Origin of the fractional derivative and fractional non-Markovian continuous-time processes

P. Van Mieghem [*]

Faculty of Electrical Engineering, Mathematics and Computer Science, P.O Box 5031, 2600 GA Delft, The Netherlands

(Received 12 April 2022; accepted 9 June 2022; published 27 June 2022)

A complex fractional derivative can be derived by formally extending the integer k in the kth derivative of a function, computed via Cauchy's integral, to complex α. This straightforward approach reveals fundamental problems due to inherent nonanalyticity. A consequence is that the complex fractional derivative is not uniquely defined. We explain in detail the anomalies (not closed paths, branch cut jumps) and try to interpret their meaning physically in terms of entropy, friction and deviations from ideal vector fields. Next, we present a class of *non-Markovian* continuous-time processes by replacing the standard derivative by a Caputo fractional derivative in the classical Chapman-Kolmogorov governing equation of a continuous-time Markov process. The fractional derivative leads to a replacement of the set of exponential base functions by a set of Mittag-Leffler functions, but also creates a complicated dependence structure between states. This fractional non-Markovian process may be applied to generalize the Markovian SIS epidemic process on a contact graph to a more realistic setting.

DOI: 10.1103/PhysRevResearch.4.023242

I. INTRODUCTION

Calculus, the theory of differentiation and integration, is a cornerstone in mathematics and nearly all applied sciences (see, e.g., Refs. [1–4]). Its history and far reaching impact, but also the beauty of the theory, is masterly narrated by Strogatz in his recent book *Infinite Powers* [5]. The fathers of calculus are Newton and Leibniz, who reached the insight due to the fine work of Galileo, Kepler, Archimedes, and others just before them. *Infinitesimal* by Alexander [6] sketches the controversial mathematical concept in a historic perspective. Already in Leibniz' time, the meaning of a "fractional" derivative was discussed, from which the field of "fractional calculus" originated, that is reviewed by Metzler and Klafter [7].

By formally extending the scope of the integer k towards a complex number α in Cauchy's integral of the kth derivative of a complex function in (3), we encountered a fundamental difficulty, due to nonanalyticity. Appendix A summarizes the essential properties of an analytic function and its relation to physics. The idea of approaching fractional derivatives from Cauchy's integral (3) already appeared shortly in Olser [8], who referred to a Russian article by Nekrassov in 1888. In a more extensive and partly review paper [9], Olser and his collaborators have followed a similar trajectory as mine, without, however, questioning the implications of *inherent* nonanalyticity. Precisely the nonanalytic character of the α-"fractional" derivative[1] leads to many possible, slightly different definitions, that obviously complicate the field. Section II discusses the consequences of the nonanalytic character of the complex "fractional" derivative and suggests that the inherent nonanalyticity of the fractional derivative operation may be related to an entropy increase in a physical system, due friction, memory effects, and other deviations from ideal vector fields, that are nicely described by analytic functions (Appendix A). Section II C proposes a solution in (11), that is, perhaps, the simplest possible *and*, in addition, consistent with the literature. In particular, confined to real α, the definition (11) reduces to the Caputo fractional derivative [10, Appendix E.2]. We conclude by advocating the view that the noninteger α treatment, representing the nonanalytic, physical world, should be the standard and that the integers $\alpha = k$, representing the analytic, ideal world, are exceptions with greatly simplified and beautiful properties.

A second part aims to generalize continuous-time Markov processes towards α-fractional non-Markovian processes. Section III formally extends the classical governing equation towards an α-fractional setting and provides a general matrix solution in terms the Mittag-Leffler function [11]. We demonstrate that the α-fractional process for $0 < \alpha < 1$ is non-Markovian, possesses a complicated stochastic dependence structure among states over time, but features all transitions between states as in the classical $\alpha = 1$ Markov process, including the final steady state. We also show the remarkable power of the so-called "semigroup" property $P(t+u) = P(t)P(u)$ of the transition probability matrix $P(t)$ of a continous-time Markov chain: it leads to the powerful

*P. F. A.VanMieghem@tudelft.nl

[1] Although a misnomer, we still use the wording "fractional" for historical reasons, due to a discussion about the meaning of the derivative of fractional order $\alpha = 1/2$ in a letter exchange between de L'Hospital and Leibniz in 1695.

Markov property, that makes Markov processes so attractive and tractable. Section IV presents another fractional matrix differential equation that also reduces for $\alpha = 1$ to the same Markov governing equation, but whose embedded Markov chain is different. We conclude in Sec. V by raising several open questions about the physical meaning of the α-fractional extension of the derivative and the probabilistic time-dependence of the α-fractional non-Markovian process. As a major motivation, the α-fractional process constitutes an extension to Markovian SIS epidemics [12] on networks with a tunable parameter α that may account for more realistic infection generation times [13] than the characteristic, exponential Markovian time.

II. COMPLEX FRACTIONAL DERIVATIVE

The kth order derivative in (A7) and (A8) can be written in terms of the Gamma function $\Gamma(z)$, whose properties are summarized in Ref. [11, Appendix A], as

$$\left.\frac{d^k f(z)}{dz^k}\right|_{z=z_0} = \frac{\Gamma(1+k)}{2\pi i}\int_{C(z_0)} \frac{f(\omega)\, d\omega}{(\omega - z_0)^{k+1}}, \tag{1}$$

where $C(z_0)$ is a contour around the point z_0 in a region of the complex plane where the function $f(z)$ is analytic (see Appendix A). The integral (1) is defined for a non-negative integer number $k \geqslant 0$. It is customary [10] in the field of fractional derivatives to use the differential operator D, rather than Leibniz' symbol d/dz or Lagrange's notation $f^{(k)}(z_0)$, thus $\frac{d^k f(z)}{dz^k}|_{z=z_0} = f^{(k)}(z_0) = D^k f(z_0)$, where D^k is the k-fold repetition of D, i.e., $D^k = \underbrace{D.D\ldots D}_{k \text{ times}}$, mainly because the integral operator $I = D^{-1}$ is the inverse operator of the differential operator and formal operator manipulations are often made. It follows directly that $D^k = D^{k-m}D^m$, which also can be deduced from (1). Since $\frac{\Gamma(k)}{(\omega - z_0)^k} f(w)|_{C(z)} = 0$ for[2] $k > 0$, partial integration of the kth derivative in (1) leads to

$$\begin{aligned}\left.\frac{d^k f(z)}{dz^k}\right|_{z=z_0} &= \frac{1}{2\pi i}\int_{C(z_0)} \frac{\Gamma(1+k) f(\omega)\, d\omega}{(\omega - z_0)^{k+1}} \\ &= \frac{1}{2\pi i}\int_{C(z_0)} \frac{\Gamma(k) f'(\omega)\, d\omega}{(\omega - z_0)^{k}}\end{aligned}$$

and repetitions give, for $0 \leqslant m \leqslant k$,

$$\left.\frac{d^k f(z)}{dz^k}\right|_{z=z_0} = \frac{1}{2\pi i}\int_{C(z_0)} \frac{\Gamma(k+1-m) f^{(m)}(\omega)\, d\omega}{(\omega - z_0)^{k+1-m}} \tag{2}$$

The right-hand side in (1) suggests to *formally* extend the scope of the integer number k to the complex number α, which then defines the left-hand side as a *complex fractional derivative*,

$$\begin{aligned}D^\alpha f(z_0) &\equiv \left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=z_0} \\ &= -\frac{1}{2i}\frac{1}{\Gamma(-\alpha)\sin\pi\alpha}\int_{C(z_0)} \frac{f(\omega)\, d\omega}{(\omega - z_0)^{\alpha+1}}\end{aligned} \tag{3}$$

where we have invoked the reflection formula $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin\pi z}$ of the Gamma function, valid for all complex z. The function $\frac{1}{\Gamma(z)}$ is an entire function in the complex variable [14], which means that $\frac{1}{\Gamma(z)}$ is analytic in the entire finite complex plane and, hence, does not possess any singularity, but only zeros at $z = -k$ (where k is a non-negative integer). Entire functions can be regarded as generalizations of polynomials to infinite degree and their theory is created primarily by Karl Weierstrass. The prefactor $\frac{1}{\Gamma(-\alpha)\sin\pi\alpha} = \frac{1}{\pi}\Gamma(\alpha+1)$ has only poles at negative integer values $\alpha = -1, -2, \ldots$, i.e., at the negative zeros of $\sin\pi\alpha$. Unfortunately, there is a fundamental problem in (3) with the extension of the integer k to the complex (or even real) α. Due to the branch cut[3] of the function $z^\alpha = e^{\alpha\ln z}$ along the negative real z axis, the function $\frac{f(\omega)}{(\omega - z_0)^{\alpha+1}}$ is not analytic around the point z_0 and the contour $C(z_0)$ cannot be closed over the branch cut. Although the function z^α reduces to one of the simplest functions when α is an integer, for complex α, the function z^α turns out to be misleadingly complicated.

A. The integral in (3)

We will employ the standard procedure [14] in the theory of complex functions and Mellin transforms to deform the contour around the branch cut. For a complex number $z_0 = r_0 e^{i\varphi}$, where the angle $\varphi = \arg z \in [-\pi, \pi]$, the branch cut of $(\omega - z_0)^\alpha = e^{\alpha\ln(\omega - z_0)} = e^{\alpha\ln|\omega - z_0|}e^{i\alpha\arg(\omega - z_0)}$ lies at an

[2]Partial integration of the Cauchy integral (A1), which corresponds to $k = 0$ in (3), and writing $f'(\omega) = \frac{df(\omega)}{d\omega}$ results in

$$f(z) = \frac{1}{2\pi i} f(\omega)\ln(\omega - z)|_{C(z)} - \int_{C(z)} \ln(\omega - z) f'(\omega)d\omega.$$

If the contour $C(z)$ is a circle $\omega = z + \varepsilon e^{i\theta}$, then

$$\begin{aligned}f(\omega)\ln(\omega - z)|_{C(z)} &= f(z + \varepsilon e^{i2\pi})(\ln(\varepsilon) + 2\pi i) \\ &\quad - f(z+\varepsilon)\ln(\varepsilon) = 2\pi i f(z+\varepsilon)\end{aligned}$$

and

$$\int_{|\omega| = |z + \varepsilon e^{i\theta}|} \ln(\omega - z) f'(\omega)d\omega = f(z+\varepsilon) - f(z),$$

which illustrates the dependence on the radius ε of the circle around z, which is an artifact of the branch cut of $\ln(z)$.

[3]The polar representation of a point $z = re^{i\varphi}$ is not altered when a multiple of $2\pi i$ is added to the phase or argument of z, i.e., $z = re^{i\varphi + 2\pi mi} = re^{i\varphi}$ for each integer m. However, if we take the logarithm of both sides of $z = re^{i(\varphi + 2\pi m)}$,

$$\ln z = \ln r + i(\varphi + 2\pi m)$$

the right-hand side does depend upon the integer m and illustrates the multi-valued nature of the logarithm $\ln z$ and requires to choose a range of length 2π for the argument φ, such as the most common choices $0 \leqslant \varphi \leqslant 2\pi$ or $-\pi \leqslant \varphi \leqslant \pi$. Since $\ln r$ does not exist for negative real r, the logarithm $\ln z$ has the negative real axis as a "forbidden region" in the complex plane, which is called a branch cut, because at the negative real axis, the Riemann sheet that corresponds to a certain integer value of m is cut and folded to interconnect with the $m-1$ and $m+1$ Riemann sheet.

angle $\arg(w - z_0) = \pi$, because the logarithm $\ln \omega$ does not exist for negative real numbers. Thus the branch cut is a line from the point z_0 over the origin $\omega = 0$ towards $\omega \to -\infty e^{i\phi}$. Without loss of generality, let us take the point z_0 equal to a real positive number t, then the branch cut is the real axis from $\omega \to -\infty$ up to $\omega \leqslant t$. In order words, we can always rotate the branch cut to coincide with the negative real axis.

1. Replacing the contour by a nonclosed path

We further replace the contour $C(t)$ in (3) by a nonclosed path $\mathcal{P}(t)$,

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} \overset{?}{\neq} -\frac{1}{2i}\frac{1}{\Gamma(-\alpha)\sin\pi\alpha}\int_{\mathcal{P}(t)}(\omega - t)^{-\alpha-1}f(\omega)d\omega$$

and question at this moment whether equality between left- and right-hand sides is possible. The path $\mathcal{P}(t)$ starts at $\omega = p - i\varepsilon$, where ε is a small positive real number and $p < t$ is a real number, travels below the real ω-axis towards the point where the line meets the circle around point $\omega = t$ with radius ρ, encircles the point $\omega = t$ along the circle $\omega - t = \rho e^{i\theta}$, where the angle varies from $\theta_- = -\pi + \arcsin\frac{\varepsilon}{\rho}$ towards $\theta_+ = \pi - \arcsin\frac{\varepsilon}{\rho}$, and travels back to $\omega = p + i\varepsilon$ now above the real ω axis. Ignoring the direction, the path $\mathcal{P}(t)$ around the point t is thus symmetric with respect to the branch cut, here, the real ω axis. We evaluate the integral along that path $\mathcal{P}(t)$,

$$\begin{aligned}\int_{\mathcal{P}(t)}&(\omega - t)^{-\alpha-1}f(\omega)d\omega\\ &= \int_p^t (x - i\varepsilon - t)^{-\alpha-1}f(x - i\varepsilon)dx\\ &\quad+\int_{\theta_-}^{\theta_+}(\rho e^{i\theta})^{-\alpha-1}f(t+\rho e^{i\theta})d(t+\rho e^{i\theta})\\ &\quad+\int_t^p (x + i\varepsilon - t)^{-\alpha-1}f(x + i\varepsilon)dx.\end{aligned}$$

