ON CONVOLUTION GROUPS OF COMPLETELY MONOTONE SEQUENCES/FUNCTIONS AND FRACTIONAL CALCULUS

LEI LI AND JIAN-GUO LIU

Abstract. We study convolution groups generated by completely monotone sequences and completely monotone functions. Using a convolution group, we define a fractional calculus for a certain class of distributions. When acting on causal functions, this definition agrees with the traditional Riemann-Liouville definition for $t > 0$ but includes some singularities at $t = 0$ so that the group property holds. Using this group, we are able to extend the definition of Caputo derivatives of order in $(0, 1)$ to a certain class of locally integrable functions without using the first derivative. The group property allows us to de-convolve the fractional differential equations to integral equations with completely monotone kernels, which then enables us to prove the general Gronwall inequality (or comparison principle) with the most general conditions. This then opens the door of a priori energy estimates of fractional PDEs. Some other fundamental results for fractional ODEs are also established within this frame under very weak conditions. Besides, we also obtain some interesting results about completely monotone sequences.

1. Introduction

A sequence $c = \{c_k\}_{k=0}^{\infty}$ is completely monotone if $(I - S)^j c_k \geq 0$ for any $j \geq 0, k \geq 0$ where $Sc_j = c_{j+1}$. A sequence is completely monotone if and only if it is the moment sequence of a Hausdorff measure (a finite nonnegative measure on $[0, 1]$) ([26]). Completely monotone sequences are closely related to infinitely divisible probability distributions on $\mathbb{N}$. In [18, 22], a nice description of completely monotone sequences is given:

Lemma 1. *A sequence c is completely monotone if and only if the generating function $F(z) = \sum_{j=0}^{\infty} c_j z^j$ is a Pick function that is analytic and nonnegative on $(-\infty, 1)$.*

A function $f : \mathbb{C}_+ \to \mathbb{C}$ (where $\mathbb{C}_+$ denotes the upper half plane, not including the real line) is Pick if it is analytic such that $Im(z) > 0 \Rightarrow Im(f(z)) \geq 0$. Note that if $f(z)$ is Pick and $Im(f(z)) = 0$ for some $Im(z) > 0$, then $Im(f(z)) = 0$ for all z. By the theory of continuation, if $f(z)$ is real on some interval (a, b), then the function can be extended to $\mathbb{C}_+ \cup (a, b) \times \{0\} \cup \mathbb{C}_-$ by reflection.

Consider the convolution $a * c$ defined by $(a * c)_k = \sum_{n_1 \geq 0, n_2 \geq 0} \delta_k^{n_1+n_2} a_{n_1} c_{n_2}$, which is associative and commutative. If we use $F_c(z)$ to mean the generating function of c, then it is clear that

$$F_{a*c}(z) = F_a(z) F_c(z). \tag{1}$$

If c is completely monotone, it is shown that there exist $c^{(r)}$, $r \in \mathbb{R}$, such that $c^{(r)} * c^{(s)} = c^{(r+s)}$ and $c^{(1)} = c$, i.e. there exists a convolution group generated by the completely monotone sequence ([18]). If $0 \leq r \leq 1$, $c^{(r)}$ is completely monotone. Further,

$$c^{(0)} = \delta_d = (1, 0, 0, \ldots),$$

2010 *Mathematics Subject Classification.* Primary 47D03, secondary 34A08, 46F10.

Key words and phrases. convolution group, fractional calculus, completely monotone sequence, completely monotone function, Riemann-Liouville derivative, Caputo derivative, fractional ODE.

is the convolution identity. An algorithm has been proposed in [18] to obtain the convolution group generated by c using its canonical sequence. The most interesting sequence is $c^{(-1)}$, the convolution inverse, which can be used for deconvolution.

Correspondingly, a function $g : (0,\infty) \to \mathbb{R}$ is completely monotone if $(-1)^n f^{(n)} \ge 0$ for $n = 0, 1, 2, \dots$. The famous Bernstein theorem says that a function is completely monotone if and only if it is the Laplace transform of a Radon measure on $[0,\infty)$ ([26, 23, 4]). Completely monotone functions appear in fractional calculus, which has drawn much attention to model memory effects in recent years ([12, 19, 25, 1]). To see this, consider the fractional integral

$$J_\gamma f = \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s)\,ds, \quad 0 < \gamma \le 1.$$

The kernel $\frac{1}{\Gamma(\gamma)} t^{\gamma-1}$ is completely monotone. The fractional integral is just the convolution between the kernel and f. We may thus expect the fractional derivative to be determined by the convolution inverse and the fractional calculus may be given by the convolution group generated by these completely monotone functions. However, it is not clear how the convolution group can be generated by a completely monotone function as this must be put in the frame of distributions, while in general the convolution between two distributions is not defined. We will aim to define the convolution group generated by the specific kernel $\frac{1}{\Gamma(\gamma)} t^{\gamma-1}$.

The Caputo derivatives ([12, 16, 5]) do not have group property, but are suitable for initial value problems and share many properties with the ordinary derivative. In the traditional definition, one has to define the γ-th order derivative ($0 < \gamma < 1$) using the first order derivative. In [1], a definition based on integration by parts is proposed and the first order derivative is not needed but the function has to possess some regularity. We will use the convolution group to generalize the Caputo derivatives so that the first order derivative is not needed either, and they are defined on a larger class of locally integrable functions. In a much weaker sense, we show that all the fundamental properties for Caputo derivatives under this new definition still hold.

The rest of the paper is organized as follows. In Section 2 we first investigate the convolution inverse of a completely monotone sequence and show that the inverse is well-behaved. Based on this, a preliminary iterative method is proposed for deconvolution. In Section 3, we introduce a specific class of distributions and generalize the traditional convolution between two distributions where one is required to have compact support to this class. A convolution group is then constructed and used to define a fractional calculus for the distributions in this class. When acting on causal functions, this definition agrees with the famous Riemann-Liouville fractional calculus for $t > 0$. At $t = 0$, some singularities must be included to make the fractional calculus a group. In Section 4, we prove a regularity result for the fractional calculus when acting on a special class of Sobolev spaces. In Section 5, using the convolution group, an extension of Caputo derivatives is proposed so that the ordinary derivative of the function is not needed in the definition. Some properties of the new Caputo derivatives are proved, which may be used for fractional ODEs (FODE) and fractional PDEs (FPDE). Especially, the fundamental theorem of the fractional calculus is valid with the most general conditions by deconvolution using the group property, which allows us to transform the differential equations with orders in $(0, 1)$ to integral equations with completely monotone kernels. In Section 6, based on the definitions and properties in Section 5, we prove some fundamental results of FODEs with quite general conditions. Especially, we show the existence and uniqueness of the FODEs using the fundamental theorem, and also show the general Gronwall inequalities. Finally, in Section 7, we define a discrete fractional calculus using a discrete convolution group generated by a specific completely monotone sequence and show that it is consistent with the Riemann-Liouville calculus.

2. Deconvolution for a completely monotone kernel

Consider the convolution equation

$$a * c = f, \tag{2}$$

where c is a completely monotone sequence and $c_0 > 0$. If we find the convolution inverse of c, the equation can be solved. We start with the properties of the convolution inverse.

2.1. The convolution inverse.

We now investigate the property of $c^{(-1)}$, whose generating function is $1/F_c(z)$. To be convenient, we use $F(z)$ to mean $F_c(z)$, the generating function of c.

Theorem 1. *Suppose c is completely monotone and $c_0 > 0$. Let $c^{(-1)}$ be its convolution inverse. Then, $F_{c^{(-1)}}$ is analytic on the open unit disk, and thus the radius of convergence of its power series around $z = 0$ is at least 1. $c_0^{(-1)} = 1/c_0$ and the sequence $(-c_1^{(-1)}, -c_2^{(-1)}, \dots)$ is completely monotone. Furthermore,*

$$0 \le -\sum_{k=1}^{\infty} c_k^{(-1)} \le \frac{1}{c_0}. \tag{3}$$

Proof. The first claim follows from that $F(z)$ has no zeros in the unit disk [18].

By Lemma 1, $F(z)$ is Pick and it is positive on $(-\infty, 1)$. $F(-\infty) = 0$ if the corresponding Hausdorff measure does not have an atom at 0 (i.e. the sequence c is minimal. See [26, Chap. IV. Sec. 14] for the definition). Since $F(-\infty)$ could be zero, we consider

$$G_\epsilon = \frac{1}{\epsilon} - \frac{1}{\epsilon + F(z)}, \ \epsilon > 0.$$

It is easy to verify that G_ϵ a Pick function, analytic and nonnegative on $(-\infty, 1)$.

Suppose G_ϵ is the generating function of $d = (d_0^\epsilon, d_1^\epsilon, \dots)$. By Lemma 1, this sequence is completely monotone. Then,

$$H_\epsilon = \frac{1}{z}[G_\epsilon(z) - G_\epsilon(0)] = \frac{F(z) - F(0)}{z(\epsilon + F(0))(\epsilon + F(z))},$$

is the generating function of the shifted sequence $(d_1^\epsilon, \dots)$, which is completely monotone. Hence, H_ϵ is also a Pick function, nonnegative and analytic on $(-\infty, 1)$.

Taking the pointwise limit of H_ϵ as $\epsilon \to 0$, we find the limit function

$$H = \frac{F(z) - F(0)}{zF(0)F(z)} \tag{4}$$

to be nonnegative on $(-\infty, 1)$. By the expression of H, it is also analytic since $F(z)$ is never zero on $\mathbb{C} \setminus [1, \infty)$. Finally, since $Im(H_\epsilon(z)) \ge 0$ for $Im(z) > 0$, then $Im(H(z))$, as the limit, is nonnegative. It follows that the sequence corresponding to H is also completely monotone. If c is in ℓ^1, $0 < H(1) = \frac{F(1)-F(0)}{F(0)F(1)} < \frac{1}{c_0}$. If $F(1) = \|c\|_1 = \infty$, we have $0 < H(z) \le \frac{F(z)}{zF(0)F(z)} = \frac{1}{z_0 c_0}$. Fix $z_0 \in (0, 1)$, then for any $z \in (z_0, 1)$, $H(z) \le \frac{1}{z_0 c_0}$. $H(z)$ is increasing in z since the sequence is completely monotone and therefore nonnegative. Letting $z \to 1^-$, by the Monotone convergence theorem, we have $H(1) \le \frac{1}{z_0 c_0}$. Taking $z_0 \to 1$, $H(1) \le \frac{1}{c_0}$.

By the explicit formula of $H(z)$, we see that it is the generating function of $-(c_1^{(-1)}, c_2^{(-1)}, \dots)$ since $1/F(z)$ is the generating function of $c^{(-1)} = (c_0^{(-1)}, c_1^{(-1)}, \dots)$. The second claim therefore follows. □

We then have the following claim:

Corollary 1. *Equation* (2) *can be solved stably. In particular,* $\forall f \in \ell^p$, $\exists a \in \ell^p$ *such that* $a * c = f$ *and*

$$\|a\|_p \le \frac{2}{c_0}\|f\|_p. \tag{5}$$

The claim follows directly from the fact that $\|c^{-1}\|_1 \le 2/c_0$ and Young's inequality.

2.2. **Computing convolution inverse and deconvolution.** The deconvolution actually can be performed directly as the corresponding matrix is lower triangular. Another method is to use the algorithm in [18] to find $c^{(r)}$. Then, the inverse is computed as $a = c^{(-1)} * f$. The algorithm for $c^{(r)}$ reads

- Determine the canonical sequence b that satisfies $(n+1)c_{n+1} = \sum_{k=0}^{n} c_{n-k}b_k$.
- Compute $c^{(r)}$ by $(n+1)c^{(r)}_{n+1} = r\sum_{k=0}^{n} c^{(r)}_{n-k}b_k$.

For a completely monotone sequence, $b_k \ge 0$ ([13]). If $c_0 = 1$, computing the canonical sequence is straightforward

$$b_n = (n+1)c_{n+1} - \sum_{k=0}^{n-1} c_{n-k}b_k. \tag{6}$$

Note that $F_b(z) = F_c'(z)/F_c(z)$.

If $c_0 = 1$, $c_0^{(-1)} = 1$ and $|c^{(-1)}_{n+1}| \le \frac{1}{n+1}\sum_{k=0}^{n} |c^{(-1)}_{n-1}|b_k$. It's clear by induction that $|c^{(-1)}_{n+1}| \le c_{n+1}$. For general c_0, we can apply the above argument to c/c_0 and have the bound

$$|c_k^{(-1)}| \le \frac{1}{c_0^2}|c_k|. \tag{7}$$

This is a pointwise bound for the convolution inverse.

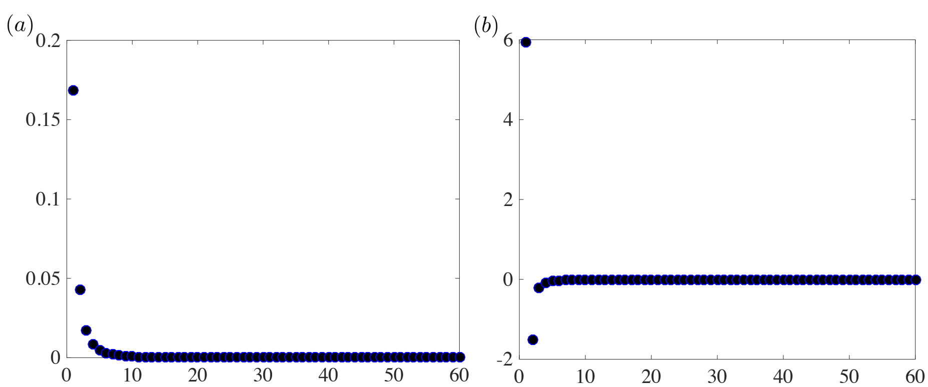

FIGURE 1. A completely monotone sequence and its convolution inverse

Every completely monotone sequence is the moment sequence of a Hausdorff measure. Fix M as a big integer and denote $h = 1/M$. $x_i = (i - 1/2)h$. Consider the discrete measures

$$\mathcal{C}_M = \Big\{\mu : \mu = h\sum_{i=1}^{M} \lambda_i \delta(x - x_i), \lambda_i \ge 0\Big\}. \tag{8}$$

The weak star closure ($\langle \mu, f\rangle = \int_{[0,1]} f d\mu$ where $f \in C[0,1]$) of $\cup_{M\ge 1}\mathcal{C}_M$ is the set of all Hausdorff measures. Hence, we can generate completely monotone sequences using

$$d_n = \sum_{i=1}^{M} h\lambda_i x_i^n,\ \ n = 0, 1, 2, \dots, \tag{9}$$

where λ_i's are generated randomly. In Fig. 1, we plot a completely monotone sequence and its convolution inverse obtained using this method. In Fig. 2 (a), we have a sequence which is of square shape; in Fig. 2 (b), we plot the convolution between the sequence in (a) and the completely monotone sequence obtained in Fig. 1. Fig. 2 (c) shows the solution $a * c = f$ by convolving the sequence in Fig. 2(b) with $c^{(-1)}$. The original sequence is recovered accurately.

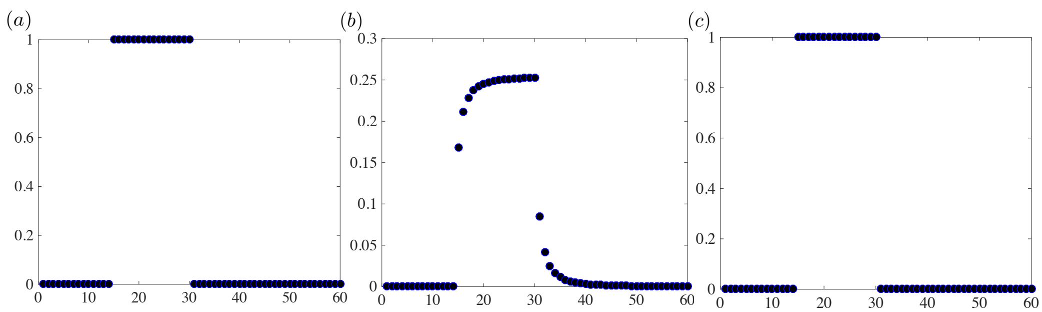

FIGURE 2. A simple example of deconvolution

2.3. **Deconvolution for a general kernel.** Consider that the sequence c is no longer completely monotone. The direct inverting is computationally inexpensive since the matrix is lower triangular. However, if $F_c(z)$ has a zero point near the origin, the generating function of $c^{(-1)}$ has a small radius of convergence. Then, an iterative method may be desired.

