Chapter 3
Caputo's Approach

It turns out that the Riemann–Liouville derivatives have certain disadvantages when trying to model real-world phenomena with fractional differential equations. We shall therefore now discuss a modified concept of a fractional derivative. As we will see below when comparing the two ideas, this second one seems to be better suited to such tasks.

3.1 Definition and Basic Properties

We commence with a preliminary definition.

Definition 3.1. Let $n \geq 0$ and $m = \lceil n \rceil$. Then, we define the operator $\widehat{D}_a^n$ by

$$\widehat{D}_a^n f := J_a^{m-n} D^m f$$

whenever $D^m f \in L_1[a,b]$.

Let us start by looking at the case $n \in \mathbb{N}$. Here we have $m = n$ and hence our definition implies

$$\widehat{D}_a^n f = J_a^0 D^n f = D^n f,$$

i.e. we recover the standard definition in the classical case.

We begin the analysis of this operator in the strictly fractional case $n \notin \mathbb{N}$ with a simple example.

Example 3.1. Let $f(x) = (x-a)^\beta$ for some $\beta \geq 0$. Then,

$$\widehat{D}_a^n f(x) = \begin{cases} 0 & \text{if } \beta \in \{0,1,2,\ldots,m-1\}, \\ \dfrac{\Gamma(\beta+1)}{\Gamma(\beta+1-n)}(x-a)^{\beta-n} & \text{if } \beta \in \mathbb{N} \text{ and } \beta \geq m \\ & \text{or } \beta \notin \mathbb{N} \text{ and } \beta > m-1. \end{cases}$$

The reader is encouraged to compare this statement with the corresponding one for Riemann–Liouville operators (Example 2.4). Notice in particular that the two

K. Diethelm, *The Analysis of Fractional Differential Equations*,
Lecture Notes in Mathematics 2004, DOI 10.1007/978-3-642-14574-2_3,

operators have different kernels, and that the domains of the two operators (exhibited here in terms of the allowed range of the parameter β) are also different.

A few additional examples of Caputo-type derivatives of certain important functions are collected for the reader's convenience in Appendix B.

Remark 3.1. We have required $m = \lceil n \rceil$ in Definition 3.1. The same condition has been imposed in the definition $D_a^n := D^m J_a^{m-n}$ of the Riemann–Liouville derivative (Definition 2.2). However, in the latter case we had seen in Lemma 2.11 that this restriction actually is not necessary; one may use any $m \in \mathbb{N}$ with $m \geq n$ in the Riemann–Liouville case. For the newly introduced operator $\widehat{D}_a^n := J_a^{m-n} D^m$ from Definition 3.1, the situation is different: Here we may not replace $m = \lceil n \rceil$ by some $m \in \mathbb{N}$ with $m > \lceil n \rceil$. This is evident by looking at the simple example $f(x) = (x-a)^{\lceil n \rceil}$. For such a function we have, according to Example 3.1,

$$\widehat{D}_a^n f(x) = \frac{\Gamma(\lceil n \rceil + 1)}{\Gamma(\lceil n \rceil + 1 - n)} (x-a)^{\lceil n \rceil - n}$$

but, for $m \in \mathbb{N}$ with $m > \lceil n \rceil$, we obtain $D^m f(x) = 0$ and hence $J_a^{m-n} D^m f(x) = 0$ too.

The key to the construction of the alternative differential operator that we are looking for is the following identity involving Riemann–Liouville derivatives on the one hand and the newly defined operator on the other hand.

Theorem 3.1. *Let $n \geq 0$ and $m = \lceil n \rceil$. Moreover assume that $f \in A^m[a,b]$. Then,*

$$\widehat{D}_a^n f = D_a^n [f - T_{m-1}[f;a]]$$

almost everywhere. Here, as in the proof of Theorem 2.20, $T_{m-1}[f;a]$ denotes the Taylor polynomial of degree $m-1$ for the function f, centered at a; in the case $m = 0$ we define $T_{m-1}[f;a] := 0$.

Note that the expression on the right-hand side of the equation exists if $D_a^n f$ exists and f possesses $m-1$ derivatives at a, the latter condition making sure that the Taylor polynomial exists. This condition is weaker than the previous condition that $f \in A^m$. (This follows since $f \in A^m$ implies (a) $f \in C^{m-1}$ and hence the existence of the required Taylor polynomial and its Riemann–Liouville derivative, and (b) the existence of $D_a^n f$ almost everywhere as can be seen by a repeated application of the ideas used in the proof of Lemma 2.12.) Therefore we will, from now on, use the latter expression. A formalization is given as follows.

Definition 3.2. Assume that $n \geq 0$ and that f is such that $D_a^n [f - T_{m-1}[f;a]]$ exists, where $m = \lceil n \rceil$. Then we define the function $D_{*a}^n f$ by

$$D_{*a}^n f := D_a^n [f - T_{m-1}[f;a]].$$

The operator D_{*a}^n is called the *Caputo differential operator of order n*.

Actually this concept has been introduced independently by many authors, including Caputo [23] and Rabotnov [157] who have based their developments on the approach given in Definition 3.1 and by Dzherbashyan and Nersesian [58] who have used Definition 3.2 as their starting point; other contributors who have dealt with such operators from various points of view include Gross [84] and Gerasimov [74], and it can even be found in a very old paper by Liouville [116, p. 10, formula (B)]. However it seems that Liouville did not see the difference between this operator and the Riemann–Liouville operator as he was mainly interested in those cases where the two operators coincide [Lützen, J., 2001, Private communication]. We follow the most common convention of naming it after Caputo only. The reader who is interested in a detailed historical account should consult the recent paper by Rossikhin [163] and the references cited therein. The notation that we have introduced here follows the generally accepted suggestion of Gorenflo and Mainardi [81].

Once again we note for $n \in \mathbb{N}$ that $m = n$ and hence

$$D^n_{*a} f = D^n_a [f - T_{n-1}[f;a]] = D^n f - D^n (T_{n-1}[f;a]) = D^n f$$

because $T_{n-1}[f;a]$ is a polynomial of degree $n-1$ that is annihilated by the classical operator D^n. So in this case we recover the usual differential operator as well. In particular, D^0_{*a} is once again the identity operator.

