On the Fourier Representations and Schwartz Distributions

Pushpendra Singh*, Amit Singhal†, Binish Fatimah‡,
Anubha Gupta$§ and Shiv Dutt Joshi¶

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, and transmitting, processing, and analyzing the signals and systems. Undergraduate engineering students are exposed to these concepts, usually in their second year, to build their foundation in the areas of signal processing and communication engineering. However, the popular signal processing literature [1–4] does not offer a clear explanation regarding the convergence or the existence of Fourier representations for certain well-known signals. Because of this subtle gap, it becomes hard for young students to assimilate the Fourier theory with clarity, and they are forced to be familiar with some of these concepts without understanding them. To bring clarity to the existence and the convergence of Fourier representation, including Fourier series and transform, lecture notes were published recently in IEEE Signal Processing Magazine's September 2022 issue [5]. This lecture note is in the continuation with technical details added from yet another mathematical topic of 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. It is considered as a completion of differential calculus, similar to how the revolutionary measure theory or Lebesgue integration theory proposed in 1903, is considered as a completion of integral calculus. Both these theories unlocked new paradigms of mathematical development. Although distribution theory is a powerful tool for understanding Fourier theory, it is ignored in engineering textbooks. In this lecture note, 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.

Fourier representation is the most important mathematical tool. It has been used in almost all fields of science, mathematics, and engineering since its inception in 1807. Many studies present recent advancements and applications of the Fourier representation [6–21]. In fact, wavelet transform that emerged as a generalization of the Fourier theory to capture joint time-scale relationship also finds umpteen applications in signal processing [22–28].

Relevance

It is well known that the Fourier transform (FT) is defined for (i) the functions of at most polynomial growth (i.e., t^n for $n \in \{0, 1, 2, 3, \cdots\}$) in the sense of tempered distributions, and (ii) the functions of at most exponential growth (i.e., $\exp(at)$ for $a \in \mathbb{C}$) in the sense of distributions corresponding to the space of Gauss functions [29]. These aspects can be easily understood with the help of distribution theory. In this lecture note, 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 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, signal processing and other literature.

2. We extend the theory of FT by proposing the space of Gauss–Schwartz functions and the corresponding tempered super-exponential distributions. Thus, we define the FT for the functions of at most tempered super-exponential growth, i.e., $\exp(\alpha t^2)$, where $\alpha \in \mathbb{C}$ such that the real part $\mathrm{Re}(\alpha) < \pi$.

*P. Singh is with the Department of ECE, National Institute of Technology Hamirpur, Hamirpur (HP) India (pushpendrasingh@iitkalumni.org).

†A. Singhal is with the Department of ECE, Netaji Subhas University of Technology, Delhi, India (amit@nsut.ac.in).

‡B. Fatimah is with MaxEye Technologies Private Ltd, Bengaluru, India (binish.fatimah@gmail.com).

§A. Gupta is with the Department of ECE, Indraprastha Institute of Information Technology-Delhi (IIIT-D), Delhi, India. (Lab: SBILab: https://sbilab.iiitd.edu.in/) (anubha@iiitd.ac.in).$Corresponding author.

¶S.D. Joshi is with the Department of EE, Indian Institute of Technology Delhi, Delhi, India (sdjoshi@ee.iitd.ac.in).

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 and is computed as

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

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*

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

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

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 $t_0 \in \mathbb{R}$. The FS representation (4) is guaranteed for $\tilde{x}(t)$, if it satisfies the Dirichlet conditions [5], i.e., $\tilde{x}(t) \in \mathcal{L}_1(\mathbb{T}) \cap BV(\mathbb{T})$. Further, according to the Carleson–Hunt theorem [30,31], if $\tilde{x}(t) \in \mathcal{L}_p(\mathbb{T})$ for $p > 1$, then its FS converges at almost all points. The convergence is understood as how the sum [32]

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

converges to the original signal $\tilde{x}(t)$ while $N \to \infty$, i.e., whether it converges uniformly, point-wise, or in norm sense. Further details regarding the convergence of FS, along with some suitable examples, can be found in [5].

