An introduction to fractional calculus

Fundamental ideas and numerics

Fabio Durastante

Università di Pisa

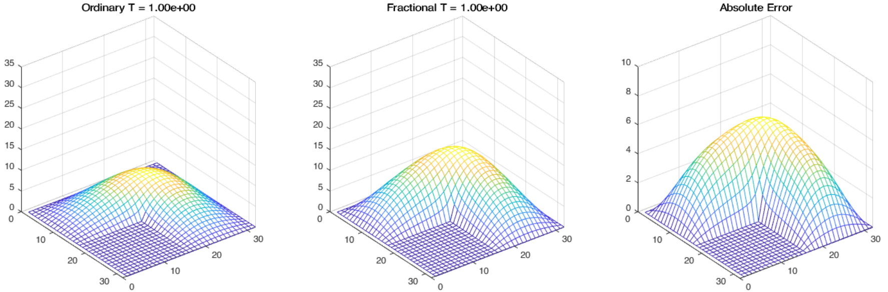
fabio.durastante@unipi.it

fdurastante.github.io

June, 2022

Subdiffusion equations

At the end of the last lecture we had observed the following behavior:



for the solution of:

$$ {}_{CA}D_t^{\alpha} u = 0.05\nabla^2 u, \quad \alpha = 0.3, 1. $$

The visual effect seemed to be a **slowing down of the diffusion**.

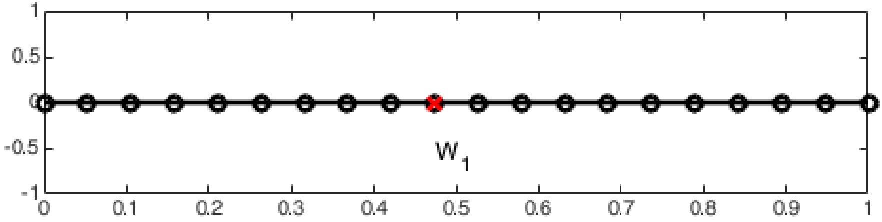
Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,

Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

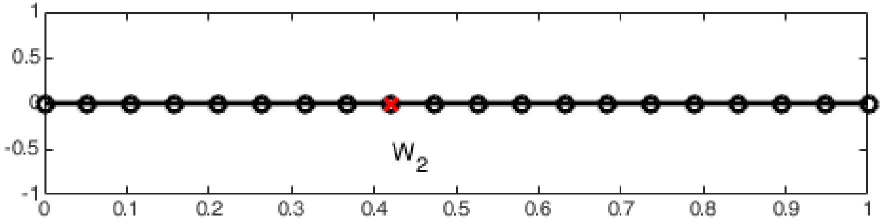
$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

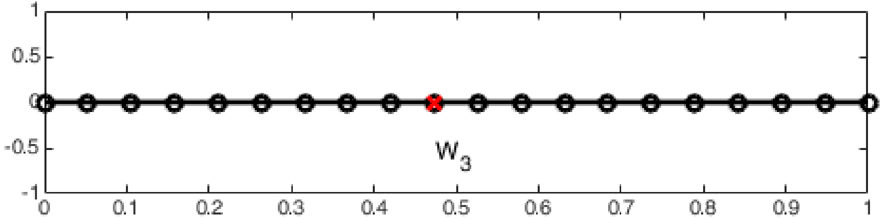
$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

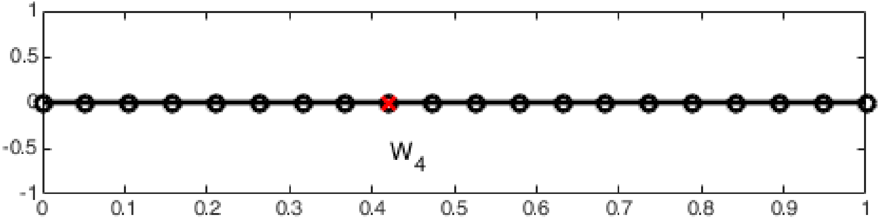
$$W_j(t + \Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

