QUARTERLY OF APPLIED MATHEMATICS
http://dx.doi.org/10.1090/qam/1479
Article electronically published on August 22, 2017

A NOTE ON DECONVOLUTION WITH COMPLETELY MONOTONE SEQUENCES AND DISCRETE FRACTIONAL CALCULUS

By

LEI LI (*Department of Mathematics, Duke University, Durham, NC 27708*)

AND

JIAN-GUO LIU (*Departments of Physics and Mathematics, Duke University, Durham, NC 27708*)

Abstract. We study in this work convolution groups generated by completely monotone sequences related to the ubiquitous time-delay memory effect in physics and engineering. In the first part, we give an accurate description of the convolution inverse of a completely monotone sequence and show that the deconvolution with a completely monotone kernel is stable. In the second part, we study a discrete fractional calculus defined by the convolution group generated by the completely monotone sequence $c^{(1)} = (1, 1, 1, \ldots)$, and show the consistency with time-continuous Riemann-Liouville calculus, which may be suitable for modeling memory kernels in discrete time series.

1. Introduction.

Many models have been proposed for the ubiquitous time-delay memory effect in physics and engineering: the generalized Langevin equation model for particles in heat bath ([7,18]), linear viscoelasticity models for soft matter ([2,12]), linear dielectric susceptibility model [1,15] for polarization to name a few. In these models, the response due to memory is given by the one-side convolution $\int_0^t g(t-s)v(s)\,ds$ following linearity, time-translation invariance and causality [11, Chap. 1], where g is the memory kernel and v is the source of memory. Causality means that the output cannot precede the input so that $g(t) = 0$ for $t < 0$. The Tichmarsh's theorem states that the Fourier transform $G(\omega)$ of g is analytic in the upper half plane, and that the real and imaginary parts of G satisfy the Kramers-Kronig relation [11,16]. Based on the principle of the fading memory [12], we consider g to be completely monotone, which by the Bernstein theorem can be expressed as the superposition of (may be infinitely many) decaying exponentials (see [14,17] for more details). If the kernel g is given by the algebraically decaying completely monotone kernels $g = \frac{\theta(t)}{\Gamma(\gamma)}t^{\gamma-1}$ where $\theta(t)$ is the Heaviside step

Received July 3, 2017.
2010 *Mathematics Subject Classification.* Primary 47D03.
Key words and phrases. Convolution group, convolution inverse, completely monotone sequence, fractional calculus, Riemann-Liouville derivative.
E-mail address: leili@math.duke.edu
E-mail address: jliu@phy.duke.edu

function and $\gamma \in (0,1)$, we are then led to the fractional integrals and the corresponding fractional derivatives, which have already been used widely in engineering for modeling memory effects [4].

In practice, the data we collect are at discrete times and we have the one-sided discrete convolution $a*c$ (see equation (2.2)). The convolution kernel c is a completely monotone sequence (see Definition 2.1) if it is the value of g at the discrete times [17]. If c is completely monotone, it is shown in [10] 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. If $0 \leqslant r \leqslant 1$, $c^{(r)}$ is completely monotone. Further, $c^{(0)} = \delta_d := (1,0,0,\ldots)$, is the convolution identity. The most interesting sequence is $c^{(-1)}$, the convolution inverse, which can be used for deconvolution. Since the data are discrete, it would also be interesting to define discrete fractional calculus using the one-sided discrete convolution.

In this short note, we first investigate the convolution inverse of a completely monotone sequence c in Section 2. We show that the ℓ_1 norm is bounded and the deconvolution is stable in any ℓ^p space. Based on this, some preliminary ideas are explored for deconvolution. In Section 3, we define a discrete fractional calculus using a discrete convolution group generated by the completely monotone sequence $c^{(1)} = (1,1,1,\ldots)$ and show that it is consistent with the time-continuous Riemann-Liouville calculus (see (3.1)).

2. Deconvolution for a completely monotone kernel. In this section, we investigate the property of convolution inverse of a completely monotone sequence and deconvolution with completely monotone sequences.

