Notes on Fractional and Irrational Calculus and Differential Equations for Engineers: Mathematics, Modeling and Numerical Methods

Bill Goodwine

January 10, 2022

Contents

Chapter 1

Introduction

1.1 Reasons to Study Fractional Calculus

Some people may study fractional calculus because it is inherently interesting to them. Other people need more or different reasons. A few reasons include:

- Fractional calculus expands the descriptive power of calculus beyond the familiar integer order derivatives and the basic concept of rate of change. This will yield more accurate descriptive equations when a system truly is fractional.
- Fractional calculus can expand scientific understanding when it models things that are traditionally very difficult to describe mathematically.
- Integer order derivatives are *local* needing information only in a neighborhood of a point. As will be seen, fractional derivatives are *non-local* and can describe systems with such a feature. For example, $F = ma$ right *now*. How something accelerates only depends on the forces right now, not the forces at any time in the past. For systems with "memory" effects, or analogous non-local spatial effects, the differential equation describing it most accurately may be fractional.

It is not surprising that "normal" integer-order calculus works well for most engineering applications over the past few centuries. Newton, who was one of the very early developers of calculus, clearly had a strong interest in mechanics [Newton, 1687]. However, there are a couple topics of more recent interest in engineering for which fractional calculus may be important. One area is bioengineering, where traditional mechanics and electromagnetism do not fully describe the system and in which non-local time effects may occur. Another is very large scale systems that may be easier to consider as having an infinite number of components, rather than keeping track of every single part. Examples of each type are given in this chapter. Finally, fractional calculus has been used as a new tool to try to match measured responses to "generally hard to model" system, *e.g.*, systems with friction, stiction, etc.

1.2 Introductory Concepts

Anyone reading this book should be familiar with the notion of the first, second and higher derivatives, *e.g.*, for $f(t) = t^3 + 5t^2 + 2$,

$$\frac{\mathrm{d}f}{\mathrm{d}t}(t) = 3t^2 + 10t \tag{1.1}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}(t) = 6t + 10, \tag{1.2}$$

etc. Also we naturally think of integrals in an antiderivative sense, *e.g.*,

$$\int f(t)\,\mathrm{d}t = \frac{1}{4}t^4 + \frac{5}{3}t^3 + 5t^2 + c \tag{1.3}$$

and we adopt a notation of $f^{(1)}(t)$ as the first derivative, $f^{(2)}(t)$ as the second derivative and $f^{(n)}(t)$ as the nth derivative as well as $f^{(-1)}(t)$ as f integrated one time, $f^{(-2)}$ as f integrated two times, *etc.*

A mathematically curious reader may already be wondering if there are any derivatives "in between" the integer ones. For example, is there a one-half derivative:

$$\frac{\mathrm{d}^{\frac{1}{2}} f}{\mathrm{d}t^{\frac{1}{2}}}(t) = f^{(\frac{1}{2})} = ?. \tag{1.4}$$

There is not an immediate obvious answer to this because of the fact that the integer order derivative (as is the integral) is defined as a limit

$$\frac{\mathrm{d}f}{\mathrm{d}t}(t) = \lim_{\Delta t \to 0} = \frac{f\left(t + \Delta t\right) - f\left(t\right)}{\Delta t} \tag{1.5}$$

and that is a discrete operation. There is not a natural half way to do it.

Basically we want to generalize the notion of the derivative. In a sense, if we define something to give the, say α derivative where $\alpha \in \mathbb{R}$, *i.e.*, α is a real number, then all we really need is that when α is an integer we get the usual definition of that integer order derivative. In between there may be lots of different options (there are!), but it makes sense to set some other basic requirements we want a fractional-order derivative to satisfy.

Example 1.1. *Consider $f(t) = t^2$ with the first and second derivatives $f^{(1)}(t) = 2t$ and $f^{(2)}(t) = 2$, respectively. We should expect that the $1/2$ derivative is, in some qualitative sense, "between" $f(t)$ and $f^{(1)}(t)$, and that the $3/2$ derivative is "between" the first and second derivatives, as is illustrated in Figures 1.1 and 1.2.*

We also expect that as we vary the order of the derivative, say from 0 to 1, for low values of the order the result is near the zeroth derivative, and for values of the fractional order near one, it is near the first derivative. This is illustrated in Figure 1.3.

This example motivates our first desirable attribute of a fractional derivative.

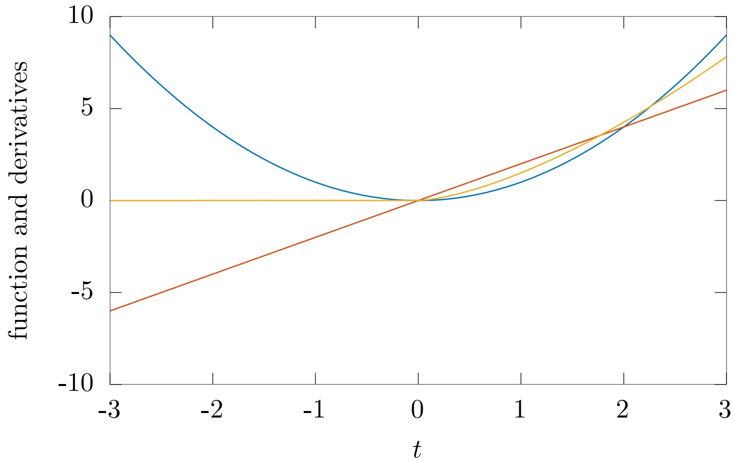

Figure 1.1: The zeroth and first derivatives of $f(t) = t^2$, (blue and red curves). A prospective half derivative (gold).

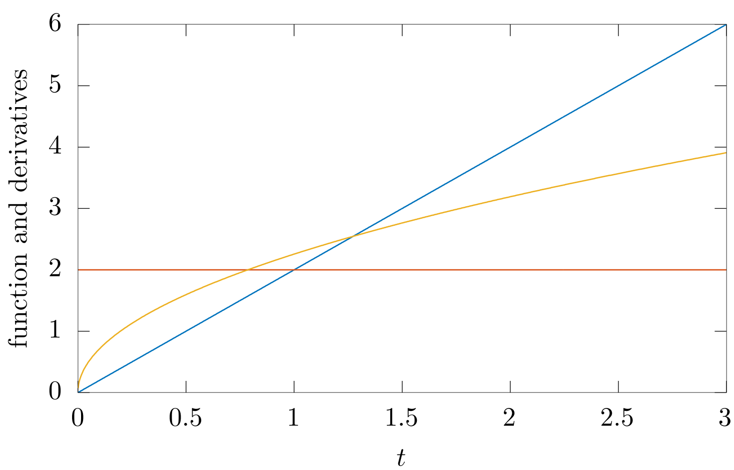

Figure 1.2: The first and second derivatives of $f(t) = t^2$ (blue and red curves) with a prospective 3/2 derivative (gold).

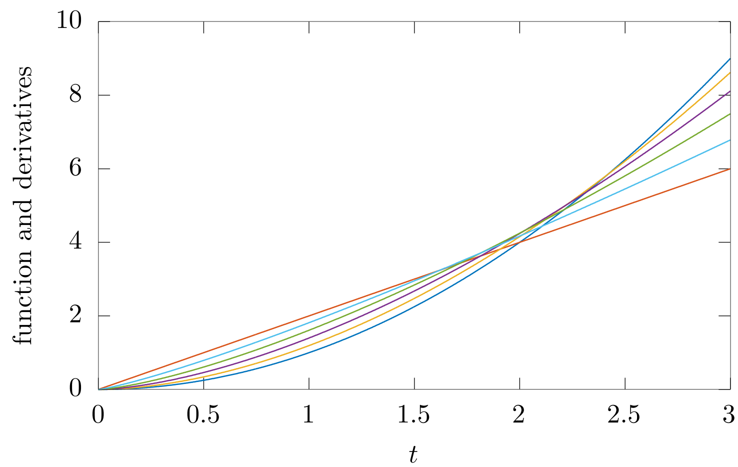

Figure 1.3: The function (blue) and the first derivative (red). The 0.2, 0.4, 0.6 and 0.8th (yellow, purple, green, light blue) order derivatives "move" from the zeroth derivative to the first derivative.

Attribute 1.1. *For a fractional derivative, $f^{(\alpha)}(t)$, if α is near an integer value, we expect the α derivative to be near that integer derivative of $f(t)$. As α varies between integer values, we expect that $f^{(\alpha)}(t)$ varies in a reasonable manner between those integer values.*

1.3 Fractional Derivatives of some Elementary Functions

1.3.1 Sine and Cosine Functions

As a first step into some real functions, we consider sine and cosine.

Example 1.2. *Consider $f(t) = \sin(t)$. The nice thing about sines and cosines are their relatively simple derivatives. In fact, from the pattern*

$$\frac{\mathrm{d}}{\mathrm{d}t}\sin(t) = \cos(t)$$
$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}\sin(t) = -\sin(t)$$

as illustrated in Figure 1.4, it is clear that the derivative for this function just shifts it to the left by $\pi/2$. So clearly we would expect that the $1/2$ derivative just shifts it to the left by $\pi/4$, or the $1/2$ integral shifts it right by $\pi/3$, etc. Therefore, a reasonable guess for the half derivative would be as illustrated in Figure 1.5.

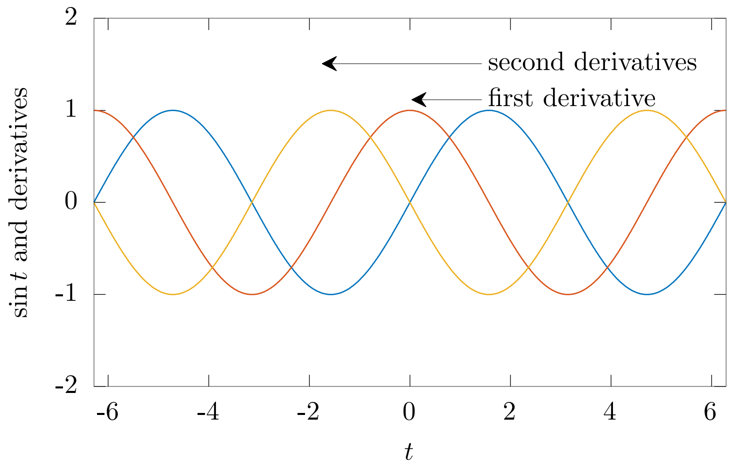

Figure 1.4: First (red) and second (gold) derivative of $\sin(t)$ (blue). The arrows indicate that derivative is just shifts the curve to the left by $\pi/2$ for each derivative.

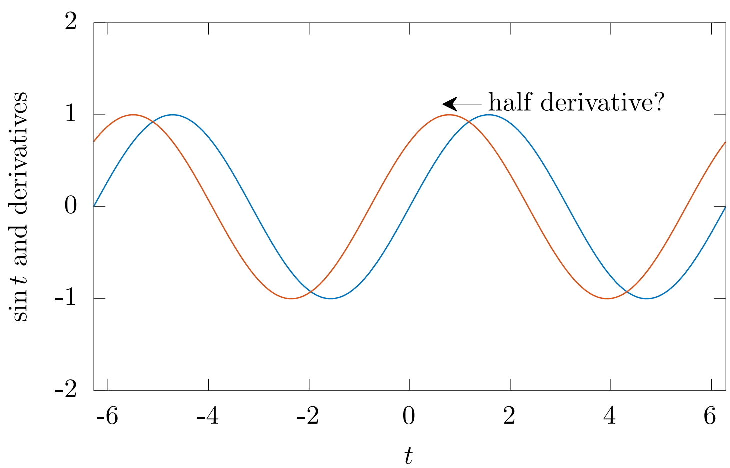

Figure 1.5: Reasonable assumption for the half derivative (red) of $\sin 9t)$ (blue).

1.3.2 Monomials and Polynomials

Being able to take the fractional derivative of sine and cosine functions is nice, but we would like to do more. The next easiest class of functions is polynomials. Consider

$$f(t) = t^k \tag{1.6}$$

which is easy to differentiate a few times and figure out the pattern:

$$\frac{\mathrm{d}f}{\mathrm{d}t}(t) = kt^{k-1} \tag{1.7}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}(t) = k(k-1)\,t^{k-2} \tag{1.8}$$

$$\frac{\mathrm{d}^3 f}{\mathrm{d}t^3}(t) = k(k-1)(k-2)\,t^{k-3} \tag{1.9}$$

$$\vdots = \vdots \tag{1.10}$$

$$\frac{\mathrm{d}^n f}{\mathrm{d}t^n}(t) = \frac{k!}{(k-n)!}t^{k-n}, \qquad n \le k. \tag{1.11}$$

Since we are looking to define a fractional derivative of t^k, we need to see what is allowed and not allowed for n to take on fractional values in Equation 1.11. The exponent of t can be a fraction (more about that later, but engineers are so used to it they may not remember where it came from). What is definitely a problem, though, is the factorial in the denominator: if k is a natural number and n is not an integer, then the factorial is not defined.

The factorial function is just a series of values, so it seems we can generalize the derivative of t^k if we can find a curve through the factorial values. In fact, of course, it has been done and it is the gamma function defined by

$$\Gamma(z) = \int_0^\infty x^{z-1}e^{-x}\,\mathrm{d}x \tag{1.12}$$

which is plotted for positive real values of z in Figure 1.6. Clearly, for integer values

$$z! = \Gamma(z+1)\,. \tag{1.13}$$

So, returning to Equation 1.11, it seems that all we need to do to define a fractional derivative is to replace the factorial in the denominator with the gamma function shifted by one. Purely for aesthetics, we might as well use a gamma function in the numerator as well, which also then would allow for a fractional k. So, we have the following seemingly legitimate definition of a fractional derivative for a monomial.

Definition 1.1. *For the monomial, t^k, $k \in \mathbb{R}$, define a*[1] *proposed fractional derivative*

$$\boxed{\frac{\mathrm{d}^\alpha}{\mathrm{d}t^\alpha}t^k = \frac{\Gamma(k+1)}{\Gamma(k+1-\alpha)}t^{k-\alpha}} \tag{1.14}$$

[1] This is *a* fractional derivative, rather than *the* fractional derivative because there are many definitions.

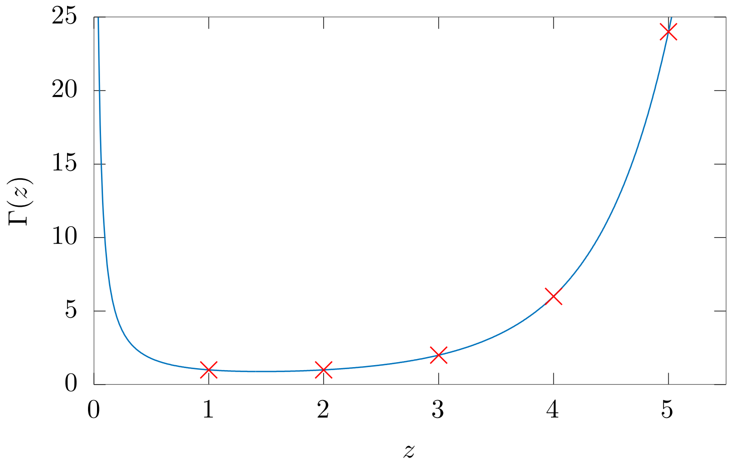

Figure 1.6: Plot of $\Gamma(z)$ (blue curve) and some factorials, $(z-1)!$ (red).

for real values of α*, i.e.,* $\alpha \in \mathbb{R}$.

This definition was used to make the fractional derivative curves in Figures 1.1 through 1.3, so in a sense it has been validated. Also, because everything we have done is linear in t, we can use this definition for monomials and extend it term-by-term to polynomials.

For fun, we will do and example with a different k and also include negative values for α to see if integral-like ideas appear.

Example 1.3. *Consider* $f(t) = t^{\frac{1}{2}} = \sqrt{t}$*. Various fractional order derivatives and integrals computed using Equation 1.14 in Definition 1.1 are illustrated in Figure 1.7.*

1.3.3 Exponentials

Since

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} e^t = e^t$$

we would like it to also hold for when n is not an integer. Similarly,

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n} e^{\alpha t} = \alpha^n e^{\alpha t}$$

we can require that the same, or something close to this, be true when n is not an integer.

At this point we can compute things we consider to be fractional derivatives and integrals of sines and cosines, monomials, by extension from monomials, polynomials if we do them term-by-term and exponentials. Before we generalize further, in order to develop a very important property of the fractional derivatives, we need to go a long way back and consider fractional exponents.

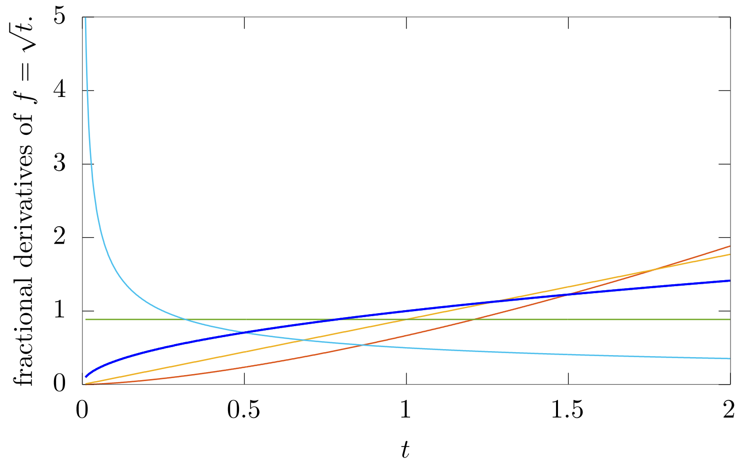

Figure 1.7: Various fractional derivatives using Definition 1.1 for $f(t) = \sqrt{t}$. The thick blue curve is the function, or zeroth derivative. The red curve is the $\alpha = -1$ derivative, which does correspond to the integral. The gold curve is the $\alpha = -1/2$ derivative, or the $1/2$ integral. The green and light blue curves are the $1/2$ and first derivatives, respectively.

