

Article

A New Look at the Initial Condition Problem

Manuel D. Ortigueira 

Centre of Technology and Systems—UNINOVA and Department of Electrical Engineering,
NOVA School of Science and Technology of NOVA University of Lisbon, Quinta da Torre,
2829-516 Caparica, Portugal; mdo@fct.unl.pt

Abstract: In this paper, some myths associated to the initial condition problem are studied and demystified. It is shown that the initial conditions provided by the one-sided Laplace transform are not those required for Riemann-Liouville and Caputo derivatives. The problem is studied and solved with generality as well as applied to continuous-time fractional autoregressive-moving average systems.

Keywords: initial conditions; fractional derivative; linear systems; Riemann-Liouville derivative; Caputo derivative; FARMA

MSC: 26A33

1. Introduction

The initial condition (IC) problem, often, gives rise to great discussion, mainly because there are several misconceptions on the subject. We are going to try to clarify the situation, to distinguish IC problem from others, such as extrapolation or prediction.

The Heaviside operational procedure [1–3], gathering the IC problem, motivated the search for its analytical justification [4]. Among the many attempts, it is very interesting to highlight Bromwich's approach that was based on the inverse Laplace transform, considered, at the time, as a formulation of the Laplace transform (LT) in the complex plane [5]. It was not until 1926 that P. Lévy showed that what is now called the Bromwich integral was really the inverse LT. Bromwich devised a procedure to insert the required IC [5]. Parallely, Carson presented a similar development, but based on the slightly modified LT, the Laplace–Carson transform (it reappeared recently with another name) [6–8]. Van der Pol [9] presented a methodology that synthesized the Bromwich and Carson procedures, having as a base the two-sided (bilateral) LT. They not only gave theoretical justification for Heaviside's method, but also introduced a simple way of inserting the IC. Regardless of these approaches, Doetsch departed from the unilateral (U) LT and presented a coherent mathematical development [10]. His book won the preference of most researchers, and his methodology continues to be followed by many mathematicians and, even, physicists and engineers, who are unaware of bilateral (B) LT. According to the derivative property of the ULT, the transform of a derivative (integer order) depends on terms involving the derivatives of the function at hand taken at $t = 0^+$. This has given rise to many discussion papers by electrical engineers that called attention to the fact that the IC comes from the past, not from the future. Consequently, a modification in the LT was required [11–13]. With that, the ULT became like a standard for introducing the IC. However, in Fractional Calculus (FC) applications, the Doetsch procedure remains in use, leading to two well-known expressions for the Riemann-Liouville (RL) and Caputo (C) derivatives [14–17]. Several authors noted the dissatisfaction caused by this use and proposed alternatives [18–23].

In this paper, we make a revision of the IC concept and introduce a clear distinction between the IC problem and another, very similar one, which is frequently confused with it [20,23,24]. To this end, we call attention to certain myths and misconceptions that we



Citation: Ortigueira, M.D. A New Look at the Initial Condition Problem. *Mathematics* **2022**, *10*, 1771. <https://doi.org/10.3390/math10101771>

Academic Editor: Yuri Luchko

Received: 23 April 2022

Accepted: 20 May 2022

Published: 23 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

often find in the FC literature and that are related to forgetting the past, a contradiction given the characteristic associated with fractional operators: having a large memory. This is, to remark, that the “beginning” is not at $t = 0$. This is merely a reference instant. Among the myths, we pay special attention to the question of the initial conditions associated with the derivatives of RL and C. We will show that the natural IC required by such derivatives are not those introduced by the one-sided LT [25]. These are incoherent.

To introduce a clarification of the problem at hand, we make a small digression through the notion of a system and a few of its properties. We define relaxed system and its associated IC, which are the basis for our main problem. We consider the alternative 0^+ vs. 0^- , for defining the IC. To solve the IC problem, we start with the continuous-time integer order autoregressive-moving average (ARMA) systems, highlighting the role of the “jump formula”, namely in avoiding discontinuities. For solving the problem for the fractional (F) ARMA systems, we use the fractional jump formula we introduced earlier [22,25]. To verify the correctness and coherence of the resulting formulae, we highlight the consistency with respect to a state-variable formulation, namely the “observable canonical form”. Through a well-known problem, we made a distinction between the IC and the extrapolation/prediction problems.

The paper is outlined as follows. A brief introduction to FC is done in Section 2. In Section 3, some myths and contradictions in FC are considered. In particular, we study the IC associated with the RL and C derivatives (Section 3.2). The IC problem is clearly formulated and solved for FARMA systems, through the use of a fractional jump formula in Section 4. In Section 4.6, we establish a difference between the IC and extrapolation/prediction problems. Finally, we present some conclusions (Section 5).

Remark 1.

- Let $\alpha > 0$. The analytic function defined by

$$E_\alpha(z) = \sum_{n=0}^{\infty} \frac{z^n}{\Gamma(n\alpha + 1)}, \quad z \in \mathbb{C},$$

is the Mittag-Leffler function (MLF). A particular case, very important in FC, is what we will call causal MLF (CMLF) given by

$$\mathcal{E}_\alpha(t, a) = \sum_{n=0}^{\infty} a^n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} u(t) \tag{1}$$

where $u(t)$ is the Heaviside unit step.

- The bilateral Laplace transform is defined by

$$\mathcal{L}f(t) = \int_{-\infty}^{\infty} f(t)e^{-st}dt, \quad s \in \mathbb{C}. \tag{2}$$

2. On the Fractional Derivatives

Systems described by fractional differential equations are becoming increasingly adopted in the 21st century, for modelling many natural and man-made phenomena [26–34]. In fact, we are currently dealing with phenomena that require a modelling beyond traditional tools. Fractional derivatives give us the possibility to mathematically express new forms of behaviour that are difficult to model with integer order derivatives. However, the success of fractional calculus has not been achieved peacefully, since its inception. In fact, from its first conception by Liouville (1832) [35,36], many problems arose and prevented its immediate acceptance. Liouville could not find a simple way of expressing a function in terms of exponentials that were the basis for their findings, since, at that time, the inverse Laplace transform was unknown. Anyway, the main definitions we find today are based on the formulae presented by Liouville, mainly the Riemann–Liouville [14], (Gerasimov, Dzherbashian)-Caputo [15,37], and Grünwald–Letnikov [14] definitions. How-

ever, and based on these derivatives, new ones have been proposed alongside these, such as Hadamard’s [15], Marchaud’s [14], or Hilfer’s [17]. Consequently, the number of currently existing fractional derivatives (FDs) is very high, which has become the biggest obstacle for the diffusion of FC in science and engineering. Trying to introduce a systematisation in the field, Oliveira and Machado, first, and Teodoro et al., more recently [38,39], listed many operators and introduced a classification, according to some specified criteria. However, several among the described operators cannot be considered as FD in agreement to the appropriate criteria proposed in [40]. This implies, first of all, that we establish a clear distinction between FD and other associated operators. A FD is a generalization of the classic derivative for any real (or complex) order (we will treat only constant order cases [41]). Given an FD of order $\alpha > 0$ there may be its right inverse that we call anti-derivative. This is important, to avoid confusion with the designation “fractional integral”, which may be applied to many operators that cannot be considered derivatives, even of negative orders. Anyway, the number of FD is high enough to create difficulties for those who intend to make applications in science or engineering. Recently, a new way at looking into the difficulty was proposed in [42]. Basically, we classify the derivatives according to:

1. Shift-invariant (unified fractional derivative) [43]

(a) Causal/anti-causal

- Grünwald–Letnikov [14,44]

$$D_+^\alpha f(t) := \lim_{h \rightarrow 0^+} h^{-\alpha} \sum_{n=0}^{+\infty} \frac{(-\alpha)_n}{n!} f(t - nh). \tag{3}$$

where $(-\alpha)_n$ is the Pochhammer representation for the raising factorial:

$$(-a)_n = \prod_{k=0}^{n-1} (-a + k),$$

with $(-a)_0 = 1$.