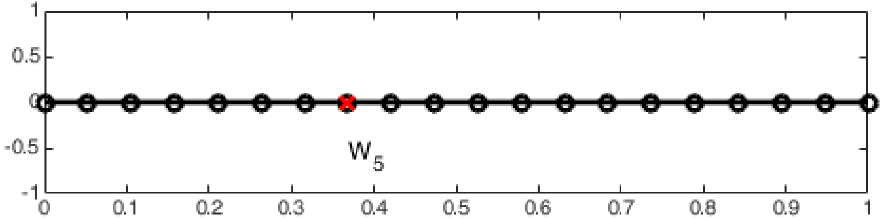
$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

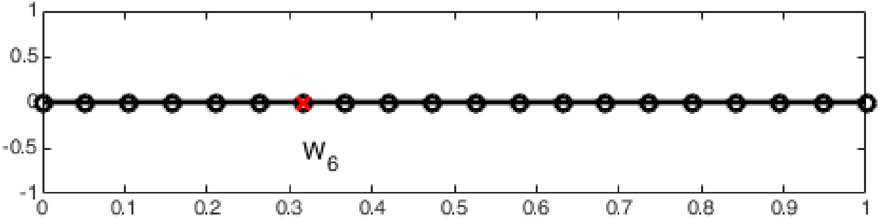
$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

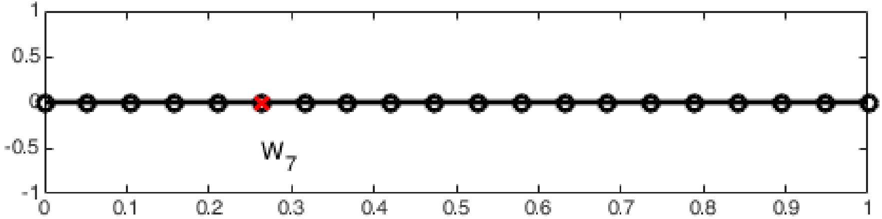
$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

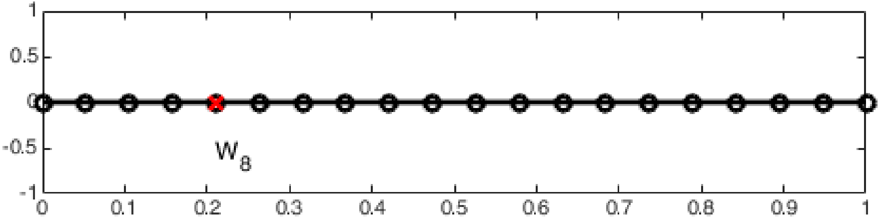
$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

$$W_j(t+\Delta t) = \frac{1}{2}W_{j-1}(t) + \frac{1}{2}W_{j+1}(t)$$



Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$

Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$

- The master equation defines the *pdf* to be at position j at time $t + \Delta t$ depending on the population of the two adjacent sites $j \pm 1$ at time t.

Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

$$W_j(t + \Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$

- The master equation defines the *pdf* to be at position j at time $t + \Delta t$ depending on the population of the two adjacent sites $j \pm 1$ at time t.
- The prefactor $1/2$ tells us that the **process is isotropic** with respect to the left/right direction.

Brownian motion (Metzler and Klafter 2000)

- Consider a 1D lattice with cell size Δx,
- In discrete time steps of span Δt a test particle jumps to one of its neighbour sites,
- The process can be modelled by the master equation

$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$

- The master equation defines the *pdf* to be at position j at time $t+\Delta t$ depending on the population of the two adjacent sites $j \pm 1$ at time t.
- The prefactor $1/2$ tells us that the **process is isotropic** with respect to the left/right direction.
- If we let $\Delta t \to 0$, $\Delta x \to 0$ and do a Taylor expansion in both Δ and Δx we get

$$\begin{aligned} W_j(t+\Delta t) =& W_j(t) + \Delta t \frac{\partial W_j}{\partial t} + O([\Delta t]^2), \qquad \text{for } \Delta t \to 0, \\ W_{j\pm 1}(t) =& W(x,t) \pm \Delta x \frac{\partial W}{\partial x} + \frac{(\Delta x)^2}{2} \frac{\partial^2 W}{\partial x^2} + O([\Delta x]^3), \qquad \text{for } \Delta x \to 0, \end{aligned}$$

