Fractional Matricial Calculus
and its application to
multiple variable SI models.

Facultad de Ciencias
Universidad Zaragoza

Universidad
Zaragoza
1542

Jorge Paz-Peñuelas Oliván

Dissertation in Mathematics.

Faculty of Science, University of Zaragoza.

Directed by Luciano Abadías Ullod.

September 2022 - December 2023.

To my mother,
who gifted me her passion for Mathematics.

Abstract

(El resumen en castellano se encuentra más abajo en la página iii.)

The current work is a final project of the Mathematics undergrad course at the university of Zaragoza. It introduces and explores the usage of fractional-order derivatives, in particular the *Caputo* derivative, for the study of a fractional-order susceptible-infected (SI) epidemiology model for the diffusion processes of a Protein Residue Network (PRN), as a recreation of the work of Abadías et Al in reference [1]. We will be working with simple symmetric connected networks to study perturbation spread, where each node i will be assigned a probability $s_i(t)$ of still being susceptible and another $x_i(t)$ of having been perturbed at time t (this makes $s_i(t) + x_i(t) = 1$).

We will begin by introducing a few preliminary results such as scalar series derivation and the scalar and matricial versions of the exponential function, and we will see that they are the solution to a system of first-order linear differential equations. We will continue by introducing the two Eulerian integrals: the Beta and Gamma functions, and prove a few properties they satisfy, namely the Stirling formula's extension to the real line and the relationship between the Beta and Gamma functions, that will prove necessary for the remainder of the work.

After these preliminary results we will proceed with defining the Mittag-Leffler function of parameter $\alpha > 0$ and variable taking values in $\mathbb{C}$ as a generalisation of the exponential, to afterwards extend it to its complex matricial analogue and prove its convergence on the whole complex plane $\mathbb{C}$. Thereafter we will focus the scope of the work back to the real line and define the Riemman-Liouville fractional order integral and the Caputo fractional order derivative along once again with a discussion of a few of their most basic properties necessary for the statements that will follow.

Once this framework has been established we will proceed with the definition of the classical SI model and state its relationship with the PRNs that reference [1] studies, together with an introduction of the adjacency matrix A we will be using to represent the network. Rather than solving this model we provide two approximations to the solution through two different transformations to the differential equation system, which we justify and discuss briefly. This will be followed by the introduction of our fractional-order adaptation of the classical SI model as done in article [1] and we will see that the two approximations done to the classic model will also give us analogous approximations to the solution of the fractional order model; which will be necessary as we won't be able to solve the new model analytically. These two approximations are a linear one and the Lee-Tenneti-Eun (LTE) transformation, named after the authors by whom it was first used, and both rely heavily on the Mittag-Leffler functions. We will prove convergence for both approximations and discuss their qualitative behaviour, mainly their limit when time is sufficiently large. Afterwards we will prove that the solution to the original (un-approximated) fractional order SI model is bounded by the LTE solution and this is simultaneously always smaller than the linear approximation, and how the former approximation represents an upper bound to the evolution of the dynamics of the PRN.

In the final discussion we will find that the appearance of the Mittag-Leffler functions in the solution due to the introduction of the fractional-order derivatives to the SI model will give us an additional parameter α which will help us better model long-range processes and phenomena in different timescales without needing to modify the model or its parameters. One of the biggest fields of application for these results is the aid in the selection of drug candidates for the treatment of illnesses; one of which is Covid as is discussed in reference [1].

Resumen en castellano

(Abstract in Spanish to meet university requirements.)

Este documento recoge el proyecto de fin de grado del grado de matemáticas por la Universidad de Zaragoza del autor. En él se presentarán las derivadas de orden fraccionario, en particular la derivada de Caputo, y su utilización en la definición del modelo epidemiológico susceptible-infectado (SI) de orden fraccionario para el estudio de los procesos de difusión que tienen lugar en las redes de residuos de proteínas (*Protein Residue Networks*), que denotaremos mediante PRN, como recreación del artículo de Abadías et Al de la referencia [1]. Trabajaremos con redes conexas simples y simétricas para estudiar la propagación de perturbaciones, donde a cada nodo i se le asignará una probabilidad $s_i(t)$ de seguir siendo susceptible y otra $x_i(t)$ para denotar la posibilidad que ya haya sido perturbado a tiempo t, de manera que $s_i(t) + x_i(t) = 1$.

Comenzaremos el trabajo con una breve sección de resultados preliminares como la derivada de una serie escalar o la definición de la función exponencial tanto escalar como matricial, y demostraremos que la última de las cuales es solución de un sistema de ecuaciones diferenciales lineales de primer orden. A continuación se proporcionarán las definiciones de las dos funciones integrales Eulerianas: las funciones Gamma y Beta, y se discutirán propiedades conocidas como la extensión de la fórmula de Stirling a números reales positivos mediante la función Gamma o el cociente que relaciona la función Beta con la Gamma.

Una vez visto esto se definirá la función Mittag-Leffler de parámetro $\alpha > 0$ como una generalización de la exponencial, y cuya variable tomará valores en el plano complejo $\mathbb{C}$, y se extenderá la definición al caso matricial como la función del producto de una matriz compleja constante por nuestra variable para probar su convergencia en la totalidad del plano. Tras ello disminuiremos la extensión de los resultados a la recta real y definiremos la integral de orden fraccionario $\alpha > 0$ como una generalización de la formula de Cauchy para la integración repetida, y la derivada fraccionaria de Caputo junto a algunas de sus propiedades más elementales que serán necesarias para las demostraciones posteriores.

Una vez se hayan establecido estos resultados comenzaremos con la definición del modelo SI clásico junto con su relación a los PRNs mencionados arriba que se estudian en el artículo [1], además de comentar la matriz de adyacencia A que representará la red a estudiar. En lugar de buscar la solución para dicho modelo proporcionaremos dos aproximaciones a esta mediante transformaciones al sistema de ecuaciones que se discutirán y explicarán brevemente. Esto dará paso a la introducción del modelo SI de orden fraccionario mediante la inclusión de la derivada de Caputo tal y como se hace en nuestro artículo de referencia [1], y veremos que las dos transformaciones que se realizaron

en el caso clásico también nos darán aproximaciones análogas a la solución del caso fraccionario; que resultarán necesarias al no ser nuestro modelo fraccionario resoluble analíticamente. Dichas aproximaciones incluyen una lineal, y la transformación Lee-Tenneti-Eun (LTE), llamada así en honor a los autores que la usaron por primera vez. Veremos que ambas aproximaciones estarán estrechamente relacionadas con la función Mittag-Leffler, y probaremos convergencia para ambas, además de estudiar cualitativamente su comportamiento, principalmente en lo referido a tiempos grandes. Tras esto realizaremos una comparativa de la solución del problema fraccionario original y las dos aproximaciones, y veremos que la aproximación proporcionada por la transformación LTE representa una cota superior de la solución original y por tanto puede entenderse como el caso de difusión de perturbaciones más veloz en la PRN.

Finalmente comentaremos la interpretación física de la aparición de la función Mittag-Leffler como solución al modelo como consecuencia de haber introducido la derivada fraccionaria de Caputo, en especial en lo que refiere al parámetro α. Veremos que dicho parámetro nos permitirá modelar comportamientos y fenómenos a distintas escalas de tiempo o efectos de distinto rango espacial sin necesidad de adaptar los parámetros del modelo SI. Uno de los mayores campos de aplicación de dichos resultados (y por tanto de las derivadas fraccionaras en sistemas de ecuaciones diferenciales) es el estudio de posibles drogas para el tratamiento de enfermedades, tal y como el Covid como se discute en el artículo [1].

Contents

Chapter 1

Preliminary Results

1.1 The exponential function

1.1.1 The complex exponential function

One way to define the exponential function would be to describe it as the function equal to its derivative, or in slightly more general terms, given $a \in \mathbb{C}$, the solution to the following differential equation.

$$\frac{\partial}{\partial z} f(z) = af(z) \ , \qquad z \in \mathbb{C} \ . \tag{1.1}$$

We want to be able to write this function as a series, as many properties in analysis are considerably more straightforward when one considers the series representation.

Proposition 1.1. *Let $f : \mathbb{C} \to \mathbb{C}$ be an analytic function with the series representation*

$$f(z) = \sum_{n=0}^{\infty} a_n (z-c)^n \ , \qquad \forall z \in B(c,R) \ , \tag{1.2}$$

where $c \in \mathbb{C}$, $R \in \mathbb{R}$ and $B(c,R) = \{z \in \mathbb{C} : |z-c| < R\}$. R is called the radius of convergence.

Then the derivative of f is given by the following series.

$$f'(z) = \sum_{n=0}^{\infty} (n+1) a_{n+1} (z-c)^n \ , \qquad \forall z \in B(c,R) \ . \tag{1.3}$$

Remark. The proposition doesn't say anything about whether the result holds on the border of the convergence disk, $\partial B(c,R) = \{z \in \mathbb{C} : |z-c| = R\}$, but in this document we will be working with series convergent on the whole plane ($R = \infty$), so these border effects won't concern us.

Proposition 1.2. *For any $a \in \mathbb{C}$, the function $e^{az} = \sum_{k=0}^{\infty} \frac{(az)^k}{k!}$, $z \in \mathbb{C}$, is the solution to equation* (1.1) *with initial value $f(0) = 1$.*

It is called the exponential function, and it converges for all $z \in \mathbb{C}$.

Proof. We will simply compute the derivative of the series as per (1.1), knowing that $R = \infty$.

$$\frac{\partial}{\partial z} f(z) = \frac{\partial}{\partial z} \sum_{k=0}^{\infty} \frac{(az)^k}{k!} = \sum_{k=0}^{\infty} (k+1) \frac{a^{k+1}}{(k+1)!} z^k = a \sum_{k=0}^{\infty} \frac{(az)^k}{k!} = af(z) \ , \qquad z \in \mathbb{C} \ ,$$

and $f(0) = 1$.

In regards to convergence of the series, we will see in the next chapter (Proposition 2.1) a generalised result which will imply convergence on the whole plane. $\square$

Remark. The solution to equation (1.1) with initial value $f(z_0) = f_0$ is $f(z) = f_0 e^{a(z-z_0)}$.

1.1.2 The matricial exponential function

Once we have seen the exponential function $f : \mathbb{C} \to \mathbb{C}$ we are going to define its matrix counterpart.

Definition (Matricial exponential function)**.** Let $A \in \mathbb{C}^{n\times n}$ be a constant complex matrix of dimension n by n, for some $n \in \mathbb{N}$. Then the matrix exponential is a function $f : \mathbb{C} \to \mathbb{C}^{n\times n}$ given by

$$f(z) = e^{zA} = \sum_{k=0}^{\infty} \frac{(zA)^k}{k!} \ , \qquad z \in \mathbb{C} \ , \tag{1.4}$$

which is convergent at all $z \in \mathbb{C}$.