Definition 2.1. A sequence $c = \{c_k\}_{k=0}^{\infty}$ is completely monotone if $(I-S)^j c_k \geqslant 0$ for any $j \geqslant 0, k \geqslant 0$ 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 $[0,1]$) ([17]). Another description is given as follows ([10, 13]):

Lemma 2.2. *A sequence c is completely monotone if and only if the generating function $F_c(z) = \sum_{j=0}^{\infty} c_j z^j$ is a Pick function that is analytic and nonnegative on $(-\infty, 1)$.*

Note that a function $f : \mathbb{C}_+ \mapsto \mathbb{C}$ (where $\mathbb{C}_+$ denotes the upper half plane, not including the real line) is Pick if it is analytic such that $\mathrm{Im}(z) > 0 \Rightarrow \mathrm{Im}(f(z)) \geqslant 0$.

Consider the one-sided convolution equation

$$a * c = f, \tag{2.1}$$

where the convolution kernel c is a completely monotone sequence and $c_0 > 0$. The discrete convolution is defined as

$$(a*c)_k = \sum_{n_1 \geqslant 0, n_2 \geqslant 0} \delta_k^{n_1+n_2} a_{n_1} c_{n_2}, \tag{2.2}$$

and δ_m^n is the Kronecker delta. This convolution is associative and commutative. Let $F_c(z)$ be the generating function of c:

$$F_c(z) = \sum_{n=0}^{\infty} c_n z^n. \tag{2.3}$$

Then, $F_{a*c}(z) = F_a(z)F_c(z)$. Given c, the convolution inverse $c^{(-1)}$ is the sequence that satisfies $c * c^{(-1)} = c^{(-1)} * c = \delta_d := (1, 0, 0, \ldots)$. The generating function of the convolution inverse $c^{(-1)}$ is $1/F_c(z)$. If we find the convolution inverse of c, the convolution equation (2.1) can be solved.

2.1. *The convolution inverse.* Now, we present our results about the convolution inverse:

THEOREM 2.3. 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 $(-c_1^{(-1)}, -c_2^{(-1)}, \ldots)$ is completely monotone. Furthermore, $0 \leqslant -\sum_{k=1}^{\infty} c_k^{(-1)} \leqslant \frac{1}{c_0}$.

Proof. The first claim follows from that $F_c(z)$ has no zeros in the unit disk [10].

By Lemma 2.2, $F_c(z)$ is Pick and it is positive on $(-\infty, 1)$. $F_c(-\infty) = 0$ if the corresponding Hausdorff measure does not have an atom at 0 (i.e. the sequence c is minimal. See [17, Chap. IV. Sec. 14] for the definition). Since $F_c(-\infty)$ could be zero, we consider

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

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

Suppose G_ϵ is the generating function of $d = (d_0^\epsilon, d_1^\epsilon, \ldots)$. By Lemma 2.2, this sequence is completely monotone. Then,

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

is the generating function of the shifted sequence $(d_1^\epsilon, \ldots)$, 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

$$H(z) = \frac{F_c(z) - F_c(0)}{zF_c(0)F_c(z)} \tag{2.4}$$

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

Further, $H(z)$ is the generating function of $-(c_1^{(-1)}, c_2^{(-1)}, \ldots)$ since $1/F_c(z)$ is the generating function of $c^{(-1)} = (c_0^{(-1)}, c_1^{(-1)}, \ldots)$. The second claim therefore follows. □

As a corollary of Theorem 2.3, we find that the deconvolution with a completely monotone sequence is stable:

Corollary 2.4. Equation (2.1) can be solved stably. In particular, $\forall f \in \ell^p$, there exists a unique $a \in \ell^p$ such that $a * c = f$ and $\|a\|_p \leqslant \frac{2}{c_0}\|f\|_p$.

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

2.2. *Computing convolution inverse and deconvolution.* To solve the convolution equation (2.1), we can use the algorithm in [10] to find the convolution group $c^{(r)}$. Then, the solution 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 $(n+1)c_{n+1}^{(r)} = r\sum_{k=0}^{n} c_{n-k}^{(r)} b_k$.

For a completely monotone sequence, the canonical sequence satisfies $b_k \geqslant 0$ ([5]). If $c_0 = 1$, computing the canonical sequence is straightforward

$$b_n = (n+1)c_{n+1} - \sum_{k=0}^{n-1} c_{n-k} b_k. \tag{2.5}$$

Note that $F_b(z) = F_c'(z)/F_c(z)$. If $c_0 = 1$, $c_0^{(-1)} = 1$ and $|c_{n+1}^{(-1)}| \leqslant \frac{1}{n+1}\sum_{k=0}^{n} |c_{n-1}^{(-1)}| b_k$. It's clear by induction that $|c_{n+1}^{(-1)}| \leqslant c_{n+1}$. For general c_0, we can apply the above argument to c/c_0 and have the pointwise bound: $|c_k^{(-1)}| \leqslant \frac{1}{c_0^2}|c_k|$.