Consider approximating the sequence c by a completely monotone sequence $d = \{d_n\}$ of the form in Equation (9). Writing d in matrix form, we have

$$d = \frac{1}{m} A\lambda = A\eta, \tag{10}$$

where $\eta = \frac{1}{m}\lambda$. A simple iterative method then reads:

$$a^{p+1} = f * d^{(-1)} - a^p * [(c - d) * d^{(-1)}]. \tag{11}$$

Clearly, the iteration converges if $\|(c - d) * d^{(-1)}\|_1 < 1$. A sufficient condition is therefore

$$\|d^{(-1)}\|_1 \|c - d\|_1 \le \frac{2}{\|\eta\|_1}\|c - A\eta\|_1 < 1, \tag{12}$$

because d is completely monotone and $d_0 = \|\eta\|_1$. As long as we can find a solution η to this optimization problem, the iterative method can be applied to solve the convolution equation (2).

3. Time-continuous groups and a new definition of fractional calculus

The fractional calculus in continuous time has been used widely in physics and engineering for memory effect, viscoelasticity, porous media etc [12, 6, 16, 19, 5, 1, 25]. Given a function $f(t)$, the fractional integral with order $\gamma > 0$ at $t > 0$ is given by Abel's formula

$$J_\gamma f(t) = \frac{1}{\Gamma(\gamma)} \int_0^t f(s)(t-s)^{\gamma-1} ds. \tag{13}$$

For the derivatives, there are two types that are commonly used: the Riemann-Liouville definition and the Caputo definition (See [16]). Let $n-1 < \gamma < n$, the Riemann-Liouville and Caputo fractional derivatives at $t > 0$ are given respectively by

$$D_{rl}^\gamma f(t) = \frac{1}{\Gamma(n-\gamma)} \frac{d^n}{dt^n} \int_0^t \frac{f(s)}{(t-s)^{\gamma+1-n}} ds, \tag{14}$$

$$D_c^\gamma f(t) = \frac{1}{\Gamma(n-\gamma)} \int_0^t \frac{f^{(n)}(s)}{(t-s)^{\gamma+1-n}} ds. \tag{15}$$

In [20], an idea using distributions to define fractional derivatives for causal functions was mentioned briefly. Inspired by the idea, we explore a group generated by some completely monotone functions in detail and define a fractional calculus for a particular class of distributions.

According to (15), the Caputo derivatives can be defined only if $f^{(n)}$ exists in some sense and this is unnatural since intuitively it can be defined for functions that are 'γ-th' order smooth only. In [1], Allen, Caffarelli and Vasseur have introduced an alternative form of Caputo derivative to avoid using the $f^{(n)}$ derivative. In Section 5, we also provide an alternative definition. Our definition will not use $f^{(n)}$ either and will cover these definitions if the function has some regularity.

In this section, we first introduce the time-continuous convolution group and then define a fractional calculus using this group. This new fractional calculus has the group property. When acting on causal functions, it agrees with the Riemann-Liouville calculus for $t > 0$. The singularities at $t = 0$ are important for the group property. At last, a group for right derivatives is mentioned briefly.

3.1. **A time-continuous convolution group.**

Consider

$$\mathscr{C}_+ = \Big\{ g_\alpha : g_\alpha = \frac{u(t)t^{\alpha-1}}{\Gamma(\alpha)} \Big\}. \tag{16}$$

Note that g_α is completely monotone for $0 < \alpha \le 1$. This set forms a semi-group of convolution for $\alpha > 0$, where $u(t)$ is the Heaviside step function. This is because

$$\int_0^t s^{\alpha-1}(t-s)^{\beta-1} ds = t^{\alpha+\beta-1} B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} t^{\alpha+\beta-1}.$$

The Abel's formula for fractional integral is given by

$$J_\alpha \varphi(t) = g_\alpha * (u(t)\varphi(t)) = \frac{u(t)}{\Gamma(\alpha)} \int_0^t \varphi(s)(t-s)^{\alpha-1} ds, \ \forall \alpha > 0. \tag{17}$$

This means the Riemann-Liouville integrals can be understood as the convolution between a member in $\mathscr{C}_+$ and a causal function $\phi = u(t)\varphi$ (i.e. $\phi = 0$ for $t < 0$).

As mentioned in the introduction, we aim to find a convolution group generalized by $\mathscr{C}_+$. To do this, we need to generalize the convolution between distributions.

First, let us introduce the following set of distributions

$$\mathscr{E} = \{ v \in \mathscr{D}'(\mathbb{R}) : \exists M_v \in \mathbb{R}, supp(v) \subset [-M_v, +\infty) \}. \tag{18}$$

$\mathscr{D}(\mathbb{R}) = C_c^\infty(\mathbb{R})$ is the set of test functions while $\mathscr{D}'(\mathbb{R})$, the dual of $\mathscr{D}$, is the set of distributions. Clearly, $\mathscr{E}$ is a linear vector space.

In general, the convolution between two distributions that are not compactly supported is not well defined. However, we can define the convolution for distributions in $\mathscr{E}$. We first choose a partition of unit for $\mathbb{R}$, $\{\phi_i\}$ (i.e. $\phi_i \in C_c^\infty$; $0 \le \phi_i \le 1$; On any compact set K, there are only finitely many

ϕ_i's that are nonzero; $\sum_i \phi_i = 1$ for all $x \in \mathbb{R}$). Such a partition exists. As an example, consider $\zeta \in C_c^\infty(-1,2)$ that is nonnegative and $\zeta = 1$ on $[0,1]$. Let $\zeta_i(x) = \zeta(x-i)$. Then, $\sum_i \zeta_i(x) > 0$ for any $x \in \mathbb{R}$, where the sum makes sense because for any x, there are only finitely many terms that are nonzero. Defining $\phi_i = \zeta_i / \sum_i \zeta_i$ yields such a partition.

Definition 1. Given $f, g \in \mathscr{E}$, we define

$$\langle f * g, \varphi \rangle = \sum_i \langle f * (\phi_i g), \varphi \rangle, \quad \forall \varphi \in \mathscr{D} = C_c^\infty, \tag{19}$$

where $f * (\phi g)$ is given by the usual definition between two distributions when one of them is compactly supported [9, Chap. 0].

Lemma 2. *The definition is independent of $\{\phi_i\}$ and agrees with the usual definition of convolution between distributions whenever one of the two distributions is compactly supported. For $f, g \in \mathscr{E}$, $f * g \in \mathscr{E}$, and there exists N_1, such that*

$$f * g = \sum_{i \geq -N_1, j \geq -N_1} (f\phi_i) * (g\phi_j),$$

where the sum makes sense because for any compact set K, there are only finitely many pairs (i,j) such that the support of $(f\phi_i) * (g\phi_j)$ has nonempty intersection with K. Moreover,

$$f * g = g * f, \tag{20}$$

$$f * (g * h) = (f * g) * h. \tag{21}$$

The proof, though tedious, is very straightforward. The key ingredient is that for $g \in \mathscr{E}$, there exists N_1 such that when $i < -N_1$, $\phi_i g = 0$ in the distribution sense. We'll omit the proof here.

Another property is as following and we omit its proof as well:

Lemma 3. *We use D to mean the distributional derivative. Then, letting $f, g \in \mathscr{E}$, we have*

$$(Df) * g = D(f * g) = f * Dg. \tag{22}$$

With the tools, we are now able extend $\mathscr{C}_+$ to a convolution group $\mathscr{C}$, under the convolution in Definition 1.

Lemma 4. *$g_0 = \delta(t)$ is the convolution identity and for $n \in \mathbb{N}$, $g_{-n} = D^n \delta$ is the convolution inverse of g_n.*

Proof. Note that g_0 and g_{-n} are compactly supported. Then, the convolution can be performed in the traditional way. That δ is the identity is obvious. For g_{-n}, noting $g_n = \frac{u(t)t^{n-1}}{(n-1)!}$, we pick $\varphi \in \mathscr{D} = C_c^\infty(\mathbb{R})$ and have

$$\langle D^n \delta * (\frac{1}{(n-1)!} u(t) t^{n-1}), \varphi \rangle = (-1)^n \frac{1}{(n-1)!} \langle u(t) t^{n-1}, D^n \varphi \rangle = \varphi(0).$$

Hence, $g_{-n} = D^n \delta$ is the convolution inverse. □

For $0 < \gamma < 1$, inspired by the fact $\mathcal{L}(g_\gamma) \sim 1/s^\gamma$ where $\mathcal{L}$ means the Laplace transform, we guess $\mathcal{L}(g_{-\gamma}) \sim s^\gamma$. Hence, we guess the convolution inverse is $\sim D(u(t)t^{-\gamma})$, where D is the distributional derivative. Actually, we have

Lemma 5. *Let $0 < \gamma < 1$, the convolution inverse of g_γ is given by*

$$g_{-\gamma} = \frac{1}{\Gamma(1-\gamma)} D\left(u(t)t^{-\gamma}\right). \tag{23}$$

Proof. We pick $\varphi \in \mathscr{D}$ and apply Lemma 3:

$$\langle D(u(t)t^{-\gamma}) * [u(t)t^{\gamma-1}], \varphi\rangle = -\langle u(t)t^{-\gamma} * u(t)t^{\gamma-1}, D\varphi\rangle$$
$$= -\langle B(1-\gamma,\gamma)u(t), D\varphi\rangle = -B(1-\gamma,\gamma)\int_0^\infty D\varphi(t)dt = B(1-\gamma,\gamma)\varphi(0).$$

This computation verifies that the claim is true. □

For $n < \gamma < n+1$, we define $g_{-\gamma} = D^n\delta * g_{n-\gamma}$.
Then, we have defined the class

$$\mathscr{C} = \{g_\alpha : \alpha \in \mathbb{R}\}. \tag{24}$$

Theorem 2. *$\mathscr{C} \subset \mathscr{E}$ and it is a convolution group under the convolution on $\mathscr{E}$ (Definition 1).*

Proof. Using the above facts and Lemma 2, Lemma 3, we find that for any $\gamma > 0$, $g_{-\gamma}$ is the convolution inverse of g_γ. The fact that $\mathscr{C}_+$ forms a semigroup, the commutativity and associativity in Lemma 2 imply that $\mathscr{C}_- = \{g_{-\gamma}\}$ forms a convolution semigroup as well.

The group property can then be verified using the semi-group property and the fact that $g_\gamma * g_{-\gamma} = \delta$. □

By the explicit expressions of the distributions, we have

Lemma 6. *For large t, $g_\alpha \sim \frac{1}{\Gamma(\alpha)}|t|^{\alpha-1}$. If $\alpha \le 0$ and is an integer, $\Gamma(\alpha) = \infty$, the distribution is compactly supported.*

3.2. **Time-continuous fractional calculus.** In this section, we use the group $\mathscr{C}$ to define the fractional calculus and illustrate that they give Riemann-Liouville fractional calculus (Equation (14)) for $t > 0$ when acting on causal functions.

3.2.1. *Fractional calculus for distributions in $\mathscr{E}$.*

Definition 2. For $\phi \in \mathscr{E}$, we define the operator $I_\alpha : \mathscr{E} \to \mathscr{E}$ by

$$I_\alpha\phi = g_\alpha * \phi. \tag{25}$$

The operators I_α give the definition of fractional calculus for distributions in $\mathscr{E}$. By the definition, it is clear that

Lemma 7. *The operators $\{I_\alpha\}$ form a group, and $I_{-n}\phi = D^n\phi$ $(n = 1, 2, 3, \dots)$ where D is the distributional derivative.*

It's clear that for $\phi \in C_c^\infty$ and $\alpha \in \mathbb{Z}$, I_α gives the usual integral (where the integral is from $-\infty$) or derivative. For example,

$$I_1\phi = u(t) * \phi = \int_{-\infty}^t \phi(s)ds,$$
$$I_{-1}\phi = (D\delta) * \phi = \delta * D\phi = \phi'.$$

For $\alpha = -\gamma, 0 < \gamma < 1$ and $\phi \in C_c^\infty$, we have

$$I_{-\gamma}\phi = \frac{1}{\Gamma(1-\gamma)}\frac{d}{dt}\int_{-\infty}^t \frac{1}{(t-s)^\gamma}\phi(s)ds = \frac{1}{\Gamma(1-\gamma)}\int_{-\infty}^t \frac{1}{(t-s)^\gamma}\phi'(s)ds.$$

Remark 1. It's possible to act the group on $\phi \notin \mathscr{E}$ but some properties mentioned may be invalid. For example, $\phi = 1 \notin \mathscr{E}$. $g_1 = u, g_{-1} = D\delta$. Both $(u * D\delta) * 1$ and $u * (D\delta * 1)$ are defined where $u(t)$ is the Heaviside function, but they are not equal. The associativity is not valid.

3.2.2. *Modified Riemann-Liouville calculus.* The above could be regarded as the fractional calculus starting from $t=-\infty$. However, we are more interested in fractional calculus starting from $t=0$.

Consider causal distributions ('zero' for $t<0$):

$$\mathscr{G}_c = \{\phi \in \mathscr{E} : supp\ \phi \subset [0,\infty)\}. \tag{26}$$

The causality is considered because the memory is usually counted from $t=0$ in many applications. We now consider the causal correspondence for a general distribution in $\mathscr{E}$. Let $u_n \in C_c^\infty(-1/n,\infty)$ where $n=1,2,\dots$ be a sequence satisfying (1) $0 \le u_n \le 1$, (2) $u_n(t)=1$ for $t \ge -1/(2n)$. Introduce the space

$$\mathscr{G} = \left\{\varphi \in \mathscr{E} : \exists \phi \in \mathscr{G}_c, u_n\varphi \xrightarrow{w^*} \phi \text{ for any such sequence}\{u_n\}\right\}. \tag{27}$$

For $\varphi \in \mathscr{G}$, the distribution ϕ is denoted as $u(t)\varphi$ where $u(t)$ is the Heaviside step function. Clearly, if $\varphi(t) \in L^1_{loc}(\mathbb{R})$, where the notation $L^1_{loc}(U)$ represents the set of all locally integrable function defined on U, $u(t)\varphi$ can be understood as the usual multiplication.

Lemma 8. *$\mathscr{G}_c \subset \mathscr{G}$. $\forall \varphi \in \mathscr{G}_c$, $u(t)\varphi = \varphi$.*

This claim is trivial as $u_n\varphi=\varphi$. This then motivates the following definition:

Definition 3. The (modified) Riemann-Liouville operators $J_\alpha : \mathscr{G} \to \mathscr{G}_c$ are given by

$$J_\alpha\varphi = I_\alpha(u(t)\varphi(t)) = g_\alpha * (u(t)\varphi(t)), \tag{28}$$

Proposition 1. *$\forall \alpha,\beta \in \mathbb{R}$, $J_\alpha J_\beta \varphi = J_{\alpha+\beta}\varphi$ and $J_0\varphi = u(t)\varphi$. If we make the domain of them to be $\mathscr{G}_c$ (i.e. the set of causal distributions), then they form a group.*

Proof. One can verify that $supp(J_\alpha\varphi) \subset [0,\infty)$. Hence, $u(t)J_\alpha\varphi = J_\alpha\varphi$. The claims follow from the properties of I_α. If $\varphi \in \mathscr{G}_c$, then φ is identified with $u(t)\varphi$. □

We are more interested in the cases where φ is locally integrable. We call them modified Riemann-Liouville because for good enough φ they agree with the traditional Riemann-Liouville operators (Equation (14)) at $t>0$ while there are some extra singularities at $t=0$. Now, let us illustrate this by checking some special cases.