With $x - i\varepsilon - t = \sqrt{(x-t)^2+\varepsilon^2}\,e^{i(-\pi+\arctan\frac{\varepsilon}{|x-t|})}$, we have

$$\int_p^t (x - i\varepsilon - t)^{-\alpha-1}f(x - i\varepsilon)dx = e^{i(\alpha+1)\pi}\int_p^t((x-t)^2+\varepsilon^2)^{-\frac{\alpha+1}{2}}e^{-i(\alpha+1)\arctan\frac{\varepsilon}{t-x}}f(x-i\varepsilon)dx$$

and similarly for the last integral with $x + i\varepsilon - t = \sqrt{(x-t)^2+\varepsilon^2}\,e^{i(\pi-\arctan\frac{\varepsilon}{|x-t|})}$

$$\int_t^p (x + i\varepsilon - t)^{-\alpha-1}f(x + i\varepsilon)dx = -e^{-i(\alpha+1)\pi}\int_p^t((x-t)^2+\varepsilon^2)^{-\frac{\alpha+1}{2}}e^{i(\alpha+1)\arctan\frac{\varepsilon}{t-x}}f(x+i\varepsilon)dx$$

while[4]

$$\begin{aligned}H_\alpha(\rho) &= \int_{\theta_-}^{\theta_+}(\rho e^{i\theta})^{-\alpha-1}f(t+\rho e^{i\theta})d(t+\rho e^{i\theta})\\ &= i\rho^{-\alpha}\int_{\theta_-}^{\theta_+}e^{-i\alpha\theta}f(t+\rho e^{i\theta})d\theta. \end{aligned}\tag{4}$$

Since the function $f(z)$ is analytic around $z_0 = t$, the right-hand side integral exists. Indeed,

$$\left|\int_{\theta_-}^{\theta_+}e^{-i\alpha\theta}f(t+\rho e^{i\theta})d\theta\right| \leqslant \int_{\theta_-}^{\theta_+}|f(t+\rho e^{i\theta})|d\theta \leqslant 2\pi M_\rho,$$

where $M_\rho = \max_{\theta\in[\theta_-,\theta_+]}|f(t+\rho e^{i\theta})|$ is finite for sufficiently small ρ and, hence,

$$|H_\alpha(\rho)| \leqslant 2\pi\rho^{-\alpha}M_\rho. \tag{5}$$

We can substitute the Taylor expansion $f(t+\rho e^{i\theta}) = \sum_{n=0}^{\infty}f_n(t)(\rho e^{i\theta})^n$ in (A6) and reverse integration and summation (because within its convergence range, a Taylor series can be repeatedly integrated and differentiated [14]),

$$\begin{aligned}\int_{\theta_-}^{\theta_+}e^{-i\alpha\theta}f(t+\rho e^{i\theta})d\theta &= \sum_{n=0}^{\infty}f_n(t)\rho^n\int_{\theta_-}^{\theta_+}e^{i(n-\alpha)\theta}d\theta\\ &= \sum_{n=0}^{\infty}f_n(t)\rho^n\frac{e^{i(n-\alpha)\theta_+}-e^{i(n-\alpha)\theta_-}}{i(n-\alpha)}.\end{aligned}$$

With $\theta_+ = -\theta_- = \pi - \arcsin\frac{\varepsilon}{\rho}$, the integral $H_\alpha(\rho)$ in (4) becomes

$$H_\alpha(\rho) = 2i\sum_{n=0}^{\infty}f_n(t)\rho^{n-\alpha}\frac{\sin\theta_+(n-\alpha)}{n-\alpha}. \tag{6}$$

Another integral representation of $H_\alpha(\rho)$ than (4) is deduced in (B2) in Appendix B. If $\alpha = k$ is an integer, then

$$\begin{aligned}\frac{1}{2i}H_k(\rho) &= \sum_{n=0}^{k-1}f_n(t)\rho^{n-k}\frac{\sin\theta_+(n-k)}{n-k} + \theta_+f_k(t)\\ &\quad+\sum_{n=k+1}^{\infty}f_n(t)\rho^{n-k}\frac{\sin\theta_+(n-k)}{n-k}\end{aligned}$$

and we observe from $\lim_{\varepsilon\to 0}\theta_+ = \pi$ that $\lim_{\varepsilon\to 0}H_k(\rho) = 2\pi i f_k(t)$ for any ρ within the convergence radius of the Taylor series, which again establishes the Cauchy integral (A1).

[4]If $\alpha = 0$ and the circle around t with radius ρ lies in an analytic region of $f(z)$ and is closed, then the Cauchy integral (A1) indicates that $\lim_{\rho\to 0}H_0(\rho) = 2\pi i f(t)$.

2. Tightening the path around the branch cut

We now consider the limit process where we tighten the path $\mathcal{P}(t)$ as close as possible around the branch cut and the point t. Clearly, both ε and ρ must tend simultaneously to zero and $\varepsilon = o(\rho)$, so that $\arcsin \frac{\varepsilon}{\rho} \to 0$, else the lines $\omega = x \pm i\varepsilon$ and the circle $\omega = t + \rho e^{i\theta}$ do not intersect anymore and the path $\mathcal{P}(t)$ is broken. Thus, if $\rho \to 0$ and if ε tends faster to zero than ρ and if $f(t) \neq 0$, then (6) shows that $H_\alpha(\rho) = 2if(t)\rho^{-\alpha}\frac{\sin(\pi - \arcsin\frac{\varepsilon}{\rho})\alpha}{\alpha} + O(\rho^{1-\alpha})$ vanishes provided $\operatorname{Re}\alpha < 0$. Moreover, when the first n higher order derivatives $f_k(t) = \frac{1}{k!}\frac{d^k f(z)}{dz^k}|_{z=t}$ in (6) vanish, i.e., $f(t) = f'(t) = \cdots = f^{(n-1)}(t) = 0$, we observe that the integral $H_\alpha(\rho)$ vanishes provided $\operatorname{Re}(\alpha) < n$, which shifts the line of validity to complex α with positive real part. Similarly as in the proof of Cauchy's integral (A1), it is essential that the integral $H_\alpha(\rho)$ vanishes, which imposes conditions on the validity of a complex fractional derivative.

In the most general setting where only analyticity of $f(z)$ is assumed, we must require that $\operatorname{Re}\alpha < 0$ when $\rho \to 0$ and $\varepsilon = o(\rho)$. Under those conditions, the limit $\varepsilon \to 0$ in the remaining integrals along straight lines reduce to

$$\lim_{\varepsilon\to 0}\int_{\mathcal{P}(t)}(\omega - t)^{-\alpha-1}f(\omega)d\omega$$
$$= (e^{i(\alpha+1)\pi} - e^{-i(\alpha+1)\pi})\int_p^t \left((x-t)^2\right)^{-\frac{\alpha+1}{2}} f(x)dx$$
$$= 2i\sin{(\alpha+1)\pi}\int_p^t \frac{f(x)}{|x-t|^{\alpha+1}}dx. \tag{7}$$

B. Closing the contour

The limit $\varepsilon \to 0$, however, does not close the path $\mathcal{P}(t)$ over the function $(\omega - t)^{-\alpha-1}f(\omega)$ into a contour $C(t)$. When walking around the branch cut in the complex plane from the point $p - i\varepsilon$ to $p + i\varepsilon$ along the path $\mathcal{P}(t)$, we move to an adjacent Riemann sheet. At the point $p + i\varepsilon$, we see the point $p - i\varepsilon$ on the "lower" Riemann sheet as if we are standing on a cliff and see the valley steep below us. The difference in height between the point $p + i\varepsilon$ and the point $p - i\varepsilon$ is called here the jump J. The jump J or discontinuity, when crossing the branch cut from $\omega = p + i\varepsilon$ towards $\omega = p - i\varepsilon$ in order to close the contour again in counter-clockwise sense at the starting point $\omega = p - i\varepsilon$, equals

$$J = \lim_{\varepsilon\to 0}(p + i\varepsilon - t)^{-\alpha-1}f(p+i\varepsilon) - (p - i\varepsilon - t)^{-\alpha-1}f(p - i\varepsilon)$$
$$= -2if(p)\lim_{\varepsilon\to 0}\,((t-p)^2 + \varepsilon^2)^{\frac{-\alpha-1}{2}}\sin\left((\alpha+1)\left(\pi - \arctan\frac{\varepsilon}{t-p}\right)\right)$$

and, with $\sin(\alpha+1)\pi = -\sin\alpha\pi$,

$$J = 2if(p)\frac{\sin\alpha\pi}{(t-p)^{\alpha+1}}. \tag{8}$$

The jump J in (8) only vanishes if α equals any integer k or if $f(p) = 0$. The remarkable fact is that, for integer $\alpha = k$, we move around the branch cut but stay in the same Riemann sheet and do not observe any cliff nor difference between $p + i\varepsilon$ and $p - i\varepsilon$ when $\varepsilon \to 0$, because the branch cut has disappeared as the negative real numbers x to an integer k power, i.e., x^k, are defined!

In order to close the path $\mathcal{P}(t)$ to define a contour $C(t)$, we may propose to add the jump J. Hence, we may argue that

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = -\frac{1}{2i}\frac{1}{\Gamma(-\alpha)\sin\pi\alpha}\int_{C(t)}(\omega - t)^{-\alpha-1}f(\omega)d\omega$$
$$= -\frac{1}{2i}\frac{1}{\Gamma(-\alpha)\sin\pi\alpha}\lim_{\varepsilon\to 0}\int_{\mathcal{P}(t)}(\omega - t)^{-\alpha-1}f(\omega)d\omega - \frac{1}{2i}\frac{J}{\Gamma(-\alpha)\sin\pi\alpha}.$$

Introducing (7) and (8) then yields, for $\operatorname{Re}\alpha < 0$ and $p < t$,

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = \frac{1}{\Gamma(-\alpha)}\int_p^t \frac{f(x)}{(t-x)^{\alpha+1}}dx - \frac{1}{\Gamma(-\alpha)}\frac{f(p)}{(t-p)^{\alpha+1}}. \tag{9}$$

In Appendix B, we demonstrate that closing the contour in a region where the integrand $(\omega - t)^{-\alpha-1}f(\omega)$ is analytic leads to entire cancellation, i.e., $\frac{d^\alpha f(z)}{dz^\alpha}|_{z=t} = 0$, and that we cannot avoid crossing the branch cut to obtain a sensible alternative for contour closure. The demonstration in Appendix B suggests that the jump J in (8) or its contribution $\frac{1}{2i}\frac{J}{\Gamma(-\alpha)\sin\pi\alpha}$ must have some physical meaning; perhaps, similar to a voltage drop over some nonlinear element (e.g., transistor) in an electric network?

Due to the nonanalytic nature of the branch cut, we observe that the right-hand side in (9) is a function of the real number $p < t$. We may thus conclude that the formal extension of the integer k in $\frac{d^k f(z)}{dz^k}|_{z=z_0}$ towards a complex number α does not uniquely define the "complex fractional derivative" $\frac{d^\alpha f(z)}{dz^\alpha}|_{z=t}$, even not subject to negative real α. In other words, there are infinitely many possible definitions that depend upon (a) the real number $p < t$ and (b) the inclusion of a path closure, e.g., via the jump J in (8), or not.

C. Avoiding closure of the integration path

Instead of requiring a closure, which is inspired by Cauchy's integral theorem and analyticity, we may take the other viewpoint that the path $\mathcal{P}(t)$ around the branch cut, not a contour $C(t)$, is the standard property for complex α. In other words, we may argue that the complex fractional derivative is better than (9) defined as

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = \frac{1}{\Gamma(-\alpha)}\int_p^t \frac{f(x)}{(t-x)^{\alpha+1}}dx \quad \text{for } \mathrm{Re}\,(\alpha) < 0. \tag{10}$$

Only for integer $\alpha = k$ analyticity holds, the branch cut disappears, the path $\mathcal{P}(t)$ is closed automatically into a contour $C(t)$ and $\lim_{\alpha\to k} \frac{d^\alpha f(z)}{dz^\alpha}|_{z=t}$ simplifies to the standard kth derivative in (1). Thus, only in the special case of integer values of α, the closed contour is independent of any starting and ending value, in contrast to the noninteger case, where (10) does depend on the point p.

As a matter of fact, the definition of the complex fractional derivative $\frac{d^\alpha f(z)}{dz^\alpha}|_{z=t}$ *must be* inherently nonanalytic and all complications due to this nonanalyticity must be regarded as properties of the complex fractional derivative. A consequence of nonanalyticity is that additional information about the point p, or even the jump J or characteristics of the path $\mathcal{P}(t)$, is needed to arrive at a well-specified definition. Contrary to what we have learned, the complex fractional derivative is the general concept and the ideal integer case should be regarded as an exception.

A final motivation for ignoring the closure of the path $\mathcal{P}(t)$, apart from simplicity, is the consistency with other definitions in the literature, as explained in Sec. II F.

D. Validity range of noninteger α

The general validity condition $\mathrm{Re}(\alpha) < 0$ in (9) and (10) points to a "fractional integral" rather than a "fractional derivative." We have shown that the validity range $\mathrm{Re}(\alpha) < n$ can only be enlarged provided the n higher order derivatives of $f(z)$ at $z = t$ all vanish. However, the analog of (2) is

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = \frac{1}{2\pi i}\int_{\mathcal{P}(t)} \frac{\Gamma(\alpha+1-m)f^{(m)}(\omega)\,d\omega}{(\omega-t)^{\alpha+1-m}} = \frac{1}{\Gamma(m-\alpha)}\int_p^t \frac{f^{(m)}(x)}{(t-x)^{\alpha+1-m}}dx \quad \text{for } \mathrm{Re}\,(\alpha) < m. \tag{11}$$

Obviously, the definition (10) is the special case of (11) for $m = 0$. The more interesting feature of (11) is that it also can deal with complex fractional differentiations, when $0 < \mathrm{Re}(\alpha) < m$. The substitution $u = t - x$ in (11) gives us

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = \frac{1}{\Gamma(m-\alpha)}\int_0^{t-p} u^{m-\alpha-1} f^{(m)}(t-u)du \quad \text{for Re }(\alpha) < m \tag{12}$$

and further motivates to choose the point $p = 0$. The simplest definition of the fractional derivative, that accounts for differentiation, is the case of $m = 1$, real α and $p = 0$ in (11), which we denoted as

$$D_0^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)}\int_0^t \frac{f'(x)}{(t-x)^\alpha}dx \quad \text{for } 0 < \alpha < 1 \tag{13}$$

and which is known as the case $m = 1$ of the Caputo fractional derivative [10, Chapter E.2]. The small validity range of $0 < \alpha < 1$ excludes normal partial integration of the integral in (13).

