



Article

Discrete-Time Fractional Difference Calculus: Origins, Evolutions, and New Formalisms

Manuel Duarte Ortigueira

NOVA School of Science and Technology, UNINOVA-CTS and LASI, NOVA University of Lisbon, Quinta da Torre, 2829-516 Caparica, Portugal; mdo@fct.unl.pt

Abstract: Differences are introduced as outputs of linear systems called differencers, being considered two classes: shift and scale-invariant. Several types are presented, namely: nabla and delta, bilateral, tempered, bilinear, stretching, and shrinking. Both continuous and discrete-time differences are described. ARMA-type systems based on differencers are introduced and exemplified. In passing, the incorrectness of the usual delta difference is shown.

Keywords: nabla difference; delta difference; ARMA; ARFIMA; bilateral difference; tempered; bilinear; stretching difference

MSC: 26A33

1. Introduction

All everyday human activities give rise to signals that carry a certain type of information about the systems that generated them. These signals are bounded functions that are collected to be studied, transmitted and manipulated in order to extract the information they carry. Discrete-Time Signal Processing (DTSP), also called Digital Signal Processing, is a set of mathematical and engineering techniques that allow the processing (collection, study, analysis, synthesis, transformation, storage, etc.) of signals performed mainly on digital devices.

Combining ideas, theories, algorithms and technologies from different quadrants, the DTSP has not stopped continuously evolving and increasing its already vast field of applications. This evolution was motivated by the enormous progress of digital technologies that allow the construction of processors, in general, more reliable and robust than analog ones and, above all, more flexible. The on-chip implementation of specialized processors (e.g., FFT) has facilitated the application of mathematical techniques that would be difficult (or impossible) to perform analogically. The DTSP plays an important role in communication systems where its mission is to handle signals, both at transmission and reception, in order to achieve an efficient and reliable flow of information between source and receiver. However, it is not only in communication systems that we find DTSP applications. In fact, its field of action has widened and includes areas such as speech, radar and sonar, seismology, bio-medicine, economics, astronomy, etc. In mathematics, it has been very useful in the study of functions and in solving differential equations. The well-known Newton series is very famous.

Mathematically, the DTSP relied on several important tools such as real and complex analysis, difference equations, discrete-time Fourier and Z transforms, algebra, etc. It benefited from the enormous development of signal theory in the 2nd half of the 20th century when signal processing techniques reached a sufficiently high degree of development. However, its origins are much earlier.

In general, we can “date” the beginning of the study of signals to the discovery of periodic phenomena that led to the introduction of the notions of year, week, day, hour, etc. With an equal degree of importance, we can consider the theory and representation



Citation: Ortigueira, M.D.

Discrete-Time Fractional Difference Calculus: Origins, Evolutions, and New Formalisms. *Fractal Fract.* **2023**, *7*, 502. <https://doi.org/10.3390/fractalfract7070502>

Academic Editor: Ricardo Almeida

Received: 6 May 2023

Revised: 12 June 2023

Accepted: 23 June 2023

Published: 25 June 2023



Copyright: © 2023 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/>).

of music made by the Pythagoreans as the first spectral analysis. It is important to note that they actually made a discrete time-frequency formulation. More recently, we refer to the discovery and study of the spectrum of sunlight by Newton (1666) and the works of mathematicians such as Euler, Gauss (who devised the first algorithm for the fast Fourier transform in 1805), Fourier (who created the basis for spectral analysis), Sturm and Liouville. These works had direct implications on the way of studying signals in the frequency domain, which did not cease to evolve and gain importance from the 1940s thanks to the works of the theoretical field of stochastic processes (Wiener and Kolmogorov): correlation, adapted filter, Wiener filter, etc. [1,2], which are notions that would become the basis of modern developments in spectral analysis (Tukey, Parzen, Akaike, Papoulis, and Burg). It was also Tukey who, with Cooley, rediscovered the algorithm that allowed the implementation of the FFT in 1965, which was a milestone in signal analysis.

The difference equations, taking the form of the ARMA (autoregressive-moving average) model, had a rapid increase in importance due to the works of Box, Jenkins, Oppenheim, Kailath, Kalman, Robinson, Rabiner, and many others in the 1980s of the 20th century [3–8]. We can place here the real beginning and affirmation of DTSP. Nevertheless, the discovery of computers was perhaps the biggest impulse given to the DTSP by the possibility of discrete implementation of processor devices, which was previously made exclusively with analog technology and to perform simulations that allow to predict, with great accuracy, the behavior of a given system. This led to an autonomization of the theory of “Discrete Signals and Systems” that became an independent branch, leading to alternative technological solutions based on digital design and realization devices [3,4,8–15]. Although the main developments were based on difference equations, the true origins were not forgotten and motivated some attempts to model and identify systems based on the delta difference [16–20].

The emergence of fractional tools has opened new doors to the modeling and estimation of everyday systems that were known to be best described by fractional systems. However, this does not mean that there was a coherent theory of fractional systems in discrete time. Probably the first attempt was made in [21], but the systems described are not really fractional, although they use fractional delays. In the last 20 years, many texts have been published on fractional differences and derivatives in discrete time, leading to different views of what fractional systems in discrete time are and how they are characterized [22–28]. The purpose of this paper is exactly to describe the mathematical basis underlying the main formulations. We introduce differences through a system approach to highlight the fact that the required definition must be valid for any function regardless of its support. This allows for a broader scope. On the other hand, it is important to make a clear distinction between time flow from left to right (causal case) or the other way around (anti-causal). Under normal conditions, they should not be mixed. To this end, we define “time sequence” as an alternative to “time scale”, avoiding the confusion that the latter might introduce. We will proceed with the definitions of nabla (causal) and delta (anti-causal) differences and enumerate their main properties. We proceed with the introduction of other formulations, such as discrete-time, bilateral, tempered differences, and the completely new bilinear differences. These differences are “invariant under translations”. We propose new “scale-invariant differences” that are connected to Hadamard derivatives. For all the presented differences, ARMA-type difference equations are proposed.

Given the importance of discrete signals inherent in this work, we review the classical sampling theorem valid for the case of shift-invariant systems [14,29–32] and another one suitable for scale-invariant systems, but they differ from similar studies in the literature [33,34].

The paper is outlined as follows. In Section 2, we present several mathematical tools useful in the paper and clarify some notions. The sampling theorems are introduced here. Section 3 is used to make a historical overview of the difference evolutions, both continuous and discrete-time. The different approaches are described. The problems created by some definitions are criticized in Section 4. The main contributions in this paper are presented

in Section 5 where several shift-invariant differencers and accumulators, say: nabla, delta, two-sided, tempered, and bilinear, are introduced. For all the definitions, continuous-time and discrete-time versions are presented. The scale-invariant differences are introduced and studied in Section 6. All the described differences are suitable for defining ARMA-type linear systems. This is exemplified in Section 7. Finally, we present a brief discussion.

2. Preliminaries

2.1. Glossary and Assumptions

In the evolution of the DTSP, several notions have been introduced without retaining a clear meaning. In fractional calculus, there is considerable confusion in the terminology adopted having, in some cases, the same name for different operators. Here, we try to clarify the meaning of some terms in order to avoid confusion. Therefore, we start with some fundamental terms. Later, we will introduce others, which are needed in the rest of the document [35].

- **Anti-causal [36]**
An anti-causal system is causal under reverse time flow. A system is *anti-causal* if the output at any instant depends only on values of the input and/or output at the present and future time instants. The delta derivative is an example of an anti-causal system.
- **Anti-difference**
The operator that is simultaneously the left and right inverse of the difference will be called *anti-difference*.
- **Backward**
Reverse time flow—from future to past.
- **Causal operator or system [37–39]**
A system is *causal* if the output at any instant depends only on the values of the input and/or output at the present and past instants. The nabla derivative is an example of a causal system.
- **Forward**
Normal time flow—from past to future.
- **Fractional**
Fractional will have the meaning of a non-integer real number.
- **Scale-invariant system**
A system is scale-invariant if a stretching or shrinking in the input produces the same stretching/shrinking in the output. It is described by the Mellin convolution [33,40]

$$y(\tau) = x(\tau) \star g(\tau) = \int_0^{\infty} x\left(\frac{\tau}{\eta}\right) g(\eta) \frac{d\eta}{\eta}.$$

- **Signal**
Bounded function that conveys some kind of information.
- **Shift-invariant system**
A system is shift-invariant if a delay or lead in the input produces the same delay/lead in the output. It is described by the usual convolution [37]

$$y(t) = x(t) * g(t) = \int_{-\infty}^{+\infty} x(t - \eta) g(\eta) d\eta.$$

- **System [37,38]**
Any operator that transforms signals into signals. We will often use the terms system and operator interchangeably.

2.2. Some Mathematical Tools

Traditionally, the delta symbol is used for the so-called “forward difference” and it comes from a long time ago [41], while nabla is attached to the “backward difference” [42],

in contradiction with the time flow. However, they are generalized for any order through the same way: the binomial theorem [43]. Let $\alpha \in \mathbb{R} - \mathbb{Z}^-$. Then,

$$(1 - z)^\alpha = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} z^k, \quad |z| < 1. \tag{1}$$

We can extend it for negative integer values of α through the Pochhammer symbol. The binomial coefficients

$$\binom{\alpha}{\beta} = \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha - \beta + 1)\Gamma(\beta + 1)}$$

assume a central role in the theory we will develop in the following. They enjoy interesting properties [36,44]:

•

$$\left| \binom{\alpha}{\beta} \right| \leq \frac{A}{\beta^{\alpha+1}}, \text{ for } \beta \rightarrow \infty,$$

with $A > 0$.

•

$$\binom{\alpha}{n} = \frac{\Gamma(\alpha + 1)}{\Gamma(\alpha - n + 1)n!} = (-1)^n \frac{(-\alpha)_n}{n!}, \quad n \in \mathbb{N},$$

where $(a)_n$ is the Pochhammer symbol for the rising factorial

$$(a)_n = \prod_{k=0}^{n-1} (a + k) = \frac{\Gamma(a + n)}{\Gamma(a)}, \text{ with } (a)_0 = 1,$$

generalized as

$$(a)_\beta = \frac{\Gamma(a + \beta)}{\Gamma(a)}, \text{ with } (a)_0 = 1.$$

•

$$\binom{\alpha}{n} = (-1)^n \binom{n - \alpha - 1}{n}.$$

•

$$\binom{\alpha + n}{n} = \binom{\alpha + n}{\alpha}.$$

•

$$\sum_{m=0}^{\infty} \binom{\alpha}{m} \binom{\beta}{n - m} = \binom{\alpha + \beta}{n}.$$

• The falling factorial is represented by [42]

$$(a)^{(n)} = \prod_{k=0}^{n-1} (a - k) = \frac{\Gamma(a + 1)}{\Gamma(a + 1 - n)}, \text{ with } (a)_0 = 1,$$

so that

$$\binom{\alpha}{n} = \frac{(a)^{(n)}}{n!}, \quad n \in \mathbb{N},$$

and

$$(a)^{(n)} = (-1)^n (-a)_n.$$

It is generalized by

$$(a)^{(\beta)} = \frac{\Gamma(a + 1)}{\Gamma(a + 1 - \beta)}, \text{ with } (a)_0 = 1,$$

and

$$(a)^{(\beta)} = (a - \beta + 1)_\beta.$$

The (bilateral) Laplace transform (LT) is given [45]:

$$\mathcal{L}[h(t)] = H(s) = \int_{-\infty}^{\infty} h(t)e^{-st} dt, \quad s \in \mathbb{C}, \quad (2)$$

that is assumed to converge in some non-void region (region of convergence—ROC) which may degenerate into the imaginary axis, giving rise to the Fourier transform (with $s = i\omega$). We define the inverse LT by the Bromwich integral

$$h(t) = \mathcal{L}^{-1}F(s) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} H(s)e^{st} ds, \quad t \in \mathbb{R}, \quad (3)$$

where $a \in \mathbb{R}$ is called abscissa of convergence. Frequently, we denote by γ the integration path. In a similar way, we define the Mellin transform (MT) by [40]

$$\mathcal{M}[h(t)] = H(v) = \int_0^{\infty} h(t)t^{-v-1} dt, \quad v \in \mathbb{C}, \quad (4)$$

with an inverse similar to (3)

$$h(t) = \mathcal{M}^{-1}H(v) = \frac{1}{2\pi i} \int_{\gamma} H(v)t^v dv, \quad t \in \mathbb{R}^+. \quad (5)$$

For the discrete-time case, we define the Z transform [14,36] by

$$\mathcal{Z}[f(n)] = F(z) = \sum_{n=-\infty}^{+\infty} f(n)z^{-n}, \quad z \in \mathbb{C}, \quad (6)$$

with the inverse given by the Cauchy integral

$$f(n) = \frac{1}{2\pi i} \oint_c F(z)z^{n-1} dz, \quad (7)$$

where c is the unit circle. With the change of variable $z = e^{i\omega}$, $-\pi < \omega \leq \pi$, we obtain the discrete-time Fourier transform.