Proof. Proof of the convergence will become apparent once we see a generalised result in chapter 2. $\square$

Proposition 1.3. *Consider $f : (0,\infty) \to \mathbb{R}^n$, with $f(t) = \big(f_1(t), \dots, f_n(t)\big)$ such that $f_i : (0,\infty) \to \mathbb{R}$ are derivable for all $i = 1, \dots, n$, $t > 0$. Consider also a matrix $A \in \mathbb{C}^{n\times n}$.*

The multivariate linear initial-value problem

$$\begin{cases} D_t f(t) = A f(t) \ , & t > 0 \ , \\ f(0) = v \ , \end{cases}$$

where we denote by D_t the derivative operator with respect to t, and $v \in \mathbb{R}^n$, has solution

$$f(t) = e^{tA} v \ , \qquad t > 0 \ .$$

Proof. Firstly, notice that each component of the solution f is given by $f_i(t) = e_i^\top e^{tA} v$, with e_i the i-th vector of the canonical basis of $\mathbb{R}^n$, and that the equation $D_t f(t) = A f(t)$ can be rewritten as

$$D_t f_i(t) = A_{i,1} f_1(t) + \dots + A_{i,n} f_n(t) \qquad \text{for } i = 1, \dots, n \ , \quad t > 0.$$

Let us compute then both sides of the equality.

$$D_t f_i\,(t) = D_t\left(e_i^\top e^{tA} v\right) = D_t\left(e_i^\top \sum_{k=0}^{\infty} t^k \frac{A^k}{k!} v\right) = D_t\left(\sum_{k=0}^{\infty} t^k \frac{\left(A^k\right)_i v}{k!}\right) \ , \quad t > 0 \ ,$$

where by $(B)_i$ we denote the i-th row of a matrix B. Note that we now have a scalar series. To be able to apply Proposition 1.1 we are going to see that it converges for all $t > 0$.

$$\sum_{k=0}^{\infty} t^k \frac{\left(A^k\right)_i v}{k!} \le \max_j \left\{\left|(v)_j\right|\right\} \sum_{k=0}^{\infty} t^k \frac{\left(A^k\right)_i \mathbf{1}}{k!} \ , \qquad t > 0,$$

where $(v)_j$ denotes the j-th component of v and $\mathbf{1}$ is the 1s vector. The above sum converges with the same radius of convergence as the exponential function. Hence

$$D_t f_i(t) = \sum_{k=0}^{\infty}(k+1)t^k \frac{\left(A^{k+1}\right)_i v}{(k+1)!} = \sum_{k=0}^{\infty} t^k \frac{\left(A^{k+1}\right)_i v}{k!} \ , \qquad t > 0 \ .$$

And the right-hand side:

$$\sum_{j=1}^{n} A_{i,j} f_j(t) = \sum_{j=1}^{n} A_{i,j} \left(e_j{}^\top e^{tA} v\right) = \sum_{j=1}^{n} A_{i,j} \sum_{k=0}^{\infty} t^k \frac{\left(A^k\right)_j v}{k!} =$$
$$\overset{\star}{=} \sum_{k=0}^{\infty} \frac{\sum_{j=1}^{n} t^k A_{i,j} \left(A^k\right)_j v}{k!} = \sum_{k=0}^{\infty} t^k \frac{\left(A^{k+1}\right)_i v}{k!} \ , \qquad t > 0 \ .$$

The swap of sum and series in $\star$ is not problematic as we know that the series converges regardless of the value of t (the product by $e_j{}^\top$ and v will only make the result a scalar rather than a matrix, but they won't influence convergence, as they are constant).

Hence we obtain that in fact $D_t f_i(t) = \sum_{j=1}^{\infty} A_{i,j} f_j(t)$ is true for every $i = 1, \dots, n$, and for every $t > 0$. Checking $f(0) = v$ is trivial (remember that the exponential of the 0 matrix is the corresponding identity matrix), and therefore the proposition is proved. □

Remark. If the initial conditions were at an arbitrary time $t_0 \in \mathbb{R}$ with the differential equation true for every $t > t_0$ then the result and proof would be analogous working with $t - t_0$ instead of t.

1.2 The Stirling formula

The Stirling Formula is a very well known formula that states that for $n \in \mathbb{N}$, as $n \to \infty$,

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n \ . \tag{1.5}$$

Definition (The Euler Gamma function)**.** The Euler gamma function $\Gamma : (0, \infty) \to \mathbb{R}$ is most commonly defined as follows

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} \mathrm{d}t \ , \qquad x > 0 \ . \tag{1.6}$$

It is an extension of the factorial to the positive real numbers, and as such has been given a broad array of equivalent definitions. Reference [2] collects many of them, as well as proving their equivalence.

Proposition 1.4. *Stirling's formula can be extended to the* Γ *function as follows,*

$$\lim_{x\to\infty} \Gamma(x+1) = \sqrt{2\pi x} \left(\frac{x}{e}\right)^x \ , \qquad \textit{for any } x > 0 \ . \tag{1.7}$$

Proof. We are going to follow reference [3] for a short proof. Let us begin with the integral definition (1.6) and perform a change of variable $t = u^2$, which gives us

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t \;=\; 2\int_0^\infty u^{2x-1}e^{-u^2}\mathrm{d}u \;, \qquad x > 0 \;.$$

With this, consider now the quotient below and a second change of variable $u = \sqrt{x} + v$ in $\star$.

$$\frac{\Gamma(x)e^x\sqrt{x}}{x^x} \;=\; 2\int_0^\infty e^{x-u^2}\left(\frac{u}{\sqrt{x}}\right)^{2x-1}\mathrm{d}u \;\overset{\star}{=}\; 2\int_{-\sqrt{x}}^\infty e^{-2v\sqrt{x}}\left(1+\frac{v}{\sqrt{x}}\right)^{2x-1} e^{-v^2}\mathrm{d}v \;.$$

If we extend the above integration interval to the whole real line we can define the function

$$\varphi_x(v) = \chi_{(-\sqrt{x},\infty)}(v)e^{-2v\sqrt{x}}\left(1+\frac{v}{\sqrt{x}}\right)^{2x-1} \;, \qquad x > 0 \;,$$

where $\chi_I(v)$ is the characteristic function of interval I, allowing us to write the integral as

$$\frac{\Gamma(x)e^x\sqrt{x}}{x^x} \;=\; 2\int_{-\infty}^\infty \varphi_x(v)e^{-v^2}\mathrm{d}v \;, \qquad x > 0 \;. \tag{1.8}$$

Let us study the function $\varphi_x(v)$. Note that it is positive, continuous and derivable on $(-\sqrt{x},\infty)$, and it is also easy to see that $\varphi_x(v) \to 0$ when $v \to \infty$ for any fixed x. For $x > 1/2$, $\lim \varphi_x(v)$ when $v \to -\sqrt{x}^+$ exists and is 0, and thus $\varphi_x(v)$ will be a bounded function. We see below that it reaches its maximum at $v = -(2\sqrt{x})^{-1}$ for $x > 1/2$,

$$\varphi_x'(v) = \varphi_x(v)\left[-2\sqrt{x} + (2x-1)\left(1+\frac{v}{\sqrt{x}}\right)^{-1}\left(\frac{1}{\sqrt{x}}\right)\right] = 0 \quad\Longleftrightarrow\quad v = -\frac{1}{2\sqrt{x}} \;, \qquad x > \frac{1}{2} \;.$$

Said maximum is

$$\varphi_x\left(-\frac{1}{2\sqrt{x}}\right) = e\left(1-\frac{1}{2x}\right)^{2x-1} < e = \Phi \;. \tag{1.9}$$

Let us now study its behaviour with respect to x. For a fixed $v \in [-\sqrt{x},\infty)$, when $v << x$ (that is, as $x \to \infty$), we have that

$$\begin{aligned}\log\varphi_x(v) &= -2v\sqrt{x} + (2x-1)\log\left(1+\frac{v}{\sqrt{x}}\right) = \\ &= -2v\sqrt{x} + (2x-1)\left[\frac{v}{\sqrt{x}} - \frac{v^2}{2x} + \dots\right] = -v^2 + O\left(\frac{1}{\sqrt{x}}\right) \;, \qquad x > 0 \;,\end{aligned}$$

where we have used the plynomial approximation of $\log(1+y)$ for y close to 0. This makes $\lim_{x\to\infty}\log\varphi_x(v) = -v^2$ and thus $\lim_{x\to\infty}\varphi_x(v) = e^{-v^2}$ as the characteristic function interval will tend to the complete real line.

As the bound Φ is constant and our integral (1.8) is of $\varphi_x(v)e^{-v^2}$ with respect to v, where we know that e^{-v^2} is integrable over $\mathbb{R}$, by the Dominated Convergence Theorem we can swap the limit with the integral sign when taking the limit of (1.8). Hence,

$$\lim_{x\to\infty}\frac{\Gamma(x)e^x\sqrt{x}}{x^x} = 2\int_{-\infty}^\infty \lim_{x\to\infty}\varphi_x(v)e^{-v^2}\mathrm{d}v = 2\int_{-\infty}^\infty e^{-2v^2}\mathrm{d}v \;, \qquad x > \frac{1}{2} \;.$$

Now we only need to compute the value of the integral.

$$2\int_{-\infty}^{\infty} e^{-2v^2}\mathrm{d}v = 4\int_{0}^{\infty} e^{-2v^2}\mathrm{d}v = 4\int_{0}^{\infty} \frac{\sqrt{2}}{4} t^{1/2} e^{-t}\mathrm{d}t = \sqrt{2}\Gamma(1/2) = \sqrt{2\pi} \ .$$

This gives us (1.7), knowing $\Gamma(x+1) = x\Gamma(x)$ for any $x > 0$. □

1.3 The Euler Beta function

Definition (The Euler Beta function)**.** The Euler Beta function, otherwise known as the Eulerian integral of first kind, is given by

$$B(x,y) = \int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t \ , \qquad \text{for any } x, y > 0 \ . \tag{1.10}$$

Proposition 1.5. *For all $x, y > 0$, it follows that*

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \ .$$

Proof. By equation (1.6) we can write

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \int_0^\infty t^{x-1}e^{-t}\mathrm{d}t \int_0^\infty \tau^{y-1}e^{-\tau}\mathrm{d}\tau = \int_{t=0}^\infty \int_{\tau=0}^\infty e^{-t-\tau}t^{x-1}\tau^{y-1}\mathrm{d}t\mathrm{d}\tau \\ &= \int_{t=0}^\infty \int_{\tau=0}^\infty (t+\tau)^{x+y-1}e^{-(t+\tau)}\left(\frac{t}{t+\tau}\right)^{x-1}\left(\frac{\tau}{t+\tau}\right)^{y-1}(t+\tau)\mathrm{d}t\mathrm{d}\tau \ , \qquad x, y > 0 \ .\end{aligned}$$

Once we have obtained the above double integral we perform a change of variables to $\mu = (t+\tau)$ and $\gamma = \frac{t}{t+\tau}$. This makes the integration domain change to $(\mu, \gamma) \in [0,\infty) \times [0,1]$, and as

$$|\mathrm{det}J| = \left\|\begin{matrix} 1 & 1 \\ \frac{\tau}{(t+\tau)^2} & -\frac{t}{(t+\tau)^2}\end{matrix}\right\| = \frac{1}{t+\tau} \ ,$$

where J is the Jacobian of the change of variables function, we get

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \int_{\mu=0}^\infty \int_{\gamma=0}^1 \mu^{x+y-1}e^{-\mu}\gamma^{x-1}(1-\gamma)^{y-1}\mathrm{d}\mu\mathrm{d}\gamma \\ &= \int_{\mu=0}^\infty \mu^{x+y-1}e^{-\mu}\mathrm{d}\mu \int_{\gamma=0}^1 \gamma^{x-1}(1-\gamma)^{y-1}\mathrm{d}\gamma \\ &= \Gamma(x+y)B(x,y) \ , \qquad x, y > 0 \ .\end{aligned}$$

From where we get the equality of the proposition. □

Chapter 2

Fractional Calculus

When one is first introduced to the concept of calculus they are initially taught about limits and how the derivative of a function is the limit of its rate of change when the interval over which it is studied tends to length zero. Only when this concept is understood is the integral introduced as the right inverse of the derivative (although it is oftentimes described as a way to recover a function from its derivative, save a constant, which refers to composition on the left of the derivative).