When $\alpha>0$ and φ is a continuous function, we have verified that (28) gives the Abel's formula of fractional integrals (Equation (17)). It would be interesting to look at the formulas for $\alpha<0$ and smooth φ:

- When $-1<\alpha<0$, we have

$$\begin{aligned} J_\alpha\varphi &= \frac{u(t)}{\Gamma(1-\gamma)} D\int_0^t \frac{1}{(t-s)^\gamma}\varphi(s)ds = \frac{1}{\Gamma(1-\gamma)}(u(t)t^{-\gamma}) * (u(t)\varphi' + \delta(t)\varphi(0)) \\ &= \frac{1}{\Gamma(1-\gamma)}\int_0^t \frac{1}{(t-s)^\gamma}\varphi'(s)ds + \varphi(0)\frac{u(t)}{\Gamma(1-\gamma)}t^{-\gamma}. \end{aligned} \tag{29}$$

  where $\gamma=-\alpha$. This is the Riemann-Liouville fractional derivative.
- When $\alpha=-1$, we have

$$J_{-1}\varphi = D\Big(u(t)\varphi(t)\Big) = u(t)\varphi'(t) + \delta(t)\varphi(0). \tag{30}$$

  We can verify easily that $J_{-1}J_1\varphi = J_1J_{-1}\varphi = \varphi$. Traditionally, the Riemann-Liouville derivatives for integer values are defined as the usual derivatives. $J_{-1}\varphi(t)$ agrees with the usual derivative for $t>0$ but it has a singularity due to the jump of $u(t)\varphi$ at $t=0$.

- When $\alpha=-1-\gamma$. By the group property, we have

$$J_\alpha\varphi=J_{-1}(J_{-\gamma}\varphi)=\frac{1}{\Gamma(1-\gamma)}D\Big(u(t)D\int_0^t\frac{1}{(t-s)^\gamma}\varphi(s)ds\Big)$$
$$=\frac{1}{\Gamma(2-|\alpha|)}D\Big(u(t)D\int_0^t\frac{1}{(t-s)^{|\alpha|-1}}\varphi(s)ds\Big).$$

This is again the Riemann-Liouville derivative for $t>0$.

In this sense, we call $\{J_\alpha\}$ the **modified Riemann-Liouville operators**. Clearly, $J_\alpha\varphi$ agrees with the traditional Riemann-Liouville calculus for $t>0$. However, at $t=0$, there is some difference. For example, J_{-1} gives an atom $\varphi(0)\delta(t)$ at the origin so that $J_1J_{-1}=J_{-1}J_1=J_0$. The singularities at $t=0$ are expected since the causal function $u(t)\varphi(t)$ usually has a jump at $t=0$.

The convolution group $\mathscr{C}$ is useful for Caputo derivatives as well. We will generalize the definition of Caputo derivatives (15) in Section 5.

3.3. **Another group for right derivatives.** Now consider another group $\widetilde{\mathscr{C}}$ generated by

$$\tilde{g}_\alpha=\frac{u(-t)}{\Gamma(\alpha)}(-t)^{\alpha-1},\quad \alpha>0. \tag{31}$$

For $0<\gamma<1$, $\tilde{g}_{-\gamma}=-\frac{1}{\Gamma(1-\gamma)}D(u(-t)(-t)^{-\gamma})$ (D means the derivative on t and $Du(-t)=-\delta(t)$). The action of this group is well-defined only if we act it on distributions that have supports on $(-\infty,M]$ or on functions that decay faster than rational functions at ∞.

This group can generate fractional derivatives that are noncausal. For example,

$$\tilde{g}_{-\gamma}*\phi=-\frac{1}{\Gamma(1-\gamma)}\frac{d}{dt}\int_t^\infty(s-t)^{-\gamma}\phi(s)ds. \tag{32}$$

This derivative is called the **right Riemann-Liouville derivative** in some literature (See e.g. [16]). The derivative at t depends on the values in the future and it is therefore noncausal.

This group is actually the dual of $\mathscr{C}$ in the following sense

$$\langle g_\alpha*\phi,\varphi\rangle=\langle\phi,\tilde{g}_\alpha*\varphi\rangle, \tag{33}$$

where both ϕ and φ are compactly supported. (If ϕ and φ are not compactly supported or do not decay at infinity, then at least one group is not well defined for them.) This dual identity actually provides a type of integration by parts.

It is interesting to write explicitly out the case $\alpha=-\gamma$.

$$\int_{-\infty}^\infty g_\alpha*\phi(t)\varphi(t)dt=\int_{-\infty}^\infty\frac{1}{\Gamma(1-\gamma)}\int_{-\infty}^t(t-s)^{-\gamma}D\phi(s)ds\varphi(t)dt$$
$$=-\int_{-\infty}^\infty\frac{1}{\Gamma(1-\gamma)}\frac{D}{Ds}\int_s^\infty(t-s)^{-\gamma}\varphi(t)dt\phi(s)ds=\int_{-\infty}^\infty\tilde{g}_\alpha*\varphi(s)\phi(s)ds.$$

Remark 2. Alternatively, one may define the operator I_α by $\langle I_\alpha\phi,\varphi\rangle=\langle\phi,\tilde{g}_\alpha*\varphi\rangle$ for $\phi\in\mathscr{D}'$ and $\varphi\in\mathscr{D}$ whenever this is well-defined. This definition however is also generally only valid for $\phi\in\mathscr{E}$. This is because $\tilde{g}_\alpha*\varphi$ is supported on $(-\infty,M]$ for some M. If $\phi\notin\mathscr{E}$, the definition does not make sense.

Remark 3. The fractional time derivatives on distributions in $\mathscr{E}$ provides a suitable frame to define fundamental solutions for fractional PDEs.

4. Regularities of the modified Riemann-Liouville operators

By the definition, it is expected that $\{J_\alpha\}$ indeed improve or reduce regularities as the ordinary integrals or derivatives do. In this section, we check this topic by considering their actions on a specific class of Sobolev spaces.

Recall that $H_0^s(0,T)$ is the closure of $C_c^\infty(0,T)$ under the norm of $H^s(0,T)$ ($H^s(0,T)$ itself equals the closure of $C^\infty[0,T]$). We would like to avoid the singularities that may appear at $t=0$ but we don't require much at $t=T$. We therefore introduce the space $\tilde{H}^s(0,T)$ which is the closure of $C_c^\infty(0,T]$ under the norm of $H^s(0,T)$ (If $\varphi \in C_c^\infty(0,T]$, *supp* $\varphi \subset C(0,T]$ and $\varphi \in C^\infty[0,T]$. $\varphi(T)$ may be nonzero.).

We now introduce some lemmas for our further discussion:

Lemma 9. *Let $s\in\mathbb{R}, s\ge 0$.*

- *The restriction mapping is bounded from $H^s(\mathbb{R})$ to $H^s(0,T)$, i.e. $\forall v\in H^s(\mathbb{R})$, then $v\in H^s(0,T)$ and there exists $C=C(s,T)$ such that $\|v\|_{H^s(0,T)}\le\|v\|_{H^s(\mathbb{R})}$.*
- *For $v\in\tilde{H}^s(0,T)$, $\exists v_n\in C_c^\infty(\mathbb{R})$ such that the following conditions hold: (i). supp $v_n\subset(0,2T)$. (ii). $\|v_n\|_{H^s(\mathbb{R})}\le C\|v_n\|_{H^s(0,T)}$, where $C=C(s,T)$. (iii). $v_n\to v$ in $H^s(0,T)$.*

We use $\mathscr{D}'(0,T)$ to mean the dual of $\mathscr{D}(0,T)=C_c^\infty(0,T)$. Clearly, $\mathscr{E}\subset\mathscr{D}'(\mathbb{R})\subset\mathscr{D}'(0,T)$.

Lemma 10. *If $v_n\to f$ in $H^s(0,T)$, ($s\ge 0$), then, $J_\alpha v_n\to J_\alpha f$ in $\mathscr{D}'(0,T)$, $\forall\alpha\in\mathbb{R}$.*

Proof. Since v_n, f are in H^s, then they are locally integrable functions.

Let $\varphi\in\mathscr{D}(0,T)$. Let $\{\phi_i\}$ a partition of unit for $\mathbb{R}$. By the definition of J_α,

$$\langle g_\alpha*(u(t)v_n),\varphi\rangle=\sum_i\langle(g_\alpha\phi_i)*(u(t)v_n),\varphi\rangle=\sum_i\langle u(t)v_n,h_\alpha^i*\varphi\rangle,$$

where $h_\alpha^i(t)=(g_\alpha\phi_i)(-t)$. Note that there are only finitely many terms that are nonzero in the sum since φ is compactly supported and g_α is supported in $[0,\infty)$. Since the support of $g_\alpha\phi_i$ is in $[0,\infty)$, then the support of $h_\alpha^i*\varphi$ is in $(-\infty,T)$. Further, $h_\alpha^i*\varphi\in C_c^\infty(\mathbb{R})$. Hence, in distribution,

$$\begin{aligned}\langle u(t)v_n,h_\alpha^i*\varphi\rangle=\int_0^T v_n(t)(h_\alpha^i*\varphi)(t)dt\to\int_0^T f(t)(h_\alpha^i*\varphi)(t)dt\\ =\langle u(t)f,h_\alpha^i*\varphi\rangle=\langle(g_\alpha\phi_i)*(u(t)f),\varphi\rangle.\end{aligned}$$

This verifies the claim. □

We now consider the action of J_α on $\tilde{H}^s(0,T)$ and we actually have:

Theorem 3. *If $\min\{s,s+\alpha\}\ge 0$, then J_α is bounded from $\tilde{H}^s(0,T)$ to $\tilde{H}^{s+\alpha}(0,T)$. In other words, if $f\in\tilde{H}^s(0,T)$, then $J_\alpha f\in\tilde{H}^{s+\alpha}(0,T)$ and there exists a constant C depending on T, s and α such that*

$$\|J_\alpha f\|_{H^{s+\alpha}(0,T)}\le C\|f\|_{H^s(0,T)},\ \forall f\in\tilde{H}^s(0,T). \tag{34}$$

About this topic, some partial results can be found in [16, 15, 11].

Proof. In the proof here, we use C to mean a generic constant, i.e. C may represent different constants from line to line, but we just use the same notation.

$\alpha=0$ is trivial as we have the identity map.

Consider $\alpha<0$ first. For $\alpha=-n$ ($n=1,\dots$), let $v\in C_c^\infty(0,\infty)$. $J_{-n}v\in C_c^\infty(0,\infty)$ because in this case, the action is the usual n-th order derivative. $\|J_{-n}v\|_{H^{s-n}(\mathbb{R})}\le C\|v\|_{H^s(\mathbb{R})}$ is clear. Taking a sequence $v_i\in C_c^\infty$ and *supp* $v_i\subset(0,2T)$ such that $\|v_i\|_{H^s(\mathbb{R})}\le C\|v_i\|_{H^s(0,T)}$, and $v_i\to f$ in $H^s(0,T)$. It then follows that $\|J_{-n}v_i\|_{H^{s-n}(\mathbb{R})}\le C\|v_i\|_{H^s(0,T)}$. Since the restriction is bounded from $H^{s-n}(\mathbb{R})$ to $H^{s-n}(0,T)$, $J_{-n}v_i$ is a Cauchy sequence in $\tilde{H}^{s-n}(0,T)$. The limit in $\tilde{H}^{s-n}(0,T)$ must be $J_{-n}f$ by Lemma 10. Hence, J_{-n} sends $\tilde{H}^s(0,T)$ to $\tilde{H}^{s-n}(0,T)$.

By the group property, it suffices to consider $-1<\alpha<0$ for fractional derivatives. Let $\gamma=|\alpha|$. We pick first $v\in C_c^\infty(0,2T)$. We have

$$J_{-\gamma}v=\frac{d}{dt}\int_0^t (t-s)^{-\gamma}v(s)ds=\frac{d}{dt}\int_0^t s^{-\gamma}v(t-s)ds=\int_0^t (t-s)^{-\gamma}v'(s)ds.$$

Since $J_{-\gamma}v=(u(t)t^{-\gamma})*(v')$ and $v'\in C_c^\infty(0,2T)$, $J_{-\gamma}v$ is C^∞ and *supp* $J_{-\gamma}v\subset(0,\infty)$. Note that the last term is the Caputo derivative. The Caputo derivative equals the Riemann-Liouville derivative for $v\in C_c^\infty(0,2T)$.

Since $|\mathcal{F}(u(t)t^{-\gamma})|\le C|\xi|^{\gamma-1}$, we find $|\mathcal{F}(J_{-\gamma}v)|\le C|\xi|^{\gamma}|\hat{v}(\xi)|$. Here, $\mathcal{F}$ represents the Fourier transform operator while $\hat{v}$ is the Fourier transform of v. Hence,

$$\int(1+|\xi|^2)^{(s-\gamma)}|\mathcal{F}(J_{-\gamma}v)|^2d\xi\le\int(1+|\xi|^2)^{s}|\hat{v}(\xi)|^2d\xi,$$

or $\|J_{-\gamma}v\|_{H^{s-\gamma}(\mathbb{R})}\le C\|v\|_{H^s(\mathbb{R})}$. By Lemma 9, the restriction is bounded

$$\|J_{-\gamma}v\|_{H^{s-\gamma}(0,T)}\le C\|v\|_{H^s(\mathbb{R})}.$$

We now take $v_i\in C_c^\infty(0,2T)$ such that $v_i\to f$ in $H^s(0,T)$ and $\|v_i\|_{H^s(\mathbb{R})}\le C\|v_i\|_{H^s(0,T)}$, then $\|J_{-\gamma}v_i\|_{H^{s-\gamma}(0,T)}\le C\|v_i\|_{H^s(0,T)}$ and $J_{-\gamma}v_i$ is a Cauchy sequence in $\tilde{H}^s(0,T)\subset H^s(0,T)$. The limit in $\tilde{H}^s(0,T)$ must be $J_{-\gamma}f$ by Lemma 10. Hence, the claim follows for $-1<\alpha<0$.

Consider $\alpha>0$ and $n\le\alpha<n+1$. Note that J_n sends $\tilde{H}^s(0,T)$ to $\tilde{H}^{s+n}(0,T)$ since this is the usual integral. We therefore only have to prove the claim for $0<\alpha<1$ by the group property.

For $0<\alpha<1$, $J_\alpha v=\int_0^t s^{\gamma-1}v(t-s)ds\in C^\infty(0,\infty)$ and *supp* $J_\alpha v\subset(0,\infty)$ for $v\in C_c^\infty(0,2T)$. We again set $\gamma=|\alpha|=\alpha$. The Fourier transform of $J_\gamma v$ is $\hat{v}/(-i\xi)^\gamma$ [16]. There is singularity at $\xi=0$ because $J_\gamma v\sim t^{\gamma-1}$ as $t\to\infty$. Since we care the behavior on $(0,T)$, we can pick a cutoff function $\zeta=\beta(x/T)$ where $\beta=1$ on $[-1,1]$ and zero for $|x|>2$. $\hat{\zeta}$ is a Schwartz function.

Noting $|\mathcal{F}(\zeta J_\gamma v)|\le|\hat{\zeta}*\hat{v}|\xi|^{-\gamma}|\le|\hat{\zeta}*|\xi|^{-\gamma}|\|\hat{v}\|_\infty\le C\|\hat{v}\|_\infty$, we find

$$\|\zeta J_\gamma v\|_{H^{s+\gamma}(\mathbb{R})}^2=\int_{\mathbb{R}}(1+|\xi|^2)^{s+\gamma}|\hat{\zeta}*(\hat{v}|\xi|^{-\gamma})|^2d\xi=\int_{|\xi|<R}+\int_{|\xi|\ge R}\le C\|\hat{v}\|_\infty^2+\int_{|\xi|\ge R}.$$

For $|\xi|\ge R$, we split the convolution $\hat{\zeta}*(\hat{v}|\xi|^{-\gamma})$ into two parts and apply the inequality $(a+b)^2\le 2(a^2+b^2)$. It then follows that

$$\int_{|\xi|\ge R}\le C\int_{|\xi|\ge R}d\xi(1+|\xi|^2)^{s+\gamma}\left(\left(\int_{|\eta|\ge|\xi|/2}|\hat{\zeta}(\xi-\eta)||\hat{v}|(\eta)|\eta|^{-\gamma}d\eta\right)^2\right.$$
$$\left.+\left(\int_{|\eta|\le|\xi|/2}|\hat{\zeta}(\xi-\eta)||\hat{v}|(\eta)|\eta|^{-\gamma}d\eta\right)^2\right)=I_1+I_2.$$

For I_1 by Holder inequality and Fubini theorem,

$$\begin{aligned}I_1&\le C\int_{|\xi|\ge R}d\xi(1+|\xi|^2)^{s+\gamma}\int_{|\eta|\ge|\xi|/2}|\hat{\zeta}(\xi-\eta)||\hat{v}(\eta)|^2|\eta|^{-2\gamma}d\eta\\&\le C\int_{|\eta|\ge R/2}d\eta|\hat{v}(\eta)|^2|\eta|^{-2\gamma}\int_\xi|\hat{\zeta}(\xi-\eta)|(1+|\xi|^2)^{s+\gamma}d\xi\\&\le C\int_{|\eta|\ge R/2}d\eta|\hat{v}(\eta)|^2|\eta|^{-2\gamma}(1+|\eta|^2)^{\gamma+s}\le C\|v\|_{H^s(\mathbb{R})}^2.\end{aligned}$$

Here, C depends on R and ζ.