2.3. On Time Sequences

A powerful approach to the continuous/discrete unification and generalization was introduced by Aulbach and Hilger through the calculus on *time scales* [46,47]. These are non-empty closed subsets \mathbb{T} of the set \mathbb{R} of real numbers. Let t be the current instant. Using the language of the time-scale calculus, the previous next instant is denoted by $\rho(t)$. Similarly, the next following point on the time scale \mathbb{T} is denoted by $\sigma(t)$. One has

$$\rho(t) = t - \nu(t), \quad \sigma(t) = t + \mu(t),$$

where $\nu(t)$ and $\mu(t)$ are called the *graininess* functions. These functions can be used to construct any time sequences. However, we will not continue this way.

Remark 1. The designation “time scale” is misleading, since the word “scale” is associated to a notion of stretching or shrinking, frequently having a relation to a speed or a rate. For example, consider a function $f(t)$ defined on \mathbb{R} and a parameter $a > 0$, which allow us to define a new function $g(t) = f(at)$. We modified the way that the flux of information is delivered. An interesting

example is given by the classic turntables where we are able to switch from a rotation speed to another one. This parameter is usually called scale. Therefore, we propose here the use of the designation time sequence.

In this work, we will consider time sequences \mathbb{T} defined by a set of discrete instants t_n , $n \in \mathbb{Z}$, and by the corresponding graininess functions. We define a direct graininess [48,49]

$$t_n = t_{n-1} + \nu_n, \quad n \in \mathbb{Z},$$

and reverse graininess

$$t_n = t_{n+1} - \mu_n, \quad n \in \mathbb{Z},$$

where we avoid representing any reference instant t_0 . These definitions of “irregular” sequences are suitable for dealing with some of the most interesting time sequences we find in practice. However, we have some difficulties in doing some kind of manipulations that are also very common. Let us consider a time sequence defined on \mathbb{R} and unbounded when $t \rightarrow \pm\infty$. For example, a time sequence defined by

$$t_n = nT + \tau_n, \quad n \in \mathbb{Z}, \quad T > 0, \quad |\tau_n| < \frac{T}{2},$$

which we can call an “almost linear sequence” [50]. However, in the most interesting engineering applications, we consider regular (uniform) sequences

$$\mathbb{T} = h\mathbb{Z} = \{ \dots, -3h, -2h, -h, 0, h, 2h, 3h, \dots \},$$

with $h \in \mathbb{R}^+$.

Remark 2. We can consider a slided time sequence by a given value, $a + h\mathbb{Z}$, $a < h$, but this corresponds to introducing another parameter that we cannot determine due to the relativistic character of any time sequence. In other words, we need another time sequence not depending on a to fix it. However, this may be an acceptable procedure for studying continuous-time functions.

Now, consider a power transformation of a time sequence:

$$\theta = q^t.$$

We generate a new (scale)–sequence which is in \mathbb{R}^+ . In particular, we will obtain sequences such as

$$\theta_n = \theta^n, \quad n \in \mathbb{Z}, \quad \theta > 0,$$

or

$$\theta_n = \theta_{n-1} \cdot \tau_n, \quad n \in \mathbb{Z}, \quad \tau_n > 0.$$

We will use these sequences when dealing with scale-invariant differences.

2.4. On the Sampling

Let $f(t)$, $t \in \mathbb{R}$ be a continuous-time bounded function. The discrete-time function obtained from $f(t)$ by retaining the values at a set of pre-specified instants is the sampled function, $f(t_n)$, $n \in \mathbb{Z}$. The procedure for obtaining such a function is called *ideal sampling*. From an operator (system) point of view, this is obtained with the help of the comb distribution. Although we can consider irregular combs, we will not do it here [50]. We intend to use uniform time sequences that lead us to the usual comb. This is a periodic repetition of the Dirac delta function [51–54]. Here, we state it in the following format.

$$c(t) = \sum_{n=-\infty}^{+\infty} \delta\left(\frac{t}{T} - n\right), \tag{8}$$

where T is the sampling interval. The Fourier transform (FT) of this function is also a periodic comb.

$$FT \left[\sum_{n=-\infty}^{+\infty} \delta(t - nT) \right] = \frac{2\pi}{T} \sum_{m=-\infty}^{+\infty} \delta(\omega - m \frac{2\pi}{T}). \tag{9}$$

The comb is called the ideal sampler because, when multiplying a given function, $x(t)$, by a comb, $c(t)$, it retains the samples of the original function, giving rise to a modulated comb:

$$x_s(t) = x(t) \cdot c(t) = \sum_{n=-\infty}^{+\infty} x(t) \delta(\frac{t}{T} - n) = T \sum_{n=-\infty}^{+\infty} x(nT) \delta(t - nT). \tag{10}$$

If $x(t)$ has a jump at $t = n_0T$, we use the half sum of the lateral limits $x(n_0T) = \frac{x(n_0T^+) + x(n_0T^-)}{2}$, which is in agreement with the inverse Laplace and Fourier integrals. Let $X(s)$ be the LT of $x(t)$. Then, the LT of $x_s(t)$ is given by [28]

$$X_s(s) = FT[x(t) \cdot c(t)] = \sum_{m=-\infty}^{+\infty} X(s - im \frac{2\pi}{T}), \tag{11}$$

stating the well-known phenomenon: sampling in a given domain implies a repetition in the transform domain, meaning that the sampling operation produces a repetition of the transform in parallel to the real axis in strips of width $\frac{2\pi}{T}$. We observe here the reason for including T in (11): the term corresponding to $m = 0$ is $X(s)$. The study we performed was based on the Laplace transform, but it can be performed also with the Fourier transform.

The choice of T depends on the objectives of the work at hand. In general, we can choose any value except if we have in mind the recovery of $X(s)$ from $X_s(s)$. In such a case, we can impose that $X(s)$ and $X_s(s)$ have the same poles in the strip defined by $|Im(s)| < \frac{\pi}{T}$. However, the most known approach is the *Whittaker–Kotel’nikov–Shannon sampling theorem* [12,14,29–32]

Theorem 1. *If a function is bandlimited to $[-\frac{\pi}{T}, \frac{\pi}{T}]$, it is completely determined by giving its ordinates at a series of points spaced T seconds apart [29–31]:*

$$f(t) = \sum_{k=-\infty}^{\infty} f(kT) \operatorname{sinc}(\frac{t}{T} - k), \tag{12}$$

where

$$\operatorname{sinc}(t) = \frac{\sin(\pi t)}{\pi t},$$

with $\operatorname{sinc}(0) = 1$ being the so-called *sinc* function that is the impulse response of the ideal lowpass filter [14].

Consider the instant $a + nT$, $n \in \mathbb{Z}$, $\gamma = a/T < 1$ and denote $f(nT)$ by f_n . We obtain

$$f(a + nT) = \sum_{k=-\infty}^{\infty} f(kT) \operatorname{sinc}(\frac{a}{T} + n - k)$$

and

$$f_{n+\gamma} = \sum_{k=-\infty}^{\infty} f(kT) \operatorname{sinc}(\gamma + n - k),$$

that states what we can call *fractional translation*, which allows expressing unknown intermediate values in terms of the uniformly spaced samples. The reverse can also be performed [21].

The *Whittaker–Kotel’nikov–Shannon sampling theorem* is based on the usual Fourier transform that has a relation with the shift-invariant systems, since it is defined in terms

of the eigenfunctions of such systems: the exponentials. To obtain a similar theorem for scale-invariant systems, we start from the Fourier–Mellin transform that we define by:

$$F(iv) = \int_0^{\infty} f(\tau)\tau^{-iv-1}d\tau, \quad (13)$$

with inverse

$$f(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(iv)\tau^{iv}dv, \quad (14)$$

obtained from (4) and (5) by letting $v = iv$.

Let $Q \in \mathbb{R}^+$. As above, consider a bandlimited scale function as the one verifying

$$F(iv) = 0, \quad |v| > \frac{\pi}{Q},$$

that has an associated Fourier series

$$F(iv) = \sum_{k=-\infty}^{\infty} F_k e^{-ikQv},$$

with

$$F_k = \frac{Q}{2\pi} \int_{-\frac{\pi}{Q}}^{\frac{\pi}{Q}} F(iu)e^{ikQu}du = Qf(e^{kQ}), \quad k \in \mathbb{Z}.$$

We have successively

$$\begin{aligned} f(\tau) &= \frac{1}{2\pi} \int_{-\frac{\pi}{Q}}^{\frac{\pi}{Q}} \sum_{k=-\infty}^{\infty} F_k e^{-ikQv} \tau^{iv} dv, \\ &= \sum_{k=-\infty}^{\infty} F_k \frac{1}{2\pi} \int_{-\frac{\pi}{Q}}^{\frac{\pi}{Q}} e^{iv(\ln(\tau)-kQ)} \tau^{iv} dv, \\ &= \sum_{k=-\infty}^{\infty} F_k \frac{\sin\left[\frac{\pi}{Q}(\ln(\tau)-kQ)\right]}{\pi(\ln(\tau)-kQ)}, \\ &= \sum_{k=-\infty}^{\infty} F_k \frac{\sin\left[\pi\left(\frac{\ln(\tau)}{Q}-k\right)\right]}{Q\pi\left(\frac{\ln(\tau)}{Q}-k\right)}. \end{aligned}$$

Therefore, the scale-invariant sampling theorem reads

Theorem 2. Let $f(\tau)$, $\tau \in \mathbb{R}^+$ be a bandlimited scale function. It is completely determined by giving its ordinates at a series of exponentially spaced points

$$f(\tau) = \sum_{k=-\infty}^{\infty} f(e^{kQ}) \operatorname{sinc}\left(\frac{\ln(\tau)}{Q} - k\right). \quad (15)$$

From this and through a comparison with Theorem 1, we conclude that the scale-invariant comb is given by:

$$s(\tau) = \sum_{k=-\infty}^{\infty} \delta\left(\frac{\ln(\tau)}{Q} - k\right), \quad (16)$$

that represents a sequence of impulses located at the points in the set $Q : \tau = e^{kQ}$, $k \in \mathbb{Z}$. These results, slightly different from similar results existing in the literature [33,34], will be used later in the corresponding difference definitions.

These two sampling theorems require that the functions have Fourier transforms with bounded support. However, this may not happen in practice, originating what is known

by “aliasing”, and several procedures exist to alleviate its effect [14,29–31]. If we do not want to recover the continuous-time function, the sampling theorems give us a chance to reduce the losses in the continuous-to-discrete conversion.

Remark 3. Traditionally, we simplify the notation by writing $f_n = f(nT)$ and $f_{n+\gamma} = f((n + \gamma)T)$. We will adopt this procedure. For the case stated in (15), we set $q = e^Q$ and $f_{q^n} = f(e^{nQ})$.

3. Historical Overview

3.1. Euler Procedure

Newton and Leibniz introduced their (different) approaches to infinitesimal calculus in the 17th century. Leibniz’s approach was based on generalizations of sums and differences. In particular, his definition of derivative was formulated in terms of limits of increments

$$f'(t) = \frac{df}{dt} = \lim_{h \rightarrow 0} \frac{f(t) - f(t - h)}{h} \quad f'(t) = \frac{df}{dt} = \lim_{h \rightarrow 0} \frac{f(t + h) - f(t)}{h}. \tag{17}$$

Euler (1768) took these formulae, removed the limit operation and used the “incremental quotients” as approximations of the derivative in solving differential equations on a set, \mathbb{T} , of pre-defined values of t , $\mathbb{T} = t_n, n = 0, 1, 2, \dots$. With this procedure, he was led to discrete functions, $f(t_n)$. It was the birth of the “difference equations”, and the procedure is the currently known Euler method. So, the difference equations gained the protagonism that belonged to the differential equations. This procedure originated the “numerical analysis” in mathematics and, in the 20th century, the discrete-time signal processing [13–15,36,37,55,56] that is a well established scientific area being responsible for important realizations in our daily life.

The former procedure, retaining the incremental ratio, was not completely abandoned; it continued being important as an intermediate step to obtain difference equations [14], and it was used in some applications [17,18]. The modern approach to differential discrete equations dates back to Hilger’s works on looking for a continuous/discrete unification [46,47,57].

To give a more precise idea and clarify the nomenclature, consider the differential equation

$$\frac{df(t)}{dt} + af(t) = \frac{dg(t)}{dt} + bg(t),$$

where $f(t)$ and $g(t)$ are real functions of real variables and $a, b \in \mathbb{R}$. Using the first incremental ratio

$$\nabla f(t) = \frac{f(t) - f(t - h)}{h}, \tag{18}$$

in the equation, we obtain

$$\nabla f(t) + af(t) = \nabla g(t) + bg(t), \tag{19}$$

that leads to

$$(1 + ah)f(t) - f(t - h) = (1 + bh)g(t) - g(t - h), \tag{20}$$

that is a difference equation. Note that only present and past values are involved: it represents a causal system. With a similar procedure for the other derivative

$$\Delta f(t) = \frac{f(t + h) - f(t)}{h}, \tag{21}$$

we obtain

$$\Delta f(t) + af(t) = \Delta g(t) + bg(t), \tag{22}$$

that gives

$$f(t+h) - (1+ah)f(t) = g(t+h) - (1+ah)g(t), \quad (23)$$

that is again a difference equation but involving only present and future values: anti-causal.