Various papers and books exist where some of the key properties of the Caputo operators have been described, see, e.g., [77, 81, 153]. Typically however they were, as stated explicitly in the abstract of [81], written "in a way accessible to applied scientists" and "avoiding unproductive generalities and excessive mathematical rigor". It seems that mathematically rigorous proofs of many important properties are not available in the literature. Therefore we try to give them here.

Proof (of Theorem 3.1). In the case $n \in \mathbb{N}$ the statement is trivial because, as we have seen above, both sides of the equation reduce to $D^n f$. We therefore only have to consider the case $n \notin \mathbb{N}$, which implies that $m > n$.

In this case we have

$$\begin{aligned} D^n_a [f - T_{m-1}[f;a]](x) &= D^m J^{m-n}_a [f - T_{m-1}[f;a]](x) \\ &= \frac{\mathrm{d}^m}{\mathrm{d}x^m} \int_a^x \frac{(x-t)^{m-n-1}}{\Gamma(m-n)} (f(t) - T_{m-1}[f;a](t))\,\mathrm{d}t. \end{aligned} \tag{3.1}$$

A partial integration of the integral is permitted and yields

$$\begin{aligned} &\int_a^x \frac{1}{\Gamma(m-n)} (f(t) - T_{m-1}[f;a](t))(x-t)^{m-n-1}\,\mathrm{d}t \\ &\quad = -\frac{1}{\Gamma(m-n+1)} \left[(f(t) - T_{m-1}[f;a](t))(x-t)^{m-n} \right]_{t=a}^{t=x} \\ &\qquad + \frac{1}{\Gamma(m-n+1)} \int_a^x (Df(t) - DT_{m-1}[f;a](t))(x-t)^{m-n}\,\mathrm{d}t. \end{aligned}$$

The term outside the integral is zero (the first factor vanishes at the lower bound, the second vanishes at the upper bound). Thus,

$$J_a^{m-n}[f - T_{m-1}[f;a]] = J_a^{m-n+1} D[f - T_{m-1}[f;a]].$$

Under our assumptions, we may repeat this process a total number of m times, and this results in

$$J_a^{m-n}[f - T_{m-1}[f;a]] = J_a^{2m-n} D^m [f - T_{m-1}[f;a]] = J_a^m J_a^{m-n} D^m [f - T_{m-1}[f;a]].$$

We note that $D^m T_{m-1}[f;a] \equiv 0$ because $T_{m-1}[f;a]$ is a polynomial of degree $m-1$. Thus, the last identity can be simplified to

$$J_a^{m-n}[f - T_{m-1}[f;a]] = J_a^m J_a^{m-n} D^m f.$$

This may be combined with (3.1) to obtain

$$D_a^n [f - T_{m-1}[f;a]](x) = D^m J_a^m J_a^{m-n} D^m f = J_a^{m-n} D^m f = \widehat{D}_a^n f$$

in view of (1.1). □

Taking into account the definition of the Caputo operator and Lemma 2.21, we obtain a direct consequence.

Lemma 3.2. *Under the assumptions of Lemma 2.21, we have*

$$D_{*a}^n f(x) = \frac{1}{\Gamma(-n)} \int_a^x (x-t)^{-n-1} \left(f(t) - T_{m-1}[f;a](t) \right) \mathrm{d}t.$$

Remark 3.2. As in the case of the Riemann–Liouville operators, we see that the Caputo derivatives are not local either.

Yet another representation for the Caputo operator can be obtained by combining its definition with Theorem 2.25:

Lemma 3.3. *Let $n > 0$, $m = \lceil n \rceil$ and $f \in C^m[a,b]$. Then, for $x \in (a,b]$,*

$$D_{*a}^n f(x) = \lim_{N \to \infty} \frac{1}{h_N^n} \sum_{k=0}^{N} (-1)^k \binom{n}{k} [f(x - kh_N) - T_{m-1}[f;a](x - kh_N)]$$

with $h_N = (x-a)/N$.

The representations of these two Lemmas have proven to be useful for numerical work [34, 121]. Other representations are known as well; we shall present some of them in Sect. 3.2. However, we first continue the investigations of the analytical aspects of Caputo operators. In this context our next goal is to express the relation between the Riemann–Liouville operator and the Caputo operator in a different way.

Lemma 3.4. *Let $n \geq 0$ and $m = \lceil n \rceil$. Assume that f is such that both $D^n_{*a} f$ and $D^n_a f$ exist. Then,*

$$D^n_{*a} f(x) = D^n_a f(x) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{\Gamma(k-n+1)} (x-a)^{k-n}.$$

Proof. In view of the definition of the Caputo derivative and Example 2.4,

$$\begin{aligned} D^n_{*a} f(x) &= D^n_a f(x) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{k!} D^n_a [(\cdot - a)^k](x) \\ &= D^n_a f(x) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{\Gamma(k-n+1)} (x-a)^{k-n}. \end{aligned}$$

□

An immediate consequence of this Lemma is

Lemma 3.5. *Assume the hypotheses of Lemma 3.4. Then,*

$$D^n_a f = D^n_{*a} f$$

holds if and only if f has an m-fold zero at a, i.e. if and only if

$$D^k f(a) = 0 \quad \text{for} \quad k = 0, 1, \ldots, m-1.$$

We may also combine Lemma 3.4 with Theorem 2.25 to deduce

Lemma 3.6. *Let $n > 0$, $m = \lceil n \rceil$ and $f \in C^m[a,b]$. Then, for $x \in (a,b]$,*

$$D^n_{*a} f(x) = \lim_{N \to \infty} \frac{1}{h_N^n} \sum_{k=0}^{N} (-1)^k \binom{n}{k} f(x - k h_N) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{\Gamma(k-n+1)} (x-a)^{k-n}$$

with $h_N = (x-a)/N$.