- Liouville [14,44]

$$D_+^\alpha f(t) = \frac{1}{\Gamma(N - \alpha)} \frac{d^N}{dt^N} \int_0^\infty f(t - \tau) \tau^{N-\alpha-1} d\tau \tag{4}$$

with $N - 1 < \alpha \leq N$.

- Liouville–Caputo [33,44]

$$D_+^\alpha f(t) = \frac{1}{\Gamma(N - \alpha)} \int_{-\infty}^t f^{(N)}(\tau) (t - \tau)^{N-\alpha-1} d\tau \tag{5}$$

with $N - 1 < \alpha \leq N$.

(b) Bilateral

$$D_\theta^\alpha f(t) := \lim_{h \rightarrow 0^+} h^{-\alpha} \sum_{n=-\infty}^{+\infty} (-1)^n \frac{\Gamma(\alpha + 1)}{\Gamma\left(\frac{\alpha+\theta}{2} - n + 1\right) \Gamma\left(\frac{\alpha-\theta}{2} + n + 1\right)} f(t - nh), \tag{6}$$

where α is the derivative order and θ the asymmetry parameter.

- Riesz potential/derivative ($\theta = 0$) [43]
- Feller potential/derivative ($\theta = 1$) [43]

2. Scale-invariant

- Hadamard [14]
- Quantum [45]

Here, we are interested in studying causal systems, hence defined by causal derivatives. The most popular are Riemann-Liouville (RL) and Caputo, (C) [16,17,33,44] which are particular cases of Liouville (L) and Liouville–Caputo (LC), obtained for functions defined in intervals, $[a, b] \in \mathbb{R}$. They are what we can name two-step derivatives, which are usually expressed by

$${}^{RL}D_{a+}^{\alpha} f(t) = \frac{1}{\Gamma(N - \alpha)} \frac{d^N}{dt^N} \int_a^t f(\tau)(t - \tau)^{N-\alpha-1} d\tau, \quad t > a \tag{7}$$

and

$${}^C D_{a+}^{\alpha} f(t) = \frac{1}{\Gamma(N - \alpha)} \int_a^t f^{(N)}(\tau)(t - \tau)^{N-\alpha-1} d\tau, \quad t > a \tag{8}$$

respectively.

Multiple-step derivatives can be defined as combinations of RL and C. This is the case of the Davidson and Essex, Cavanati, and Hilfer derivatives [42].

Any of these derivatives can be used to define differential equations, linear or nonlinear, and the corresponding general solutions depend on some IC. In the derivative definitions, there is nothing that tells us which IC are required. This depends, surely, on the structure of the system at hand [22]. It is important to note that, from the bilateral Laplace transform (BLT) [25] point of view, all the causal derivatives above referred to are equivalent and verify

$$\mathcal{L}D_{+}^{\alpha} f(t) = s^{\alpha} \mathcal{L}f(t), \quad \text{Re}(s) > 0. \tag{9}$$

where \mathcal{L} refers to the BLT. No IC appear.

3. Some Myths and Contradictions of Fractional Calculus

3.1. *Methaphysique Derivatives*

Many historical introductions to fractional calculus attribute the origin of the idea of the derivative of non-integer order to l'Hôpital, assuming that, in a missing letter to Leibniz, they would ask “what if $n = 1/2$?” This would motivate Leibniz to reply, in a letter dated 30 September 1695: It will be an apparent paradox, from which one day useful consequences will be drawn. However, as S. Dugowson [46] points out, there is no evidence that such a letter ever existed. Furthermore, he attributes the origin of the idea exactly to Leibniz himself in a letter to J. Bernoulli, sent on 28 February 1695, where the designation “metaphysical derivatives” appears, for the first time. Furthermore, they continued the discussion in subsequent letters. Liouville [35], also, attributed the idea to Leibniz and presented, for the first time, some derived definitions. One of these was, incorrectly, called the Caputo derivative.

The idea of “methaphysique derivatives” was recovered by S. Dugowson, who titled his Ph.D. thesis “Les différentielles métaphysiques”, due to the appearance of some “abnormal” results. In fact, and for the RL case, there is a well-known formula for the derivative of the power function [14]:

$${}^{RL}D^{\alpha} [t^{\beta}] = \frac{\Gamma(\beta + 1)}{\Gamma(\beta - \alpha + 1)} t^{\beta-\alpha}, \quad t > 0, \tag{10}$$

where $\beta > -1$. With $\alpha = 1/2$, we obtain

$$\frac{d^{1/2}}{dt^{1/2}} [t^{-1/2}] = 0, \quad \frac{d^{1/2}}{dt^{1/2}} [1] = \frac{1}{\sqrt{\pi t}}, \quad \frac{d^{1/2}}{dt^{1/2}} [t] = \sqrt{\frac{t}{\pi}} \tag{11}$$

These “strange” results led B. West [47] to write “These three fractional derivatives alert us to the fact that we have entered into a world in which the rules for quantitative analysis are different from what we have always believed, but they are not arbitrary. It remains to be seen if this mathematical world can explain the complexity of the physical, biological and social worlds in which we live”.

It is interesting to note how many people accept the above results, without questioning them, which means that they are entering a really different world, with its own rules that can go against the usual laws. Furthermore, B. West adds “of course, neither of these curious findings is consistent with the ordinary calculus and has to do with the nonlocal nature of fractional derivatives”. In fact, these results are not coherent with the ordinary calculus, since they express a partial view, like a torn photo. This is the result of forgetting the nonlocal characteristics of the FC, mainly from the past, from which nothing is said. Such a fact is very strange, since we know that the fractional derivatives are nonlocal. Therefore, *the past is important*.

To obtain the complete picture, consider that we take a fractional derivative defined on \mathbb{R} [48]. To be coherent, we can use the forward Liouville derivative, from which the RL is a particular case. Let $u(t)$ be the Heaviside function and rewrite the Equation (10) as

$${}^L D^\alpha [t^\beta u(t)] = \frac{\Gamma(\beta + 1)}{\Gamma(\beta - \alpha + 1)} t^{\beta - \alpha} u(t), \quad t \in \mathbb{R}, \tag{12}$$

with the help of distribution theory, the validity of this formula can be extended for any $\alpha, \beta \in \mathbb{R}$ [44]. It is not a difficult task to show that

$$D^\alpha u(t) = \frac{t^{-\alpha}}{\Gamma(1 - \alpha)} u(t) \tag{13}$$

is valid for any real α and for Grünwald–Letnikov, Liouville, and Liouville–Caputo derivatives [44,48]. Anyway, if we let $\delta(t) = Du(t) = \frac{t^{-1}}{\Gamma(0)}$ [49] and consider the above-used pair, $\alpha = -\beta = \frac{1}{2}$, we obtain

$$\begin{aligned} \frac{d^{1/2}}{dt^{1/2}} [t^{-1/2} u(t)] &= \frac{\Gamma(\frac{1}{2})}{\Gamma(0)} t^{-1} = \sqrt{\pi} \delta(t) \\ \frac{d^{1/2}}{dt^{1/2}} u(t) &= \frac{1}{\sqrt{\pi t}} u(t) \\ \frac{d^{1/2}}{dt^{1/2}} [tu(t)] &= \sqrt{\frac{t}{\pi}} u(t) \end{aligned} \tag{14}$$

These relations show the complete picture cut in Equation (10).