Brownian motion (Metzler and Klafter 2000)

We now substitute the expansions

$$\begin{aligned} W_j(t+\Delta t) =& W_j(t) + \Delta t \frac{\partial W_j}{\partial t} + O([\Delta t]^2), \qquad \text{for } \Delta t \to 0, \\ W_{j\pm 1}(t) =& W(x,t) \pm \Delta x \frac{\partial W}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 W}{\partial x^2} + O([\Delta x]^3), \qquad \text{for } \Delta x \to 0, \end{aligned}$$

in

$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$

obtaining

$$W(x,t) + \Delta \mathrm{t} \frac{\partial W}{\partial t} + O\left(\Delta \mathrm{t}^2\right) = W(x,t) + \frac{1}{2}\Delta \mathrm{x}^2 \frac{\partial^2 W}{\partial x^2} + O\left(\Delta \mathrm{x}^3\right)$$

Brownian motion (Metzler and Klafter 2000)

We now substitute the expansions

$$
\begin{aligned}
W_j(t+\Delta t) =& W_j(t) + \Delta t \frac{\partial W_j}{\partial t} + O([\Delta t]^2), \qquad \text{for } \Delta t \to 0, \\
W_{j\pm 1}(t) =& W(x,t) \pm \Delta x \frac{\partial W}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 W}{\partial x^2} + O([\Delta x]^3), \qquad \text{for } \Delta x \to 0,
\end{aligned}
$$

in

$$W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)$$

obtaining

$$\frac{\partial W}{\partial t} = \frac{\Delta x^2}{2\Delta t}\frac{\partial^2 W}{\partial x^2} + O\left(\Delta \mathrm{x}^3 + \Delta t\right)$$

Brownian motion (Metzler and Klafter 2000)

We now substitute the expansions

$$
\begin{aligned}
W_j(t+\Delta t) =& W_j(t) + \Delta t \frac{\partial W_j}{\partial t} + O([\Delta t]^2), \qquad \text{for } \Delta t \to 0, \\
W_{j\pm 1}(t) =& W(x,t) \pm \Delta x \frac{\partial W}{\partial x} + \frac{(\Delta x)^2}{2}\frac{\partial^2 W}{\partial x^2} + O([\Delta x]^3), \qquad \text{for } \Delta x \to 0,
\end{aligned}
$$

in

$$
W_j(t+\Delta t) = \frac{1}{2} W_{j-1}(t) + \frac{1}{2} W_{j+1}(t)
$$

obtaining

$$
\frac{\partial W}{\partial t} = K_1 \frac{\partial^2 W}{\partial x^2}, \qquad K_1 = \lim_{\substack{\Delta x \to 0 \\ \Delta t \to 0}} \frac{\Delta x^2}{2\Delta t} < \infty.
$$

Brownian motion

$$\frac{\partial W}{\partial t} = K_1 \frac{\partial^2 W}{\partial x^2}$$

Let us call X the random variable measuring the distance covered in two consecutive jumps

- Assume that the *pdf* of X (appropriately normalised) has existing moments

$$\overline{X} = \sum_i X_i, \qquad \overline{X^2},$$

and mean time-span Δt between any two individual jump events.

Brownian motion

$$\frac{\partial W}{\partial t} = K_1 \frac{\partial^2 W}{\partial x^2}$$

Let us call X the random variable measuring the distance covered in two consecutive jumps

- Assume that the *pdf* of X (appropriately normalised) has existing moments

$$\overline{X} = \sum_i X_i, \qquad \overline{X^2},$$

  and mean time-span Δt between any two individual jump events.