For complex $\alpha = \sigma + iT$, where $\sigma < m$ and T are real, (12) becomes

$$\Gamma(m-\sigma-iT)\left.\frac{d^{\sigma+iT} f(z)}{dz^{\sigma+iT}}\right|_{z=t} = \int_0^{t-p} u^{m-\sigma-1}e^{-iT\ln u} f^{(m)}(t-u)du,$$

where an oscillatory function $e^{-iT\ln u}$ is added in the integrand and the integral tends to zero if $T\to\infty$. If $f^{(m)}(t-u) \geqslant 0$ for all $x \in [p, t]$, it holds for all real T that

$$\left|\left.\frac{d^{\sigma+iT} f(z)}{dz^{\sigma+iT}}\right|_{z=t}\right| \leqslant \frac{\Gamma(m-\sigma)}{|\Gamma(m-\sigma-iT)|}\left.\frac{d^\sigma f(z)}{dz^\sigma}\right|_{z=t}.$$

Generally with $F = \int_0^{t-p} u^{m-\sigma-1}|f^{(m)}(t-u)|du$, we find a not so tight,[5] though easy upper bound

$$\left|\left.\frac{d^{\sigma+iT} f(z)}{dz^{\sigma+iT}}\right|_{z=t}\right| \leqslant \frac{F}{|\Gamma(m-\sigma-iT)|}.$$

Since $\frac{1}{|\Gamma(b\pm ir)|} = \frac{r^{\frac{1}{2}-b}}{\sqrt{2\pi}}e^{\frac{\pi}{2}r}(1+O(\frac{1}{r}))$, as deduced in Ref. [11, art. 54], the bound becomes

$$\left|\left.\frac{d^{\sigma+iT} f(z)}{dz^{\sigma+iT}}\right|_{z=t}\right| \leqslant \frac{|T|^{\frac{1}{2}-m+\sigma}}{\sqrt{2\pi}}e^{\frac{\pi}{2}|T|}\left(1+O\left(\frac{1}{T}\right)\right)F,$$

which hints that the complex fractional derivative may increase exponentially with the imaginary part $T = \mathrm{Im}(\alpha)$.

E. Alternative representations for the fractional derivative in (10)

In Appendix B, we deduce an alternative integral representation to the common integral (10). Indeed, it follows from (B1) that

$$\frac{1}{\Gamma(-\alpha)}\int_p^t \frac{f(x)}{(t-x)^{\alpha+1}}dx = -\frac{1}{2i}\frac{H_\alpha(|t-p|)}{\Gamma(-\alpha)\sin\pi\alpha}.$$

[5]The oscillatory function $e^{-iT\ln u} = \cos T\ln u + i\sin T\ln u$ is removed, which causes considerable cancellations for increasing T, resulting in a smaller absolute value of the integral.

The integral of $H_\alpha(\rho)$ in (4) indicates, with $\frac{1}{\Gamma(-\alpha)\sin\pi\alpha} = \frac{1}{\pi}\Gamma(\alpha+1)$ and (10), that, for Re $(\alpha)<0$.

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = -\frac{\Gamma(\alpha+1)}{2\pi|t-p|^\alpha}\int_{-\pi}^{\pi} e^{-i\alpha\theta} f(t+|t-p|e^{i\theta})d\theta \tag{14}$$

The integral in (14) is close to a Fourier-type integral, void from branch cut features that now appear in front of the integral. In addition to the integral representation (4) of $H_\alpha(\rho)$, the series in (6) provides a series representation of the α-fractional derivative

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = -\frac{\Gamma(\alpha+1)}{|t-p|^\alpha}\sum_{n=0}^{\infty} f_n(t)|t-p|^n \frac{\sin\pi(n-\alpha)}{\pi(n-\alpha)} \quad \text{for Re}\,(\alpha)<0, \tag{15}$$

where $|t-p|<R$ and R is the radius of convergence of the Taylor series in (A6) around $z_0=t$. Newton's classical equidistant interpolating polynomial (see, e.g., Ref. [15, p. 275–276]) applied to the real derivatives $\{f_k(t)\}_{0\leqslant k\leqslant n}$ gives a *finite* series

$$p_n(\alpha;t) = \frac{\Gamma(\alpha+1)}{n!\Gamma(\alpha-n)}\sum_{j=0}^{n}\binom{n}{j}\frac{(-1)^{n-j}f_j(t)}{\alpha-j}$$

approximation of the fractional derivative $\frac{d^\alpha f(z)}{dz^\alpha}|_{z=t} \approx p_n(\alpha;t)$ for real $\alpha\in[0,n]$.

F. Comparison with other definitions of fractional derivatives

Many trials in history to define a "fractional derivative" are reviewed in Ref. [10, Appendix E] [7, Appendix A]. None of them incorporates an addition as the jump J due to the closure of the path $P(t)$. Perhaps, the most attractive alternative to the kth derivative (1) of the Cauchy integral is the k-fold integral

$$F_k(t,a) = \int_a^t du_1\int_a^{u_1}du_2\ldots\int_a^{u_{k-1}}du_k f(u_k), \tag{16}$$

which can be reduced to a single integral, by repeated partial integration, for $k\geqslant 1$,

$$F_k(t,a) = \frac{1}{(k-1)!}\int_a^t (t-u)^{k-1}f(u)du. \tag{17}$$

The expression (17) was known to Cauchy, but is contributed to Riemann and Liouville [16, p. 43],[10, Appendix E]. If we replace the k-fold integration by the integral operator I^k in (16), then (17) can be written, following the notation in Ref. [10], as

$$F_k(t,a) = \left(I_a^k f\right)(t) = \frac{1}{\Gamma(k)}\int_a^t (t-u)^{k-1}f(u)du.$$

Extension of the scope of the integer k to a real, positive number β yields

$$\left(I_a^\beta f\right)(t) = \frac{1}{\Gamma(\beta)}\int_a^t (t-u)^{\beta-1}f(u)du. \tag{18}$$

In order to avoid the branch cut, we must require that $t\geqslant a$. Comparing the fractional integral (18) to the definition in (10), we replace $\beta\to-\alpha$ and $p\to a$,

$$\left.\frac{d^{-\beta}f(z)}{dz^{-\beta}}\right|_{z=t} = \frac{1}{\Gamma(\beta)}\int_a^t (t-u)^{\beta-1}f(u)du$$

and find agreement.

Based on the semigroup property $D^{\alpha+\beta}=D^\alpha D^\beta$ of the differential operator, Mainardi and Gorenflu [17] discuss the peculiar difference between the Riemann-Liouville fractional derivative, defined for $m-1<\alpha<m$ as

$$D_{RL}^\alpha f(t) = \frac{d^m}{dt^m}\left[\frac{1}{\Gamma(m-\alpha)}\int_0^t \frac{f^{(m)}(x)}{(t-x)^{\alpha+1-m}}dx\right]$$

and the Caputo fractional derivative, which equals (11) for $p=0$ and real α. Both definitions (just as all our definitions) coincide for integer $\alpha=k$ with the common kth derivative. Our analysis shows that the difference between the Caputo and the Riemann-Liouville fractional derivative arises due the nonclosure of path $\mathcal{P}(t)$. The definition (9) of the complex fractional derivative, incorporating the closure of the path by the jump over the branch cut, satisfies $D.D^\alpha = D^{\alpha+1}$. Indeed, differentiating the right-hand of (9) with respect to t yields

$$\begin{aligned}\frac{d}{dt}\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} &= \frac{1}{\Gamma(-\alpha)}\lim_{x\to t}\frac{f(x)}{(t-x)^{\alpha+1}}\\ &\quad-\frac{(\alpha+1)}{\Gamma(-\alpha)}\int_p^t\frac{f(x)}{(t-x)^{\alpha+2}}dx\\ &\quad+\frac{(\alpha+1)}{\Gamma(-\alpha)}\frac{f(p)}{(t-p)^{\alpha+2}}.\end{aligned}$$

The limit only vanishes if $\mathrm{Re}(\alpha)<-1$. In that case, we arrive, using the functional equation $\Gamma(z+1)=z\Gamma(z)$ of the Gamma function, again at (9) with α replaced by $\alpha+1$, demonstrating that $D.D^\alpha = D^{\alpha+1}$ holds.

G. Laplace transform of $\frac{d^\alpha f(z)}{dz^\alpha}|_{z=t}$ in (10)

The single-sided[6] Laplace transform for complex z is defined (see, e.g., Refs. [18,19, Chapter VII], [20]) as

$$\varphi(z) = \int_0^\infty e^{-zt}f(t)dt \tag{19}$$

and in operator form $\varphi(z)=\mathcal{L}_z[f(t)]$, simplified to $\varphi(z)=\mathcal{L}[f(t)]$. The inverse Laplace transform, in operator form $f(t)=\mathcal{L}_t^{-1}[\varphi(z)]$, simplified to $f(t)=\mathcal{L}^{-1}[\varphi(z)]$, is

$$f(t) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}\varphi(z)e^{zt}dz, \tag{20}$$

[6]The double-sided Laplace transform extends the lower integration bound of the single-sided Laplace transform over all negative real values of t and is $\varphi(z)=\int_{-\infty}^{\infty}e^{-zt}f(t)dt$.

where c is the smallest real value of $\mathrm{Re}(z)$ for which the integral in (19) converges.

The Laplace transform of the derivative[7] $\mathcal{L}[f'(t)] = \int_0^\infty e^{-zt} f'(t)dt = z\mathcal{L}[f(t)] - f(0)$. If we replace f by $f^{(k-1)}$, then we find with $y_k = \mathcal{L}[f^{(k)}(t)]$ the recursion $y_k = zy_{k-1} - f^{(k-1)}(0)$, which becomes after l iterations

$$y_k = z^l y_{k-l} - \sum_{j=1}^{l} z^{j-1} f^{(l-j)}(0).$$

Since $y_0 = \mathcal{L}[f(t)]$, choosing $l = k$ yields the general form

$$\mathcal{L}[f^{(k)}(t)] = z^k \mathcal{L}[f(t)] - \sum_{j=1}^{k} z^{j-1} f^{(k-j)}(0).$$

Using the Riemann-Liouville integral in (17) for k-fold integration with $a = 0$ and the important convolution property $\mathcal{L}[(h * g)(t)] = \mathcal{L}[\int_0^t h(t-x)g(x)dx] = \mathcal{L}[h(t)]\mathcal{L}[g(t)]$ results in

$$\mathcal{L}[F_k(t,0)] = \mathcal{L}\left[\frac{1}{(k-1)!}\int_0^t (t-u)^{k-1} f(u)du\right] = \frac{\mathcal{L}[f(t)]}{z^k}$$

illustrating an asymmetry between repeated integration (without function evaluations at $t = 0$) and repeated differentiation with higher-order derivatives at $t = 0$, due to the choice—in our setting—of the point $p = 0$.

Invoking the convolution property and $\mathcal{L}[t^{-\beta}] = \int_0^\infty t^{-\beta} e^{-zt} dt = \frac{\Gamma(1-\beta)}{z^{1-\beta}}$ valid for $\mathrm{Re}(\beta) < 1$, the Laplace transform (19) of the right-hand side of (10) with $p = 0$ is

$$\mathcal{L}\left[\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t}\right] = z^\alpha \mathcal{L}[f(t)] \quad \text{for } \mathrm{Re}\,(\alpha) < 0. \tag{21}$$

The Laplace transform of the generalization (11) with $p = 0$ is

$$\mathcal{L}\left[\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t}\right] = \frac{1}{\Gamma(m-\alpha)}\mathcal{L}[t^{-\alpha-1+m}]\mathcal{L}[f^{(m)}(t)]$$

and, for $\mathrm{Re}(\alpha) < m$

$$\mathcal{L}\left[\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t}\right] = z^\alpha\left\{\mathcal{L}[f(t)] - \sum_{l=0}^{m-1} z^{-l-1} f^{(l)}(0)\right\} \tag{22}$$

where, in contrast to (21), higher-order derivatives at $t = 0$ appear.

H. Physical interpretation

The general fractional derivative (11) as well as the Caputo integral (13) are instances of a convolution $(h * g)(t) = \int_0^t h(t-x)g(x)dx$ of two real functions, where, for the Caputo integral in (13), $h(t) = \frac{t^{-\alpha}}{\Gamma(1-\alpha)}$ and $g(t) = f'(t)$. The concept of a convolution is deeply rooted in integral transforms such as Fourier, Laplace, and Mellin transforms, with many applications in, e.g., probability theory but, perhaps, mainly in signal processing. The integral in (13) weighs the current time t most highly and the past times $x < t$ up to the time $x = 0$ are depreciated by a power law decay $\frac{(t-x)^{-\alpha}}{\Gamma(1-\alpha)}$. Hence, the value of the Caputo fractional derivative (13) for nonintegral $\alpha \in [0, 1]$ is a weighted average over the time interval $[0, t]$. For the integer value of $\alpha = 1$, the behavior is radically different due to the zero of $\frac{1}{\Gamma(1-\alpha)}$ and the divergence of $\int_0^t \frac{f'(x)}{(t-x)^\alpha}dx$ for $\alpha \to 1$. Only the present time t dominates and the entire past at times $x < t$ is washed away. Therefore the Caputo fractional derivative for $\alpha = 1$ is defined as $D_0^1 f(t) = f'(t)$. For the integer value $\alpha = 0$, (13) reduces to $D_0^0 f(t) = f(t) - f(0)$ and an identity is retrieved, provided that $f(0) = 0$. These limit cases explain why the validity range $0 < \alpha < 1$ is mentioned in the Caputo definition (13).

Estrada [21] has provided a physical interpretation of the Caputo fractional derivative (13). He relates the Caputo fractional derivative (13) to memory in a physical process, by interpreting t as time. Estrada's main argument stems from a numerical solution of the integral in (13) due to Odibat [22, Theorem 3], which discretizes the interval $[0, t]$ into a sequel of points $0 = t_0 < t_1 < \cdots < t_m = t$ and approximates the integral as a sum of function evaluations at the points t_j where the integer $0 \leqslant j \leqslant m$. He deduces the same conclusions as the convolution argument above.