Now, let us show a simple example to illustrate the deconvolution with completely monotone sequences. 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

$$\mathcal{C}_M = \Big\{\mu : \mu = h\sum_{i=1}^{M} \lambda_i \delta(x - x_i), \lambda_i \geqslant 0\Big\}. \tag{2.6}$$

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

$$d_n = \sum_{i=1}^{M} h\lambda_i x_i^n,\ n = 0, 1, 2, \ldots, \tag{2.7}$$

where $\lambda_i > 0$ are generated randomly (for example uniformly from $[0, 1]$).

In Fig. 1 (a), we have a sequence which is of square shape; in Fig. 1 (b), we plot the convolution between the sequence in (a) and the completely monotone sequence obtained using (2.7). Fig. 1 (c) shows the solution $a * c = f$ by convolving the sequence in Fig. 1(b) with $c^{(-1)}$. The original sequence is recovered accurately.

If the sequence c is no longer completely monotone, the generating function of $c^{(-1)}$ may have a small radius of convergence and an iterative method may be desired to

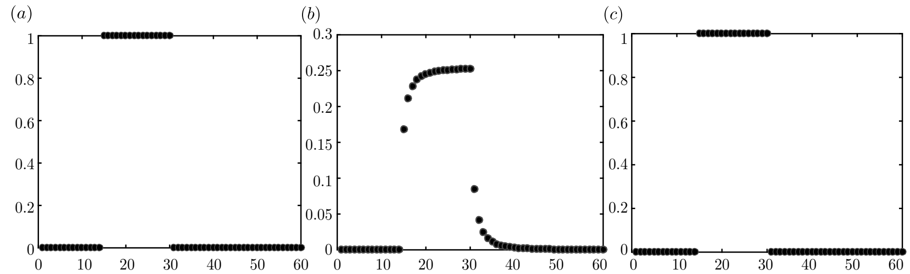

FIG. 1. A simple example of deconvolution

solve (2.1). Consider approximating the sequence c by a completely monotone sequence $d = \{d_n\}$ of the form in equation (2.7). Writing d in matrix form, we have

$$d = \frac{1}{m} A\lambda = A\eta, \tag{2.8}$$

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

$$a^{p+1} = f * d^{(-1)} - a^p * [(c-d) * d^{(-1)}],\ p = 0, 1, 2, \ldots, \tag{2.9}$$

where a^0 is arbitrary. Clearly, the iteration converges if $\|(c-d)*d^{(-1)}\|_1 < 1$. A sufficient condition is therefore

$$\|d^{(-1)}\|_1 \|c - d\|_1 \leqslant \frac{2}{\|\eta\|_1} \|c - A\eta\|_1 < 1, \tag{2.10}$$

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 (2.1).

3. A discrete convolution group and discrete fractional calculus. In this section, we introduce a special discrete convolution group generated by a completely monotone sequence and define discrete fractional calculus. We show that the discrete fractional calculus is consistent with the Riemann-Liouville fractional calculus ([4, 6, 8]) with appropriate time scaling. The discrete convolution group proposed may be suitable for modeling memory effects in discrete time series.

The traditional Riemann-Liouville fractional calculus for a function in $C^1[0, T), T > 0$ with index $|\alpha| \leqslant 1$ is defined as

$$(J_\alpha f)(t) = \begin{cases} \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f(s) ds, & \alpha > 0, \\ f(t), & \alpha = 0, \\ \frac{1}{\Gamma(1+\alpha)} \frac{d}{dt} \int_0^t \frac{f(s)}{(t-s)^{|\alpha|}} ds, & \alpha \in (-1, 0), \\ f'(t), & \alpha = -1. \end{cases} \tag{3.1}$$

In [8], a slightly different Riemann-Liouville calculus is proposed. The new definition introduces some singularities at $t = 0$ such that the resulted Riemann-Liouville calculus forms a group. However, for $t > 0$, the modified definition of a smooth function agrees with the traditional definition.