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 as

$$\text{Analysis Equation:} \quad \widehat{x}(\omega) = c_1 \int_{-\infty}^{\infty} x(t) \, \exp(-j\omega t) \, \mathrm{d}t, \qquad \text{for } x \in \mathcal{L}_1(\mathbb{R}) \tag{7}$$

$$\text{Synthesis Equation:} \quad x(t) = c_2 \int_{-\infty}^{\infty} \widehat{x}(\omega) \, \exp(j\omega t) \, \mathrm{d}\omega, \qquad \text{for } \widehat{x} \in \mathcal{L}_1(\mathbb{R}) \tag{8}$$

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

$$\widehat{x}(f) = \int_{-\infty}^{\infty} x(t) \, \exp(-j2\pi f t) \, \mathrm{d}t \qquad \text{and} \qquad x(t) = \int_{-\infty}^{\infty} \widehat{x}(f) \, \exp(j2\pi f t) \, \mathrm{d}f, \tag{9}$$

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 (9) as $\widehat{x}(t) \rightleftharpoons x(-f)$, which is very useful because it can provide FT that may be difficult to compute directly. From (9), we can write

$$\widehat{x}(0) = \int_{-\infty}^{\infty} x(t)\,\mathrm{d}t \qquad \text{and} \qquad x(0) = \int_{-\infty}^{\infty} \widehat{x}(f)\,\mathrm{d}f. \tag{10}$$

Also, we observe that

$$|\widehat{x}(f)| = \left|\int_{-\infty}^{\infty} x(t)\,\exp(-j2\pi ft)\,\mathrm{d}t\right| \leq \int_{-\infty}^{\infty} |x(t)|\,\mathrm{d}t = \|x\|_1 \qquad \text{and} \tag{11}$$

$$|x(t)| = \left|\int_{-\infty}^{\infty} \widehat{x}(f)\,\exp(j2\pi ft)\,\mathrm{d}f\right| \leq \int_{-\infty}^{\infty} |\widehat{x}(f)|\,\mathrm{d}f = \|\widehat{x}\|_1. \tag{12}$$

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 $x \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$.

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 [33,34] 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

$$\chi_{[-n,n]}(t) = \begin{cases} 1, & -n \leq t \leq n \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

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 [5,33,34] 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 $\|\widehat{x}\|_\infty \leq \|x\|_1$.
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 $1 \leq p \leq 2$ from the Riesz-Thorin Theorem $\mathcal{F} : \mathcal{L}_p(\mathbb{R}) \to \mathcal{L}_q(\mathbb{R})$, $\frac{1}{p} + \frac{1}{q} = 1$, $1 \leq p \leq 2$ with $\|\widehat{x}\|_q \leq \|x\|_p$ 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 $\|\widehat{x}\|_q \leq \|x\|_p$ 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 (9). If a signal is not an energy signal, then the FT integral (9) 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 [35], which we present in the next section.

Schwartz Distributions and Fourier Transform in Distributional Sense

Definition 3 *Let $\mathcal{D} = C_c^\infty(\mathbb{R})$ denotes the space of test functions that are infinitely differentiable and compactly supported.*

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, and $\mathbb{N}_0 = \mathbb{N}\bigcup\{0\} = \{0, 1, 2, 3, \dots\}$.*

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 lecture note 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 $c \in \mathbb{C}$: (i) $T(\phi_1 + \phi_2) = T(\phi_1) + T(\phi_2)$, (ii) $T(c\phi_1) = cT(\phi_1)$, and (iii) If $\phi_n \to \phi$, then $T(\phi_n) \to T(\phi)$. In general, the distribution of a test function $T(\phi)$ 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 $\mathcal{L}_2(\mathbb{R})$ [35]. 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.

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:

$$T_x(\phi) = \langle T_x, \phi\rangle = \int_{-\infty}^{\infty} x(t)\phi(t)\,\mathrm{d}t < \infty,\ \forall\phi \in \mathcal{D}, \tag{14}$$

where $T_x \in \mathcal{D}'$. In general, a function x determines a distribution T_x by (14). 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 (14). 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 $T_x = T_y \Leftrightarrow \langle x, \phi\rangle = \langle y, \phi\rangle \Leftrightarrow \langle x - y, \phi\rangle = 0 \Leftrightarrow x - y = 0 \Leftrightarrow x = y$.