1.4 The Law of Indices

Engineers deal with fractional and negative exponents so often that it is easy to lose track of why they actually make sense. The exponent is defined for natural numbers (integers greater than zero) as the number of times the base is multiplied by itself, *i.e.*,

$$t^n = \underbrace{t \times t \times t \cdots t \times t}_{n \text{ times}}. \tag{1.15}$$

An obvious property of this is that for two natural number exponents

$$\begin{aligned} (t^n) \times (t^m) &= \left(\underbrace{t \times t \times t \cdots t \times t}_{n \text{ times}} \right) \times \left(\underbrace{t \times t \times t \cdots t \times t}_{m \text{ times}} \right) &(1.16)\\ &= \underbrace{t \times t \times t \cdots t \times t}_{n+m \text{ times}} &(1.17)\\ &= t^{n+m} &(1.18) \end{aligned}$$

which also immediately leads to

$$(t^n)^p = t^{n \times p}. \tag{1.19}$$

This notion of adding indices can be used to *define* negative and fractional exponents by requiring that Equations 1.18 and 1.19 hold for all rational values as well (negative values, zero and fractional values).

For negatives values, consider n and m to be positive integers with $n > m$, and if we require that $(t^n) \times (t^{-m}) = t^{n-m}$, then the only way for the $-m$ to take away powers is for it to mean division, or in too much detail

$$t^n \times t^{-m} = \frac{\overbrace{t \times t \times t \cdots t \times t}^{n \text{ times}}}{\underbrace{t \times t \cdots t \times t}_{m \text{ times}}} = \underbrace{t \times t \cdots t \times t}_{n-m \text{ times}}. \tag{1.20}$$

For fractional values Equation we can use Equation 1.19 so that

$$t^{\frac{n}{m}} = y \qquad \Longrightarrow \qquad \left(t^{\frac{n}{m}}\right)^m = y^m \qquad \Longrightarrow \qquad t^n = y^m \tag{1.21}$$

which gives the meaning that y is the number that if you raise it to the mth power gives t to the nth power, e.g. in the simple case of $1/2$, y is the number that if you square you get t.

These exercises in indices are important, because they hold for integer order derivatives

$$\frac{\mathrm{d}^n}{\mathrm{d}t^n}\left(\frac{\mathrm{d}^m}{\mathrm{d}t^m} f(t)\right) = \frac{\mathrm{d}^{n+m} f}{\mathrm{d}t^{n+m}}(t). \tag{1.22}$$

By insisting that the same property hold when n and m are fractional, will help in generalizing the derivative to non-integer values for a larger class of functions that simple sines, cosines and polynomials. In fact, we may as well call it an attribute.

Attribute 1.2. *For real values of α and β*

$$\frac{\mathrm{d}^\alpha}{\mathrm{d}t^\alpha}\left(\frac{\mathrm{d}^\beta}{\mathrm{d}t^\beta} f(t)\right) = \frac{\mathrm{d}^{\alpha+\beta} f}{\mathrm{d}t^{\alpha+\beta}}(t). \tag{1.23}$$

Even in integer order calculus, integration and differentiation are not exactly inverses because an indefinite integral will have a constant of integration. In other words, if we take $f(t)$ and differentiate it and then integrate it, we get $f(t)+c$, but if we integrate and then differentiate, we get $f(t)$. The idea is clear enough, but it turns out that this complication does affect things.

1.5 Examples

1.5.1 Mechanical System and Frequency Domain Example

Multiplication by s in the frequency domain corresponds to differentiation by t in the time domain if we use the usual variables in the Laplace transform

$$\mathcal{L}\{f(t)\} = \int_{0^-}^{\infty} f(t)e^{-st}\,\mathrm{d}t = F(s), \tag{1.24}$$

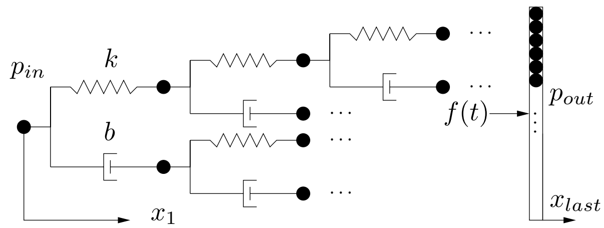

Figure 1.8: Infinite tree of springs and dampers.

i.e.,

$$\mathcal{L}\left\{\frac{\mathrm{d}f}{\mathrm{d}t}(t)\right\} = sF(s) - f(0) \tag{1.25}$$

or assuming zero initial conditions

$$\mathcal{L}\left\{\frac{\mathrm{d}f}{\mathrm{d}t}(t)\right\} = sF(s). \tag{1.26}$$

Higher derivatives are just increased exponents on the s, *e.g.*,

$$\mathcal{L}\left\{\frac{\mathrm{d}^n f}{\mathrm{d}t^n}(t)\right\} = s^n F(s) \tag{1.27}$$

again assuming zero initial conditions.

Of course, the half derivative then would correspond to s raised to the one-half power:

$$\mathcal{L}\left\{\frac{\mathrm{d}^{\frac{1}{2}} f}{\mathrm{d}t^{\frac{1}{2}}}(t)\right\} = s^{\frac{1}{2}} F(s). \tag{1.28}$$

It turns out that irrational transfer functions can arise rather easily in two types of cases:

1. systems with an infinite number of components, and
2. systems with non-local interactions.

The following example illustrates the first case. Non-locality will be inherent in the more general definitions of fractional derivatives we develop subsequently, so examples will be deferred until later.

Example 1.4. *Consider the tree network of springs and dampers illustrated in Figure 1.8. The position of the left-most node is $x_1(t)$ and the right-most node by $x_{last}(t)$. Note all the nodes on the right are in the same position, so in effect they make one node. The equations of motion for this system can be determined by applying the relatively simple series and parallel rules for the springs and dampers. In order to change the network, a force must be exerted on one end, and an an equal and opposite force on the other, $f(t)$.*

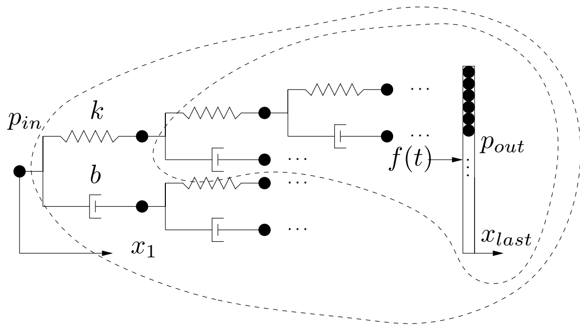

Figure 1.9: Self-similar network where the transfer function from the beginning to end of each outlined region must be equal.

It turns out that if we consider the network to be infinitely large, *i.e.*, an infinite number of bifurcating generations, it is easier to determine the transfer function describing we want, which is

$$G(s) = \frac{X_{last}(s) - X_1(s)}{F(s)} \tag{1.29}$$

and which describes the relationship between the applied force and deflection of the network. The reason it is easier, is that the network is *self-similar*. If there are an infinite number of generations, if we look at any specific node, then the transfer function from that node to the right end, is the same as any other node to the right end. In other words, from any node, there is an infinite tree growing to the right.

As such, the transfer function from the first node to the end, is equal to the transfer function from one of the nodes in the second generation to the end, as is illustrated in Figure 1.9.

Let $G_\infty(s)$ represent the infinite transfer function from any node to the end, and let the transfer function corresponding to the individual components be

$$G_1(s) = \frac{1}{k} \qquad G_2(s) = \frac{1}{bs}. \tag{1.30}$$

If there are an infinite number of generations then

$$G_\infty(s) = \frac{1}{\frac{1}{G_1(s)+G_\infty(s)} + \frac{1}{G_2(s)+G_\infty(s)}},$$

and solving this for $G_\infty(s)$ gives

$$G_\infty(s) = \sqrt{G_1(s)G_2(s)} = \sqrt{\frac{1}{kbs}} = \frac{1}{\sqrt{kb}}\frac{1}{\sqrt{s}}.$$

where the $G_\infty(s)$ on the left hand side is the transfer function for the entire network, and the two $G_\infty(s)$ terms in the denominator are the transfer functions

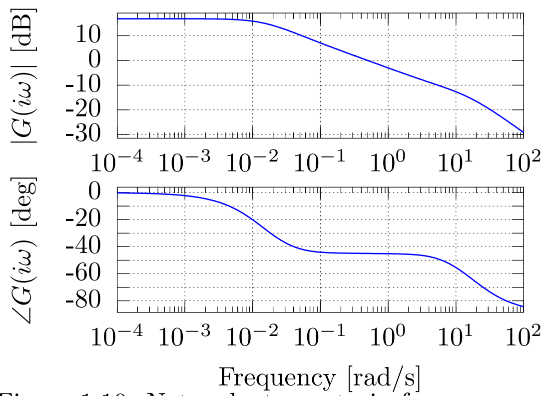

Figure 1.10: Network stress-strain frequency response.

from the two nodes in the second generation to the end. So this system should be characterized by half-order dynamics because $\sqrt{s} = s^{1/2}$.

Consistent with the idea above about fractional order systems "converging" to integer order ones, we plot the Bode plot for a network of this type where there are five generations with $k = 1$ and The Bode plot for this system is illustrated in Figure 1.10, which is characterized by two half-order dynamics features. First, the slope of the high frequency portion of the magnitude plot is -10dB/decade and the phase is -45°. Because first order terms are characterized by a slope of -20dB/decade and a phase of -90°, these features make sense as half order effects.

1.5.2 Ultrasound Example

Reference and summaries Holm's paper.

Chapter 2

Fractional Derivative Definitions

There are many definitions of fractional derivatives. In this chapter, we present a few of them along with their properties and compare and contrast them.

2.1 Summary of Important Functions

In engineering there is a relatively limited collection of functions that are so useful that their properties become second nature. Fractional calculus adds to that collection, and this section presents some of them along with some of their most important properties. First we consider some basic computations.

2.1.1 Preliminaries

A few computations are needed before we study the functions.

Gaussian Integral

The Gaussian integral is

$$\boxed{\int_{-\infty}^{\infty} e^{-z^2}\,\mathrm{d}z = \sqrt{\pi},} \tag{2.1}$$

and is the area under the curve in Figure 2.1. Clearly, also

$$\int_{0}^{\infty} e^{-z^2}\,\mathrm{d}z = \frac{\sqrt{\pi}}{2}.$$

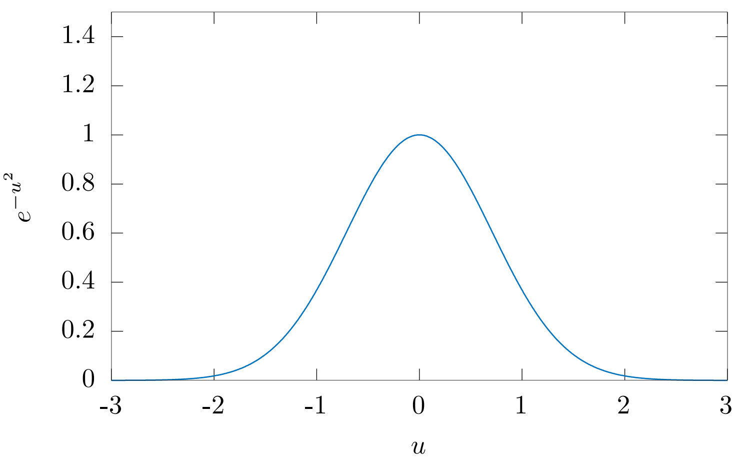

Figure 2.1: Gaussian integral.

To see this, consider the square of the integral and switch to polar coordinates

$$\begin{aligned}
\left(\int_{-\infty}^{\infty} e^{-z^2}\,\mathrm{d}z\right)^2 &= \left(\int_{-\infty}^{\infty} e^{-z_1^2}\,\mathrm{d}z_1\right)\left(\int_{-\infty}^{\infty} e^{-z_2^2}\,\mathrm{d}z_2\right)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-z_1^2-z_2^2}\,\mathrm{d}z_1\,\mathrm{d}z_2\\
&= \int_0^{2\pi}\int_0^{\infty} re^{-r^2}\,\mathrm{d}r\,\mathrm{d}\theta\\
&= 2\pi\int_0^{\infty} re^{-r^2}\,\mathrm{d}r\\
&= \pi\int_{-\infty}^{0} e^{u}\,\mathrm{d}u \qquad (u=-r^2)\\
&= \pi,
\end{aligned}$$

which shows

$$\int_{-\infty}^{\infty} e^{-z^2}\,\mathrm{d}z = \sqrt{\pi}.$$

Notation

We will need notation for the "floor" and "ceiling" operations. For $\alpha \in \mathbb{R}$, $\lceil\alpha\rceil$ represents the smallest integer greater than α. For example, $\lceil 3.05\rceil = 4$. Correspondingly, $\lfloor\alpha\rfloor$ is the largest integer less than α, so $\lfloor 4.99\rfloor = 4$.

Because there will be multiple definitions of a fractional derivative, we need a way to distinguish them. In general, D^α will stand for the derivative operator of order α, and $D^{-\alpha}$ will represent integrating by a fractional number of times.

As will be seen, because most definitions will involve integrals, the limits of integration will also impact the operation, so we need to have those in the definition, so ${}_tD_0^\alpha$ will be the derivative operator if α is positive and the 0 and t will be the limits of integration (described subsequently).

2.1.2 The Gamma Function

The gamma function will appear just about everywhere where we deal with fractional derivatives. We have already seen an example. The integral representation of the gamma function is

$$\boxed{\Gamma(t) = \int_0^\infty e^{-z} z^{t-1}\,\mathrm{d}z.} \tag{2.2}$$

In the case where t is an integer, they way to compute the integral by hand would be to do so repeated by parts to work the exponent of z in the integrand down to zero:

$$\begin{aligned}
\Gamma(t) &= \int_0^\infty e^{-z} z^{t-1}\,\mathrm{d}z \\
&= \left[z^{t-1}\left(-e^{-z}\right)\right]_0^\infty + (t-1)\int_0^\infty e^{-z} z^{t-2}\,\mathrm{d}z \\
&= 0 - 0 + (t-1)\,\Gamma(t-1).
\end{aligned}$$

Comparing the last line to the right hand side of the line above it gives a recursion relation analogous to $n\,(n-1)! = n!$,

$$\boxed{\Gamma(t) = (t-1)\,\Gamma(t-1).} \tag{2.3}$$

Also, continuing to integrate by parts and knowing that the boundary terms will always continue to be zero, we have

$$\begin{aligned}
\Gamma(t) &= (t-1)\,\Gamma\,(t-1) \\
&= \left[(t-1)\,(t-2)\right]\Gamma\,(t-2) \\
&\;\vdots \\
&= \left[(t-1)\,(t-2)\cdots 1\right]\Gamma(1) \\
&= (t-1)!,
\end{aligned}$$

which proves that $\boxed{\Gamma(t) = (t-1)!, t \in \mathbb{Z}}$, where $\mathbb{Z}$ is the set of natural numbers.

While the gamma function provides a nice generalization of the factorial for positive t, it is singular at zero and negative integer values as is illustrated in Figure 2.2.[1] This is a feature we will have to expect to see in fractional

[1]In fact, this integral definition we used in Equation 1.12 is only valid for positive arguments. Other definitions, such as a series one, has to be used for zero and negative values for the definition to be complete and rigorous.

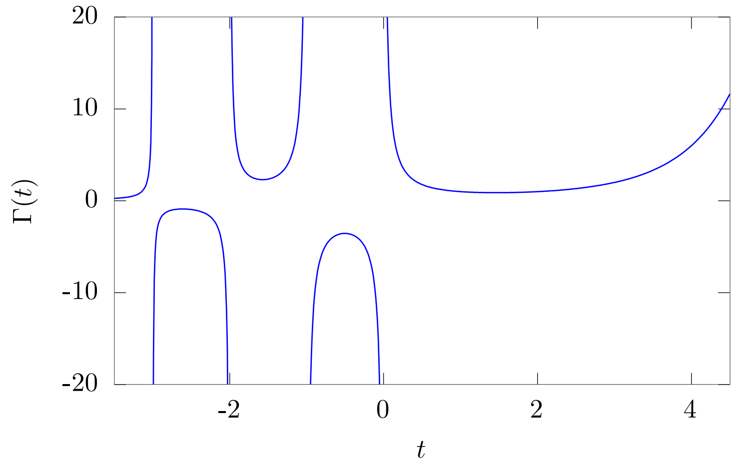

Figure 2.2: Gamma function for positive and negative real t values.

derivatives that use the gamma function. Singularities are usually considered "bad things" but they actually make some sense in this context as the following example illustrates.

Example 2.1. *Consider $f(t) = t$ and the fractional derivatives computed using Equation 1.14 that are illustrated in Figure 2.3. Note that the singularity of the gamma function at $t = 0$ can be seen as a way for the fractional derivatives between the zeroth and first derivatives to move between the two.*

The value of the gamma function at some special values should be cataloged.

- $\boxed{\Gamma(1) = 1.}$ This can be directly computed

$$\Gamma(1) = \int_0^\infty e^{-z} z^{1-1} \, \mathrm{d}z = -e^{-z}\big|_0^\infty = 1.$$

- $\boxed{\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.}$ This can also be directly computed using the Gaussian integral from Equation 2.1

$$\begin{aligned}
\Gamma\left(\frac{1}{2}\right) &= \int_0^\infty e^{-z} z^{\frac{1}{2}-1} \, \mathrm{d}z = \int_0^\infty e^{-z} z^{-\frac{1}{2}} \, \mathrm{d}z \\
&= 2\int_0^\infty e^{-u^2} \, \mathrm{d}u \qquad (u^2 = z) \\
&= \sqrt{\pi}.
\end{aligned}$$

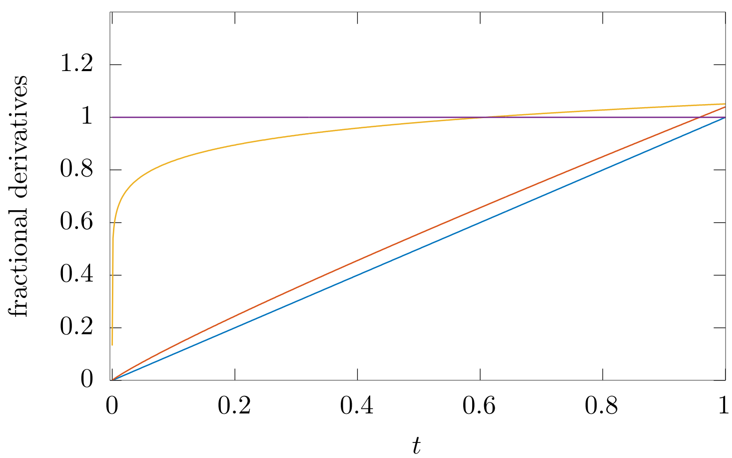

Figure 2.3: Plot of $f(t) = t$ and its 0.1, 0.9 and first derivative.