It can be shown that the RL and C derivatives of the Heaviside unit step are as follows

1. Riemann-Liouville derivative

$${}^{RL} D_+^\alpha u(t) = \begin{cases} \frac{1}{\Gamma(1 - \alpha)} \frac{d}{dt} \int_{0^+}^t u(\tau) (t - \tau)^{-\alpha} d\tau & t \geq 0 \\ \text{undefined} & t < 0 \end{cases} = \begin{cases} \frac{t^{-\alpha}}{\Gamma(1 - \alpha)} & t \geq 0 \\ \text{undefined} & t < 0 \end{cases}$$

2. Caputo derivative

This case must be studied with care. If the integration starts at 0^+ , as usually done, $u'(t) = 0$

$${}^C D_+^\alpha u(t) = \begin{cases} \frac{1}{\Gamma(1 - \alpha)} \int_{0^+}^t \frac{du(\tau)}{d\tau} (t - \tau)^{-\alpha} d\tau & t \geq 0 \\ \text{undefined} & t < 0 \end{cases} = \begin{cases} 0 & t \geq 0 \\ \text{undefined} & t < 0 \end{cases}$$

The derivative of the unit step being zero is a negative result that was used in [50], to show that the Caputo derivative is useless for modelling circuits with fractional capacitors, since the results are contradicted by laboratory experiments.

This has a connection with a phrase we find frequently, the “Caputo (C) derivative has the advantage that its derivative of a constant is zero, while the Riemann-Liouville derivative (RL) is not zero”, see [33,51,52]

$${}^C D^\alpha [1] = 0 \tag{15}$$

together with Equation (11). Therefore, and as seen above, there is great confusion between the constant function and the Heaviside unit step: they share the future, but have different pasts.

3.2. RL and C Initial Conditions

3.2.1. Incoherences

The solution of the linear constant coefficient, defined in terms of RL or C derivatives, is, frequently, obtained by using ULT that are stated as [53,54]

$$\mathcal{L} \left[{}^{RL} D^\alpha f(t) \right] = s^\alpha F(s) - \sum_{k=0}^{N-1} s^k D^{\alpha-k-1} f(0^+) \tag{16}$$

and

$$\mathcal{L} \left[{}^C D^\alpha f(t) \right] = s^\alpha F(s) - \sum_{k=0}^{N-1} s^{\alpha-k-1} D^k f(0^+), \tag{17}$$

where $N = \lceil \alpha \rceil$. These relations describe the usual approach, to introduce the IC to solve differential equations. Since, the IC associated with the C derivative are expressed as integer order derivatives, this has been the reason for the preference given to it. However, these relations lead to several contradictions.

To understand what is at stake, consider a linear system defined by the following differential equation [25]:

$$D^{2\alpha} f(t) + aD^\alpha f(t) + bf(t) = 0, \quad t > 0, \tag{18}$$

where $a, b \in \mathbb{R}$ and $0 < \alpha < 1$. We want to compute the output $f(t)$ for $t > 0$, using the unilateral Laplace transform (ULT) with the RL or C derivatives. In both cases, we have to treat two different cases, corresponding to $\alpha \leq 1/2$ and to $\alpha > 1/2$. For solving the Equation (18), the following IC are involved

$$\begin{aligned} \text{ULT - RL} : & f^{(\alpha-1)}(0^+) && \text{if } \alpha \leq \frac{1}{2} \\ \text{ULT - RL} : & f^{(\alpha-1)}(0^+), f^{(2\alpha-1)}(0^+), f^{(2\alpha-2)}(0^+) && \text{if } \frac{1}{2} < \alpha \leq 1 \\ \text{ULT - C} : & f(0^+) && \text{if } \alpha \leq \frac{1}{2} \\ \text{ULT - C} : & f(0^+), f'(0^+) && \text{if } \frac{1}{2} < \alpha \leq 1. \end{aligned} \tag{19}$$

For another comparison, let us introduce two state variables $v_1(t) = f(t)$ and $v_2(t) = D^\alpha f(t)$, so that Equation (18), with $g(t) = 0$, can be rewritten as

$$\begin{bmatrix} D^\alpha v_1(t) \\ D^\alpha v_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}, \quad t > 0. \tag{20}$$

Let $v = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}$. To solve state Equation (20), using the above procedure, we only need the following IC

$$\begin{aligned} \text{ULT - RL} : & v^{(\alpha-1)}(0^+) \\ \text{ULT - C} : & v(0^+) \end{aligned} \quad 0 < \alpha \leq 1,$$

respectively, or, attending to the definition of the vector v , the following IC are required.

$$\begin{aligned}
 ULT - RL &: f^{(\alpha-1)}(0^+), f^{(2\alpha-1)}(0^+) \quad \text{if } 0 < \alpha \leq 1 \\
 ULT - C &: f(0^+), f^{(\alpha)}(0^+) \quad \text{if } 0 < \alpha \leq 1.
 \end{aligned}
 \tag{21}$$

As observed, there is an evident contradiction between Equations (19) and (21).

3.2.2. Outfit Results

Consider the differential equation

$$D^\alpha f(t) + Af(t) = 0 \quad A \in \mathbb{R}, t > 0, \tag{22}$$

where D^α represents either RL or C derivatives. We are going to obtain a modified version, so that it accommodates the IC.

Theorem 1. *Let $\alpha > 0$. The solution of the equation*

$${}^{RL}D^\alpha f(t) - f(0)\frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} + Af(t) = 0 \quad A \in \mathbb{R}, \tag{23}$$

is given aside a constant by the CMLF, $\mathcal{E}_\alpha(t)$,

$$f(t) = f(0)\mathcal{E}_\alpha(t, -A) = f(0) \sum_{n=0}^{\infty} (-1)^n A^n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} u(t). \tag{24}$$

Proof. Firstly, assume that the solution of Equation (22) has the form:

$$f(t) = \sum_{n=0}^{\infty} a_n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} u(t) \tag{25}$$

where the series is uniformly convergent in $[0, t]$. Using Equation (10)

$${}^{RL}D^\alpha \left[\frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} \right] = \frac{t^{(n-1)\alpha}}{\Gamma((n-1)\alpha + 1)}, \quad t > 0,$$

so that,

$${}^{RL}D^\alpha f(t) = f(0) \sum_{n=0}^{\infty} a_{n+1} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} u(t) + a_0 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} u(t) \tag{26}$$

Substituting Equations (25) and (26) into Equation (22) and setting $a_0 = f(0)$, we get

$$a_{n+1} = -Aa_n, \quad n = 0, 1, 2, \dots$$

which leads to Equation (24). A non-compensated term, $a_0 \frac{t^{-\alpha}}{\Gamma(-\alpha+1)} u(t)$, appears. It has to be removed. \square

Therefore,

- The CMLF solves Equation (22) for the RL derivative,
- The natural IC is $f(0)$ which originates the appearance of the term ${}^{RL}D^\alpha u(t) = f(0) \frac{t^{-\alpha}}{\Gamma(-\alpha+1)} u(t)$ in contradiction with Equation (16),
- Relation Equation (23) can be written as

$${}^{RL}D^\alpha [f(t) - f(0)u(t)] + Af(t) = 0 \quad t \geq 0, \tag{27}$$

which is very interesting and will be reconsidered later.