- Then the **central limit theorem** assures that exists

$$V = \frac{\overline{X}}{\Delta t} \text{ (Mean velocity)} \quad K = \frac{\overline{X^2} - \overline{X}^2}{2\Delta t} \text{ (Diffusion coefficient)}$$

  and that

$$W(x,t) = \frac{1}{2\sqrt{\pi K_1 t}} \exp\left(-x^2/4K_1 t\right).$$

Brownian motion: the Fourier domain

We can rewrite

$$W(x,t) = \frac{1}{2\sqrt{\pi K_1 t}} \exp\left(-{}^{x^2}\!/_{4K_1 t}\right).$$

in the **Fourier domain** as

$$W(k,t) = \exp(-K_1 k^2 t), \qquad W_0(x) = \lim_{t \to 0^+} W(x,t) = \delta(x),$$

that solve the Fourier transformed diffusion equation

$$\frac{\partial W}{\partial t} = -K_1 k^2 W(k,t),$$

that is a **relaxation equation**, for a fixed wavenumber k.

From the discrete to the continuous

The **C**ontinuous **T**ime **R**andom **W**alk model (CTRW):

- Both the **length of a given jump**, and the **waiting time** elapsing between two successive jumps are drawn from a pdf $\psi(x, t)$

From the discrete to the continuous

The **C**ontinuous **T**ime **R**andom **W**alk model (CTRW):

- Both the **length of a given jump**, and the **waiting time** elapsing between two successive jumps are drawn from a pdf $\psi(x, t)$
- The jump length pdf

$$\lambda(x) = \int_0^{+\infty} \psi(x, y)\, \mathrm{d}t,$$

Jump length

$\lambda(x)\mathrm{d}x$ produces the probability for a jump length in the interval $(x, x + \mathrm{d}x)$.

From the discrete to the continuous

The **C**ontinuous **T**ime **R**andom **W**alk model (CTRW):

- Both the **length of a given jump**, and the **waiting time** elapsing between two successive jumps are drawn from a pdf $\psi(x, t)$
- The jump length pdf

$$\lambda(x) = \int_0^{+\infty} \psi(x, y)\, \mathrm{d}t,$$

- The waiting time pdf

$$w(t) = \int_{-\infty}^{+\infty} \psi(x, t)\, \mathrm{d}x$$

Waiting time

$w(t)\mathrm{d}t$ produces the probability for a waiting time in the interval $(t, t + \mathrm{d}t)$.

From the discrete to the continuous

The **C**ontinuous **T**ime **R**andom **W**alk model (CTRW):

- Both the **length of a given jump**, and the **waiting time** elapsing between two successive jumps are drawn from a pdf $\psi(x, t)$
- The jump length pdf

$$\lambda(x) = \int_0^{+\infty} \psi(x, y)\, \mathrm{d}t,$$

- The waiting time pdf

$$w(t) = \int_{-\infty}^{+\infty} \psi(x, t)\, \mathrm{d}x$$

- If the jump length and waiting time are **independent random variables** then:

$$\psi(x, t) = w(t)\lambda(x)$$

Characterization of CTRW

To categorise different CTRW one can look at the quantities

$$T = \int_0^{+\infty} t w(t)\,\mathrm{d}t, \text{ (Characteristic waiting time)},$$

and

$$\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x \text{ (Jump length variance)},$$

specifically, are they **finite**? Do they **diverge**?

Characterization of CTRW

To categorise different CTRW one can look at the quantities

$$T = \int_0^{+\infty} t w(t)\,\mathrm{d}t, \text{ (Characteristic waiting time)},$$

and

$$\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x \text{ (Jump length variance)},$$

specifically, are they **finite**? Do they **diverge**?
The master (Langevin) equation for this process is then given by

$$\eta(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}x' \int_0^{+\infty} \mathrm{d}t'\, \eta(x',t')\psi(x-x',t-t') + \delta(x)\delta(t),$$