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*.*

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 observe that

$$\lim_{\sigma\to 0} \int_{-\infty}^{\infty} g(t,\sigma)\phi(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \left(\lim_{\sigma\to 0} g(t,\sigma)\right)\phi(t)\,\mathrm{d}t \tag{15}$$

$$= \int_{-\infty}^{\infty} \delta(t)\phi(t)\,\mathrm{d}t = \phi(0) = \langle \delta, \phi \rangle, \tag{16}$$

where $\lim_{\sigma\to 0} g(t,\sigma) = \delta(t)$ is concentrated at $t = 0$. Similarly, one can define δ_μ concentrated at μ as

$$\delta(t-\mu) = \lim_{\sigma\to 0} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right). \tag{17}$$

From the above discussions, it is clear that (i) $\delta(t-\mu) = 0$ for $t \neq \mu$, (ii) $\delta(t-\mu) = \infty$ for $t = \mu$, and (iii) $\int_{-\infty}^{\infty} \delta(t-\mu)\,\mathrm{d}t = 1$. On using (17) in (16), we observe that the limiting function of (17) exhibits convolution operation in (16) as $(\delta * \phi)(t) = \phi(t)$.
The derivative of a distribution or the distributional derivative of a function is defined as:

$$\begin{aligned}\langle x', \phi \rangle &= \int_{-\infty}^{\infty} x'(t)\phi(t)\,\mathrm{d}t = x(t)\phi(t)|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} x(t)\phi'(t)\,\mathrm{d}t, \\ &= -\int_{-\infty}^{\infty} x(t)\phi'(t)\,\mathrm{d}t = (-1)\langle x, \phi' \rangle,\end{aligned} \tag{18}$$

because $\phi(\pm\infty) = 0$ with $\phi(t)$ being a function of rapid decay. This result can be generalized as

$$\langle x^{(n)}, \phi \rangle = (-1)^n \langle x, \phi^{(n)} \rangle \text{ for } \phi \in \mathcal{D} \tag{19}$$

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

$$\langle T^{(n)}, \phi \rangle = (-1)^n \langle T, \phi^{(n)} \rangle \implies T^{(n)}(\phi) = (-1)^n T(\phi^{(n)}). \tag{20}$$

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.

Example 1 Let us consider the derivative of a unit step function defined as

$$u(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0. \end{cases} \tag{21}$$

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 (19) 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 [35].

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

Tempered distributions and Fourier transform

A nice property of tempered distributions contained in $\mathcal{S}'$ is that the FT $\mathcal{F} : \mathcal{S}' \to \mathcal{S}'$ defined on $\mathcal{S}'$ is a linear isomorphism because FT $\mathcal{F} : \mathcal{S} \to \mathcal{S}$ defined on $\mathcal{S}$ is a linear isomorphism [35]. However, in general, one cannot compute the FT of a regular distribution.

If $x, \phi \in \mathcal{S}$, then using Fubini's Theorem:

$$\begin{aligned}
\langle \widehat{x}, \phi \rangle &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x(t) \exp(-j2\pi f t)\, \mathrm{d}t \right) \phi(f)\, \mathrm{d}f, \\
&= \int_{-\infty}^{\infty} x(t) \left(\int_{-\infty}^{\infty} \phi(f) \exp(-j2\pi f t)\, \mathrm{d}f \right) \mathrm{d}t, \\
&= \langle x, \widehat{\phi} \rangle. \qquad (22)
\end{aligned}$$

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, $\widehat{T}$, is defined by $\langle \widehat{T}, \phi \rangle = \langle T, \widehat{\phi} \rangle$ for $\phi \in \mathcal{S}$.*

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 $\phi \in \mathcal{S} \Leftrightarrow \widehat{\phi} \in \mathcal{S}$. Some typical examples of Schwartz function are $\phi_1(t) = \exp\bigl(-a\sqrt{1+t^2}\bigr)$ and $\phi_2(t) = \exp\bigl(-at^2\bigr)$ 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 $\widehat{\phi}$ is not in $\mathcal{D}$ 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.