We will follow a different approach to fractional calculus. Our main goal is to define a fractional derivative consistent with the existing *whole* derivative so that it can be used to solve fractional differential problems. To do so, we will first define the fractional integral and then use it to consider a derivative of order $\alpha > 0$ as the integral of order $m - \alpha$ of the m-th whole derivative, with $m \in \mathbb{N}$ such that $m - 1 < \alpha \leq m$.

Before we get into that, however, we introduce below the Mittag-Leffler function, which will be closely related to the solution to the fractional differential equations.

2.1 The Mittag-Leffler function

Definition (The classical Mittag-Leffler function, $E_\alpha(z)$)**.** Given $\alpha > 0$, the Mittag-Leffler function (*ML function* for short) of order α is the function defined by the following power series,

$$E_\alpha(z) := \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + 1)} , \qquad z \in \mathbb{C} . \tag{2.1}$$

We will see shortly that it is in fact a well defined function on the whole complex plane, which makes it an entire function. It is easy to see that $E_1(z) = e^z$, given that $\Gamma(k+1) = k!$ for $k \in \mathbb{N}_0$. This makes the ML function a generalisation of the exponential function.

Proposition 2.1. *Given $n \in \mathbb{N}$ and $A \in \mathbb{C}^{n \times n}$, the matricial Mittag-Leffler function $E_\alpha(zA)$ with $\alpha > 0$, $z \in \mathbb{C}$, whose power series representation is*

$$\sum_{k=0}^{\infty} \frac{(zA)^k}{\Gamma(\alpha k + 1)} , \qquad z \in \mathbb{C} , \tag{2.2}$$

converges absolutely and uniformly for any $z \in \mathbb{C}$.

Proof. A series of matricial terms will converge to a matrix of the same dimension if all the elements converge. In our case that is if

$$\left(E_\alpha(Az)\right)_{i,j} = \sum_{k=0}^{\infty} \frac{z^k \left(A^k\right)_{i,j}}{\Gamma(\alpha k+1)} \ , \qquad z \in \mathbb{C} \ ,$$

converges for all $i, j = 1, \ldots, n$.

Notice that if we try to prove the result by the quotient criterion,

$$\lim_{k\to\infty} \frac{\left| z^{k+1}\left(A^{k+1}\right)_{i,j} / \Gamma\left(\alpha(k+1)+1\right)\right|}{\left| z^{k}\left(A^{k}\right)_{i,j} / \Gamma(\alpha k+1)\right|} = \lim_{k\to\infty} |z| \frac{\Gamma\left(\alpha k+1\right)}{\Gamma\left(\alpha(k+1)+1\right)} \frac{\left|\left(A^{k+1}\right)_{i,j}\right|}{\left|\left(A^{k}\right)_{i,j}\right|} \ ,$$

we run into the problem that $\left(A^k\right)_{i,j}$ could be 0 for some combination of k, i, j, but not necessarily all of them, giving us trouble with the behaviour of the limit. We are therefore going to use Cauchy-Hadamard's theorem, which states that the radius of convergence of a power series is given by

$$R = \frac{1}{\limsup_{k\to\infty} |a_k|^{1/k}} \ ,$$

where a_k are the coefficients of the series. The theorem also states that if the limit in the denominator is ∞ then $R = 0$, and if it is 0 then $R = \infty$. Thus, to prove that the series converges on the whole plane we need to see that the limit is 0.

In our case $a_k \equiv a_{k,i,j} = \frac{(A^k)_{i,j}}{\Gamma(\alpha k+1)}$, which makes the radius of convergence depend on i and j. We then would have to choose the radius of convergence of the matricial series R to be the smallest radius of convergence of the individual elements: $\min_{i,j}\left\{R_{i,j}\right\}$. However, we can see by induction that for all $i, j = 1, \ldots, n$ we have that

$$\left|\left(A^k\right)_{i,j}\right| \leq n^{k-1} \max_{a,b}\{|A_{a,b}|\}^k \ . \tag{2.3}$$

Trivially, for $k = 1$, $\left|A_{i,j}\right| \leq \max_{a,b}\{|A_{a,b}|\}$ for all i, j. Assuming that the property holds for $k-1$:

$$\left|\left(A^k\right)_{i,j}\right| = \left|\sum_{l=1}^{n} A_{i,l}\left(A^{k-1}\right)_{l,j}\right| \leq \sum_{l=1}^{n} \left|A_{i,l}\right| \left|\left(A^{k-1}\right)_{l,j}\right| \leq \sum_{l=1}^{n} \max_{a,b}\left\{\left|A_{a,b}\right|\right\} n^{k-2} \max_{a,b}\left\{\left|A_{a,b}\right|\right\}^{k-1}$$
$$\leq n^{k-2} \max_{a,b}\left\{\left|A_{a,b}\right|\right\}^k \sum_{l=1}^{n} 1 \quad = \quad n^{k-1} \max_{a,b}\left\{\left|A_{a,b}\right|\right\}^k \ .$$

Thus we can provide an upper bound of $\left|\left(A^k\right)_{i,j}\right|$ which will provide us with a lower bound to $R_{i,j}$ for all i, j. If this bound makes the limit of $\left|a_{k,i,j}\right|^{1/k}$ tend to 0 we will have found R. Observe

$$\limsup_{k\to\infty} \left|a_{k,i,j}\right|^{1/k} = \limsup_{k\to\infty} \left|\frac{\left(A^k\right)_{i,j}}{\Gamma(\alpha k+1)}\right|^{1/k} \leq \limsup_{k\to\infty} \frac{\left|n^{k-1} \max_{a,b}\left\{\left|A_{a,b}\right|\right\}^k\right|^{1/k}}{\Gamma(\alpha k+1)^{1/k}} \ .$$

Note that as $\alpha > 0$, we have that $\alpha k + 1 > 0$ for all $k \in \mathbb{N}_0$, which allows us to apply Stirling's approximation as seen in Proposition 1.4, eq (1.7), giving us

$$\begin{aligned}\limsup_{k\to\infty} \left|a_{k,i,j}\right|^{1/k} &\leq \limsup_{k\to\infty} \frac{n^{\frac{k-1}{k}} \max_{a,b}\left\{|A_{a,b}|\right\}}{\left(\sqrt{2\pi\alpha k}\left(\frac{\alpha k}{e}\right)^{\alpha k}\right)^{1/k}} \\ &= n \max_{a,b}\left\{|A_{a,b}|\right\} \limsup_{k\to\infty}\left[(2\pi)^{-\frac{1}{2k}}(\alpha k)^{-\left(\alpha+\frac{1}{2k}\right)}e^{\alpha}\right] = 0 \ ,\end{aligned}$$

for all $i, j = 1, \ldots, n$.

Hence $R = \infty$ for all elements of the matrix and thus the series representation of the matricial Mittag-Leffler function is valid on the whole complex plane. Further, the series will converge absolutely and uniformly on any compact set contained in $\mathbb{C}$, and thus effectively on the whole plane. □

For a more complete practical consideration of the Mittag-Leffler function see [4, Appendix E].

2.2 The *Riemann-Liouville* fractional order integral

Now we are going to turn our attention to functions defined over time, that is, for variables taking values in $\mathbb{R}_+ := \{t \in \mathbb{R} : t > 0\}$. To keep consistency with the physical application of this work we will call the independent variable t.

We now introduce the concept of fractional integral as a generalisation of Cauchy's formula of repeated integration (this is known as the *Riemmann-Liouville fractional integral*), along some of its properties.

Definition (Riemann-Liouville Fractional Integral of order $\alpha > 0$)**.** Given $\alpha > 0$, the *R-L* fractional integral of a function $f : \mathbb{R}_+ \to \mathbb{R}$ is given by

$$I_t^\alpha f(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-\tau)^{\alpha-1} f(\tau) \mathrm{d}\tau = \left(h_\alpha * f\right)(t) \ , \qquad t > 0 \ , \tag{2.4}$$

where $*$ denotes the convolution product and $h_\alpha(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}$ for $t \in \mathbb{R}_+$ is the *Gel'fand-Shilov* function.

For completion, we also define I_t^0 as the identity operator to extend the definition to allow $\alpha \geq 0$. This is consistent with (2.4) as $h_0(t) = \delta(t)$, the Dirac delta distribution, which is a particular case of [4, eq 1.31] proven in reference [5].

Note that this definition requires f to be locally integrable in $\mathbb{R}_+$. We will denote the set of locally integrable functions over $\mathbb{R}_+$ as $\mathscr{L}^1_{\text{loc}}(\mathbb{R}_+)$

Proposition 2.2. *The R-L fractional integral definition is consistent with the usual integral in the sense of the addition semi-group property, which is, given $\alpha, \beta \geq 0$,*

$$I_t^\alpha \circ I_t^\beta = I_t^{\alpha+\beta} \ , \qquad \forall t > 0 \ .$$

Proof. It will be easier to prove the result using the convolution product definition.

Let $f \in \mathscr{L}^1_{\text{loc}}(\mathbb{R}_+)$. Then note that, expressed as a convolution product, the composition applied to f is as follows,

$$\left(I_t^\alpha \circ\ I_t^\beta\right) f(t) = \left(I_t^\alpha \circ \left(\ I_t^\beta \circ f\right)\right)(t) = \left(h_\alpha * \left(h_\beta * f\right)\right)(t) \ , \qquad t > 0 \ .$$

The convolution product is associative, so $\left(h_\alpha * \left(h_\beta * f\right)\right)(t) = \left(\left(h_\alpha * h_\beta\right) * f\right)(t)$ and thus we only need to prove that $h_\alpha * h_\beta = h_{\alpha+\beta}$ due to the local integrability of f. Note

$$\begin{aligned}(h_\alpha * h_\beta)(t) &= \int_0^t \frac{(t-\tau)^{\alpha-1}}{\Gamma(\alpha)} \frac{\tau^{\beta-1}}{\Gamma(\beta)} \mathrm{d}\tau = \frac{t^{\alpha+\beta-2}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^t \left(1-\frac{\tau}{t}\right)^{\alpha-1} \left(\frac{\tau}{t}\right)^{\beta-1} \mathrm{d}\tau \\ &= \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 (1-y)^{\alpha-1} (y)^{\beta-1} \mathrm{d}y \overset{\text{Proposition 1.5}}{=} \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \\ &= \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)} = h_{\alpha+\beta}(t) \ , \qquad t > 0 \ .\end{aligned}$$

where we have used the change of variable $y = \tau/t$ and the Eulerian integral of first kind. Thus this result is valid for any $\alpha, \beta > 0$. □

Remark. Note that the additive semigroup property $I_t^\alpha \circ I_t^\beta = I_t^{\alpha+\beta}$ implies that the fractional integral operator is commutative $I_t^\alpha \circ I_t^\beta = I_t^\beta \circ I_t^\alpha$ on "good enough" functions (both the function and its fractional order primitive have to be locally integrable), as we described in the proof. This is also justified by the commutativity of the convolution product.