Characterization of CTRW

To categorise different CTRW one can look at the quantities

$$T = \int_0^{+\infty} t w(t)\,\mathrm{d}t, \text{ (Characteristic waiting time)},$$

and

$$\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x \text{ (Jump length variance)},$$

specifically, are they **finite**? Do they **diverge**?
The master (Langevin) equation for this process is then given by

$$\eta(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}x' \int_0^{+\infty} \mathrm{d}t'\, \eta(x',t')\psi(x-x',t-t') + \delta(x)\delta(t),$$

Pdf of having arrived at position x at time t – $\eta(x,t)$ –

Characterization of CTRW

To categorise different CTRW one can look at the quantities

$$T = \int_0^{+\infty} t w(t)\,\mathrm{d}t, \text{ (Characteristic waiting time)},$$

and

$$\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x \text{ (Jump length variance)},$$

specifically, are they **finite**? Do they **diverge**?
The master (Langevin) equation for this process is then given by

$$\eta(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}x' \int_0^{+\infty} \mathrm{d}t'\, \eta(x',t') \psi(x-x', t-t') + \delta(x)\delta(t),$$

Pdf of having arrived at position x at time t – $\eta(x,t)$ – having just arrived at x' at time t' – $\eta(x',t')$ –

Characterization of CTRW

To categorise different CTRW one can look at the quantities

$$T = \int_0^{+\infty} t w(t) \, \mathrm{d}t, \text{ (Characteristic waiting time)},$$

and

$$\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x) \, \mathrm{d}x \text{ (Jump length variance)},$$

specifically, are they **finite**? Do they **diverge**?
The master (Langevin) equation for this process is then given by

$$\eta(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}x' \int_0^{+\infty} \mathrm{d}t' \, \eta(x',t') \psi(x-x', t-t') + \delta(x)\delta(t),$$

Pdf of having arrived at position x at time t – $\eta(x,t)$ – having just arrived at x' at time t' – $\eta(x',t')$ – with initial condition $\delta(x)$.

Characterization of CTRW

Then if we use

$$\eta(x,t) = \int_{-\infty}^{+\infty} \mathrm{d}x' \int_0^{+\infty} \mathrm{d}t' \, \eta(x',t')\psi(x-x',t-t') + \delta(x)\delta(t),$$

we can write the pdf of being in x at time t as

$$W(x,t) = \int_0^t \eta(x,t')\Psi(t-t'), \mathrm{d}t, \qquad \Psi(t) = 1 - \int_0^t w(t')\,\mathrm{d}t',$$

where the latter is the cumulative probability assigned to the probability of no jump event during the time interval $t - t'$.

Fact

If both T and Σ^2 are finite the long-time limit corresponds to Brownian motion, e.g., $w(t) = \tau^{-1} exp(-t/\tau)$, $T = \tau$, $\lambda(x) = (4\pi\sigma^2)^{-1/2} \exp(-x^2/4\sigma^2)$, $\Sigma^2 = 2\sigma^2$, we recover the standard diffusion equation.

The CTRW in the Fourier-Laplace domain

We take

$$W(x,t) = \int_0^t \eta(x,t')\Psi(t-t'), \mathrm{d}t, \qquad \Psi(t) = 1 - \int_0^t w(t')\,\mathrm{d}t',$$

and rewrite it again in the **Fourier-Laplace domain** (Fourier for the space variable, Laplace for the time one) as

$$W(k,u) = \frac{1-w(u)}{u}\frac{W_0(k)}{1-\psi(k,u)}, \qquad W_0(k) = \int_{-\infty}^{+\infty} W_0(x)e^{-i2\pi kx}\,\mathrm{d}x.$$

In the **Brownian case**

$$w(u) \sim 1 - u\tau + O(\tau^2), \quad \lambda(k) \sim 1 - \sigma^2 k^2 + O(k^4), \quad W_0(x) = \delta(x)$$

then

$$W(k,u) = \frac{1}{u + K_1 k^2}, \quad K_1 = {}^{\sigma^2}/_{\tau}.$$

The case of long rests

Long rests

The **characteristic waiting time** $T = \int_0^{+\infty} tw(t)\,\mathrm{d}t$ **diverges**, but the jump length variance $\Sigma^2 = \int_{-\infty}^{+\infty} x^2\lambda(x)\,\mathrm{d}x$ is finite.