When it comes to the composition of Riemann–Liouville integrals and Caputo differential operators, we find that the Caputo derivative is also a left inverse of the Riemann–Liouville integral:

Theorem 3.7. *If f is continuous and $n \geq 0$, then*

$$D^n_{*a} J^n_a f = f.$$

Proof. Let $\phi = J^n_a f$. By Theorem 2.5, we have $D^k \phi(a) = 0$ for $k = 0, 1, \ldots, m-1$, and thus (in view of Lemma 3.5 and Theorem 2.14)

$$D^n_{*a} J^n_a f = D^n_{*a} \phi = D^n_a \phi = D^n_a J^n_a f = f.$$

□

Once again, we find that the Caputo derivative is not the right inverse of the Riemann–Liouville integral:

Theorem 3.8. *Assume that $n \geq 0$, $m = \lceil n \rceil$, and $f \in A^m[a,b]$. Then*

$$J_a^n D_{*a}^n f(x) = f(x) - \sum_{k=0}^{m-1} \frac{D^k f(a)}{k!}(x-a)^k.$$

Proof. By Theorem 3.1 and Definition 3.1, we have

$$D_{*a}^n f = \widehat{D}_a^n f = J_a^{m-n} D^m f.$$

Thus, applying the operator J_a^n to both sides of this equation and using the semigroup property of fractional integration, we obtain

$$J_a^n D_{*a}^n f = J_a^n J_a^{m-n} D^m f = J_a^m D^m f.$$

By the classical version of Taylor's theorem (cf. Theorem 2.C), we have that

$$f(x) = \sum_{k=0}^{m-1} \frac{D^k f(a)}{k!}(x-a)^k + J_a^m D^m f(x).$$

Combining these two equations we derive the claim. □

A fractional analogue of Taylor's theorem follows immediately:

Corollary 3.9 (Taylor expansion for Caputo derivatives). *Under the assumptions of Theorem 3.8,*

$$f(x) = \sum_{k=0}^{m-1} \frac{D^k f(a)}{k!}(x-a)^k + J_a^n D_{*a}^n f(x).$$

The relations shown in Theorem 3.8 and Corollary 3.9 have major implications when it comes to the solution of differential equations involving the two types of differential operators. Specifically, assume that h is a given function with the property that there exists some function g such that $h = D_a^n g$. Then, the solution of the Riemann–Liouville differential equation

$$D_a^n f = h$$

is given by

$$f(x) = g(x) + \sum_{j=1}^{\lceil n \rceil} c_j (x-a)^{n-j}$$

with arbitrary constants c_j. This follows by the same techniques that one would employ for differential equations of integer order because the equation is linear and inhomogeneous. By construction, g is a solution of the inhomogeneous equation, and by Example 2.4 each of the terms in the sum solves the corresponding homogeneous equation.

Similarly, if h_* is a given function with the property that $h_* = D^n_{*a} g_*$ and if we want to solve

$$D^n_{*a} f_* = h_*,$$

then we find

$$f_*(x) = g_*(x) + \sum_{j=1}^{\lceil n \rceil} c_j^* (x-a)^{\lceil n \rceil - j},$$

again with arbitrary constants c_j^*. Thus, in order to obtain a unique solution, it is most natural to prescribe the values $f_*(a), Df_*(a), \ldots, D^{\lceil n \rceil - 1} f_*(a)$ in the Caputo setting, whereas in the Riemann–Liouville case one would rather prescribe fractional derivatives of f at a. This will be explored in a more detailed fashion in the following chapters. For the moment we note that the Caputo version is usually preferred when physical models are described because the physical interpretation of the prescribed data is clear, and therefore it is in general possible to provide these data, e.g. by suitable measurements. This is not true for the fractional order initial conditions required for the Riemann–Liouville environment. For example, in applications like the modelling of viscoelastic materials in mechanics [184], $f(x)$ is typically a displacement at time x, and so $f'(x)$ and $f''(x)$ would be the corresponding velocity and acceleration, respectively – quantities that are well understood and easily measured. On the other hand, in spite of recently attempted explanations [154], a fractional derivative of a displacement remains an object whose physical nature is unclear, and so no measurement methods for such a quantity are readily available.

Apart from this reason (which is mainly motivated by arguments in connection with applications) there are actually other reasons coming from the "pure" side of mathematics for preferring the Caputo derivative over the Riemann–Liouville operator [97, 108], but we shall not dwell on this topic here. Rather, we shall continue by stating another representation for the Caputo derivative of a function under a quite natural assumption. To this end we require the following definition that is a special case of a concept established in [73] (see also [167, p. 426]).

Definition 3.3. Let $n > 0$ and let v be an entire function with the power series expansion $v(x) = \sum_{k=0}^{\infty} c_k x^k$. Then, the operator $\mathscr{D}^n$ that maps this function v to the function $\mathscr{D}^n v$ with

$$\mathscr{D}^n v(x) := \sum_{k=1}^{\infty} c_k \frac{\Gamma(kn+1)}{\Gamma(kn+1-n)} x^{k-1}$$

is called the *Gel'fond-Leont'ev operator* of order n.

The Caputo derivatives of certain functions can be expressed with the help of these operators in a very convenient way (see, e.g., [106, p. 139] or [167, p. 426]):

Theorem 3.10. *Let $n > 0$ and let v be an entire function with the power series expansion $v(x) = \sum_{k=0}^{\infty} c_k x^k$. Moreover let $f(x) := v(x^n)$ for $x \geq 0$. Then,*

$$D_{*0}^{n} f(x) = \mathscr{D}^n v(x^n).$$

Proof. This result follows using a straightforward computation using the defintions of the function f and the Caputo differential operator, the power series expansion of v and the fact that, because v is entire, we may interchange the infinite series operator and the Caputo differential operator (i.e., we may apply the Caputo operator to the series in a term-by-term manner). □

The following result establishes another significant difference between Riemann–Liouville and Caputo derivatives. A comparison with, e.g., Example 2.4 for $f(x) = 1$ and $n > 0$, $n \notin \mathbb{N}$, reveals that we are not allowed to replace D_{*a}^{n} by D_{a}^{n} here.

Lemma 3.11. *Let $n > 0$, $n \notin \mathbb{N}$ and $m = \lceil n \rceil$. Moreover assume that $f \in C^m[a,b]$. Then, $D_{*a}^{n} f \in C[a,b]$ and $D_{*a}^{n} f(a) = 0$.*

Proof. By definition and Theorem 3.1, $D_{*a}^{n} f = J_a^{m-n} D^m f$. The result follows from Theorem 2.5 because $D^m f$ is assumed to be continuous. □

We may relax the conditions on f slightly.