For I_2 part, we note that $|\hat{\zeta}(\xi-\eta)| \le C|\xi|^{-N}$ if R is large enough, since $\hat{\zeta}$ is a Schwartz function.

$$\int_{|\eta|\le|\xi|/2} |\hat{\zeta}(\xi-\eta)||\hat{v}|(\eta)|\eta|^{-\gamma}d\eta \le C\|\hat{v}\|_\infty |\xi|^{-N}\int_{|\eta|\le|\xi|/2}|\eta|^{-\gamma}d\eta \le C\|\hat{v}\|_\infty|\xi|^{-N+1-\gamma}.$$

Hence, $I_2 \le C\|\hat{v}\|_\infty^2$.

Overall, we have

$$\|\zeta J_\gamma v\|_{H^{s+\gamma}(\mathbb{R})} \le C(\|\hat{v}\|_\infty + \|v\|_{H^s(\mathbb{R})}) \le C\|v\|_{H^s(\mathbb{R})}.$$

Note that v is supported in $(0,2T)$ and $\|\hat{v}\|_\infty = \|v\|_{L^1(0,2T)}$, which is bounded by it's $L^2(0,2T)$ norm and thus $H^s(\mathbb{R})$ norm. The constant C depends on T and ζ. Using again that the restriction map is bounded, we find that

$$\|J_\gamma v\|_{H^{s+\gamma}(0,T)} \le C\|\zeta J_\gamma v\|_{H^{s+\gamma}(\mathbb{R})} \le C\|v\|_{H^s(\mathbb{R})}.$$

The claim is true for $C_c^\infty(0,2T)$. Again, using an approximation sequence $v_i \in C_c^\infty(0,2T)$, $\|v_i\|_{H^s(\mathbb{R})} \le C\|v_i\|_{H^s(0,T)}$ implies that it is true for $\tilde{H}^s(0,T)$ also. □

Enforcing φ to be in $\tilde{H}^s(0,T)$ removes the singularities at $t=0$. This then allows us to obtain the regularity estimates above and the Caputo derivatives will be the same as Riemann-Liouville derivatives. If $v \in \tilde{H}^0(0,T) = L^2(0,T)$, then the value of $J_\gamma v$ at $t=0$ is well-defined for $\gamma > 1/2$, which should be zero (See also [15]), because the Holder inequality implies $\int_0^t (t-s)^{\gamma-1}v(s)ds \le C(v)t^{\gamma-1/2}$. Actually, $\tilde{H}^\gamma(0,T) \subset C^0[0,T]$ if $\gamma > 1/2$.

5. An extension of Caputo derivatives

By observing the calculation like (29) above, the Caputo derivatives $D_c^\gamma\varphi$ ($\gamma > 0$) (Equation (15)) may be defined using $J_{-\gamma}\varphi$ and the terms like $\varphi(0)\frac{1}{\Gamma(1-\gamma)}t^{-\gamma}$, and hence may be generalized to a function φ such that only $\varphi^{(m)}, m \le [\gamma]$ exist in some sense, where $[\gamma]$ means the largest integer that does not exceed γ. We then do not need to require that $\varphi^{([\gamma]+1)}$ exists. In this paper, we only deal with $0<\gamma<1$ cases as they are mostly used in practice. (For general $\gamma > 1$, one has to remove singular terms related to $\varphi(0), \dots, \varphi^{[\gamma]}(0)$, the jumps of the derivatives of $u(t)\varphi$, from $J_{-\gamma}\varphi$.) We prove some basic properties of the extended Caputo derivatives according to our definition, which will be used for the analysis of fractional ODEs (FODEs) in Section 6, and may be possibly used for fractional PDEs (FPDEs).

For our discussion, we first introduce some results from real analysis:

Lemma 11. *Suppose both $f, g \in L^1(\mathbb{R})$, then $F(x,y) = f(x-y)g(y)$ is integrable with respect to y for almost every x. Let $h(x) = \int F(x,y)dy$. Then, $h(x)$ is integrable and $\int |h(x)|dx \le \|f\|_1\|g\|_1$. It follows that $F(x,y) \in L^1(\mathbb{R}^2)$.*

For the first part of claim, refer to [21, 12.4, exercises]. $F \in L^1(\mathbb{R}^2)$ follows from Tonelli's theorem by considering $|F| = |f(x-y)||g(y)|$.

Lemma 12. *Suppose $f, g \in L^1_{loc}[0,\infty)$, then $h(x) = \int_0^x f(x-y)g(y)dy$ is defined for almost every $x \in [0,\infty)$ and $h \in L^1_{loc}[0,\infty)$.*

Proof. Fix $M > 0$. Denote $\Omega = \{(x,y) : 0 \le y \le x \le M\}$. $F(x,y) = |f(x-y)||g(y)|$ is measurable and nonnegative on Ω. Tonelli's theorem ([21, 12.4]) indicates that

$$\iint_D F(x,y)dA = \int_0^M \int_0^x |f(x-y)||g(y)|dydx = \int_0^M |g(y)|\int_y^M |f(x-y)|dxdy \le C(M),$$

for some $C(M) \in (0, \infty)$. This means that $F(x, y)$ is integrable on D. Hence, $h(x)$ is defined for almost every $x \in [0, M]$ and $\int_0^M |h(x)|dx < \infty$. Since M is arbitrary, the claim follows. $\square$

We introduce

$$X = \left\{ \varphi \in L^1_{loc}[0, \infty) : \exists C \in \mathbb{R}, \lim_{t \to 0^+} \frac{1}{t} \int_0^t |\varphi - C| dt = 0 \right\}. \tag{35}$$

Recall that $L^1_{loc}[0, \infty)$ is the set of locally integrable functions on $[0, \infty)$, i.e. the functions are integrable on any compact set $K \subset [0, \infty)$.

Clearly, X is a subspace of $H^1_{loc}[0, \infty)$ and $C^0[0, \infty) \subset X$. It is easy to see that C is unique for every $\varphi \in X$. We denote

$$\varphi(0+) := C. \tag{36}$$

For convenience, we also introduce the following set for $0 < \gamma < 1$:

$$Y_\gamma = \left\{ f \in L^1_{loc}[0, \infty) : \lim_{T \to 0^+} \frac{1}{T} \int_0^T \left| \int_0^t (t-s)^{\gamma - 1} f(s) ds \right| dt = 0 \right\}, \tag{37}$$

and also

$$X_\gamma = C + J_\gamma Y_\gamma = \Big\{ \varphi : \exists C \in \mathbb{R}, f \in Y_\gamma, s.t. \varphi = C + J_\gamma(f) \Big\}. \tag{38}$$

Recall that

$$J_\gamma(f) = g_\gamma * (u(t) f) = \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s) ds,$$

where the integral is in Lebesgue sense, as $f \in L^1_{loc}[0, \infty)$ by Lemma 12.

By the definition of Y_γ, it is almost trivial to conclude that:

Lemma 13. *Y_γ and X_γ are subspaces of $L^1_{loc}[0, \infty)$. If $f \in Y_\gamma$, then $J_\gamma f(0+) = 0$ and $X_\gamma \subset X$.*

Remark 4. If $f \geq 0, a.e.$, $\lim_{T \to 0^+} \frac{1}{T} \int_0^T |\int_0^t (t-s)^{\gamma-1} f(s) ds| dt = 0$ is equivalent to $\lim_{T \to 0^+} \frac{1}{T} \int_0^T (T-s)^\gamma f(s) ds = 0$. Hence, $t^{-\gamma} \notin Y_\gamma$ and $t^{-\gamma + \delta} \in Y_\gamma, \forall \delta > 0$. Whether Y_γ is strictly bigger than the space determined by $\lim_{T \to 0^+} \frac{1}{T} \int_0^T (T-s)^\gamma |f(s)| ds = 0$ or not is an interesting real analysis question.

Now, we introduce our definition of Caputo derivatives:

Definition 4. For $0 < \gamma < 1$, we define the Caputo derivative of order γ, $0 < \gamma < 1$ as $D_c^\gamma : X \to \mathscr{E}$,

$$D_c^\gamma \varphi = J_{-\gamma} \varphi - \varphi(0+) g_{1-\gamma} = J_{-\gamma} \varphi - \varphi(0+) \frac{u(t)}{\Gamma(1-\gamma)} t^{-\gamma}. \tag{39}$$

Recall $J_{-\gamma} \varphi = g_{-\gamma} * (u(t) \varphi(t))$.

Note that we have used explicitly the convolution operator $J_{-\gamma}$ in the definition. The convolution structure here enables us to establish the fundamental theorem (Theorem 4) below using deconvolution so that we can rewrite fractional differential equations using integral equations with completely monotone kernels.

Suppose f is a distribution supported in $[0, \infty)$. We then formally denote $(u(t) t^{\gamma-1}) * f$ by $\int_0^t (t-s)^{\gamma-1} f(s) ds$, though the convolution is generally in distribution sense. Below, we say a distribution f is locally integrable function if we can find a locally integrable function $\tilde{f}$ such that $\langle f, \varphi \rangle = \int \tilde{f} \varphi dt, \forall \varphi \in \mathscr{D}(\mathbb{R})$.

Note that if φ does not have regularities, $D_c^\gamma \varphi$ is generally a distribution in $\mathscr{D}'(\mathbb{R})$. If $\varphi \in H_0^\gamma(0, T)$, $D_c^\gamma \varphi$ a function in $H_0^0(0, T) = L^2(0, T) \subset L^1(0, T)$ as we have seen in Section 4.

Lemma 14. *By the definition, we have the following claims:*

(1) $\forall \varphi \in X$, $D_c^\gamma \varphi = J_{-\gamma}(\varphi - \varphi(0+))$. *For any constant* C, $D_c^\gamma C = 0$.
(2) $D_c^\gamma : X \to \mathscr{E}$ *is a linear operator.*
(3) $\forall \varphi \in X$, $0 < \gamma_1 < 1$ *and* $\gamma_2 > \gamma_1 - 1$, *we have*

$$J_{\gamma_2} D_c^{\gamma_1} \varphi = \begin{cases} D_c^{\gamma_1 - \gamma_2} \varphi, & \gamma_2 < \gamma_1, \\ J_{\gamma_2 - \gamma_1}(\varphi - \varphi(0+)), & \gamma_2 \geq \gamma_1. \end{cases}$$

(4) *Suppose* $0 < \gamma_1 < 1$. *If* $f \in Y_{\gamma_1}$, $D_c^{\gamma_2} J_{\gamma_1} f = J_{\gamma_1 - \gamma_2} f$ *for* $0 < \gamma_2 < 1$.
(5) *If* $D_c^{\gamma_1} \varphi \in X$, *then for* $0 < \gamma_2 < 1$, $0 < \gamma_1 + \gamma_2 < 1$,

$$D_c^{\gamma_2} D_c^{\gamma_1} \varphi = D_c^{\gamma_1 + \gamma_2} \varphi - D_c^{\gamma_1} \varphi(0+) g_{1-\gamma_2}.$$

(6) $J_{\gamma - 1} D_c^\gamma \varphi = J_{-1}\varphi - \varphi(0+)\delta(t)$. *If we define this to be* D_c^1, *then for* $\varphi \in C^1[0,\infty)$, $D_c^1 \varphi = \varphi'$.

Proof. The first follows from $g_{-\gamma} * (u(t)) = g_{-\gamma} * g_1 = g_{1-\gamma}$. The second is obvious. The third claim follows from $J_{\gamma_2} D_c^{\gamma_1} \varphi = J_{\gamma_2}(J_{-\gamma_1}\varphi - \varphi(0+) g_{1-\gamma_1}) = J_{\gamma_2 - \gamma_1}\varphi - \varphi(0+) g_{1-\gamma_1+\gamma_2}$, which holds by the group property. For the fourth, we just note that $J_{\gamma_1} f(0+) = 0$ and use the group property for J_α. The fifth statement follows easily from the third statement. The last claim follows from $J_{\gamma-1}(J_{-\gamma}\varphi - \varphi(0+) g_{1-\gamma}) = J_{-1}\varphi - \varphi(0+) g_0$ and Equation (29). □

Now, we verify that our definition agrees with (15) if φ has some regularity:

Proposition 2. *For* $\varphi \in X$, *if the distributional derivative on* $(0,\infty)$ $D_+\varphi$ *is a locally integrable function, then*

$$D_c^\gamma \varphi = \frac{1}{\Gamma(1-\gamma)} \int_0^t \frac{D_+\varphi(s)}{(t-s)^\gamma} ds, \ 0 < \gamma < 1, \tag{40}$$

where the convolution integral can be understood in the Lebesgue sense. Further, $D_c^\gamma \varphi \in L^1_{loc}[0,\infty)$.

Proof. First of all, we show that in distribution sense

$$D(u(t)\varphi) = \delta(t)\varphi(0+) + u(t)D_+\varphi,$$

where $u(t)D_+\varphi$ is well defined since $D_+\varphi$ is locally integrable.

To show this, we first mollify φ and define $\varphi^\epsilon = \varphi * \eta_\epsilon$ where $\eta_\epsilon = \frac{1}{\epsilon}\eta(\frac{t}{\epsilon})$ and $0 \leq \eta \leq 1$ satisfies: (i). $\eta \in C_c^\infty(\mathbb{R})$ with $supp(\eta) \subset (-M, 0)$ for some $M > 0$. (ii). $\int \eta dt = 1$. φ^ϵ is clearly smooth. Then, we have

$$D(u(t)\varphi^\epsilon) = \delta(t)\varphi^\epsilon(0) + u(t)D(\varphi^\epsilon),$$

in the distribution sense, which can be verified easily. Now, take $\epsilon \to 0$. Let $C = \varphi(0+)$ ($\varphi(0+)$ is just the constant in the definition of X). Then, $|\varphi^\epsilon(0) - C| = |\int_0^M \eta(-x)\varphi(\epsilon x)dx - C| \leq \sup|\eta| \int_0^M |\varphi(\epsilon x) - C| dx = \sup|\eta| \frac{M}{\epsilon M} \int_0^{M\epsilon} |\varphi(y) - C| dy \to 0$. Hence, $\varphi^\epsilon(0) \to C$.

Since $u(t)\varphi^\epsilon \to u(t)\varphi$ in $L^1_{loc}(\mathbb{R})$, $D(u(t)\varphi^\epsilon) \to D(u(t)\varphi)$ in $\mathscr{D}'(\mathbb{R})$. For $u(t)D(\varphi^\epsilon)$, by the convolution property we have $D(\varphi^\epsilon) = (D\varphi)^\epsilon$ in $\mathscr{D}'(\mathbb{R})$. Since $D_+\varphi$ is locally integrable, we can define its values to be zero on $(-\infty, 0]$ and then it becomes a distribution in $\mathscr{D}'(\mathbb{R})$. We still denote it as $D_+\varphi$. If $supp(\eta) \subset (-M, 0)$, then $u(t)(D\varphi)^\epsilon \to u(t)D_+\varphi$ in $\mathscr{D}'(\mathbb{R})$. (Note carefully that if φ is a causal function, $D\varphi$ as a distribution in $\mathscr{D}'(\mathbb{R})$ generally has an atom $C\delta(t)$, but here we are using $D_+\varphi$ in $\mathscr{D}'(0,\infty)$ that does not include the singularity.) This then verifies the distributional identity.

By the definition of $J_{-\gamma}$ (Definition 3) and applying Lemma 3,

$$\begin{aligned} J_{-\gamma}\varphi = \frac{1}{\Gamma(1-\gamma)} D\left(u(t)t^{-\gamma}\right) * (u(t)\varphi) &= \frac{1}{\Gamma(1-\gamma)} \left(u(t)t^{-\gamma}\right) * D(u(t)\varphi) \\ &= \frac{1}{\Gamma(1-\gamma)} \left(u(t)t^{-\gamma}\right) * (\delta(t)\varphi(0+) + u(t)D_+\varphi). \end{aligned}$$

The first term gives $\varphi(0+)\frac{1}{\Gamma(1-\gamma)}u(t)t^{-\gamma}$.