Now, let us repeat the reasoning for the C derivative. We have:

Theorem 2. Let $\alpha > 0$. The solution of the equation

$${}^C D^\alpha f(t) + Af(t) = 0 \quad A \in \mathbb{R}, t \geq 0, \tag{28}$$

is given by the function

$$f(t) = a_0 + \sum_{n=1}^{\infty} (-1)^n A^n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)}, \quad t \geq 0. \tag{29}$$

where a_0 is an undetermined constant.

Proof. The proof is similar to the followed above. Using the derivation rules of the power function, we get

$${}^C D^\alpha \sum_{n=0}^{\infty} a_n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} = \sum_{n=0}^{\infty} a_{n+1} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)}, \quad t > 0$$

which, inserted in Equation (28), leads to

$$a_{n+1} = -Aa_n, \quad n = 0, 1, 2, \dots$$

but a_0 is indetermined. It can be set equal to $f(0)$, but any other value can be used. This result enters in contradiction with Equation (17). \square

In this case, the non-compensated term does not appear (it is zero), but this originates an indeterminacy.

To continue, we return back to equation

$$D^{2\alpha} f(t) + aD^\alpha f(t) + bf(t) = 0, \quad t > 0,$$

and try to solve it using the procedure above used. Therefore, let $\alpha \leq 1$ and assume that $f(t)$ is, again, given by a series, as in Equation (25), and attend to Equation (26) to obtain

$${}^{RL} D^\alpha f(t) = \sum_{n=0}^{\infty} \gamma_{n+1} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} + \gamma_0 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)}, \quad t > 0,$$

and

$${}^{RL} D^{2\alpha} f(t) = \sum_{n=0}^{\infty} \gamma_{n+2} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} + \gamma_0 \frac{t^{-2\alpha}}{\Gamma(-2\alpha + 1)} + \gamma_1 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} \quad t > 0,$$

that, when inserted into the equation, lead to

$$\begin{aligned} & \sum_{n=0}^{\infty} \gamma_{n+2} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} + \gamma_0 \frac{t^{-2\alpha}}{\Gamma(-2\alpha + 1)} + \gamma_1 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} + \\ & a \sum_{n=0}^{\infty} \gamma_{n+1} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} + a\gamma_0 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} + b \sum_{n=0}^{\infty} \gamma_n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} = 0, \quad t > 0 \end{aligned}$$

If we remove the negative power terms (that are a function of the IC), we deduce that

$$\gamma_{n+2} = -a\gamma_{n+1} - b\gamma_n, \quad n = 0, 1, 2, \dots$$

needs the IC γ_0 and γ_1 , which are the coefficients of the negative power terms. Therefore, Equation (18) is transformed into

$$D^{2\alpha} f(t) - \gamma_0 \frac{t^{-2\alpha}}{\Gamma(-2\alpha + 1)} - \gamma_1 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} + aD^\alpha f(t) - a\gamma_0 \frac{t^{-\alpha}}{\Gamma(-\alpha + 1)} + bf(t) = g(t), \quad t > 0,$$

that can be rewritten as

$$\left[D^{2\alpha} f(t) - \gamma_0 D^{2\alpha} u(t) - \gamma_1 D^{2\alpha} \frac{t^\alpha}{\Gamma(\alpha + 1)} \right] + a[D^\alpha f(t) - \gamma_0 D^\alpha u(t)] + b f(t) = g(t), \quad t > 0,$$

leading to

$$D^{2\alpha} \left[f(t) - \gamma_0 u(t) - \gamma_1 \frac{t^\alpha u(t)}{\Gamma(\alpha + 1)} \right] + a[D^\alpha f(t) - \gamma_0 u(t)] + b f(t) = g(t), \quad t > 0. \quad (30)$$

For the C derivative, it is a simple task to see that

$${}^C D^\alpha f(t) = \sum_{n=0}^{\infty} \gamma_{n+1} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)}, \quad t > 0,$$

and

$${}^C D^{2\alpha} f(t) = \sum_{n=0}^{\infty} \gamma_{n+2} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)}, \quad t > 0,$$

that, when inserted into the equation, give

$$\sum_{n=0}^{\infty} \gamma_{n+2} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} + a \sum_{n=0}^{\infty} \gamma_{n+1} \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} + b \sum_{n=0}^{\infty} \gamma_n \frac{t^{n\alpha}}{\Gamma(n\alpha + 1)} = 0, \quad t > 0,$$

and again

$$\gamma_{n+2} = -a\gamma_{n+1} - b\gamma_n, \quad n = 0, 1, 2, \dots$$

while letting γ_0 and γ_1 be indeterminated. However, setting $\gamma_0 = f(0)$ and $\gamma_1 = D^\alpha f(0)$ seems to be the natural solution.

We conclude that:

1. We need two IC, independently of $\alpha > 1/2$ or not.
2. Instead of the IC used in the previous sub-section, we need $f(0)$ and $D^\alpha f(0)$, for both RL and C derivatives.

Both these facts contradict the usual procedures.

4. Redefining the Problem

4.1. Systems and Differential Equations

The designation *system* is widely used in different areas of science and engineering, although not always with the same meaning. In systems theory, we designate, by system, an entity that performs a given *task*. Its nature can be diverse—a machine or a company, a biological structure or a computer program, a circuit (of any fluid) or a communication system, etc. A system can, also, be a combination of these entities, interacting with each other to perform its *objective*. Traditionally, such entities were physical objects that could have a mathematical representation, called a model. For example, the differential equation

$$\frac{y(t)}{dt} + ay(t) = x(t), \quad t \in \mathbb{R} \quad (31)$$

is the model for a lowpass RC circuit or for the speed of a ball on a pool table. However, in the last 50 years, circumstances have changed, due to the action of signal processing and the spread of microprocessors, first, and computers, later, which allowed the implementation, in real time, of many mathematical algorithms, leading to the replacement of hardware-based systems by software-based equivalents. Therefore, the model itself became a system, and today we use the designations “model” and “system” interchangeably, when referring to the mathematical representation or its computational implementation. Next, we will use “system” with this expanded meaning. Systems react to input (stimulus) actions by providing outputs, according to their goals. Both input and output mathematical

representations are functions that we call *signals*. Next, we will assume that our signals are piecewise continuous bounded functions, defined in \mathbb{R} .

Definition 1. A system is defined, mathematically, as an application in the set of signals, that is, a transformation of a signal, $x(t)$, into another one, $y(t)$.

Let $T[.]$ be an operator that symbolically represents such a transformation, then

$$y(t) = T[x(t)] \tag{32}$$

$x(t)$ is the input or excitation and $y(t)$ is the output or response.

If a system does not produce an output before the application of an input, we say that it is a *causal (not anticipatory) system*. Therefore, the output of a causal system depends on previous inputs and outputs (memory) as well as actual input. If a system depends on future, instead of past, memories, it is called *anti-causal*.

Definition 2. A system is at rest or relaxed, in a given (non empty) time interval, if it is static, meaning that it has no dynamic behaviour: both input and output are null.

It is, what we can call, a switched-off system. However, this does not mean that its internal components are empty: we may have accumulated “energy” that manifests itself in the output, when we turn on the system. These are the initial conditions. For example: a closed tank with water, a charged capacitor inserted in an open circuit, a compressed spring, a suspended body, and so on. These systems have non-null IC, which manifests their existence, when we restart (switch-on) using the systems.