- $\boxed{\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}.}$ This can be computed using Equation 2.3

$$\Gamma\left(\frac{3}{2}\right) = \left(\frac{3}{2} - 1\right)\Gamma\left(\frac{3}{2} - 1\right) = \frac{1}{2}\sqrt{\pi}.$$

- $\boxed{\Gamma\left(\frac{5}{2}\right) = \frac{3\sqrt{\pi}}{4}.}$ Similar to the previous one

$$\Gamma\left(\frac{5}{2}\right) = \left(\frac{5}{2} - 1\right)\Gamma\left(\frac{5}{2} - 1\right) = \frac{3}{2}\frac{\sqrt{\pi}}{2} = \frac{3}{4}\sqrt{\pi}.$$

- $\boxed{\Gamma\left(-\frac{1}{2}\right) = -2\sqrt{\pi}.}$ Similarly, for some negative values, this follows from the recursion relation in Equation 2.3

$$\Gamma\left(\frac{1}{2}\right) = \left(\frac{1}{2} - 1\right)\Gamma\left(\frac{1}{2} - 1\right) \qquad \Longleftrightarrow \qquad \Gamma\left(-\frac{1}{2}\right) = \frac{\Gamma\left(\frac{1}{2}\right)}{-\frac{1}{2}} = -2\sqrt{\pi}.$$

The gamma function also appears in fractional-order inverse Laplace transforms. Similarly to what we did previously, we will start with integer-order computations and then generalize. Consider the usual Laplace transform of

$f(t) = t^n$ where $n \in \mathbb{Z}$

$$\begin{aligned}
\mathcal{L}\left\{t^n\right\} &= \int_{0^-}^{\infty} t^n e^{-st}\,\mathrm{d}t \\
&= \left(-t^n \frac{1}{s} e^{-st}\right)\Bigg|_0^{\infty} + \frac{n}{s}\int_{0^-}^{\infty} t^{n-1} e^{-st}\,\mathrm{d}t \\
&= \frac{n}{s}\int_{0^-}^{\infty} t^{n-1} e^{-st}\,\mathrm{d}t \\
&\vdots \\
&= \frac{n!}{s^{n+1}}\int_{0^-}^{\infty} e^{-st}\,\mathrm{d}t \\
&= \frac{n!}{s^{n+1}}.
\end{aligned}$$

A perfectly reasonable, albeit not mathematically rigorous, inference at this point would be

$$\boxed{\mathcal{L}\left\{t^\alpha\right\} = \frac{\Gamma(\alpha+1)}{s^{\alpha+1}}.}$$

In fact, we can compute it directly

$$\begin{aligned}
\mathcal{L}\left\{t^\alpha\right\} &= \int_{0^-}^{\infty} t^\alpha e^{-st}\,\mathrm{d}t \\
&= \int_{0^-}^{\infty} \left(\frac{u}{s}\right)^\alpha \frac{e^{-u}}{s}\,\mathrm{d}u \qquad (u = st) \\
&= \frac{1}{s^{\alpha+1}}\int_{0^-}^{\infty} e^{-u} u^\alpha\,\mathrm{d}u \\
&= \frac{\Gamma(\alpha+1)}{s^{\alpha+1}}, \quad \alpha > -1.
\end{aligned}$$

2.1.3 Binomial Coefficient

In the integer case, the binomial coefficient for $n, k \in \mathbb{Z}$ is

$$\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!\,(n-k)!}. \tag{2.4}$$

Common interpretations include entries in Pascal's triangle and "n choose k" because it is the number of different ways to choose k elements from a set with n elements. It will be convenient, but perhaps not so rigorous, to consider the binomial coefficient to be zero "outside" of Pascal's triangle. This makes sense for the probability: *e.g.*, there are zero ways to choose 5 elements from a set of 4.

We will use it because it naturally arises in finite difference expansions, and we will naturally want to allow the entries to be non-integers. So we will have

k	n=4	n=3.95	n=3.5
0	1	1	1
1	4	3.95	3.5
2	6	5.82625	4.375
3	4	3.78706	2.1875
4	1	0.899427	0.273438
5	0	-0.00899427	-0.0273438
6	0	0.001574	0.00683594
7	0	-0.000460957	-0.00244141
8	0	0.00017574	0.00106812
9	0	-0.00007908	-0.000534058
10	0	0.00003994	0.000293732

Table 2.1: Binomial coefficients with $n = 4$, $n = 3.95$ and $n = 3.5$ for various integer values of k

the generalization

$$\binom{\alpha}{\beta} = \frac{\Gamma(\alpha+1)}{\Gamma(\beta+1)\,\Gamma(\alpha-\beta+1)}, \tag{2.5}$$

where $\alpha, \beta \in \mathbb{R}$, *i.e.*, they can be fractional.

The following example illustrates the manner in which this will play a role in a subsequent fractional derivative definition.

Example 2.2. *Consider the binomial coefficient with $n = 4$ for various values of k, as shown in the first row of Table 2.1, where we have adopted the convention that the value is zero where the factorials are not defined. These values are also plotted in Figure 2.4.*

Note that for a small change in n from 4 to $3;95$ the values of the binomial coefficient are close to the values for 4. However, while the values for increasing β are tending towards zero in absolute value, they are not equal to zero and the rate of convergence is not all that fast.

2.1.4 The Error Function and Complementary Error Functions

These functions will be important as solutions to equations like

$$\frac{\mathrm{d}^{\frac{1}{2}}x}{\mathrm{d}t^{\frac{1}{2}}}(t) + x(t) = 1.$$

The error function is defined by

$$\boxed{\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}}\int_0^t e^{-z^2}\,\mathrm{d}z.} \tag{2.6}$$

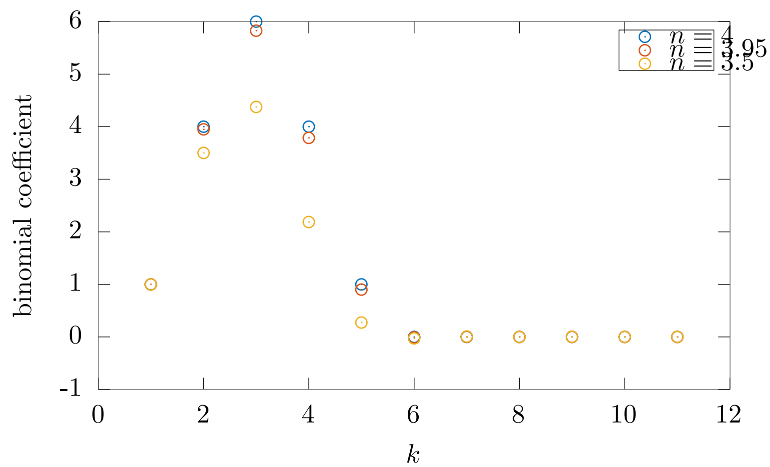

Figure 2.4: Binomial coefficient values for $n = 4, 3.95 and 3.5$ for various k values.

Note that it is like the Gaussian integral, but only over a subset of the range of the definite integral. The *complementary error function* is the integral over the remaining part of the domain

$$\boxed{\operatorname{erfc}(t) = \frac{2}{\sqrt{\pi}} \int_t^\infty e^{-z^2}\,\mathrm{d}z.} \tag{2.7}$$

Plots of both the error function and the complementary error function appear in Figure 2.5. It is clear from the definitions and the plots that

$$\operatorname{erfc}(t) = 1 - \operatorname{erf}(t). \tag{2.8}$$

Evaluating erf at specific values of t:

- $\boxed{\operatorname{erf}(0) = \frac{2}{\sqrt{\pi}} \int_0^0 e^{-u^2}\,\mathrm{d}u = 0.}$
- $\boxed{\operatorname{erf}(\infty) = \frac{2}{\sqrt{\pi}} \int_0^\infty e^{-u^2}\,\mathrm{d}u = 1.}$
- $\boxed{\operatorname{erf}(-\infty) = -1.}$

2.1.5 Mittag-Leffler Functions

Mittag-Leffler Functions are generalizations of the exponential function, and play a role in solutions to constant-coefficient, homogeneous linear fractional-order ordinary differential equations analogous to the exponential for integer

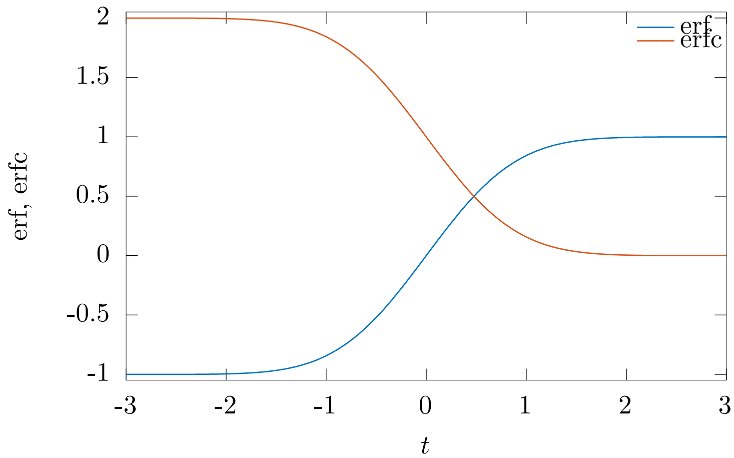

Figure 2.5: The error function and complementary error function.

order differential equations. As will be shown subsequently, just as $x(t) = ce^{-at}$ is the solution to

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) + ax(t) = 0$$

the function $ct^{\alpha-1}E_{\alpha,\alpha}\left(-at^{\alpha}\right)$, where $E_{\alpha,\alpha}$ is the Mittag-Leffler function to be defined shortly, is the solution to

$$\frac{\mathrm{d}^{\alpha}x}{\mathrm{d}t^{\alpha}}(t) + ax(t) = 0.$$

Recall the Taylor series of the exponential function about $t = 0$ is

$$e^t = 1 + t + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots = \sum_{k=0}^{\infty} \frac{t^k}{k!}.$$

These days we can not help but replace factorials with gamma functions. However, just doing that in the previous equation does not generalize anything because

$$\sum_{k=0}^{\infty} \frac{t^k}{\Gamma\left(k+1\right)} = e^t,$$

and nothing is really changed.

The *one parameter* and *two parameter* Mittag-Leffler functions put a coefficient in front of the k and 1 in the gamma function

$$\boxed{E_{\alpha}(t) = \sum_{k=1}^{\infty} \frac{t^k}{\Gamma\left(\alpha k+1\right)}, \quad \alpha > 0} \tag{2.9}$$

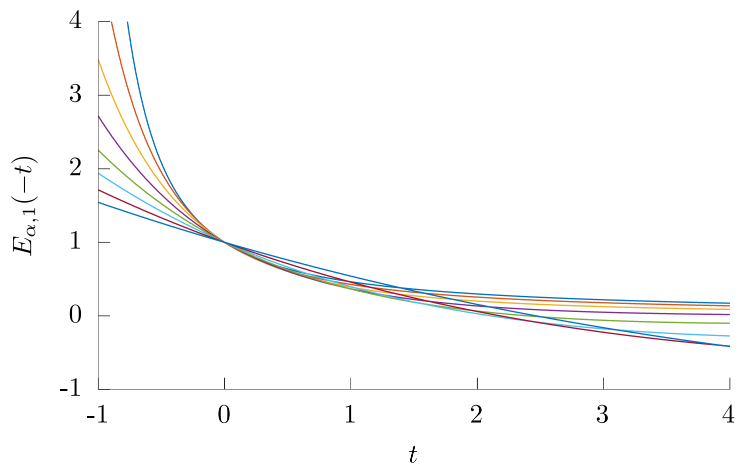

Figure 2.6: Mittag-Leffler functions, $E_{\alpha,1}(-t)$ for $\alpha = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$ and 2. Looking at the left part of the plot near $t = -1$, $\alpha = 0.25$ is the top curve, and they are in order down to $\alpha = 2$ for the bottom curve.

and

$$\boxed{E_{\alpha,\beta}(t) = \sum_{k=1}^{\infty} \frac{t^k}{\Gamma(\alpha k + \beta)}, \quad \alpha, \beta > 0.} \tag{2.10}$$

In order to gain some insight into these functions, let us see what the effect of varying the two parameters does. Figure 2.6 plots $E_{\alpha,t}(-t)$ for various values of α. Observe that for negative values, smaller α values are "stronger" whereas for positive values of t the opposite is basically the case. All of the curves go through the value of 1 at $t = 0$. The curves are more "curved" than the exponential for α values less than one, and less curved for α values greater than one.

Figure 2.7 illustrates $E_{1,\beta}(-t)$ for various β values. The trend to observe is that around time $t = 0$, the lowest curve corresponds to the smallest β values, and each subsequent curve increased from that one correspond to increasing β values.

There are certain combinations of α and β where $E_{\alpha,\beta}(t)$ is equal to a known function. Specifically

- $\boxed{E_{1,1}(t) = E_1(t) = e^t.}$
- $\boxed{E_{\frac{1}{2},1}(t) = E_{\frac{1}{2}}(t) = x^{t^2}\,\mathrm{erfc}(-t)}$
- $\boxed{E_{1,2}(t) = \frac{e^t - 1}{t}}$

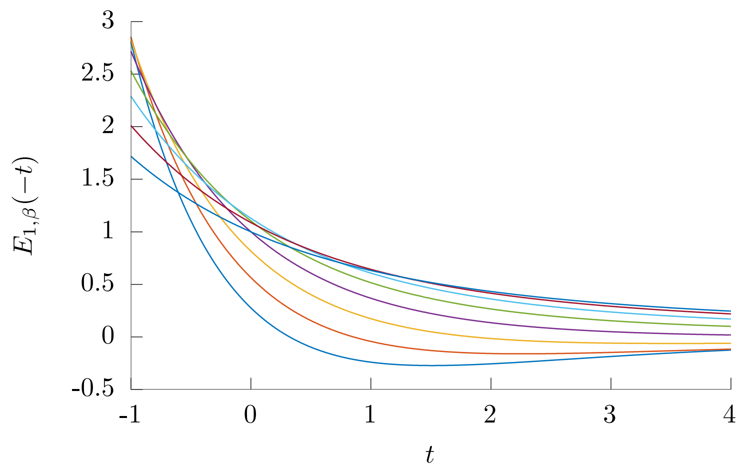

Figure 2.7: Mittag-Leffler functions, $E_{1,\beta}(-t)$ for $\beta = 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75$ and 2. Near $t = 0$ the lowest curve is for $\beta = 0.25$ and increased values in β correspond to the curves above that in order.

- $\boxed{E_{2,1}(t^2)\cosh(t)}$
- $\boxed{E_{2,2}(t^2) = \frac{\sinh(t)}{t}.}$

2.2 Fractional Integration and Fractional Derivatives

From last chapter, it was obvious that one element to generalize from integer-order derivatives to allow for fractional or real-ordered derivatives, was the gamma function in cases where the only barrier to allowing a derivative to take real values was a factorial. This section covers a couple other similar tools that we will need shortly.

2.2.1 Fractional Integration

It turns out we will more easily find a general formula for a fractional number of integrations, as opposed to differentiation. That is no problem, though, because, for example, if we want the 1/3 derivative, we can integrate a function 2/3 times and then compute the integer-order first derivative of the result, the law of indices (through the Fundamental Theorem of Calculus) gives that the result what we want.

The following theorem contains what is commonly called *Cauchy's formula for repeated integration.*

Theorem 2.1. *Let $f(t)$ be continuous. Then the nth repeated integral of $f(t)$ is given by*

$$\begin{aligned} f^{(-n)}(t) &= \int_a^t \int_a^{\sigma_1} \int_a^{\sigma_2} \int_a^{\sigma_3} \cdots \int_a^{\sigma_{n-1}} f(\sigma_n) d\sigma_n d\sigma_{n-1} \cdots d\sigma_1 \\ &= \frac{1}{(n-1)!} \int_a^t (x-z)^{n-1} f(z)\, \mathrm{d}z. \end{aligned} \tag{2.11}$$

Proof. This is fairly apparent. In the case where n is an integer, integrate the right hand side by parts $n-1$ times to obtain the left hand side. □

This theorem should make some intuitive sense. If you had to evaluate the single integral, the way to to it would be to integrate by parts n times to eliminate the $(x-z)$ term, which would give the multiple integral form of it.

If we ask how can we integrate a function a fractional number of times, though, it is similar to what was done in the first chapter. If we have

$$f^{(-n)}(t) = \frac{1}{(n-1)!} \int_a^t (t-z)^{n-1} f(z)\, \mathrm{d}z$$

the only term containing the order of integration, n where n can not be a fraction is, again, the factorial. So we can just replace it with a gamma function

$$\boxed{{}_a\mathrm{D}_t^{(-\alpha)}(t) = \frac{1}{\Gamma(\alpha)} \int_a^t (t-z)^{\alpha-1} f(z)\, \mathrm{d}z} \tag{2.12}$$

where the new notation for the operator D will be used going forward.

Because it is so common to have initial conditions specified at time $t = 0$, we will adopt the notation

$$\boxed{\mathrm{D}^{(-\alpha)}(t) = {}_0\mathrm{D}_t^{(-\alpha)}(t)}$$

i.e., we will not bother with adding to the notation when the limits of integration are from 0 to t.

Let us compute some fractional-order integrals of some common functions.