Consider now $(u(t)t^{-\gamma})*(u(t)D_+\varphi)$. It is clear that

$$(u(t)t^{-\gamma})*(u(t)D_+\varphi)=\lim_{T\to\infty}(u(t)t^{-\gamma}\chi(t\le T))*(u(t)D_+\varphi\chi(t\le T)),\quad in\ \mathscr{D}'.$$

With the truncation, each function becomes in $L^1(\mathbb{R})$. By Lemma 11, the convolution $h_T=(u(t)t^{-\gamma}\chi(t\le T))*(u(t)D_+\varphi\chi(t\le T))\in L^1(\mathbb{R})$ and

$$h_T(t)=\int_0^t\frac{1}{(t-s)^\gamma}D_+\varphi(s)ds,\quad 0<t<T,$$

where the integral is in Lebesgue sense. Hence, as $T\to\infty$, we find that $(u(t)t^{-\gamma})*(u(t)D_+\varphi)$ is a measurable function and for almost every t, Equation (40) holds and the integral is a Lebesgue integral. By Lemma 12, $D_c^\gamma\varphi\in L^1_{loc}[0,\infty)$.

Then, by Definition 4, we obtain Equation (40). □

Remark 5. Note that if the distributional derivative is locally integrable, then it is also called the weak derivative in PDE theory. Sometimes φ is differentiable almost everywhere and the conventional derivative is denoted as φ'. Even if φ' may be defined almost everywhere, one should be careful not to use φ' in the integrand. However, if φ is nice enough, $D\varphi$ will be identical to φ'. On one hand, if the weak derivative exists and the function is differentiable almost everywhere, then $D\varphi=\varphi'$ *a.e.* On the other hand, if φ is integrable on $[0,T], T>0$ and differentiable **everywhere** such that φ' is in $L^1(0,T)$, then φ is absolutely continuous on $[0,T]$ and hence the weak derivative exists.

Regarding the Caputo derivatives, we introduce several results that may be applied for fractional ODEs (FODE) and fractional PDEs (FPDE). The first is the fundamental theorem of fractional calculus:

Theorem 4. *Suppose $\varphi\in X$ and denote $f=D_c^\gamma\varphi$ $(0<\gamma<1)$ which is supported in $[0,\infty)$. Then,*

$$u(t)\varphi(t)=\varphi(0+)+J_\gamma(f)=\varphi(0+)+\frac{1}{\Gamma(\gamma)}\int_0^t(t-s)^{\gamma-1}f(s)ds. \tag{41}$$

where we only use the integral to mean $\frac{1}{\Gamma(\gamma)}(u(t)t^{\gamma-1})*f$. If f is locally integrable, the integral can be understood in Lebesgue sense.

Proof. By the definition (Equation (39)), $f(t)+\varphi(0+)\frac{u(t)}{\Gamma(1-\gamma)}t^{-\gamma}=J_{-\gamma}(\varphi(t))$. Then, the convolution group property yields:

$$u(t)\varphi(t)=J_\gamma\left(f(t)+\varphi(0+)\frac{u(t)}{\Gamma(1-\gamma)}t^{-\gamma}\right)=g_\gamma*f+\varphi(0+)\frac{1}{\Gamma(\gamma)\Gamma(1-\gamma)}\int_0^t(t-s)^{\gamma-1}s^{-\gamma}ds.$$

Since $\int_0^t(t-s)^{\gamma-1}s^{-\gamma}ds=B(\gamma,1-\gamma)=\Gamma(\gamma)\Gamma(1-\gamma)$, the second term is just $\varphi(0+)$.

If $f\in L^1_{loc}[0,\infty)$, by the truncation technique in the proof of Proposition 2, the convolution can be written in Lebesgue integral. □

This theorem is fundamental for fractional differential equations because this allows us to transform the fractional differential equations to integral equations with completely monotone kernels (which are nonnegative). Then, we are able to establish the general Gronwall inequalities or the comparison principle (Theorem 6), which in turn opens the door of a priori estimate of fractional PDEs.

Using Theorem 4, we conclude that

Corollary 2. *Suppose $\varphi(t) \in X$, $\varphi \ge 0$ and $\varphi(0+) = 0$. If the Caputo derivative $D_c^\gamma \varphi$ $(0 < \gamma < 1)$ is locally integrable, and $D_c^\gamma \varphi \le 0$, then $\varphi(t) = 0$. (The local integrability assumption can be dropped if we understand the inequality in the distribution sense which we will introduce in Section 6.)*

Now, we consider functions whose Caputo derivatives are $L^1_{loc}[0,\infty)$. Actually, we have

Proposition 3. *Let $f \in L^1_{loc}[0,\infty)$. Then, $D_c^\gamma \varphi = f$ has solutions $\varphi \in X$ if and only if $f \in Y_\gamma$. If $f \in Y_\gamma$, the solutions are in X_γ and they can be written as*

$$\varphi = C + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s)ds, \quad \forall C \in \mathbb{R}.$$

Further, $\forall \varphi \in X_\gamma$, $D_c^\gamma \varphi \in Y_\gamma$.

Proof. Suppose that $D_c^\gamma \varphi = f$ has a solution $\varphi \in X$. Since $f \in L^1_{loc}[0,\infty)$, by Theorem 4, we have

$$\varphi(t) = \varphi(0+) + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f_1(s)ds,$$

and the integral is in Lebesgue sense. Since $\varphi \in X$, we have $\lim_{T\to 0} \frac{1}{T}\int_0^T |\varphi(s) - \varphi(0+)|ds = 0$. It follows that

$$\frac{1}{T}\int_0^T \left| \int_0^t (t-s)^{\gamma-1} f_1(s)ds \right| dt \to 0, \quad T \to 0.$$

Hence, $f_1 \in Y_\gamma$. This implies that there are no solutions for $D_c^\gamma \varphi = f$ if $f \in L^1_{loc} \setminus Y_\gamma$. (For example, $D_c^\gamma \varphi = t^{-\gamma}$ has no solutions in X.)

For the other direction, now assume $f \in Y_\gamma$. We first note $D_c^\gamma \varphi = 0$ implies that $\varphi = C$ by Theorem 4. One then can check that $J_\gamma f$, which is in X_γ by definition, is a solution to the equation $D_c^\gamma \varphi = f$. Hence any solution can be written as $J_\gamma f + C$. The other direction and the second claim are shown.

We now show the last claim. If $\varphi \in X_\gamma$, by definition (Equation (38)), $\exists f \in Y_\gamma, C \in \mathbb{R}$ such that

$$\varphi = C + \frac{1}{\Gamma(\gamma)} \int_0^t (t-s)^{\gamma-1} f(s)ds.$$

Since $f \in Y_\gamma$, $J_\gamma f(0+) = 0$ by Lemma 13. This means $C = \varphi(0+)$. Now, apply $J_{-\gamma}$ on both sides. Note that $J_{-\gamma} C = C g_{-\gamma} * g_1 = C g_{1-\gamma} = C \frac{u(t)}{\Gamma(1-\gamma)} t^{-\gamma}$. By the group property, we find that

$$D_c^\gamma \varphi = J_{-\gamma} J_\gamma f = f.$$

∎

Remark 6. For the purpose of applications in FPDE analysis, we note that a solution to the equation $D_c^\gamma \varphi = f$ where $f \in L^1_{loc}$ is understood in the distribution sense or in weak sense: Find $\varphi \in X_\gamma$, such that $\forall \zeta \in C_c^\infty$:

$$-\frac{1}{\Gamma(1-\gamma)} \left\langle \int_0^t (t-s)^{-\gamma} \varphi(s)ds, D\zeta \right\rangle - \frac{\varphi(0+)}{\Gamma(1-\gamma)} \int_0^t (t-s)^{-\gamma} \zeta(s)ds = \int_0^t f(t-s)\zeta(s)ds.$$

Motivated by the discussion in Section 4, we say $f \in \tilde{H}^s_{loc}(0,\infty)$ if it is in $\tilde{H}^s(0,T)$ for any $T > 0$. We have another claim about the equation $D_c^\gamma \varphi = f$ where the solutions are in $C^0[0,\infty)$:

Proposition 4. *Suppose $f \in L^1_{loc}[0,\infty) \cap \tilde{H}^s_{loc}(0,\infty)$. If s satisfies (i). $s \ge 0$ when $\gamma > 1/2$ or (ii). $s > \frac{1}{2} - \gamma$ when $\gamma \le 1/2$, then $\exists \varphi \in C^0[0,\infty)$ such that $D_c^\gamma \varphi = f$.*

Let us focus on the mollifying effect on the Caputo derivatives. Let $\eta \in C_c^\infty$, $0 \le \eta \le 1$ and $\int \eta dt = 1$. We define $\eta_\epsilon = \frac{1}{\epsilon}\eta(\frac{t}{\epsilon})$. For a distribution $\varphi \in \mathscr{D}'(\mathbb{R})$, it is well known that

$$\varphi^\epsilon = \varphi * \eta_\epsilon \in C^\infty(\mathbb{R}) \tag{42}$$

and that $supp(\varphi^\epsilon) \subset supp(\varphi) + supp(\eta_\epsilon)$ where $A + B = \{x + y : x \in A, y \in B\}$.

Proposition 5. *If $supp(\eta) \subset (-\infty, 0)$, then $\forall \varphi \in X$, $D_c^\gamma(\varphi^\epsilon) \to D_c^\gamma \varphi$ in $\mathscr{D}'(\mathbb{R})$ as $\epsilon \to 0^+$. Also,*

$$D_c^\gamma \varphi^\epsilon = \int_0^t \frac{D\varphi^\epsilon}{(t-s)^\gamma} ds = \frac{\varphi^\epsilon(t) - \varphi^\epsilon(0)}{t^\gamma} + \gamma \int_0^t \frac{\varphi^\epsilon(t) - \varphi^\epsilon(s)}{(t-s)^{\gamma+1}} ds. \tag{43}$$

If $E(\cdot)$ is a continuous convex function, then

$$D_c^\gamma E(\varphi^\epsilon) \le E'(\varphi^\epsilon) D_c^\gamma \varphi^\epsilon. \tag{44}$$

Proof. For any $\varphi \in X$, it is clear that $u(t)\varphi^\epsilon \to u(t)\varphi$ in $L^1_{loc}[0, \infty)$ and hence in distribution. Using the definition of J_α and the definition of convolution on $\mathscr{E}$, one can readily check $J_{-\gamma}\varphi^\epsilon \to J_{-\gamma}\varphi$ in $\mathscr{D}'(\mathbb{R})$.

For $\varphi^\epsilon(0) \to \varphi(0+)$, we need $supp(\eta) \subset (-\infty, 0)$. There then exists $M > 0$ such that $\eta(t) = 0$ if $t < -M$. Let $C = \varphi(0+)$. Then, $|\varphi^\epsilon(0) - C| = |\int_0^M \eta(-x)\varphi(\epsilon x)dx - C| \le \sup|\eta| \int_0^M |\varphi(\epsilon x) - C|dx = \sup|\eta| \frac{M}{\epsilon M} \int_0^{M\epsilon} |\varphi(y) - C|dy \to 0$. Hence, $\varphi^\epsilon(0) \to C$. This then shows that the first claim is true.

For the alternative expressions of $D_c^\gamma \varphi^\epsilon$, we have used Proposition 2 and integration by parts. These are valid since $\varphi^\epsilon \in C^\infty$.

Multiplying $E'(\varphi^\epsilon(t))$ on $\frac{\varphi^\epsilon(t) - \varphi^\epsilon(0)}{t^\gamma} + \gamma \int_0^t \frac{\varphi^\epsilon(t) - \varphi^\epsilon(s)}{(t-s)^{\gamma+1}} ds$ and using the inequality

$$E'(a)(a - b) \ge E(a) - E(b)$$

since $E(\cdot)$ is convex, we find

$$E'(\varphi^\epsilon) D_c^\gamma \varphi^\epsilon \ge \frac{E(\varphi^\epsilon(t)) - E(\varphi^\epsilon(0))}{t^\gamma} + \gamma \int_0^t \frac{E(\varphi^\epsilon(t)) - E(\varphi^\epsilon(s))}{(t-s)^{\gamma+1}} ds.$$

Since E is convex and φ^ϵ is smooth, the second integral converges for almost every t. Further, $D_+E(\varphi^\epsilon)$ exists and $D_+E(\varphi^\epsilon) = E'(\varphi^\epsilon)D\varphi^\epsilon$. Then, the right hand side must be $\int_0^t \frac{D_+E(\varphi^\epsilon)}{(t-s)^\gamma}$. By Proposition 2, we end the last claim. □

Remark 7. It is interesting to note that we have to choose η such that $supp(\eta) \subset (-\infty, 0)$ to mollify. Using other mollifiers, we may not get the correct limit. This reflects that Caputo derivatives only model the dynamics of memory from $t = 0^+$ and the singularities at $t = 0$ for Riemann-Liouville derivatives are removed totally. It is exactly this nature that makes Caputo derivatives to have many similarities with the ordinary derivative and suitable for initial value problems.

Some obvious conclusions we can make:

Corollary 3. *With the assumptions in Proposition 5, $\forall \varphi \in X$:*

- *If there exists a sequence ϵ_k, such that $D_c^\gamma \varphi^{\epsilon_k}$ converges in L^1_{loc}. Then, the limit is $D_c^\gamma \varphi$ and $D_c^\gamma \varphi \in L^1_{loc}[0, \infty)$.*
- *If φ is $\gamma + \delta$ ($\delta > 0$) Holder continuous (see* [7, Chap. 5]*), then $D_c^\gamma \varphi^\epsilon \to D_c^\gamma \varphi$ uniformly on $[0, T]$ and $D_c^\gamma \varphi = \frac{\varphi(t) - \varphi(0)}{t^\gamma} + \gamma \int_0^t \frac{\varphi(t) - \varphi(s)}{(t-s)^{\gamma+1}} ds$.*

The alternative expressions for $D_c^\gamma \varphi^\epsilon$ are more useful for FPDEs. Unfortunately, for $\varphi \in X$, these forms are generally not valid. For example, φ must have some regularity for the integral form, which is used in [1], to make sense in the Lebesgue sense. It is possible to show that if φ has better regularity, then $D_c^\gamma \varphi^\epsilon$ converges to $D_c^\gamma \varphi$ in better spaces, but we are not going to do this here.

Lastly, we consider the Laplace transform. If $\varphi \in X$, $D_c^\gamma \varphi$ is only a distribution. Recalling that its support is in $[0,\infty)$, we then define the Laplace transform of $D_c^\gamma\varphi$ as

$$\mathcal{L}(D_c^\gamma\varphi) = \lim_{M\to\infty}\langle D_c^\gamma\varphi, \zeta_M e^{-st}\rangle, \tag{45}$$

where $\zeta_M(t) = \zeta_0(t/M)$. $\zeta_0 \in C_c^\infty, 0 \le \zeta_0 \le 1$ satisfies: (i) *supp* $\zeta_0 \subset [-2,2]$ (ii) $\zeta_0 = 1$ for $t \in [-1,1]$. This definition clearly agrees with the usual definition of Laplace transform if the usual Laplace transform of function φ exists.