Definition 3. We define the initial conditions of a relaxed system as the effects of past inputs and outputs accumulated in the system, which originate a corresponding output, when the system is activated.

When the system is not relaxed, the associated problem is different: it is an extrapolation or prediction problem, in the sense pointed out by Kolmogorov [55] and Wiener [56]. Later, we will return to the subject.

4.2. 0^+ or 0^- ?

Most interesting systems are defined by (fractional) differential equations. In the following, we shall be concerned with linear systems defined by the fractional autoregressive-moving average model [57]. The equations used in the above section are particular cases. The standard definition of derivative is

$$Df(t) = f'(t) = \lim_{h \rightarrow 0} \frac{f(t) - f(t - h)}{h}, \tag{33}$$

or

$$Df(t) = f'(t) = \lim_{h \rightarrow 0} \frac{f(t + h) - f(t)}{h}. \tag{34}$$

where we assume $h \in \mathbb{R}^+$. As it is obvious, the first is causal, contrarily to the second, which is anti-causal. Consider the differential equation

$$\frac{dy(t)}{dt} + ay(t) = \frac{dx(t)}{dt}$$

that we intend to solve numerically, using Equation (33), by removing the limit computation but using a small h . Assume that the input is $x(nh)$, $n = 0, 1, \dots$. We have, successively, for $t = 0, h, 2h \dots$

$$\begin{aligned}
 y(0)[1 + ah] - y(0 - h) &= x(0) - x(0 - h) \rightarrow \mathbf{y(0)} = \frac{\mathbf{y(0 - h)} - \mathbf{x(0 - h)}}{\mathbf{1 + ah}} + \frac{x(0)}{1 + ah} \\
 y(h)[1 + ah] - y(0) &= x(h) - x(0) \rightarrow \mathbf{y(h)} = \frac{\mathbf{y(0 - h)} - \mathbf{x(0 - h)}}{(\mathbf{1 + ah})^2} + \\
 &\quad \frac{x(0)}{(1 + ah)^2} + \frac{x(h) - x(0)}{1 + ah} \\
 y(2h)[1 + ah] - y(h) &= x(2h) - x(h) \rightarrow \mathbf{y(2h)} = \frac{\mathbf{y(0 - h)} - \mathbf{x(0 - h)}}{(\mathbf{1 + ah})^3} + \\
 &\quad \frac{x(0)}{(1 + ah)^3} + \frac{x(h) - x(0)}{(1 + ah)^2} + \frac{x(2h) - x(h)}{1 + ah}
 \end{aligned}$$

For a generic instant nh , a term dependent on $x(-h)$ and $y(-h)$ appears. These are the IC and show clearly that they must be taken in the past. When h decreases to 0, such IC read $x(0^-)$ and $y(0^-)$. Now, repeat the procedure with the derivative Equation (34). Assume now that the input is $x(nh)$, $n = 0, -1, -2 \dots$. We have, successively, for $t = 0, -h, -2h \dots$

$$\begin{aligned}
 y(0)[ah - 1] + y(0 + h) &= x(0 + h) - x(0) \rightarrow \mathbf{y(0)} = \frac{\mathbf{x(0 + h)} - \mathbf{y(0 + h)}}{\mathbf{ah - 1}} - \frac{x(0)}{ah - 1} \\
 y(-h)[ah - 1] + y(0) &= x(-h) - x(0) \rightarrow \mathbf{y(-h)} = -\frac{\mathbf{x(0)} - \mathbf{y(0)}}{(\mathbf{ah - 1})^2} - \frac{x(-h)}{(ah - 1)} - ah \frac{x(0)}{(ah - 1)^2}
 \end{aligned}$$

It is no use to continue, since the situation is similar to the previous one. Now, we observe that the IC are $x(h)$ and $y(h)$, and when h decreases to 0, such IC read $x(0^+)$ and $y(0^+)$. This procedure can be repeated for a system involving other derivatives. Therefore, the IC of causal systems must be taken at $t = 0^-$. This means that most results obtained with $t = 0^+$ may be incorrect.

4.3. Another Look at the IC of the Integer Order Systems

Consider a well-known class of systems described by differential equations of the type

$$\sum_{k=0}^N a_k D^k y(t) = \sum_{k=0}^M b_k D^k x(t) \quad t \in \mathbb{R} \tag{35}$$

where D^k means the classic integer k -order derivative, and N, M means the orders. The parameters, $a_k, b_k, k = 0, 1, \dots$, are real numbers. Without losing generality, we set $a_N = 1$. By abuse of language, we will use the designation continuous-time (CT) ARMA systems for this kind of system [58]. They have a large number of applications in engineering, since a long time ago, and have been increasing their importance in economics and finance. The question we face is, again, how to modify the above equation to make the IC appear.

Theorem 3. Consider system (35), which is to be observed for $t \geq 0$, only. The modified formulation, to include the IC, reads

$$\begin{aligned}
 \sum_{k=0}^N a_k [y(t) \cdot u(t)]^{(k)} &= \sum_{k=0}^M b_k [x(t) \cdot u(t)]^{(k)} \\
 + \sum_{k=1}^N a_k \sum_{m=0}^{k-1} y^{(m)}(0) \delta^{(k-m)}(t) &- \sum_{k=1}^M b_k \sum_{m=0}^{k-1} x^{(m)}(0) \delta^{(k-m)}(t).
 \end{aligned} \tag{36}$$

Proof. Assume that we are not observing the system, until the instant $t = 0^-$. This means that our observation window is the unit step. Analytically, this is equivalent to multiplying both members of the equation by $u(t)$. We will get terms with the general form $u(t)D^k x(t)$. To manipulate the resulting equation, we need to relate $u(t)D^k x(t)$ to $D^k [x(t)u(t)]$. This is done by the (“saltus”) jump formula [59,60]

$$D^k[x(t)u(t)] = D^k[x(t)]u(t) + \sum_{n=0}^{k-1} D^{k-1-n}x(0)\delta^{(n)}(t) \tag{37}$$

The result (36) appears, immediately. This new equation involves the IC directly, without any transform. □

It is interesting to take a look at Equation (37) and try to understand its action. For $k = 1, 2, \dots$, we have, successively

$$D[x(t)]u(t) = D[x(t)u(t)] - x(0)\delta(t) = D[x(t)u(t) - x(0)u(t)]$$

meaning that we removed the jump at the origin, resulting from the multiplication by the unit step. Going on:

$$D^2[x(t)]u(t) = D^2[x(t)u(t)] - x'(0)\delta(t) - x(0)\delta'(t) = D\{D[x(t)u(t) - x(0)u(t)] - Dx(0)u(t)\}$$

Again, we removed the jump in the first order derivative. Therefore, the jump formula comes from a recursive removal of the successive jumps resulting from each derivative computation. It is interesting to remark that Equation (36) is almost equal to the formula we obtain using the ULT.

4.4. Fractional Order Systems

Returning to Equation (27), we observe that the principle found in the previous subsection remains valid: *before doing any derivative computation, we need to remove the jump.*

Theorem 4. Consider an increasing sequence of positive real numbers $\gamma_k, k = 0, 1, \dots, N$. The fractional jump formula reads [22]

$$f(t)^{(\gamma_N)} \cdot u(t) = [f(t) \cdot u(t)]^{(\gamma_N)} - \sum_{m=0}^{N-1} y^{(\gamma_m)}(0)\delta^{(\gamma_N-\gamma_m-1)}(t). \tag{38}$$

where $f(t)^{(\gamma_k)}$ stands for a fractional derivative defined on \mathbb{R} : GL, L, or LC.