Lemma 3.12. *Let $n > 0$, $n \notin \mathbb{N}$ and $m = \lceil n \rceil$. Moreover let $f \in A^m[a,b]$ and assume that $D_{*a}^{\hat{n}} f \in C[a,b]$ for some $\hat{n} \in (n,m)$. Then, $D_{*a}^{n} f \in C[a,b]$ and $D_{*a}^{n} f(a) = 0$.*

Proof. By definition and Theorems 3.1 and 2.2,

$$D_{*a}^{n} f = J_a^{m-n} D^m f = J_a^{\hat{n}-n} J_a^{m-\hat{n}} D^m f = J_a^{\hat{n}-n} D_{*a}^{\hat{n}} f.$$

Thus the claim follows by virtue of Theorem 2.5. □

Remark 3.3. In order to assess the consequences of these two Lemmata, we point out the following fact. Assume, for example, the hypotheses of Lemma 3.11 and additionally that $n > 2$. Then the Lemma asserts that all derivatives $D_{*a}^{n} f$, $0 < n < 3$, are continuous, and thus, in particular, $D_{*a}^{\ell} f$ and $D_{*a}^{\ell+1} f$ are continuous whenever $0 < \ell < 2$. This does not mean, however, that $D_{*a}^{\ell} f \in C^1[a,b]$ for these ℓ. For a counterexample, we refer to Exercise 3.1. As a consequence of this observation, we obtain that we cannot deduce the identity $D D_{*a}^{\ell} f = D_{*a}^{\ell+1} f$ to be true under the assumptions of our Lemmata because the function on the right-hand side is continuous whereas the one on the left-hand side need not have this property. Hence we find that the Caputo differential operators do not form a semigroup in general.

In this context the following observation is important.

Lemma 3.13. *Let $f \in C^k[a,b]$ for some $a < b$ and some $k \in \mathbb{N}$. Moreover let $n, \varepsilon > 0$ be such that there exists some $\ell \in \mathbb{N}$ with $\ell \leq k$ and $n, n+\varepsilon \in [\ell-1, \ell]$. Then,*

$$D_{*a}^{\varepsilon} D_{*a}^{n} f = D_{*a}^{n+\varepsilon} f.$$

Remark 3.4. Two comments concerning this Lemma need to be made:

(a) Such a result cannot be expected to hold in general if Riemann–Liouville derivatives were used instead of Caputo derivatives. As an example, consider the function f with $f(x) = 1$ and let $a = 0$, $n = 1$ and $\varepsilon = 1/2$. If we were to use Riemann–Liouville derivatives, the left-hand side would be $(D_0^{1/2} f')(x) = D_0^{1/2} 0 = 0$, whereas the right-hand side is

$$D_0^{3/2} f(x) = D^2 J_0^{1/2} f(x) = \frac{1}{\Gamma(-1/2)} x^{-3/2}.$$

(b) The condition requiring the existence of the number ℓ with the properties mentioned in the Lemma is essential. To see what can happen without it, consider the example $n = \varepsilon = 7/10$ (i.e. $7/10 = n < 1 < n + \varepsilon = 7/5$), $a = 0$ and $f(x) = x$. Then, the right-hand side is $D_{*0}^{7/5} f(x) = (J_0^{3/5} f'')(x) = J_0^{3/5} 0 = 0$, but since

$$(D_{*0}^{7/10} f)(x) = \frac{1}{\Gamma(13/10)} x^{3/10},$$

the left-hand side takes the value

$$D_{*0}^{7/10} (D_{*0}^{7/10} f)(x) = \frac{1}{\Gamma(3/5)} x^{-2/5}.$$

Proof (of Lemma 3.13). The statement is trivial in the case $n = \ell - 1$ and $n + \varepsilon = \ell$, so we only treat the other situations explicitly. Then we first observe that our assumptions imply $0 < \varepsilon < 1$. Thus, by Lemma 3.5 we find that

$$D_{*a}^{\varepsilon} z = D_a^{\varepsilon} z$$

whenever $z(a) = 0$. We consider three cases:

1. $n + \varepsilon \in \mathbb{N}$: In this case we have that $\lceil n \rceil = n + \varepsilon$ and hence $\lceil n \rceil - n = \varepsilon$. Moreover, by Lemma 3.11, $D_{*a}^n f(a) = 0$. Thus

$$\begin{aligned} D_{*a}^{\varepsilon} D_{*a}^n f &= D_a^{\varepsilon} D_{*a}^n f = D_a^{\varepsilon} J_a^{\lceil n \rceil - n} D^{\lceil n \rceil} f \\ &= D_a^{\varepsilon} J_a^{\varepsilon} D^{\lceil n \rceil} f = D^{\lceil n \rceil} f = D^{n+\varepsilon} f = D_{*a}^{n+\varepsilon} f. \end{aligned}$$

2. $n \in \mathbb{N}$: Here we have, using Theorem 3.1,

$$D_{*a}^{\varepsilon} D_{*a}^n f = D_{*a}^{\varepsilon} D^n f = J_a^{1-\varepsilon} D^{n+1} f = D_{*a}^{n+\varepsilon} f.$$

3. Otherwise, we have $\lceil n \rceil = \lceil n + \varepsilon \rceil$, and thus we find by similar arguments that

$$\begin{aligned} D_{*a}^{\varepsilon} D_{*a}^n f &= D_a^{\varepsilon} D_{*a}^n f = D_a^{\varepsilon} J_a^{\lceil n \rceil - n} D^{\lceil n \rceil} f \\ &= D^1 J_a^{1-\varepsilon} J_a^{\lceil n \rceil - n} D^{\lceil n \rceil} f = D^1 J_a^1 J_a^{\lceil n+\varepsilon \rceil - (n+\varepsilon)} D^{\lceil n+\varepsilon \rceil} f \\ &= D_{*a}^{n+\varepsilon} f. \end{aligned}$$

$\square$

The previous result has dealt with the concatenation of two Caputo differential operators. In some instances however it may also be useful to concatenate a Caputo operator with a differential operator of Riemann–Liouville type:

Theorem 3.14. *Let $f \in C^\mu[a,b]$ for some $\mu \in \mathbb{N}$. Moreover let $n \in [0,\mu]$. Then,*

$$D_a^{\mu-n} D_{*a}^n f = D^\mu f.$$

Notice that the operator D^μ appearing on the right-hand side of the claim is a classical (integer-order) differential operator.