To be convenient in discussion below, we will denote the following set

$$\mathcal{E}(\mathcal{L}) = \left\{\varphi \in X : \exists L > 0, s.t. \lim_{T\to\infty}\|e^{-Lt}\varphi\|_{L^\infty[T,\infty)} = 0\right\}. \tag{46}$$

Proposition 6. *If $\varphi \in \mathcal{E}(\mathcal{L})$, then $\mathcal{L}(D_c^\gamma\varphi)$ is defined for $Re(s) > L$ and is given by*

$$\mathcal{L}(D_c^\gamma\varphi) = \mathcal{L}(\varphi)s^\gamma - \varphi(0+)s^{\gamma-1}. \tag{47}$$

Proof. $\zeta_M e^{-st} \in C_c^\infty$. Then, it follows that

$$\begin{aligned}\langle D_c^\gamma\varphi, \zeta_M e^{-st}\rangle &= \left\langle g_{-\gamma} * (u(t)\varphi) - \frac{\varphi(0+)u(t)t^{-\gamma}}{\Gamma(1-\gamma)}, \zeta_M e^{-st}\right\rangle \\ &= -\frac{1}{\Gamma(1-\gamma)}\left\langle (u(t)t^{-\gamma}) * (u(t)\varphi), \zeta_M' e^{-st} - s\zeta_M e^{-st}\right\rangle - \frac{\varphi(0)}{\Gamma(1-\gamma)}\int_0^\infty t^{-\gamma}\zeta_M e^{-st}dt.\end{aligned}$$

Note that *supp* $\zeta_M' \cap [0,\infty) \subset [M, 2M]$:

$$\left|\int_M^{2M}\int_0^t (t-\tau)^{-\gamma}\varphi(\tau)d\tau\zeta_M' e^{-st}dt\right| \le \frac{\sup|\zeta_0'|}{M}\int_M^{2M}\int_0^t (t-\tau)^{-\gamma}e^{-L\tau}|\varphi(\tau)|d\tau e^{-(Re(s)-L)t}dt.$$

By the assumption, there exists T_0 such that $|e^{-L\tau}\varphi(\tau)| < 1, a.e.$ if $\tau > T_0$. Hence, if $M > 2T_0$, the inner integral is controlled by $T_0^{-\gamma}\int_0^{T_0}|\varphi(\tau)|d\tau + \int_{T_0}^t (t-\tau)^{-\gamma}d\tau \le C(1+t^{1-\gamma})$. Since $\lim_{M\to\infty}\int_M^{2M}(1+t^{1-\gamma})e^{-\epsilon t}dt = 0$ for any $\epsilon > 0$, we find that the term associated with ζ_M' tends to zero as $M \to \infty$.

For the second term,

$$\begin{aligned}\frac{s}{\Gamma(1-\gamma)}\left\langle (u(t)t^{-\gamma}) * (u(t)\varphi), \zeta_M e^{-st}\right\rangle &= \frac{s}{\Gamma(1-\gamma)}\int_0^\infty\int_0^t (t-\tau)^{-\gamma}\varphi(\tau)d\tau\zeta_M(t)e^{-st}dt \\ &= \frac{s}{\Gamma(1-\gamma)}\int_0^\infty \varphi(\tau)e^{-s\tau}\int_\tau^\infty (t-\tau)^{-\gamma}\zeta_M(t)e^{-(t-\tau)s}dt.\end{aligned}$$

As $M \to \infty$, one finds that

$$\int_0^\infty t^{-\gamma}\zeta_M(t+\tau)e^{-ts}dt \to \Gamma(1-\gamma)s^{\gamma-1},$$

for every $\tau > 0$. Since $Re(s) > L$, the dominate convergence theorem implies that the first term goes to $\mathcal{L}(\varphi)s^\gamma$.

Similarly, the last term converges to $-\varphi(0+)s^{\gamma-1}$. □

To conclude, there is no group property for Caputo derivatives. However, the Caputo derivatives remove the singularities at $t = 0$ compared with the Riemann-Liouville derivatives and have many properties that are similar to the ordinary derivative so that they are suitable for initial value problems.

6. Fractional ordinary differential equations

Some analysis about fractional ODEs could be found in [6, 5]. In this section, we prove some results about fractional ODEs using the Caputo derivatives, whose new definition and properties have been discussed in Section 5. The assumptions here are sufficiently weak and conclusions are general.

We start with a simple linear FODE:

Proposition 7. *Let $0<\gamma<1$, $\lambda\neq 0$, and suppose $b(t)$ is continuous such that there exists $L>0$, $\limsup_{t\to\infty} e^{-Lt}|b(t)|=0$. Then, there is a unique solution of the equation*

$$D_c^\gamma v=\lambda v+b(t),\quad v(0)=v_0$$

in $\mathcal{E}(\mathcal{L})\subset X$ and is given by

$$v(t)=v_0 e_{\gamma,\lambda}(t)+\frac{1}{\lambda}\int_0^t b(t-s)e'_{\gamma,\lambda}(s)ds, \tag{48}$$

where $e_{\gamma,\lambda}(t)=E_\gamma(\lambda t^\gamma)$ and

$$E_\gamma(z)=\sum_{n=0}^\infty \frac{z^n}{\Gamma(\gamma n+1)} \tag{49}$$

is the Mittag-Leffler function [14].

Proof. For $\gamma\in(0,1)$, we have by Proposition 6

$$\mathcal{L}(D_c^\gamma v)=s^\gamma V(s)-v_0 s^{\gamma-1},$$

and $V(s)=\mathcal{L}(v)$.

By the assumption on $b(t)$, the Laplace transform $B(s)=\mathcal{L}(b)$ exists. Hence, any continuous solution of the initial value problem that is in $\mathcal{E}(\mathcal{L})$ must satisfy

$$V(s)=v_0\frac{s^{\gamma-1}}{s^\gamma-\lambda}+\frac{B(s)}{s^\gamma-\lambda}.$$

Using the equality ([19, Appendix])

$$\int_0^\infty e^{-st}E_\gamma(s^\gamma z t^\gamma)dt=\frac{1}{s(1-z)}, \tag{50}$$

and denoting $e_{\gamma,\lambda}(t)=E_\gamma(\lambda t^\gamma)$, we have that

$$\mathcal{L}(e_{\gamma,\lambda})=\frac{s^{\gamma-1}}{s^\gamma-\lambda},\quad \mathcal{L}(e'_{\gamma,\lambda})=\frac{\lambda}{s^\gamma-\lambda}.$$

Taking the inverse Laplace transform of $V(\cdot)$, we get (48). Note that though $e'_{\gamma,\lambda}$ blows up at $t=0$, it is integrable near $t=0$ and the convolution is well-defined. Further, by the asymptotic behavior of b and the decaying rate of $e'_{\gamma,\lambda}$, the solution is again in $\mathcal{E}(\mathcal{L})$. The existence part is proved.

Since the Laplace transform of functions that are in $\mathcal{E}(\mathcal{L})$ is unique, the uniqueness part is proved. ☐

Remark 8. For the existence of solutions in X, the condition $\limsup_{t\to\infty} e^{-Lt}|b(t)|=0$ can be removed, since for any $t>0$, we can redefine b beyond t so that $\limsup_{t\to\infty} e^{-Lt}|b(t)|=0$. The value of $v(t)$ keeps unchanged by the re-definition according to Formula (48). Hence, (48) gives a solution for any continuous function b in X. The uniqueness in X instead of in $\mathcal{E}(\mathcal{L})$ will be established in Theorem 5 below.

We now consider a general FODE:

Theorem 5. *Suppose $f(t,v)$ is continuous and for any $T>0$, there exists $L_T>0$ such that*

$$\sup_{0\le t\le T}|f(t,v_1)-f(t,v_2)|\le L_T|v_1-v_2|.$$

Let $0<\gamma<1$. Then, there is a unique solution $v(\cdot)$ of the equation

$$D_c^{\gamma}v=f(t,v),\quad v(0)=v_0, \tag{51}$$

in X, where $v(0)$ is understood as Equation (36). Moreover, $v(\cdot)\in C^0[0,\infty)$. Further, the solution is continuous with respect to the initial value, indeed,

$$\lim_{\delta\to 0}|v^{\delta}(t)-v(t)|=0,\quad \forall t>0, \tag{52}$$

where $v^{\delta}(\cdot)$ is the solution with $v^{\delta}(0)=v_0+\delta$.

Proof. The proof is just like the proof of existence and uniqueness theorem for ODEs using Picard iteration.

We first show uniqueness. Suppose both $v_1,v_2\in X$ are solutions. Let $w=v_1-v_2$. Then, by the linearity of D_c^{γ}, we have in distribution sense that

$$D_c^{\gamma}w=f(t,v_1)-f(t,v_2).$$

Since f is Lipschitz continuous and both $v_1,v_2\in L^1_{loc}[0,\infty)$, $f(t,v_1)-f(t,v_2)\in L^1_{loc}[0,\infty)$. By Theorem 4, we have in Lebesgue sense that

$$u(t)w=\frac{1}{\Gamma(\gamma)}\int_0^t(t-s)^{\gamma-1}(f(s,v_1(s)-f(s,v_2(s)))ds.$$

Fix $T>0$. For all $t\le T$,

$$u(t)|w(t)|\le\frac{L_T}{\Gamma(\gamma)}\int_0^t(t-s)^{\gamma-1}|w(s)|ds=L_Tg_{\gamma}*(u(t)|w|).$$

Since $g_{\alpha}\ge 0$ when $\alpha>0$ and $u(t)g_{\alpha}=g_{\alpha}$, we convolve both sides with $g_{n_0\gamma}$ and have

$$u(t)w_1(t)\le L_Tg_{\gamma}*(u(t)w_1),$$

where $u(t)w_1=g_{n_0\gamma}*(u(t)|w|)\ge 0$. Since $g_{n_0\gamma}=C_1t^{n_0\gamma-1}$, if n_0 is large enough, w_1 is continuous on $[0,T]$.

Then, by iteration and group property, we have

$$u(t)w_1\le L_T^ng_{n\gamma}*(u(t)w_1)\le\frac{L_T^n}{\Gamma(n\gamma-1)}\sup_{0\le t\le T}|w_1|\int_0^t(t-s)^{n\gamma-1}ds.$$

Since $\Gamma(n\gamma-1)$ grows exponentially, this tends to zero. Hence, $w_1=0$ on $[0,T]$. Because T is arbitrary, $w_1=0$ for all $t\ge 0$. Then, convolve both sides with $g_{-n_0\gamma}$ on $w_1=g_{n_0\gamma}*(u(t)|w|)=0$ and we find $|w|=0$ in distribution. Hence, $v_1=v_2$.

Now, we show that there is a continuous solution. Consider the sequence constructed by

$$v^n=v_0+g_{\gamma}*(u(t)f(t,v^{n-1})),\ v^0=v_0. \tag{53}$$

Consider $E^n=|v^n-v^{n-1}|$. We find

$$E^n\le L_Tg_{\gamma}*(u(t)E^{n-1}),$$
$$E^1=|g_{\gamma}*(u(t)f(t,v^0))|\in C^0[0,T].$$

We use the bound $E^1\le M_T$. Note that $u(t)=g_1$ and by group property, we find

$$E^n\le M_TL_T^{n-1}g_{1+(n-1)\gamma},n=1,2,\dots$$

Since

$$\sum_{n=1}^{\infty} L_T^{n-1} g_{1+(n-1)\gamma} = \sum_n \frac{L_T^{n-1} t^{(n-1)\gamma}}{\Gamma((n-1)\gamma)} < C(T),$$

as it is just a Mittag-Leffler function up to some constants. Therefore, $\sum_n |v^n - v^{n-1}|$ converges uniformly on $[0,T]$. This shows that $v^n \to v$ uniformly on $[0,T]$. v is then continuous. Hence,

$$v(t) = v_0 + g_\gamma * (u(t)f(t,v)), \quad v(0) = v_0,$$

where v is continuous. This means that $v(\cdot)$ is a solution.

For continuity on initial value, we let $u = v - v_0$. Since the Caputo derivative of a constant is zero, the equation is reduced to

$$D_c^\gamma u = f(t, u + v_0), \quad u(0) = 0.$$

For this question, once again, construct the sequence u^{n+1} like in Equation (53) and show u^{n+1} is continuous on v_0. Performing similar argument, $u^{n+1} \to u$ uniformly on $[0,T]$. Then, u is continuous on v_0. □

Before further discussion, we introduce the concept of negative distributions.

Definition 5. We say $f \in \mathscr{D}'(\mathbb{R})$ is a negative distribution if $\forall \varphi \in \mathscr{D}(\mathbb{R}) = C_c^\infty(\mathbb{R})$ with $\varphi \ge 0$, we have

$$\langle f, \varphi \rangle \le 0. \tag{54}$$

We say $f_1 \le f_2$ for $f_1, f_2 \in \mathscr{D}'(\mathbb{R})$ if $f_1 - f_2$ is negative. We say $f_1 \ge f_2$ if $f_2 - f_1$ is negative.

The following lemma is well-known and we omit the proof:

Lemma 15. *If $f \in L^1_{loc}[0,\infty) \subset \mathscr{E} \subset \mathscr{D}'(\mathbb{R})$ is a negative distribution. Then, $f \le 0$ almost everywhere with respect to Lebesgue measure.*

Lemma 16. *Suppose $f_1, f_2 \in \mathscr{E}$, supp$(f_1) \subset [0,\infty)$ and supp$(f_2) \subset [0,\infty)$. If $f_1 \le f_2$ (we mean $f_1 - f_2$ is a negative distribution), and that both $h_1 = (u(t)t^{\gamma-1}) * f_1$ and $h_2 = (u(t)t^{\gamma-1}) * f_2$ are functions in $L^1_{loc}[0,\infty)$, then*

$$h_1 \le h_2, a.e.$$

Proof. Suppose the conclusion is not true. Then by Lemma 15, there exists $\varphi \in C_c^\infty(\mathbb{R})$, $\varphi \ge 0$ such that

$$\int_0^\infty (h_1 - h_2)\varphi dx > 0.$$

This means

$$\langle (u(t)t^{\gamma-1}) * f_1, \varphi \rangle > \langle (u(t)t^{\gamma-1}) * f_2, \varphi \rangle.$$

Let $\{\phi_i\}$ be a partition of unit for $\mathbb{R}$. Then, by Definition 1, we have

$$\sum_i \Big(\langle [\phi_i(u(t)t^{\gamma-1})] * f_1, \varphi \rangle - \langle [\phi_i(u(t)t^{\gamma-1})] * f_2, \varphi \rangle \Big) > 0.$$

There are only finitely many terms that are nonzero in this sum. Hence, there must exist i_0 such that

$$\Big\langle [\phi_{i_0}(u(t)t^{\gamma-1})] * f_1, \varphi \Big\rangle - \Big\langle [\phi_{i_0}(u(t)t^{\gamma-1})] * f_2, \varphi \Big\rangle > 0.$$

Denote $\zeta_{i_0} = \phi_{i_0}(u(t)t^{\gamma-1})$ and $\tilde{\zeta}_{i_0} = \zeta_{i_0}(-t)$. Then,

$$\langle f_1 - f_2, \tilde{\zeta}_{i_0} * \varphi \rangle > 0.$$

$\tilde{\zeta}_{i_0}$ is a positive integrable function with compact support and $\varphi \ge 0$ is compactly supported smooth function. Then, $\tilde{\zeta}_{i_0} * \varphi \ge 0$ and is $C_c^\infty(\mathbb{R})$. This is a contradiction since we have assumed $f_1 \le f_2$. □

Now, we introduce the general Gronwall inequality (or the comparison principle), which is important for a priori energy estimates of FPDEs:

Theorem 6. *Let $f(t,v)$ be a continuous function such that it satisfies the conditions in Theorem 5 and $\forall t \geq 0$, $x \leq y$ implies $f(t,x) \leq f(t,y)$. Let $0 < \gamma < 1$. Suppose $v_1(t)$ is continuous satisfying*

$$D_c^\gamma v_1 \leq f(t, v_1),$$

where this inequality means $D_c^\gamma v_1 - f(t,v_1)$ is a negative distribution. Suppose also that $v_2 \in X$ is the solution of the equation

$$D_c^\gamma v_2 = f(t, v_2), \quad v_2(0) \geq v_1(0),$$

which exists and is continuous by Theorem 5. Then, $v_1(t) \leq v_2(t)$ for all $t > 0$.

Correspondingly, if

$$D_c^\gamma v_1 \geq f(t, v_1),$$

and $v_2 \in X$ solves

$$D_c^\gamma v_2 = f(t, v_2), \quad v_2(0) \leq v_1(0).$$

Then, $v_1(t) \geq v_2(t)$ for all $t > 0$.