III. FRACTIONAL DERIVATIVE FOR CONTINUOUS-TIME MARKOV PROCESSES

We generalize continuous-time Markov processes [23, Chapter 10] to a "fractional setting." The governing Chapman-Kolmogorov equation of a continuous-time Markovian process with N states is

$$\frac{ds(t)}{dt} = -Qs(t), \tag{23}$$

where $-Q$ is the infinitesimal generator in [23, (10.13) after transpose], equal to minus the $N \times N$ weighted Laplacian matrix Q of the underlying Markov graph, in which a node defines a state in the Markov process. We confine the analysis to an infinitesimal generator Q that is fixed and does not depend upon time t. The $N \times 1$ probability state vector[8] is

$$s(t) = (\Pr[X(t) = 1], \Pr[X(t) = 2], \ldots, \Pr[X(t) = N])$$

and $s_j(t) = \Pr[X(t) = j]$ denotes the probability that the Markov process $X(t)$ is in state j at time t. Since the Markov process $X(t)$ at time t must be in one of the N possible states, axiom 1 in probability theory [23, p. 8] states that $u^T s(t) = \sum_{j=1}^N \Pr[X(t) = j] = 1$, where $u = (1, 1, \ldots, 1)$ is the all-one vector. Hence, it follows from the governing equation (23) after left-multiplication by u^T that $u^T Q = 0$, which is the characteristic property of the weighted, possibly asymmetric Laplacian Q.

[7] We assume that $f(t)$ is analytic at $t = 0$, else $f(0)$ should be replaced by $\lim_{t\downarrow 0} f(t) = f(0^+)$, because then the direction of the path towards the point $t = 0$ matters and the left-limit $\lim_{t\uparrow 0} f(t) = f(0^-)$ and right-limit may be different and the function $f(t)$ may not be continuous at $t = 0$.

[8] In contrast to Markov theory [23], where vectors are usually written as $1 \times N$ row vectors, we follow here linear algebra as in Ref. [15] and write $N \times 1$ column vectors.

A. The "fractional α" process

Inspired by Refs. [24–29], we formally extend the governing equation (23) to the Caputo fractional derivative (13) with $0 < \alpha < 1$

$$D_0^\alpha s_\alpha(t) = -Q s_\alpha(t). \tag{24}$$

Georgiou *et al.* [30] have taken a different approach towards "fractional α extension" and have assumed an underlying embedded Markov chain, which is a discrete-time birth-death process [23, Sec. 11.2], where Mittag-Leffler transition times, specified below in (40)–(42), between states are added to create a continuous-time "semi"-Markov process. In contrast to Refs. [30,31], our aim is to avoid incorporating any other assumption such as a semi-Markov property, but just unravel from (24) its underlying physics.

Taking the Laplace transform of both sides in (24) and using (22) yields, for $0 < \alpha < 1$,

$$Q\mathcal{L}[s_\alpha(t)] = z^{\alpha-1}s_\alpha(0) - z^\alpha \mathcal{L}[s_\alpha(t)]. \tag{25}$$

For simplicity, we first proceed by assuming that the weighted Laplacian Q is diagonalizable and the eigenvalue decomposition [15] is

$$Q = \sum_{k=1}^N \mu_k x_k y_k^T,$$

where μ_k with $\mathrm{Re}(\mu_k) \geqslant 0$ is the non-negative eigenvalue belonging to the right eigenvector x_k and the left eigenvector y_k of the weighted Laplacian Q and we assume the ordering $|\mu_1| \geqslant |\mu_2| \geqslant \cdots \geqslant \mu_N = 0$. Then, Eq. (25) becomes

$$z^{\alpha-1}s_\alpha(0) - z^\alpha \mathcal{L}[s_\alpha(t)] = \sum_{k=1}^N \mu_k x_k y_k^T \mathcal{L}[s_\alpha(t)].$$

Left-multiplying both sides by y_m^T and invoking orthogonality [15, art. 140 on p. 213] of the eigenvectors $x_m^T y_k = \delta_{km}$, where δ_{km} is the Kronecker delta, results in

$$\begin{aligned} z^{\alpha-1}y_m^T s_\alpha(0) - z^\alpha y_m^T \mathcal{L}[s_\alpha(t)] &= \sum_{k=1}^N \mu_k \left(y_m^T x_k\right) y_k^T \mathcal{L}[s_\alpha(t)] \\ &= \mu_m y_m^T \mathcal{L}[s_\alpha(t)] \end{aligned}$$

so that

$$y_m^T \mathcal{L}[s_\alpha(t)] = \frac{z^{\alpha-1}}{z^\alpha + \mu_m} y_m^T s_\alpha(0).$$

The "classical" Laplace transform of the Mittag-Leffler function $E_{a,b}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(b+ak)}$ (see, e.g., Ref. [11, Eq. (59)], [10]), which restricts z by $|z^a| > |x|$,

$$\int_0^\infty e^{-zt} t^{b-1} E_{a,b}(xt^a)dt = \frac{z^{a-b}}{z^a - x} \tag{26}$$

indicates that the inverse Laplace transform is

$$\mathcal{L}^{-1}\left[\frac{z^{\alpha-1}}{z^\alpha + \mu_m}\right] = E_{\alpha,1}(-\mu_m t^\alpha) = E_\alpha(-\mu_m t^\alpha),$$

where $E_\alpha(z) = E_{\alpha,1}(z)$ with $E_\alpha(0) = 1$. Inverse Laplace transformation thus leads to

$$y_m^T s_\alpha(t) = y_m^T s_\alpha(0) E_\alpha(-\mu_m t^\alpha).$$

Since any $N \times 1$ vector can be written as a linear combination $s_\alpha(t) = \sum_{m=1}^N (y_m^T s_\alpha(t)) x_m$ of the independent eigenvectors $\{x_m\}_{0 \leqslant m \leqslant N}$ that span the N-dimensional space, we arrive at

$$s_\alpha(t) = \sum_{m=1}^N y_m^T s_\alpha(0) E_\alpha(-\mu_m t^\alpha) x_m. \tag{27}$$

For $\alpha = 1$ with $E_1(z) = e^z$, the solution (27) reduces to the classical vector solution $s(t) = \sum_{m=1}^N (y_m^T s(0)) e^{-\mu_m t} x_m$ with exponential, decaying (and possibly oscillating if $\mu_m \in \mathbb{C}$) functions in time t with rates equal to the eigenvalues μ_m with $\mathrm{Re}(\mu_m) \geqslant 0$ of the weighted Laplacian Q.

The classical solution of (23) for *any* $N \times N$ matrix Q is

$$s(t) = e^{-Qt}s(0),$$

which suggests, for $0 < \alpha < 1$, that the "fractional α process" is described by

$$s_\alpha(t) = E_\alpha(-Qt^\alpha)s_\alpha(0). \tag{28}$$

In Appendix C, we demonstrate the correctness of (28) for any $N \times N$ matrix Q. Hence, the fractional differential equation (24) with general solution (28) holds for *any* infinitesimal generator $-Q$ of a continuous-time Markov process, also if Q is not diagonalizable and may be defective requiring a Jordan form. The Taylor series of $E_\alpha(-Qt^\alpha) = \sum_{k=0}^\infty \frac{Q^k(-t^a)^k}{\Gamma(1+ak)}$ illustrates that the matrices Q and $E_\alpha(-Qt^\alpha)$ commute, i.e., $Q.E_\alpha(-Qt^\alpha) = E_\alpha(-Qt^\alpha).Q$ and commuting matrices possess the same set of eigenvectors [15, p. 213], if the set of eigenvectors is independent. We refer to Garrappa and Popolizio [32] for an efficient numerical computation of the matrix $E_\alpha(-Qt^\alpha)$.

B. Probabilistic interpretations

1. Transition probability matrix $_\alpha P(t)$

The law[9] of total probability [23, p. 23] states that, for $0 \leqslant u$ and $0 \leqslant t$,

$$\Pr[X_\alpha(t+u) = j] = \sum_{l=1}^N \Pr[X_\alpha(t+u) = j | X_\alpha(u) = l] \Pr[X_\alpha(u) = l]. \tag{30}$$

Matrix multiplication of $s_\alpha(t) = E_\alpha(-Qt^\alpha)s_\alpha(0)$ in (28) results in

$$(s_\alpha(t))_j = \Pr[X_\alpha(t) = j] = \sum_{l=1}^N (E_\alpha(-Qt^\alpha))_{jl}(s_\alpha(0))_l.$$

[9]If $\Omega = \cup_k B_k$, where Ω is the sample space and if in addition, for any pair j, k and $j \neq k$, the event sets $B_k \cap B_j = \emptyset$ are not overlapping, then the *law of total probability* is

$$\Pr[A] = \sum_k \Pr[A|B_k] \Pr[B_k]. \tag{29}$$

Comparison of both relations by choosing $u = 0$ in (30) and recalling that $(s_\alpha(t))_j = \Pr[X_\alpha(t) = j]$ shows that

$$\Pr[X_\alpha(t) = j | X_\alpha(0) = l] = (E_\alpha(-Qt^\alpha))_{jl}. \tag{31}$$

Hence, the $N \times N$ transition probability matrix ${}_\alpha P(t)$ with elements $({}_\alpha P(t))_{jk} = \Pr[X_\alpha(t) = j | X_\alpha(0) = k]$ equals

$${}_\alpha P(t) = E_\alpha(-Qt^\alpha)$$

and illustrates that $E_\alpha(-Qt^\alpha)$ is a non-negative matrix and, in particular, a stochastic matrix. The conditional probability (31) indicates that, given that the α-fractional process $X_\alpha(t)$ starts in state l at $t = 0$, then the probability that $X_\alpha(t)$ is in state j at time t equals the (j, l) matrix element of $N \times N$ matrix $E_\alpha(-Qt^a)$.

Application of axiom 1 in probability theory [23, p. 8], $u^T s_\alpha(t) = 1$ for all times t, to $s_\alpha(t) = E_\alpha(-Qt^\alpha)s_\alpha(0)$ in (28) shows that $1 = u^T E_\alpha(-Qt^\alpha)s_\alpha(0)$. Denoting the $N \times 1$ vector $w = (E_\alpha(-Qt^\alpha))^T u$ indicates that $1 = w^T s_\alpha(0)$, which is again an instance of axiom 1 and thus $w = u$. This result directly follows from $Q^T u = 0$ and the Mittag-Leffler Taylor series, because $(E_\alpha(-Qt^\alpha))^T u = Iu + \sum_{k=1}^{\infty} \frac{(Q^k)^T u\, (-t^a)^k}{\Gamma(1+ak)} = u$. Thus we conclude that

$$(E_\alpha(-Qt^\alpha))^T u = u, \tag{32}$$

which is a particular case of the eigenvalue equation $(E_\alpha(-Qt^\alpha))^T y_k = \xi_k y_k$, stating that the all-one vector u is the left eigenvector of the matrix $E_\alpha(-Qt^\alpha)$ belonging to eigenvalue $\xi_\alpha = 1$. The Perron-Frobenius theorem (see, e.g., Ref. [15, p. 235]) for a non-negative matrix then states that the eigenvalue $\xi_\alpha = 1$ is the largest one in absolute value. For any column j of the transition probability matrix ${}_\alpha P(t) = E_\alpha(-Qt^\alpha)$, the eigenvalue equation (32) means that

$$\sum_{k=1}^{N} (E_\alpha(-Qt^\alpha))_{kj} = 1.$$

In other words, the column sum of ${}_\alpha P(t) = E_\alpha(-Qt^\alpha)$ is one at any time t.

2. Dependence on more than only the previous state

Another application of the law of total probability for non-negative real t and u yields

$$\begin{aligned}
&\Pr[X_\alpha(t+u) = j] \\
&= \sum_{k=1}^{N}\sum_{m=1}^{N} \Pr[X_\alpha(t+u) = j | \{X_\alpha(u) = k, X_\alpha(0) = m\}] \\
&\quad \times \Pr[X_\alpha(u) = k, X_\alpha(0) = m].
\end{aligned}$$

Given that the α-fractional process $X_\alpha(t)$ starts in one state l at time $t = 0$, then $\Pr[X_\alpha(0) = m] = \delta_{ml}$ and the above simplifies, with $\Pr[X_\alpha(u) = k, X_\alpha(0) = m] = \Pr[X_\alpha(u) = k | X_\alpha(0) = m] \Pr[X_\alpha(0) = m]$, to

$$\begin{aligned}
&\Pr[X_\alpha(t+u) = j | X_\alpha(0) = l] \\
&= \sum_{k=1}^{N} \Pr[X_\alpha(t+u) = j | \{X_\alpha(u) = k, X_\alpha(0) = l\}] \\
&\quad \times \Pr[X_\alpha(u) = k | X_\alpha(0) = l].
\end{aligned} \tag{33}$$

Introducing (31) in (33) indicates that

$$\begin{aligned}
&(E_\alpha(-Q(t+u)^\alpha))_{jl} \\
&= \sum_{k=1}^{N} \Pr[X_\alpha(t+u) = j | \{X_\alpha(u) = k, X_\alpha(0) = l\}] \\
&\quad \times (E_\alpha(-Qu^\alpha))_{kl},
\end{aligned} \tag{34}$$

which defines the $N \times N$ matrix $E_\alpha(-Q(t+u)^\alpha)$ as a matrix product of the $N \times N$ matrix $E_\alpha(-Qu^\alpha)$ and a tensor of a conditional probabilities that depends on the times $t + u$ and u as well as of the starting time $t = 0$, which manifests for $0 < \alpha < 1$ the intricate dependence among process states at different times. If $u = 0$, then we retrieve (31).