2.3 Fractional derivatives

Now that we have seen how we can define a fractional integral, let's introduce the concept of a fractional derivative. As we mentioned at the start of the chapter, the usual integral I is the right inverse of the derivative D, as $D \circ I = \text{id}$ but $I \circ D$ gives us a collection of functions as solution, as well as requiring different properties of the function we apply the operators to.

Following this train of thought the *Riemann-Liouville* fractional derivative of order $\alpha > 0$ is introduced as ${}_{RL}D_t^\alpha = D_t^m \circ I_t^{m-\alpha}$ where $m \in \mathbb{N}$ such that $m - 1 < \alpha \leq m$ and where D_t^m denotes the m-th time derivative operator (for $m = 1$ we will simply write D_t). Alternatively we can swap the derivative and the fractional integral to define the *Caputo* fractional derivative as ${}_{C}D_t^\alpha = I_t^{m-\alpha} \circ D_t^m$, giving us a considerably different operator with different properties. For a short comparison of the two see [4, section 1.2]. In this work we will consider exclusively the *Caputo* derivative, so we will denote it by D_t^α.

2.3.1 The *Caputo* fractional order derivative

Definition (The Caputo fractional derivative of order $\alpha > 0$)**.** Given $\alpha > 0$ with $m \in \mathbb{N}$ such that $m - 1 < \alpha \leq m$, the *Caputo* fractional derivative is defined as

$$D_t^\alpha f(t) = I_t^{m-\alpha} f^{(m)}(t) \ , \qquad t > 0 \ .$$

Notice that this definition requires the m-th derivative of f to be locally integrable. Then $D_t^\alpha f(t)$ will be given by

$$D_t^\alpha f(t) = \begin{cases} \dfrac{1}{\Gamma(m-\alpha)} \displaystyle\int_0^t \frac{f^{(m)}(\tau)}{(t-\tau)^{\alpha+1-m}} \mathrm{d}\tau \ , & \text{if } m-1 < \alpha < m \ , \\ \dfrac{\mathrm{d}^m f}{\mathrm{d}t^m}(t) \ , & \text{if } \alpha = m \ , \end{cases} \qquad \text{for } t > 0 \ .$$

To avoid making the language too dense we will shorten *Caputo fractional derivative* to simply C-derivative. Also, for the remainder of the document we will work with values of $0 < \alpha < 1$, as those are the ones we will consider when dealing with the SI model.

C-derivative of polynomials

Let us see now how the Caputo derivative acts on polynomials. Firstly note that

$$D_t^\alpha (f+g)(t) = I_t^{1-\alpha} \circ D_t (f+g)(t) = I_t^{1-\alpha} \left(f' + g'\right)(t) = D_t^\alpha f(t) + D_t^\alpha g(t) \ , \qquad t > 0 \ ,$$

and that, for $K \in \mathbb{R}$

$$D_t^\alpha (Kf)(t) = I_t^{1-\alpha} \circ D_t (Kf)(t) = K I_t^{1-\alpha} \circ D_t (f)(t) = K D_t^\alpha f(t) \ , \qquad t > 0 \ .$$

So it's enough to study the derivative's effect on monomials. To generalise the result we will consider any power $\gamma \in [0, \infty)$. Further, as we justified above, we are only interested in $0 < \alpha < 1$.

Proposition 2.3. *For $0 < \alpha < 1$ and $\gamma > 0$,*

$$D_t^\alpha \left(\frac{t^\gamma}{\Gamma(\gamma+1)} \right) = \frac{t^{\gamma-\alpha}}{\Gamma(\gamma-\alpha+1)} \ , \qquad t > 0 \ .$$

In the case of $\gamma = 0$ then $D_t^\alpha(t^0) = 0$ for any α.

Proof. Let's see first the case of $\gamma = 0$. Trivially $D_t^\alpha(t^0) = I_t^{1-\alpha} \circ D_t(1) = I_t^{1-\alpha}(0) = 0$.

Take now $\gamma > 0$. Then

$$\begin{aligned} D_t^\alpha \left(\frac{t^\gamma}{\Gamma(\gamma+1)} \right) = I_t^{1-\alpha} \left(\frac{\gamma t^{\gamma-1}}{\Gamma(\gamma+1)} \right) &= \frac{\gamma}{\Gamma(\gamma+1)\Gamma(1-\alpha)} \int_0^t (t-\tau)^{-\alpha} \tau^{\gamma-1} \mathrm{d}\tau \\ &= \frac{\gamma t^{-\alpha+\gamma-1}}{\Gamma(\gamma+1)\Gamma(1-\alpha)} \int_0^t \left(1 - \frac{\tau}{t}\right)^{-\alpha} \left(\frac{\tau}{t}\right)^{\gamma-1} \mathrm{d}\tau \\ &= \frac{\gamma t^{-\alpha+\gamma}}{\Gamma(\gamma+1)\Gamma(1-\alpha)} \int_0^1 (1-y)^{(1-\alpha)-1} y^{\gamma-1} \mathrm{d}y \\ &= \frac{\gamma t^{-\alpha+\gamma}}{\Gamma(\gamma+1)\Gamma(1-\alpha)} \frac{\Gamma(1-\alpha)\Gamma(\gamma)}{\Gamma(1-\alpha+\gamma)} = \frac{t^{\gamma-\alpha}}{\Gamma(\gamma-\alpha+1)} \ , \qquad t > 0 \ . \end{aligned}$$

Which is well defined as $\gamma - \alpha + 1 > \gamma > 0$. □

Remark. With a simple check one can see that this result is consistent with the whole derivative of a monomial.

A note on sign conservation

We state here this important proposition as we will be using it frequently during the next chapter.

Proposition 2.4. *Let $u : (0, \infty) \to \mathbb{R}$ be a derivable function. Then*

$$D_t^\alpha u(t) \geq 0 \ , \quad \forall t > 0 \text{ for any one } 0 < \alpha < 1 \quad \Longrightarrow \quad u(t) \geq u(0) \ , \quad \forall t > 0 \ .$$

Proof. Before we begin the proof remark that the convolution product of two positive functions is necessarily positive, as all contributions to the integral are. Also, note that $D_t^\alpha u(t) \geq 0$ can be written as $\left(h_{1-\alpha} * u'\right)(t) \geq 0$. With this in mind, we have

$$u(t) - u(0) = \int_0^t u'(\tau)\mathrm{d}\tau = \left(h_1 * u'\right)(t) \overset{\star}{=} \left(h_\alpha * h_{1-\alpha} * u'\right)(t) \geq 0 \ , \qquad t > 0 \ ,$$

where $\star$ is due to the property $(h_\beta * h_\gamma)(t) = h_{\beta+\gamma}(t)$ we saw in the proof of Proposition 2.2. □

Note that the above does not prove that if the C-derivative of a function u is positive then u is increasing, but rather that it is positive if it began at $u(0) = 0$. It is an important distinction, as one could think that simply shifting the time origin and partitioning the time interval $(0, t]$ into many other sufficiently small ones ($\{(0, t_1], (t_1, t_2], \ldots, (t_{n-1}, t]\}$ for $0 < t_1 < \cdots < t_{n-1} < t$) we could prove monotonicity. This would be wrong, as the C-derivative, in particular due to the RL-fractional integral, is not a local operator, but one that takes into account the whole interval $(0, t]$.

Chapter 3

The fractional order SI model

Following the work in the article by Abadias et Al (reference [1]) we want to model the effect of an inhibitor on a network of amino acids of a protein, which we will call a protein residue network, or PRN in short. We know that said networks share information through a process called allostery, which can communicate perturbations of the order of 1Å up to a distance of 100Å, and it has been seen that these diffusion processes can be faithfully modelled after epidemiological contagion. Hence, to study our PRNs we will use an adaptation of the Susceptible-Infected (SI) model which will include the fractional order C-derivative we introduced in the previous chapter. We will see the advantages of this inclusion in section 3.3. This choice of model is justified further in the introduction of [1] and the references therein.

3.1 Classic SI model

To implement the model we will consider n amino acids, each of which can either be susceptible to being perturbed or have already been perturbed by the inhibitor. These two values will be represented by probabilities, which we will call $s_i(t)$ and $x_i(t)$ respectively. Thus $s_i(t), x_i(t) \in [0,1]$ and $s_i(t) + x_i(t) = 1$. An amino acid i will become perturbed at rate $\beta > 0$ upon contact with an already perturbed amino acid j to which it is connected. This dynamic is described by the following equations

$$\begin{cases} D_t s_i(t) = -\beta s_i(t) x_j(t) \ , \quad t > 0 \ , \\ D_t x_i(t) = \ \ \beta s_i(t) x_j(t) \ \ , \quad t > 0 \ , \end{cases} \xRightarrow{s_i(t)+x_i(t)=1} \qquad D_t x_i(t) = \beta \left(1 - x_i(t)\right) x_j(t) \ , \quad t > 0 \ . \tag{3.1}$$

The PRN will be represented by an adjacency matrix A, which will be symmetric and whose entries will be ones and zeroes to represent presence and lack of interaction between nodes. The nodes represent the α-carbon of the amino acids of the modelled protein, and they will be considered as interacting if the distance between them is smaller than a cut-off radius (in [1] it is taken as 7 Å). It is important to remark that the matrix's diagonal entries will all be null, as we will assume that an amino acid cannot influence itself. Thus we can conclude that all eigenvalues of A will be real and their sum will be 0. By construction A will not be the 0 matrix, so there will at least be one non-zero positive eigenvalue; making A have a positive spectral radius. The creation of this matrix is more thoroughly detailed in [1, Sec. 2.1]. This allows to write the above formula as

$$D_t x_i(t) = \beta \big(1 - x_i(t)\big) \sum_{j=1}^{n} A_{ij} x_j(t) \ , \qquad t > 0 \ . \tag{3.2}$$

As we are going to be considering initial value problems (*IVP* from now on) we need to consider initial conditions. We will choose all interactions to start at time $t_0 = 0$, and write x_0 to refer to the state of the system at t_0. That is, $x_0 = x(0) \in [0,1]^n$. This, in addition to writing (3.2) in vector form, gives us our complete problem.

$$\begin{cases} D_t x(t) = \beta \Big(\mathbb{I}_n - \operatorname{diag}\big(x(t)\big)\Big) A x(t) \ , \quad t > 0 \ . \\ \quad x(0) = x_0 \ . \end{cases} \tag{3.3}$$

This IVP has been proved to be sufficiently well behaved in that the solution x will not exit the $[0,1]^n$ set if it starts within it, which is consistent with the fact that it describes a collection of probabilities. It is also known that its two steady points are the 0s and 1s vector, representing lack of and full contagion respectively, and that any node i that begins with perturbation probability 1 will remain there. Also, the solution is monotonically increasing as long as it is non trivial, and will tend asymptotically to full contagion. The proofs for these statements are referenced in [1], but the reader can convince themselves of their veracity quite easily looking at equation (3.1).