Proof. If n is an integer then both differential operators on the left-hand side reduce to integer-order operators and hence we obtain the desired result by an application of the definition of the iterated operators, viz. Definition 1.1 (c).

If n is not an integer then we may invoke Theorem 3.1 to conclude that

$$D_{*a}^n f = \widehat{D}_a^n f = J_a^{\lceil n\rceil - n} D^{\lceil n\rceil} f.$$

Combining this with the definition of the Riemann–Liouville derivative and using the semigroup property of fractional integration and eq. (1.1) we find

$$\begin{aligned} D_a^{\mu-n} D_{*a}^n f &= D^{\mu-\lceil n\rceil+1} J_a^{n+1-\lceil n\rceil} J_a^{\lceil n\rceil-n} D^{\lceil n\rceil} f = D^{\mu-\lceil n\rceil+1} J_a^1 D^{\lceil n\rceil} f \\ &= D^{\mu-\lceil n\rceil} D^{\lceil n\rceil} f = D^\mu f. \end{aligned}$$ □

It is actually possible to explore the smoothness properties of $D_{*0}^n f$ under smoothness assumptions on f in more detail than in Lemma 3.11:

Theorem 3.15. *If $f \in C^\mu[a,b]$ for some $\mu \in \mathbb{N}$ and $0 < n < \mu$ then*

$$D_{*a}^n f(x) = \sum_{\ell=0}^{\mu-\lceil n\rceil-1} \frac{f^{(\ell+\lceil n\rceil)}(a)}{\Gamma(\lceil n\rceil - n + \ell + 1)} (x-a)^{\lceil n\rceil - n + \ell} + g(x)$$

with some function $g \in C^{\mu-\lceil n\rceil}[a,b]$. Moreover, the $(\mu - \lceil n\rceil)$th derivative of g satisfies a Lipschitz condition of order $\lceil n\rceil - n$.

Proof. This is a direct consequence of the definition of the Caputo differential operator and Theorems 3.1 and 2.5. □

The main computational rules for the Caputo derivative are similar, but not identical, to those for the Riemann–Liouville derivative.

Theorem 3.16. *Let $f_1, f_2 : [a,b] \to \mathbb{R}$ be such that $D_{*a}^n f_1$ and $D_{*a}^n f_2$ exist almost everywhere and let $c_1, c_2 \in \mathbb{R}$. Then, $D_{*a}^n(c_1 f_1 + c_2 f_2)$ exists almost everywhere, and*

$$D_{*a}^n(c_1 f_1 + c_2 f_2) = c_1 D_{*a}^n f_1 + c_2 D_{*a}^n f_2.$$

Proof. This linearity property of the fractional differential operator is an immediate consequence of the definition of D^n_{*a}. □

For the formula of Leibniz, we only state the case $0 < n < 1$ explicitly.

Theorem 3.17 (Leibniz' formula for Caputo operators). *Let $0 < n < 1$, and assume that f and g are analytic on $(a-h, a+h)$. Then,*

$$\begin{aligned} D^n_{*a}[fg](x) &= \frac{(x-a)^{-n}}{\Gamma(1-n)} g(a)(f(x)-f(a)) + (D^n_{*a}g(x)) f(x) \\ &\quad + \sum_{k=1}^{\infty} \binom{n}{k} \left(J_a^{k-n} g(x)\right) D^k_{*a} f(x). \end{aligned}$$

Proof. We apply the definition of the Caputo derivative and find

$$D^n_{*a}[fg] = D^n_a[fg - f(a)g(a)] = D^n_a[fg] - f(a)g(a)D^n_a[1].$$

Next we use Leibniz' formula for Riemann–Liouville derivatives and find

$$D^n_{*a}[fg] = f(D^n_a g) + \sum_{k=1}^{\infty} \binom{n}{k} (D^k_a f)(J_a^{k-n} g) - f(a)g(a)D^n_a[1].$$

Now we add and subtract $f \cdot g(a)(D^n_a[1])$ and rearrange to obtain

$$\begin{aligned} D^n_{*a}[fg] &= f(D^n_a[g - g(a)]) + \sum_{k=1}^{\infty} \binom{n}{k} (D^k_a f)(J_a^{k-n} g) \\ &\quad + g(a)(f - f(a))D^n_a[1] \\ &= f \times (D^n_{*a} g) + \sum_{k=1}^{\infty} \binom{n}{k} (D^k_{*a} f)(J_a^{k-n} g) + g(a)(f - f(a)) \times D^n_a[1] \end{aligned}$$

where we have used the fact that, for $k \in \mathbb{N}$, $D^k_a = D^k = D^k_{*a}$. To finally complete the proof it only remains to use the explicit expression for $D^n_a[1]$ from Example 2.4. □

Remark 3.5. For Faà di Bruno's formula (the chain rule) for Caputo operators we may combine eq. (2.6), i.e. the corresponding rule for Riemann–Liouville operators, and Lemma 3.4. This yields that, once again under suitable assumptions on the functions f and g that we shall not specify explicitly, it has the form

$$\begin{aligned} & D^n_{*a}[f(g(\cdot))](x) \\ &= \sum_{k=1}^{\infty} \binom{n}{k} \frac{k!(x-a)^{k-n}}{\Gamma(k-n+1)} \sum_{\ell=1}^{k} (D^\ell f)(g(x)) \sum_{(a_1,\dots,a_k)\in A_{k,\ell}} \prod_{r=1}^{k} \frac{1}{a_r!} \left(\frac{D^r g(x)}{r!}\right)^{a_r} \\ &\quad + \frac{(x-a)^{-n}}{\Gamma(1-n)} f(g(x)) - \sum_{k=0}^{\lceil n\rceil -1} \frac{D^k[f(g(\cdot))](a)}{\Gamma(k-n+1)} (x-a)^{k-n} \end{aligned}$$

where, as in the Riemann–Liouville case discussed in Theorem 2.19, $(a_1,\dots,a_k) \in A_{k,\ell}$ means that

$$a_1,\dots,a_k \in \mathbb{N}_0, \qquad \sum_{r=1}^{k} r a_r = k \quad \text{and} \quad \sum_{r=1}^{k} a_r = \ell.$$