3. Power of the semigroup property in Markov processes

In the classical $\alpha = 1$ Markovian case, the solution (28) reduces to $s(t) = e^{-Qt}s(0)$ and substitution $t \to t + u$ yields

$$s(t+u) = e^{-Q(t+u)}s(0) = e^{-Qt}e^{-Qu}s(0) = e^{-Qt}s(u),$$

where the important "semigroup property" $e^{-Q(t+u)} = e^{-Qt}e^{-Qu}$ is invoked, which is lacking for $E_\alpha(-Q(t+u)^\alpha)$ for $\alpha \neq 1$ as shown in Appendix D. The jth vector component $s_j(t+u)$ is

$$\begin{aligned}
\Pr[X(t+u) = j] &= \sum_{k=1}^{N} (e^{-Qt})_{jk} s_k(u) \\
&= \sum_{k=1}^{N} (e^{-Qt})_{jk} \Pr[X(u) = k].
\end{aligned}$$

Conditioned to the starting state l,

$$\Pr[X(t+u) = j | X(0) = l] = \sum_{k=1}^{N} (e^{-Qt})_{jk} \Pr[X(u) = k | X(0) = l]$$

and comparing with the general result (33) shows that

$$\Pr[X(t+u) = j | \{X(u) = k, X(0) = l\}] = (e^{-Qt})_{jk},$$

while $\Pr[X(t) = j | X(0) = k] = (e^{-Qt})_{jk}$ in (31) indicates that

$$\Pr[X(t+u) = j | \{X(u) = k, X(0) = l\}] = \Pr[X(t) = j | X(0) = k].$$

In other words, the left-hand side does not depend upon (a) the state l, which establishes the famous Markov property that the current state at time $t + u$ only dependent on the previous state at time u of the Markov process and not on earlier states before time u and (b) the variable u, which leads to stationarity $\Pr[X(t+u) = j | X(u) = k] = \Pr[X(t) = j | X(0) = k]$ of the Markov process with time-independent infinitesimal generator Q. Thus the semigroup property defines the Markov process completely in terms the conditional probabilities $\Pr[X(t) = j | X(0) = k]$, i.e., elements of the transition probability matrix $P(t)$. The probability of any other process trajectory $\Pr[X(t_1) = k_1, X(t_2) = k_2, \ldots, X(t_m) = k_m]$ can be reduced (see e.g. [23, p. 180]) by the Markov property as a product of conditional probabilities and the initial probability vector $s(0)$ with kth component $\Pr[X(0) = k]$. We further illustrate the power of the Markov property in (37) in Sec. III E. In conclusion, the semigroup property implies the Markov property, which makes Markov processes so attractive.

Unfortunately, for $0 < \alpha < 1$, we cannot determine $\Pr[X_\alpha(t+u) = j | \{X_\alpha(u) = k, X_\alpha(0) = l\}]$ explicitly, only implicitly via (34). Appendix D shows that there does not exist a one-parameter semigroup property for the Mittag-Leffler function with $\alpha \neq 1$. A consequence is that joint probabilities as $\Pr[X_\alpha(t_1) = k_1, X_\alpha(t_2) = k_2, \ldots, X_\alpha(t_m) = k_m]$ require more knowledge than only the conditional probabilities in the transition probability matrix ${}_\alpha P(t)$. Likely,[10] the governing equation (24) alone is insufficient to completely specify the α-fractional process. Additional information, such as a semi-Markov property as in e.g., Refs. [30,31], seems needed. Beghin and Orsingher [26] have proposed three different fractional Poisson processes that all reduce to the classical Poisson process for $\alpha = 1$.

C. Caputo fractional change in time

The governing equation (24) gives the Caputo fractional change in time of $\Pr[X_\alpha(t) = j]$

$$D_0^\alpha (s_\alpha(t))_j = -\sum_{k=1}^{N} q_{jk} (s_\alpha(t))_k.$$

Conditioned to the event that the α-fractional process $X_\alpha(t)$ started in state l, we have

$$\begin{aligned} D_0^\alpha \Pr[X_\alpha(t) = j | X_\alpha(0) = l] &= -\sum_{k=1}^{N} q_{jk} \Pr[X_\alpha(t) = k | X_\alpha(0) = l] \\ &= -q_{jj} \Pr[X_\alpha(t) = j | X_\alpha(0) = l] \\ &\quad - \sum_{k=1; k\neq j}^{N} q_{jk} \Pr[X_\alpha(t) = k | X_\alpha(0) = l]. \end{aligned}$$

Using $u^T Q = 0$, i.e., $q_{jj} = -\sum_{k=1;k\neq j}^{N} q_{jk} = q_j$, meaning that the self-rate q_j into state j equals the rate, i.e., change in transition probability, of the process out of state j towards any other state k, a balance equation is obtained,

$$\begin{aligned} D_0^\alpha \Pr[X_\alpha(t) = j | X_\alpha(0) = l] &= \sum_{k=1; k\neq j}^{N} q_{jk} (\Pr[X_\alpha(t) = j | X_\alpha(0) = l] \\ &\quad - \Pr[X_\alpha(t) = k | X_\alpha(0) = l]). \end{aligned}$$

The difference in probability for the α-fractional process to be in state j compared to be in state k multiplied by the rate q_{jk} that the process wants to leave state j towards state k defines the Caputo fractional derivative for all fractional strengths α. We observe that the right-hand side equation for the "α fractional process" has the same structure as for the "classical $\alpha = 1$ Markov process," which we further treat in Sec. III F.

[10] If the function $F_{a,b}(x, y)$ in Appendix D were known in analytic closed form, more structure in the sequence of states in the process $X_\alpha(t)$ might be deduced and an analogous α property to the Markov property, which reduces to the Markov property for $\alpha = 1$, might exist.

D. The steady state of the α-fractional process

The right-hand side of the governing equation (24) for the classical $\alpha = 1$ Markov vanishes for the steady-state vector $\pi = \lim_{t\to\infty} s(t)$, which obeys $Q\pi = 0$, implying that π is the right eigenvector of infinitesimal generator Q belonging to the zero eigenvalue $\mu_N = 0$, whereas u is the corresponding left eigenvector. Corresponding to the Perron left eigenvector u with eigenvalue equation (32) is the right eigenvector π that obeys

$$E_\alpha(-Qt^\alpha)\pi = \pi. \tag{35}$$

Assuming diagonalization or an independent set of eigenvectors, orthogonality $x_k^T y_m = \delta_{km}$ then shows that $u^T x_m = 0$ for any $m < N$ and $\pi^T u = 1$. Hence, relation (27) becomes with $u^T s_\alpha(0) = 1$,

$$s_\alpha(t) = \pi + \sum_{m=2}^{N} y_m^T s_\alpha(0) E_\alpha(-\mu_m t^\alpha) x_m \quad \text{for } \alpha > 0$$

and shows that, for large t and $\alpha > 0$,

$$\Pr[X_\alpha(t) = j] = \pi_j + O(E_\alpha(-\mu_{N-1} t^\alpha)).$$

The probability that the α-fractional process $X_\alpha(t)$ is in state j for large time t equals the steady-state probability $\pi_j = \lim_{t\to\infty} \Pr[X(t) = j]$ of the classical $\alpha = 1$ process (and of the embedded Markov chain) and the tendency in time towards the steady probability π_j is given by $E_\alpha(-\mu_{N-1}t^\alpha) \sim \frac{1}{\Gamma(1-\alpha)} \frac{1}{\mu_{N-1}t^\alpha}$ (see, e.g., Ref. [11, art. 31 on p. 35]), where the second smallest eigenvalue μ_{N-1} of the (weighted) Laplacian is known as the algebraic connectivity [15].

The left-hand side of the governing equation (24) shows, invoking the Caputo integral in (13), that

$$0 = \lim_{t\to\infty} D_0^\alpha s_\alpha(t) = \frac{1}{\Gamma(1-\alpha)} \lim_{t\to\infty} \int_0^t \frac{s'_\alpha(x)}{(t-x)^\alpha} dx,$$

which illustrates that it is essential that $s'_\alpha(x) \sim \frac{1}{\Gamma(-\alpha)} \frac{1}{\mu_{N-1} x^{\alpha+1}} \to 0$ faster than $(t-x)^\alpha$ with time $x \to t$ as $t \to \infty$.

E. Sojourn times

The probability that the α-fractional process $X_\alpha(t)$ remains in state j during the interval $[0, t]$ is

$$\Pr[X_\alpha(u) = j, \forall u \in [0, t] | X_\alpha(0) = j].$$

The sojourn time ${}_\alpha\tau_j$ is the (random) time that the process remains in state j. As in Ref. [23, p. 211], we consider for an initial state j, the probability ${}_\alpha H_n$ that the process remains in state j during an interval $[0, t]$. The idea is to first sample the continuous-time interval with step t/n and afterwards proceed to the limit $n \to \infty$, which corresponds to a sampling with infinitesimally small step,

$$ {}_{\alpha}H_n = \Pr\Big[X_\alpha(0) = j, X_\alpha\Big(\frac{t}{n}\Big) = j, \dots, X_\alpha\Big(t - \frac{t}{n}\Big) = j, X_\alpha(t) = j\Big]. $$

Introducing conditional probabilities, i.e., $\Pr[A \cap B] = \Pr[A|B]\Pr[B]$, yields

$$ \begin{aligned} {}_{\alpha}H_n &= \Pr\left[X_\alpha(t) = j | X_\alpha\Big(t - \frac{t}{n}\Big) = j, \dots, X_\alpha\Big(\frac{t}{n}\Big) = j, X_\alpha(0) = j\right] \\ &\quad\times \Pr\left[X_\alpha\Big(t - \frac{t}{n}\Big) = j, \dots, X_\alpha\Big(\frac{t}{n}\Big) = j, X_\alpha(0) = j\right]. \end{aligned} $$

After repetitions, we obtain

$$ \begin{aligned} {}_{\alpha}H_n &= \prod_{m=0}^{n-1} \Pr\left[X_\alpha\left(\frac{(m+1)t}{n}\right) = j \,\middle|\, \cap_{l=0}^{m} X_\alpha\left(\frac{lt}{n}\right) = j\right] \\ &\quad\times \Pr[X_\alpha(0) = j], \end{aligned} \tag{36} $$

which is difficult to evaluate further, without knowledge about the depedence structure of α-fractional process. Therefore most papers cited above have made additional assumptions that confine the α-fractional process.

For the $\alpha = 1$ standard Markov process $X_1(t) = X(t)$, the Markov property

$$ \begin{aligned} &\Pr\left[X\left(\frac{(m+1)t}{n}\right) = j \,\middle|\, \cap_{l=0}^{m} X\left(\frac{lt}{n}\right) = j\right] \\ &= \Pr\left[X\left(\frac{(m+1)t}{n}\right) = j \,\middle|\, X\left(\frac{mt}{n}\right) = j\right] \end{aligned} \tag{37} $$

meaning that the present state only depends upon the previous state, is applied

$$ H_n = \prod_{m=0}^{n-1} \Pr\left[X\left(\frac{(m+1)t}{n}\right) = j \,\middle|\, X\left(\frac{mt}{n}\right) = j\right] \Pr[X(0) = j]. $$

Next, stationarity is invoked, which means that only the time difference between events matter, not the time at which an event occurs,

$$ \Pr\left[X\left(\frac{t}{n} + \frac{mt}{n}\right) = j \,\middle|\, X\left(\frac{mt}{n}\right) = j\right] = \Pr\left[X\left(\frac{t}{n}\right) = j \,\middle|\, X(0) = j\right], $$

resulting in $H_n = [P_{jj}(\frac{t}{n})]^n \Pr[X(0) = j]$, because the transition probability matrix $P(t)$ has element $P_{ij}(t) = \Pr[X(t) = i|X(0) = j]$ and $P(t) = e^{-Qt}$. Invoking the Taylor expansion

$$ \begin{aligned} P_{jj}\left(\frac{t}{n}\right) &= P_{jj}(0) + \left.\frac{dP_{jj}(x)}{dx}\right|_{x=0} \frac{t}{n} + O\left(\left(\frac{t}{n}\right)^2\right) \\ &= 1 - q_{jj}\frac{t}{n} + O\left(\frac{1}{n^2}\right) \end{aligned} $$

results, with the definition $q_j = q_{jj}$, in

$$ H_n = \left(1 - q_j\frac{t}{n} + O\left(\frac{1}{n^2}\right)\right)^n \Pr[X(0) = j]. $$

Using $\lim_{n\to\infty}(1 - q_j\frac{t}{n})^n = e^{-q_j t}$, the probability that the process remains in state j at least for a duration t equals

$$ \begin{aligned} \lim_{n\to\infty} H_n &= \Pr[X(u) = j, \forall u \in [0, t] | X(0) = l] \\ &= e^{-q_j t}\Pr[X(0) = j] \end{aligned} $$

Conditioned to the initial state, we finally arrive at

$$ \Pr[X(u) = j, 0 \leqslant u \leqslant t | X(0) = j] = \Pr[\tau_j > t] = e^{-q_j t}. \tag{38} $$

The Markovian sojourn time theorem [23, p. 210] has been proved: the sojourn times τ_j of a continuous-time Markov process in a state j are independent, exponential random variables with mean $\frac{1}{q_j}$. The reverse of the Theorem also holds, based on exploitation of the memoryless property that is intertwined with the exponential function (see, e.g., Ref. [23, p. 43, 144–145]). The derivation demonstrates that the Markov property (37) does not hold for the α-fractional process $X_\alpha(t)$ with $0 < \alpha < 1$, else we would find exponential sojourn times for all α, which would contradict the Mittag-Lefler solution (28).