The case of long rests

Long rests

The **characteristic waiting time** $T = \int_0^{+\infty} t w(t)\,\mathrm{d}t$ **diverges**, but the jump length variance $\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x$ is finite.

- To realize this we can select

$$w(t) \sim A_\alpha \left(\tau/t\right)^{1+\alpha}, \qquad 0 < \alpha < 1,$$

The case of long rests

Long rests

The **characteristic waiting time** $T = \int_0^{+\infty} t w(t)\,\mathrm{d}t$ **diverges**, but the jump length variance $\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x$ is finite.

- To realize this we can select

$$w(t) \sim A_\alpha \left(\tau/t\right)^{1+\alpha}, \qquad 0 < \alpha < 1,$$

- For the jump pdf we use again the Gaussian jump length

$$\lambda(x) = \left(4\pi\sigma^2\right)^{-1/2} \exp\left(-x^2/4\sigma^2\right).$$

The case of long rests

Long rests

The **characteristic waiting time** $T = \int_0^{+\infty} t w(t)\,\mathrm{d}t$ **diverges**, but the jump length variance $\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x$ is finite.

- To realize this we can select

$$w(t) \sim A_\alpha \left({}^{\tau}/_{t}\right)^{1+\alpha}, \qquad 0 < \alpha < 1,$$

- For the jump pdf we use again the Gaussian jump length

$$\lambda(x) = (4\pi\sigma^2)^{-1/2} \exp(-{}^{x^2}/_{4\sigma^2}).$$

- To get the form of the equation we first go to the Laplace domain:

$$w(u) \sim 1 - (u\tau)^\alpha,$$

The case of long rests

Long rests

The **characteristic waiting time** $T = \int_0^{+\infty} t w(t)\,\mathrm{d}t$ **diverges**, but the jump length variance $\Sigma^2 = \int_{-\infty}^{+\infty} x^2 \lambda(x)\,\mathrm{d}x$ is finite.

- To realize this we can select

$$w(t) \sim A_\alpha \left({}^{\tau}/_{t}\right)^{1+\alpha}, \qquad 0 < \alpha < 1,$$

- For the jump pdf we use again the Gaussian jump length

$$\lambda(x) = \left(4\pi\sigma^2\right)^{-1/2} \exp\left(-{}^{x^2}/_{4\sigma^2}\right).$$

- To get the form of the equation we first go to the Laplace domain:

$$w(u) \sim 1 - (u\tau)^\alpha,$$

- and then obtain the expression for $W(k,u)$ in the Fourier-Laplace space

$$W(k,u) = {}^{W_0(k)}/_{u} \big/ \left(1 + K_\alpha u^{-\alpha} k^2\right).$$

The case of long rests

To get an expression of the equation we use the Laplace transform for fractional integrals:

$$\mathcal{L}\left\{ I_{[0,t]}^{-\alpha} W(x,t) \right\} = u^{-\alpha} W(x,u), \qquad \alpha \geq 0,$$

and together with

$$W(k,u) = \frac{W_0(k)/u}{(1 + K_\alpha u^{-\alpha} k^2)}.$$

we infer the fractional integral equation

$$W(x,t) - W_0(x) = I_{[0,t]} K_\alpha \frac{\partial^2}{\partial x^2} W(x,t).$$

The case of long rests

To get an expression of the equation we use the Laplace transform for fractional integrals:

$$\mathcal{L}\left\{ I_{[0,t]}^{-\alpha} W(x,t) \right\} = u^{-\alpha} W(x,u), \qquad \alpha \geq 0,$$

and together with

$$W(k,u) = \frac{W_0(k)/u}{(1 + K_\alpha u^{-\alpha} k^2)}.$$

we infer the fractional integral equation, and apply derivative w.r.t. to time

$$\frac{\partial}{\partial t}\left(W(x,t) - W_0(x)\right) = \frac{\partial}{\partial t}\left(I_{[0,t]} K_\alpha \frac{\partial^2}{\partial x^2} W(x,t) \right).$$