Example 2.3. *Consider $f(t) = t$. We know that $\mathrm{D}^{(-1)}\, t = 1/2t^2$. The half*

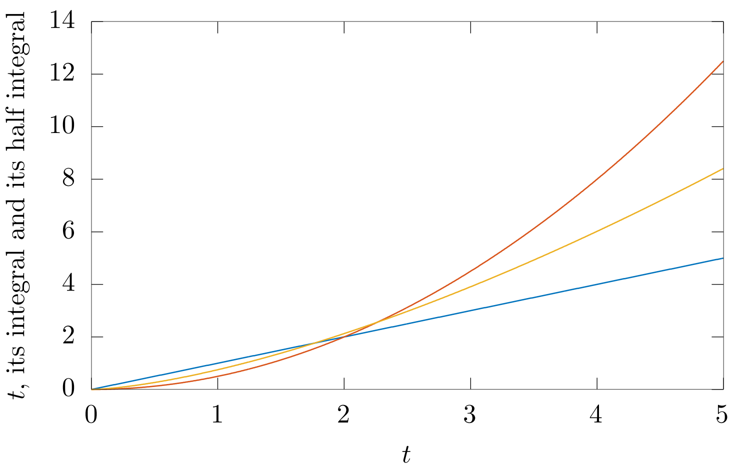

Figure 2.8: The function $f(t) = t$ (blue), ${}_0\mathrm{D}_t^{(-1)}\, t$ (red) and ${}_0\mathrm{D}_t^{(-1/2)}\, t$ (yellow).

integral should be something "in between" t and t^2. In detail

$$
\begin{aligned}
\mathrm{D}^{-\frac{1}{2}}\, t &= \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^t (t-z)^{(1/2-1)}\, z \,\mathrm{d}z \\
&= -\frac{1}{\sqrt{\pi}} \int_t^0 \frac{t-u}{\sqrt{u}} \,\mathrm{d}u \qquad (u = t-z) \\
&= \frac{1}{\sqrt{\pi}} \int_0^t \frac{t}{\sqrt{u}} - \sqrt{u} \,\mathrm{d}u \\
&= \frac{1}{\sqrt{\pi}} \left[2tu^{\frac{1}{2}} - \frac{2}{3} u^{\frac{3}{2}} \right]_0^t \\
&= \frac{4}{3\sqrt{\pi}} t^{\frac{3}{2}}.
\end{aligned}
$$

Figure 2.8 illustrates t, ${}_0\mathrm{D}_t^{(-1/2)}\, t$ and ${}_0\mathrm{D}_t^{(-1)}\, t$.

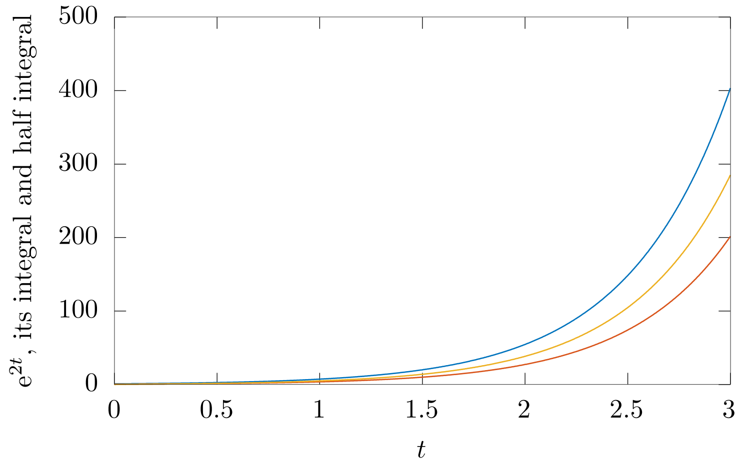

Figure 2.9: The function $f(t) = \mathrm{e}^{2t}$ and ${}_0\mathrm{D}_t^{(-1)}\,\mathrm{e}^{2t}$ (red) and ${}_0\mathrm{D}_t^{(-1/2)}\,\mathrm{e}^{2t}$ (yellow).

Example 2.4. *Consider $f(t) = \mathrm{e}^{2t}$. Then*

$$\begin{aligned}
\mathrm{D}^{(-1/2)}\,\mathrm{e}^{2t} &= \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^t (t-z)^{(-1/2)}\,\mathrm{e}^{2z}\,\mathrm{d}z \\
&= \frac{2}{\sqrt{\pi}} \int_0^{\sqrt{t}} \mathrm{e}^{2\left(t-u^2\right)}\,\mathrm{d}u \qquad \left(u = \sqrt{t-z}\right) \\
&= \mathrm{e}^{2t}\,\frac{1}{\sqrt{\pi}} \int_0^{\sqrt{t}} \mathrm{e}^{\left(\sqrt{2}u\right)^2}\,\mathrm{d}u \\
&= \mathrm{e}^{2t}\,\frac{\sqrt{2}}{\sqrt{\pi}} \int_0^{\sqrt{2t}} \mathrm{e}^{-v^2}\,\mathrm{d}v \qquad \left(v = \sqrt{2}u\right) \\
&= \frac{1}{\sqrt{2}}\,\mathrm{e}^{2t}\,\mathrm{erf}\,\sqrt{2t}
\end{aligned}$$

Figures 2.9 and 2.10 illustrate e^{2t}, ${}_0\mathrm{D}_t^{(-1/2)}\,\mathrm{e}^{2t}$ and ${}_0\mathrm{D}_t^{(-1)}\,\mathrm{e}^{2t}$. Note that because of the square root, the error function part of the solution is not defined for negative values of t.

Example 2.5. *Consider $f(t) = \cos 3t$. We have*

$$\mathrm{D}^{-1/2}\cos 3t = \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^t (t-z)^{(-1/2)}\cos(3z)\,\mathrm{d}z.$$

There is a closed-form solution to this integral in terms of the Fresnel integrals, but an easier check in this case is probably a numerical approach. Numerical

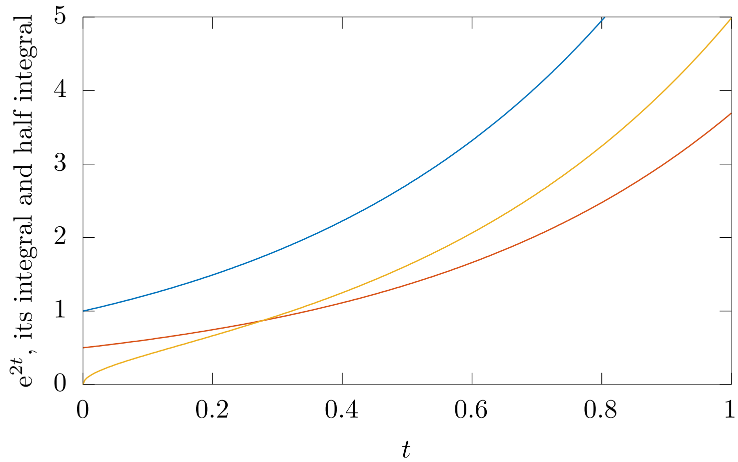

Figure 2.10: The function $f(t) = \mathrm{e}^{2t}$ and ${}_0\mathrm{D}_t^{(-1)}\,\mathrm{e}^{2t}$ (red) and ${}_0\mathrm{D}_t^{(-1/2)}\,\mathrm{e}^{2t}$ (yellow).

methods are considered in greater detail subsequently. For now, we can simply evaluate the integral using the `integrate()` *function in octave or Matlab. The plot of this integral is illustrated in Figure 2.11 and displays the characteristics one would expect of the half integral of the cosine function, including both the magnitude and the phase shift.*

Note that the initial part of the solution is not exactly a shifted cosine function, however. In fact, the fractional derivative is equal to zero at zero whereas a simple phase shift would result in a non-zero value. Specifically, if the fractional integral only shifted the function in phase and scaled the magnitude in a manner expected by the coefficient of t in the argument of the cosine function, we would have expected the value at zero to be equal to $2/3 * cos(\pi/4) = \sqrt{2}/3 \approx 0.4714$.

The octave code that generated this figure is:

```
t = linspace(0,10,1000);
soln = zeros(1,length(t));
f = @(z,t) ((t - z).^(-1/2)).*cos(3*z);
for i=1:length(t)
  soln(i) = 1/gamma(1/2)*integral(@(z) f(z,t(i)),0,t(i));
end
plot(t,cos(3*t),'linewidth',2);
hold on;
plot(t,1/3*sin(3*t),'linewidth',2);
plot(t,soln,'linewidth',2);
```

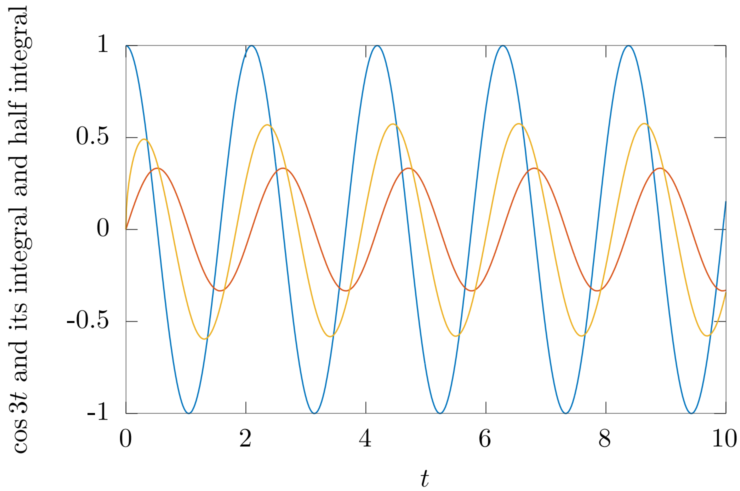

Figure 2.11: The function $f(t) = \cos 3t$ and ${}_0\mathrm{D}_t^{(-1)} \cos 3t$ (red) and ${}_0\mathrm{D}_t^{(-1/2)} \cos 3t$ (yellow).

2.2.2 Fractional Derivative Definitions

We will consider two basic approaches to generalizing the derivative to fractional orders. The first will use the fractional integration idea above along with the fundamental theorem of calculus. The second is an extension of the usual limit definition of the derivative that has a nice extension to the finite difference method in numerical methods.

At this point we have the ability to integrate a function by a fractional amount, *e.g.*, $\mathrm{D}^{(-1/2)} f(t)$ is the one-half integral of f. The basic idea is that we can use the fact that integrals and derivatives are inverse operations and, for example, integrate a function by, say, $2/3$ and then differentiate once to get the $1/3$ derivative. The difference between the first two definitions we will consider is simply whether we differentiate first and then fractionally integrate, or vice versa. That seemingly small difference actually has large consequences.

Riemann-Liouville Fractional Derivative

For the Riemann-Liouville derivative definition, we integrate first and then differentiate, specifically

$$\begin{aligned} {}^{RL}_{\;\;0}\mathrm{D}_t^{\alpha} f(t) &= {}^{RL}\mathrm{D}^{\alpha} f(t) \\ &= \frac{\mathrm{d}^{\lceil \alpha \rceil}}{\mathrm{d}t^{\lceil \alpha \rceil}} \mathrm{D}^{-(\lceil \alpha \rceil - \alpha)} f(t) \\ &= \frac{\mathrm{d}^{\lceil \alpha \rceil}}{\mathrm{d}t^{\lceil \alpha \rceil}} \left(\frac{1}{\Gamma\left(\lceil \alpha \rceil - \alpha\right)} \int_0^t (t-z)^{\lceil \alpha \rceil - \alpha - 1} f(z) \,\mathrm{d}z \right). \end{aligned} \tag{2.13}$$

Remark 2.1. *Note that the nonlocal nature of a fractional derivative is apparent* *from this definition. Information from the function evaluated over the entire range of the integral will effect the fractional derivative.*

Example 2.6. *Consider $f(t) = t$ and assume we want to compute ${}^{RL}\mathrm{D}^{1/2}\, t$. So for this problem, $\alpha = 1/2$ and $\lceil\alpha\rceil = 1$. In Example 2.3 we computed $\mathrm{D}^{(-1/2)}\, t = 4/\left(3\sqrt{\pi}\right) t^{3/2}$. So we have from Equation 2.13*

$$\begin{aligned} {}^{RL}\mathrm{D}^{1/2}\, t &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\Gamma\left(1/2\right)}\,\mathrm{D}^{-(1-1/2)}\, t\right) \\ &= \frac{4}{3\sqrt{\pi}}\frac{\mathrm{d}}{\mathrm{d}t}t^{\frac{3}{2}} \\ &= \frac{2}{\sqrt{\pi}}t^{\frac{1}{2}}. \end{aligned}$$

Note that this is the same as the $1/2$ derivative that we could compute from Definition 1.1.

We will do a slightly different example, because its fractional derivative will be different from what we will get with the next definition.

Example 2.7. *Add one to the function from the previous example:*

$$\begin{aligned} {}^{RL}\mathrm{D}^{1/2}\,(t+1) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\Gamma\left(1/2\right)}\,\mathrm{D}^{-(1-1/2)}\,(t+1)\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{4}{3\sqrt{\pi}}t^{\frac{3}{2}} + \frac{2}{\sqrt{\pi}}t^{\frac{1}{2}}\right) \\ &= \frac{2}{\sqrt{\pi}}\sqrt{t} + \frac{1}{\sqrt{\pi}\sqrt{t}}. \end{aligned}$$

See Example 2.8 for the substitution to integrate the second term in the integral.

In integer-order calculus, adding a constant to a function has no effect on its derivative. However, for fractional derivatives, this is not the case. The second term came directly from the added constant in $f(t)$.

Caputo Fractional Derivative

The Caputo definition simply switches the order of fractional-order integration and integer-order differentiation

$$\begin{aligned} {}^{C}_{0}\mathrm{D}^{\alpha}_{t}\, f(t) &= {}^{C}\mathrm{D}^{\alpha}\, f(t) \\ &= \mathrm{D}^{-(\lceil\alpha\rceil-\alpha)}\left(\frac{\mathrm{d}^{\lceil\alpha\rceil}f}{\mathrm{d}t^{\lceil\alpha\rceil}}(t)\right) \\ &= \frac{1}{\Gamma\left(\lceil\alpha\rceil-\alpha\right)}\int_0^t (t-z)^{\lceil\alpha\rceil-\alpha-1}\,\frac{\mathrm{d}^{\lceil\alpha\rceil}f}{\mathrm{d}z^{\lceil\alpha\rceil}}(z)\,\mathrm{d}z. \end{aligned} \tag{2.14}$$

Example 2.8. *Let us repeat Example 2.6 using the Caputo definition:*

$$\begin{aligned}
{}^{C}\mathrm{D}^{1/2}\, t &= \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^t (t-z)^{-\frac{1}{2}} \frac{\mathrm{d}z}{\mathrm{d}z}\, \mathrm{d}z \\
&= \frac{1}{\sqrt{\pi}} \int_0^t \frac{1}{\sqrt{t-z}}\, \mathrm{d}z \\
&= \frac{1}{\sqrt{\pi}} \int_t^0 -\frac{1}{\sqrt{u}}\, \mathrm{d}u \qquad (u = t - z) \\
&= \frac{1}{\sqrt{\pi}} \int_0^t \frac{1}{\sqrt{u}}\, \mathrm{d}u \\
&= \frac{1}{\sqrt{\pi}} \left[2\sqrt{u}\right]_0^t \\
&= \frac{2}{\sqrt{\pi}} t^{\frac{1}{2}}.
\end{aligned}$$

This is the same as the Riemann-Liouville definition.

Example 2.9. *Let us repeat Example 2.7 using the Caputo definition:*

$$\begin{aligned}
{}^{C}\mathrm{D}^{1/2}\, t &= \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^t (t-z)^{-\frac{1}{2}} \frac{\mathrm{d}}{\mathrm{d}z} (z+1)\, \mathrm{d}z \\
&= \frac{1}{\sqrt{\pi}} \int_0^t \frac{1}{\sqrt{t-z}}\, \mathrm{d}z \\
&= \frac{2}{\sqrt{\pi}} t^{\frac{1}{2}}.
\end{aligned}$$

This is not *the same as the Riemann-Liouville definition! It is due to the fact that the constant term in the function was eliminated by differentiating first using the Caputo definition, but was retained when it was integrated first using the Riemann-Liouville definition. Figure 2.12 illustrates the difference between these two fractional derivatives.*

Grünwald-Letnikov Fractional Derivative

The third definition of a fractional derivative we will consider is appealing for two reasons. First, it is a limit definition, which corresponds in a sense to our usual consideration of an integer-order derivative. Second, because it is a limit as $\Delta t \to 0$, we can directly adopt it in numerical methods by eliminating the limit and simply taking a small Δt. To understand the basis for the definition, we will do what we have done a lot so far, which is look for a pattern.

Consider the usual definition of the first derivative of the function $f(t)$ at time t

$$\frac{\mathrm{d}f}{\mathrm{d}t}(t) = \lim_{\Delta t \to 0} \frac{f(t) - f(t - \Delta t)}{\Delta t}$$

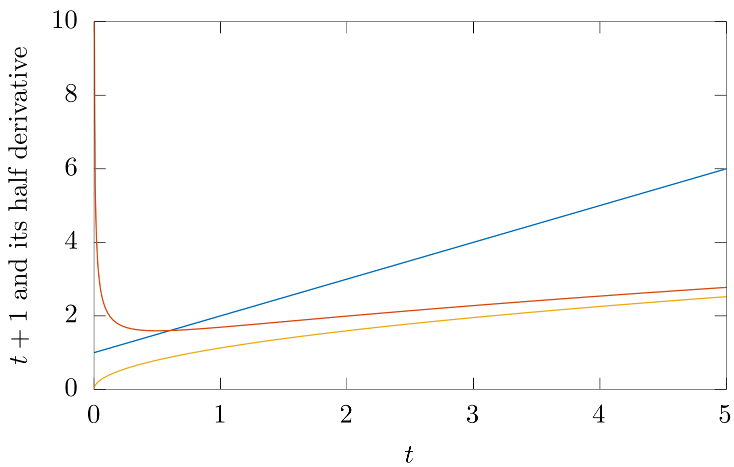

Figure 2.12: The function $f(t) = t+1$ (blue), the 1/2 derivative computed using the Riemann-Liouville definition (red) and the 1/2 derivative computed using the Caputo definition (yellow).

and the second derivative

$$\begin{aligned}\frac{\mathrm{d}^2 f}{\mathrm{d}t^2}(t) &= \lim_{\Delta t \to 0}\left(\frac{\frac{f(t)-f(t-\Delta t)}{\Delta t} - \frac{f(t-\Delta t)-f(t-2\Delta t)}{\Delta t}}{\Delta t}\right) \\ &= \lim_{\Delta t \to 0}\left(\frac{f(t) - 2f(t-\Delta t) + f(t-2\Delta t)}{(\Delta t)^2}\right).\end{aligned}$$

Continuing to compute higher-order derivatives makes it easy to see the pattern that gives

$$\frac{\mathrm{d}^n f}{\mathrm{d}t^n}(t) = \lim_{\Delta t \to 0} \frac{\sum_{k=0}^{n} (-1)^k \binom{n}{k} f(t - k\Delta t)}{(\Delta t)^n}$$

where $t = n\Delta t$. We have the usual business of factorials inside the binomial coefficient, which we have already generalized in Equation 2.5, which gives us the following definition.