It would be tempting to suggest for $0 < \alpha \leqslant 1$ that

$$ \begin{aligned} &\Pr[X_\alpha(r) = j, 0 \leqslant r \leqslant t | X_\alpha(0) = j] \\ &= \Pr[{}_{\alpha}\tau_j > t] = E_\alpha(-q_j t^\alpha) \end{aligned} \tag{39} $$

but we could not prove the guess that the sojourn time ${}_{\alpha}\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ is a scaled Mittag-Leffler random variable M_α. As demonstrated in Ref. [11, art. 41], the Mittag-Leffler random variable M_α is defined by the probability generating function (pgf)

$$ \varphi_{M_\alpha}(s) = E[e^{-sM_\alpha}] = \int_0^\infty e^{-st} f_{M_\alpha}(t)dt = \frac{1}{s^\alpha + 1}. \tag{40} $$

Inverse Laplace transformation determines the probability density function (pdf)

$$ f_{M_\alpha}(t) = t^{\alpha-1}E_{\alpha,\alpha}(-t^\alpha) \quad \text{for } 0 < \alpha < 1 \tag{41} $$

of the random variable M_α with Mittag-Leffler distribution for $0 < \alpha < 1$ equal to

$$ F_{M_\alpha}(t) = \Pr[M_\alpha \leqslant t] = \int_0^t f_{M_\alpha}(u)du = 1 - E_\alpha(-t^\alpha) \tag{42} $$

and with mean $E[M_\alpha] = -\varphi'_{M_\alpha}(0) = \lim_{s\to0}\frac{\alpha s^{\alpha-1}}{(s^\alpha+1)^2} = \infty$. Figure 1 plots the Mittag-Leffler distribution $F_{M_\alpha}(t)$ in (42) for $\alpha \in [0.01, 1]$. In fact, for $0 < \alpha < 1$, the pgf (40) is not analytic at $s = 0$, implying that the Taylor series around $s = 0$ does not exist, nor any derivative. Hence, the Mittag-Leffler

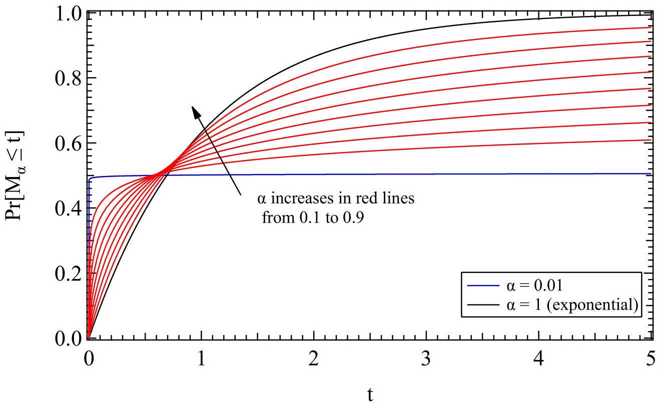

FIG. 1. The Mittag-Leffler distribution (42) for various values of real $\alpha \in [0, 1]$. The limit cases are shown: $\alpha = 1$ (exponential distribution) and $\alpha \to 0$ (constant number $M_0 = 1/2$).

random variable M_α, defined by the pgf (40) and pdf (41) for $0 < \alpha < 1$, does not possess any finite moment $E[M_\alpha^k]$. In the limit $\alpha \to 1$, the Mittag-Leffler random variable M_α becomes an exponential random variable with mean 1, while, for $\alpha = 0$, $M_0 = 1/2$ is not a random variable. In Ref. [11, art. 42], we show that the Mittag-Leffler random variable M_α can be written as product $M_\alpha = RW$ of a stable distribution R and a Weibull distribution W.

Let us return to our claim (39) that the sojourn time in state j is ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$. Hence, the average sojourn time in state j is $E[{}_\alpha\tau_j] = (\frac{1}{q_j})^{\frac{1}{\alpha}} E[M_\alpha]$. For $0 < \alpha < 1$, the factor $(\frac{1}{q_j})^{\frac{1}{\alpha}} > \frac{1}{q_j}$ if $q_i < 1$, otherwise $(\frac{1}{q_j})^{\frac{1}{\alpha}} < \frac{1}{q_j}$ reveals longer sojourn times for low self-rates ($0 \leqslant q_j < 1$) and smaller sojourn times for states with high self-rates $q_j > 1$ than in the corresponding $\alpha = 1$ Markov process, implying that the α-fractional process slows down the slower transition dynamics, but speeds up the faster dynamics. Since the average $E[M_\alpha] \to \infty$ of a Mittag-Leffler random variable diverges for $0 < \alpha < 1$, one may argue, assuming independent sojourn times as in the Markovian sojourn time theorem, that the α-fractional process will stay in a state, on average, infinitely long before jumping to another state, in which the process again will remain infinitely long time on average, and etc. If the sojourn time ${}_\alpha\tau_j \to \infty$ were infinite, the α-fractional process would not change anymore and the corresponding state j can be regarded as an absorbing state, in which the process stays infinitely long. Unless the state j is an absorbing state in the $\alpha = 1$ Markov process, the steady-state analysis in Section III D tells us that the α-fractional process does not generate a different steady-state vector than the right eigenvector π of infinitesimal generator Q. Hence, except for absorbing states ($q_k = 0$), all other sojourn times ${}_\alpha\tau_j$ must be finite, which is physically and intuitively rather obvious. However, it may hint that the claim (39) must be altered by an additional conditioning, ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha 1_{\{M_\alpha < \infty\}}$, requiring that the Mittag-Leffler random variable M_α must be finite. On the other hand, Delvenne *et al.* [33] have shown, for the classical derivative and governing equation (23), that a time-varying weighted Laplacian $Q(t)$ and its topological structure of the underlying Markov graph may result in nonexponential decay times towards the steady state or equilibrium. Thus, even for the classical derivative not every single detail is fully understood. The computation of ${}_\alpha H_n$ in (36) exhibits a complicated process dependence over time and the fractional derivative operator introduces long-term dependence or memory in the dynamics, as explained in Section II H. Hence, we may argue that the α-fractional process may contain an underlying—yet unknown—dependence structure that causes all sojourn times, except for absorbing states, to be finite and that additional conditioning is not needed, but automatically is embedded in the α-fractional process just as the semigroup property in the $\alpha = 1$ Markov process. A final consideration that supports the claim ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ is perturbation theory: for $\alpha < 1$, but $\alpha \uparrow 1$, the α-fractional process tends arbitrarily closely to the $\alpha = 1$ Markov process.

F. The embedded Markov chain does not depend on α

The *embedded* Markov chain of a continuous-time Markov process is the corresponding discrete Markov chain that follows the same state transitions, but that abstracts the time when transitions occur as well as the sojourn times [23, Section 10.4]. The process evolution equation (28) indicates that the transition probability matrix at zero time equals, ${}_\alpha P(0) = E_\alpha(0) = I$, the identity matrix.

Let us denote

$${}_\alpha V_{ji}(t) = \Pr\left[X_\alpha(t) = j | X_\alpha(t) \neq i, X_\alpha(0) = i\right],$$

which describes the probability that, if a transition occurs, the process moves from state i to a different state $j \neq i$. Using the definition of conditional probability,

$$\begin{aligned}
{}_\alpha V_{ji}(t) &= \frac{\Pr\left[\{X_\alpha(t) = j\} \cap \{X_\alpha(t) \neq i\} | X_\alpha(0) = i\right]}{\Pr\left[X_\alpha(t) \neq i | X_\alpha(0) = i\right]} \\
&= \frac{{}_\alpha P_{ji}(t)}{1 - {}_\alpha P_{ii}(t)}.
\end{aligned}$$

In the limit $t \downarrow 0$, we use (31) and obtain[11]

$$\begin{aligned}
\lim_{t\downarrow 0} {}_\alpha V_{ji}(t) &= \lim_{t\downarrow 0} \frac{\Pr\left[X_\alpha(t) = j | X_\alpha(0) = i\right]}{1 - \Pr\left[X_\alpha(t) = i | X_\alpha(0) = i\right]} \\
&= \lim_{t\downarrow 0} \frac{\frac{\Pr[X_\alpha(t) = j | X_\alpha(0) = i]}{t}}{\frac{1 - \Pr[X_\alpha(t) = i | X_\alpha(0) = i]}{t}} \\
&= \lim_{t\downarrow 0} \frac{(E'_\alpha(-Qt^\alpha))_{ji}}{(E'_\alpha(-Qt^\alpha))_{ii}} = \lim_{t\downarrow 0} \frac{\sum_{k=1}^\infty \frac{((-Q)^k)_{ji} t^{\alpha k - 1}}{\Gamma(\alpha k)}}{\sum_{k=1}^\infty \frac{((-Q)^k)_{ii} t^{\alpha k - 1}}{\Gamma(\alpha k)}} \\
&= \lim_{t\downarrow 0} \frac{q_{ji} - \sum_{k=2}^\infty \frac{\Gamma(\alpha)((-Q)^k)_{ji} t^{\alpha(k-1)}}{\Gamma(\alpha k)}}{q_{jj} - \sum_{k=2}^\infty \frac{\Gamma(\alpha)((-Q)^k)_{ii} t^{\alpha(k-1)}}{\Gamma(\alpha k)}} = \frac{q_{ji}}{q_j}.
\end{aligned}$$

Hence,

$$V_{ji} = \lim_{t\downarrow 0} ({}_\alpha V_{ji}(t)) = \frac{q_{ji}}{q_j}$$

[11]The derivative in (C1) indicates $\lim_{t\to 0} s'_\alpha(t) = -\infty$ for $0 < \alpha < 1$ and $\lim_{t\to 0} s'(t) = -Qs(0)$ for the Markovian $\alpha = 1$ case. Only in the fractional regime $0 < \alpha < 1$, the probability vector $s_\alpha(t)$ jumps down, just after the start at time $t = 0$.

does not depend on the fractional strength α! By the definition of q_i due to $u^T Q = 0$, we see that $\sum_{i=1,i\neq j}^{N} V_{ji} = 1$ and thus $V_{ii} = 0$, demonstrating that, given a transition occurs, it is a transition out of state i to another state j. The quantities V_{ji} correspond to the transition probabilities of the embedded Markov chain corresponding to the $\alpha = 1$ continuous Markov chain. In the embedded Markov chain specified by the transition probability matrix V, there are no self-transitions ($V_{ii} = 0$).

G. Summary

We conclude that all transitions in the α-fractional Markov process for $0 < \alpha < 1$ are the same as in the classical $\alpha = 1$ Markov process and exhibited by the embedded Markov chain with transition probability matrix V. Only the time when a transition occurs does depend on α and the dependence of the process $X_\alpha(t)$ over time t is complicated and not Markovian.

The fractional differential equation (24) with general solution (28) constitutes a class of non-Markovian processes, which directly generalizes the Markovian SIS epidemic process [12] on a contact graph with N nodes and a Markov graph on 2^N states to the $0 < \alpha < 1$ regime. In particular, the infection generation time seems well modeled [13] by a Weibull random variable W and the Mittag-Leffler random variable $M = RW$ bears relationship to a Weibull random variable W.

IV. PHYSICAL DIMENSIONS OF QUANTITIES

The matrix function $E_\alpha(-Qt^\alpha)$ in (28) implicitly assumes that the involved quantities, represented by t, α and the matrix Q are dimensionless. If the parameter t represents time and t is dimensionless, then $t = \frac{\tilde{t}}{T}$, where $\tilde{t}$ is an amount of time in units of a time interval T. Clearly, the dimension of $\tilde{t}$ and T is in seconds s, denoted by $[\tilde{t}] = [T] = s$. Since the argument x of a function $f(x)$ must be dimensionless and α is a real dimensionless number, it must hold that $[Q][t^\alpha] = 1$, which is automatically satisfied for dimensionless quantities.

While the claim ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ in (39) may be mathematically correct, it implicitly assumes dimensionless quantities. Explicitly and using tildes for quantities with physical dimension, the dimensionless time is $t = \frac{\tilde{t}}{T}$, the dimensionless sojourn time is ${}_\alpha\tau_j = \frac{{}_\alpha\tilde{\tau}_j}{T}$ and the dimensionless infinitesimal generator is $Q = T\widetilde{Q}$, where the rates $\tilde{q}_{ij}$ have dimension $[\tilde{q}_{ij}] = s^{-1} = [T]^{-1}$. Then, we have

$$E_\alpha(-Qt^\alpha) = E_\alpha(-T^{1-\alpha}\widetilde{Q}\tilde{t}^\alpha)$$

and similarly the sojourn time ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ in dimensionless quantities transforms to ${}_\alpha\tilde{\tau}_j = T^{1-\frac{1}{\alpha}}\ (\frac{1}{\tilde{q}_j})^{\frac{1}{\alpha}} M_\alpha$. We observe for $\alpha \neq 1$ that different powers in the time unit T, in which we measure the time, occur. In other words, the dimensionless form Qt^α and/or ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ disguises the physical time unit T as well as the fact that T appears as a function of the fractional strength α. Alternatively, if we forget dimensionless quantities and regard that the dimension of the sojourn time ${}_\alpha\tau_j$ is in seconds, $[{}_\alpha\tau_j] = s$, the self-rate q_j, which expresses a number of events per unit time, has dimension $[q_j] = s^{-1}$, while the Mittag-Leffler random variable M_α is dimensionless $[M_\alpha] = 1$, then taking the dimension of the claim, $[{}_\alpha\tau_j] = [(\frac{1}{q_j})^{\frac{1}{\alpha}}][M_\alpha]$, leads to an inconsistency for $\alpha \neq 1$.

A. Modifying the fractional differential equation

The inconsistency is removed by replacing in the starting differential (24) the matrix $Q \to Q^\alpha$, which results in a new fractional differential equation

$$D_0^\alpha s_\alpha(t) = -Q^\alpha s_\alpha(t) \tag{43}$$

with solution

$$s_\alpha(t) = E_\alpha(-(Qt)^\alpha)s_\alpha(0).$$

If we now write the dimensionless quantities as scaled quantities with physical dimension as above, we obtain $Qt = \tilde{t}\widetilde{Q}$ and the sojourn time ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ becomes ${}_\alpha\tilde{\tau}_j = T(\frac{1}{T^\alpha(Q^\alpha)_j})^{\frac{1}{\alpha}} M_\alpha = (\frac{1}{(Q^\alpha)_j})^{\frac{1}{\alpha}} M_\alpha$. We observe that the physical time unit T has disappeared. Thus, based on a dimension analysis of physical quantities, we believe that (43) is the "physically" correct governing differential equation.