If we sample the function at a given discrete time grid $\mathbb{T} : t_n, n \in \mathbb{Z}$, (19) and (22) become discrete-time differential equations, while (20) and (23) transform into discrete-time difference equations.

3.2. Differences and Fractional Calculus

A fractional difference was introduced for the first time in 1832 by Liouville [58,59] who used it for defining a fractional derivative. His first formula, derived directly to be a generalization of the delta derivative, assumed the form:

$$\bar{\Delta}_L^\alpha f(t) = h^{-\alpha} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t + (\alpha - k)h). \quad (24)$$

This formula constituted a second step into defining fractional-order derivatives. However, Liouville recognized that such a formula was not exactly what he was looking for and introduced also another one, much more interesting, that we can write as

$$\nabla^\alpha f(t) = h^{-\alpha} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t - kh). \quad (25)$$

He observed that if $f(t) = e^{st}$, $t \in \mathbb{R}$, the summation was convergent if $Re(s) > 0$. This formula was the base of the fractional derivative definition used by Grünwald [60] and Letnikov [61,62]. It is important to note that the Liouville definition was assumed to be valid for functions of exponential type and defined on \mathbb{R} or \mathbb{C} , while Grünwald and Letnikov worked on $[t_0, t] \in \mathbb{R}$.

In the study of operators based on his operational methods, Heaviside obtained the same result and made a study of the binomial series which he generalized [43,63,64]. This approach was retaken later by E. Post (1930) [65] and by P. Butzer and U. Westphal (1974) [66]. In addition, in 1974, J. Diaz and T. Osier obtained a particular case of (24) with $h = 1$ [67]. In a later section, we will perform the analysis of such a definition.

It is important to highlight that while searching for a definition convergent for $Re(s) < 0$, Liouville arrived at the delta-type difference

$$\Delta_L^\alpha f(t) = (-h)^{-\alpha} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t + kh). \quad (26)$$

However, this route to fractional calculus never gained the favor of the vast majority of researchers, so it was basically considered as an approximation and used in numerical methods. Very interesting are the integral representations for (25) and (26) [25,67] and the two-sided differences [68,69]. The approach and applications introduced by V. Tarasov are also very important [23,24].

Due to its (bad) influence, we are going to study the “difference” (24) that states the first attempt Liouville made to define a fractional derivative using an infinite summation. It seems that in an independent way, Diaz and Osier proposed as a delta difference the expression [67]

$$\Delta_{D+O}^\alpha f(t) = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t + \alpha - k). \quad (27)$$

As it is easy to observe, this formula, while agreeing with the requests in the positive integer order, fails in the fractional order, since it uses simultaneously past and future values. This fact has implications in the discrete-time differences deduced from it.

To make a fair analysis, let us apply the LT to (24). We obtain

$$F_{\Delta_{D+O}^\alpha}(s) = e^{\alpha s} \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} e^{-ks} F(s). \tag{28}$$

The series converges only for $Re(s) > 0$, which implies causality, contrarily to what we were expecting, since the delta difference should be anti-causal. Therefore, the formula (24) does not make sense, and not good consequences should be expected from it.

3.3. Discrete-Time Differences

A less likely appearance of a discrete difference happened in a different context, the summation of series using Cesàro’s method [43,70,71], where S. Chapman introduced a fractional delta discrete difference as being given by

$$\Delta_C^\alpha v_n = \lim_{p \rightarrow \infty} \sum_{k=0}^{p-n+1} (-1)^k \binom{\alpha}{k} v_{n+k}, \quad n \in \mathbb{N}, \tag{29}$$

that was considered by G. Isaacs as an alternative to Diaz and Osler’s approach, which was mainly due to the validity of the associativity of orders [72]. However, in 1962, Isaacs proposed another formulation reading

$$\Delta_I^N v_n = \sum_{k=n}^{\infty} \binom{k-n-N-1}{k-n} v_k, \quad n \in \mathbb{N}. \tag{30}$$

It is important to highlight that Isaacs’ formulae express anti-causal differences as we should expect. Among several interesting results, he presented what can be considered a discrete Leibniz rule

$$\Delta_I^N (v_n w_n) = \sum_{k=N}^{\infty} \binom{N}{k} \Delta_I^k v_n \Delta_I^{N-k} w_{n+k}.$$

Meanwhile, discrete signal (time-series) analysis, which has benefited from major developments since World War II, received a spectacular boost with the publication of the landmark book *Time Series Analysis: Forecasting and Control* by George E. P. Box and Gwilym M. Jenkins [4]. Here, the authors presented a coherent and mathematically well-founded study of the ARMA models and their evolution, the Autoregressive-Moving Integrated Average models (ARIMA) [10,73]. The “integrated” factor pointed already to what was going to happen next: the insertion of a “fractionally integrated” term, leading to the ARFIMA models [74–80]. To the formalization of this new model, the concept of fractional differencing was introduced through (1) [74,75]

$$(1 - z^{-1})^\alpha = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} z^{-k} = \sum_{k=0}^{\infty} \frac{(-\alpha)_k}{k!} z^{-k},$$

for $|z| > 1$, meaning that

$$\nabla^\alpha v_n = \sum_{k=0}^{\infty} \frac{(-\alpha)_k}{k!} v_{n-k}, \quad n \in \mathbb{Z}, \tag{31}$$

that corresponds to make $h = 1$ in the nabla (Liouville–)Grünwald–Letnikov difference. These two formulae were also introduced earlier by Cargo and Shisha in the study of the zeros of polynomials [81]. The fractionally differenced and integrated models have gained roots and continue being used today [10,15,82]. In engineering, namely in signal processing, a fractional generalization of the ARMA model has been proposed [25,27].

A brief glance at (31) may give us an impression of repulsion due to the summation to infinity. However, such an impression is incorrect, because if the function is null for $n < n_0 \in \mathbb{Z}$, the summation becomes finite. For example, if $v_n = 0$, $n < 0$,

$$\nabla^\alpha v_n = \sum_{k=0}^n (-1)^k \frac{(-\alpha)_n}{n!} v_{n-k}, \quad n \in \mathbb{N}. \quad (32)$$

However, H. Gray and N. Zhang did not notice this fact and, in 1988, proposed an alternative through a new definition of the fractional difference. They had in mind to preserve many properties of fractional derivative definitions, namely the exponent law and the important Leibniz rule. To obtain it, they started from the summation formula that they repeated using the Cauchy procedure to obtain the n -fold summation

$$\nabla^{-N} f(t) = \frac{1}{\Gamma(N)} \sum_{k=k_0}^t (t-k+1)_{N-1} f(k).$$

This definition can be extended to $N \in \mathbb{Z}_0^-$. It is important to refer that such a formula was already known from the theory of Z transform [9], since it corresponds to the inverse of a multiple pole at the point $z = 1$ in the complex plane (discrete integrator or accumulator). Gray and Zhang were led to the following definition:

Let $\alpha \in \mathbb{R}$ and $f(t)$ are defined over the set $\mathbb{T} = k_0 - N, k_0 - N + 1, \dots, 0, 1, \dots, t \in \mathbb{Z}$. The α -order summation over $\mathbb{T}_0 = k_0, k_0 + 1, \dots, t \in \mathbb{Z}$ is defined by

$$\nabla^{-\alpha} f(t) = \frac{\nabla^N}{\Gamma(\alpha + N)} \sum_{k=k_0}^t (t-k+1)_{N+\alpha-1} f(k), \quad (33)$$

where $N = \max(0, N_0) \in \mathbb{Z}$ such that $0 < \alpha + N < 1$. The α -order derivative is defined by the substitution $-\alpha \rightarrow \alpha$. This definition is coherent with the usual integer-order difference. It is interesting to note that the above definition mimics the Riemann–Liouville definition of fractional derivative [36,44,83]. It can be shown that, setting $N - 1 < \alpha \leq N \in \mathbb{Z}_0^+$, we obtain

$$\begin{aligned} \nabla^\alpha f(t) &= \frac{\nabla^N}{\Gamma(N - \alpha)} \sum_{k=k_0}^t (t-k+1)_{N-\alpha-1} f(k) \\ &= \frac{\nabla^N}{\Gamma(N - \alpha)} \sum_{k=0}^{t-k_0} (k+1)_{N-\alpha-1} f(t-k). \end{aligned} \quad (34)$$

Consider the $N = 0$ case and note that

$$(k+1)_{-\alpha-1} = \frac{\Gamma(-\alpha+k)}{\Gamma(k+1)} = (-1)^k \binom{\alpha}{k}$$

meaning that (33) coincides with (26), provided that we remove the constraint on the domain of $f(t)$ and let it be defined over \mathbb{Z} . On the other hand, we can prove that if $N > 0$,

$$\nabla^N \binom{\alpha}{k} = \binom{\alpha+N}{k}.$$

In such a case, we can write, for any α

$$\begin{aligned} \nabla^\alpha f(t) &= \frac{\nabla^N}{\Gamma(N - \alpha)} \sum_{k=0}^{\infty} (k+1)_{-\alpha-1} f(t-k) \\ &= \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t-k). \end{aligned} \quad (35)$$

This last expression coincides with (26) when $h = 1$. We can conclude that the work of H. Gray and N. Zhang shows a different but equivalent to Liouville's and Chapman's way to introduce fractional differences. They presented the causal (nabla) version, but it is not very difficult to obtain the corresponding anti-causal (Δ). It is important to remark that they introduced, for the first time, the discrete-time Riemann–Liouville-type difference. The Caputo-type difference was introduced later for both nabla and delta cases [84–86]. Meanwhile, K. Miller and B. Ross [87] presented an approach to delta difference that we will consider next. After these publications, it seems there existed a time gap of almost 11 years without any novelty on the subject excepting the revision of the difference concepts having in consideration the notions and unification introduced by Hilger [46,47] and the discrete-time approximations to continuous-time derivatives by I. Podlubny [88]. In [57], Bohner and Peterson revised the concepts of differences and derivatives in terms of the notion of *time scale*.

4. A Critical View of Some Aspects Related to Differences

4.1. A "Fractional Delta Difference" that Is Not a Delta Difference

Concerning the fractional difference, only in 2007, F. Atici and P. Eloe returned to the subject by considering the delta difference first [89,90] and the nabla later [91]. Although they introduce a similar formulation to Gray and Zhang's, it seems they were unaware of it, at least initially. They intended to follow closely the continuous-time fractional derivative definitions. The starting point was the definition of Miller and Ross, which [87] they recovered [89,90]. However, this definition was based off the continuous-time difference of Diaz and Osler that we studied above. These works were followed by many others in the last 15 years. This had, as a consequence, a parallel introduction of several similar but different formalizations, even from the same author, of both nabla and delta differences, as the Riemann–Liouville-like and Caputo-like formulations, their compositions, and difference equations [22,84,85,89,90,92–103].

As referred, the approach of Atici and Eloe started by using the definition of Miller and Ross for the fractional discrete delta sum given by

$$(\Delta_a^{-\alpha} f)(t) = \frac{1}{\Gamma(\alpha)} \sum_{u=a}^{t-\alpha} (t - \sigma(u))^{\alpha-1} f(u) \quad (36)$$

where $\alpha > 0$, $\sigma(t) = t + 1$, and $t = a + \alpha + k$, $k = 0, 1, 2, \dots, N - 1$. The function $f(t)$ was defined in $\mathbb{T}_a = a, a + 1, a + 2, \dots$. As shown by D. Mozyrska [99], it can assume the form

$$\Delta^{-\alpha} f(t) = \sum_{m=0}^k \binom{k-m+\alpha-1}{k-m} f(a+m) \quad (37)$$

For the nabla difference, again in disagreement with Gray and Zhang, they proposed [90,91,104]

$$\nabla^{-\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \sum_{u=a}^t \frac{\Gamma(t-u+\alpha-1)}{\Gamma(t-u)} f(u)$$

As above, set $u = a + m$ and $t = a + 1 + k$ to obtain

$$\nabla^{-\alpha} f(t) = \frac{1}{\Gamma(\alpha)} \sum_{n=0}^k \frac{\Gamma(k-m+\alpha-1)}{\Gamma(k-m+1)} f(a+m) = \frac{1}{\Gamma(\alpha)} \sum_{n=0}^k \frac{\Gamma(n+\alpha-1)}{\Gamma(n+1)} f(a+k-m).$$

For other versions, see [84,92,95,98,104,105].

Remark 4. Strangely, the domain of the above delta sum is not \mathbb{T}_a but $\mathbb{T}_{a+\alpha}$. This means that the value at a given instant depends on past values, which is a causal behavior contrary to what we were expecting, since it is assumed to be a delta difference, which is anti-causal. In addition, most applications to engineering, economics, or statistics involve discrete-time systems where

all the components, expressed in terms of sums and differences, are defined on the same time sequence. However, a lot of people accepted the formulation of Miller and Ross as correct for the delta difference with corresponding results. In fact, similar situations are repeated in other formulations [84,92,95,98,104,106]. This has a very important consequence: it is impossible to define ARMA-type equations using these differences.