We must highlight the fact that the $\gamma_k, k = 0, 1, \dots, N$ sequence is imposed by the particular application at hand, and not by any transform. The fractional jump formula provides a rule for modifying a given differential equation, to make the IC appear explicitly. This can be done, also, in nonlinear equations.

By curiosity, note that, if we use the BLT, we obtain [25]

$$\mathcal{L}[D^{\gamma_N} f(t) \cdot \varepsilon(t)] = s^{\gamma_N} \mathcal{L}[f(t) \cdot \varepsilon(t)] - \sum_{m=0}^{N-1} f^{(\gamma_m)}(0)s^{\gamma_N-\gamma_m-1}. \tag{39}$$

However, this is a consequence of Equation (38), not of the transform. We could use the Fourier transform, also.

Therefore, consider a general FARMA-type system, generalizing Equation (35):

$$\sum_{k=0}^N a_k D^{\alpha_k} y(t) = \sum_{k=0}^M b_k D^{\beta_k} x(t), \tag{40}$$

where the parameters α_k and β_k are the derivative orders that we assume form strictly increasing sequences of positive numbers. Using the fractional jump formula, we can reformulate it to include the IC:

$$\sum_{k=0}^N a_k [y(t) \cdot u(t)]^{(\alpha_k)} = \sum_{k=0}^M b_k [x(t) \cdot u(t)]^{(\beta_k)} + \sum_{k=1}^N a_k \sum_{m=0}^{k-1} y^{(\alpha_m)}(0) \delta^{(\alpha_k - \alpha_m)}(t) - \sum_{k=1}^M b_k \sum_{m=0}^{k-1} x^{(\beta_m)}(0) \delta^{(\beta_k - \beta_m)}(t). \tag{41}$$

For the commensurate case, we obtain

$$\sum_{k=0}^N a_k [y(t) \cdot u(t)]^{(\alpha k)} = \sum_{k=0}^M b_k [x(t) \cdot u(t)]^{(\alpha k)} + \sum_{k=1}^N a_k \sum_{m=0}^{k-1} y^{(\alpha m)}(0) \delta^{(\alpha k - \alpha m)}(t) - \sum_{k=1}^M b_k \sum_{m=0}^{k-1} x^{(\alpha m)}(0) \delta^{(\alpha k - \alpha m)}(t). \tag{42}$$

If $\alpha = 1$, we recover Equation (36). The above formula gives us the general solution to the IC problem for any FARMA-like system.

Therefore, the solution of a given IC problem implies a redefinition of the corresponding differential equation, so that the IC appear explicitly. Only after this change, can we use a transform.

4.5. From the Observable Canonical Form

We are going to show, using a different approach, that the above methodology is correct. To do it, we use a state-space formulation.

In Section 3.2, we showed that, for the solution of a simple equation of the type $y^{(\alpha)}(t) + ay(t) = v(t)$, only one IC is required. This remains valid for the state-variable formulation [22].

Theorem 5. Let \mathbf{A} be an $n \times n$ non singular matrix and $\alpha > 0$. The solution of the equation

$$D^{(\alpha)}[\mathbf{x}(t) - \mathbf{x}(0)u(t)] = -\mathbf{A}\mathbf{x}(t), \quad t > 0 \tag{43}$$

is given by

$$\mathbf{x}(t) = \mathbf{x}(0)\mathcal{E}_\alpha(t, -\mathbf{A}). \tag{44}$$

For proof, see [22].

Consider a FARMA(n,n) commensurate system

$$D^{n\alpha}y + a_n D^{(n-1)\alpha}y + a_{n-1} D^{(n-2)\alpha}y + \dots + a_0 y = b_n D^\alpha u + b_{n-1} D^{(n-1)\alpha}v + b_{n-2} D^{(n-2)\alpha}v + \dots + b_0 v$$

and rewrite it as

$$D^{n\alpha}y = b_n D^{n\alpha}v + D^{(n-1)\alpha}(b_{n-1}v - a_{n-1}y) + D^{(n-2)\alpha}(b_{n-2}v - a_{n-2}y) + \dots + (b_0v - a_0y)$$

or

$$y = b_n v + D^{-\alpha}(b_{n-1}v - a_{n-1}y) + D^{-2\alpha}(b_{n-2}v - a_{n-2}y) + \dots + D^{-n\alpha}(b_0v - a_0y)$$

The consideration of equal orders of AR and MA does not introduce any restrictions, as we can always define additional null coefficients. We can give these formulae a matrix format, by introducing a suitable state vector \mathbf{x} :

$$D^\alpha \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_0 - a_0 b_n \\ b_1 - a_1 b_n \\ \vdots \\ \vdots \\ b_{n-1} - a_{n-1} b_n \end{bmatrix} v$$

$$y = [0 \ 0 \ \dots \ 0 \ 1] \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix} + b_n v$$

The state vector components, $x_k, k = 1, 2, \dots, n$, are defined by

$$x_k(t) = D^{(n-k)\alpha} [b_{n-k}v(t) - a_{n-k}y(t)]. \tag{45}$$

This means that the initial state vector is constituted by successive derivatives of input and output functions at $t = 0$.

$$x_k(0) = b_{n-k}D^{(n-k)\alpha}v(0) - a_{n-k}D^{(n-k)\alpha}y(0).$$

This confirms the correctness of the procedure introduced in the previous sub-section.

4.6. A Bucket of Cold Water?

Consider a simple system, obtained from Equation (42), with $M = 0, N = 1$, and null IC:

$$y^{(\alpha)}(t) + Ay(t) = x(t)$$

It is not very hard to find the impulse response (α -exponential function [15])

$$h(t) = \sum_1^{\infty} (-A)^{n-1} \frac{t^{n\alpha-1}}{\Gamma(n\alpha)} u(t)$$

and the step response

$$r_u(t) = \sum_1^{\infty} (-A)^{n-1} \frac{t^{n\alpha}}{\Gamma(n\alpha) + 1} u(t) = \frac{1 - \mathcal{E}_\alpha(t, -A)}{A} u(t)$$

Let $x(t) = u(t + T) - u(t)$, with $T > 0$. The output is

$$y(t) = \begin{cases} \frac{\mathcal{E}_\alpha(t, -A) - \mathcal{E}_\alpha(t + T, -A)}{A} & t \geq 0 \\ \frac{u(t + T) - \mathcal{E}_\alpha(t + T, -A)}{A} & t < 0. \end{cases}$$

We can consider the upper term as the transient response for $t > 0$, for which there is no input. This situation has been treated as an IC problem [20,23,24]. However, it is not. We have a switched-on system with a dynamic behaviour that produced an output until $t = 0^-$. So, we have an extrapolation or prediction problem, as studied by Kolmogorov [55] and Wiener [56]. There are many approaches to the solution, depending on the type of past observations

1. Acquisitions of past input and output in a given interval, we can
 - Design a predictor by the Wiener–Hopf method of a Kalman filter [61],
 - Estimate a continuous-time ARMA model [62,63],
 - Use other functional extrapolation methods, such as polynomial or by splines.
2. Sampling the input and output signals
 - Design a linear predictor [64–67],
 - Use a discrete-time Kalman filter [61],
 - Estimate an ARMA model [68,69].

There are other alternatives. In practice, the option by one or another is, also, related to the deterministic or random characteristics of the used signals. This theme goes beyond the objectives of this work.