3.2 Nonclassical Representations of Caputo Operators

In the previous section we have developed a number of different representations for Caputo differential operators under various assumptions on the function to be differentiated; see, e.g., Definition 3.2, Definition 3.1 in combination with Theorem 3.1, Lemma 3.2, Lemma 3.3, or Lemma 3.4. Essentially, all these representations consist of a combination of a convolution integral and some sort of a differential operator (or the limit of a discrete version of this). This approach immediately reveals the non-locality of the Caputo operator and provides a natural approach to handling this property. However, in some applications it has turned out that other ways to express the non-locality are more helpful. Two such alternative representations have recently been developed independently by Yuan and Agrawal [194] and Singh and Chatterjee [26, 177]. Our treatment of these two closely related methods is based on the generalizations provided and analyzed in [38].

We first recall the details of the method proposed by Yuan and Agrawal [194]. They have only discussed the case $0 < n < 1$. An extension to $1 < n < 2$ has been provided by Trinks and Ruge [185]. We will not impose any such restriction on the size of n. However, our techniques do require that $n \notin \mathbb{N}$ throughout this section. Since this only excludes cases that are not truly fractional anyway, this is not a substantial limitation. It can easily be seen that our approach reduces to the original scheme of Yuan and Agrawal if $0 < n < 1$.

The approach is tailored to functions $f \in C^{\lceil n \rceil}[a,b]$. In view of Theorem 3.1 this feature allows us to use the representation of Definition 3.1 for the Caputo derivative. Then we define an auxiliary bivariate function $\phi : (0,\infty) \times [a,b] \to \mathbb{R}$ by

$$\phi(w,x) := (-1)^{\lfloor n \rfloor} \frac{2 \sin \pi n}{\pi} w^{2n-2\lceil n \rceil+1} \int_a^x f^{(\lceil n \rceil)}(\tau) \mathrm{e}^{-(x-\tau)w^2} \,\mathrm{d}\tau. \tag{3.2}$$

With this notation we obtain the following generalization of a result presented in [194, §2]:

Theorem 3.18. *Under the above assumptions,*

$$D_{*a}^{n} f(x) = \int_0^\infty \phi(w,x)\,\mathrm{d}w. \tag{3.3}$$

In addition, for fixed $w > 0$ the function $\phi(w,\cdot)$ satisfies the differential equation

$$\frac{\partial}{\partial x}\phi(w,x) = -w^2\phi(w,x) + (-1)^{\lfloor n\rfloor}\frac{2\sin\pi n}{\pi}w^{2n-2\lceil n\rceil+1}f^{(\lceil n\rceil)}(x) \tag{3.4}$$

subject to the initial condition $\phi(w,a) = 0$.

Notice that the differential equation (3.4) is effectively an ordinary differential equation since we assume w to be a fixed parameter. Moreover it is a differential equation of order 1 and hence of classical (not fractional in the strict sense) type. In addition we note that the differential equation is linear and inhomogeneous and that it has constant coefficients. Therefore it is a simple matter to compute the solution of the initial value problem explicitly, and of course this computation reproduces the representation (3.2).

Proof. Bearing in mind the definition of the Gamma function (Definition 1.2), Theorem D.3 and the obvious identity

$$\sin\pi(n-\lceil n\rceil+1) = (-1)^{\lfloor n\rfloor}\sin\pi n \qquad (n\notin\mathbb{N})$$

we obtain that

$$\begin{aligned}
D^n_{*a}f(x) &= \frac{1}{\Gamma(\lceil n\rceil-n)}\int_a^x (x-\tau)^{\lceil n\rceil-n-1}f^{(\lceil n\rceil)}(\tau)\,\mathrm{d}\tau \\
&= \frac{1}{\Gamma(n-\lceil n\rceil+1)\Gamma(\lceil n\rceil-n)} \\
&\quad\times\int_a^x\int_0^\infty \mathrm{e}^{-z}z^{n-\lceil n\rceil}\,\mathrm{d}z\,(x-\tau)^{\lceil n\rceil-n-1}f^{(\lceil n\rceil)}(\tau)\,\mathrm{d}\tau \\
&= \frac{\sin\pi(n-\lceil n\rceil+1)}{\pi}\int_a^x\int_0^\infty \mathrm{e}^{-z}\left(\frac{z}{x-\tau}\right)^{n-\lceil n\rceil+1}\frac{1}{z}f^{(\lceil n\rceil)}(\tau)\,\mathrm{d}z\,\mathrm{d}\tau \\
&= (-1)^{\lfloor n\rfloor}\frac{\sin\pi n}{\pi}\int_a^x\int_0^\infty \mathrm{e}^{-z}\left(\frac{z}{x-\tau}\right)^{n-\lceil n\rceil+1}\frac{1}{z}f^{(\lceil n\rceil)}(\tau)\,\mathrm{d}z\,\mathrm{d}\tau.
\end{aligned}$$

We may now apply the substitution $z = (x-\tau)w^2$ in the inner integral and note that Fubini's Theorem allows us to interchange the order of the integrations since $f^{(\lceil n\rceil)}$ is assumed to be continuous. This yields

$$\begin{aligned}
D^n_{*a}f(x) &= (-1)^{\lfloor n\rfloor}\frac{2\sin\pi n}{\pi}\int_a^x\int_0^\infty \mathrm{e}^{-(x-\tau)w^2}w^{2n-2\lceil n\rceil+1}f^{(\lceil n\rceil)}(\tau)\,\mathrm{d}w\,\mathrm{d}\tau \\
&= \int_0^\infty (-1)^{\lfloor n\rfloor}\frac{2\sin\pi n}{\pi}w^{2n-2\lceil n\rceil+1}\int_a^x \mathrm{e}^{-(x-\tau)w^2}f^{(\lceil n\rceil)}(\tau)\,\mathrm{d}\tau\,\mathrm{d}w
\end{aligned}$$

and, recalling the definition (3.2) of ϕ, we deduce (3.3).