Example 1. To have an idea of the incorrectness of the delta difference definition, assume that $a = 0$ and $f(t) = e^{-t}$, $t \in \mathbb{Z}$. Let us compute three cases:

1. Order 1 difference

$$\Delta f(t) = e^{-t-1} - e^{-t} = (e^{-1} - 1)e^{-t}.$$

For a given t , the difference depends on the present and future values.

2. Order -1 difference

$$\Delta^{-1} f(t) = -e^{-t} - e^{-t-1} - e^{-t-2} \dots = -e^{-t} \sum_{n=0}^{\infty} e^{-n} = (e^{-1} - 1)^{-1} e^{-t}.$$

Again, for a given t , the -1 -order difference (sum) depends on the present and future values.

3. Order $-1/2$ difference

with $t = 1/2 + k$, we have

$$\begin{aligned} \Delta^{-1/2} f(t) &= \sum_{m=0}^k \binom{k-m-1/2}{k-m} e^{-m} \\ &= e^{-k} \sum_{m=0}^k \binom{n-1/2}{n} e^m \\ &= e^{-t+1/2} \sum_{m=0}^k (-1)^m \binom{-1/2}{m} e^m \\ &= e^{-(t+1/2)} + 1/2 e^{-(t-1/2)} - \frac{1}{8} e^{-(t-3/2)} \dots \end{aligned}$$

Contrarily to the above examples, the $-1/2$ -order derivative depends on one future value and infinite past values. Therefore, an operator that we were expected to be anti-causal is essentially causal.

4.2. One for All or One for Each

In some discrete-time formulations, as well as in the usual continuous-time fractional calculus, it is current to attach to a given difference/derivative definition the support of the function at hand. This means that for any function, there is a particular difference/derivative definition. It is a one to one case. This was the procedure used by Gray and Zhang. However, it creates difficulties, since we cannot add functions and differences/derivatives with different durations. This is strange because it is assumed that the differences/derivatives are linear operators.

Alternatively, we can define general differences/derivatives over the whole possible domain (\mathbb{R} or \mathbb{Z}) and only particularize at the moment of the computation. To understand the situation, consider two functions

$$f(n) = \begin{cases} 0 & n > a \\ n & a \leq n \leq b \\ 0 & n < b. \end{cases}$$

$$g(n) = n, \quad a \leq n \leq b.$$

At first glance, it looks like they are the same. However, they express different situations:

1. $f(n)$ expresses a situation where there is a past and a future. It is like some system that exists, is in stand-by first, acts for some time, and returns to the previous state. It is the situation corresponding to many physical, biological, and social systems.
2. $g(n)$, on the contrary, has no past and will have no future: something is born, lives for some time and disappears.

Assume that we want to define and compute a Z transform of such functions. In the first, we use the usual definition

$$F(z) = \sum_{n=-\infty}^{\infty} f(n)z^{-n}, \quad (38)$$

that gives

$$F(z) = \sum_{n=a}^b nz^{-n}.$$

On the other hand, for $g(n)$, we have to make a new definition

$${}_aG_b(z) = \sum_{n=a}^b nz^{-n}, \quad (39)$$

valid only for all the functions defined in the interval $a \leq n \leq b$. However, $F(z)$ and ${}_aG_b(z)$ are complex variable functions analytic in the same region and to which corresponds the same inversion integral that leads to the same inverse. In fact, the Cauchy integral

$$h(n) = \frac{1}{2\pi i} \oint_c H(z)z^{n-1} dz$$

defines a function $h(n)$ on \mathbb{Z} . This is interesting: the inverse transform imposes an extrapolation of $g(t)$ so that its domain is \mathbb{Z} , although its support is $a \leq n \leq b$. This means that we should use (38) as a definition of Z transform and consider functions of the type $f(n)$, even if its support is finite. Therefore, we do not need to specify the support in defining differences, derivatives, and transforms.

4.3. The Riemann–Liouville and Caputo-like Procedures

In many works and due to the influence of traditional fractional calculus, it has become commonplace to use Riemann–Liouville and Caputo-type procedures. These, instead of doing a direct difference computation, use a convolution between an integer-order difference and a fractional sum. However, while being mathematically correct, it is not a suitable way of doing computations, since

1. We are increasing unnecessarily the number of operations; this is very important in computational implementations, because we are increasing the numerical error [14,107];
2. We throw most of the computational burden on negative-order binomial coefficients that behave asymptotically like $\frac{1}{n^{-\alpha+1}}$, so decreasing very slowly or even increasing.

In continuous fractional calculus, the Riemann–Liouville integral becomes singular when the order is positive (derivative case), as Liouville recognized in his first paper [108], having proposed both procedures that transfer the singular behavior to the derivative of integer order. However, such a difficulty does not appear in discrete-time formulations, since the computational burden falls heavily on the binomial coefficients, directly or otherwise, which can be computed through the Pochhammer symbol that is always non-singular. Therefore, there is neither particular reason nor advantage in using those procedures.

5. Shift-Invariant Differencers and Accumulators

5.1. Causal

Consider two bounded piecewise continuous functions $f(t), g(t), t \in \mathbb{R}$ with $f(-\infty) = g(-\infty) = 0$. For simplicity, assume they are of exponential order so that we assure they have Laplace transforms, $F(s), G(s), s \in \mathbb{C}$, analytic over suitable ROC.

Definition 1. Let us define a nabla or causal differencer as a linear system whose output, at a given instant, is the difference between the input at that instant and at a previous one:

$$g(t) = f(t) - f(t - h), \tag{40}$$

where $h > 0$ is the delay.

The output will be called *nabla difference* and represented by $\nabla f(t)$. From this definition, we can draw some conclusions:

1. It is a moving-average system, which is sometimes called a “feedforward” system;
2. Its impulse response is given by:

$$\phi(t) = \delta(t) - \delta(t - h),$$

so that

$$g(t) = \phi(t) * f(t),$$

implying that

$$G(s) = (1 - e^{-sh})F(s). \tag{41}$$

3. The transfer function is

$$H(s) = 1 - e^{-sh}, \tag{42}$$

having \mathbb{C} as the ROC.

4. We can associate in series as many systems as we can in such a way that the output of the $(n - 1)$ -th system is the input of the next one, n -th

$$f_n(t) = g_{n-1}(t), \quad g_0(t) = f(t)$$

and

$$g_n(t) = f_n(t) - f_n(t - h) = g_{n-1}(t) - g_{n-1}(t - h).$$

The transfer function of the association is given by

$$G(s) = [1 - e^{-sh}]^n F(s),$$

that inverted gives the n -th order nabla difference

$$\nabla^n f(t) = \sum_{k=0}^n (-1)^k \binom{n}{k} f(t - kh). \tag{43}$$

Now, let us return back to (40) and invert the role of the functions: assume that the input is $g(t)$ and the output is $f(t)$ (feedback system)

$$f(t) = f(t - h) + g(t)$$

that can be reused to give:

$$f(t) = g(t) + g(t - h) + f(t - 2h) = g(t) + g(t - h) + g(t - 2h) + f(t - 3h) = \dots$$

It is not hard to see that

$$f(t) = \nabla^{-1}g(t) = \sum_{n=0}^{\infty} g(t - nh), \tag{44}$$

with LT

$$F(s) = G(s) \sum_{n=0}^{\infty} e^{-nsh} = [1 - e^{-sh}]^{-1} G(s), \quad \text{Re}(s) > 0. \tag{45}$$

Relation (44) shows why this operator is called an accumulator or sum. The series association of n equal accumulators gives the n -th order nabla sum:

$$\nabla^{-n}g(t) = \sum_{k=0}^{\infty} (-1)^k \binom{-n}{k} g(t - kh). \tag{46}$$

Definition 2. This result together with (43) suggests that the α -order nabla differencer/accumulator be given by

$$\nabla^{\alpha} f(t) = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t - kh), \tag{47}$$

where α is any real number.

The corresponding LT is

$$\mathcal{L}[\nabla^{\alpha} f(t)] = [1 - e^{-sh}]^{\alpha} F(s), \quad \text{Re}(s) > 0. \tag{48}$$

Therefore, there are some facts that we must emphasize:

1. A difference/sum is the output of a system: differencer/accumulator;
2. The system structure is independent of the inputs;
3. If the order is not a positive integer, even if the input function has finite support, the output has infinite support; in particular, if $f(t)$ has support $[a, b] \subset \mathbb{R}$, $g(t)$ is not identically null above any real value: the support is $[a, \infty]$. This is a very important fact that is frequently forgotten or dismissed.
4. If the input is a right-hand function, so is the output; in particular, if $f(t) = 0, t < 0$, then $\nabla^{\alpha} f(t) = 0$, for negative t and for $t \geq 0$, we have

$$\nabla^{\alpha} f(t) = \sum_{k=0}^n (-1)^k \binom{\alpha}{k} f(t - kh), \tag{49}$$

where $n = \lfloor t/h \rfloor$ is the the great integer such that $n \leq t/h$.

5. The ROC of the transfer function is defined by $\text{Re}(s) > 0$, as expected, since we are dealing with a causal system.

5.2. Anti-Causal

Consider the conditions of the previous subsection, but now with $f(\infty) = g(\infty) = 0$.

Definition 3. We define delta or anti-causal differencer as a linear system whose output, at a given instant, is the difference between the input at that instant and at a future one:

$$g(t) = f(t) - f(t + h), \tag{50}$$

where $h > 0$ is the lead.

The output will be called the *delta difference* and represented by $\Delta f(t)$. Note that it is the symmetric of the current definition (21). Traditionally, the symmetric of (50) is considered: $g(t) = f(t + h) - f(t)$. To simplify and highlight the similarity to the nabla difference, we found it preferable to omit the $-$ sign.

From this definition and the previous results, we conclude that:

1. It is also a moving-average system;
2. The LT gives

$$G(s) = (1 - e^{+sh})F(s). \tag{51}$$

3. The transfer function is

$$H(s) = 1 - e^{+sh}, \tag{52}$$

having \mathbb{C} as ROC.

4. The association in a series of n systems as above has a transfer function given by

$$G(s) = [1 - e^{sh}]^n F(s),$$

that inverted gives the n -th-order delta difference

$$\Delta^n f(t) = \sum_{k=0}^n (-1)^k \binom{n}{k} f(t + kh). \tag{53}$$

5. In a similar way, the delta accumulator is

$$\Delta^{-1}g(t) = \sum_{n=0}^{\infty} g(t + nh). \tag{54}$$

The series association of n accumulators is expressed by

$$\nabla^{-n}g(t) = \sum_{k=0}^{\infty} (-1)^k \binom{-n}{k} g(t - kh) \tag{55}$$

and has LT

$$\mathcal{L}[\Delta^{-n}g(t)] = [1 - e^{sh}]^{-n} G(s), \quad Re(s) < 0. \tag{56}$$

Definition 4. This result, together with (43), suggests that the α -order delta differencer/accumulator be defined by

$$\Delta^\alpha f(t) = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t + kh), \tag{57}$$

where α is any real number.

The corresponding LT is

$$\mathcal{L}[\Delta^\alpha f(t)] = [1 - e^{sh}]^\alpha F(s), \quad Re(s) < 0. \tag{58}$$

This and the nabla differences have similar characteristics. Therefore, and in particular, we can say relatively to the delta difference:

1. If the order is not a positive integer, even if the input function has finite support, the output has infinite support; in particular, if $f(t)$ has support $[a, b] \subset \mathbb{R}$, $g(t)$ is not identically null below any real value; the support is $[-\infty, b]$.
2. If the input is a left-hand function, so is the output; in particular, if $f(t) = 0, t > 0$, then $\Delta^\alpha f(t) = 0$, for positive t and for $t \leq 0$, we have

$$\Delta^\alpha f(t) = \sum_{k=0}^n (-1)^k \binom{\alpha}{k} f(kh - |t|), \tag{59}$$

where $n = \lfloor -t/h \rfloor$ is the the great integer such that $n \leq -t/h$.

3. The ROC of the transfer function is defined by $Re(s) < 0$, as expected, since we are dealing with an anti-causal system.
4. We can account for the $(-)$ sign we removed above by inserting the factor $(-1)^\alpha$ into (57).

5.3. Properties

The nabla and delta differencers have similar properties that we will describe for the first. The most important are [25]

- *Additivity and commutativity of the orders*
 $\nabla^\alpha [\nabla^\beta f(t)] = \nabla^\beta [\nabla^\alpha f(t)] = \nabla^{\alpha+\beta} f(t).$

- *Neutral element*

This comes from the last property by putting $\beta = -\alpha$, $\nabla^\alpha [\nabla^{-\alpha} f(t)] = \nabla^0 f(t) = f(t)$. This is very important because it states the existence of the inverse, which is in coherence with the previous sub-sections.

- *Inverse element*

From the last result, we conclude that there is always an inverse element: for every α -order difference, there is always a $-\alpha$ -order difference given by the same formula.