Since the space $\mathcal{S}$ is dense in $\mathcal{L}_2$ [35], 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 $\widehat{\phi}(f) = \int_{-\infty}^{\infty} \phi(t) \exp(-j2\pi f t)\, \mathrm{d}t$, its n-th derivative, $\widehat{\phi}^{(n)}(f) = \int_{-\infty}^{\infty} (-j2\pi t)^n \phi(t) \exp(-j2\pi f t)\, \mathrm{d}t$, 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

$$(-j2\pi t)^n \phi(t) \rightleftharpoons \widehat{\phi}^{(n)}(f), \qquad (23)$$

$$\phi^{(n)}(t) \rightleftharpoons (j2\pi f)^n \widehat{\phi}(f), \qquad (24)$$

$$|\widehat{\phi}(f)| = |\mathcal{F}\{\phi^{(n)}(t)\}/(j2\pi f)^n| \leq \|\phi^{(n)}\|_1/(2\pi|f|)^n, \qquad (25)$$

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

Example 2 The FT of delta function using (10) 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, $\cos(2\pi f_0 t) = \frac{1}{2}\bigl(\exp(j2\pi f_0 t) + \exp(-j2\pi f_0 t)\bigr) \rightleftharpoons \frac{1}{2}\bigl(\delta(f - f_0) + \delta(f + f_0)\bigr)$ and $\sin(2\pi f_0 t) = \frac{1}{2j}\bigl(\exp(j2\pi f_0 t) - \exp(-j2\pi f_0 t)\bigr) \rightleftharpoons \frac{1}{2j}\bigl(\delta(f - f_0) - \delta(f + f_0)\bigr)$.

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

$$\begin{aligned}
&\left\langle \widehat{(t^n)}, \phi \right\rangle = \left\langle t^n, \widehat{\phi} \right\rangle = \int_{-\infty}^{\infty} t^n \widehat{\phi}(t)\, \mathrm{d}t = \frac{\phi^{(n)}(0)}{(j2\pi)^n} = \left\langle \left(\frac{j}{2\pi} \right)^n \delta^{(n)}, \phi \right\rangle \\
&\implies t^n \rightleftharpoons \left(\frac{j}{2\pi} \right)^n \delta^{(n)}(f) \quad \text{and} \qquad (26) \\
&\delta^{(n)}(t) \rightleftharpoons (j2\pi f)^n \qquad \text{for all}\, n \in \mathbb{N}_0. \qquad (27)
\end{aligned}$$

Here, we have used $\left\langle \delta^{(n)}, \phi \right\rangle = (-1)^n \phi^{(n)}(0)$ and $\phi^{(n)}(0)/(j2\pi)^n = \int_{-\infty}^{\infty} f^n \widehat{\phi}(f)\, \mathrm{d}f$ from (24). 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 4 *FT of $\frac{1}{t^{n+1}}$ for $n \geq 0$ in the distributional sense*: A signum function is defined as

$$\operatorname{sgn}(t) = \begin{cases} 1, & t > 0 \\ 0, & t = 0 \\ -1 & t < 0 \end{cases} \tag{28}$$

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 (24) 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

$$\operatorname{sgn}(t) \rightleftharpoons \widehat{\operatorname{sgn}}(f) = \frac{1}{j\pi f}. \tag{29}$$

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 (24), we can write $\frac{(-1)^n}{\pi}\frac{n!}{t^{n+1}} \rightleftharpoons -j\operatorname{sgn}(f)(j2\pi f)^n$. This implies that

$$\frac{1}{t^{n+1}} \rightleftharpoons \frac{\pi}{n!}(-j)^{n+1}\operatorname{sgn}(f)(2\pi f)^n, \quad n = 0, 1, 2, \ldots, \tag{30}$$

$$\frac{\pi}{n!}j^{n+1}\operatorname{sgn}(t)(2\pi t)^n \rightleftharpoons \frac{1}{f^{n+1}}, \quad n = 0, 1, 2, \ldots \tag{31}$$