Next we differentiate the definition (3.2) of ϕ with respect to x. The classical rules for the differentiation of parameter integrals with respect to the parameter immediately give (3.4). Finally the fact that $\phi(w,a)=0$ for $w>0$ is also a direct consequence of (3.2) because the integrand of the integral on the right-hand side of (3.2) is continuous. □

The approach proposed by Chatterjee [26] and investigated further in [177] is also based on expressing the fractional derivative of the given function f in the form of an integral over $(0,\infty)$ whose integrand can be computed as the solution of a first-order initial value problem. Specifically, it is based on the following analogue of Theorem 3.18. The result essentially states that we may replace the integrand ϕ by a function ϕ^* which can be characterized as the solution of a different first-order initial value problem.

Theorem 3.19. *Let $f \in C^{\lceil n\rceil}[a,b]$. Moreover, for fixed $w>0$, let $\phi^*(w,\cdot)$ be the solution of the differential equation*

$$\frac{\partial}{\partial x}\phi^*(w,x) = -w^{1/(n-\lceil n\rceil-1)}\phi^*(w,x) + \frac{(-1)^{\lfloor n\rfloor}\sin \pi n}{\pi(n-\lceil n\rceil+1)} f^{(\lceil n\rceil)}(x) \tag{3.5}$$

subject to the initial condition $\phi^*(w,a)=0$. Then, we have

$$\phi^*(w,x) = \frac{(-1)^{\lfloor n\rfloor}\sin \pi n}{\pi(n-\lceil n\rceil+1)} \int_0^x f^{(\lceil n\rceil)}(\tau)\exp\left(-(x-\tau)w^{1/(n-\lceil n\rceil+1)}\right)\mathrm{d}\tau \tag{3.6}$$

and

$$D^n_{*a}f(x) = \int_0^\infty \phi^*(w,x)\,\mathrm{d}w. \tag{3.7}$$

Proof. The proof of this result is almost identical to the proof of Theorem 3.18; one only needs to replace the substitution $z=(x-\tau)w^2$ by $z=(x-\tau)w^{1/(n-\lceil n\rceil+1)}$ and use the functional equation of the Gamma function, $u\Gamma(u)=\Gamma(u+1)$. We leave the details to the reader. □

In many applications of these representations [41, 118, 171, 177, 185] it is important to have some additional knowledge about the behaviour of the function ϕ in eq. (3.2) or the function ϕ^* in eq. (3.6), respectively. The most important of these properties are summarized in the following theorems. Here, the symbol $\alpha(v)\sim\beta(v)$ means that there exist two strictly positive constants A and B such that $|\alpha(v)/\beta(v)|\in[A,B]$ as v tends to the indicated limit. We begin with the function ϕ arising in the original Yuan-Agrawal representation that we had given in our Theorem 3.18.

Theorem 3.20. *Let $x\in(a,b)$ be fixed and $0<n\notin\mathbb{N}$, and assume that there exists some $C>0$ such that $|f^{(\lceil n\rceil)}(\tilde{x})|>C$ for all $\tilde{x}\in[a,b]$.*

(a) The function $\phi(\cdot,x)$ defined in (3.2) behaves as

$$\phi(w,x)\sim w^{2n-2\lceil n\rceil+1} \quad \text{as } w\to 0. \tag{3.8}$$

(b) Moreover,

$$\phi(w,x) \sim w^{2n-2\lceil n\rceil-1} \quad \text{as } w \to \infty. \tag{3.9}$$

(c) We have $\phi(\cdot,x) \in C^\infty(0,\infty)$.

Remark 3.6. The condition that $f^{(\lceil n\rceil)}$ be bounded away from zero is a technical condition required in order to keep the proof simple and to keep the result valid for *all* $x \in (a,b)$. Using more complicated techniques, one could show that the same asymptotic behaviour is present for *almost all* $x \in (a,b)$ under substantially weaker conditions. Thus it is justified to say that the asymptotic behaviour described in Theorem 3.20 is the behaviour that one may reasonably expect for the function ϕ unless the given function f is of a highly exceptional nature.

Proof. For part (a), a partial integration gives

$$\begin{aligned}
&\int_a^x f^{(\lceil n\rceil)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau \\
&= f^{(\lceil n\rceil-1)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\Big|_{\tau=a}^{\tau=x} - w^2\int_a^x f^{(\lceil n\rceil-1)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau \\
&= f^{(\lceil n\rceil-1)}(x) - f^{(\lceil n\rceil-1)}(a)\mathrm{e}^{-xw^2} - w^2\int_0^x f^{(\lceil n\rceil-1)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau.
\end{aligned}$$

Since x is fixed, the rightmost integral obviously remains bounded as $w \to 0$, and hence we conclude

$$\lim_{w\to 0}\int_a^x f^{(\lceil n\rceil)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau = f^{(\lceil n\rceil-1)}(x) - f^{(\lceil n\rceil-1)}(0). \tag{3.10}$$

Inserting this relation into the definition (3.2) of ϕ we obtain the first claim.