- *Associativity of the orders*

$$\nabla^\gamma [\nabla^\alpha \nabla^\beta] f(t) = \nabla^{\gamma+\alpha+\beta} f(t) = \nabla^{\alpha+\beta+\gamma} f(t) = \nabla^\alpha [\nabla^{\beta+\gamma}] f(t).$$

It is a consequence of the additivity.

- *Derivative of the product*

$$\nabla^\alpha [f(t)g(t)] = \sum_{m=0}^{\infty} \binom{\alpha}{m} \nabla^m f(t) \nabla^{\alpha-m} g(t - mh). \tag{60}$$

The delta case is slightly different as expected

$$\Delta^\alpha [f(t)g(t)] = \sum_{m=0}^{\infty} \binom{\alpha}{m} \Delta^m f(t) \Delta^{\alpha-m} g(t + mh).$$

5.4. Discrete-Time Differences

In the Introduction, we mentioned the importance of discrete-time signals. In Section 2.4, we showed how to obtain them by sampling continuous-time signals. However, we must highlight an important fact: discrete-time signals exist by themselves, without there needing to be a continuous-time signal from which they resulted. There are many signals that are inherently discrete. This means that in each case, there is necessarily a clock that defines the time sequence in which we work. Therefore, the delay/lead, h , must be equal or a multiple of the sampling interval used to obtain the discrete-time formulation for differences. For simplicity, we choose the sampling interval equal to h so that $\mathbb{T} = \{t = nh, n \in \mathbb{Z}\}$, giving

$$\nabla^\alpha f_k = \sum_{n=0}^{\infty} \frac{(-\alpha)_n}{n!} f_{k-n}, \quad k \in \mathbb{Z}, \tag{61}$$

and

$$\Delta^\alpha f_k = \sum_{n=0}^{\infty} \frac{(-\alpha)_n}{n!} f_{k+n}. \quad k \in \mathbb{Z}, \tag{62}$$

that can assume different forms:

$$\nabla^\alpha f_k = \sum_{n=-\infty}^k \frac{(-\alpha)_{k-n}}{(k-n)!} f_n, \quad k \in \mathbb{Z} \tag{63}$$

and

$$\Delta^\alpha f_k = \sum_{n=k}^{\infty} \frac{(-\alpha)_{n-k}}{(n-k)!} f_n, \quad k \in \mathbb{Z}. \tag{64}$$

As it is obvious, the Formulae (61)–(64) express input/output relations that are discrete-time convolutions between the input f_n and the binomial coefficients (impulse responses) to produce the outputs that are the differences. Such relations are mediated by the called impulse responses that are the outputs, h_n , when the input is the Kronecker delta defined by

$$\delta_n = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0. \end{cases}$$

In passing, we define the discrete-time step by

$$\varepsilon_n = \begin{cases} 1 & n \geq 0 \\ 0 & n < 0. \end{cases}$$

Therefore, the impulse responses of the nabla and delta differences are, respectively,

$$h_n = \frac{(-\alpha)_n}{n!} \varepsilon_n$$

and

$$g_n = \frac{(-\alpha)_{-n}}{(-n)!} \varepsilon_{-n},$$

where $n \in \mathbb{Z}$, which is in agreement with the causality. It is important to remark that

1. Both responses have finite duration if $\alpha \in \mathbb{N}$, and the systems are called FIRs (finite impulse systems) [14].
2. If $\alpha \notin \mathbb{N}$, both responses extend to infinite, and the corresponding systems are IIR (infinite impulse response).
3. If $f_k = 0, k < 0$, then $\nabla^\alpha f_k = 0$, for negative k , and for $k \geq 0$, we have

$$\nabla^\alpha f_k = \sum_{n=0}^k \frac{(-\alpha)_n}{n!} f_{k-n}, \quad k \in \mathbb{Z} \tag{65}$$

or,

$$\nabla^\alpha f_k = \sum_{n=0}^k \frac{(-\alpha)_{k-n}}{(k-n)!} f_n, \quad k \in \mathbb{Z}. \tag{66}$$

4. Similarly, if $f_k = 0, k > 0$, then $\Delta^\alpha f_k = 0$ for positive k and we obtain for $k \leq 0$

$$\Delta^\alpha f_k = \sum_{n=0}^{|k|} \frac{(-\alpha)_n}{n!} f_{n-|k|}. \quad k \in \mathbb{Z} \tag{67}$$

or

$$\Delta^\alpha f_k = \sum_{n=0}^{|k|} \frac{(-\alpha)_{n-k}}{(n-k)!} f_n, \quad k \in \mathbb{Z}. \tag{68}$$

5. It is a simple task to obtain formulae for functions with other supports.
6. The Z transforms of the above discrete-time differences can be obtained from the corresponding LT by setting $z = e^{sh}$. For example, the Z transform of the nabla difference (61) is

$$\mathcal{Z}[\nabla^\alpha f_k] = (1 - z^{-1})^\alpha F(z), \tag{69}$$

in the suitable ROC in the region defined by $|z| > 1$.

7. If, in any particular application, a time sequence with the form $\mathbb{T} = a + nh, ; n \in \mathbb{Z}, a \in \mathbb{R}$ is used, we can make a substitution $f(a + nh)$ for $f(nh)$.

5.5. Two-Sided Differences

There is another differencer, two-sided (bilateral), which is given by

$$\Theta_0^1 f(t) = f(t + h/2) - f(t - h/2), \tag{70}$$

that originates two bilateral fractional differences, but we will not study both here [68,69,109]. We can easily obtain one of them by combining a delta with a nabla difference.

Definition 5. Let $t = kh, k \in \mathbb{Z}$. We define a bilateral differencer through

$$\Theta_\theta^\beta f_k = \nabla^a \Delta^b f_k = \sum_{n=-\infty}^{+\infty} \frac{(-1)^n \Gamma(\beta + 1)}{\Gamma\left(\frac{\beta+\theta}{2} - n + 1\right) \Gamma\left(\frac{\beta-\theta}{2} + n + 1\right)} f_{k-n}, \tag{71}$$

with $\beta = a + b$ (difference order) and $\theta = a - b$ (asymmetry parameter).

For $\beta \in \mathbb{Z}^-$, we obtain singular cases that were treated in [69]. Suitable choices of these paramters allow us to recover the above introduced differences:

- With $\theta = \beta$, we obtain (47);
- Similarly, with $\theta = -\beta$ and variable change, we obtain (57);

Some particular cases of (71) are very interesting:

1. Riesz-type difference, $\theta = 0$,

$$\Theta_0^\beta f_k = \sum_{n=-\infty}^{+\infty} \frac{(-1)^n \Gamma(\beta + 1)}{\Gamma\left(\frac{\beta}{2} - n + 1\right) \Gamma\left(\frac{\beta}{2} + n + 1\right)} f_{k-n}. \tag{72}$$

2. Feller-type difference, $\theta = \pm 1$,

$$\Theta_1^\beta f_k = \sum_{n=-\infty}^{+\infty} \frac{(-1)^n \Gamma(\beta + 1)}{\Gamma\left(\frac{\beta+1}{2} - n + 1\right) \Gamma\left(\frac{\beta-1}{2} + n + 1\right)} f_{k-n}. \tag{73}$$

3. Two-sided discrete Hilbert transform, $\beta = 0$,

$$\Theta_\theta^0 f_k = \sum_{n=-\infty}^{+\infty} (-1)^n \cdot \frac{1}{\Gamma\left(\frac{\theta}{2} - n + 1\right) \Gamma\left(\frac{-\theta}{2} + n + 1\right)} f_{k-n}. \tag{74}$$

With $\theta = 1$, we obtain the usual discrete-time formulation of the Hilbert transform [14].

Remark 5. It must be noticed that in the general case stated in (71), the difference $\Theta_\theta^\beta f_k$ does not have Z transform, since the ROC degenerates in the unit circle ($z = e^{i\omega}$), meaning that it has discrete-time Fourier transform, $F(e^{i\omega})$ [68,109]:

$$\mathcal{F}\left[\Theta_\theta^\beta f_k\right] = |2 \sin(\omega/2)|^\beta e^{-i\omega\theta} e^{i \operatorname{sgn}(\omega)\theta\pi/2} F(e^{i\omega}), \quad |\omega| < \pi, \tag{75}$$

where $\operatorname{sgn}(\cdot)$ is the signum function.

5.6. The Tempered Differences

We introduced above the three standard definitions for the differences of order one (40), (50) and (70). With a slight modification, we obtain their tempered versions. We only need to make adaptations of the results described in [110]. The concept of a tempered fractional difference was introduced, for the first time, by Sabzikar et al. [111], starting from the Grünwald–Letnikov derivative. Their results are similar to those that we are going to present.

Definition 6. Let $\lambda \in \mathbb{R}$. We can define three tempered differences (TDs):

- Nabla TD

$$\nabla_{\lambda, f}(t) = f(t) - e^{-\lambda h} f(t - h), \tag{76}$$

that has LT

$$\mathcal{L}[\nabla_{\lambda} f(t)] = [1 - e^{-\lambda h} e^{-sh}] F(s), \tag{77}$$

for any $s \in \mathbb{C}$.

- Delta TD

$$\Delta_{\lambda} f(t) = f(t) - e^{\lambda h} f(t + h). \tag{78}$$

Its LT is valid for any $s \in \mathbb{C}$ and given by

$$\mathcal{L}[\Delta_{\lambda} f(t)] = [1 - e^{\lambda h} e^{sh}] F(s), \tag{79}$$

As above, we removed a (-) sign.

- Two-sided TD

$$\Theta_{\theta, \lambda} f(t) = \left[e^{\lambda \frac{h}{2}} f\left(t + \frac{h}{2}\right) - e^{-\lambda \frac{h}{2}} f\left(t - \frac{h}{2}\right) \right] \tag{80}$$

It is straightforward to invert the relations (76) and (78), and so, we obtain the first-order anti-differences

$$\nabla_{\lambda}^{-1} f(t) = \sum_{n=0}^{\infty} e^{-n\lambda h} f(t - nh) \tag{81}$$

and

$$\Delta_{\lambda}^{-1} f(t) = \sum_{n=0}^{\infty} e^{n\lambda h} f(t + nh) \tag{82}$$

that can be generalized for any real order.

Definition 7. For $\alpha \in \mathbb{R}$, we can write [110]

$$\nabla_{\lambda}^{\alpha} f(t) = \sum_{n=0}^{\infty} \frac{(-\alpha)_n}{n!} e^{-n\lambda h} f(t - nh), \tag{83}$$

and

$$\Delta_{\lambda}^{\alpha} f(t) = \sum_{n=0}^{\infty} \frac{(-\alpha)_n}{n!} e^{n\lambda h} f(t + nh). \tag{84}$$

Definition 8. To obtain the discrete-time versions, we set $t = nh$ and $\mu = e^{\lambda h}$ so that we can write [110]

$$\nabla_{\lambda}^{\alpha} f_k = \sum_{n=0}^{\infty} \frac{(-\alpha)_n}{n!} \mu^{-n} f_{k-n}, \tag{85}$$

and

$$\Delta_{\lambda}^{\alpha} f_k = \sum_{n=0}^{\infty} \frac{(-\alpha)_n}{n!} \mu^n f_{k+n}, \tag{86}$$

valid for any $\alpha \in \mathbb{R}$.

The bilateral tempered fractional differences are somehow more involved. They can be obtained from the results introduced in [112]. We will not do it here.

5.7. Bilinear Differences

One of the most interesting methods of obtaining discrete-time systems from continuous-time templates is through the conformal transformations, so that we pass from the LT context, (s) , to the one of Z transform, (z) [28]. The traditional procedures can be described by

$$s = 1 - z^{-1},$$

for the nabla case and

$$s = z - 1,$$

for the delta case. Basically, z^{-1} and z correspond to delay and lead, respectively. This procedure is equivalent to the one described in Section 5.4. It has a great drawback: the imaginary axis in the s plane is transformed into the Hilger circles $|z - 1| = 1$ and $|z + 1| = 1$ [25]. This fact brings some inconvenience [28].

Another alternative to obtain continuous to discrete conversion is through the bilinear (Tustin) transformation [28,113]

$$s = \frac{2}{h} \frac{1 - z^{-1}}{1 + z^{-1}},$$

which leads to discrete-time formulations similar to the GL-like above obtained. In [114], new discrete-time fractional derivatives based on the bilinear transformation were introduced and studied. In agreement with such a theory, we propose new differences here.

Definition 9. Let us define a bilinear nabla differencer as a linear system whose output, at a given instant, is given by

$$g(t) + g(t - h) = f(t) - f(t - h) \quad (87)$$

and the bilinear delta differencer by

$$g(t) + g(t + h) = f(t) - f(t + h). \quad (88)$$

In terms of the LT, we can write, respectively,

$$G(s) = \frac{1 - e^{-sh}}{1 + e^{-sh}} F(s) \quad (89)$$

and

$$G(s) = \frac{1 - e^{sh}}{1 + e^{sh}} F(s). \quad (90)$$

It is obvious that we can formulate symmetric bilinear differences from the LT

$$G(s) = \frac{e^{sh/2} - e^{-sh/2}}{e^{sh/2} + e^{-sh/2}} F(s). \quad (91)$$

To avoid any repetition of concepts, we will consider the nabla case only. The above formulae suggest immediately how we make a continuous-to-discrete conversion by using the substitution $e^{sh} \rightarrow z$.