Definition 2.1. *The Grünwald-Letnikov fractional derivative is given by*

$$\frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t) = \lim_{\Delta t \to 0} \frac{\sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t - k\Delta t)}{(\Delta t)^\alpha}. \tag{2.15}$$

Note that the sum goes back over all values of $f(t)$ from t to $t = -\infty$, i.e., all of the history of $f(t)$ contributes to the definition.

If all values of $f(t)$ are zero for $t < 0$, then we can write

$$\boxed{\frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t) = \lim_{\Delta t \to 0} \frac{\sum_{k=0}^{\left\lfloor \frac{t}{\Delta t} \right\rfloor} (-1)^k \binom{\alpha}{k} f(t - k\Delta t)}{(\Delta t)^\alpha}} \tag{2.16}$$

so the sum only goes back over values of $f(t)$ to zero. In the limit the sum will still contain an infinite number of terms, but those infinite number of evaluations of $f(t)$ will only be between the current t and 0.

Note that the nonlocal nature of the fractional-derivative is also present in this definition because the sum incorporates values of the function back to $t = 0$, regardless of the order. In cases where the order is an integer order, then many of the binomial coefficients will be zero, making the definition local. In the limit, fractional derivatives will contain an infinite number of terms in the summation, much like the integrals do for the Riemann-Liouville and Caputo definitions.

We will use this definition frequently because taking a small Δt should give us decent numerical approximations to fractional derivatives, and, as we will see shorty, allow us to compute numerical solutions to fractional-order differential equations.

It may seem like this definition is the best because it will alleviate the need to compute a lot of integrals that may not have solutions in terms of simple functions. However, there is a hidden complication. Note that k, the index of the summation, will become increasingly large as t gets large, and will generally be large if Δt is small, which we want for good numerical accuracy. However, referring back to the generalization of the binomial coefficient given by Equation 2.5, the denominator will contain two gamma functions with possibly large arguments. Recall that the gamma function can be thought of as the generalization of the factorial, so this will grow very large very quickly. Numerical issues arise surprisingly quickly.

Unless we specify to the contrary, in this course we will assume everything has a value of zero prior to $t = 0$ and hence use the definition in Equation 2.16.

Example 2.10. *Use the Grünwald-Letnikov definition to compute a numerical approximation for the $1/2$ derivative of*

$$f(t) = \begin{cases} 0, & t < 0 \\ \cos 3t, & t \geq 0 \end{cases} \tag{2.17}$$

for $t \in (0, 10]$.

The period of oscillation of $f(t)$ is approximately 2, so if we take $\Delta t = 1/100$, each period should contain approximately 200 data points, which, as a first guess, should give a reasonable approximation. So we have

$$\frac{\mathrm{d}^{\frac{1}{2}} f}{\mathrm{d}t^{\frac{1}{2}}}(t) \approx \sum_{k=0}^{\lfloor 100t \rfloor} \frac{(-1)^k \binom{\frac{1}{2}}{k} \cos 3\left(t - \frac{k}{100}\right)}{\frac{1}{100}}$$

The octave code that produced this result is

```
t = linspace(0,10,1001);
dt = t(2)-t(1);
alpha = 1.1;
deriv = 0;
f = cos(3*t);
coefs = 0;
coefs(1) = bincoeff(alpha,0);
deriv(1) = 0;
for n = 2:length(t)
   coefs(n) = (-1)^(n-1)*bincoeff(alpha,(n-1));
   sum = dot(fliplr(f(1:n)),coefs)/dt^alpha;
   deriv(n) = sum;
end
plot(t,f,'linewidth',2);
hold on;
plot(t,-3*sin(3*t),'linewidth',2);
plot(t,deriv,'linewidth',2);
xlabel('$t$');
ylabel('half derivative of $\cos 3 t$');
```

2.3 Operational Calculus

This section deals with operational calculus, specifically oriented towards fractional cases, which also includes the more common notions of Laplace transforms.

2.3.1 History and Basic Ideas

2.3.2 Rigorous Operational Calculus

2.3.3 Laplace Transforms

In many ways, Laplace transforms will be one of the more powerful tools. First we will consider the Laplace transform of some of the functions described in the previous sections, and it will be apparent that they are related to some fractional powers of s. Then we will consider fractional differentiation and integration in the Laplace transform context.

First, though, recall the definition of the Laplace transform

$$\boxed{F(s) = \mathcal{L}\left\{f(t)\right\} = \int_{0^-}^{\infty} f(t)\, \mathrm{e}^{-st}\, dt} \tag{2.18}$$

and the inverse

$$f(t) = \frac{1}{2\pi\imath} \lim_{b\to\infty} \int_{a-\imath b}^{a+\imath b} F(s) e^{st} ds. \tag{2.19}$$

2.3.4 Laplace Transform Pairs

A list of the Laplace transform of some functions appears in Table 2.2. We will work out some of the entries and leave the others as exercises.

First, consider $f(t) = t^n$ where n is a positive integer. Using the definition and integrating by parts gives

$$\begin{aligned}
\mathcal{L}\{t^n\} &= \int_0^\infty t^n \,\mathrm{e}^{-st}\, dt \\
&= \left[t^n \,\mathrm{e}^{-st}\right]_{0^-}^\infty + \frac{n}{s}\int_{0^-}^\infty t^{n-1}\,\mathrm{e}^{-st}\, dt \\
&= \frac{n}{s}\int_{0^-}^\infty t^{n-1}\,\mathrm{e}^{-st}\, dt \\
&= \frac{n}{s}\left(\left[t^{n-1}\,\mathrm{e}^{-st}\right]_{0^-}^\infty + \frac{n-1}{s}\int_{0^-}^\infty t^{n-1}\,\mathrm{e}^{-st}\, dt\right) \\
&= \vdots \\
&= \frac{n!}{s^{n+1}}.
\end{aligned}$$

From this, we have a pretty clear generalization for $f(t) = t^\alpha$ where α is not necessarily an integer, specifically

$$\mathcal{L}\{t^\alpha\} = \frac{\Gamma(\alpha+1)}{s^{\alpha+1}}.$$

This leads easily to a few special cases for specific values of α:

- $\mathcal{L}\left\{\frac{1}{\sqrt{t}}\right\} = \frac{\sqrt{\pi}}{\sqrt{s}}$

- $\mathcal{L}\left\{\sqrt{t}\right\} = \frac{\sqrt{\pi}}{2s^{\frac{3}{2}}}$

	$f(t)$	$F(s)$
1	1	$\frac{1}{s}$
2	$t^n, \quad n \in \mathbb{Z}$	$\frac{n!}{s^{n+1}}$
3	$t^\alpha, \quad \alpha \in \mathbb{R}$	$\frac{\Gamma(\alpha+1)}{s^{\alpha+1}}$
4	$\frac{1}{\sqrt{t}}$	$\sqrt{\frac{\pi}{s}}$
5	$\sqrt{t}$	$\frac{\sqrt{\pi}}{2s^{3/2}}$
6	$t^{\beta-1}\,\mathrm{E}_{\alpha,\beta}\left(\pm a t^\alpha\right)$	$\frac{s^{\alpha-\beta}}{s^\alpha \mp a}$
7	$\frac{1}{\sqrt{t}}\,\mathrm{E}_{\frac{1}{2},\frac{1}{2}}\left(\pm a\sqrt{t}\right)$	$\frac{1}{\sqrt{s}\mp a}$
8	$\mathrm{E}_\alpha\left(\pm a t^\alpha\right)$	$\frac{s^\alpha}{s(s^\alpha \mp a)}$

Table 2.2: Laplace Transform Pairs.

Let us consider the Laplace transform of Mittag-Leffler functions:

$$\begin{aligned}
\mathcal{L}\left\{t^{\beta-1}\,\mathrm{E}_{\alpha,\beta}\left(a t^\alpha\right)\right\} &= \int_{0^-}^{\infty} t^{\beta-1}\,\mathrm{E}_{\alpha,\beta}\left(a t^\alpha\right)\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \int_{o^-}^{\infty} t^{\beta-1}\sum_{k=0}^{\infty}\frac{\left(a t^\alpha\right)^k}{\Gamma\left(\alpha k+\beta\right)}\,\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \int_{o^-}^{\infty}\sum_{k=0}^{\infty}\frac{a^k t^{\alpha k+\beta-1}}{\Gamma\left(\alpha k+\beta\right)}\,\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \sum_{k=0}^{\infty}\frac{a^k}{\Gamma\left(\alpha k+\beta\right)}\int_{o^-}^{\infty} t^{\alpha k+\beta-1}\,\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \sum_{k=0}^{\infty}\frac{a^k}{\Gamma\left(\alpha k+\beta\right)}\frac{\Gamma\left(\alpha k+\beta\right)}{s^{\alpha k+\beta}} \\
&= \frac{1}{s^\beta}\sum_{k=0}^{\infty}\left(\frac{a}{s^\alpha}\right)^k \\
&= \frac{1}{s^\beta}\frac{1}{1-\frac{a}{s^\alpha}} \\
&= \frac{s^{\alpha-\beta}}{s^\alpha - a}.
\end{aligned}$$

Remark 2.2.

1. *Note that convergence properties are needed in order to switch the order of integration and the sum. While the conditions are met to make this valid, we have not shown them here.*

2. *Recall the basic geometric series property for the last step*

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$$

for $|r| < 1$.

A few special cases that may arise frequently are also in the table.

2.3.5 Integration and Differentiation in the Frequency Domain

Now we will consider calculus in the frequency domain. First, recall the basic property of Laplace transforms that multiplication in the time domain is equal to convolution in the frequency domain, specifically

$$\begin{aligned}
\mathcal{L}\left\{\int_0^t f_1(t-\tau)f_2(\tau)\,\mathrm{d}\tau\right\} &= \int_{0^-}^{\infty}\int_0^t f_1(t-\tau)f_2(\tau)\,\mathrm{d}\tau\,\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \int_{0^-}^{\infty}\int_0^{\infty} f_1(t-\tau)\mathbb{1}(t-\tau)f_2(\tau)\,\mathrm{d}\tau\,\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \int_0^{\infty}\int_{0^-}^{\infty} f_1(t-\tau)\mathbb{1}(t-\tau)f_2(\tau)\,\mathrm{e}^{-st}\,\mathrm{d}t\,\mathrm{d}\tau \\
&= \int_0^{\infty}\left(\int_{0^-}^{\infty} f_1(t-\tau)\mathbb{1}(t-\tau)\,\mathrm{e}^{-st}\,\mathrm{d}t\right)f_2(\tau)\,\mathrm{d}\tau \\
&= \int_0^{\infty}\left(\int_{-\tau}^{\infty} f_1(\lambda)\mathbb{1}(\lambda)\,\mathrm{e}^{-s(\lambda+\tau)}\,\mathrm{d}\lambda\right)f_2(\tau)\,\mathrm{d}\tau \\
&= \int_0^{\infty}\left(\int_0^{\infty} f_1(\lambda)\,\mathrm{e}^{-s\lambda}\,\mathrm{d}\lambda\right)\mathrm{e}^{-s\tau}\,f_2(\tau)\,\mathrm{d}\tau \\
&= \left(\int_0^{\infty} f_1(\lambda)\,\mathrm{e}^{-s\lambda}\,\mathrm{d}\lambda\right)\left(\int_0^{\infty}\mathrm{e}^{-s\tau}\,f_2(\tau)\,\mathrm{d}\tau\right) \\
&= F_1(s)F_2(s).
\end{aligned}$$

This is a fundamental fact important outside the fractional context. We will use it to derive the formulas we want; namely, that division by s^α corresponds to the fractional-order integral operation and correspondingly for multiplication and differentiation.

Consider the Laplace transform of the fractional-order integral given in Equation 2.12

$$\begin{aligned}
\mathcal{L}\left\{{}_0\mathrm{D}_t^{-\alpha}f(t)\right\} &= \int_0^{\infty}\frac{1}{\Gamma(\alpha)}\int_0^t (t-z)^{\alpha-1}f(z)\,\mathrm{d}z\,\mathrm{e}^{-st}\,\mathrm{d}t \\
&= \frac{1}{\Gamma(\alpha)}\mathcal{L}\left\{t^{\alpha-1}\right\}\mathcal{L}\left\{f(t)\right\} \\
&= \frac{1}{\Gamma(\alpha)}\left(\frac{\Gamma(\alpha-1+1)}{s^{\alpha-1+1}}\right)(F(s)) \\
&= \frac{1}{s^\alpha}F(s)
\end{aligned}$$

which is exactly what we would expect. Fractional integration of order α in the frequency domain is division by s^α. It is important enough to put it in a box and give it a number:

$$\boxed{\mathcal{L}\left\{\mathrm{D}^{(-\alpha)}\, f(t)\right\} = \frac{1}{s^{\alpha}} F(s), \qquad \alpha > 0.} \tag{2.20}$$

For fractional differentiation in the frequency domain, we will get a different answer for the Riemann-Liouville and Caputo definitions. For the Riemann-Liouville fractional derivative we have

$$\begin{aligned}
\mathcal{L}\left\{ {}^{RL}\mathrm{D}^{\alpha}\, f(t)\right\} &= \mathcal{L}\left\{\frac{\mathrm{d}^{\lceil\alpha\rceil}}{\mathrm{d}t^{\lceil\alpha\rceil}}\left(\mathrm{D}^{-(\lceil\alpha\rceil-\alpha)}\, f(t)\right)\right\} \\
&= s^{\lceil\alpha\rceil}\left(\frac{F(s)}{s^{\lceil\alpha\rceil-\alpha}}\right) - s^{\lceil\alpha\rceil-1}\left[\mathrm{D}^{-(\lceil\alpha\rceil-\alpha)}\, f(t)\right]_{t=0} \\
&\quad - s^{\lceil\alpha\rceil-2}\left[\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{D}^{-(\lceil\alpha\rceil-1)}\, f(t)\right)\right]_{t=0} \\
&\quad - \cdots - \left[\frac{\mathrm{d}^{\lceil\alpha\rceil-1}}{\mathrm{d}t^{\lceil\alpha\rceil-1}}\left(\mathrm{D}^{-(\lceil\alpha\rceil-1)}\, f(t)\right)\right]_{t=0}.
\end{aligned}$$

Note that the first term simplifies to $s^{\alpha}F(s)$ just like we want. However, the initial conditions involve fractional integrals of $f(t)$, which may (or may not!) be convenient for a given problem. In the usual situation in controls where we assume that all initial conditions, including fractional initial conditions, are zero, we have

$$\mathcal{L}\left\{ {}^{RL}\mathrm{D}^{\alpha}\, f(t)\right\} = s^{\alpha}F(s).$$

For the Caputo derivative, we have

$$\begin{aligned}
\mathcal{L}\left\{ {}^{C}\mathrm{D}^{\alpha}\, f(t)\right\} &= \mathcal{L}\left\{\mathrm{D}^{-(\lceil\alpha\rceil-\alpha)}\left(\frac{\mathrm{d}^{\lceil\alpha\rceil}f}{\mathrm{d}t^{\lceil\alpha\rceil}}(t)\right)\right\} \\
&= \mathcal{L}\left\{\frac{1}{\Gamma\left(\lceil\alpha\rceil-\alpha\right)}\int_0^t (t-z)^{\lceil\alpha\rceil-\alpha-1}\,\frac{\mathrm{d}^{\lceil\alpha\rceil}f}{\mathrm{d}z^{\lceil\alpha\rceil}}(z)\,\mathrm{d}z\right\} \\
&= \frac{1}{\Gamma\left(\lceil\alpha\rceil-\alpha\right)}\mathcal{L}\left\{(t-z)^{\lceil\alpha\rceil-\alpha-1}\right\}\mathcal{L}\left\{\frac{\mathrm{d}^{\lceil\alpha\rceil}f}{\mathrm{d}t^{\lceil\alpha\rceil}}(t)\right\} \\
&= \frac{1}{s^{\lceil\alpha\rceil-\alpha}}\left(s^{\lceil\alpha\rceil}F(s) - s^{\lceil\alpha\rceil-1}f(0) - s^{\lceil\alpha\rceil-2}\frac{\mathrm{d}f}{\mathrm{d}t}(0) - \right. \\
&\qquad \left. \cdots - \frac{\mathrm{d}^{\lceil\alpha\rceil-1}f}{\mathrm{d}t^{\lceil\alpha\rceil-1}}(0)\right)
\end{aligned}$$

Note that the Laplace transform of the Caputo derivative involves evaluating *integer-order* initial conditions, in contrast to the fractional-order initial conditions from the Riemann-Liouville derivative.

Chapter 3

Fractional-Order Differential Equations

Example 3.1. *Determine the solution to*

$$\frac{\mathrm{d}^{\frac{1}{2}}x}{\mathrm{d}t^{\frac{1}{2}}}(t) + ax(t) = f(t)$$

where we use the Riemann-Liouville fractional derivative, or

$$^{RL}\mathrm{D}^{\frac{1}{2}}\, x(t) + ax(t) = f(t). \tag{3.1}$$

Computing the Laplace transform of each side of the equation gives

$$s^{\frac{1}{2}}X(s) - \left[\mathrm{D}^{-\frac{1}{2}}\, x(t)\right]_{t=0} + aX(s) = F(s).$$

For the time being, let us assume that the initial condition term is not zero, and call it c. Solving for $X(s)$ gives

$$X(s) = \frac{c}{\sqrt{s}+a} + \frac{F(s)}{\sqrt{s}+a}.$$

If $f(t) = 0$ and $c = 1$, then

$$X(s) = \frac{1}{\sqrt{s}+a}$$

and the Laplace transform table gives that

$$x(t) = \frac{1}{\sqrt{t}}\,\mathrm{E}_{\frac{1}{2},\frac{1}{2}}\left(a\sqrt{t}\right).$$

The solution to this equation is illustrated in Figure 3.1. Note for increasing a the solution decays more rapidly. A graph of $x(t) = \mathrm{e}^{-2t}$ is also shown for comparison.

