomposition for non-stationary time-series and image analysis," *Signal Processing*, vol. 207, p. 108961, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016516842300035X

[3] J. Martinez, R. Heusdens, and R. C. Hendriks, "A generalized Fourier domain: Signal processing framework and applications," *Signal Processing*, vol. 93, no. 5, pp. 1259–1267, 2013. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165168412003775

[4] T. Zhang, L. Zhang, Z. Chen, Y. Xing, and H. Gao, "Fourier properties of symmetric-geometry computed tomography and its linogram reconstruction with neural network," *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4445–4457, 2020.

[5] H. Hu, D. Zhu, and F. Hu, "A novel imaging method using fractional Fourier transform for near-field synthetic aperture radiometer systems," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.

[6] X. Cunxiang, Z. Limin, and Z. Zhaogen, "Quasi-LFM radar waveform recognition based on fractional Fourier transform and time-frequency analysis," *Journal of Systems Engineering and Electronics*, vol. 32, no. 5, pp. 1130–1142, 2021.

[7] M. Alon, Y. Shor, A. Moshkin, T. Dagan, and S. Singer, "Envelope tracking power amplifier using short-time Fourier transform," *IEEE Microwave and Wireless Components Letters*, vol. 31, no. 6, pp. 575–578, 2021.

[8] D. Wei and J. Yang, "Non-uniform sparse Fourier transform and its applications," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4468–4482, 2022.

[9] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals and Systems*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[10] J. G. Proakis and D. G. Manolakis, *Digital Signal Processing Principles Algorithms and Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, 1996.

[11] S. Haykin and M. Moher, *Introduction to Analog and Digital Communications*. Hoboken, NJ, USA: Wiley, 2007.

[12] B. P. Lathi, *Signal Processing and Linear Systems*. Carmichael, CA, USA: Berkeley-Cambridge Press, 1998.

[13] P. Singh, A. Gupta, and S. D. Joshi, "General parameterized Fourier transform: A unified framework for the Fourier, Laplace, Mellin and Z transforms," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1295–1309, 2022.

[14] P. Singh, A. Singhal, B. Fatimah, A. Gupta, and S. D. Joshi, "Proper definitions of Dirichlet conditions and convergence of Fourier representations [lecture notes]," *IEEE Signal Processing Magazine*, vol. 39, no. 5, pp. 77–84, 2022.

[15] R. S. Strichartz, *A guide to distribution theory and Fourier transforms*. World Scientific Publishing Company, 1994.

[16] M. Stéphane, *A Wavelet Tour of Signal Processing, Third Edition*. Boston: Academic Press, 2009.

[17] L. Grafakos, *Classical Fourier Analysis, Graduate Texts in Mathematics, 249, Third Edition*. Springer, New York, 2008.

[18] K. B. Howell, "A new theory for Fourier analysis. Part I. the space of test functions," *Journal of Mathematical Analysis and Applications*, vol. 168, no. 2, pp. 342–350, 1992.

[19] L. Carleson, "On convergence and growth of partial sums of Fourier series," *Acta Math*, vol. 116, no. 1, pp. 137–157, 1966.

[20] R. A. Hunt, "On the convergence of Fourier series," *Proc. Conf. on orthogonal expansions and their continuous analogues, Edwardsville, IL, 1967, Illinois Univ. Press, Carbondale*, pp. 235–255, 1968.

[21] A. Kolmogoroff, "Une série de fourier-lebesgue divergente presque partout," *Fundamenta mathematicae*, vol. 4, no. 1, pp. 324–328, 1923.

[22] J. S. Walker, *Fourier series in Encyclopedia of Physical Science and Technology*. San Diego, CA, USA: Academic, 2002.

[23] C. Candan, "Proper definition and handling of dirac delta functions [lecture notes]," *IEEE Signal Processing Magazine*, vol. 38, no. 3, pp. 186–203, 2021.

[24] C. Calcaterra, "Linear combinations of Gaussians with a single variance are dense in L_2," in *Proceedings of the World Congress on Engineering*, vol. 2, 2008.