Anyway, let us treat the above case as if it was an IC problem. As the IC, taken at $t = 0^-$, is

$$y(0^-) = \frac{1 - \mathcal{E}_\alpha(T, -A)}{A}$$

the corresponding response would be the solution of

$$[y_1(t) - y(0^-)u(t)]^{(\alpha)} + Ay_1(t) = 0,$$

in agreement with the procedure in Equation (24),

$$y_1(t) = \frac{u(t) - \mathcal{E}_\alpha(T, -A)}{A} \mathcal{E}_\alpha(t, -A)$$

However, the ideal solution we get above is

$$y_0(t) = \frac{\mathcal{E}_\alpha(t, -A) - \mathcal{E}_\alpha(t + T, -A)}{A} = \frac{1 - \frac{\mathcal{E}_\alpha(t+T, -A)}{\mathcal{E}_\alpha(t, -A)}}{A} \mathcal{E}_\alpha(t, -A)$$

Obviously, we are not expecting that they are equal. However, and for an order near 1, the IC extrapolator $y_1(t)$ is reasonably good. To see it, let $\alpha = 1$, so that $\mathcal{E}_1(t, -A) = e^{-At}u(t)$. In such a case,

$$y_0(t) = y_1(t) = \frac{1 - e^{-AT}}{A} e^{-At}u(t)$$

is exact.

5. Conclusions

From the theory just presented, we can draw some important conclusions, namely: the IC provided by the unilateral LT are their own IC, not necessarily from any system; the required IC depend on the structure of the system in question, mainly on the set of derivative orders; and the IC and extrapolation/prediction problems are not equivalent.

Funding: This work was partially funded by National Funds through the FCT-Foundation for Science and Technology within the scope of the CTS Research Unit—Center of Technology and Systems/UNINOVA/FCT/NOVA, under the reference UIDB/00066/2020.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ARMA	autoregressive-moving average
BLT	bilateral Laplace transform
C	Caputo
CMLF	causal Mittag-Leffler function
FARMA	fractional autoregressive-moving average
FC	Fractional calculus
GL	Grünwald–Letnikov
IC	initial-conditions
L	Liouville
LC	Liouville–Caputo
MLF	Mittag-Leffler function
RL	Riemann–Liouville
ULT	unilateral Laplace transform

References

1. Heaviside, O. *Electrical Papers*; Macmillan Co.: London, UK; New York, NY, USA, 1892; Volume 1.
2. Heaviside, O. *Electromagnetic Theory*; The Electrician Printing and Publishing Co.: London, UK, 1892; Volume 1.
3. Heaviside, O. *Electrical Papers*; Macmillan Co.: London, UK; New York, NY, USA, 1894; Volume 2.
4. Lützen, J. Heaviside's operational calculus and the attempts to rigorise it. *Arch. Hist. Exact Sci.* **1979**, *21*, 161–200. [[CrossRef](#)]
5. Bromwich, T.I. Normal coordinates in dynamical systems. *Proc. Lond. Math. Soc.* **1917**, *2*, 401–448. [[CrossRef](#)]
6. Carson, J.R. On a general expansion theorem for the transient oscillations of a connected system. *Phys. Rev.* **1917**, *10*, 217. [[CrossRef](#)]
7. Carson, J.R. Theory of the transient oscillations of electrical networks and transmission systems. *Trans. Am. Inst. Electr. Eng.* **1919**, *38*, 345–427. [[CrossRef](#)]
8. Carson, J.R. The Heaviside operational calculus. *Bell Syst. Tech. J.* **1922**, *1*, 43–55. [[CrossRef](#)]
9. Van der Pol, B. A simple proof and an extension of Heaviside's operational calculus for invariable systems. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1929**, *7*, 1153–1162. [[CrossRef](#)]
10. Doetsch, G. *Theorie und Anwendung der Laplace-Transformation*; Springer: Berlin, Germany, 1937.
11. Oppenheim, A.V.; Willsky, A.S.; Hamid, S. *Signals and Systems*, 2nd ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 1997.
12. Roberts, M. *Signals and Systems: Analysis Using Transform Methods and Matlab*, 2nd ed.; McGraw-Hill: New York, NY, USA, 2003.
13. Lundberg, K.H.; Miller, H.R.; Trumper, D.L. Initial conditions, generalized functions, and the Laplace transform: Troubles at the origin. *IEEE Control Syst. Mag.* **2007**, *27*, 22–35.
14. Samko, S.G.; Kilbas, A.A.; Marichev, O.I. *Fractional Integrals and Derivatives*; Gordon and Breach: Yverdon, Switzerland, 1993.
15. Kilbas, A.A.; Srivastava, H.M.; Trujillo, J.J. *Theory and Applications of Fractional Differential Equations*; Elsevier: Amsterdam, The Netherlands, 2006.
16. Petráš, I. *Fractional-Order Nonlinear Systems: Modeling, Analysis and Simulation*; Series Nonlinear Physical Science; Springer: Heidelberg, Germany, 2011.
17. Kochubei, A.; Luchko. *Handbook of Fractional Calculus with Applications: Basic Theory*; De Gruyter Berlin: Berlin, Germany, 2019; Volume 1.
18. Lorenzo, C.F.; Hartley, T.T. Initialization in fractional order systems. In Proceedings of the 2001 European Control Conference (ECC), Porto, Portugal, 4–7 September 2001; IEEE: Piscataway, NJ, USA, 2001; pp. 1471–1476.
19. Lorenzo, C.F.; Hartley, T.T. Initialization of Fractional-Order Operators and Fractional Differential Equations. *J. Comput. Nonlinear Dyn.* **2008**, *3*, 021101. [[CrossRef](#)]
20. Sabatier, J.; Merveillaut, M.; Malti, R.; Oustaloup, A. How to impose physically coherent initial conditions to a fractional system? *Commun. Nonlinear Sci. Numer. Simul.* **2010**, *15*, 1318–1326. [[CrossRef](#)]
21. Heymans, N.; Podlubny, I. Physical interpretation of initial conditions for fractional differential equations with Riemann-Liouville fractional derivatives. *Rheol. Acta* **2006**, *45*, 765–771. [[CrossRef](#)]
22. Ortigueira, M.D.; Coito, F.J. System initial conditions vs. derivative initial conditions. *Comput. Math. Appl.* **2010**, *59*, 1782–1789. [[CrossRef](#)]
23. Sabatier, J.; Farges, C. Comments on the description and initialization of fractional partial differential equations using Riemann-Liouville's and Caputo's definitions. *J. Comput. Appl. Math.* **2018**, *339*, 30–39. [[CrossRef](#)]
24. Trigeassou, J.C.; Maamri, N. Initial conditions and initialization of linear fractional differential equations. *Signal Process.* **2011**, *91*, 427–436. [[CrossRef](#)]
25. Ortigueira, M.D.; Machado, J.T. Revisiting the 1D and 2D Laplace transforms. *Mathematics* **2020**, *8*, 1330. [[CrossRef](#)]
26. Heaviside, O. *Electromagnetic Theory 2*; "The Electrician" Printing and Publishing Company, Limited: London, UK, 1925; Volume 2.
27. Westerlund, S. *Dead Matter Has Memory!*; Causal Consulting: Kalmar, Sweden, 2002.
28. Magin, R.L. *Fractional Calculus in Bioengineering*; Begell House: Danbury, CT, USA, 2006.
29. Mainardi, F. *Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models*; Imperial College Press: London, UK, 2010.
30. Tarasov, V.E. *Fractional Dynamics: Applications of Fractional Calculus to Dynamics of Particles, Fields and Media*; Nonlinear Physical Science; Springer: Beijing, China; Heidelberg, Germany, 2011.
31. Machado, J.T. And I say to myself: "What a fractional world!". *Fract. Calc. Appl. Anal.* **2011**, *14*, 635–654. [[CrossRef](#)]
32. Ionescu, C.M. *The Human Respiratory System: An Analysis of the Interplay between Anatomy, Structure, Breathing and Fractal Dynamics*; BioEngineering; Springer: London, UK, 2013.
33. Herrmann, R. *Fractional Calculus: An Introduction for Physicists*, 3rd ed.; World Scientific: Singapore, 2018.
34. Martynyuk, V.; Ortigueira, M.; Fedula, M.; Savenko, O. Methodology of electrochemical capacitor quality control with fractional order model. *AEU-Int. J. Electron. Commun.* **2018**, *91*, 118–124. [[CrossRef](#)]
35. Liouville, J. Mémoire sur le calcul des différentielles à indices quelconques. *J. L'École Polytech. Paris* **1832**, *13*, 71–162.
36. Liouville, J. Mémoire sur quelques questions de Géométrie et de Mécanique, et sur un nouveau genre de calcul pour résoudre ces questions. *J. L'École Polytech. Paris* **1832**, *13*, 1–69.
37. Podlubny, I. *Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications*; Academic Press: San Diego, CA, USA, 1999; pp. 1–3.