Proof. Denote by v_2^ϵ the solution of the equation with $v_2(0) + \delta$. Define $T^\epsilon = \inf\{t > 0 : v_2^\epsilon(t) \leq v_1(t)\}$. Since both v_1 and v_2^ϵ are continuous and $v_2^\epsilon(0) > v_1(0)$, $T^\epsilon > 0$. We claim that $T^* = \infty$. For otherwise, we have $v_2^\epsilon(T^*) = v_1(T^*)$ and $v_1(t) < v_2(t)$ for $t < T^*$. Note that $f(t, v_1)$ is a continuous function. By using Theorem 4 and Lemma 16:

$$\begin{aligned} v_1(T^*) = v_1(0) + \frac{1}{\Gamma(\gamma)} \int_0^{T^*} (T^* - s)^{\gamma-1} D_c^\gamma v_1 ds &\leq v_1(0) + \frac{1}{\Gamma(\gamma)} \int_0^{T^*} (T^* - s)^{\gamma-1} f(s, v_1(s)) ds \\ &< v_2(0) + \epsilon + \frac{1}{\Gamma(\gamma)} \int_0^{T^*} (T^* - s)^{\gamma-1} f(s, v_2(s)) ds = v_2^\epsilon(T^*). \end{aligned}$$

(*The first integral is understood as $[u(t)t^{\gamma-1}] * D_c^\gamma v_1$. However, the obtained distribution is a continuous function and Lemma 16 guarantees that we can have the first inequality.*) This is a contradiction. Hence, $v_1(t) < v_2^\epsilon(t)$ for all $t > 0$.

Taking $\epsilon \to 0^+$ and using the continuity on initial value yield the claim. Similarly arguments hold for the second claim, except that we perturb $v_2(0)$ to $v_2(0) - \epsilon$ to construct v_2^ϵ. □

We now show another result that may be useful for FPDEs:

Proposition 8. *Suppose f is Lipschitz continuous and nondecreasing, satisfying $f(0) \geq 0$. Then, the solution of $D_c^\gamma v = f(v), v(0) \geq 0$ is continuous and nondecreasing.*

Proof. v is continuous by Theorem 5. It is clear that $f(v) \geq 0$ whenever $v \geq 0$. We first show that $v(t) \geq v(0)$ for all $t \geq 0$.

Let v^ϵ be the solution with initial data $v(0)+\epsilon > v(0)$. v^ϵ is continuous. Define $T^* = \inf\{t : v^\epsilon(t) \leq v(0)\}$. $T^* > 0$ because $v^\epsilon(0) > v(0)$. We show that $T^* = \infty$. If this is not true, $v^\epsilon(T^*) = v(0)$ and $v^\epsilon(t) > v(0) \geq 0$ for all $t < T^*$. Applying Theorem 4 for $t = T^*$ says $v^\epsilon(T^*) > v(0)$, a contradiction. Hence, $v^\epsilon \geq v(t)$ for all $t > 0$. Taking $\epsilon \to 0$ yields $v(t) \geq v(0)$ for all $t \geq 0$, since the solution is continuous on initial data by Theorem 5.

Now, consider the function sequence

$$D_c^\gamma v^n = f(v^{n-1}), v^n(0) = v(0), \quad v^0 = v(0) \geq 0.$$

All functions are continuous. Since $v(t) \geq v^0$, then $f(v) \geq f(v^0)$, and it follows that $v \geq v^1$ by Theorem 4. Doing this iteratively, we find that $v \geq v^n$ for all $n \geq 0$.

Theorem 4 shows that

$$v^1 = v(0) + f(v(0))\frac{1}{\Gamma(1+\gamma)}t^\gamma \ge v^0.$$

By Theorem 4 again, it follows that $v_2 \ge v_1$ and hence $v^n \ge v^{n-1}$ for all $n \ge 1$. Therefore, v^n is increasing in n and bounded above by v. Then, $v^n \to \tilde{v}$ uniformly on $[0,T]$. $\tilde{v} \in C^0[0,T]$. Since T is arbitrary, $\tilde{v} \in C^0[0,\infty)$ and thus in $L^1_{loc}[0,\infty)$. In distribution sense, we therefore have $v^n \to \tilde{v}$ and thus $D_c^\gamma v^n \to D_c^\gamma \tilde{v}$. Hence, $\tilde{v}$ is a solution of $D_c^\gamma v = f(v)$ with $\tilde{v}(0) = v(0)$. Since the solution is unique in X, it must be v.

This said, now we show that v^n is increasing in t. This is clear by induction if we note this fact: "If $h(t) \ge 0$ is a nondecreasing locally integrable function, then $g_\gamma * h$ is nondecreasing in t", which can be verified by direct computation. v, as the the limit of increasing functions, is increasing. □

The results in Proposition 5 can be generalized to $\varphi \in \mathbb{R}^m$ easily and we conclude the following:

Proposition 9. *Suppose that $E(\cdot) \in C^1(\mathbb{R}^m, \mathbb{R})$ is convex and that ∇E is Lipschitz continuous. Let $v : [0,\infty) \to \mathbb{R}^m$ solve the FODE:*

$$D_c^\gamma v = -\nabla_v E(v) \tag{55}$$

such that $D_c^\gamma E(v)$ is a measurable function. Assume also $\exists \eta \in C_c^\infty(-\infty, 0)$ and a sequence $\epsilon_k \to 0$ such that $D_c^\gamma E(v^{\epsilon_k}), D_c^\gamma v^{\epsilon_k}$ (See (42)) each converges in $L^1_{loc}[0,\infty)$. Then

$$E(v(t)) \le E(v(0)).$$

Similarly, with the same conditions except that the equation is of the form:

$$D_c^\gamma v = J\nabla_v E(v), \tag{56}$$

where J is an anti-Hermitian constant operator, then

$$E(v(t)) \le E(v(0)).$$

Proof. If $E(\cdot) \in C^1$, then $\nabla_v E(v)$ is continuous. By Theorem 5, v is continuous. Then, $D_c^\gamma v$ is continuous by the equation. $v^\epsilon \to v$ uniformly on $[0,T]$ and thus bounded. Then, $E(v^\epsilon) \to E(v)$ uniformly on $[0,T]$. Since T is arbitrary, $E(v^\epsilon) \to E(v)$ in $L^1_{loc}[0,\infty)$ and thus in distribution. Hence, $D_c^\gamma E(v^\epsilon) \to D_c^\gamma E(v)$ in distribution.

Passing limit on the subsequences for

$$D_c^\gamma E(v^\epsilon) \le \nabla_v E(v^\epsilon) \cdot D_c^\gamma v^\epsilon,$$

by the conditions given, left hand side converges to $D_c^\gamma E(v)$ in $L^1_{loc}[0,\infty)$ and the right hand side converges to $\nabla_v E(v) D_c^\gamma u$ in $L^1_{loc}[0,\infty)$ because $\nabla_v E(v^\epsilon) \to \nabla_v E(v)$ uniformly on any $[0,T]$ interval. Then, it follows that

$$D_c^\gamma E(v) \le \nabla_v E(v) \cdot D_c^\gamma v = -|\nabla_u E(v)|^2.$$

The solution to $D_c^\gamma E(v) = 0$ is a constant $E(v) = E(v(0))$. By Theorem 6 (The function is $f(t, E(v)) = 0$), we conclude that

$$E(v(t)) \le E(v(0)).$$

For the second case, we have

$$D_c^\gamma E(v) \le \nabla_v E(v) \cdot D_c^\gamma u = \nabla_v E \cdot (J\nabla_v E) = 0.$$

Theorem 6 again yields

$$E(v(t)) \le E(v(0)).$$

□

The physical background of fractional Hamiltonian system $D_c^\gamma v = J\nabla_v E(v)$ has been discussed in [24, 3]. The fractional Hamiltonian system can be rewritten as $v(t) = v(0) + J_\gamma(J\nabla_v E(v))$ by Theorem 4, which is of the Volterra type $v_0 \in v(t) + b * (Av)$. The general Volterra equations with completely positive kernels and m-accretive A operators have been discussed in [2] and the solutions have been shown to converge to the equilibrium. In the fractional Hamiltonian system, $-J\nabla_v E(v)$ is not m-accretive and it is not clear whether the solutions converge to one equilibrium satisfying $\nabla_v E(v) = 0$ or not. However, for the following simple example, the energy function E indeed dissipates and the solutions converge to the equilibrium:

Example: Consider $E(p,q) = \frac{1}{2}(p^2 + q^2)$ and

$$D_c^\gamma q = \frac{\partial E}{\partial p} = p,$$
$$D_c^\gamma p = -\frac{\partial E}{\partial q} = -q,$$

where we assume $\gamma < 1/2$, and the initial conditions are $p(0) = p_0, q(0) = q_0$.

Applying Theorem 5 for $u(t) = (p,q)$, we find that both p and q are continuous functions. By Lemma 14, $D_c^\gamma(D_c^\gamma q) = D_c^{2\gamma} q - p_0 g_{1-\gamma}$. Applying D_c^γ on the first equation yields

$$D_c^{2\gamma} q - p_0 g_{1-\gamma} = D_c^\gamma p = -q.$$

By Proposition 7, we find

$$q(t) = q_0\beta_{2\gamma}(t) - p_0 g_{1-\gamma} * (\beta_{2\gamma}') = q_0\beta_{2\gamma}(t) - p_0 D_c^\gamma \beta_{2\gamma},$$

where $\beta_{2\gamma} = E_{2\gamma}(-t^{2\gamma})$ is defined as in Proposition 7. Note that $g_{1-\gamma} * (\beta_{2\gamma}') = D_c^\gamma \beta_{2\gamma}$ is due to Proposition 2.

Using the second equation, we find

$$p(t) = p_0 + J_\gamma(-q) = p_0\beta_{2\gamma} - q_0 J_\gamma \beta_{2\gamma}.$$

Actually, from the equation of p, $D_c^{2\gamma} p = -p - q(0) g_{1-\gamma}$, we find

$$p(t) = p_0\beta_{2\gamma}(t) + q_0 D_c^\gamma \beta_{2\gamma}.$$

Since $\beta_{2\gamma}$ solves the equation $D_c^{2\gamma} v = -v$, we see $D_c^\gamma \beta_{2\gamma} = -J_\gamma \beta_{2\gamma}$. Those two expressions for $p(t)$ are identical. Hence, we find that

$$E(t) = E(0)(\beta_{2\gamma}^2 + (J_\gamma \beta_{2\gamma})^2). \tag{57}$$

(Note that $\beta_{2\gamma}$ and $J_\gamma\beta_{2\gamma}$ are the solutions to the following two equations respectively:

$$D_c^{2\gamma} v = -v, \quad v(0) = 1,$$
$$D_c^{2\gamma} v = -v + g_{1-\gamma}, \quad v(0) = 0.$$

Unlike the corresponding ODE system where the two components are both solutions to $v'' = -v$, here since $D_c^{2\gamma} = -v$ only has one solution for an initial value, the two functions are from different equations.) By the series expression of $E_{2\gamma} = E_{2\gamma,1}$ ([14]):

$$\beta_{2\gamma} = E_{2\gamma}(-t^{2\gamma}) = \sum_{n=0}^{\infty}(-1)^n g_{2n\gamma+1}, \quad J_\gamma\beta_{2\gamma} = \sum_{n=0}^{\infty}(-1)^n g_{(2n+1)\gamma+1} = t^\gamma E_{2\gamma,\gamma+1}(-t^{2\gamma}). \tag{58}$$

According to the asymptotic behavior listed in [10, Eq. (7)],

$$E_{2\gamma,\rho}(-t^{2\gamma}) \sim -\sum_{k=1}^{p} \frac{(-1)^k t^{-2\gamma k}}{\Gamma(\rho - 2\gamma k)}, \quad 0 < \gamma < 1. \tag{59}$$

As examples, $E_{1/2,1}(-t^{1/2}) = e^t \left(1 - \frac{2}{\sqrt{\pi}} \int_0^{t^{1/2}} \exp(-s^2) ds\right)$ and $E_{1/2,1}(-t^{1/2}) \sim 1/t^{1/2}$. $E_{1,1}(-t) = e^{-t}$ decays exponentially while in Equation (59), $\Gamma(1-k) = \infty$ for all $k = 1, 2, 3, \ldots$. Note that one cannot take the limit $\gamma \to 1$ in Equation (59) and conclude that $E_{2,1}(-t^2)$ and $tE_{2,2}(-t^2)$ both decay exponentially. Actually, $\beta_2 = E_{2,1}(-t^2) = \cos(t)$ and $J_1\beta_2 = tE_{2,2}(-t^2) = \sin(t)$. The singular limit is due to an exponentially small term $C_1 \exp(C_2(\gamma - 1)t)$ in $E_{2\gamma,\rho}$.

If $\gamma < 1/2$, $\Gamma(\beta - 2\gamma) \neq \infty$ and the leading order behavior is $t^{-2\gamma}$. Hence, $J_\gamma \beta_{2\gamma} \sim t^{-\gamma}$ and E decays like $t^{-2\gamma}$. Following the same method, one can use the last statement in Lemma 14 to solve the case $\gamma = 1/2$ and find that $\gamma = 1/2$ case is still right: $\beta_1 = E_{2,1}(-t) = e^{-t}$ decays exponentially fast while $t^{1/2}E_{2,2}(-t)$ decays like $t^{-1/2}$. Since E decays to zero, the solution must converge to $(0, 0)$.

Whether this is true for $1/2 < \gamma < 1$ is interesting, which we leave for future study.

Remark 9. According to Theorem 5, the system has a unique solution for any $(p(0), q(0)) \in \mathbb{R}^2$ and $0 < \gamma < 1$. We have only considered $0 < \gamma < 1/2$ here just because we can find the solution easily for these cases. If one defines the Caputo derivatives for $1 < \alpha < 2$ ($\alpha = 2\gamma$) consistently, we guess that the expressions above are still correct.

7. A discrete convolution group and discrete fractional calculus

We now introduce a special discrete convolution group generated by a completely monotone sequence to define discrete fractional calculus and show that the discrete fractional calculus is consistent with the Riemann-Liouville fractional calculus (Definition 3) with appropriate time scaling.

To motivate this, let us consider a smooth function $f(t)$. We sample this function with step size k and obtain a sequence $a = \{a_i\}_{i=0}^\infty$ with $a_i = f(ik)$.

Using numerical approximations ([17]) for the fractional calculus, we find the following sequence for fractional integral J_γ, $0 < \gamma \leq 1$:

$$(c_\gamma)_j = \frac{k^\gamma}{\gamma\Gamma(\gamma)}((j+1)^\gamma - j^\gamma).$$

Then, $J_\gamma f \approx c_\gamma * a$. c_γ is completely monotone.

These sequences $\{c_\gamma\}$ do not form a convolution semi-group. However, each sequence generates a group. Let $\{c_\gamma^{(\alpha)}\}$ be the group generated by c_γ, with $c_\gamma^{(\gamma)} = c_\gamma$. We hope that $\{c_\gamma^{(\alpha)}\}$ is a reasonable convolution group to define discrete fractional calculus. Below, we focus on the case $\gamma = 1$.

In general, the sequence may not be from sampling of f and we may not have the concept of time step k. This then motivates the following discrete fractional calculus for a general given sequence:

7.1. A definition of discrete fractional calculus.

Consider the sequence

$$c^{(\alpha)} := c_1^{(\alpha)}, \forall \alpha \in \mathbb{R}, \tag{60}$$

whose generating function is $F(z) = (1-z)^{-\alpha}$. Note that $c^{(1)} = (1, 1, 1, \ldots)$ and $c^{(\alpha)}, 0 \leq \alpha \leq 1$ are completely monotone.

Definition 6. For a sequence $a = (a_0, a_1, \ldots)$, we define the discrete fractional operators $J_\alpha^d : \mathbb{R}^\mathbb{N} \to \mathbb{R}^\mathbb{N}$ as

$$J_\alpha^d a = c^{(\alpha)} * a. \tag{61}$$

By the discussion in [18], $\{c^{(\alpha)}\}$ form a discrete convolution group, and it follows that $\{J_\alpha^d : \alpha \in \mathbb{R}\}$ form a group.

7.2. **Consistency with the continuous convolution group.** Even though we define the discrete convolution group without worrying its relationship with the Riemann-Liouville fractional calculus, we show that they are actually consistent. Given a function time-continuous function $f(t)$, we pick a time step $k>0$ and define the sequence a with $a_i=f(ik)$ $(i=0,1,2,\dots)$. We consider

$$T_\alpha f = k^\alpha J_\alpha^d a. \tag{62}$$

We now show that for $t>0$ $(T_\alpha f)_n$ converges to $J_\alpha f(t)$ as $k=t/n\to 0^+$ when f is smooth. To be convenient, we'll only consider $|\alpha|\le 1$. We will start with some lemmas.