To motivate the discrete fractional calculus, we take a grid $t_i = ik : i = 0, 1, 2, \dots$ where k is the step size. Evaluating f at the grid points yields a sequence $a = \{a_i\}_{i=0}^{\infty}$, $a_i = f(ik)$. Using numerical approximations ([9]) for the fractional calculus, we find the following sequence for fractional integral J_γ, $0 < \gamma \leqslant 1$:

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

Then, $J_\gamma f \approx k^\gamma c_\gamma * a$. The sequences $\{c_\gamma\}$ do not form a convolution semi-group. However, each sequence generates a convolution group. Let $\{c_\gamma^{(\alpha)} : \alpha \in \mathbb{R}\}$ be the group generated by c_γ, with $c_\gamma^{(\gamma)} = c_\gamma$. It is desirable that $\{c_\gamma^{(\alpha)} : \alpha \in \mathbb{R}\}$ can be used to define discrete fractional calculus.

We focus on the case $\gamma = 1$ and we have $c^{(1)} := c_1^{(\alpha)} = (1, 1, \dots)$, with generating function $F_1(z) = (1-z)^{-1}$. The convolution group generated by $c^{(1)}$ is denoted by $c^{(\alpha)} := c_1^{(\alpha)} : \alpha \in \mathbb{R}$ and the generating function is $F_\alpha(z) = (1-z)^{-\alpha}, \forall \alpha \in \mathbb{R}$. $c^{(\alpha)}, 0 < \alpha \leqslant 1$ are completely monotone.

Definition 3.1. For a sequence $a = (a_0, a_1, \dots)$, we define the discrete fractional operators $I_\alpha : \mathbb{R}^{\mathbb{N}} \mapsto \mathbb{R}^{\mathbb{N}}$ as $a \mapsto I_\alpha a := c^{(\alpha)} * a$.

Clearly, $\{I_\alpha : \alpha \in \mathbb{R}\}$ form a group.

3.1. *Consistency with the time continuous fractional calculus.* In this subsection, we show that the discrete fractional calculus is consistent with Riemann-Liouville fractional calculus if $|\alpha| \leqslant 1$.

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

$$T_\alpha f = k^\alpha I_\alpha a. \tag{3.2}$$

We now show that for $t > 0$ $(T_\alpha f)_n$ converges to $J_\alpha f(t)$ as $k = t/n \to 0^+$:

Theorem 3.2. *Suppose $f \in C^2[0, \infty)$. Fix $t > 0$, and define $k = t/n$. Then, $|(T_\alpha f)_n - (J_\alpha f)(t)| \to 0$ as $n \to \infty$ for $|\alpha| \leqslant 1$.*

We first introduce some useful lemmas and then prove this theorem. The following is from [3]:

Lemma 3.3. *The m-th term of $c^{(\alpha)}$ has the following asymptotic behavior as $m \to \infty$:*

$$c_m^{(\alpha)} \sim \frac{m^{\alpha-1}}{\Gamma(\alpha)} \left(1 + \frac{\alpha(\alpha-1)}{2m} + O(\frac{1}{m^2})\right), \tag{3.3}$$

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

Lemma 3.4. *For $|\alpha| < 1$, let $A_m = \sum_{i=0}^m c_i^{(\alpha)}$ be the partial sum of $c^{(\alpha)}$ and R be the convolution between $c^{(\alpha)}$ and $(1, 2, \dots)$. Then, as $m \to \infty$, we have:*

$$A_m = \frac{m^\alpha}{\Gamma(1+\alpha)} \left(1 + O(\frac{1}{m})\right),\ R_m = \sum_{i=0}^m (m-i) c_i^{(\alpha)} = \frac{m^{1+\alpha}}{\Gamma(2+\alpha)} \left(1 + O(\frac{1}{m})\right). \tag{3.4}$$

Proof. $\alpha = 0$ is trivial. Suppose $\alpha \neq 0$. $A = \{A_m\}_{m=0}^\infty$ is the convolution between $c^{(\alpha)}$ and $c^{(1)}$ and $A = c^{(\alpha+1)}$ by the group property. Similarly, since $c^{(2)} = (1, 2, 3, \dots)$, $R := \{R_m\}_{m=0}^\infty = c^{(\alpha+2)}$. Applying Lemma 3.3 yields the claims. □

Proof of Theorem 3.2. 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 $\alpha=1$, $(T_\alpha f)_n=\sum_{m=0}^n kf(t-mk)$. It is well known that $|(T_\alpha f)_n-\int_0^t f(s)ds|=O(k)$.