38. De Oliveira, E.C.; Tenreiro Machado, J.A. A review of definitions for fractional derivatives and integrals. *Math. Probl. Eng.* **2014**, *2014*, 238459. [[CrossRef](#)]
39. Teodoro, G.S.; Machado, J.T.; De Oliveira, E.C. A review of definitions of fractional derivatives and other operators. *J. Comput. Phys.* **2019**, *388*, 195–208. [[CrossRef](#)]
40. Ortigueira, M.D.; Machado, J.T. What is a fractional derivative? *J. Comput. Phys.* **2015**, *293*, 4–13. [[CrossRef](#)]
41. Ortigueira, M.D.; Valério, D.; Machado, J.A.T. Variable order fractional systems. *Commun. Nonlinear Sci. Numer. Simul.* **2019**, *71*, 231–243. [[CrossRef](#)]
42. Valério, D.; Ortigueira, M.D.; Lopes, A.M. How Many Fractional Derivatives Are There? *Mathematics* **2022**, *10*, 737. [[CrossRef](#)]
43. Ortigueira, M.D. Two-sided and regularised Riesz-Feller derivatives. *Math. Methods Appl. Sci.* **2021**, *44*, 8057–8069. [[CrossRef](#)]
44. Ortigueira, M.D. *Fractional Calculus for Scientists and Engineers*; Lecture Notes in Electrical Engineering; Springer: Dordrecht, The Netherlands; Heidelberg, Germany, 2011.
45. Ortigueira, M.D. The fractional quantum derivative and its integral representations. *Commun. Nonlinear Sci. Numer. Simul.* **2010**, *15*, 956–962. [[CrossRef](#)]
46. Dugowson, S. *Les Différentielles Métaphysiques (Histoire et Philosophie de la Généralisation de L'ordre de Dérivation)*. Ph.D. Thesis, Université Paris Nord, Villetaneuse, France, 1994.
47. West, B.J. *Fractional Calculus View of Complexity: Tomorrow's Science*; CRC Press: Boca Raton, FL, USA, 2015.
48. Ortigueira, M.; Machado, J. Which Derivative? *Fractal Fract.* **2017**, *1*, 3. [[CrossRef](#)]
49. Gel'fand, I.M.; Shilov, G.E. *Generalized Functions: Properties and Operations*; Academic Press: New York, NY, USA, 1964.
50. Jiang, Y.; Zhang, B. Comparative study of Riemann–Liouville and Caputo derivative definitions in time-domain analysis of fractional-order capacitor. *IEEE Trans. Circuits Syst. II Express Briefs* **2019**, *67*, 2184–2188. [[CrossRef](#)]
51. Tavazoei, M.S.; Haeri, M.; Attari, M.; Bolouki, S.; Siami, M. More details on analysis of fractional-order Van der Pol oscillator. *J. Vib. Control* **2009**, *15*, 803–819. [[CrossRef](#)]
52. Dokoumetzidis, A.; Magin, R.; Macheras, P. Fractional kinetics in multi-compartmental systems. *J. Pharmacokinet. Pharmacodyn.* **2010**, *37*, 507–524. [[CrossRef](#)]
53. Valério, D.; da Costa, J.S. *An Introduction to Fractional Control*; Control Engineering; IET: Stevenage, UK, 2012.
54. Kochubei, A.; Luchko. *Handbook of Fractional Calculus with Applications: Fractional Differential Equations*; De Gruyter Berlin: Berlin, Germany, 2019; Volume 2.
55. Kolmogoroff, A. Interpolation und Extrapolation von stationären zufälligen Folgen. *Bull. Acad. Sci. URSS Sér. Math.* **1941**, *5*, 3–14.
56. Wiener, N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*; MIT Press: Cambridge, MA, USA, 1949; Volume 113.
57. Ortigueira, M.D.; Valério, D. *Fractional Signals and Systems*; De Gruyter: Berlin, Germany; Boston, MA, USA, 2020.
58. Mahata, K.; Fu, M. On the indirect approaches for CARMA model identification. *Automatica* **2007**, *43*, 1457–1463. [[CrossRef](#)]
59. Ferreira, J. *Introduction to the Theory of Distributions*; Pitman Monographs and Surveys in Pure and Applied Mathematics; Pitman: London, UK, 1997.
60. Hoskins, R. *Delta Functions: An Introduction to Generalised Functions*; Woodhead Publishing Limited: Cambridge, UK, 2009.
61. Kailath, T. Lectures on wiener and kalman filtering. In *Lectures on Wiener and Kalman Filtering*; Springer: Berlin/Heidelberg, Germany, 1981; pp. 1–143.
62. Brockwell, P. Recent results in the theory and applications of CARMA processes. *Ann. Inst. Stat. Math.* **2014**, *66*, 647–685. [[CrossRef](#)]
63. Boularouk, Y.; Djeddour, K. New approximation for ARMA parameters estimate. *Math. Comput. Simul.* **2015**, *118*, 116–122. [[CrossRef](#)]
64. Wold, H.O. On prediction in stationary time series. *Ann. Math. Stat.* **1948**, *19*, 558–567. [[CrossRef](#)]
65. Makhoul, J. Linear prediction: A tutorial review. *Proc. IEEE* **1975**, *63*, 561–580. [[CrossRef](#)]
66. Ortigueira, M.D.; Matos, C.J.; Piedade, M.S. Fractional discrete-time signal processing: Scale conversion and linear prediction. *Nonlinear Dyn.* **2002**, *29*, 173–190. [[CrossRef](#)]
67. Vaidyanathan, P.P. The theory of linear prediction. *Synth. Lect. Signal Process.* **2007**, *2*, 1–184. [[CrossRef](#)]
68. Fan, H.; Soderstrom, T.; Mossberg, M.; Carlsson, B.; Zou, Y. Estimation of continuous-time AR process parameters from discrete-time data. *IEEE Trans. Signal Process.* **1999**, *47*, 1232–1244. [[CrossRef](#)]
69. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.