B. Consequences

All results so far, *but also in all previously published papers that start from (24)*, remain mathematically correct, but are physically less convincing, because of complications with the physical dimensions. In order to pass the physical dimension test, a replacement of the matrix $Q \to Q^\alpha$ is needed, with the corresponding eigenvalues $\mu_k \to \mu_k^\alpha$, but $q_j \to (Q^\alpha)_{jj} \neq q_j^\alpha$, although both have the same dimension $[(Q^\alpha)_{jj}] = [q_j^\alpha]$. If the matrix $Q = \sum_{k=1}^N \mu_k x_k y_k^T$ is diagonalizable and the eigenvalues μ_k are not negative real numbers, then $Q^\alpha = \sum_{k=1}^N \mu_k^\alpha x_k y_k^T$ exists, but is computationally more demanding than Q.

We prefer as claim for the sojourn times the more pleasing form ${}_\alpha\tau_j = \frac{1}{q_j} M_\alpha$ above ${}_\alpha\tilde{\tau}_j = (\frac{1}{(Q^\alpha)_{jj}})^{\frac{1}{\alpha}} M_\alpha$. Expression ${}_\alpha\tau_j = \frac{1}{q_j} M_\alpha$ based on the fractional differential equation (43) is simpler than ${}_\alpha\tau_j = (\frac{1}{q_j})^{\frac{1}{\alpha}} M_\alpha$ deduced from the fractional differential equation (24). Moreover, the form ${}_\alpha\tau_j = \frac{1}{q_j} M_\alpha$ cancels the effect described in Sec. III E of slowing down the slow dynamics and speeding up the faster dynamics (i.e., transitions between Markov states). Another consequence of the fractional differential equation (43) is that the embedded, discrete-time Markov chain discussed in Sec. III F is not anymore the same for all fractional strengths α and the transition probability matrix will depend upon α, i.e. ${}_\alpha V_{ji} = \frac{(Q^\alpha)_{ji}}{(Q^\alpha)_{jj}}$.

Although the fractional extension of continuous-time Markov process, described by (43), converges to the same $\alpha = 1$ Markov process as the fractional differential equation (24), our arguments illustrate that both fractional processes feature a different behavior.

V. CONCLUSIONS

After our generalization of the scope of the integer k to complex α in the kth derivative of the Cauchy integral, we are still puzzled by what the deviation from analyticity means.

Since analyticity relates to simple vector fields in an ideal setting, the fractional derivative in α may account for entropy or friction in a system. Perhaps, the jump J is a measure of the entropy or friction? Also, the way in order to reach the state of a process must depend on many intermediate and past states, which points to the impact of memory in the dynamics process. In that sense, fractional calculus in $\alpha \in (0, 1)$ describes non-Markovian processes, because a Markovian process, whose present state only depends upon the previous one, is essentially memoryless. Indeed, in each state, a Poisson process with certain rate operates and only the transitions to different states require the knowledge of the previous state.

The continuous-time α-fractional extension of the Markov process, by either fractional differential equation (24) or (43) is demonstrated *not* to be a Markov process. While the states now depend upon one another over the entire duration of the process, the extension via (24) follows all Markov transitions between states and the embedded Markov chain is the same, but the fractional extension via (43) possesses a different embedded Markov chain. A dimension analysis indicates that the latter extension via (43) is physically preferable, although both extensions are mathematically correct. The determination of the sojourn time ${}_{\alpha}\tau_j$ in either fractional extension is an open problem and the way how exponential sojourn times [23, Theorem 10.2.3 on p. 210] transfer from the classical $\alpha = 1$ case to the fractional $0 < \alpha < 1$ region. The distribution of the sojourn time ${}_{\alpha}\tau_j$ is needed to construct simulators that precisely mimic the α-fractional non-Markov process.

A main motivation to develop non-Markovian processes is inspired by realistic epidemics, such as Corona, on contact networks. Apart from the human mobility process—often neglected in analyses—that changes the temporal contact graph, the viral infection process does not seem [13] to be close to a Markovian process. The $0 < \alpha < 1$ fractional generalization with tunable parameter α is believed to agree closer with reality than the Markovian ($\alpha = 1$) setting and stands on the agenda of future research.

ACKNOWLEDGMENT

I am grateful to Francesco Mainardi and Enrico Scalas for pointing me to several papers and for interesting discussions. This research has been funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101019718).

APPENDIX A: ANALYTIC FUNCTIONS

The theory of analytic functions can be regarded as an extension of calculus to complex numbers and to the complex plane. Complex numbers arose from the problem in finding the roots of polynomials such as $z^2 + a = 0$ for a real positive number a. During many years, complex numbers have puzzled the great scholars such as Euler, but it was Gauss who created insight and proposed to regard complex numbers in the complex plane. A complex number can be written as $z = x + iy$, where x and y are real numbers and $i = \sqrt{-1}$ is the famous imaginary unit. In the complex plane, we can draw a real axis, on which the number x is specified and an imaginary axis for the real number y, perpendicular to the real axis and the intersection of both axes is called the origin or the point $z = 0$. Perhaps inspired by Gauss, Riemann [34], who was a PhD student of Gauss, and Cauchy established the famous Cauchy integral theorem, that was already known to Gauss.

Let us denote the function $f(z)$ that maps a complex number z to another complex number $\zeta = f(z)$ in the complex plane. If $f(z)$ is an analytic function in some closed region of the complex plane, then Cauchy's integral formula is

$$f(z) = \frac{1}{2\pi i}\int_{C(z)} \frac{f(\omega)}{\omega - z} d\omega \tag{A1}$$

where the integration is over any contour $C(z)$ around the point z that lies in the region where the function is analytic. The contour $C(z)$ is closed, i.e., the begin is the same as the end point, and is a continuous (i.e., piecewise smooth and simple) curve around, not through, the point z and is positively oriented (i.e., counter clockwise sense). In fact, the Cauchy integral (A1) defines the analytic nature of the function $f(z)$ completely [14]. Most likely, the Cauchy integral (A1) was invented or discovered from physics. The vector field, for example created by a electrical charge fixed at a certain point in a large volume, describes the force experienced by another charged particle and has the nice property that the movement of that other charged particle around the fixed charge along a closed curve does not involve any overall labor. Thus, while at each point of the curve, energy is required to move the charged particle to another point of the curve, the sum of the entire energy to move the particle along the closed curve is zero! Of course, the example is a simplification of reality in that we neglect friction, which is nearly always present due to second law of thermodynamics, stating that the degree of disorder or entropy cannot decrease in a closed system. In that sense, analytic functions bear resemblance to certain, simple vector fields in an ideal physical setting.

1. Cauchy's integral theorem

If we denote the complex number $\zeta = f(z) = u + iv$, then u and v are real functions of the real numbers x and y, that satisfy $f(x + iy) = u(x, y) + iv(x, y)$ where the dependence on x and y is expressed explicitly. Taking the derivative with respect to x gives us

$$\frac{\partial f(x+iy)}{\partial x} = \frac{\partial}{\partial x}u(x, y) + i\frac{\partial}{\partial x}v(x, y)$$

and the chain rule tells that $\frac{\partial f(x+iy)}{\partial x} = \frac{\partial f(z)}{\partial z}\frac{\partial z}{\partial x} = f'(z)$. Similarly for the y variable, $\frac{\partial f(x+iy)}{\partial y} = \frac{\partial f(z)}{\partial z}\frac{\partial z}{\partial y} = if'(z)$ and

$$if'(z) = \frac{\partial}{\partial y}u(x, y) + i\frac{\partial}{\partial y}v(x, y)$$

Equating both equations,

$$\frac{\partial}{\partial x}u(x, y) + i\frac{\partial}{\partial x}v(x, y) = \frac{\partial}{\partial y}v(x, y) - i\frac{\partial}{\partial y}u(x, y)$$

and separating the real and imaginary part lead to the Cauchy-Riemann equations [34, art. 4, p. 4]

$$\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}$$
$$\frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}. \tag{A2}$$

Additionally taking the derivative with respect to x and y of each Cauchy-Riemann equation yields the Laplacian partial differential equation,

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0 \tag{A3}$$

as well as $\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}=0$. Thus the both real functions $u(x,y)$ and $v(x,y)$ satisfy the Laplace equation in the continuous real variables x and y and are, therefore, called *harmonic functions*. The Laplace equation is intimately connected to potential fields with simple and nice physical properties (see, e.g., [3,35,36]). An important observation is that the Laplace operator $\Delta=\nabla^2=\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}$ is a linear operator. Linearity seems fundamental in our world: most basic fundamental governing equations, such as Maxwell's laws, the Schrödinger equation, Newton's equations of motion, are linear.

Cauchy's integral theorem for an analytic function $f(z)$ inside and on the contour C is

$$\int_C f(z)dz=0 \tag{A4}$$

and manifests zero labor along a contour in a domain where the function is analytic (i.e., well-behaved). We give Cauchy's proof,[12] because the relation to the physical interpretation is nowadays omitted in the theory of complex functions.

Proof of (A4). As above, we write $f(z)dz=(u(x,y)+iv(x,y))(dx+idy)$ so that

$$\int_C f(z)dz=\int_C u(x,y)dx-v(x,y)dy$$
$$+i\int_C v(x,y)dx+u(x,y)dy$$

Invoking Green's theorem,[13] we find

$$\int_C u(x,y)dx-v(x,y)dy=-\iint_D\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right)dxdy,$$

$$\int_C v(x,y)dx+u(x,y)dy=\iint_D\left(\frac{\partial u}{\partial x}-\frac{\partial v}{\partial y}\right)dxdy.$$

The Cauchy-Riemann equations (A2) then demonstrate (A4). ■

The Cauchy integral formula (A1) follows as a consequence of the Cauchy integral theorem (A4).

Proof of (A1). The arbitrary contour $C(z)$ around the point z encloses a region where the integrand $\frac{f(\omega)}{\omega-z}$ in (A1) is analytic, except for the point at $\omega=z$. We deform the contour $C(z)$ into a new contour $C'(z)$ by adding, just before its closure for example, a line from the contour towards the point z, followed by a circle around the point z with radius ε in clockwise sense and returning in the other direction along the same line back to the contour. Since the deformed contour $C'(z)$ now encloses a region between original contour $C(z)$ and the circle around z that is analytic, the Cauchy integral theorem (A4) states that that contribution is zero. Hence, the evaluation of the integral around the contour $C'(z)$ reduces to the computation of the integral along the path towards the circle and back to the contour $C(z)$, which amounts to a net zero contribution, and the path around the circle

$$g(z)=\int_{|\omega-z|=\varepsilon}\frac{f(\omega)}{\omega-z}d\omega$$
$$=\int_0^{2\pi}\frac{f(z+\varepsilon e^{i\theta})}{z+\varepsilon e^{i\theta}-z}d(z+\varepsilon e^{i\theta})=i\int_0^{2\pi}f(z+\varepsilon e^{i\theta})d\theta.$$

When the radius $\varepsilon\to 0$, only the point z is enclosed and we find that $g(z)=2\pi i f(z)$, which establishes Cauchy integral formula (A1). ■

2. Taylor series of a complex function $f(z)$

A fundamental concept related to analytic functions is the notion of a Taylor series. The power series or Taylor expansion of a complex function $f(z)$ around a point z_0 is defined as

$$f(z)=\sum_{k=0}^{\infty}f_k(z_0)\,(z-z_0)^k \tag{A6}$$

with

$$f_k(z_0)=\frac{1}{k!}\left.\frac{d^k f(z)}{dz^k}\right|_{z=z_0} \tag{A7}$$

$$=\frac{1}{2\pi i}\int_{C(z_0)}\frac{f(\omega)\,d\omega}{(\omega-z_0)^{k+1}} \tag{A8}$$

where the contour $C(z_0)$ encloses z_0 and $f(z)$ is analytic on and inside $C(z_0)$. The integral (A8) follows from the Cauchy integral (A1) by differentiation with respect to z. The theory of complex functions [14,38–40] heavily relies on the Taylor series, because the necessary and sufficient condition that a function should be expansible in a power or Taylor series is that it should be analytic in a region of the complex plane.

APPENDIX B: CLOSING THE CONTOUR BY ANOTHER PATH BACKWARDS

Another way to close the path $P(t)$ is by returning from the point $p+i\varepsilon$ back to the starting point $p-i\varepsilon$ by a path L_p that

[12] A proof without requiring techniques from vector calculus or the continuity of partial derivatives is found in Ref. [14], based on Goursat's arguments.