Since $\operatorname{sgn}(t) \in \mathcal{S}'$, it implies that $\widehat{\operatorname{sgn}}(f) = \frac{1}{j\pi f} \in \mathcal{S}'$ and hence, (30) and (31) 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 5 Let us consider signals $x(t)$ and $y(t)$ that are differentiable even and differentiable odd functions, respectively. 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 1 and Example 4 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 6 Here, we consider the example of a train of delta functions, which is well-known and is widely used in signal processing applications:

$$\tilde{x}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_0). \tag{32}$$

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

$$\tilde{x}(t) = \frac{1}{T_0}\sum_{k=-\infty}^{\infty} \exp(jk\omega_0 t) \rightleftharpoons \frac{1}{T_0}\sum_{k=-\infty}^{\infty} \delta(f - kf_0), \tag{33}$$

where $\omega_0 = 2\pi f_0$ and $f_0 = 1/T_0$. One can also compute the FT of (32) as

$$\sum_{n=-\infty}^{\infty} \delta(t - nT_0) \rightleftharpoons \sum_{n=-\infty}^{\infty} \exp(-j2\pi f n T_0). \tag{34}$$

Thus, using the theory of distributions, we obtain

$$\sum_{n=-\infty}^{\infty} \exp(-j2\pi f n T_0) = \frac{1}{T_0} \sum_{k=-\infty}^{\infty} \delta(f - kf_0). \tag{35}$$

Example 7 The next logical question would be to explore the way to obtain the FT of $\exp(t)$ in the distributional sense using (27) because

$$\exp(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{j}{2\pi}\right)^n \delta^{(n)}(f), \tag{36}$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{j}{2\pi}\right)^n \delta^{(n)}(t) \rightleftharpoons \exp(-f) = \sum_{n=0}^{\infty} \frac{(-f)^n}{n!}, \tag{37}$$

$$\exp(-t) = \sum_{n=0}^{\infty} \frac{(-1)^n t^n}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-j}{2\pi}\right)^n \delta^{(n)}(f), \tag{38}$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{-j}{2\pi}\right)^n \delta^{(n)}(t) \rightleftharpoons \exp(f) = \sum_{n=0}^{\infty} \frac{f^n}{n!}. \tag{39}$$

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, $\phi(t) = \exp\big(-\sqrt{1+t^2}\big)$. One can observe that $\exp(t)\exp\big(-\sqrt{1+t^2}\big) \not\to 0$ as $t \to \infty$. In fact, $\exp(t)\exp\big(-\sqrt{1+t^2}\big) \to 0$ as $t \to -\infty$, and $\exp(t)\exp\big(-\sqrt{1+t^2}\big) \to 1$ as $t \to \infty$.

Now, the main questions are: Can we make FTs (36)–(39) 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.

The proposed space of Gauss functions in the theory of distributions

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

Definition 12 *Let $\mathcal{G} = \mathcal{G}(\mathbb{R}) = \{\phi \in C^\infty(\mathbb{R}) \,|\, \big(c_1 t^k + c_2\exp(at)\big)\,\phi^{(m)}(t) \to 0$ 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 $\phi(t) = c\exp\big(-\sigma(t-t_0)^2\big)$ 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., $\phi(t) = \exp\big(-\sqrt{1+t^2}\big) \in \mathcal{S}$, but $\exp\big(-\sqrt{1+t^2}\big) \notin \mathcal{G}$, and (ii) all other functions which have a higher decay than Gaussian type decay, e.g., $\exp\big(-t^4\big)$, and compactly supported functions. Thus, the space $\mathcal{G}$ has a linear combination of set of functions $\{t^k\exp(at)\exp\big(-\sigma t^2\big)\}$ for $k \in \mathbb{N}_0$, $a \in \mathbb{C}$ and $\sigma \in (0,\infty)$. 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 $\phi(t) = c\exp\big(-\sigma t^2\big)$ 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, $\phi(t) = \exp\big(-\sigma t^2\big)$, 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 $\phi(t) = c\,\exp\!\big(-\sigma t^2\big)$ 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

$$\exp(at) = \sum_{n=0}^{\infty} \frac{a^n t^n}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{a^n}{n!}\left(\frac{j}{2\pi}\right)^n \delta^{(n)}(f). \tag{40}$$