Consider $0<\alpha<1$. Let $n\gg 1$, $1\ll M\ll n$ and $t_M=(M-1)k$. We break the summation for $(T_\alpha f)_n$ at $m=M$ and apply Lemma 3.3 for the terms with $m\geqslant M$:

$$(T_\alpha f)_n = k^\alpha\sum_{m=0}^{M-1} c_m^{(\alpha)} f((n-m)k)+k^\alpha\sum_{m=M}^{n}\frac{m^{\alpha-1}}{\Gamma(\alpha)}f((n-m)k)+O(M^{\alpha-1}k^\alpha).$$

Since $f((n-m)k)=f(t)-f'(\xi)mk$ and $f(t-s)=f(t)-f'(\tilde{\xi})s$, by Lemma 3.4,

$$\begin{aligned}
&\left|k^\alpha\sum_{i=0}^{M-1}c_m^{(\alpha)}f((n-m)k)-\frac{1}{\Gamma(\alpha)}\int_0^{t_M}f(t-s)s^{\alpha-1}ds\right|\\
&\quad\leqslant|f(t)|\left|k^\alpha\sum_{m=0}^{M-1}c_m^{(\alpha)}-\frac{t_M^\alpha}{\Gamma(1+\alpha)}\right|\\
&\qquad+\sup|f'|Mk^{\alpha+1}\sum_{m=0}^{M-1}c_m^{(\alpha)}+C\sup|f'|\int_0^{t_M}s^\alpha ds\\
&\quad\leqslant C(M^{\alpha-1}k^\alpha+M^{1+\alpha}k^{1+\alpha}).
\end{aligned}$$

Finally, by the error for rectangle rule for quadrature,

$$\begin{aligned}
&\left|k^\alpha\sum_{m=K}^{n}\frac{m^{\alpha-1}}{\Gamma(\alpha)}f((n-m)k)-\int_{t_M}^{t}\frac{f(t-s)}{\Gamma(\alpha)}s^{\alpha-1}ds\right|\\
&\quad\leqslant Ck\sup_{s\in(t_M,t)}\frac{d}{ds}(f(t-s)s^{\alpha-1})\leqslant C(Mk)^{\alpha-2}k.
\end{aligned}$$

Choosing $M\sim k^{-1/2}$, we find $M^{\alpha-1}k^\alpha\sim k^{(1+\alpha)/2}$, $(Mk)^{1+\alpha}\sim k^{(1+\alpha)/2}$ and $M^{\alpha-2}k^{\alpha-1}\sim k^{\alpha/2}$. Then, as $k\to 0$,

$$\left|(T_\alpha f)_n-\frac{1}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}f(s)ds\right|\leqslant C(k^{(1+\alpha)/2}+k^{\alpha/2})\to 0.$$

CASE 2 ($-1\leqslant\alpha<0$). If $\alpha=-1$, $c^{(\alpha)}=(1,-1,0,0,\ldots)$. It is then clear that:

$$(T_{-1}f)_n=k^{-1}(f(nk)-f((n-1)k))=f'(nk)+O(k)=J_{-1}f(t)+O(k).$$

Consider that $\alpha\in(-1,0)$ and $\gamma=|\alpha|$. The continuous Riemann-Liouville fraction derivative (3.1) equals

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

where b is chosen such that $b^\gamma = \Gamma(1-\gamma) = -\gamma\Gamma(-\gamma) \geqslant 1$. Since

$$k^{-\gamma} f(t) - k^{-\gamma} f(t-k/b) = O(k^{1-\gamma})$$

and

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

we find

$$\begin{aligned} &|(T_{-\gamma} f)_n - (J_{-\gamma} f)(t)| \\ &\quad \leqslant \left| \frac{1}{k^\gamma} \sum_{i=1}^{n} c_i^{(-\gamma)} f((n-i)k) + \frac{\gamma}{\Gamma(1-\gamma)} \int_{k/b}^{t} \frac{f(t-s)}{s^{\gamma+1}} ds \right| + O(k^{1-\gamma}). \end{aligned} \tag{3.5}$$