Definition 10. We define the bilinear nabla fractional difference through

$$\mathcal{Z}[\nabla_b^\alpha f_k] = \left[\frac{1 - z^{-1}}{1 + z^{-1}} \right]^\alpha F(z), \quad |z| > 1. \quad (92)$$

Attending to the results in [114], we can write

$$\nabla_b^\alpha f_n = \sum_{k=0}^{\infty} \psi_k^\alpha f_{n-k}, \quad n \in \mathbb{Z}, \quad (93)$$

where [114]

$$\psi_k^\alpha = \sum_{m=0}^k \frac{(-\alpha)_m}{m!} \frac{(-1)^{k-m} (\alpha)_{k-m}}{(k-m)!}, k \in \mathbb{Z}_0^+ \tag{94}$$

In particular, for $\alpha = \pm 1$, we have

$$\begin{aligned} \bullet \psi_k^1 &= \begin{cases} 0 & k < 0 \\ 1 & k = 0 \\ 2(-1)^k & k > 0 \end{cases} \\ \bullet \psi_k^{-1} &= \begin{cases} 0 & k < 0 \\ 1 & k = 0 \\ 2 & k > 0 \end{cases} \end{aligned}$$

Although these new differences seem to be rather involved, they are easily implemented with the help of the fast Fourier transform (FFT) [114].

We can also obtain two-sided bilinear differences, but the procedure is rather involved and not very useful. However, the frequency domain representation is easily obtained from (91).

6. Scale-Invariant Differences

In previous sections, we dealt preferably with the shift-invariant differences. Here, we will consider others that have deep relations with the scale-invariant derivatives [40].

Consider two bounded piecewise continuous functions $f(t), g(t)$, $t \in \mathbb{R}^+$ with $f(0), g(0) = 0$. For simplicity, assume they are of polynomial order, so that we assure they have Mellin transforms, $F(v), G(v)$, $v \in \mathbb{C}$, analytic over suitable ROC.

Definition 11. Let $q > 1$. We define a stretching differencer as a linear system whose output, at a given scale, is the difference between the input at different scales:

$$g(\tau) = f(\tau) - f(\tau q^{-1}), \quad \tau \in \mathbb{R}^+, \tag{95}$$

where q is the scale constant.

The output will be called the *stretching difference* and represented by $\triangleleft f(t)$. Letting $q < 1$, we obtain the *shrinking difference*, $\triangleright f(t)$. Therefore, their properties are similar. We will study the first only, because the other is easy to obtain. From this definition, we can draw some conclusions:

1. Its impulse response is given by:

$$\phi(\tau) = \delta(\tau - 1) - \delta(\tau/q - 1),$$

so that

$$g(\tau) = \phi(\tau) \star f(\tau),$$

implying that

$$G(v) = (1 - q^{-v})F(v). \tag{96}$$

2. The transfer function is

$$H(v) = 1 - q^{-v}, \tag{97}$$

having \mathbb{C} as the ROC.

3. As in the shift-invariant case, if associated in series n systems, the resulting system defines the n -th order stretching difference that has a transfer function given by

$$H(v) = [1 - q^{-v}]^n,$$

from which we obtain the n -th order stretching difference

$$\triangleleft^n f(\tau) = \sum_{k=0}^n (-1)^k \binom{n}{k} f(\tau q^{-k}). \tag{98}$$

Now, let us return back to (40) and invert the role of the functions: assume that the input is $g(\tau)$ and the output is $f(\tau)$

$$f(\tau) = f(\tau q^{-1}) + g(\tau),$$

that can be reused to give:

$$f(\tau) = g(\tau) + g(\tau q^{-1}) + f(\tau q^{-2}) = g(\tau) + g(\tau q^{-1}) + g(\tau q^{-2}) + f(\tau q^{-3}) = \dots$$

It is not hard to see that

$$f(\tau) = \triangleleft^{-1} g(\tau) = \sum_{n=0}^{\infty} g(\tau q^{-n}), \tag{99}$$

with MT

$$F(v) = G(v) \sum_{n=0}^{\infty} q^{-nv} = [1 - q^{-v}]^{-1} G(v), \quad \text{for } \text{Re}(v) > 0. \tag{100}$$

Relation (99) shows why this operator is again an accumulator. The series association of n equal accumulators gives the n -th order stretching sum:

$$\triangleleft^{-n} g(\tau) = \sum_{k=0}^{\infty} (-1)^k \binom{-n}{k} g(\tau q^{-k}). \tag{101}$$

Definition 12. This result together with (98) suggests that the α -order stretching difference/accumulator must be given by

$$\triangleleft^\alpha f(\tau) = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(\tau q^{-k}), \tag{102}$$

where α is any real number.

As we observe, this difference uses an exponential domain in agreement with our considerations above (Section 2.3). The corresponding MT is

$$\mathcal{M}[\triangleleft^\alpha f(\tau)] = [1 - q^{-v}]^\alpha F(v), \quad \text{for } \text{Re}(v) > 0. \tag{103}$$

Therefore, the shrinking difference is defined by

$$\triangleright^\alpha f(\tau) = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(\tau q^k), \tag{104}$$

having MT

$$\mathcal{M}[\triangleright^\alpha f(\tau)] = [1 - q^v]^\alpha F(v), \quad \text{for } \text{Re}(v) < 0. \tag{105}$$

Remark 6. The relations (100) and (105) show that there is a scale-invariant system that produces the difference as an output.

To obtain the discrete-scale versions, we only need to make a sampling in agreement with Theorem 2. Let $\tau = q^n$. For the stretching difference, we obtain

$$\triangleleft^\alpha f_{q^n} = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f_{q^{n-k}}, \tag{106}$$

while for the shrinking one, it is

$$\triangleright^\alpha f_{q^n} = \sum_{k=0}^\infty (-1)^k \binom{\alpha}{k} f_{q^{n+k}}. \tag{107}$$

As we can see, they are formally similar to the nabla and delta differences; they only differ in the sampling sequence used: linear or exponential. Thus, from a purely discrete point of view, we have no way of making any distinction between linearly or exponentially sampled functions. This means that if we want to define stretching and shrinking differences in discrete time, we have to break the continuous connection: we have to work exclusively with the sequences defined in \mathbb{Z} . So, having a sequence $f_n, n \in \mathbb{Z}$, we wonder how to define stretched and shrunk sequences so that we can introduce differences. In [115], ways to produce stretched and shrunk sequences were presented. In principle, we could use them to define differences, but this procedure has difficulties, since the operations of stretching and shrinking involve all knowledge of the underlying sequence, and the scale transformation system is two-sided.

However, we can use the traditional “decimation” operation used in signal processing to define a stretching difference [5,8,14].

Definition 13. Let N be a positive integer (decimation parameter). We define a stretching difference through

$$\triangleleft_N f_n = f_n - f_{Nn}. \tag{108}$$

We can show immediately that if M is a positive integer, then

$$\triangleleft_N^M f_n = \sum_{m=0}^M (-1)^m \binom{M}{m} f_{N^m n}. \tag{109}$$

If $f_n = n^a, n \in \mathbb{N}, a \in \mathbb{R}$, then

$$\triangleleft_N^M n^a = \sum_{m=0}^M (-1)^m \binom{M}{m} N^{am} n^a = [1 - N^a]^M n^a. \tag{110}$$

Proceeding as above, we obtain

$$\triangleleft_N^{-M} f_n = \sum_{m=0}^M \frac{(M)_m}{m!} f_{N^m n}, \tag{111}$$

that allows us to write

$$\triangleleft_N^M n^a = \sum_{m=0}^M (-1)^m \frac{(M)_m}{m!} N^{am} n^a = [1 - N^a]^{-M} n^a. \tag{112}$$

Definition 14. These relations suggest we write

$$\triangleleft_N^\alpha f_n = \sum_{m=0}^\infty \frac{(-\alpha)_m}{m!} f_{N^m n}, \tag{113}$$

so that

$$\triangleleft_N^\alpha n^a = \sum_{m=0}^\infty \frac{(-\alpha)_m}{m!} N^{am} n^a = [1 - N^a]^\alpha n^a. \tag{114}$$

This last relation seems to point to a definition of a “discrete Mellin transform”, as

$$F(v) = \sum_{k=1}^\infty f_k k^v, \quad v \in \mathbb{C},$$

which is different from other proposed in the past [33,34]. We do not go further in this way.

The properties of the stretching discrete difference just proposed are readily deduced from the results in Section 5.3. From such a definition, the reason for not defining a shrinking difference is evident.

7. The ARMA-Type Difference Linear Systems

Definition 15. We define an ARMA-type difference linear system through the following equation [25,28]

$$\sum_{k=0}^N a_k \mathcal{D}^{\alpha_k} y(t) = \sum_{k=0}^M b_k \mathcal{D}^{\beta_k} x(t), \tag{115}$$

where a_k and b_k ($k = 0, 1, \dots$) with $a_N = 1$ are real numbers. The operator \mathcal{D} is any difference defined previously. The orders N and M are any positive integers. The positive real numbers α_k and β_k with $k = 0, 1, \dots$, form strictly increasing sequences.

The most interesting systems are those with commensurate orders:

$$\sum_{k=0}^N a_k \mathcal{D}^{k\alpha} y(t) = \sum_{k=0}^M b_k \mathcal{D}^{k\alpha} x(t). \tag{116}$$

In the shift-invariant cases, the exponentials are the eigenfunctions and the eigenvalues are the transfer functions [27,36,116]. In the scale-invariant system, such a role is played by the powers [40]. The corresponding (Laplace, Z, Mellin) transforms give the *impulse response* or Green function of the system.

Example 2. Consider the ARMA(1,1) system:

$$\nabla^\alpha y(t) + ay(t) = b_1 \nabla^\alpha x(t) + b_0 x(t)$$

and let $x(t) = e^{st}$, $t \in \mathbb{R}$. The output will be $y(t) = H(s)e^{st}$, $t \in \mathbb{R}, c \in \mathbb{C}$, with

$$H(s) = \frac{b_1(1 - e^{-sh})^\alpha + b_0}{(1 - e^{-sh})^\alpha + a}, \quad \text{Re}(s) > 0.$$

The impulse response is not easily obtained. From

$$\begin{aligned} H(s) &= \frac{b_1(1 - e^{-sh})^\alpha + b_0 + b_1a - b_1a}{(1 - e^{-sh})^\alpha + a} \\ &= b_1 + \frac{b_0 - b_1a}{(1 - e^{-sh})^\alpha + a} = b_1 + (b_0 - b_1a) \sum_{m=0}^{\infty} a^m (1 - e^{-sh})^{-(m+1)\alpha}, \\ &= b_1 + (b_0 - b_1a) \sum_{m=0}^{\infty} a^m \sum_{k=0}^{\infty} \frac{((m+1)\alpha)_k}{k!} e^{-ksh}. \end{aligned}$$

we get

$$h(t) = b_1 \delta(t) + (b_0 - b_1a) \sum_{m=0}^{\infty} a^m \sum_{k=0}^{\infty} \frac{((m+1)\alpha)_k}{k!} \delta(t - kh).$$

The output for any $x(t)$ is given by the usual convolution

$$y(t) = b_1 x(t) + (b_0 - b_1a) \sum_{m=0}^{\infty} a^m \sum_{k=0}^{\infty} \frac{((m+1)\alpha)_k}{k!} x(t - kh).$$

In the discrete-time case, we set $t = nh$, $n \in \mathbb{Z}$ to give

$$\nabla^\alpha y_n + ay_n = b_1 \nabla^\alpha x_n + b_0 x_n.$$

If the input is the exponential $z^n, n \in \mathbb{Z}, z \in \mathbb{C}$, the output is $y_n = H(z)z^n, n \in \mathbb{Z}, z \in \mathbb{C}$, where

$$H(z) = \frac{b_1(1 - z^{-1})^\alpha + b_0}{(1 - z^{-1})^\alpha + a}, \quad |z| > 1.$$

We can obtain the Z transform inverse of this function approximately through the FFT. However, proceeding as above, we obtain

$$h_n = b_1\delta_n + (b_0 - b_1a) \sum_{m=0}^{\infty} a^m \sum_{k=0}^{\infty} \frac{((m + 1)\alpha)_k}{k!} \delta_{n-k}.$$

As in the continuous-time case, the output corresponding to a given input is obtained through the discrete convolution

$$y_n = b_1x_n + (b_0 - b_1a) \sum_{m=0}^{\infty} a^m \sum_{k=0}^{\infty} \frac{((m + 1)\alpha)_k}{k!} x_{n-k}.$$

Example 3. Consider the scale-invariant AR(1) difference system defined by the equation

$$\triangleright_3^{1/2}y_n = \triangleright_3^{1/2}x_n - x_n$$

and let $x_n = n^{-2}, n \in \mathbb{Z}_0^-$.