Example 8 Similarly, we can obtain FT of $\sin(1/t)$, $\cos(1/t)$ and $\exp(j/t)$ using (30) as

$$\sin\!\left(\frac{1}{t}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\frac{1}{t^{2n+1}} \rightleftharpoons \sum_{n=0}^{\infty} \frac{j(-1)^n}{(2n+1)!}\frac{\pi}{(2n)!}\,\mathrm{sgn}(f)(2\pi f)^{2n}, \tag{41}$$

$$\cos\!\left(\frac{1}{t}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\frac{1}{t^{2n}} \rightleftharpoons \delta(f) + \sum_{n=1}^{\infty} \frac{1}{(2n)!}\frac{\pi}{(2n-1)!}\,\mathrm{sgn}(f)(2\pi f)^{2n-1}, \tag{42}$$

$$\exp\!\left(\frac{j}{t}\right) = \sum_{n=0}^{\infty} \frac{j^n}{n!}\frac{1}{t^n} \rightleftharpoons \delta(f) + \sum_{n=1}^{\infty} \frac{1}{n!}\frac{\pi}{(n-1)!}\,\mathrm{sgn}(f)(2\pi f)^{n-1}. \tag{43}$$

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.

Tempered superexponential distribution

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

$$\langle \widehat{x}, \phi\rangle = \langle x, \widehat{\phi}\rangle = \sqrt{\frac{\pi}{\sigma}} \int_{-\infty}^{\infty} \exp\!\big(\alpha t^2\big)\exp\!\big(-\pi^2 t^2/\sigma\big)\mathrm{d}t = \sqrt{\frac{\pi}{\sigma}} \int_{-\infty}^{\infty} \exp\!\big(-t^2((\pi^2/\sigma) - \alpha)\big)\,\mathrm{d}t < \infty, \tag{44}$$

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 $\mathcal{L}_1(\mathbb{R})$. 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. 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, $c_1c_2\exp\big(-2\pi t^2\big) \notin \mathcal{G}_s$; (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 9 Let us consider a function $\phi(t) = \exp\big(-\pi t^2\big) \in \mathcal{G}_s$. The time shifting of this function by t_0 corresponds to amplitude scaling and multiplication by an exponential because $\phi(t-t_0) = \exp\big(-\pi t_0^2\big)\,\exp(2\pi t_0 t)\,\exp\big(-\pi t^2\big)$. 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 1. 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

$$\exp\big(\alpha t^2\big) = \sum_{n=0}^{\infty} \frac{\alpha^n t^{2n}}{n!} \rightleftharpoons \sum_{n=0}^{\infty} \frac{\alpha^n}{n!}\left(\frac{j}{2\pi}\right)^{2n} \delta^{(2n)}(f), \tag{45}$$