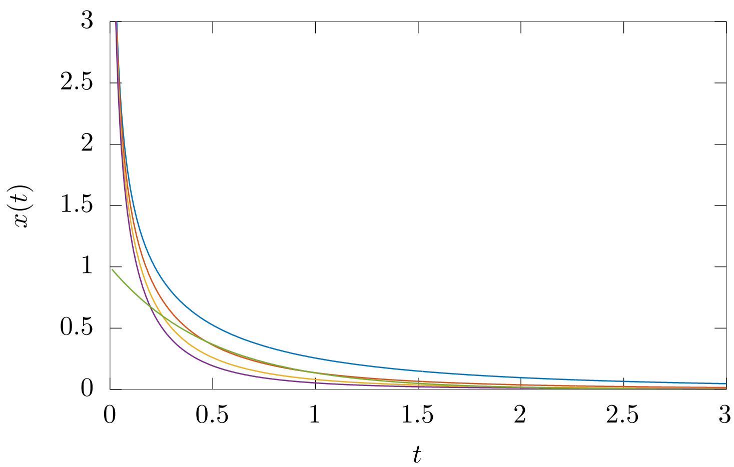

Figure 3.1: Solutions to Equation 3.1 for various $a = 1/2$ (blue), $a = 1$ (red), $a = 3.2$ gold and $a = 2$ (purple) and $f(t) = 0$ and $x(t) = e^{-2t}$ (green) for comparison.

Now, assume that $f(t) = 1$, so that $F(s) = 1/s$, in which case

$$X(s) = \frac{1}{\sqrt{s}+a} + \frac{1}{s\left(\sqrt{s}+a\right)}. \tag{3.2}$$

From Table 2.2, for the second term, we need that $\alpha = 1/2$ and in order to get the other s term in the denominator, we need $\beta - \alpha = 1$, so $\beta = 3/2$, which gives

$$x(t) = \frac{1}{\sqrt{t}} \, \mathrm{E}_{\frac{1}{2},\frac{1}{2}}\left(a\sqrt{t}\right) + \sqrt{t}\, \mathrm{E}_{\frac{1}{2},\frac{3}{2}}\left(a\sqrt{t}\right).$$

Figure 3.2 illustrates the solutions when $c = 0$, i.e., it is the"step response" portion of the solution. Figure 3.3 illustrates the full solution including the term with $c = 1$.

Example 3.2. *As a second example, consider a mass attached to a wall and subjected to a force as illustrated in Figure 3.4. We will consider the attachment to the wall as having some sort of mechanical impendence, which could be a spring, or a damper or perhaps a fractional-order type network such as the 1/2-order tree network of springs and dampers considered earlier in Section 1.5.1. The top part of the figure illustrates the half-order connection, and the bottom part represents the more general situation where the order of the relationship between the force resisting the motion and the displacement is of order γ.*

Newton's law on the mass gives

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}(t) + q\frac{\mathrm{d}^\gamma x}{\mathrm{d}t^\gamma}(t) = f(t).$$

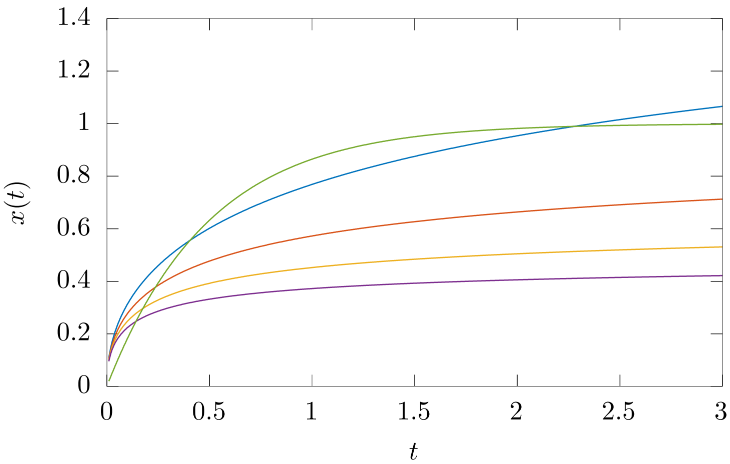

Figure 3.2: Solutions to Equation 3.1 for various $a = 1/2$ (blue), $a = 1$ (red), $a = 3.2$ gold and $a = 2$ (purple), $f(t) = 1$ and $c = 0$ and $x(t) = 1 - e^{-2t}$ (green) for comparison.

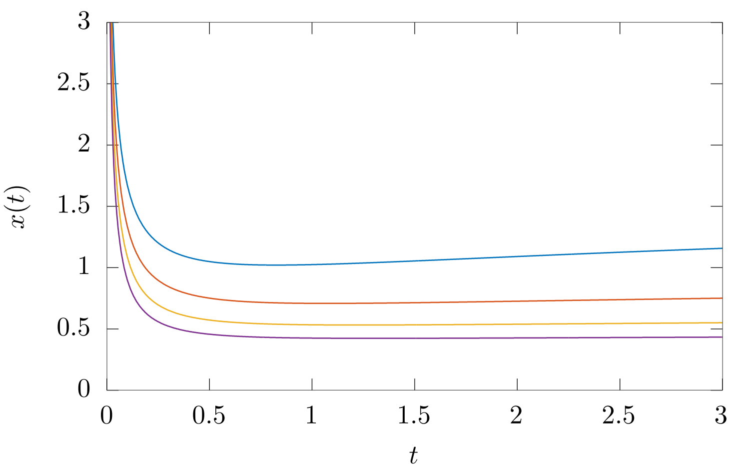

Figure 3.3: Solutions to Equation 3.1 for various $a = 1/2$ (blue), $a = 1$ (red), $a = 3.2$ gold and $a = 2$ (purple), $f(t) = 1$ and $c = 1$.

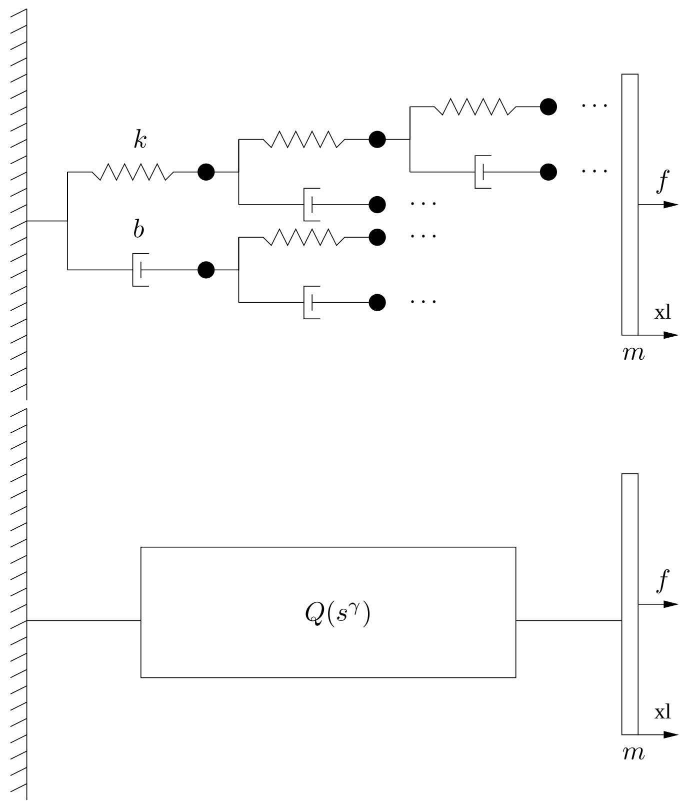

Figure 3.4: Mass attached to a wall.

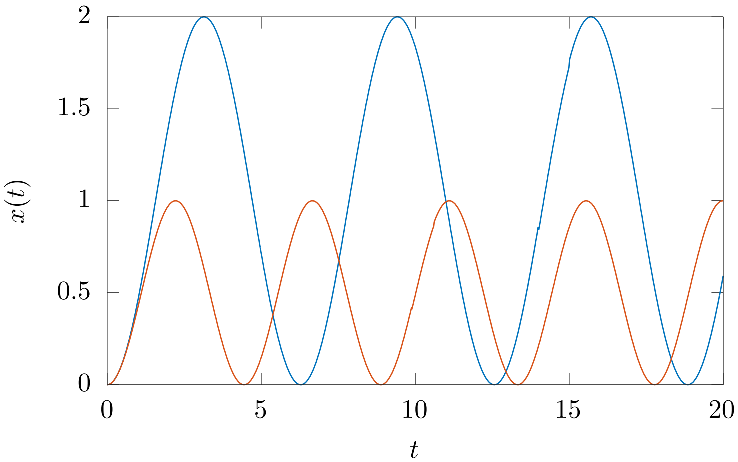

Figure 3.5: Solution to Equation 3.3 in the case where $\gamma = 0$, corresponding to a spring element connecting the mass to the wall. The blue curve is for $q/m = 1$ and the red curve is for $q/m = 2$.

In the case where the connection is a spring, we would use $q = k$ and $\gamma = 0$, and when it is a damper, we would use $q = b$ and $\gamma = 1$. Divide both sides by m, let $f(t)$ be a unit step input, assume zero initial conditions of all orders and take the Laplace transform of both sides, which gives

$$X(s) = \frac{1}{s\left(s^2 + \frac{q}{m}s^\gamma\right)} = \frac{1}{s^{\gamma+1}\left(s^{2-\gamma} + \frac{q}{m}\right)}. \tag{3.3}$$

Referring to Table 2.2, $\alpha = 2 - \gamma$ and $\alpha - \beta = -(\gamma + 1)$, so $\beta = 3$, and the time domain response is

$$x(t) = t^2 \, \mathrm{E}_{2-\gamma,3}\left(-\frac{q}{m}t^{2-\gamma}\right).$$

To validate this, let us plot the solution for $\gamma = 0$, which would correspond to the connection being a spring, in which case we expect a purly oscillatory response. Additionally, if we increase q/m we expect a higher frequency response. Both of these attributes of the solution are illustrated in Figure 3.5. In the case where $\gamma = 1$, the connection to the wall is a damper, in which case we would expect the steady-state solution to be a constant velocity where the damper force and applied force are in equilibrium. This case is illustrated in Figure 3.6.

Finally, the solutions for $\gamma \in \{0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75\}$ are illustrated in Figure 3.7. We will revisit this problem in the next chapter when we consider numerical methods.

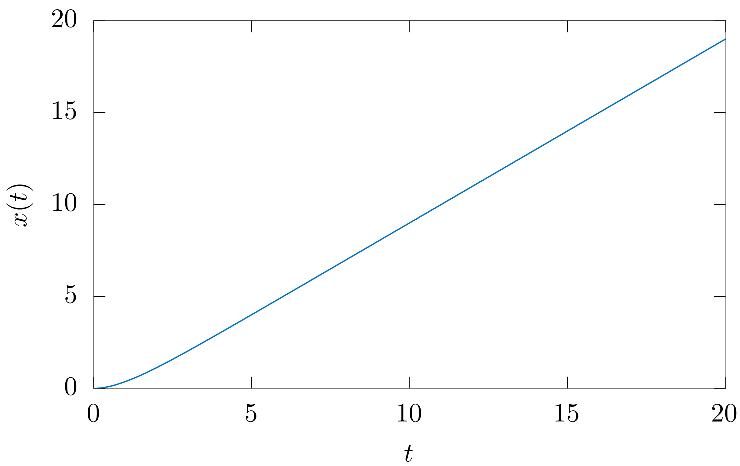

Figure 3.6: Solution to Equation 3.3 in the case where $\gamma = 1$, corresponding to a damper element connecting the mass to the wall.

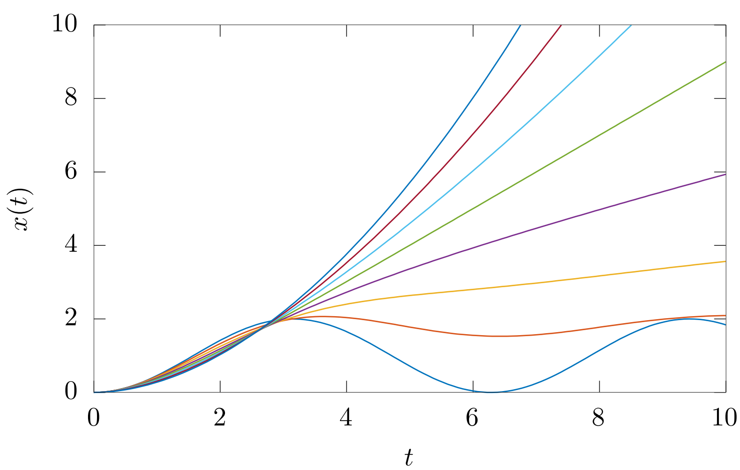

Figure 3.7: Solution to Equation 3.3 for $\gamma = 0, 0.25, 0.5, \ldots, 1.75$.

Chapter 4

Numerical Methods for Fractional-Order Differential Equations

Numerical methods are challenging for fractional-order systems because of the non-locality of the fractional derivative.

4.1 Direct Application of the Grünwald-Letnikov Fractional Derivative

Because the Grünwald-Letnikov definition of the fractional derivative contains a limit, we can use it by taking the limiting term to be small as a computational approximation of the fractional derivative. The approach itself is fairly straightforward, but numerical issues arise fairly quickly. We will illustrate both the approach and its limitations with an example.

Recall that the Grünwald-Letnikov definition of the fractional derivative is

$$\frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t) = \lim_{\Delta t \to 0} \frac{\sum_{k=0}^{\infty} (-1)^k \binom{\alpha}{k} f(t - k\Delta t)}{(\Delta t)^\alpha}$$

where the binomial coefficient is generalized to non-integer arguments as

$$\binom{\alpha}{k} = \frac{\Gamma(\alpha + 1)}{\Gamma(k+1)\,\Gamma(\alpha - k + 1)}.$$

In the case where all initial conditions of all orders are zero, the upper limit of the summation can be changed

$$\frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t) = \lim_{\Delta t \to 0} \frac{\sum_{k=0}^{\left\lfloor \frac{t}{\Delta t} \right\rfloor} (-1)^k \binom{\alpha}{k} f(t - k\Delta t)}{(\Delta t)^\alpha}.$$

This, of course, leads to the approximation

$$\frac{\mathrm{d}^\alpha f}{\mathrm{d}t^\alpha}(t) \approx \frac{\sum_{k=0}^{\left\lfloor \frac{t}{\Delta t} \right\rfloor} (-1)^k \binom{\alpha}{k} f(t - k\Delta t)}{(\Delta t)^\alpha}, \qquad \Delta t \ll 1.$$

We will return to an example from the Chapter 3 to implement this.

Example 4.1. *Solve*

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}(t) + q\frac{\mathrm{d}^\gamma x}{\mathrm{d}t^\gamma}(t) = f(t)$$

assuming all zero initial conditions. Because this is a numerical method, we need to choose numerical values for all the paramters, so let $m = q = 1$ *and* $\gamma = 1/2$ *and let* $f(t)$ *be a unit step input:*

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2}(t) + \frac{\mathrm{d}^{\frac{1}{2}} x}{\mathrm{d}t^{\frac{1}{2}}} = 1.$$

At this point we can take one of two approaches. The first is to convert this second-order differential equation into two first-order equations and use Euler's method to compute an approximate numerical solution. So, let $x_1 = x$ *and* $x_2 = \dot{x}$ *so that*

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ 1 - \frac{\mathrm{d}^{\frac{1}{2}} x_1}{\mathrm{d}t^{\frac{1}{2}}} \end{bmatrix}$$

which gives the approximation using Euler's method

$$\begin{bmatrix} x_1(t + \Delta t) \\ x_2(t + \Delta t) \end{bmatrix} = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} x_2(t) \\ 1 - \sum_{k=0}^{\left\lfloor \frac{t}{\Delta t} \right\rfloor} \frac{(-1)^k \binom{\gamma}{k} x_1(t - k\Delta t)}{(\Delta t)^\gamma} \end{bmatrix} \Delta t.$$

The second, which we will adopt, is to express the second derivative using finite differences, i.e., basically use the Grünwald-Letnikov definition for the second derivative term,

$$\frac{x(t) - 2x(t - \Delta t) + x(t - 2\Delta t)}{(\Delta t)^2} + \sum_{k=0}^{\left\lfloor \frac{t}{\Delta t} \right\rfloor} \frac{(-1)^\gamma \binom{\gamma}{k} x(t - k\Delta t)}{(\Delta t)^\gamma} = 1$$

and solve the resulting expression for $x(t)$

$$x(t) = \frac{(\Delta t)^{2+\gamma}}{(\Delta t)^2 + (\Delta t)^\gamma}\left[1 - \sum_{k=1}^{\left\lfloor \frac{t}{\Delta t} \right\rfloor} \left(\frac{(-1)^k \binom{\gamma}{k} x(t - k\Delta t)}{(\Delta t)^\gamma}\right) + \frac{2x(t - \Delta t) - x(t - 2\Delta t)}{(\Delta t)^2}\right]$$

Chapter 5

Fractional Control

Fraction-order control, especially fractional-order PID control, is becoming increasingly popular. There are two main topics we will consider: using a controller that has fractional-order derivatives in it and controlling fractional-order systsems. Fractional-order PID is fairly self-evident. Both the I and D terms in the controller may have non-integer order, such that the control input to a system, $u(t)$ is given by

$$u(t) = k_p e(t) + k_I \, {}_0\mathrm{D}_t^{-\lambda} \, e(t) + k_I \, {}_0\mathrm{D}_t^{\mu} \, e(t)$$

where $e(t)$ is the error signal. If λ and μ are both one, then this is what we normally would call PID control and if they are not one, it is what we call fractional PID control.

It makes some sense that since for normal PID control we have the three gains to tune and for fractional PID we have five parameters, the three gains and the two fractional orders, that with more degrees of freedom we can often design a better controller. How to effectively tune five parameters is an issue. Furthermore, a drawback to fractional PID is that the controller needs to either compute or approximate the fractional terms, which can be expensive and therefore slow because they are non local. As such, there has been a lot of effort to determine good appriximations for the fractional controller terms that retain the benefit of fractional PID but are fast to compute.

Controls in general is primarily concerned with stability and performance. We will consider stability first.

5.1 Stability of Fractional-Order Transfer Functions

Many of the topics covered in undergraduate controls courses are ultimately considering stability. For an integer order transfer function

$$G(s) = \frac{s^m + a_1 s^{m-1} + \cdots + a_m}{b_0 s^n + b_1 s^{n-1} + \cdots + b_n}$$

solutions will have decaying exponential terms if and only if all the roots of the denominator have negative real part. Many of the usual tools in controls are essentially directed towards that.