Accordingly to what we wrote in the last section, the solution is, for $n \in \mathbb{N}$,

$$y_n = x_n - \sum_{m=0}^{\infty} \frac{(1/2)_m}{m!} x_{3^m n}.$$

So,

$$y_n = n^{-2} - \sum_{m=0}^{\infty} \frac{(1/2)_m}{m!} 3^{-2m} n^{-2} = n^{-2} [1 - (1 - 1/9)^{-1/2}] = \frac{2\sqrt{2} - 3}{2\sqrt{2}} n^{-2}.$$

8. Which Difference?

We described the most interesting differences. We ask ourselves which one should be used in a particular application. In some uses, there is doubt. For example, in most engineering applications, such as digital signal processing, which is the basis of many electrical, mechanical, or biomedical systems, the nabla derivative is the appropriate one. In fact, these systems are causal. The bilinear nabla systems are similar, but they have one important advantage: they are suitable for frequency domain implementations with the fast Fourier transform. The reduction in the number of operations should be a goal in applications. If the independent variable is not time, but any other, such as space, the delta or bilateral derivative can also be used. Although often employed, the delta difference is not suitable for most applications; its anti-causal characteristic gives rise to unstable systems.

In dealing with stochastic processes, such as the fractional Brownian motion or other generalized discrete fractional processes, the nabla or the two-sided can be used [117–119].

The integer-order differences and systems are characterized by having short memory and an exponentially decreasing impulse response, which is unsuitable for many applications. On the other hand, the fractional systems have long memory, since they have power decrease impulse responses, which brings too big of an influence of the past into the present. The tempered systems offer an intermediate solution. In engineering, the frequency domain analysis helps in deciding which system is more suitable. This fact makes the area of filter design one of the most important in applications.

The scale-invariant differences and systems, considered separately from the shift-invariant analogues, only recently were introduced. The corresponding discrete versions were proposed above for the first time. Therefore, there are no prescribed applications.

9. Discussion

Differences are basic building blocks for defining derivatives, but they can be used in many applications to solve differential equations and model many systems. In most situations, shift-invariant differences are used, although scale-invariant versions are also useful. Here, they have been studied separately. General continuous-time cases have been introduced, although the main interest has been placed on the discrete versions.

When looking for a discrete-time equivalent to continuous-time differential systems, there is a difference system that works as an intermediate [28]. Normally, such a system is not used for anything else. Discrete differences and systems are fundamental in computational applications that are the primary means for practical implementations. For this reason, when introducing differences, we adopted a system point of view to emphasize that differences are outputs of linear systems, which implies that they are defined independently of the inputs, notably their duration. Moreover, outputs are generally of infinite duration, even if the input support is finite. In addition to the classic differences, we introduce new ones such as bilateral and tempered differences.

The option by system approach for introducing differences allowed us to define ARMA-type linear systems, enlarging the classic procedure used in time-series analysis and processing which supported many important applications in engineering, economics, statistics, and so on. It is important to remark that many interesting functions we find in applications are acquired under a discrete-time form without having any analytical formula. This implies that we have to deal with functions (signals) defined in the set of integers. Anyway, implicit in any application, there is a time sequence imposed by an underlying clock which imposes a working domain that we cannot change. This aspect was frequently dismissed in the past, which is a fact that led to some “abnormalities” such as the loss of (anti-)causality. This happened, for example, with the assumed “delta difference” that is not really a delta difference, since it should be anti-causal, but it is bilateral. This fact is expected to make a review of some associated concepts and tools.

The theory we have just described leaves open issues that deserve consideration in future research. In fact, we have not considered the modeling/identification problems [16–20]. To do so, the frequency-domain approach will certainly be a safe way forward. This can be applied by adapting and extending the studies in [25,114] by defining suitable transforms. It is important to highlight the fact that such transforms use the eigenfunctions of the differences as kernels. Another open issue is the interaction between the ARMA difference systems above introduced and the stochastic processes, as referred above. This will imply the search for estimation methods that may involve spectral estimation.

Funding: The author was partially funded by National Funds through the Foundation for Science and Technology of Portugal under the projects UIDB/00066/2020.

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

References

1. Kolmogoroff, A. Interpolation und Extrapolation von stationären zufälligen Folgen. *Bull. Acad. Sci. URSS Math. [Izvestia Akad. Nauk. SSSR]* **1941**, *5*, 3–14.
2. Wiener, N. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications*; MIT Press: Cambridge, MA, USA, 1949; Volume 113.
3. Jenkins, G.M.; Priestley, M. The spectral analysis of time-series. *J. R. Stat. Soc. Ser. B (Methodol.)* **1957**, *19*, 1–12. [[CrossRef](#)]
4. Box, G.; Jenkins, G. *Time Series Analysis: Forecasting and Control*; Holden-Day: San Francisco, CA, USA, 1970.
5. Oppenheim, A.V.; Schaffer, R.W. *Discrete-Time Signal Processing*, 3rd ed.; Prentice Hall Press: Upper Saddle River, NJ, USA, 2009.
6. Kailath, T. *Linear Systems*; Information and System Sciences Series; Prentice-Hall: Upper Saddle River, NJ, USA, 1980.
7. Kailath, T. Lectures on Wiener and Kalman filtering. In *Lectures on Wiener and Kalman Filtering*; Springer: Vienna, Austria, 1981; pp. 1–143.
8. Rabiner, L.R.; Gold, B. *Theory and Application of Digital Signal Processing*; Prentice-Hall: Englewood Cliffs, NJ, USA, 1975.
9. Jury, E.I. *Analysis and Synthesis of Sampled-Data Control Systems*; Columbia University: New York, NY, USA, 1953.
10. Pollock, D.S.G.; Green, R.C.; Nguyen, T. *Handbook of Time Series Analysis, Signal Processing, and Dynamics*; Elsevier: Amsterdam, The Netherlands, 1999.

11. Robinson, E.A.; Treitel, S. *Geophysical Signal Analysis*; Society of Exploration Geophysicists: Tulsa, OK, USA, 2000.
12. Papoulis, A. *Signal Analysis*; McGraw-Hill: New York, NY, USA, 1977; pp. 1–435.
13. Ifeachor, E.C.; Jervis, B.W. *Digital Signal Processing: A Practical Approach*; Pearson Education: Harlow, Essex, England, 2002.
14. Proakis, J.G.; Manolakis, D.G. *Digital Signal Processing: Principles, Algorithms, and Applications*; Prentice Hall: Upper Saddle River, NJ, USA, 2007.
15. Brockwell, P.J.; Davis, R.A. *Time Series: Theory and Methods*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.
16. Neuman, C.P. Properties of the delta operator model of dynamic physical systems. *IEEE Trans. Syst. Man Cybern.* **1993**, *23*, 296–301. [[CrossRef](#)]
17. Premaratne, K.; Salvi, R.; Habib, N.; LeGall, J. Delta-operator formulated discrete-time approximations of continuous-time systems. *IEEE Trans. Autom. Control* **1994**, *39*, 581–585. [[CrossRef](#)]
18. Poor, H.V. Delta-operator based signal processing: Fast algorithms for rapidly sampled data. In Proceedings of the 36th IEEE Conference on Decision and Control, San Diego, CA, USA, 10–12 December 1997; Volume 1, pp. 872–877.
19. Gessing, R. Identification of shift and delta operator models for small sampling periods. In Proceedings of the 1999 IEEE American Control Conference (Cat. No. 99CH36251), San Diego, CA, USA, 2–4 June 1999; Volume 1, pp. 346–350.
20. Fan, H.; Liu, X. Delta Levinson and Schur-type RLS algorithms for adaptive signal processing. *IEEE Trans. Signal Process.* **1994**, *42*, 1629–1639. [[CrossRef](#)]
21. Ortigueira, M.D. Introduction to fractional linear systems. Part 2. Discrete-time case. *IEE Proc. Vis. Image Signal Process.* **2000**, *147*, 71–78. :20000273. [[CrossRef](#)]
22. Goodrich, C.; Peterson, A.C. *Discrete Fractional Calculus*; Springer International Publishing AG: Cham, Switzerland, 2015.
23. Tarasov, V.E. Exact discrete analogs of derivatives of integer orders: Differences as infinite series. *J. Math.* **2015**, *2015*, 134842. [[CrossRef](#)]
24. Tarasov, V.E. Lattice fractional calculus. *Appl. Math. Comput.* **2015**, *257*, 12–33. [[CrossRef](#)]
25. Ortigueira, M.D.; Coito, F.J.V.; Trujillo, J.J. Discrete-time differential systems. *Signal Process.* **2015**, *107*, 198–217. [[CrossRef](#)]
26. El-Khazali, R.; Machado, J.T. Closed-Form Discretization of Fractional-Order Differential and Integral Operators. In Proceedings of the Fractional Calculus: ICFDA 2018, Amman, Jordan, 16–18 July 2018; Springer: Singapore, 2019; pp. 1–17.
27. Ortigueira, M.D.; Machado, J.T. The 21st century systems: An updated vision of discrete-time fractional models. *IEEE Circuits Syst. Mag.* **2022**, *22*, 6–21. [[CrossRef](#)]
28. Ortigueira, M.D.; Magin, R.L. On the Equivalence between Integer-and Fractional Order-Models of Continuous-Time and Discrete-Time ARMA Systems. *Fractal Fract.* **2022**, *6*, 242. [[CrossRef](#)]
29. Butzer, P.; Engels, W.; Ries, S.; Stens, R. The Shannon sampling series and the reconstruction of signals in terms of linear, quadratic and cubic splines. *SIAM J. Appl. Math.* **1986**, *46*, 299–323. [[CrossRef](#)]
30. Gensun, F. Whittaker–Kotelnikov–Shannon sampling theorem and aliasing error. *J. Approx. Theory* **1996**, *85*, 115–131. [[CrossRef](#)]
31. Unser, M. Sampling-50 years after Shannon. *Proc. IEEE* **2000**, *88*, 569–587. [[CrossRef](#)]
32. Marvasti, F. *Nonuniform Sampling: Theory and Practice*; Springer Science & Business Media: New York, NY, USA, 2012.
33. Bertrand, J.; Bertrand, P.; Ovarlez, J. The Mellin Transform. In *The Transforms and Applications Handbook*, 2nd ed.; Poularikas, A.D., Ed.; CRC Press: Boca Raton, FL, USA, 2000.
34. De Sena, A.; Rocchesso, D. A fast Mellin and scale transform. *EURASIP J. Adv. Signal Process.* **2007**, *2007*, 89170. [[CrossRef](#)]
35. Ortigueira, M.D.; Machado, J.A.T. Fractional Derivatives: The Perspective of System Theory. *Mathematics* **2019**, *7*, 150. [[CrossRef](#)]
36. Ortigueira, M.D.; Valério, D. *Fractional Signals and Systems*; De Gruyter: Berlin, Germany; Boston, MA, USA, 2020.
37. Oppenheim, A.V.; Willsky, A.S.; Hamid, S. *Signals and Systems*, 2nd ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 1997.
38. Shmaliy, Y. *Continuous-Time Systems*; Springer: Dordrecht, The Netherlands, 2007.
39. Gulgowski, J.; Stefański, T.P. Generalization of Kramers–Kronig relations for evaluation of causality in power-law media. *Commun. Nonlinear Sci. Numer. Simul.* **2021**, *95*, 105664. [[CrossRef](#)]
40. Ortigueira, M.D.; Bohannan, G.W. Fractional Scale Calculus: Hadamard vs. Liouville. *Fractal Fract.* **2023**, *7*, 296. [[CrossRef](#)]
41. Lacroix, S.F. *Traité des Différences et des Séries*; Duprat: Paris, France, 1800.
42. Householder, A.S. *Principles of Numerical Analysis*; McGraw-Hill Book Company: New York, NY, USA, 1953.
43. Hardy, G.H. *Divergent Series*; American Mathematical Soc.: Ann Arbor, MI, USA, 2000; Volume 334.
44. Samko, S.G.; Kilbas, A.A.; Marichev, O.I. *Fractional Integrals and Derivatives*; Gordon and Breach: Yverdon, Switzerland, 1993.
45. Ortigueira, M.D.; Machado, J.T. Revisiting the 1D and 2D Laplace transforms. *Mathematics* **2020**, *8*, 1330. [[CrossRef](#)]
46. Aulbach, B.; Hilger, S. A unified approach to continuous and discrete dynamics. *Qualitative Theory of Differential Equations. In Colloquia Mathematica Societatis János Bolyai*; North-Holland: Amsterdam, The Netherlands, 1990; Volume 53, pp. 37–56.
47. Hilger, S. Analysis on Measure Chains—A Unified Approach to Continuous and Discrete Calculus. *Results Math.* **1990**, *18*, 18–56. [[CrossRef](#)]
48. Ortigueira, M.D.; Torres, D.F.; Trujillo, J.J. Exponentials and Laplace transforms on nonuniform time scales. *Commun. Nonlinear Sci. Numer. Simul.* **2016**, *39*, 252–270. [[CrossRef](#)]
49. Şan, M.; Ortigueira, M.D. Unilateral Laplace Transforms on Time Scales. *Mathematics* **2022**, *10*, 4552. [[CrossRef](#)]
50. Ortigueira, M.D. The comb signal and its Fourier transform. *Signal Process.* **2001**, *81*, 581–592. [[CrossRef](#)]
51. Ferreira, J. *Introduction to the Theory of Distributions*; Pitman Monographs and Surveys in Pure and Applied Mathematics; Pitman: London, UK, 1997.