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

Summary

Fourier Series

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

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

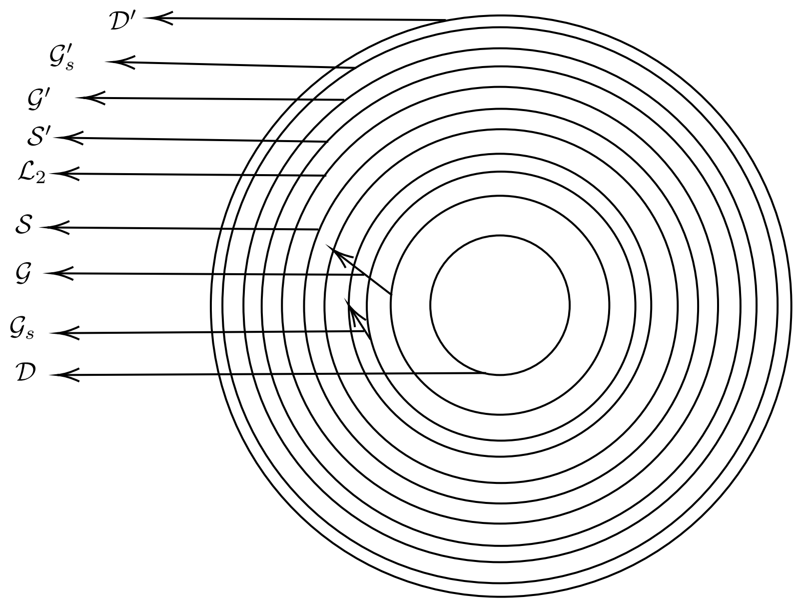

Figure 1: Venn diagram representation of the linear spaces wherein $\mathcal{D} \subset \mathcal{S} \subset \mathcal{L}_2 \subset \mathcal{S}' \subset \mathcal{G}' \subset \mathcal{G}_s' \subset \mathcal{D}'$ and $\mathcal{G}_s \subset \mathcal{G} \subset \mathcal{S} \subset \mathcal{L}_2 \subset \mathcal{S}' \subset \mathcal{G}' \subset \mathcal{G}_s' \subset \mathcal{D}'$. Moreover, $\mathcal{D}$ is neither a subset of $\mathcal{G}_s$ nor a subset of $\mathcal{G}$.

Fourier Transform

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

1. $x(t) \in \mathcal{L}_1(\mathbb{R}) \cap BV(\mathbb{R})$.
2. $x(t) \in \mathcal{L}_1(\mathbb{R}) \cap \mathcal{L}_2(\mathbb{R})$.
3. $x(t) \in \mathcal{L}_2(\mathbb{R})$.
4. $x(t) \in \mathcal{L}_p([a,b])$ for $p > 1$, where $x(t) = 0$, $t \notin [a,b]$.
5. If $x(t) \in \mathcal{S}'(\mathbb{R})$ has at most polynomial growth, FT exists in the sense of tempered distribution.
6. If $x(t) \in \mathcal{G}'(\mathbb{R})$ has at most exponential growth, FT exists in the sense of exponential distribution.
7. If $x(t) \in \mathcal{G}_s'(\mathbb{R})$ has at most superexponential growth, FT exists in the sense of tempered superexponential distribution.

What we have learned

This lecture note provided a detailed description of the various conditions that guarantee the existence of Fourier representation for a given signal, along with some suitable examples. A brief discussion is presented to understand the basics of distribution theory. The space of Gauss–Schwartz functions and corresponding distributions are 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 both the students and researchers working in the related areas.

References

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

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

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

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

[5] P. Singh, A. Singhal, B. Fatimah, A. Gupta, S. D. Joshi, Proper definitions of Dirichlet conditions and convergence of Fourier representations [lecture notes], IEEE Signal Processing Magazine 39 (5) (2022) 77–84. doi:10.1109/MSP.2022.3172620.

[6] P. Singh, S. D. Joshi, R. K. Patney, 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 473: 20160871 (2017) 1–27.

[7] P. Singh, Novel Fourier quadrature transforms and analytic signal representations for nonlinear and non-stationary time series analysis, Royal Society Open Science 5: 181131 (2018) 1–26.

[8] P. Singh, S. D. Joshi, Some studies on multidimensional Fourier theory for Hilbert transform, analytic signal and AM–FM representation, Circuits, Systems, and Signal Processing 38 (12) (2019) 5623–5650. doi:10.1007/s00034-019-01133-x.

[9] P. Singh, Breaking the limits: Redefining the instantaneous frequency, Circuits Syst Signal Process 37 (2018) 3515–3536. doi:10.1007/s00034-017-0719-y.

[10] P. Singh, A. Gupta, S. D. Joshi, General parameterized Fourier transform: A unified framework for the Fourier, Laplace, Mellin and Z transforms, IEEE Transactions on Signal Processing 70 (2022) 1295–1309. doi:10.1109/TSP.2022.3152607.

[11] A. Gupta, S. D. Joshi, P. Singh, On the approximate discrete KLT of fractional Brownian motion and applications, Journal of the Franklin Institute 355 (2018) 8989–9016.

[12] N. Ansari, A. Sen Gupta, A. Gupta, Shallow water channel estimation with energy efficient transmitted signal design, The Journal of the Acoustical Society of America 145 (5) (2019) 2955–2970.

[13] N. Ansari, A. Gupta, A. S. Gupta, Shallow water acoustic channel estimation using two-dimensional frequency characterization, The Journal of the Acoustical Society of America 140 (5) (2016) 3995–4009.