3.1.1 Approximating the solution

IVP (3.3) is a non-linear system of differential equations without an (apparent) analytic solution, and thus we are looking to approximate it. To do so we return to the individual equations and rewrite them as

$$\frac{1}{1 - x_i(t)} D_t x_i(t) = \beta \sum_{j=1}^{n} A_{ij} x_j(t) \qquad \Longrightarrow \qquad D_t \Big(-\log\big(1 - x_i(t)\big)\Big) = \beta \sum_{j=1}^{n} A_{ij} x_j(t) \ , \quad t > 0 \ ,$$

where we have noticed that the left-hand side is the derivative of the minus logarithm of $1 - x_i(t)$. This allows for a change of variable $y_i(t) = g\big(x_i(t)\big) = -\log\big(1 - x_i(t)\big)$ for $t \in \mathbb{R}_+$, which makes the above

$$D_t y_i(t) = \beta \sum_{j=1}^{n} A_{ij} f(y_j(t)) \ , \qquad t > 0 \ ,$$

where $f(y) = 1 - e^{-y}$ is the inverse of g. Notice that this change of variable limits our choice of initial conditions to the set $[0,1)^n$, but as we have discussed $x_i(0) = 1$ is a stationary point anyways.

These can be written in vector form, giving us two equivalent statements

$$D_t g\big(x(t)\big) = \beta A x(t) \ , \qquad t > 0 \ , \tag{3.4}$$

$$D_t y(t) = \beta A f(y(t)) \ , \qquad t > 0 \ . \tag{3.5}$$

Remark. Here g and f have been extended to be functions from $[0,1)^n$ to $[0,\infty)^n$ and from $[0,\infty)^n$ to $[0,1)^n$ respectively, rather than their original one-dimensional definition. Thus $g(x) = -\log(\mathbf{1} - x)$, where the logarithm is applied element-wise, and $f(x) = \mathbf{1} - \exp(-y)$, where the exponential is also applied element-wise. $\mathbf{1}$ denotes the ones vector.

Firstly we are going to consider a linear approximation to the solution by taking the first order Taylor Polynomial of g centred at $x_c = \mathbf{0}$ on equation (3.4). This is, it will be an approximation of the solution for a wholly susceptible PRN with a small chance of proteins becoming perturbed. As $g(x) = -\log(\mathbf{1} - x)$ we have

$$T_g(x) = -\log(\mathbf{1} - x_c)\Big|_{x_c=\mathbf{0}} + \operatorname{diag}\left(\frac{1}{\mathbf{1} - x_c}\right)\Big|_{x_c=\mathbf{0}} x + O\big(x^2\big) \xrightarrow{x\to 0} x \ .$$

Again the inverse and square of a vector denotes the vector of operations taken element-wise.

We arrive at a linear approximation which, together with the initial condition, gives us the following IVP.

$$\begin{cases} D_t\tilde{x}(t) = \beta A\tilde{x}(t) \ , & t > 0 \ , \\ \tilde{x}(0) = x_0 \ , \end{cases} \tag{3.6}$$

which by Proposition 1.3 has solution

$$\tilde{x}(t) = e^{\beta t A} x_0 \ , \qquad t \geq 0 \ . \tag{3.7}$$

We know that this solution diverges for at least some x_0 as A has a positive spectral radius, which makes this approximation acceptable only in the earliest stages of the dynamics.

Let us see an improvement on this approximation, studying instead equation (3.5). We will work with function f's first order Taylor polynomial centred at y_c. Note that our new variable $y \in [0, \infty)^n$, and we will allow any y_c from said set. As $f'(y_i) = e^{-y_i}$ we have

$$T_f(y) = f(y_c) + Df(y_c)(y - y_c) + O(y^2) \approx \underbrace{\mathbf{1} - \exp(-y_c)}_{x_c} + \underbrace{e^{-\operatorname{diag}(y_c)}}_{\mathbb{I}_n - \operatorname{diag}(x_c)}(y - y_c) \ .$$

If we choose to centre the polynomial on $y_c = y_0 = g(x_0)$ we can approximate (3.5) to

$$D_t\hat{y}(t) = \beta A x_0 + \beta\hat{A}\left(\hat{y}(t) - y_0\right) = \beta\hat{A}\hat{y} + \beta A b_{x_0} \ , \qquad t > 0 \ ,$$

where $\hat{A} = A\left(\mathbb{I}_n - \operatorname{diag}(x_0)\right)$ and $b_{x_0} = x_0 + \left(\mathbb{I}_n - \operatorname{diag}(x_0)\right)\log(\mathbf{1} - x_0)$.

This approximation, therefore, gives us the following IVP

$$\begin{cases} D_t\hat{y}(t) = \beta\hat{A}\hat{y} + \beta A b_{x_0} \ , & t > 0 \ , \\ \hat{y}(0) = y_0 = g(x_0) \ . \end{cases} \tag{3.8}$$

Proposition 3.1. *The solution to* (3.8) *is given by*

$$\hat{y}(t) = e^{\beta t\hat{A}} g(x_0) + \sum_{k=1}^{\infty} \frac{(\beta t)^k}{k!}\hat{A}^{k-1} A b_{x_0} \ , \qquad t \geq 0 \ . \tag{3.9}$$

Proof. Equation (3.8) is a first order linear differential system of equations and thus will have a unique analytical solution. The expresion of $\hat{y}(t)$ can be deduced using the variation of constants method for ODEs, which is a fairly well known result and thus we won't get into it here. □

The approximation to the solution to (3.3) will be easily recovered as $\hat{x} = f(\hat{y})$. We can see that $\hat{y}$ is positive for all positive values of t, and that it diverges, at least for some values of x_0, when t tends to infinity, which implies that $\hat{x}$ remains within the $[0, 1)^n$ set and tends to $\mathbf{1}$ when t goes to infinity. This is an important improvement on (3.7).

3.2 Inclusion of the fractional derivative

Once we have seen the results of the classical model we want to expand them using fractional derivatives. One could think that such extension of (3.1) would simply give us

$$D_t^\alpha x_i(t) \overset{*}{=} \beta^\alpha \big(1 - x_i(t)\big)\, x_j(t) \ , \qquad t > 0 \ ,$$

where the α exponent of β has been added to keep the equation dimensionally correct. This equation, however, is rather uninteresting analytically as there isn't too much we can do with it. Instead, we will work with the following adaptation:

$$\int_0^t h_{1-\alpha}(t-\tau)\frac{x_i'(\tau)}{1 - x_i(\tau)}\mathrm{d}\tau = \beta^\alpha x_j(t) \ , \qquad t > 0 \ . \tag{3.10}$$

This, together with the adjacency matrix and in vector form gives us the Fractional order SI model.

$$D_t^\alpha \Big(-\log\big(\mathbf{1} - x(t)\big)\Big) = D_t^\alpha g\left(x(t)\right) = \beta^\alpha A x(t) \ , \qquad t > 0 \ , \tag{3.11}$$

where we have used $g(x) = -\log(\mathbf{1} - x)$ as defined in the previous section. Alternatively, calling $y = g(x)$ and having f with $f(y) = \mathbf{1} - \exp(-y) = x$ be the inverse of g again we can rewrite the equation (3.11) as

$$D_t^\alpha y(t) = \beta^\alpha A f(y(t)) \ , \qquad t > 0 \ , \tag{3.12}$$

which in reality is nothing but the fractional order derivative equivalent of equation (3.5).

We will take the same initial conditions as with the classic case, $t_0 = 0$ and $x(0) = x_0 \in [0,1]^n$, which will allow us to write the fractional order IVP that describes the dynamic in our PRN as

$$\begin{cases} D_t^\alpha g\left(x(t)\right) = \beta^\alpha A x(t) \ , \qquad t > 0 \ , \\ \qquad x(0) = x_0 \ . \end{cases} \tag{3.13}$$

We will also call $y(0) = g(x_0) = y_0 \in [0, \infty)^n$ as initial conditions when working with variable y.

Behaviour of the solution

Let us invest some time to talk about the behaviour of the solution $x(t)$ to (3.13).

Firstly notice that the right-hand side of (3.11) will be non-negative as long as all components of $x(t)$ are non-negative, which will be the case as we are looking for solutions representing probabilities as we have discussed. Thus, by Proposition 2.4, $-\log\big(\mathbf{1} - x(t)\big) \succeq 0$, giving us that $x(t) \succeq x(0)$ for all $t \geq 0$. Thus there exists $\varepsilon > 0$ for which $x_i(s)$ is non-decreasing for $0 < s < \varepsilon$. Now assume that $x_i(0) = 1$ for some $i \in \{1, \dots, n\}$. If x_i increases then $\frac{x_i'(s)}{1-x_i(s)} < 0$ for said $s > 0$. By (3.10) we have that $\frac{x_i'(\tau)}{1-x_i(\tau)} \geq 0$ for $s > 0$ sufficiently close to 0, so we conclude that $x_i'(t) = 0$, which will hold for all $t > 0$ and hence $x_i^* = 1$ is an equilibrium point for $x_i(t)$. By the same argument with $s_i(t)$ (see (3.1)) $x_i^* = 0$ is also an equilibrium point. It can be seen that x_i won't have 0 derivative for any other value in $(0,1)$, so 1 and 0 are the only equilibrium points.

With this, we can easily justify that for $x_0 \in (0,1)^n$ then $x(t) \in (0,1)^n$ for all $t > 0$: the solution remains bounded and consistent with probabilities. Similarly, as $x_i(t) \geq x_i(0)$ for all $t > 0$, if $x_i(0) \in (0,1)$ then $x_i(t)$ will necessarily tend to 1 as it's the only accessible equilibrium point. This makes $x(t) \to \mathbf{1}$ asymptotically as $t \to \infty$.

3.2.1 Linear approximation

As with the procedure we followed to approximate the solution in the classic case, we will first provide a linear approximation to the solution of the fractional order SI model working with equation (3.11). By the first order Taylor polynomial approximation of g we obtain the equation

$$D_t^\alpha \tilde{x}(t) = \beta^\alpha A \tilde{x}(t) \ , \qquad t > 0 \ .$$

Proposition 3.2. *For $0 < \alpha < 1$, the initial value problem given by the equation above with the same initial conditions as the original problem*

$$\begin{cases} D_t^\alpha \tilde{x}(t) = \beta^\alpha A \tilde{x}(t) \ , & t > 0 \ , \\ \tilde{x}(0) = x_0 \ , \end{cases} \tag{3.14}$$

has solution

$$\tilde{x}(t) = E_\alpha\left((\beta t)^\alpha A\right) x_0 = \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \ x_0 \ , \qquad t \geq 0 \ . \tag{3.15}$$

Proof. Firstly note that for any $\alpha > 0$, $E_\alpha\left((\beta t)^\alpha A\right)\big|_{t=0} = \mathbb{I}_n$ and thus $\tilde{x}(0) = x_0$, so the initial condition is satisfied.