The case of long rests

To get an expression of the equation we use the Laplace transform for fractional integrals:

$$\mathcal{L}\left\{I_{[0,t]}^{-\alpha} W(x,t)\right\} = u^{-\alpha} W(x,u), \qquad \alpha \geq 0,$$

and together with

$$W(k,u) = \frac{W_0(k)/u}{(1 + K_\alpha u^{-\alpha} k^2)}.$$

we infer the fractional integral equation

$$\frac{\partial W}{\partial t} = {}_{RL}D_{[0,t]}^{\alpha} K_\alpha \frac{\partial^2}{\partial x^2} W(x,t).$$

The case of long rests

To get an expression of the equation we use the Laplace transform for fractional integrals:

$$\mathcal{L}\left\{ I^{-\alpha}_{[0,t]} W(x,t) \right\} = u^{-\alpha} W(x,u), \qquad \alpha \geq 0,$$

and together with

$$W(k,u) = \frac{W_0(k)/u}{(1 + K_\alpha u^{-\alpha} k^2)}.$$

we infer the fractional integral equation

$$\frac{\partial W}{\partial t} = {}_{RL}D^{\alpha}_{[0,t]} K_\alpha \frac{\partial^2}{\partial x^2} W(x,t).$$

We can compute also the mean squared displacement

$$\langle x^2(t) \rangle = \mathcal{L}^{-1}\left\{ \lim_{k \to 0} -\frac{d^2}{dk^2} W(k,u) \right\} = \frac{2K_\alpha}{\Gamma(1+\alpha)} t^\alpha.$$

The case of long rests

We have obtained a Fractional Differential Equation:

$$\frac{\partial W}{\partial t} = {}_{RL}D^{\alpha}_{[0,t]} K_\alpha \frac{\partial^2}{\partial x^2} W(x,t), \qquad 0 < \alpha < 1$$

but this is not the model we started looking at, that was

$${}_{CA}D^{\alpha}_{[0,t]} W = K_\alpha \frac{\partial^2}{\partial x^2} W(x,t), \qquad 0 < \alpha < 1$$

? Are they related?

The case of long rests

We have obtained a Fractional Differential Equation:

$$\frac{\partial W}{\partial t} = {}_{RL}D^{\alpha}_{[0,t]} K_{\alpha} \frac{\partial^2}{\partial x^2} W(x,t), \qquad 0 < \alpha < 1$$

but this is not the model we started looking at, that was

$${}_{CA}D^{\alpha}_{[0,t]} W = K_{\alpha} \frac{\partial^2}{\partial x^2} W(x,t), \qquad 0 < \alpha < 1$$

❓ Are they related? It turns out that this is indeed the case (Sokolov and Klafter 2005), the proof involves doing some work in inverting Fourier-Laplace transform.

The case of long rests

We have obtained a Fractional Differential Equation:

$$\frac{\partial W}{\partial t} = {}_{RL}D^{\alpha}_{[0,t]} K_\alpha \frac{\partial^2}{\partial x^2} W(x,t), \qquad 0 < \alpha < 1$$

but this is not the model we started looking at, that was

$$ {}_{CA}D^{\alpha}_{[0,t]} W = K_\alpha \frac{\partial^2}{\partial x^2} W(x,t), \qquad 0 < \alpha < 1$$

❓ Are they related? It turns out that this is indeed the case (Sokolov and Klafter 2005), the proof involves doing some work in inverting Fourier-Laplace transform.

We now have an *interpretation* of what a Fractional Derivative with respect to time is. We will come back to this when we will speak about fractional derivative with respect to space.