52. Gelfand, I.M.; Shilov, G.P. *Generalized Functions*; Academic Press: New York, NY, USA, 1964; Volume 3, English translation.
53. Hoskins, R.; Pinto, J. *Theories of Generalised Functions: Distributions, Ultradistributions and Other Generalised Functions*; Woodhead Publishing Limited: Cambridge, UK, 2010.
54. Hoskins, R. *Delta Functions: An Introduction to Generalised Functions*; Woodhead Publishing Limited: Cambridge, UK, 2009.
55. Roberts, M. *Signals and Systems: Analysis Using Transform Methods and Matlab*, 2nd ed.; McGraw-Hill: New York, NY, USA, 2003.
56. Vaidyanathan, P.P. The theory of linear prediction. *Synth. Lect. Signal Process.* **2007**, *2*, 1–184.
57. Bohner, M.; Peterson, A. *Dynamic Equations on Time Scales: An Introduction with Applications*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2001.
58. Liouville, J. Mémoire sur le calcul des différentielles à indices quelconques. *J. l'École Polytech. Paris* **1832**, *13*, 71–162.
59. 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.
60. Grünwald, A.K. Ueber "begrentz" Derivationen und deren Anwendung. *Z. Math. Phys.* **1867**, *12*, 441–480.
61. Letnikov, A. Note relative à l'explication des principes fondamentaux de la théorie de la différentiation à indice quelconque (A propos d'un mémoire). *Mat. Sb.* **1873**, *6*, 413–445.
62. Rogosin, S.; Dubatovskaya, M. Fractional Calculus in Russia at the End of XIX Century. *Mathematics* **2021**, *9*, 1736. [[CrossRef](#)]
63. Heaviside, O., III. On Operators in Physical Mathematics. Part I. *Proc. R. Soc. Lond.* **1893**, *52*, 504–529.
64. Heaviside, O., VIII. On operations in physical mathematics. Part II. *Proc. R. Soc. Lond.* **1894**, *54*, 105–143.
65. Post, E.L. Generalized differentiation. *Trans. Am. Math. Soc.* **1930**, *32*, 723–781. [[CrossRef](#)]
66. Butzer, P.L.; Westphal, U. An access to fractional differentiation via fractional difference quotients. In *Fractional Calculus and Its Applications: Proceedings of the International Conference Held at the University of New Haven, June 1974*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 116–145.
67. Diaz, J.; Osler, T. Differences of fractional order. *Math. Comput.* **1974**, *28*, 185–202. [[CrossRef](#)]
68. Ortigueira, M.D. Fractional central differences and derivatives. *J. Vib. Control* **2008**, *14*, 1255–1266. [[CrossRef](#)]
69. Ortigueira, M.D. Two-sided and regularised Riesz-Feller derivatives. *Math. Methods Appl. Sci.* **2021**, *44*, 8057–8069. [[CrossRef](#)]
70. Chapman, S. On non-integral orders of summability of series and integrals. *Proc. Lond. Math. Soc.* **1911**, *2*, 369–409. [[CrossRef](#)]
71. Kuttner, B. On Differences of Fractional Order. *Proc. Lond. Math. Soc.* **1957**, *s3-7*, 453–466. [[CrossRef](#)]
72. Isaacs, G.L. Exponential laws for fractional differences. *Math. Comput.* **1980**, *35*, 933–936. [[CrossRef](#)]
73. Granger, C. New classes of time series models. *J. R. Stat. Soc. Ser. D (Stat.)* **1978**, *27*, 237–253. [[CrossRef](#)]
74. Granger, C.W.; Joyeux, R. An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.* **1980**, *1*, 15–29. [[CrossRef](#)]
75. Hosking, J.R.M. Fractional differencing. *Biometrika* **1981**, *68*, 165–176. [[CrossRef](#)]
76. Gonçalves, E. Une généralisation des processus ARMA. *Ann. d'Économie Stat.* **1987**, 109–145. [[CrossRef](#)]
77. Elder, J.; Elliott, R.J.; Miao, H. Fractional differencing in discrete time. *Quant. Financ.* **2013**, *13*, 195–204. [[CrossRef](#)]
78. Graves, T.; Gramacy, R.; Watkins, N.; Franzke, C. A brief history of long memory: Hurst, Mandelbrot and the road to ARFIMA, 1951–1980. *Entropy* **2017**, *19*, 437. [[CrossRef](#)]
79. Dingari, M.; Reddy, D.M.; Sumalatha, V. Time series analysis for long memory process of air traffic using arfima. *Int. J. Sci. Technol. Res.* **2019**, *8*, 395–400.
80. Monge, M.; Infante, J. A Fractional ARIMA (ARFIMA) Model in the Analysis of Historical Crude Oil Prices. *Energy Res. Lett.* **2022**, *4*. [[CrossRef](#)]
81. Cargo, G.; Shisha, O. Zeros of polynomials and fractional order differences of their coefficients. *J. Math. Anal. Appl.* **1963**, *7*, 176–182. [[CrossRef](#)]
82. Burnecki, K.; Weron, A. Algorithms for testing of fractional dynamics: A practical guide to ARFIMA modelling. *J. Stat. Mech. Theory Exp.* **2014**, *2014*, P10036. [[CrossRef](#)]
83. Kilbas, A.A.; Srivastava, H.M.; Trujillo, J.J. *Theory and Applications of Fractional Differential Equations*; Elsevier: Amsterdam, The Netherlands, 2006.
84. Abdeljawad, T. On Riemann and Caputo fractional differences. *Comput. Math. Appl.* **2011**, *62*, 1602–1611. [[CrossRef](#)]
85. Abdeljawad, T. Dual identities in fractional difference calculus within Riemann. *Adv. Differ. Equ.* **2013**, *2013*, 36. [[CrossRef](#)]
86. Ostalczyk, P. Remarks on five equivalent forms of the fractional-order backward-difference. *Bull. Pol. Acad. Sci. Tech. Sci.* **2014**, *62*, 271–278. [[CrossRef](#)]
87. Miller, K.; Ross, B. Fractional difference calculus. In *Proceedings of the International Symposium on Univalent Functions, Fractional Calculus and Their Applications, Nihon University, Koriyama, Japan, May 1988*; Ellis Horwood: Chichester, West Sussex, England, UK, 1989; pp. 139–152.
88. Podlubny, I. Matrix approach to discrete fractional calculus. *Fract. Calc. Appl. Anal.* **2000**, *3*, 359–386.
89. Atici, F.M.; Eloe, P.W. A transform method in discrete fractional calculus. *Int. J. Differ. Equ.* **2007**, *2*, 165–176.
90. Atici, F.; Eloe, P. Initial value problems in discrete fractional calculus. *Proc. Am. Math. Soc.* **2009**, *137*, 981–989. [[CrossRef](#)]
91. Atici, F.M.; Eloe, P. Discrete fractional calculus with the nabla operator. *Electron. J. Qual. Theory Differ. Equ. [Electron. Only]* **2009**, *2009*, 1–12. [[CrossRef](#)]
92. Bastos, N.R.; Torres, D.F. Combined Delta-Nabla Sum Operator in Discrete Fractional Calculus. *arXiv* **2010**, arXiv:1009.3883.

93. Bastos, N.R.; Ferreira, R.A.; Torres, D.F. Necessary optimality conditions for fractional difference problems of the calculus of variations. *arXiv* **2010**, arXiv:1007.0594.
94. Ferreira, R.A.; Torres, D.F. Fractional h-difference equations arising from the calculus of variations. *Appl. Anal. Discret. Math.* **2011**, *5*, 110–121. [[CrossRef](#)]
95. Holm, M. Sum and difference compositions in discrete fractional calculus. *Cubo* **2011**, *13*, 153–184. [[CrossRef](#)]
96. Bastos, N.R. Fractional calculus on time scales. *arXiv* **2012**, arXiv:1202.2960.
97. Mohan, J.J.; Deekshitulu, G. Fractional order difference equations. *Int. J. Differ. Equ.* **2012**, *2012*, 780619. [[CrossRef](#)]
98. Mozyrska, D.; Girejko, E. Overview of fractional h-difference operators. In *Proceedings of the Advances in Harmonic Analysis and Operator Theory: The Stefan Samko Anniversary, Lisbon and Aveiro, Portugal, in June–July, 2011*; Operator Theory: Advances and Applications; Springer: Basel, Switzerland, 2013; Volume 229, pp. 253–268.
99. Mozyrska, D. Multiparameter fractional difference linear control systems. *Discret. Dyn. Nat. Soc.* **2014**, *2014*, 183782. [[CrossRef](#)]
100. Atici, F.M.; Dadashova, K.; Jonnalagadda, J. Linear fractional order h-difference equations. *Int. J. Differ. Equ. (Spec. Issue Honor. Profr. Johnny Henderson)* **2020**, *15*, 281–300.
101. Wang, Q.; Xu, R. A review of definitions of fractional differences and sums. *Math. Found. Comput.* **2023**, *6*, 136–160. [[CrossRef](#)]
102. Wei, Y.; Zhao, L.; Zhao, X.; Cao, J. Enhancing the Mathematical Theory of Nabla Tempered Fractional Calculus: Several Useful Equations. *Fractal Fract.* **2023**, *7*, 330. [[CrossRef](#)]
103. Joshi, D.D.; Bhalekar, S.; Gade, P.M. Controlling fractional difference equations using feedback. *Chaos Solitons Fractals* **2023**, *170*, 113401. [[CrossRef](#)]
104. Abdeljawad, T.; Atici, F.M. On the definitions of nabla fractional operators. In *Abstract and Applied Analysis*; Hindawi Publishing Corporation: New York, NY, USA, 2012; Volume 2012, p. 406757.
105. Bastos, N.R.; Ferreira, R.A.; Torres, D.F. Discrete-time fractional variational problems. *Signal Process.* **2011**, *91*, 513–524. [[CrossRef](#)]
106. Alzabut, J.; Grace, S.R.; Jonnalagadda, J.M.; Santra, S.S.; Abdalla, B. Higher-Order Nabla Difference Equations of Arbitrary Order with Forcing, Positive and Negative Terms: Non-Oscillatory Solutions. *Axioms* **2023**, *12*, 325. . [[CrossRef](#)]
107. Graham, R.L.; Knuth, D.E.; Patashnik, O.; Liu, S. Concrete mathematics: A foundation for computer science. *Comput. Phys.* **1989**, *3*, 106–107. [[CrossRef](#)]
108. 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.
109. Ortigueira, M.D. Riesz potential operators and inverses via fractional centred derivatives. *Int. J. Math. Math. Sci.* **2006**, *2006*, 48391:1–48391:12. [[CrossRef](#)]
110. Ortigueira, M.D.; Bengochea, G.; Machado, J.A.T. Substantial, tempered, and shifted fractional derivatives: Three faces of a tetrahedron. *Math. Methods Appl. Sci.* **2021**, *44*, 9191–9209. [[CrossRef](#)]
111. Sabzikar, F.; Meerschaert, M.M.; Chen, J. Tempered fractional calculus. *J. Comput. Phys.* **2015**, *293*, 14–28. [[CrossRef](#)]
112. Ortigueira, M.D.; Bengochea, G. Bilateral tempered fractional derivatives. *Symmetry* **2021**, *13*, 823. [[CrossRef](#)]
113. Tustin, A. A method of analysing the behaviour of linear systems in terms of time series. *J. Inst. Electr. Eng.—Part IIA Autom. Regul. Servo Mech.* **1947**, *94*, 130–142. [[CrossRef](#)]
114. Ortigueira, M.D.; Machado, J.T. New discrete-time fractional derivatives based on the bilinear transformation: Definitions and properties. *J. Adv. Res.* **2020**, *25*, 1–10. [[CrossRef](#)]
115. 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](#)]
116. Ortigueira, M.D.; Machado, J.A.T. The 21st Century Systems: An updated vision of Continuous-Time Fractional Models. *Circuits Syst. Mag.* **2022**, *22*, 36–56. [[CrossRef](#)]
117. Ortigueira, M.D.; Batista, A.G. A fractional linear system view of the fractional Brownian motion. *Nonlinear Dyn.* **2004**, *38*, 295–303. [[CrossRef](#)]
118. Ortigueira, M.D.; Batista, A.G. On the relation between the fractional Brownian motion and the fractional derivatives. *Phys. Lett. A* **2008**, *372*, 958–968. [[CrossRef](#)]
119. Dissanayake, G.S.; Peiris, M.S.; Proietti, T. Fractionally Differenced Gegenbauer Processes with Long Memory: A Review. *Stat. Sci.* **2018**, *33*, 413–426. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.