Recall that by Proposition 2.1 the series for $E_\alpha\left((\beta t)^\alpha A\right)$ converges absolutely and uniformly on the whole complex plane, and thus in particular for all $t \geq 0$. Now let us take the C-derivative of (3.15) to see that it is in fact a solution to (3.14):

$$D_t^\alpha \tilde{x}(t) = D_t^\alpha \left(\sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \ x_0 \right) = I_t^{1-\alpha} \circ D_t \left(\sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \right) x_0 \ , \qquad t > 0 \ . \tag{3.16}$$

Let us consider each element of the matrix individually, as both the derivative and integral of a matrix are taken element-wise. That is

$$e_i^T \left(\sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \right) e_j = \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} \left(e_i^T A^k e_j\right)}{\Gamma(\alpha k + 1)} = \sum_{k=0}^{\infty} \frac{\beta^{\alpha k} \left(e_i^T A^k e_j\right)}{\Gamma(\alpha k + 1)} (t^\alpha)^k \equiv \sum_{k=0}^{\infty} b_{ijk} s^k \ , \qquad t \geq 0 \ ,$$

which is an absolutely and uniformly convergent scalar series on $s = t^\alpha \in [0, \infty) \subset \mathbb{C}$ for all $t \geq 0$. If we take its derivative by Proposition 1.1 we can swap it with the summation, giving us another convergent scalar series for all $t \geq 0$. In particular, this new series will be analytic and thus continuous. Therefore we know that the R-L integral of order $1 - \alpha$ of this series will converge as it is the convolution product of two continuous functions. That is,

$$I_t^{1-\alpha} \sum_{k=0}^{\infty} D_t \left(b_{ijk} (t^\alpha)^k \right) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \left(\sum_{k=1}^{\infty} b_{ijk} \alpha k \tau^{\alpha k - 1} (t - \tau)^{-\alpha} \right) \mathrm{d}\tau \ , \qquad t > 0 \ .$$

By Fubini's theorem, to prove that we can swap the integral with the series it is enough to see that

$$\sum_{k=1}^{\infty} \frac{\left|b_{ijk} \alpha k\right|}{\Gamma(1-\alpha)} \left(\int_0^t \tau^{\alpha k - 1} (t - \tau)^{-\alpha} \mathrm{d}\tau \right) < \infty \ , \qquad \text{for all } t > 0 \ . \tag{3.17}$$

Let's work step by step. Firstly note that on account of the proof of Proposition 2.2 we have that the integral above is equal to $\Gamma(\alpha k)\Gamma(1-\alpha)h_{\alpha(k-1)+1}(t)$. This makes the series of (3.17) become

$$\sum_{k=1}^{\infty}\left|b_{ijk}\alpha k\right|\Gamma(\alpha k)h_{\alpha(k-1)+1}(t) = \sum_{k=1}^{\infty}\left|b_{ijk}\right|\frac{\Gamma(\alpha k+1)}{\Gamma(\alpha(k-1)+1)}t^{\alpha(k-1)} =$$
$$= \sum_{k=0}^{\infty}\left|b_{ij(k+1)}\right|\frac{\Gamma(\alpha(k+1)+1)}{\Gamma(\alpha k+1)}t^{\alpha k}\ , \qquad t>0\ .$$

Let's study now the coefficients to find an upper bound to prove convergence. Notice that, calling $\Lambda = \max_{a,b}\left\{\left|A_{a,b}\right|\right\}$ we have

$$\left|e_i^T A^{k+1}e_j\right| \leq n\Lambda\left|e_i^T A^k e_j\right|\ ,$$

thanks to which we can write our bound:

$$\left|b_{ij(k+1)}\right|\frac{\Gamma(\alpha(k+1)+1)}{\Gamma(\alpha k+1)} = \frac{\beta^{\alpha(k+1)}\left|e_i^T A^{k+1}e_j\right|}{\Gamma(\alpha k+1)} \leq \beta n\Lambda\frac{\beta^{\alpha k}\left|e_i^T A^k e_j\right|}{\Gamma(\alpha k+1)} = \beta n\Lambda\left|b_{ijk}\right|\ .$$

Implementing this into our original summation

$$\sum_{k=0}^{\infty}\left|b_{ij(k+1)}\right|\frac{\Gamma(\alpha(k+1)+1)}{\Gamma(\alpha k+1)}t^{\alpha k} \leq \sum_{k=0}^{\infty}\beta n\Lambda\left|b_{ijk}\right|t^{\alpha k} = \beta n\Lambda\sum_{k=0}^{\infty}\left|b_{ijk}\right|t^{\alpha k} < \infty\ , \qquad \text{for all } t>0\ .$$

This proves the convergence of the series of (3.17) and thus we can apply Fubini's theorem individually on every matrix element of (3.16).

In the light of this result we can write the following,

$$D_t^\alpha\tilde{x}(t) = D_t^\alpha\left(\sum_{k=0}^{\infty}\frac{(\beta t)^{\alpha k}A^k}{\Gamma(\alpha k+1)}\right)x_0 = \sum_{k=0}^{\infty}D_t^\alpha\left(\frac{(\beta t)^{\alpha k}A^k}{\Gamma(\alpha k+1)}\right)x_0 =$$
$$= \sum_{k=0}^{\infty}\beta^{\alpha k}A^k D_t^\alpha\left(\frac{t^{\alpha k}}{\Gamma(\alpha k+1)}\right)x_0\ , \qquad t>0\ .$$

By Proposition 2.3, that is equal to

$$\sum_{k=1}^{\infty}\beta^{\alpha k}A^k\frac{t^{\alpha(k-1)}}{\Gamma\left(\alpha(k-1)+1\right)}\ x_0 = \beta^\alpha A\sum_{k=0}^{\infty}\frac{(\beta t)^{\alpha k}A^k}{\Gamma\left(\alpha k+1\right)}\ x_0 = \beta^\alpha A E_\alpha\left((\beta t)^\alpha A\right)x_0 = \beta^\alpha A\tilde{x}(t)\ ,$$

which holds for every $t>0$ and is what we set out to prove. □

Proposition 3.3. *$\tilde{x}(t)$ diverges for some $x_0 \in [0,1]^n$.*

Proof. When we introduced the matrix A that describes the PRN we deduced that its spectral radius was positive. Let us call λ_1 that eigevalue and v_1 the eigenvector corresponding to λ_1. Then we can consider the contribution of x_0 to v_1 as the standard scalar product $\langle v_1, x_0\rangle = \sum_{j=1}^n v_{1,j}x_{0,j}$ and thus

$$E_\alpha\left((\beta t)^\alpha A\right)\langle v_1,x_0\rangle v_1 = \sum_{k=0}^{\infty}\frac{(\beta t)^{\alpha k}A^k}{\Gamma(\alpha k+1)}\langle v_1,x_0\rangle v_1 = \sum_{k=0}^{\infty}\frac{(\beta t)^{\alpha k}\lambda_1^k}{\Gamma(\alpha k+1)}\langle v_1,x_0\rangle v_1\ , \qquad t\geq 0\ .$$

We know that v_1 cannot be the trivial vector, and it will have at least one positive component (if all are non-positive $-v_1$ will also be an eigenvector of eigenvalue λ_1). With v_1 we can also build initial conditions $x_0 \in [0,1]^n$ such that $\langle v_1, x_0\rangle > 0$: taking x_0 as v_1 with any negative components made 0 (we could write $x_{0i} = \max\{v_{1i}, 0\}$). If under these conditions we take the limit of the above expression when t tends to ∞ we get

$$\lim_{t\to\infty} \tilde{x}_i(t) = \lim_{t\to\infty} \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k}\lambda_1^k}{\Gamma(\alpha k+1)} \langle v_1, x_0\rangle v_{1,i} = \infty \ ,$$

and thus the solution diverges. □

Proposition 3.4. *Let $B \in [0,\infty)^{n\times n}$ be a non-trivial symmetric matrix with all elements in its diagonal 0. Then $w(t) = E_\alpha\left((\beta t)^\alpha B\right) w_0$ diverges when t tends to ∞ for every $w_0 \in [0,\infty)^n \setminus \mathbf{0}$.*

Proof. As B is symmetric all its eigenvalues $\lambda_1, \ldots, \lambda_n$ are real, and given that its trace is 0, $\sum \lambda_i = 0$. Now, they cannot be all 0 as B is not the zero matrix, so there must be at least one positive and at least one negative eigenvalue.

Let λ be a negative eigenvalue of B and v its corresponding eigenvector. Then we can see that $E_\alpha(tB)\, v$ tends to $\mathbf{0}$ when t tends to ∞:

$$E_\alpha(tB)\, v = \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} B^k}{\Gamma(\alpha k+1)} v = \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k}\lambda^k}{\Gamma(\alpha k+1)} v \xrightarrow{t\to\infty} 0v = \mathbf{0} \ .$$

Let now $0 > \lambda_p \leq \cdots \leq \lambda_n$ be all the negative eigenvalues of B, and $v_p, \ldots, v_n$ their corresponding eigenvectors, with $2 \leq p \leq n$. Then any non-trivial linear combination of them $v = \sum_{i=p}^n \alpha_i v_i$ with $\alpha_i \in \mathbb{R}$ must have at least one positive component and one negative component, that is, neither v nor $-v$ are in $[0,\infty)^n \setminus \mathbf{0}$.

Let us see so by assuming without loss of generality that $v \in [0,\infty)^n \setminus \mathbf{0}$ and reaching a contradiction. Then the k-th component of Bv is $(Bv)_k = \sum_{j=1}^n B_{kj} v_j \geq 0$ as all components of B are positive. This means that

$$\langle Bv, v\rangle = \sum_{j=1}^{n} (Bv)_j (v)_j \geq 0 \ .$$

However, as B is symmetric its eigenvectors are orthogonal, which means that

$$\langle Bv, v\rangle = \left\langle B\sum_{i=p}^{n} \alpha_i v_i, \sum_{l=p}^{n} \alpha_l v_l \right\rangle = \left\langle \sum_{i=p}^{n} \alpha_i \lambda_i v_i, \sum_{l=p}^{n} \alpha_l v_l \right\rangle = \sum_{i=p}^{n}\sum_{l=p}^{n} \alpha_i \lambda_i \alpha_l \langle v_i, v_l\rangle = \sum_{i=p}^{n} \lambda_i \alpha_i^2 < 0 \ ,$$

which cannot be 0 as all $\lambda_i < 0$ and $\sum_{i=p}^n \alpha_i^2 > 0$ or v would be the 0 vector. This clashes with our previous result and we arrive at a contradiction.

We conclude then that any non-trivial linear combination of negative-eigenvector eigenvalues will fall outside $[0,\infty)^{n\times n}$, implying that any $w_0 \in [0,\infty)^n \setminus \mathbf{0}$ must have a contribution from a positive-eigenvalue eigenvector, which will make $w(t)$ diverge when t tends to infinity as we have seen in Proposition 3.3.

For completion, if $w_0 = 0$ then $w(t) = 0$ for all $t \geq 0$ and thus the solution will not diverge. □

Corollary 3.5. *$\tilde{x}(t)$ diverges for **all** $x_0 \in [0,1]^n \setminus \mathbf{0}$.*

Proof. A satisfies all conditions asked on B in the previous proposition, so we have that $\tilde{x}(t) = E_\alpha\left((\beta t)^\alpha A\right) x_0$ will diverge for all $x_0 \in [0,\infty)^n \setminus \mathbf{0}$, and in particular for all $x_0 \in [0,1]^n \setminus \mathbf{0}$. □

This implies that the solution to the linear approximation does not hold the property of consistency with probabilities ($x(t) \in [0,1]^n$ for any $x_0 \in [0,1]^n$), as we know it must. Thus, this approximation will only really work in the earliest stages of the dynamic of an almost-uninfected PRN (recall that the Taylor approximation of g was made around the 0 vector). We will provide a better approximation in the next subsection.