[13] Green's theorem. Let C be a contour in a plane and let D be the region bounded by the contour C. If u and v are functions of x and y defined on an open region containing D and possessing continuous partial derivatives there, then

$$\oint_C udx+vdy=\iint_D\left(\frac{\partial v}{\partial x}-\frac{\partial u}{\partial y}\right)dxdy. \tag{A5}$$

Green's theorem is the two-dimensional instance of Stokes' more general vector theorem [37]. The integral $\oint_C$ explicitly shows here that the path of integration must be closed and that C must be a contour.

travels back in a region where the integrand $(\omega - t)^{-\alpha-1} f(\omega)$ is analytic. Indeed, Cauchy's integral theorem (A4) then states that the precise shape of the path L_p is not important and that all paths from $p + i\varepsilon$ to $p - i\varepsilon$ give the same integral $\int_{L_p} (\omega - t)^{-\alpha-1} f(\omega) d\omega$. An easy path L_p consists of a circle at t with radius $|t - p|$ and, with the integral $H_\alpha(\rho)$ in (4), we find

$$\int_{L_p} (\omega - t)^{-\alpha-1} f(\omega) d\omega = -H_\alpha(|t - p|).$$

The contribution to the fractional derivative is then

$$-\frac{1}{2i} \frac{1}{\Gamma(-\alpha) \sin \pi\alpha} \int_{L_p} (\omega - t)^{-\alpha-1} f(\omega) d\omega$$
$$= \frac{1}{2i} \frac{H_\alpha(|t - p|)}{\Gamma(-\alpha) \sin \pi\alpha}.$$

An alternative to (9) *may sound* as

$$\left.\frac{d^\alpha f(z)}{dz^\alpha}\right|_{z=t} = \frac{1}{\Gamma(-\alpha)} \int_p^t \frac{f(x)}{(t - x)^{\alpha+1}} dx + \frac{1}{2i} \frac{H_\alpha(|t - p|)}{\Gamma(-\alpha) \sin \pi\alpha}. \tag{B1}$$

However, since the entire contour $C(z)$ now encloses an analytic region, Cauchy's integral theorem (A4) tells us that $\frac{d^\alpha f(z)}{dz^\alpha}|_{z=t} = 0$ in (B1) and that, in addition to the series (6), we find that

$$H_\alpha(|t - p|) = -2i \sin \pi\alpha \int_p^t \frac{f(x)}{(t - x)^{\alpha+1}} dx. \tag{B2}$$

APPENDIX C: PROOF OF (28) FOR ANY $N \times N$ MATRIX Q

Since the Mittag-Leffler function $E_{a,b}(z)$ is an entire function, the Taylor series $E_{a,b}(z) = \sum_{k=0}^\infty \frac{z^k}{\Gamma(b+ak)}$ converges for all finite complex z and the matrix function $E_\alpha(-Qt^a) = \sum_{k=0}^\infty \frac{(-Qt^a)^k}{\Gamma(1+ak)}$ is thus defined for any matrix Q and all times $t \geqslant 0$. Substituting the suggestion (28) into the matrix differential equation $D_0^\alpha s_\alpha(t) = -Qs_\alpha(t)$ in (24) with the explicit form of the Caputo fractional derivative (13) first needs the normal derivative $s'_\alpha(x) = \frac{dE_\alpha(-Qx^\alpha)}{dx} s_\alpha(0)$,

$$s'_\alpha(x) = \sum_{k=1}^\infty \frac{(-Q)^k x^{\alpha k-1}}{\Gamma(\alpha k)} s_\alpha(0). \tag{C1}$$

Using (C1), the Caputo fractional derivative (13) then becomes

$$\frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{s'_\alpha(x)}{(t - x)^\alpha} dx$$
$$= \frac{1}{\Gamma(1-\alpha)} \left(\sum_{k=1}^\infty \frac{(-Q)^k}{\Gamma(\alpha k)} \int_0^t \frac{x^{k\alpha-1}}{(t - x)^\alpha} dx \right) s(0).$$

The integral

$$\int_0^t \frac{x^{k\alpha-1}}{(t - x)^\alpha} dx = t^{-\alpha} \int_0^t \left(1 - \frac{x}{t}\right)^{-\alpha} x^{k\alpha-1} dx$$
$$= t^{-\alpha} t^{k\alpha} \int_0^1 (1 - u)^{-\alpha} u^{k\alpha-1} du$$

is

$$\int_0^t \frac{x^{k\alpha-1}}{(t - x)^\alpha} dx = t^{\alpha(k-1)} \frac{\Gamma(1-\alpha)\Gamma(k\alpha)}{\Gamma(1 + \alpha(k-1))},$$

where we have used the Beta-function integral, valid for $\text{Re}(z) > 0$ and $\text{Re}(q) > 0$,

$$B(z, q) = \frac{\Gamma(z)\Gamma(q)}{\Gamma(z + q)} = \int_0^1 u^{z-1}(1 - u)^{q-1} du.$$

Hence, we have

$$D_0^\alpha s_\alpha(t)$$
$$= \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{s'_\alpha(x)}{(t - x)^\alpha} dx$$
$$= \frac{1}{\Gamma(1-\alpha)} \left(\sum_{k=1}^\infty \frac{(-Q)^k}{\Gamma(\alpha k)} t^{\alpha(k-1)} \frac{\Gamma(1-\alpha)\Gamma(k\alpha)}{\Gamma(1 + \alpha(k-1))} \right) s(0)$$
$$= -Q \sum_{k=1}^\infty \frac{(-Q)^{k-1} t^{\alpha(k-1)}}{\Gamma(1 + \alpha(k-1))} s(0)$$
$$= -Q \sum_{k=0}^\infty \frac{(-Qt^\alpha)^k}{\Gamma(1 + \alpha k)} s(0)$$
$$= -QE_\alpha(-Qt^\alpha)s(0) = -Qs_\alpha(t),$$

which satisfies the fractional differential equation (24). ■

The solution of the general fractional matrix differential equation

$$D_{0;m}^\alpha s_\alpha(t) = -Qs_\alpha(t),$$

where $D_{0;m}^\alpha$ is the m-Caputo fractional derivative with $0 < \alpha < m$, which equals (11) with $p = 0$, and with given initial vectors $\{s_\alpha^{(n)}(0)\}_{0 \leqslant n < m}$ is

$$s_\alpha(t) = \sum_{n=0}^{m-1} t^n E_{\alpha,n+1}(-Qt^\alpha) s_\alpha^{(n)}(0)$$

and can be verified similarly as above.

APPENDIX D: FRACTIONAL SEMIGROUP PROPERTY

A fundamental property of the exponential function is the so-called semigroup property, $e^{x+y} = e^x e^y$. Since $E_{a,b}(z)$ reduces to e^z for $a = b = 1$, the question arises whether there exist a "fractional semigroup property" for the Mittag-Leffler function in the form

$$E_{a,b}(x + y) = F_{a,b}(E_{a,b}(x), E_{a,b}(y)), \tag{D1}$$

where $F_{a,b}(x, y) = F_{a,b}(y, x)$ is a two-variable, symmetric function in the parameters a and b. For $a = 1$ and $b = 1$, we obtain $F_{1,1}(x, y) = xy$. If $x = y = 0$, then $\frac{1}{\Gamma(b)} = F_{a,b}(\frac{1}{\Gamma(b)}, \frac{1}{\Gamma(b)})$. Let $u = E_{a,b}(x)$ and $v = E_{a,b}(y)$, then formally assuming that the inverse function exists, $x = E_{a,b}^{-1}(u)$ and $y = E_{a,b}^{-1}(v)$, the unknown function in (D1) equals

$$F_{a,b}(u, v) = E_{a,b}\big(E_{a,b}^{-1}(u) + E_{a,b}^{-1}(v)\big).$$

Since $E_{a,b}(z)$ is entire function in z, there are regions in the complex plane where the inverse function $E_{a,b}^{-1}(z)$ exists, which leads us to conclude that the two-variable function $F_{a,b}(u, v)$ exists, but we are not capable to specify $F_{a,b}(u, v)$ in closed form. However, not much is known about the inverse function $E_{a,b}^{-1}(z)$ nor about a zero ζ of the Mittag-Leffler function $E_{a,b}(z)$ that satisfies $\zeta = E_{a,b}^{-1}(0)$.

There does *not* exist a one-variable function $G_{a,b}(z)$. For, in that case the two-variable function in (D1) would reduce to a one-variable "fractional semigroup property,"

$$E_{a,b}(x+y) = G_{a,b}(E_{a,b}(x)E_{a,b}(y)),$$

which reduces for $a = b = 1$ to $e^{x+y} = G_{1,1}(e^x e^y)$, so that $G_{1,1}(z) = z$. Differentiation with respect to x,

$$E'_{a,b}(x+y) = E'_{a,b}(x)E_{a,b}(y) \left.\frac{dG_{a,b}(z)}{dz}\right|_{z=E_{a,b}(x)E_{a,b}(y)}$$

and to y

$$E'_{a,b}(x+y) = E'_{a,b}(y)E_{a,b}(x) \left.\frac{dG_{a,b}(z)}{dz}\right|_{z=E_{a,b}(x)E_{a,b}(y)}$$

yields

$$\left.\frac{dG_{a,b}(z)}{dz}\right|_{z=E_{a,b}(x)E_{a,b}(y)} = \frac{E'_{a,b}(x+y)}{E'_{a,b}(x)E_{a,b}(y)} = \frac{E'_{a,b}(x+y)}{E'_{a,b}(y)E_{a,b}(x)},$$

which implies that $E'_{a,b}(x)E_{a,b}(y) = E'_{a,b}(y)E_{a,b}(x)$ or that $\frac{d \ln E_{a,b}(y)}{dy} = \frac{d \ln E_{a,b}(x)}{dx}$ for all x and y. This condition amounts to

$$\frac{d \ln E_{a,b}(x)}{dx} = c$$

and $\ln E_{a,b}(x) = cx + k$, equivalent to $E_{a,b}(x) = \frac{1}{\Gamma(b)}e^{cx}$, which is not correct for all real $a > 0$ and b. Hence, we conclude that there does not exist a one-variable functional form $E_{a,b}(x+y) = G_{a,b}(z)|_{z=E_{a,b}(x)E_{a,b}(y)}$ for $a \neq 1$.

[1] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, 1st ed., Wiley Classic Library Vol. I, (Interscience, New York, 1953), translated and revised from the german original of Springer in 1937 edition.

[2] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, 1st ed., Wiley Classic Library Vol. II, (Interscience, New York, 1953), translated and revised from the german original of Springer in 1937 edition.

[3] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, New York, 1978).

[4] L. M. B. C. Campos, *Complex Analysis with Applications to Flows and Fields*, Mathematics and Physics for Science and Technology (CRC Press, Taylor & Francis Group, Boca Raton, 2011).

[5] S. H. Strogatz, *Infinite Powers: How Calculus Reveals the Secrets of the Universe* (Houghton Mifflin Harcourt, Boston, 2019).

[6] A. Alexander, *Infinitesimal: How a dangerous Mathematical Theory Shaped the Modern World* (Scientific American /Ferrar, Strauss and Geroux, New York, 2014).

[7] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, Phys. Rep. **339**, 1 (2000).

[8] T. J. Osler, Fractional derivatives and Leibniz rule, Am. Math. Mon. **78**, 645 (1971).

[9] J. L. Lavoye, T. J. Osler, and R Tremblay, Fractional derivatives and special functions, SIAM Rev. **18**, 240 (1976).

[10] R. Gorenflo, A. A. Kilbas, F. Mainardi, and S. V. Rogosin, *Mittag-Leffler Functions, Related Topics and Applications*, 2nd ed. (Springer, Berlin, 2020).

[11] P. Van Mieghem, The Mittag-Leffler function, arXiv:2005.13330.

[12] P. Van Mieghem, J. Omic, and R. E. Kooij, Virus spread in networks, IEEE/ACM Transactions on Networking **17**, 1 (2009).

[13] P. Van Mieghem and Q. Liu, Explicit non-Markovian SIS mean-field epidemic threshold for Weibull and Gamma infections but Poisson curings, Phys. Rev. E **100**, 022317 (2019).

[14] E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, Amen House, London, 1964).

[15] P. Van Mieghem, *Graph Spectra for Complex Networks* (Cambridge University Press, Cambridge, UK, 2011).

[16] G. F. D. Duff and D. Naylor, *Differential Equations of Applied Mathematics* (John Wiley & Sons, Hoboken, 1966).

[17] F. Mainardi and R. Gorenflo, Time-fractional derivatives in relaxation processes: a tutorial survey, Fract. Calc. Appl. Anal. **10**, 269 (2007).

[18] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, 2nd ed. (Oxford University Press, Ely House, London W. I, 1948).

[19] M. A. Evgrafov, *Analytic Functions* (W. B. Saunders Company, 1966; Reprinted by Dover Publications, Inc., New York, dover 2019 edition, 2019).

[20] D. V. Widder, *The Laplace Transform* (Princeton University Press, Princeton, 1946).

[21] E. Estrada, Communicability in time-varying networks with memory, New J. Phys. **24**, 063017 (2022).

[22] Z. Odibat, Approximations of fractional integrals and Caputo fractional derivates, Appl. Math. Comput. **178**, 527 (2006).

[23] P. Van Mieghem, *Performance Analysis of Complex Networks and Systems* (Cambridge University Press, Cambridge, UK, 2014).

[24] F. Mainardi, R. Gorenflo, and E. Scalas, A fractional generalization of the Poisson processes, Vietnam Journal of Mathematics **32**, 53 (2004).

[25] A. Mura, M. S. Taqqu, and F. Mainardi, Non-markovian diffusion equations and processes: Analysis and simulations, Physica A **387**, 5033 (2008).

[26] L. Beghin and E. Orsingher, Fractional Poisson processes and related planar random motions, Electron. J. Probab. **14**, 1790 (2009).

[27] S. de Nigris, A. Hastir, and R. Lambiotte, Burstiness and fractional diffusion on complex networks, Eur. Phys. J. B **89**, 114 (2016).

[28] L. Abadias, G. Estrada-Rodriguez, and E. Estrada, Fractional-order susceptible-infected model: definition and applications to the study of Covid-19 main protease, Fract. Calc. Appl. Anal. **23**, 635 (2020).

[29] F. Diaz-Diaz and E. Estrada, Time and space generalized diffusion equation on graph/networks,, Chaos, Solitons and Fractals **156**, 111791 (2022).

[30] N. Georgiou, I. Z. Kiss, and E. Scalas, Solvable non-Markovian dynamic network, Phys. Rev. E **92**, 042801 (2015).

[31] N. Georgiou and E. Scalas, Bounds for mixing times for finite semi-markov processes with heavy-tail jump distribution, Fract. Calc. Appl. Anal. **25**, 229 (2022)
[32] R. Garrappa and M. Popolizio, Computing the matrix Mittag-Leffler function with applications to fractional calculus, J. Sci. Comput. **77**, 129 (2018).
[33] J.-C. Delvenne, R. Lambiotte, and L. E. C. Rocha, Diffusion on networked systems is a question of time or structure, Nat. Commun. **6**, 7366 (2015).
[34] B. Riemann, Grundlagen für eine allgemeine Theorie der Functionen einer veränderlichen complexen Grösse, PhD. thesis of 1851. Adalbert Rente, Göttingen, 1867.
[35] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Massachusetts, 1963), Vol. 1.
[36] R. P. Feynman, R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics* (Addison-Wesley, Massachusetts, 1963), Vol. 2.
[37] R. C. Buck, *Advanced Calculus* (McGraw-Hill Book Company, New York, 1978).
[38] E. T. Whittaker and G. N. Watson, *A Course of Modern Analysis* (Cambridge University Press, Cambridge, UK, 1996).
[39] G. Sansone and J. Gerretsen, *Lectures on the Theory of Functions of a Complex Variable* (P. Noordhoff, Groningen, 1960), Vols. 1 and 2.
[40] A. I. Markushevich, *Theory of Functions of a Complex Variable* (Chelsea Publishing Company, New York, 1985), Vols. I–III.