Lemma 17. *The m-th term of $c^{(\alpha)}$ has the following asymptotic behavior as $m\to\infty$:*

$$c_m^{(\alpha)} = [z^m](1-z)^{-\alpha} \sim \frac{m^{\alpha-1}}{\Gamma(\alpha)}\left(1+\frac{\alpha(\alpha-1)}{2m}+O(\frac{1}{m^2})\right), \tag{63}$$

for $\alpha\ne 0,-1,-2,\dots$.

Here, $[z^m]F(z)$ means the coefficient of z^m in the Taylor series of $F(z)$ about center 0. One can refer to [8] for the proof.

Lemma 18. *For $|\alpha|<1$, let $S_n=\sum_{i=0}^n c_i^{(\alpha)}$. Then, as $m\to\infty$, we have:*

$$S_m = \frac{m^\alpha}{\Gamma(1+\alpha)}\left(1+O(\frac{1}{m})\right), \tag{64}$$

$$R_m = \sum_{i=0}^m (m-i)c_i^{(\alpha)} = \frac{m^{1+\alpha}}{\Gamma(2+\alpha)}\left(1+O(\frac{1}{m})\right). \tag{65}$$

Proof. $\alpha=0$ is trivial. Suppose $\alpha\ne 0$. $\{S_n\}$ is the convolution between $c^{(\alpha)}$ and $c^{(1)}$ and $S=c^{(\alpha+1)}$ by the group property. Hence, the generating function for S is $(1-z)^{-1-\alpha}$. Similarly, since $c^{(2)}=(1,2,3,\dots)$, $R=c^{(\alpha+2)}$. Applying Lemma 17 yields the claims. ☐

Now we state the consistency result:

Theorem 7. *Suppose $f\in C^2[0,\infty)$. For any $t>0$, define $k=t/n$. Then, $|(T_\alpha f)_n-(J_\alpha f)(t)|\to 0$ as $n\to\infty$ for $|\alpha|\le 1$.*

Proof. Below, we only show the consistency and we are not trying to find the best estimate for the convergence rate.

$\alpha=0$, $(T_0 f)_n=f(t)$ and the claim is trivial.

Case 1: $\alpha>0$

If $\alpha=1$, we have $c^{(1)}=k(1,1,1,1,1,\dots)$ is the sequence for integration. We find that $T_\alpha f=\sum_{i=0}^n kf(t-ik)$. This is $O(k)$ approximation for the integral.

Consider $0<\alpha<1$. Fix $t>0$ and $k=t/n$ with $n\gg 1$. Let $1\ll K\ll n$. By Lemma 17, we have

$$(T_\alpha f)_n = k^\alpha\sum_{i=0}^{K-1} c_i^{(\alpha)} f((n-i)k) + k^\alpha\sum_{i=K}^n \frac{i^{\alpha-1}}{\Gamma(\alpha)}f((n-i)k) + k^\alpha\sum_{i=K}^n O(\frac{1}{i^{2-\alpha}}).$$

The last term is easily estimated: $k^\alpha\sum_{i=K}^n O(\frac{1}{i^{2-\alpha}})=O(K^{\alpha-1}k^\alpha)$.

Consider the first term. We have $f((n-i)k)=f(t)-f'(\xi)ik$ and $f(t-s)=f(t)-f'(\tilde{\xi})s$. Then, it follows by Lemma 17 and Lemma 18

$$\Big|k^{\alpha}\sum_{i=0}^{K-1}c_i^{(\alpha)}f((n-i)k)-\frac{1}{\Gamma(\alpha)}\int_0^{(K-1)k}f(t-s)s^{\alpha-1}ds\Big|\le |f(t)|\Big|k^{\alpha}\sum_{i=0}^{K-1}c_i^{(\alpha)}-\frac{1}{\Gamma(\alpha)}\int_0^{(K-1)k}s^{\alpha-1}ds\Big|$$
$$+\sup|f'|Kk^{\alpha+1}\sum_{i=0}^{K-1}c_i^{(\alpha)}+C\int_0^{(K-1)k}\sup|f'|s^{\alpha}ds\le C(K^{\alpha-1}k^{\alpha}+K^{1+\alpha}k^{1+\alpha}).$$

Lastly a very rough estimate is

$$\left|k^{\alpha}\sum_{i=K}^{n}\frac{i^{\alpha-1}}{\Gamma(\alpha)}f((n-i)k)-\frac{1}{\Gamma(\alpha)}\int_{(K-1)k}^{t}f(t-s)s^{\alpha-1}ds\right|\le C(Kk)^{\alpha-2}k.$$

Choosing $K\sim k^{-1/2}$, all terms tend to zero as $k\to 0$.

Case 2: $-1\le\alpha<0$. $\gamma=|\alpha|$.

If $\alpha=-1$, the sequence is $(1,-1,0,0,\ldots)$. $(T_{-1}f)_n=f'(nk)+O(k)=f'(t)+O(k)$ by the standard finite difference.

In the case that $f(t)\in\tilde{H}^1(0,T)\cap C^2[0,\infty)$, $f(0)=0$. $(T_{-1}f)_i=f'(ik)-f'(0)\delta_{i0}+O(k)$ for any $i\le n$. The consistency for $-1<\alpha<0$ follows then from the group property

$$(T_{\alpha}f)_n=(T_{-\gamma}f)_n=[T_{1-\gamma}(T_{-1}f)]_n=[T_{1-\gamma}f']_n-f'(0)c_n^{(1-\gamma)}k^{1-\gamma}+O(k).$$

We still have $O(k)$ as the error because $\|k^{1+\alpha}c^{(1+\alpha)}\|_1\le 2k^{1+\alpha}$ by Theorem 1, and $\|(k^{1+\alpha}c^{(1+\alpha)})*O(k)\|_{\infty}\le Ck$. By the result for integral, we find that $T_{1-\gamma}(f')-\int_0^t(t-s)^{-\gamma}f'(s)ds\to 0$ as $k\to 0$. Since $f(0)=0$, the latter is equal to $(J_{-\gamma}f)(t)$. The consistency is shown.

Consider a general smooth function f that we do not have $f(0)=0$. $(T_{-1}f)_0=f(0)/k$. The first term is very singular and the above group argument fails. We now verify directly that the consistency is also true. Consider that $\alpha\in(-1,0)$ and $\gamma=|\alpha|$. By definition,

$$(T_{-\gamma}f)_n=\frac{f(t)}{k^{\gamma}}+\frac{1}{k^{\gamma}}\sum_{i=1}^{n}c_i^{(-\gamma)}f((n-i)k).$$

The continuous Riemann-Liouville fraction derivative equals

$$\begin{aligned}(J_{-\gamma}f)(t)&=f(0)\frac{1}{\Gamma(1-\gamma)}t^{-\gamma}+\frac{1}{\Gamma(1-\gamma)}\int_0^t\frac{f'(s)}{(t-s)^{\gamma}}ds\\&=\frac{f(t-k/b)}{k^{\gamma}}+\frac{1}{\Gamma(1-\gamma)}\left[\int_{t-k/b}^{t}\frac{f'(s)}{(t-s)^{\gamma}}ds-\gamma\int_0^{t-k/b}\frac{f(s)}{(t-s)^{\gamma+1}}ds\right],\end{aligned}$$

where b is chosen such that $b^{\gamma}=\Gamma(1-\gamma)\ge 1$.

It is not hard to see

$$\frac{f(t)}{k^{\gamma}}-\frac{f(t-k/b)}{k^{\gamma}}-\frac{1}{\Gamma(1-\gamma)}\int_{t-k/b}^{t}\frac{f'(s)}{(t-s)^{\gamma}}ds=O(k^{1-\gamma}).$$

We need to estimate

$$\Big|\frac{1}{k^{\gamma}}\sum_{i=1}^{n}c_i^{(-\gamma)}f((n-i)k)-\frac{1}{\Gamma(-\gamma)}\int_{k/b}^{t}\frac{f(t-s)}{s^{\gamma+1}}ds\Big|. \tag{66}$$

We first observe that expression (66) is of order $O(k)$ and $O((k/b)^{1-\gamma})$ if $f(s)=1$ and $f(s)=t-s$ respectively.

Consider that $f=1$. By Lemma 18 and noting $b^\gamma=-\gamma\Gamma(-\gamma)$, we have

$$k^{-\gamma}\sum_{i=1}^{n}c_i^{(-\gamma)}=k^{-\gamma}\sum_{i=0}^{n}c_i^{(-\gamma)}-k^{-\gamma}=k^{-\gamma}\left(\frac{n^{-\gamma}}{\Gamma(1-\gamma)}-1\right)+O\left(\frac{1}{(nk)^\gamma n}\right)$$
$$=\frac{1}{\Gamma(-\gamma)}\int_{k/b}^{t}\frac{1}{s^{\gamma+1}}ds+O\left(\frac{1}{(nk)^\gamma n}\right).$$

For $f(s)=t-s$, we need to estimate $|k^{-\gamma}\sum_{i=1}^n c_i^{(-\gamma)}(ik)-\frac{1}{\Gamma(-\gamma)}\int_{k/b}^t s^{-\gamma}ds|$. By Lemma 18 again, we find

$$k^{-\gamma}\sum_{i=1}^{n}c_i^{(-\gamma)}(n-i)k-\frac{1}{\Gamma(-\gamma)}\int_{k/b}^{t}\frac{t-s}{s^{\gamma+1}}ds=O((k/b)^{1-\gamma}).$$

By what has been just computed $k^{-\gamma}\sum_{i=1}^n c_i^{(-\gamma)}nk-\frac{1}{\Gamma(-\gamma)}\int_{k/b}^t\frac{t}{s^{\gamma+1}}ds=O(k)$ and thus

$$\left|k^{-\gamma}\sum_{i=1}^{n}c_i^{(-\gamma)}ik-\frac{1}{\Gamma(-\gamma)}\int_{k/b}^{t}s^{-\gamma}ds\right|=O((k/b)^{1-\gamma}).$$

Using these two facts, we can therefore assume $f(t)=f'(t)=0$ in Equation (66) with introducing error $O(k^{1-\gamma})$.

For $1\le i\le K-1$, $|f((n-i)k)|\le C(ik)^2$. Since $c_i^{(-\gamma)}$ is all negative for $i\ge 1$ and by Lemma 18:

$$|k^{-\gamma}\sum_{i=1}^{K-1}c_i^{(-\gamma)}\frac{1}{2}f''(\xi)(ik)^2|\le k^{-\gamma}Kk^2C_1\sum_{i=1}^{K-1}|ic_i^{(-\gamma)}|\le CKk^{2-\gamma}K^{1-\gamma}\le C(Kk)^{2-\gamma}.$$

Also, for $k/b\le s\le (K-1)k$, $|f(t-s)|\le Cs^2$, and hence

$$\left|\int_{k/b}^{(K-1)k}\frac{f(t-s)}{s^{\gamma+1}}ds\right|\le C(Kk)^{2-\gamma}.$$

By Lemma 17,

$$\left|k^{-\gamma}\sum_{i=K}^{n}\left(c_i^{(-\gamma)}-\frac{i^{-1-\gamma}}{\Gamma(-\gamma)}\right)f((n-i)k)\right|\le CK^{-1-\gamma}.$$

Hence,

$$\left|k^{-\gamma}\sum_{i=K}^{n}c_i^{(-\gamma)}f((n-i)k)-\frac{1}{\Gamma(-\gamma)}\int_{(K-1)k}^{t}\frac{f(t-s)}{s^{\gamma+1}}ds\right|\le$$
$$CK^{-1-\gamma}+\left|k^{-\gamma}\sum_{i=K}^{n}\frac{i^{-1-\gamma}}{\Gamma(-\gamma)}f((n-i)k)-\frac{1}{\Gamma(-\gamma)}\int_{(K-1)k}^{t}\frac{f(t-s)}{s^{\gamma+1}}ds\right|.$$

The last term, by a very rough estimate, could be bounded by $(Kk)^{-2-\gamma}k$. Hence, if $K=k^{\epsilon-\frac{1+\gamma}{2+\gamma}}$, the last term, $(Kk)^{2-\gamma}$ and $K^{-1-\gamma}$ all tend to zero as $k\to 0$. □

Remark 10. In the case $\alpha=-1$ and $f(0)\neq 0$, for which the alternative group argument in the proof fails, $(T_\alpha f)_0=\frac{f(0)}{k}$. This actually approximates the singular term in $J_{-1}f$, $\delta(t)f(0)$.

Acknowledgements

L. Li is grateful to Xianghong Chen and Xiaoqian Xu for discussion on Sobolev spaces. The work of J.-G Liu is partially supported by KI-Net NSF RNMS grant No. 1107291 and NSF DMS grant No. 1514826.

References

1. M. Allen, L. Caffarelli, and A. Vasseur, *A parabolic problem with a fractional-time derivative*, Archive for Rational Mechanics and Analysis **221** (2016), no. 2, 603–630.
2. Ph Clément and JA Nohel, *Asymptotic behavior of solutions of nonlinear volterra equations with completely positive kernels*, SIAM Journal on Mathematical Analysis **12** (1981), no. 4, 514–535.
3. J. Cresson and P. Inizan, *About fractional hamiltonian systems*, Physica Scripta **2009** (2009), no. T136, 014007.
4. P. Degond, J. Liu, and R.L. Pego, *Coagulation-fragmentation model for animal group-size statistics*, arXiv preprint arXiv:1510.06077 (2015).
5. K. Diethelm, *The analysis of fractional differential equations: An application-oriented exposition using differential operators of Caputo type*, Springer, 2010.
6. K. Diethelm and N.J. Ford, *Analysis of fractional differential equations*, Journal of Mathematical Analysis and Applications **265** (2002), no. 2, 229–248.
7. L.C. Evans, *Partial differential equations*, (2010).
8. P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University press, 2009.
9. G.B. Folland, *Introduction to partial differential equations*, Princeton university press, 1995.
10. R. Gorenflo, J. Loutchko, and Y. Luchko, *Computation of the Mittag-Leffler function $E_{\alpha,\beta}(z)$ and its derivative*, Fract. Calc. Appl. Anal, Citeseer, 2002.
11. R. Gorenflo, Y. Luchko, and M. Yamamoto, *Operator theoretic approach to the Caputo derivative and the fractional diffusion equations*, arXiv preprint arXiv:1411.7289 (2014).
12. R Gorenflo and F Mainardi, *Fractional calculus: Integral and differential equations of fractional order*, 1997.
13. B.G. Hansen and F.W. Steutel, *On moment sequences and infinitely divisible sequences*, J. Math. Anal. Appl. **136** (1988), no. 1, 304–313.
14. H.J. Haubold, A.M. Mathai, and R.K. Saxena, *Mittag-Leffler functions and their applications*, Journal of Applied Mathematics **2011** (2011).
15. B. Jin, R. Lazarov, J. Pasciak, and W. Rundell, *Variational formulation of problems involving fractional order differential operators*, Math. Comput. (2015).
16. A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, vol. 204, Elsevier Science Limited, 2006.
17. Y. Lin and C. Xu, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys. **225** (2007), no. 2, 1533–1552.
18. J. Liu and R. Pego, *On generating functions of Hausdorff moment sequences*, Transactions of the American Mathematical Society (2016).
19. F. Mainardi, P. Paradisi, and R. Gorenflo, *Probability distributions generated by fractional diffusion equations*, arXiv preprint arXiv:0704.0320 (2007).
20. D. Matignon, *Stability results for fractional differential equations with applications to control processing*, Computational engineering in systems applications, vol. 2, 1996, pp. 963–968.
21. H.L. Royden, *Real analysis*, 1988.
22. S. Ruscheweyh, L. Salinas, and T. Sugawa, *Completely monotone sequences and universally prestarlike functions*, Israel J. Math. **171** (2009), no. 1, 285–304.
23. R.L. Schilling, R. Song, and Z. Vondracek, *Bernstein functions: theory and applications*, vol. 37, Walter de Gruyter, 2012.
24. A.A. Stanislavsky, *Hamiltonian formalism of fractional systems*, The European Physical Journal B-Condensed Matter and Complex Systems **49** (2006), no. 1, 93–101.
25. M. Taylor, *Remarks on fractional diffusion equations*, Preprint.
26. D.V. Widder, *The Laplace transform*, Princeton University Press, 1941.

Department of Mathematics, Duke University, Durham, NC 27708, USA.
E-mail address: `leili@math.duke.edu`

Departments of Physics and Mathematics, Duke University, Durham, NC 27708, USA.
E-mail address: `jliu@phy.duke.edu`