3.2.2 LTE approximation

Let us again provide a more adequate approximation. Following the example of the classic model, we will now work with (3.12) and use the first order Taylor polynomial of f. This transformation is referred to as Lee-Tenneti-Eun (LTE) in reference [1], due to the authors by whom it was first used.

Proposition 3.6. *For $0 < \alpha < 1$, the LTE transformation, together with the initial conditions $y_0 = g(x_0)$, where $x_0 \in [0,1)^n$, gives us the following initial value problem*

$$\begin{cases} D_t^\alpha \hat{y}(t) = \beta^\alpha \hat{A}\hat{y}(t) + \beta^\alpha A b_{x_0} \;, & t > 0 \;, \\ \quad \hat{y}(0) = y_0 = g(x_0) \;, & \end{cases} \tag{3.18}$$

where, if we call $\Omega = \left(\mathbb{I}_n - diag(x_0)\right)$, we define $\hat{A} = A\Omega$ and $b_{x_0} = x_0 + \Omega \log(\mathbf{1} - x_0)$ with the logarithm taken element-wise, similar to what we defined for the classic case.

Its solution is

$$\hat{y}(t) = E_\alpha\left((\beta t)^\alpha \hat{A}\right) g(x_0) + \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} b_{x_0} \;, \qquad t \geq 0 \;. \tag{3.19}$$

Proof. Let's first prove that $\hat{y}(t)$, in particular the second series term, converges. We will consider the matricial sum (before multiplying by b_{x_0}). We have

$$\sum_{k=0}^{\infty} \left| \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} \right| = \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \left|\hat{A}^k\right| |A|}{\Gamma\left(\alpha(k+1)+1\right)} \overset{A \in \{0,1\}^{n\times n}}{=} \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} \;, \qquad t \geq 0 \;,$$

where the absolute value denotes the operation element-wise.

As $\Omega = \left(\mathbb{I}_n - \text{diag}(x_0)\right)$ is a diagonal matrix whose entries are given by a number between 0 and 1, Ω^k will also be diagonal and all diagonal entries of Ω^k will be bound by 1. Thus we can write $\Omega^k \preceq \mathbb{I}_n$ (that is, the bound $\leq$ holds element-wise) and we can therefore bound the series by

$$\sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} \preceq \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} A^{k+1}}{\Gamma\left(\alpha(k+1)+1\right)} = \sum_{k=1}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma\left(\alpha k+1\right)} = E_\alpha\left((\beta t)^\alpha A\right) - \mathbb{I}_n \;, \qquad t \geq 0 \;,$$

which we know converges absolutely and uniformly for all $t \geq 0$.

Having seen this we check that the initial condition is satisfied. Note that the second term of (3.19) will be null for $t = 0$, so we only have the contribution of the first. Again, the Mittag-Leffler function at $t = 0$ corresponds with the identity matrix, and $t^\alpha\big|_{t=0}$ will be well defined as $\alpha > 0$. Hence $\hat{y}(0) = \mathbb{I}_n g(x_0) + 0 = g(x_0)$.

Let's see now that $\hat{y}(t)$ satisfies the differential equation of (3.18). We know that the C-derivative of a sum is the sum of C-derivatives (see the considerations prior to Proposition 2.3), and by the proof of Proposition 3.2 we know that $D_t^\alpha E_\alpha\left((\beta t)^\alpha \hat{A}\right) = \beta^\alpha \hat{A} E_\alpha\left((\beta t)^\alpha \hat{A}\right)$. Thus it is enough to compute the C-derivative of the second term of (3.19).

Knowing that the series converges absolutely and uniformly one can prove that the C-derivative can be swapped with the summation following a similar procedure to that of the proof of Proposition 3.2 . Hence

$$D_t^\alpha \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} b_{x_0} = \sum_{k=0}^{\infty} \beta^{\alpha(k+1)} D_t^\alpha \left(\frac{t^{\alpha(k+1)}}{\Gamma\left(\alpha(k+1)+1\right)}\right) \hat{A}^k A b_{x_0} \ , \qquad t > 0 \ ,$$

and by Proposition 2.3 once again the above is equal to

$$\sum_{k=0}^{\infty} \beta^{\alpha(k+1)} \left(\frac{t^{\alpha k}}{\Gamma\left(\alpha k+1\right)}\right) \hat{A}^k A b_{x_0} = \beta^\alpha A b_{x_0} + \beta \hat{A} \sum_{k=1}^{\infty} \left(\frac{(\beta t)^{\alpha k} \hat{A}^{k-1} A}{\Gamma\left(\alpha k+1\right)}\right) b_{x_0} \ , \qquad t > 0 \ ,$$

which finally gives us that

$$D_t^\alpha \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} b_{x_0} = \beta \hat{A} \sum_{k=0}^{\infty} \left(\frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)}\right) b_{x_0} + \beta^\alpha A b_{x_0} \ , \qquad t > 0 \ .$$

Therefore the C-derivative of the whole expression for $\hat{y}(t)$ in (3.19) is

$$D_t^\alpha \left(E_\alpha\left((\beta t)^\alpha \hat{A}\right) g(x_0) + \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} b_{x_0}\right) =$$

$$= \beta^\alpha \hat{A} \left(E_\alpha\left((\beta t)^\alpha \hat{A}\right) g(x_0) + \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} b_{x_0}\right) + \beta^\alpha A b_{x_0} \ , \qquad t > 0 \ .$$

From where we conclude that (3.19) satisfies (3.18) and thus it is its solution. □

The approximation to the solution of the original problem (3.11) can be recovered as $\hat{x} = f(\hat{y})$. We will study its behaviour in subsection 3.2.3.

Proposition 3.7. *If one considers a PRN where there is certainty that no protein has been perturbed (that is, where $x_0 \in [0,1)^n$) the solution $\hat{y}(t)$ to* (3.18) *can be rewritten in a more illustrative way:*

$$\hat{y}(t)\Bigg|_{x_0 \in [0,1)^n} = g(x_0) + \left[E_{\alpha,1}\left((\beta t)^\alpha \hat{A}\right) - \mathbb{I}_n\right] \Omega^{-1} x_0 \ , \qquad t \geq 0 \ . \tag{3.20}$$

Proof. Remember that $\Omega = \mathbb{I}_n - \mathrm{diag}(x_0)$, which means that Ω^{-1} will exist only for $x_0 \in [0,1)^n$ and will be given by a diagonal matrix with elements $\left(\Omega^{-1}\right)_{i,i} = (1 - x_{0i})^{-1}$. Under these conditions one can write (3.19) as

$$\begin{aligned}\hat{y}(t)\Big|_{x_0\in[0,1)^n} &= E_\alpha\left((\beta t)^\alpha \hat{A}\right) g(x_0) + \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)}\hat{A}^{k+1}}{\Gamma\left(\alpha(k+1)+1\right)}\Omega^{-1}b_{x_0} \\ &= E_\alpha\left((\beta t)^\alpha \hat{A}\right) g(x_0) + \sum_{k=1}^{\infty} \frac{(\beta t)^{\alpha k}\hat{A}^{k}}{\Gamma\left(\alpha k+1\right)}\Omega^{-1}b_{x_0} \\ &= E_\alpha\left((\beta t)^\alpha \hat{A}\right) \left[g(x_0) + \Omega^{-1}b_{x_0}\right] - \Omega^{-1}b_{x_0} \ , \qquad t \geq 0 \ .\end{aligned}$$

Taking into account $\Omega^{-1}b_{x_0} = \Omega^{-1}\left[x_0 + \Omega\log(\mathbf{1}-x_0)\right] = \Omega^{-1}x_0 - g(x_0)$ we arrive at equation (3.20). □

Note that both $g(x_0)$ and $\Omega^{-1}x_0$ are vectors positive in all components, in addition to $\hat{A}$ having a positive spectral radius. By the same arguments as Proposition 3.3 we can expect $\hat{y}(t)$ to diverge when t tends to ∞ for some values of x_0, meaning that $\hat{x}(t) = f\left(\hat{y}(t)\right)$ will tend to $\mathbf{1}$. This is exactly the dynamic we expected from the classic system, so this approximation represents a consistent result. We will see in the next section that $\hat{y}(t)$ will diverge for every $x_0 \in [0,1)^n$, as we did for $\tilde{x}(t)$, and that $\hat{x}(t)$ will represent the fastest-perturbation-spread scenario: $x(t) \preceq \hat{x}(t)$ for any $x_0 \in [0,1]^n$.

3.2.3 Comparison of the solutions

Theorem 3.8. *Let $0 < \alpha < 1$, and let us denote by $x(t)$ the solution to our original unapproximated initial value problem* (3.13) *and an initial condition $x_0 \in [0,1]^n$.*

Let also $\tilde{x}(t)$ and $\hat{x}(t) = f\left(\hat{y}(t)\right)$ be the solutions to the linear approximation (IVP 3.14) and the LTE transformation (IVP 3.18) respectively, all with the same initial conditions x_0.

Then

$$x(t) \preceq \hat{x}(t) \preceq \tilde{x}(t) \ , \qquad t \geq 0 \ . \tag{3.21}$$

Proof. Let us begin with the first inequality. Let g and f be given as they have been throughout the document (see between eq (3.11) and eq (3.12)), and define $u : [0,\infty) \to \mathbb{R}^n$ as $u(t) = g\left(\hat{x}(t)\right) - g\left(x(t)\right) = \hat{y}(t) - y(t)$. As both the IVPs have the same initial conditions, $u(0) = \mathbf{0}$. We want to see that $D_t^\alpha u(t) \succeq 0$ for all $t \in \mathbb{R}_+$, so that by Proposition 2.4 applied individually on each component we will have $\hat{y}(t) \succeq y(t)$ and therefore, as $f(y) = \mathbf{1} - \exp(-y)$ is a monotonously increasing function in every variable, that $\hat{x}(t) \succeq x(t)$ for all $t > 0$.

Thus we have to see whether $D_t^\alpha y(t) \preceq D_t^\alpha \hat{y}(t)$. By (3.12) and (3.18) we have that

$$D_t^\alpha y(t) = \beta^\alpha A f(y(t)) \ , \quad \text{and} \quad D_t^\alpha \hat{y}(t) = \beta^\alpha \hat{A}\hat{y}(t) + \beta^\alpha A b_{x_0} \ , \qquad \text{both for } t > 0 \ ,$$

respectively, where (3.18) is the LTE approximation of (3.12) which we got by taking the first order Taylor Polynomial of f centred at y_0. We can then simply study the hessian of f, which by the definition of f will either be definite positive, definite negative, or the zero matrix (recall that our vectorial f was the extension of a scalar f acting separately on every variable). The second derivative of scalar $f(t) = 1 - e^{-y}$ is $f''(y) = -e^{-y} < 0$ for all $y \in [0, \infty)$, making the hessian of our vectorial f definite negative, and thus f a concave function. This means that

$$\beta^\alpha A f\left(y(t)\right) \preceq \beta^\alpha \hat{A} y(t) + \beta^\alpha A b_{x_0} \ , \quad \text{for all } t \geq 0 \ ,$$

giving us the sought after inequality: $D_t^\alpha y(t) \preceq D_t^\alpha \hat{y}(t)$ and thus $x(t) \preceq \hat{x}(t)$ for any $t > 0$.