For the proof of (b), we write

$$\begin{aligned}
&w^2\int_a^x f^{(\lceil n\rceil)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau \\
&= w^2\int_a^{x-w^{-1}} f^{(\lceil n\rceil)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau + w^2\int_{x-w^{-1}}^x f^{(\lceil n\rceil)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau \\
&= f^{(\lceil n\rceil)}(\xi_1)w^2\int_a^{x-w^{-1}} \mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau + f^{(\lceil n\rceil)}(\xi_2)w^2\int_{x-w^{-1}}^x \mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau \\
&= f^{(\lceil n\rceil)}(\xi_1)(\mathrm{e}^{-w} - \mathrm{e}^{-w^2}) + f^{(\lceil n\rceil)}(\xi_2)(1-\mathrm{e}^{-w})
\end{aligned}$$

with some $\xi_1 \in [a, x-w^{-1}]$ and $\xi_2 \in [x-w^{-1}, x]$ because of the Mean Value Theorem. Now, as $w \to \infty$, $f^{(\lceil n\rceil)}(\xi_1)$ remains bounded whereas $\mathrm{e}^{-w} - \mathrm{e}^{-w^2} \to 0$. Thus the first summand on the right-hand side vanishes. For the second summand we have $1-\mathrm{e}^{-w} \to 1$ and $f^{(\lceil n\rceil)}(\xi_2) \to f^{(\lceil n\rceil)}(x)$ because $\xi_2 \in [x-w^{-1}, x]$. Thus, we conclude

$$\lim_{w\to\infty} w^2\int_0^x f^{(\lceil n\rceil)}(\tau)\mathrm{e}^{-(x-\tau)w^2}\,\mathrm{d}\tau = f^{(\lceil n\rceil)}(x).$$

Inserting this relation into the definition of ϕ, we arrive at

$$\phi(w,x) = (-1)^{\lfloor n \rfloor} \frac{2\sin \pi n}{\pi} w^{2n-2\lceil n \rceil -1} [f^{(\lceil n \rceil)}(x) + o(1)] \tag{3.11}$$

which completes the proof of (b).

Finally, part (c) follows directly from the definition of ϕ that we had given in eq. (3.2). □

We can also provide a corresponding result for the function ϕ^* used in Chatterjee's representation (Theorem 3.19). The behaviour of this function ϕ^*, which is defined in eq. (3.6), can be described as follows.

Theorem 3.21. *Let $x \in (a,b)$ be fixed and $0 < n \notin \mathbb{N}$, and assume that there exists some $C > 0$ such that $|f^{(\lceil n \rceil)}(\tilde{x})| > C$ for all $\tilde{x} \in [0,X]$.*

(a) The function $\phi^*(\cdot,x)$ described in eq. (3.6) behaves as

$$\phi^*(w,x) \sim 1 \quad \text{as } w \to 0. \tag{3.12}$$

(b) Moreover,

$$\phi^*(w,x) \sim w^{-1/(n-\lceil n \rceil+1)} \quad \text{as } w \to \infty. \tag{3.13}$$

(c) We have $\phi^*(\cdot,x) \in C^\infty(0,\infty)$.

The proof proceeds along the same lines as the proof of Theorem 3.20.

Theorems 3.20 and 3.21 allow us to compare the analytical properties of the function ϕ used by Yuan and Agrawal and the function ϕ^* proposed by Chatterjee. First of all we note that both functions possess infinitely many derivatives (in the classical sense) with respect to the first variable. However, there are significant differences in the asymptotic behaviour of the functions as the first variable tends to either end of the interval $(0,\infty)$ over which the functions need to be integrated in order to compute the Caputo derivative $D^n_{*a}f$.

To be precise, for $w \to 0$ the function $\phi(w,x)$ exhibits an asymptotic behaviour of the form $w^{2n-2\lceil n \rceil -1}$ according to Theorem 3.20 (a). The exponent of w here is always strictly between -1 and $+1$. This asserts the integrability of $\phi(w,x)$ with respect to w near $w = 0$ at least in the improper sense. However, we can expect a smooth behaviour near this end point of the integration interval only if the exponent is an integer, and this is the case if and only if $n = k + 1/2$ with some $k \in \mathbb{N}_0$. For all other values of n the behaviour is less regular. This irregularity needs to be taken into account carefully when one tries to use this approach in a numerical algorithm [38, 118]. Theorem 3.21 (a) demonstrates that the function ϕ^* is easier to handle in this respect since here we always have that $\phi^*(w,x)$ remains bounded by nonzero constants from above and below.

The behaviour of the integrands $\phi(w,x)$ and $\phi^*(w,x)$ for $w \to \infty$ also exhibits substantial differences. As shown in Theorem 3.20 (b), the Yuan-Agrawal integrand $\phi(w,x)$ behaves as $w^{2n-2\lceil n \rceil -1}$. The exponent of w here is always contained in the interval $(-3,-1)$. This is just about fast enough to make sure that the improper integral exists. On the other hand, according to Theorem 3.21 (b), the Chatterjee integrand $\phi^*(w,x)$ behaves in a way that depends on n in a somewhat more complicated fashion: If $n = k + \varepsilon$ with some $k \in \mathbb{N}_0$ and $0 < \varepsilon < 1$ then the exponent in question is $-1/\varepsilon$. This is always less than -1, and hence the improper integral converges. In this respect we have no difference to the Yuan-Agrawal method. However, if ε is close to 0 then the exponent of course remains negative but it may be arbitrarily large in modulus, leading to a much faster (but still algebraic) decay of the integrand. In particular, in contrast to the Yuan-Agrawal method there is no lower bound on the exponent as n runs through all the admissible numbers. For numerical work, a rapidly decaying integrand is preferable, so at least for $n = k + \varepsilon$ with $k \in \mathbb{N}$ and ε close to 0 the approach via Theorem 3.19 has some advantages over the path via Theorem 3.18. It should be pointed out though that, from the point of view of approximation theory (see, e.g., the survey article [122]), an ideal integrand (i.e. an integrand that can be handled very nicely by a numerical algorithm) would decay exponentially as $w \to \infty$, i.e. much faster than we can ever hope even for ϕ^*.

Exercises

Exercise 3.1. Let $f(x) = \cos \lambda x$ for some $\lambda > 0$.

(a) Determine the functions $D_0^n f(x)$ and $D_{*0}^n f(x)$ for arbitrary $n > 0$.
(b) For which values of x are these derivatives defined?
(c) Investigate these derivatives with respect to continuity and differentiability.
(d) Draw a sketch of the derivatives for some values of n.

Exercise 3.2. Prove the identities stated in Appendix B.

Exercise 3.3. Work out the details of the proofs of Theorems 3.19 and 3.21.

http://www.springer.com/978-3-642-14573-5

The Analysis of Fractional Differential Equations
An Application-Oriented Exposition Using Differential Operators of Caputo Type
Diethelm, K.
2010, VIII, 247p. 20 illus., 10 illus. in color., Softcover
ISBN: 978-3-642-14573-5