[14] A. S. Udawat, P. Singh, An automated detection of atrial fibrillation from single-lead ECG using HRV features and machine learning, Journal of Electrocardiology (2022). doi:https://doi.org/10.1016/j.jelectrocard.2022.07.069.

[15] B. Fatimah, A. Singhal, P. Singh, A multi-modal assessment of sleep stages using adaptive Fourier decomposition and machine learning, Computers in Biology and Medicine 148 (2022) 105877. doi:https://doi.org/10.1016/j.compbiomed.2022.105877.

[16] B. Fatimah, P. Singh, A. Singhal, D. Pramanick, P. S., R. B. Pachori, Efficient detection of myocardial infarction from single lead ECG signal, Biomedical Signal Processing and Control 68 (2021) 102678. doi:https://doi.org/10.1016/j.bspc.2021.102678.

[17] B. Fatimah, P. Singh, A. Singhal, R. B. Pachori, Hand movement recognition from sEMG signals using Fourier decomposition method, Biocybernetics and Biomedical Engineering 41 (2) (2021) 690–703. doi:https://doi.org/10.1016/j.bbe.2021.03.004.

[18] A. Singhal, P. Singh, B. Lall, S. D. Joshi, Modeling and prediction of covid-19 pandemic using Gaussian mixture model, Chaos, Solitons & Fractals 138 (2020) 110023. doi:https://doi.org/10.1016/j.chaos.2020.110023.

[19] B. Fatimah, P. Singh, A. Singhal, R. B. Pachori, Detection of apnea events from ECG segments using Fourier decomposition method, Biomedical Signal Processing and Control 61 (2020) 102005. doi:https://doi.org/10.1016/j.bspc.2020.102005.

[20] A. Singhal, P. Singh, B. Fatimah, R. B. Pachori, An efficient removal of power-line interference and baseline wander from ECG signals by employing Fourier decomposition technique, Biomedical Signal Processing and Control 57 (2020) 101741. doi:https://doi.org/10.1016/j.bspc.2019.101741.

[21] P. Singh, A. Singhal, B. Fatimah, A. Gupta, S. D. Joshi, AF-MNS: A novel AM-FM based measure of non-stationarity, IEEE Communications Letters 25 (3) (2021) 990–994. doi:10.1109/LCOMM.2020.3041722.

[22] A. Gupta, S. D. Joshi, S. Prasad, A new approach for estimation of statistically matched wavelet, IEEE transactions on signal processing 53 (5) (2005) 1778–1793.

[23] A. Gupta, S. D. Joshi, S. Prasad, A new method of estimating wavelet with desired features from a given signal, Signal Processing 85 (1) (2005) 147–161.

[24] A. Gupta, S. Joshi, Two-channel nonseparable wavelets statistically matched to 2-D images, Signal processing 91 (4) (2011) 673–689.

[25] N. Ansari, A. Gupta, Image reconstruction using matched wavelet estimated from data sensed compressively using partial canonical identity matrix, IEEE Transactions on image Processing 26 (8) (2017) 3680–3695.

[26] N. Ansari, A. Gupta, M-RWTL: learning signal-matched rational wavelet transform in lifting framework, IEEE Access 6 (2018) 12213–12227.

[27] N. Ansari, A. Gupta, WNC-ECGlet: Weighted non-convex minimization based reconstruction of compressively transmitted ECG using ECGlet, Biomedical Signal Processing and Control 49 (2019) 1–13.

[28] S. Gehlot, A. Gupta, N. Ansari, WTL-I: Mutual information based wavelet transform learning for hyperspectral imaging, Frontiers in Signal Processing (2022) 24.

[29] K. B. Howell, A new theory for Fourier analysis. Part I. the space of test functions, Journal of Mathematical Analysis and Applications 168 (2) (1992) 342–350. doi:10.1016/0022-247X(92)90162-7.

[30] L. Carleson, On convergence and growth of partial sums of Fourier series, Acta Math 116 (1) (1966) 137–157. doi:10.1007/BF02392815.

[31] 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 (1968) 235–255.

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

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

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

[35] R. S. Strichartz, A guide to distribution theory and Fourier transforms, World Scientific Publishing Company, 1994. doi:https://doi.org/10.1142/5314.