Let us see the second inequality now. Firstly, we know that

$$D_t^\alpha \hat{x}(t) = D_t^\alpha f(\hat{y}(t)) = \left(h_{1-\alpha} * e^{-\operatorname{diag}(\hat{y})} D_t \hat{y}\right)(t) \preceq \left(h_{1-\alpha} * D_t \hat{y}\right)(t) = D_t^\alpha \hat{y}(t) \ , \qquad t > 0 \ ,$$

where by $D_t \hat{y}$ we denote the component-wise whole derivative vector of $\hat{y}$, which we know by (3.19) exists and is non-negative for all $t \in \mathbb{R}_+$. Thus the inequality $e^{-\operatorname{diag}(\hat{y})} D_t \hat{y} \preceq D_t \hat{y}$.

Let us work then with $D_t^\alpha \hat{y}(t)$, which by (3.18) and (3.19) is

$$\begin{aligned}
D_t^\alpha \hat{y}(t) = \beta^\alpha \hat{A} \hat{y}(t) + \beta^\alpha A b_{x_0} &= \beta^\alpha \hat{A} \left[E_\alpha \left((\beta t)^\alpha \hat{A} \right) g(x_0) + \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^k A}{\Gamma\left(\alpha(k+1)+1\right)} b_{x_0} \right] + \beta^\alpha A b_{x_0} \\
&= \beta^\alpha \hat{A} E_\alpha \left((\beta t)^\alpha \hat{A} \right) g(x_0) + \beta^\alpha \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha(k+1)} \hat{A}^{k+1}}{\Gamma\left(\alpha(k+1)+1\right)} A b_{x_0} + \beta^\alpha A b_{x_0} \\
&= \beta^\alpha \hat{A} E_\alpha \left((\beta t)^\alpha \hat{A} \right) g(x_0) + \beta^\alpha \left[E_\alpha \left((\beta t)^\alpha \hat{A} \right) - \mathbb{I}_n \right] A b_{x_0} + \beta^\alpha A b_{x_0} \\
&= \beta^\alpha E_\alpha \left((\beta t)^\alpha \hat{A} \right) A \left[\Omega g(x_0) + b_{x_0} \right] = \beta^\alpha E_\alpha \left((\beta t)^\alpha \hat{A} \right) A x_0 \ , \qquad t > 0 \ .
\end{aligned}$$

On the other hand, by (3.14) and (3.15) we know that $D_t^\alpha \tilde{x}(t) = \beta^\alpha A E_\alpha \left((\beta t)^\alpha A \right) x_0$ for $t \geq 0$. Recall that we defined $\Omega = \operatorname{diag}(\mathbf{1} - x_0)$, so $0 \preceq \Omega \preceq \mathbb{I}_n$ and thus $0 \preceq \hat{A} \preceq A$ and $0 \preceq \hat{A}^k \preceq A^k$ for $k \in \mathbb{N}_0$. Hence

$$\frac{(\beta t)^{\alpha k} \hat{A}^k}{\Gamma(\alpha k + 1)} \preceq \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \quad \forall k \in \mathbb{N}_0 \qquad \Longrightarrow \qquad \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} \hat{A}^k}{\Gamma(\alpha k + 1)} \preceq \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \ , \qquad t \geq 0 \ .$$

and thus

$$E_\alpha \left((\beta t)^\alpha \hat{A} \right) \preceq E_\alpha \left((\beta t)^\alpha A \right) \ , \qquad t \geq 0 \ .$$

From where we conclude that $D_t^\alpha \hat{x}(t) \preceq D_t^\alpha \hat{y}(t) \preceq D_t^\alpha \tilde{x}(t)$ with $t > 0$. Once again by Proposition 2.4 component-wisely, taking $u(t) = \tilde{x}(t) - \hat{x}(t)$, we can write $\hat{x}(t) \preceq \tilde{x}(t)$, which completes the proof for $x(t) \preceq \hat{x}(t) \preceq \tilde{x}(t)$ for $t > 0$. The case for $t = 0$ holds trivially, hence giving us equation (3.21). □

Let us see now that $\hat{x}(t)$ will not tend to $\mathbf{0}$ for any $x_0 \in [0, 1]^n$.

Proposition 3.9. *$\hat{y}(t)$ diverges for all $x_0 \in [0, 1)^n \setminus \mathbf{0}$.*

Proof. Let us define $\widehat{A} = \Omega \hat{A} = \Omega A \Omega$. This makes $\widehat{A}$ a symmetric matrix in $[0, 1]^{n \times n} \subset [0, \infty)^{n \times n}$ with its diagonal elements 0, making $\widehat{A}$ satisfy the hypothesis on B of Proposition 3.4. Notice that $\widehat{A} \preceq \hat{A} \preceq A$ aswell.

Having seen this, we define $\widehat{x}(t)$ as the solution to the linear ivp given below.

$$\begin{cases} D_t^\alpha \widehat{x}(t) = \beta^\alpha \widehat{A}\widehat{x}(t) \ , \quad t > 0 \ , \\ \quad \widehat{x}(0) = x_0 \ , \end{cases}$$

which by Proposition 3.2 is

$$\widehat{x}(t) = E_\alpha\left((\beta t)^\alpha \widehat{A}\right) x_0 \ , \quad t \geq 0 \ .$$

By the above, and as we have seen in the proof of the previous theorem we have

$$\begin{aligned} D_t^\alpha \widehat{x}(t) &= \beta^\alpha E_\alpha\left((\beta t)^\alpha \widehat{A}\right) \widehat{A} x_0 \ , \qquad t > 0 \ , \\ D_t^\alpha \hat{y}(t) &= \beta^\alpha E_\alpha\left((\beta t)^\alpha \hat{A}\right) A x_0 \ , \qquad t > 0 \ , \end{aligned}$$

meaning that we can provide a lower found to the C-derivative of $\hat{y}(t)$: $D_t^\alpha \widehat{x}(t) \preceq D_t^\alpha \hat{y}(t)$ for $t > 0$, as $\widehat{A} \preceq \hat{A} \preceq A$. By Theorem 2.4 this implies that $\widehat{x}(t) \preceq \hat{y}(t)$ for all $t \geq 0$, and by Theorem 3.4 we know that $\widehat{x}(t)$ diverges for all $x_0 \in [0, \infty)^n \setminus \mathbf{0}$. Hence $\hat{y}(t)$ also will for any $x_0 \in [0, 1)^n \setminus \mathbf{0}$. □

For completeness, in the case $x_i(0) = (x_0)_i = 1$ we would have $\left(\hat{y}(0)\right)_i = \infty$.

Corollary 3.10. $\left\|\tilde{x}(t) - x(t)\right\| \xrightarrow{t} \infty$ *and* $\left\|\hat{x}(t) - x(t)\right\| \xrightarrow{t} 0$ *for all initial conditions* $x_0 \in [0, 1]^n$.

Proof. We have seen in Corollary 3.5 and Proposition 3.9 that both $\tilde{x}(t)$ and $\hat{y}(t)$ diverge for all initial conditions, which means that $\hat{x}(t)$ tends to $\mathbf{1}$. We also saw in page 16 that the solution $x(t)$ to our original PRN problem converges monotonically to $\mathbf{1}$ when t is sufficiently large. Hence the limits above. □

These results imply that $\hat{x}$ will behave in all the ways that we managed to characterise the solution to our original problem (3.13), and thus it can be interpreted as a fastest-spread-case scenario of the perturbation in our PRN rather than an approximation of its solution. Thus we will be able to work with it instead of the solution to the original IVP, which we haven't been able to find analytically.

3.3 Why fractional derivatives?

Why is there a need to include fractional derivatives? What benefit do we get from working with them? Let us discuss now how this work improves the solution to the classic SI model.

When one considers a connected network given by an adjacency matrix $A \in \{0, 1\}^{n \times n}$, A_{ij} can be seen as the amount of paths of length one that allow getting from node i to node j. Similarly, it is known that if we take the k-th power of matrix A then the element $\left(A^k\right)_{ij}$ will instead be the number of paths of length k from node i to j. Given this interpretation, one could consider the matricial exponential function $e^{\nu A}$ as a weighted sum of all possible paths of any length, where paths of length k are weighted by $w_k = \nu^k/k! = \nu^k/\Gamma(k+1)$. This heavily penalises long paths due to how rapidly the factorial increases, being the weight $0.5\nu^2$ for length 2 but $0.04\nu^4$ for length 4 or even $0.0014\nu^6$ for length 6, an effect which becomes even stronger if $\nu < 1$.

Further, comparing the above with the approximations to the classic SI-model (eq (3.7) or (3.9)), one can see that $\nu = \beta t$ is proportional to the time, so the solution will be useful only within a certain timescale. It would be necessary to change parameters of the dynamics (namely β) to describe processes spanning different times.

These two properties clash with the biological diffusion phenomena that PRNs undergo. Without getting here into details, we would find useful to be able to model effects taking place in different timescales or with different spatial reaches.

The appearance of the Mittag-Leffler functions in the solution of the model through the usage of the C-derivative introduces the parameter $0 < \alpha < 1$ in the solution, allowing us to adjust its value to accommodate for both these effects without having to change the model parameters. Say, as

$$E_\alpha\left((\beta t)^\alpha A\right) = \sum_{k=0}^{\infty} \frac{(\beta t)^{\alpha k} A^k}{\Gamma(\alpha k + 1)} \ , \qquad t \geq 0 \ ,$$

we can tweak the value of α to change the weights given to each path length, which will become $w_k = \nu^{\alpha k}/\Gamma(\alpha k+1)$. As an example, with an α value of 0.5 the weights become ν, $0.5\nu^2$ and $0.16\nu^3$ for lengths 2, 4 and 6 respectively. Notice that not only the individual penalisation for length has reduced, but the ratio of penalisation as well, allowing us to model further-reaching effects. Similarly with the contribution of time through ν.

Despite our found solutions to the approximations, and the upper bound proved for the original IVP describing the PRN, not being exactly the Mittag-Leffler function above, we have seen that it plays a crucial part in the expressions of $\tilde{x}(t)$ and more importantly $\hat{x}(t) = f\left(\hat{y}(t)\right)$. Therefore these effects will directly influence the behaviour of our descriptions even if not exactly as described here, and now we can understand the meaning of $\hat{y}(t)$ as written in Proposition 3.7 more deeply and venture what the implications on $\hat{x}(t)$ are.

In all, the introduction of a fractional order model and with it the Mittag-Leffler function allows us to have a finer control on the spatial reach and the timescale of the phenomena we would wish to study thanks to the additional parameter it introduces to the solution of the classical model.

References

[1] L. Abadias, G. Estrada-Rodriguez, and E. Estrada. Fractional-order susceptible-infected model: Definition and applications to the study of covid-19 main protease. *Fractional Calculus and Applied Analysis*, 23(3):635–655, 2020.

[2] R. Pérez-Marco. On the definition of Euler Gamma function. *L'Enseignement Mathématique* , 1/2:135–160, May 2022.

[3] J. M. Patin. A very short proof of stirling's formula. *The American Mathematical Monthly*, 96(1):41–42, 1989.

[4] F. Mainardi. *Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models*. Imperial College Press, 2010.

[5] I. M. Gel'fand and G. E. Shilov. *Generalized functions. Properties and operations.*, volume 1. Academic Press New York and London, 1964.