We first show that the right hand side of (3.5) goes to zero for constant and linear functions. By the first equation of (3.4) in Lemma 3.4 and noting $b^\gamma = -\gamma\Gamma(-\gamma)$, we have

$$k^{-\gamma} \sum_{i=1}^{n} c_i^{(-\gamma)} = k^{-\gamma} \left(\frac{n^{-\gamma}}{\Gamma(1-\gamma)} - 1 \right) + O\left(\frac{1}{(nk)^\gamma n} \right) = \frac{1}{\Gamma(-\gamma)} \int_{k/b}^{t} \frac{1}{s^{\gamma+1}} ds + O(k). \tag{3.6}$$

Hence, the right hand side of (3.5) goes to zero for constant functions. Similarly, by the second equation of (3.4), $k^{-\gamma} \sum_{i=1}^{n} c_i^{(-\gamma)} (n-i)k - \frac{1}{\Gamma(-\gamma)} \int_{k/b}^{t} \frac{t-s}{s^{\gamma+1}} ds = O((k/b)^{1-\gamma})$, and then

$$\left| k^{-\gamma} \sum_{i=1}^{n} c_i^{(-\gamma)} ik - \frac{1}{\Gamma(-\gamma)} \int_{k/b}^{t} s^{-\gamma} ds \right| = t \times O(k) + O((k/b)^{1-\gamma}) = O(k^{1-\gamma}). \tag{3.7}$$

The right hand side of (3.5) goes to zero for linear functions. Combining (3.6) and (3.7), we can assume without loss of generality that $f(t) = f'(t) = 0$ in equation (3.5) (actually, one can consider the function $\tilde{f}(s) = f(s) - f(t) - f'(t)(s-t)$).

Choose M such that $1 \ll M \ll n$ and set $t_M = (M-1)k$ again.

We first estimate the integral for $s \in (k/b, t_M)$ and the summation from 1 to $M-1$. Since $f(t) = f'(t) = 0$, one has $|f(t-s)| \leqslant Cs^2$, and hence

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

Similarly, since $f(nk) = f'(nk) = 0$ and $c_i^{(-\gamma)}$ is negative for $i \geqslant 1$,

$$\left| k^{-\gamma} \sum_{i=1}^{M-1} c_i^{(-\gamma)} f((n-i)k) \right| \leqslant Ck^{2-\gamma} \sum_{i=1}^{M-1} i^2 |c_i^{-\gamma}| \leqslant CMk^{2-\gamma} \left| \sum_{i=1}^{M-1} ic_i^{(-\gamma)} \right| \leqslant C(Mk)^{2-\gamma}.$$

Note that (3.7) also implies $|\sum_{i=1}^{M-1} ic_i^{(-\gamma)}| = O(M^{1-\gamma})$, which has been used for the last inequality.

Now, we move onto the summation from M to n, and $s \in (t_M, t)$. By Lemma 3.3 and applying the error analysis for rectangle rule of quadrature,

$$\begin{aligned}
&\left| k^{-\gamma} \sum_{i=M}^{n} c_i^{(-\gamma)} f((n-i)k) - \frac{1}{\Gamma(-\gamma)} \int_{t_M}^{t} \frac{f(t-s)}{s^{\gamma+1}} ds \right| \\
&\quad \leqslant \left| k^{-\gamma} \sum_{i=M}^{n} \left(c_i^{(-\gamma)} - \frac{i^{-1-\gamma}}{\Gamma(-\gamma)} \right) f((n-i)k) \right| \\
&\qquad + \left| k^{-\gamma} \sum_{i=M}^{n} \frac{i^{-1-\gamma}}{\Gamma(-\gamma)} f((n-i)k) - \frac{1}{\Gamma(-\gamma)} \int_{t_M}^{t} \frac{f(t-s)}{s^{\gamma+1}} ds \right| \\
&\quad \leqslant C M^{-1-\gamma} k^{-\gamma} + (Mk)^{-2-\gamma} k.
\end{aligned}$$

Taking $M = k^{-\epsilon - \frac{1+\gamma}{2+\gamma}}$ for some small $\epsilon > 0$, $(Mk)^{-2-\gamma}k$, $(Mk)^{2-\gamma}$ and $M^{-1-\gamma}k^{-\gamma}$ all tend to zero as $k \to 0$. Hence, the right hand side of (3.5) goes to zero for all $C^2[0, \infty)$ functions. □

REMARK 3.5. In the case $\alpha = -1$ and $f(0) \neq 0$, $(T_\alpha f)_0 = \frac{f(0)}{k}$. This actually approximates the singular term $\delta(t)f(0)$ in the modified Riemann-Liouville derivative $J_{-1}f$ in [8].