- Constructing and analyzing the *Routh array* is a means to determine the number of right half plane poles of the transfer function (but it does not fully factor the characteristic equation).
- The *root locus* method is a way to plot how the pole locations of a transfer function change as some parameter in the system changes. It is especially useful when the transfer function is made up of interactions among relatively simple, *i.e.*, factorable, sub-transfer functions with some feedback. Modern computational tools make it not particularly valuable as a means to simply determine the pole locations of a transfer function, *but knowing how poles and zeros affect pole locations in a system provide valuable insighe into design controllers.*
- *Bode plots* are a frequency-based tool that can provide insight into stability for a more limited class of systems (non-minimum phase), but are especially useful when a low-order model of a system is difficult to determine from first principles but the input-output behavior is relatively easy to determine experimentally.
- *Nyquist plots* are also frequency-based with fewer restrictions than Bode plots.

An extremely important point that is always made, but often forgotten because all the work ever given in a controls course automatically satisfy it, is that the class of transfer functions for which such tools work are *rational and proper.* Rational means they are ratios of polynomials (autmatically *not* satisfied in the fractional case) and proper means that the order of the denominator is greater than the numerator.

In order to develop some convenient tools to study stability for fractional-order transfer functions, we first need to understand the asymptotic properties of the Mittag-Leffler functions, *i.e.*, we know that in the integer-order case that poles with positive real parts correspond to terms in the solution with exponentials that increase in time, so what is the generalization of that to $\mathrm{E}_{\alpha,\beta}$?

Theorem 5.1. *For* $0 < \alpha < 2$ *and all* β*,*

$$\lim_{t\to\infty} \mathrm{E}_{\alpha,\beta}\left(at\right) \approx \frac{1}{\alpha} t^{\frac{1-\beta}{\alpha}} \,\mathrm{e}^{(at)^{\frac{1}{\alpha}}} .$$

The proof is beyond the scope of this book, but see [Valério and de Costa, 2013, Gorenflo et al., 2014]. Here we can check this for a few different cases of α, β and a.

Remark 5.1. *For stability this theorem provides some easy boundaries:*

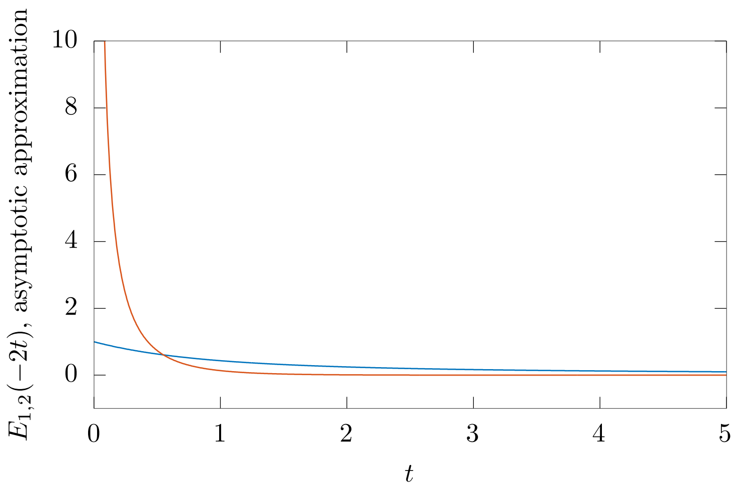

Figure 5.1: Mittag-Leffler function $\mathrm{E}_{1,2}(-2t)$ (blue) and asymptotic approximation (red).

- *If $a \in \mathbb{R}$, then we need $a < 0$ as usual.*
- *If $a \in \mathbb{C}$, then we need the real part of a to be negative as usual.*
- *The value of β does not matter.*
- *The theorem only applies for $0 < \alpha < 2$.*

Note that as $\alpha \to 2$, the region in the complex plan where the poles must be for stability shrinks to nothing. In the case where $a \in \mathbb{R}$, this makes sense because the square root of a negative number is purly imaginary and cube roots will have a positive real part. When $a > 0$, one of the roots will always be positive.

Example 5.1.

1. *Figure 5.1 plots $\mathrm{E}_{1,2}(-2t)$ and the corresponding asymptotic expression from Theorem 5.1.*
2. *Figure 5.2 plots $\mathrm{E}_{3/2,1/2}(-t)$ and the corresponding asymptotic expression from Theorem 5.1.*
3. *Figure 5.3 plots $\mathrm{E}_{1,2}(t)$ and the corresponding asymptotic expression from Theorem 5.1.*

Recall the main Laplace transform table entry for Mittag-Leffler functions:

$$\mathcal{L}\left\{t^{\beta-1}\,\mathrm{E}_{\alpha,\beta}\left(\pm a t^{\alpha}\right)\right\} = \frac{s^{\beta-\alpha}}{s^{\alpha} \mp a}.$$

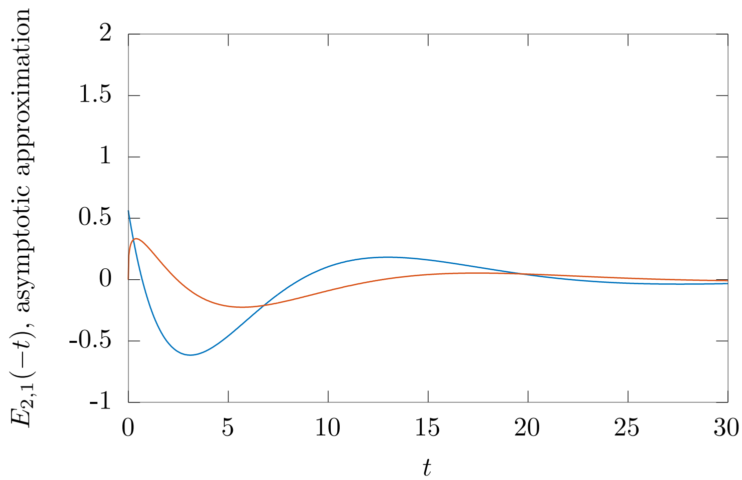

Figure 5.2: Mittag-Leffler function $\mathrm{E}_{3/2,1/2}\,(-t)$ (blue) and asymptotic approximation (red).

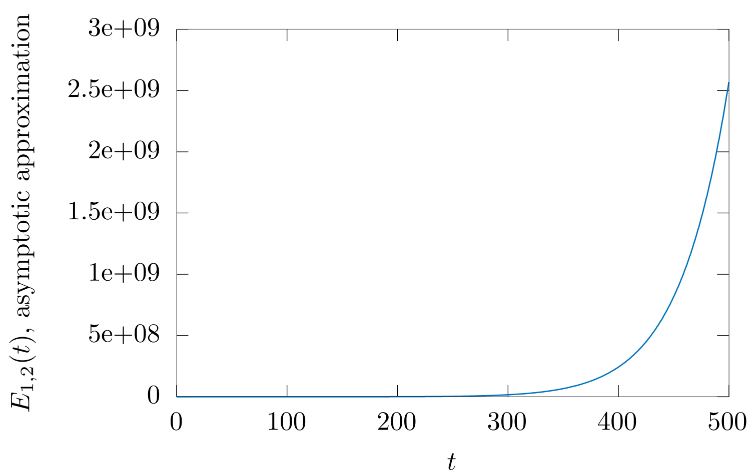

Figure 5.3: Mittag-Leffler function $\mathrm{E}_{1,2}\,(t)$ (blue) and asymptotic approximation (red).

The thing to note here is that the denominator term, $s^\alpha + a$ plays a very similar role as the integer-order poles we are used to considering. In other words, by inspection we can tell that a denominator $s - 3$ corresponds to an unstable component of the solution and $s + 2$ corresponds to a stable one. Theorem 5.1 essentially verifies the same thing for a denominator term like $s^{1/2} + 2$ in the fractional case as long as $0 < \alpha < 2$.

Definition 5.1. *A fractional transfer function is of the form*

$$G(s) = \frac{\sum_{k=1}^{m} b_k s^{\beta_k}}{\sum_{k=1}^{n} a_k s^{\alpha_k}}$$

where the $\alpha_k \geq 0$ are called the denominator orders and correspodingly the $\beta_k \geq 0$ are the numerator orders. Often we will normalize it by requiring one of the coefficients to be zero, say $a_1 = 0$.

5.1.1 Commensurable Transfer Functions

A particularly convenient situation arises in the case of *commensurable transfer functions.*

Definition 5.2. *A transfer function is said to be commensurable when all the orders α_k and β_k are integer multiples of a common divisor $\alpha > 0$. In that case the transfer function is rational in s^α. If we let $\lambda = s^\alpha$ then*

$$G(s) = \frac{\sum_{k=1}^{m} b_k s^{\beta_k}}{\sum_{k=1}^{n} a_k s^{\alpha_k}} = \frac{\sum_{k=0}^{m} b_k s^{\alpha k}}{\sum_{k=0}^{n} a_k s^{\alpha k}} = \frac{\sum_{k=0}^{m} b_k \lambda^k}{\sum_{k=0}^{n} a_k \lambda^k}.$$

Of course, if all the coefficients of the numerator and denominatore are integer multiples of α, they will also be integer multipes of $\alpha/2$, so it is convenient to chose the largest α.

Example 5.2. *The transfer function*

$$G(s) = \frac{s + 3s^{\frac{1}{2}} + 5}{s^2 + 2s^{\frac{3}{2}} + s^{\frac{1}{2}} + 2}$$

is commensurable with $a = 1/2$ because if we let $\lambda = 1/2$

$$G(s) = \frac{\lambda^2 + 3\lambda + 5}{\lambda^4 + 2\lambda^3 + \lambda + 2}.$$

Stability of commensurable transfer functions has a transparent extension to integer-order transfer functions, but we first need to characterize the idea of left- and right-half planes slightly differently. Another way to say that all the poles of an integer-order transfer function must be in the left half of the complex plane is to require that the absolute value of the angle of the poles (as complex numbers) is greater than $\pi/2$, or more formally, we have the following theorem.

Theorem 5.2. *The commensurable transfer function*

$$G(s) = \frac{\sum_{k=1}^{m} b_k s^{ak}}{\sum_{k=1}^{n} a_k s^{ak}} = \frac{\sum_{k=1}^{m} b_k \lambda^k}{\sum_{k=1}^{n} a_k \lambda^k}$$

is stable if the solutions to $a_n \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 = 0$*, denoted by* λ_k *are such that*

$$|\angle \lambda_k| > a\frac{\pi}{2}.$$

Proof. (for now follows very closely what is in [Valério and de Costa, 2013]).

If there are no repeated roots, then $G(s)$ can be written as the usual sort of partial fraction expansion

$$G(s) = \sum_{k=1}^{n} \frac{\rho_k}{\lambda - \lambda_k} = \sum_{k=1}^{n} \frac{\rho_k}{s^\alpha - \lambda_k}$$

so the impulse response will be of the form

$$x(t) = \sum_{k=1}^{n} \rho_k t^{\alpha-1}\, \mathrm{E}_{\alpha,\alpha}\left(\lambda_k t^\alpha\right).$$

The asymptotic behavior of the solution is, as $t \to +\infty$

$$x(t) \approx \sum_{k=1}^{n} \rho_k t^{\alpha-1} \frac{1}{\alpha} \left(\lambda_k t^\alpha\right)^{\frac{1-\alpha}{\alpha}} \mathrm{e}^{(\lambda_k t^\alpha)^{\frac{1}{\alpha}}} = \sum_{k=1}^{n} \frac{\rho_k}{\alpha} \lambda^{\frac{1-\alpha}{\alpha}} \mathrm{e}^{t \lambda_k^{\frac{1}{\alpha}}}.$$

This will be stable as long as the real part of $\lambda_k^{1/\alpha} < 0$. Since

$$\begin{aligned} \lambda_k^{1/\alpha} &= \left[|\lambda_k| \left(\cos\angle\lambda_k + \mathrm{i}\sin\angle\lambda_k\right)\right]^{1/\alpha} \\ &= |\lambda_k|^{1/\alpha} \left(\cos\frac{\angle\lambda_k}{\alpha} + \mathrm{i}\sin\frac{\angle\lambda_k}{\alpha}\right). \end{aligned}$$

The cosine term will be negative when its argument is greater than $\pi/2$, so we have

$$\frac{\angle\lambda_k}{a} > \frac{\pi}{2}$$

or

$$\angle\lambda_k > a\frac{\pi}{2}.$$

which is the desired result. ∎

Remark 5.2. *In the integer order case,* $a = 1$ *and the angle condition is the same as requiring all the poles to be in the left half plane.*

Consider the following program that uses the Grünwald-Letnikov numerical approximation to numerically compute the solution to

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) + a\frac{d^{\frac{1}{2}}x}{\mathrm{d}t^{\frac{1}{2}}}(t) + bx(t) = 1$$

assuming all initial conditions of all orders are zero. Note that the `bincoeff()` function uses the `gammaln()` function to extend the domain over which it returns a non-zero value. We will use this program to compute numerical step responses for the next couple examples.

```
%% controls roots stability examples
% numericall solve x' + a D^{1/2} x + b x = 1 with all zero initial
% conditions of all orders
close all; clear all;

a = -2; b = 5; alpha = 1/2;
t = linspace(0,5,6000);
dt = t(2)-t(1);
x = 0;
coefs = (-1)^1*bincoeff(alpha,1);

for k=2:length(t)
     thesum = dot(fliplr(x),coefs);
    x(k) = (1 - a*thesum/(dt^alpha) + x(k-1)/dt - b*x(k-1))/(1/dt + a/(dt^alpha));
    coefs(k) = (-1)^(k)*bincoeff(alpha,k);
end
plot(t,x);

function y = bincoeff(alpha,k)
     if(alpha < k)
          y = real((1/pi) * exp (gammaln (alpha + 1) - gammaln (k + 1)...
               + gammaln (k - alpha) ...
               + log (sin (pi * (alpha - k + 1)))));
     else
          y =  real(exp (gammaln (alpha + 1) - gammaln (k + 1) ...
               - gammaln (alpha - k + 1)));
     end
end
```

Example 5.3. *Is the solution*

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) - 2\frac{d^{\frac{1}{2}}x}{\mathrm{d}t^{\frac{1}{2}}}(t) + 5x(t) = 1$$

where all initial conditions of all orders are zero stable?

Since all initial conditions are zero, the Laplace transform of the equation is the same if we use the Riemann-Liouville or Caputo fractional derivatives, and

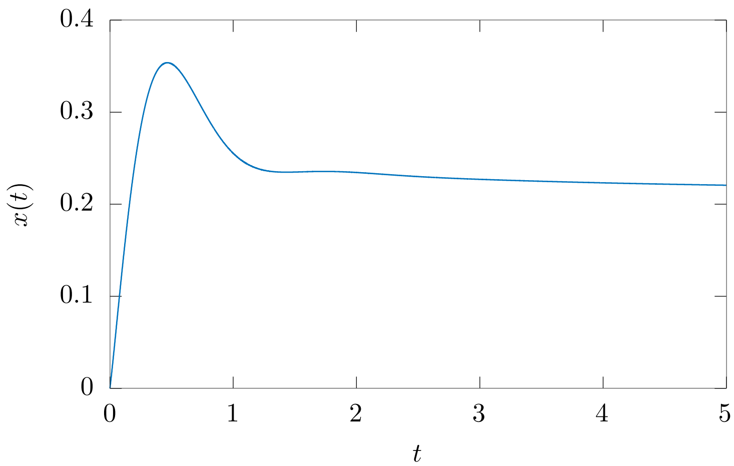

Figure 5.4: Step response for Example 5.3.

we get

$$X(s) = \frac{1}{s - 2s^{\frac{1}{2}} + 5}.$$

This transfer function is commensurate of order $a = 1/2$ and will be stable of all roots of

$$\lambda^2 - 2\lambda + 5 = 0$$

are such that $|\angle\lambda_i| > \frac{1}{2}\frac{\pi}{2}$. In this case $\lambda = 1 \pm 2\mathrm{i}$ which has an angle greater than $45°$, so the system is stable.

A numerically-computed solution is in Figure 5.4.

Example 5.4. *Is the solution*

$$\frac{\mathrm{d}x}{\mathrm{d}t}(t) - 4\frac{d^{\frac{1}{2}}x}{\mathrm{d}t^{\frac{1}{2}}}(t) + 5x(t) = 1$$

where all initial conditions of all orders are zero stable?

Since all initial conditions are zero, the Laplace transform of the equation is the same if we use the Riemann-Liouville or Caputo fractional derivatives, and we get

$$X(s) = \frac{1}{s - 4s^{\frac{1}{2}} + 5}.$$

This transfer function is commensurate of order $a = 1/2$ and will be stable of all roots of

$$\lambda^2 - 4\lambda + 5 = 0$$

are such that $|\angle\lambda_i| > \frac{1}{2}\frac{\pi}{2}$. In this case $\lambda = 2 \pm \mathrm{i}$ which has an angle less than $45°$, so the system is unstable.

A numerically-computed solution is in Figure 5.5.

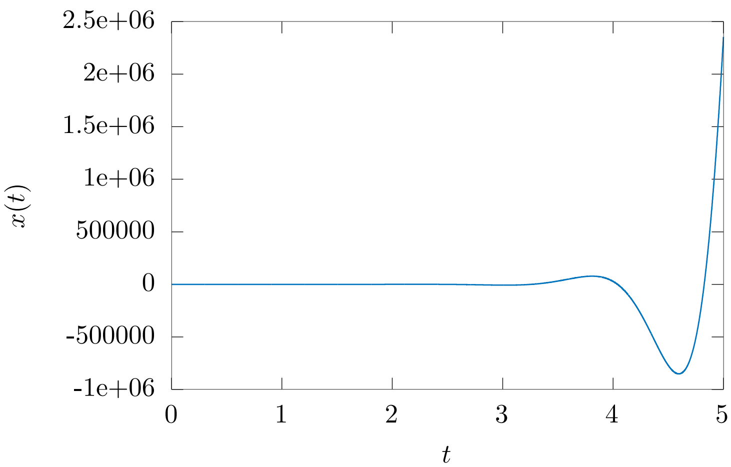

Figure 5.5: Step response for Example 5.4.

Because we normally do not factor separate complex terms, it makes sense to consider stability of a “second-order” type commensurable system. Analogous to the ubiquitous $\omega_n^2/\left(s^2+2\zeta\omega_n s+\omega_n^2\right)$ transfer function, consider

$$\begin{aligned} G(s) &= \frac{k}{ms^{2\alpha}+bs^{\alpha}+k} \\ &= \frac{\omega_n^2}{s^{2\alpha}+2\zeta\omega_n s^{\alpha}+\omega_n^2} \\ &= \frac{1}{\left(\frac{s^\alpha}{\omega_n}\right)^2+2\zeta\left(\frac{s^\alpha}{\omega_n}\right)+1}. \end{aligned} \tag{5.1}$$

The solutions of $\lambda^2/\omega_n^2+2\zeta\lambda/\omega_n+1=0$ are

$$\lambda=\omega_n\left(-\zeta\pm\sqrt{\zeta^2-1}\right) \qquad \Longleftrightarrow \qquad s^\alpha=\omega_n\left(-\zeta\pm\sqrt{\zeta^2-1}\right).$$

From this we can infer the following:

1. If $\zeta>1$ both roots are real and both are negative. Hence the transfer function is stable as long as $0<\alpha<2$

2. If $\zeta<-1$, then at least one root is positive and hence the transfer function is unstable.

3. If $\zeta=0$, then the roots are $\pm i\omega_n$. Hence the transfer function will be stable if $\alpha<1$ and unstable if $\alpha>1$.

4. If $0<\zeta<1$ and $0<\alpha\leq 1$, then the roots are complex with negative real part. Since $\alpha <\leq 1$ the entire left half plane is in the stability region and hence the transfer function is stable.

5. If $-1 < \zeta < 0$ and $0 < \alpha < 1$, then the roots are complex with a positive real part. To be stable, α must be small enough so that the roots are in the stability region. The tangent of the angle is given by the ratio of the imagainary to the real part of the pole, so for stability we require

$$\tan^{-1}\frac{\sqrt{1-\zeta^2}}{-\zeta} > \alpha\frac{\pi}{2} \quad \Longrightarrow \frac{\sqrt{1-\zeta^2}}{-\zeta} > \tan\frac{\alpha\pi}{2}.$$

   Because $\zeta < 0$ the left hand side is positive

$$\frac{1-\zeta^2}{\zeta^2} > \tan^2\frac{\alpha\pi}{2} \qquad \Longleftrightarrow \qquad \cos^2\frac{\alpha\pi}{2} > \zeta^2.$$

   Because $\zeta < 0$, we finally have that the transfer function is stable if

$$\zeta > -\cos\frac{\alpha\pi}{2}.$$

6. A similar argument shows that if $0 < \zeta < 1$ and $1 < \alpha < 2$, the transfer function is stable if

$$\zeta > -\cos\frac{\alpha\pi}{2}.$$

7. Finally, if $-1 < \zeta < 0$ and $1 \leq \alpha < 2$ the roots are complex and in the right half plane. Because $\alpha > 1$ the stability region is "less than" the left half plane. Hence the transfer function is always unstable.

Figure 5.6 illustrates the stability regions for a commensurable transfer fuction where the area above the curve is stable and the area under the curve is unstable.

5.2 Frequency Response

Bode plots for fractional-order systems are computationally straight-forward to construct by simply defining set of discrete frequencies and evaluating $|G(\mathrm{i}\omega)|$ and $\angle G(\mathrm{i}\omega)$ at those frequencies.

Example 5.5. *Consider*

$$G(s) = \frac{1}{s^{\frac{1}{2}} + 10}. \tag{5.2}$$

The Bode plot for that transfer function is illustrated in Figure 5.7.

Observe that for high frequencies the slope of the magnitude curve is -10db/decade and the phase is $-45°$.

The code that generated Figure 5.7 is:

```
w = logspace(-1,6,1000);
s = i*w;
G = 1./(s.^(1/2) + 10);
gh = figure(7, 'paperposition', size);
```

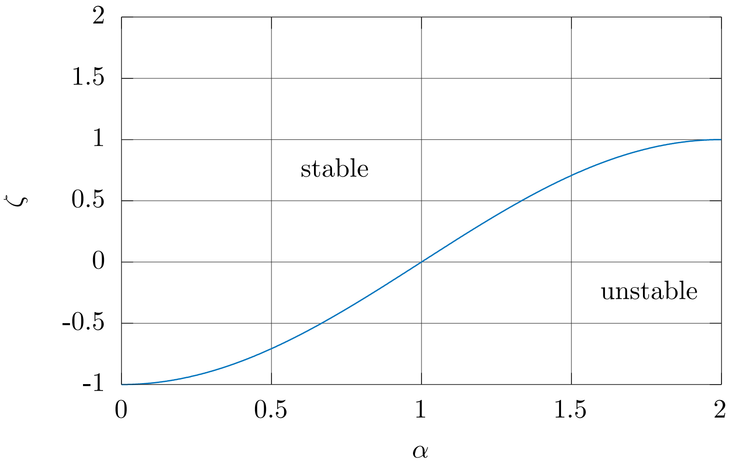

Figure 5.6: Stability regions for transfer functions of the form given by Equation 5.1.

```
subplot(2,1,1);
semilogx(w,20*log10(abs(G)),'linewidth',2);
ylabel('$\left|␣G(\iu␣\omega)␣\right|$')
grid on;
subplot(2,1,2)
semilogx(w,180*angle(G)/pi,'linewidth',2);
grid on;
xlabel('$\omega$');
ylabel('$\angle␣G(\iu␣\omega)$');
print('fracfreqex1.tex','-depslatex');
```

5.3 Integer-Order Approximations for Fractional-Order Transfer Functions

It may be the case that we want to use an integer-order transfer function that mimics a fractional order one. That could be, for example, if we design a fractional-order controller. In that case, when we implement the controller we may use an integer-order approximation that has approximately the same dynamics.

Recall that the way to draw the magnitude plot for

$$G(s) = \frac{p}{s+p}$$

the usual approach is to treat the magnitude as 0dB for frequencies less than p and then for frequencies greater than p, as a line with a slope of -20dB/decade

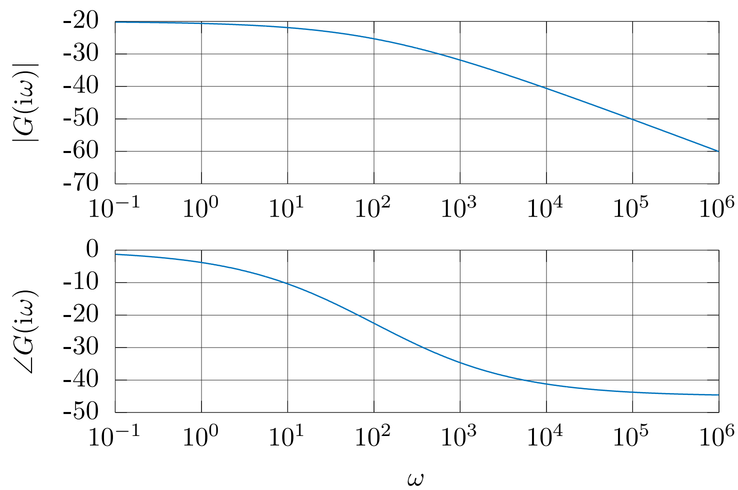

Figure 5.7: Frequency response for system in Equation 5.2.

(on a log-dB scale). So, when the frequency crosses from a lower to higher frequency than a pole, the slope decreases by 20dB/decade. Correspondingly when the frequency crosses a zero, the slope increases by 20dB/decade. Thus if we properly space poles and zeros we can construct a stair-case type magnitude plot that, on average, has a slope of, say, -10dB/decade. In fact, if the poles are evenly spaced logarithmically with the zeros half way between each pole (again logarithmically), we would expect a slope of -10dB/decade. Analogous reasoning will lead to a phase of $-45°$.

Example 5.6. *Consider*

$$G(s) = \frac{(s+0.0316)\,(s+0.316)\,(s+3.16)\,(s+31.6)}{(s+0.01)\,(s+0.1)\,(s+1)\,(s+10)\,(s+100)} \tag{5.3}$$

5.4 Fractional-Order PID

As mentioned at the beginning of this chapter, we can generalize PID control to the fractional case by allowing the derivative and integral terms to have a fractional order, *i.e.*,

$$u(t) = k_p e(t) + k_I \, {}_0\mathrm{D}_t^{-\lambda}\, e(t) + k_I \, {}_0\mathrm{D}_t^{\mu}\, e(t)$$

where $u(t)$ is the input and $e(t)$ is the error signal. Fractional-order PID should generally allow for better performance because integer-order PID is among the options available, *i.e.*, the special case where we choose $\lambda = \mu = 1$. However, we do need a method to search for all five parameters. The main benefit of PID control, and hence its ubiquity, is that it is fairly straight-forward to tune on a real system. In other words, we do not need a mathematical model to design the

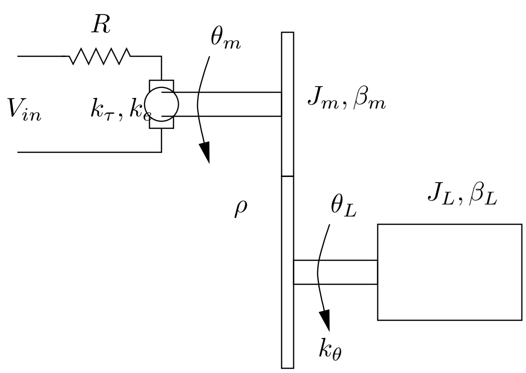

Figure 5.8: PID system.

controller if the approach is to implement PID in hardware and tune the three gains by experiment on the real system. Specifically, for a second-order plant, the three gains have very straight-forward effects: increasing the proporational gain increases the natural frequency, increasing the derivative gain increases the damping and increasing the integral gain decreases the time to eliminate steady-state error.

In this chapter we take an alternative approach when a good mathematical model is avaiable, which is to implement an optimization algorithm to search for the best gains. Because optimization is not generally covered in many controls courses, first we will present a problem implementing this in full detail.

5.4.1 Optimization of a PID Controller

The example in this section is taken from [Dingyu Xue et al., 2006]. Consider a DC motor driving a load as illustrated in Figure 5.8. An input voltage drives a circuit with a dc motor. The motor shaft has an inertia J_m and its motion is resisted by viscous friction described by the coefficient β_m. The angle of the motor shaft is given by θ_m. The shaft is attached to a gearbox with a gear ratio of ρ, and the output of the gear box drives the load. The inertia of the load is J_L and the motion of the load is resisted by viscous friction with a coefficient of β_L. The shaft connecting the output of the gearbox to the load is flexible and acts like a torsional spring with spring constant k_θ.

Problem Statement: Use PID control given by Equation 5.3 where $u(t)$ is the voltage into the circuit and $e(t)$ is the difference between the desired angle of the load and the actual angle of the load. Find the optimal gains, k_p, k_d and k_i such that:

- The cost function

$$J = \int_0^{60} \left(\theta_{desired}(t) - \theta_L(t)\right)^2 \mathrm{d}t \tag{5.4}$$

  is minimized;

- $|V_{in}(t)| \leq 220$
- The torque in the output shaft of the gearbox driving the load always has a magnitude less than 100Nm.

This is a *constrained optmization problem*. We will use the Matlab function `fmincon()` to find a solution.

Kirchhoff's voltage law gives

$$V_{in}(t) = i(t)R + k_e\dot{\theta}_m(t)$$

where $i(t)$ is the current through the motor. Hence, the output torque from the motor is given by

$$\tau_m(t) = k_\tau \left(\frac{V_{in}(t) - k_e\dot{\theta}_m(t)}{R} \right).$$

Newton's law on the motor shaft is given by

$$J_m\ddot{\theta}_m(t) = k_\tau \left(\frac{V_{in}(t) - k_e\dot{\theta}_m(t)}{R} \right) - \beta_m\dot{\theta}_m(t) + \frac{k_\theta}{\rho}\left(\theta_L(t) - \frac{\theta_m(t)}{\rho}\right)$$

and similarly on the load

$$J_l\ddot{\theta}_L(t) = -k_\theta\left(\theta_L(t) - \frac{\theta_m(t)}{\rho}\right) - \beta_L\dot{\theta}_L(t).$$

If we define the state vector as

$$x(t) = \begin{bmatrix} \theta_L(t) \\ \dot{\theta}_L(t) \\ \theta_m(t) \\ \dot{\theta}_m(t) \end{bmatrix}$$

then we can express the equations of motion in the standard state space form of $\dot{x}(t) = Ax(t) + BV_{in}(t)$ as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} \theta_L(t) \\ \dot{\theta}_L(t) \\ \theta_m(t) \\ \dot{\theta}_m(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 1 \\ -\frac{k_\theta}{J_L} & -\frac{\beta_L}{J_L} & \frac{k_\theta}{\rho J_L} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_\theta}{\rho J_m} & 0 & -\frac{k_\theta}{\rho^2 J_m} & -\left(\frac{k_\tau k_e}{J_m R} + \frac{\beta_M}{J_m}\right) \end{bmatrix} \begin{bmatrix} \theta_L(t) \\ \dot{\theta}_L(t) \\ \theta_m(t) \\ \dot{\theta}_m(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{k_\tau}{J_m R} \end{bmatrix} V_{in}(t). \tag{5.5}$$

If the only state we can measure is $\theta_L(t)$ then in the standard form $y(t) = Cx(t) + DV_{in}(t)$ we have

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} x(t).$$

From [Dingyu Xue et al., 2006] let

- $k_\theta = 1280.2$

- $k_\tau = 10$
- $k_e = 10$
- $J_m = 0.5$
- $J_L = 25$
- $\rho = 20$
- $\beta_m = 0.1$
- $\beta_L = 25$
- $R = 20$.

Using those values and constructing A, B, C and D in Matlab, and then using `G = ss2tf(A,B,C,D)` gives the transfer function from the input voltage to the angular position of the load as

$$G(s) = \frac{\Theta_L(s)}{V_{in}(s)} = \frac{2.56}{s^4 + 11.2s^3 + 67.81s^2 + 528.7s}$$

which is a type 1 system, which also makes sense.

If we already have PID gains, or we wish to tune the controller by hand, then we can use set numerical values for the gains and use `Con = tf([kd + kp + ki],[1 0])` to define the PID controller transfer function, from which `step(feedback(Con*G,1))` will plot the step response.

Because our goal is to compare integer-order PID control with fractional-order PID control, we need an objective basis for comparison. As such, we will set up an optimization problem to find the "best" PID gains in each case, and compare the results. We will use `(fmincon()` in Matlab. From the documentation, the function attempts to find the minimum of a problem specified by

$$\min_x f(x) \text{such that} \begin{cases} c(x) & \leq 0 \\ c_{eq}(x) & = 0 \\ Ax & \leq b \\ A_{eq}x & = b_{eq} \\ lb & \leq x \leq ub \end{cases} \tag{5.6}$$

For the constraints

- $c(x) \leq 0$ are the nonlinear inequality constraints
- $c_{eq}(x) = 0$ are the nonlinear equality constraints
- $Ax \leq b$ are the linear inequality constraints
- $A_{eq}x = b_{eq}$ are the linear equality constraints
- $lb \leq x \leq ub$ are the bound constraints.

The Matlab syntax is

`fmincon(fun,x0,A,b,Aeq,beq,lb,ub,nonlin)`

where

- `fun` is the function to be minimized
- `x0` is the initial guess for the solution
- `A` is the matrix in the linear inequality constraints
- `b` is the right hand side of the linear inequality consnstraints
- `Aeq` is the matrix in the linear equality constraints
- `beq` is the right hand side of the linear equality constraints
- `lb` and `ub` are the lower and upper bounds of the search space
- `nonlin` is the function of the nonlinear inequality and equality constraints (the are combined!).

For our problem, the function to be minimixed is given by Equation 5.4. Note that we are searching for optimal gains, so the gains are represented by x in the previous equations describing `fmincon()`, which of course can be confusing because we are using x to represent the states in our problem. We need a matlab function that evaluates Equation 5.4. So the function has to determine the step response for the given gains, and then compute the integral of the square of the difference between the desired angle and actual angle of the load.

So the function will have to call either `step()` or `ode45()` to compute the step response. With some foresight knowing we will have to compute the input voltage and the torque in the flexible shaft, we will solve the problem in state space using `ode45()`. The function (`sossserr()` "sum of squares for the state space system of the error") will need to know the equations of motion, time and the gains:

```
function ret = sossserr(ks,A,B,C,D,t)
    [t,y] = ode45(@(t,x)sspidrhs(t,x,ks,A,B,C,D),t,[0; 0; 0; 0; 0]);
    ret = trapz(((1-y(:,1)).^2)*(t(2)-t(1)));
end
```

The right hand side of the differential equations needs an additional state appended to be able to compute the integral of the error. So a fifth state which is the integral of the error is appended, so its derivative is simply the error:

```
function xdot = sspidrhs(t,x,ks,A,B,C,D)
    error = 1 - x(1);
    errordot = 0 - x(2);
    input = ks(1)*error + ks(2)*errordot + ks(3)*x(5);
    xdot(1:4) = A*x(1:4)+B*input;
```

```
    xdot(5) = error;
    xdot = xdot';
end
```

Using these functions, we now can find the best PID gains to minimize the integral of the square of the error. However, if we simply try to minimize J, then the algorithm to determine the gains will not coverge because arbitrarily high gains will drive the solution to the desired value of one. Hence, we must add bounds to the search using `lb` and `ub` or add in some other constraints. The real constraints are the voltage and torque limits, but because those are relatively difficult to implement, for now we will just put bounds on the allowable gain values.

Bibliography

[Dingyu Xue et al., 2006] Dingyu Xue, Chunna Zhao, and YangQuan Chen (2006). Fractional order pid control of a dc-motor with elastic shaft: a case study. In *2006 American Control Conference*, pages 3182–3187.

[Gorenflo et al., 2014] Gorenflo, R., Kilbas, A. A., Mainiardi, F., and Rogosin, S. V. (2014). *Mittag-Leffler Functions, Related Topics and Applications*. Springer-Verlag, Berlin Heidelberg.

[Newton, 1687] Newton, I. (1687). *Philosophiae Naturaliz Principia Mathematica*. J. Societatis Regiae ac Typis J. Streater.

[Valério and de Costa, 2013] Valério, D. and de Costa, J. S. (2013). *An Introduction to Fractional Control*. Institution of Engineering and Technology, London, United Kingdom.