Acknowledgments. The work of J.-G Liu was partially supported by KI-Net NSF RNMS11-07444 and NSF DMS-1514826.

References

[1] W. Cai, *Computational methods for electromagnetic phenomena*, Cambridge University Press, Cambridge, 2013. Electrostatics in solvation, scattering, and electron transport; With a foreword by Weng Cho Chew. MR3027264

[2] B. D. Coleman and W. Noll, *Foundations of linear viscoelasticity*, Rev. Modern Phys. **33** (1961), 239–249, DOI 10.1103/RevModPhys.33.239. MR0158605

[3] P. Flajolet and R. Sedgewick, *Analytic combinatorics*, Cambridge University Press, Cambridge, 2009. MR2483235

[4] R. Gorenflo and F. Mainardi, *Fractional calculus: integral and differential equations of fractional order*, Fractals and fractional calculus in continuum mechanics (Udine, 1996), CISM Courses and Lect., vol. 378, Springer, Vienna, 1997, pp. 223–276. MR1611585

[5] B. G. Hansen and F. W. Steutel, *On moment sequences and infinitely divisible sequences*, J. Math. Anal. Appl. **136** (1988), no. 1, 304–313, DOI 10.1016/0022-247X(88)90133-3. MR972601

[6] A. A. Kilbas, H. M. Srivastava, and J. J. Trujillo, *Theory and applications of fractional differential equations*, North-Holland Mathematics Studies, vol. 204, Elsevier Science B.V., Amsterdam, 2006. MR2218073

[7] R. Kubo, *The fluctuation-dissipation theorem*, Reports on progress in physics **29** (1966), no. 1, 255.

[8] L. Li and J.-G. Liu, *A generalized definition of Caputo derivatives and its application to fractional ODEs*, arXiv Preprint arXiv:1612.05103v2 (2017).

[9] Y. Lin and C. Xu, *Finite difference/spectral approximations for the time-fractional diffusion equation*, J. Comput. Phys. **225** (2007), no. 2, 1533–1552, DOI 10.1016/j.jcp.2007.02.001. MR2349193

[10] J.-G. Liu and R. L. Pego, *On generating functions of Hausdorff moment sequences*, Trans. Amer. Math. Soc. **368** (2016), no. 12, 8499–8518, DOI 10.1090/tran/6618. MR3551579

[11] H. M. Nussenzveig, *Causality and dispersion relations*, Academic Press, New York-London, 1972. Mathematics in Science and Engineering, Vol. 95. MR0503032

[12] G. Del Piero and L. Deseri, *On the concepts of state and free energy in linear viscoelasticity*, Arch. Rational Mech. Anal. **138** (1997), no. 1, 1–35, DOI 10.1007/s002050050035. MR1463802

[13] S. Ruscheweyh, L. Salinas, and T. Sugawa, *Completely monotone sequences and universally prestar-like functions*, Israel J. Math. **171** (2009), 285–304, DOI 10.1007/s11856-009-0050-9. MR2520111
[14] R. L. Schilling, R. Song, and Z. Vondraček, *Bernstein functions*, 2nd ed., De Gruyter Studies in Mathematics, vol. 37, Walter de Gruyter & Co., Berlin, 2012. Theory and applications. MR2978140
[15] O. Stenzel, *The physics of thin film optical spectra*, Springer, 2005.
[16] J. S. Toll, *Causality and the dispersion relation: logical foundations*, Phys. Rev. (2) **104** (1956), 1760–1770. MR0083623
[17] D. V. Widder, *The Laplace Transform*, Princeton Mathematical Series, v. 6, Princeton University Press, Princeton, N. J., 1941. MR0005923
[18] R. Zwanzig, *Nonequilibrium statistical mechanics*, Oxford University Press, New York, 2001. MR2012558