“Exponential” Fractional Integrators

We start from the FDE

$$\begin{cases} {}_{CA}D^{\alpha}_{[t_0,t]} u(t) + \lambda y(t) = f(t), \\ u(0) = u_0, \end{cases} \qquad \alpha \in \mathbb{R}_{>0}, \quad \lambda \in \mathbb{R}, \; u(t) : [t_0, T] \to \mathbb{R}.$$

Then we rewrite the solution as

$$u(t) = e_{\alpha,1}(t - t_0; \lambda) u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t - s; \lambda) f(s) \, \mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1} E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

for $E_{\alpha,\beta}(z)$ the Mittag-Leffler (ML) function with two parameters.

"Exponential" Fractional Integrators

We start from the FDE

$$\begin{cases} {}_{CA}D^{\alpha}_{[t_0,t]} u(t) + \lambda y(t) = f(t), \\ u(0) = u_0, \end{cases} \qquad \alpha \in \mathbb{R}_{>0}, \quad \lambda \in \mathbb{R},\; u(t) : [t_0, T] \to \mathbb{R}.$$

Then we rewrite the solution as

$$u(t) = e_{\alpha,1}(t - t_0; \lambda) u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t - s; \lambda) f(s)\,\mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1} E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

for $E_{\alpha,\beta}(z)$ the Mittag-Leffler (ML) function with two parameters.

- We can use this formulation to build different PI rules,

“Exponential” Fractional Integrators

We start from the FDE

$$\begin{cases} {}_{CA}D^{\alpha}_{[t_0,t]}u(t) + \lambda y(t) = f(t), \\ u(0) = u_0, \end{cases} \qquad \alpha \in \mathbb{R}_{>0}, \quad \lambda \in \mathbb{R},\ u(t) : [t_0, T] \to \mathbb{R}.$$

Then we rewrite the solution as

$$u(t) = e_{\alpha,1}(t - t_0; \lambda)u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t - s; \lambda) f(s)\,\mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1}E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

for $E_{\alpha,\beta}(z)$ the Mittag-Leffler (ML) function with two parameters.

- We can use this formulation to build different PI rules,
- We can use it to address the problem

$${}_{CA}D^{\alpha}_{[t_0,t]}U(t) + Ay(t) = F(U(t)), \quad U(0) = U_0.$$

Evaluation of the ML function

For both the approaches we need reliable ways for **computing** the **ML function** on both the **real line** and with **matrix argument**.

Evaluation of the ML function

For both the approaches we need reliable ways for **computing** the **ML function** on both the **real line** and with **matrix argument**.

Scalar case: Inversion of the Laplace transform via the Optimal Parabola Contour selection algorithm (Garrappa 2015),

Evaluation of the ML function

For both the approaches we need reliable ways for **computing** the **ML function** on both the **real line** and with **matrix argument**.

Scalar case Inversion of the Laplace transform via the Optimal Parabola Contour selection algorithm (Garrappa 2015),

Matrix argument To apply algorithm for matrix-function evaluation we may need also the value of the derivative of the ML function, e.g., Schur-Parlett type algorithm (Garrappa and Popolizio 2018; Higham and Liu 2021).

In general, we expect to mostly need matrix function–times–vector operations:

$$\mathbf{y} = E_{\alpha,\beta}(A)\mathbf{v}, \qquad A \in \mathbb{R}^{n\times n}, \quad \mathbf{y}, \mathbf{v} \in \mathbb{R}^n.$$

Evaluation of the ML function

For both the approaches we need reliable ways for **computing** the **ML function** on both the **real line** and with **matrix argument**.

Scalar case: Inversion of the Laplace transform via the Optimal Parabola Contour selection algorithm (Garrappa 2015),

Matrix argument: To apply algorithm for matrix-function evaluation we may need also the value of the derivative of the ML function, e.g., Schur-Parlett type algorithm (Garrappa and Popolizio 2018; Higham and Liu 2021).

In general, we expect to mostly need matrix function–times–vector operations:

$$\mathbf{y} = E_{\alpha,\beta}(A)\mathbf{v}, \qquad A \in \mathbb{R}^{n\times n}, \quad \mathbf{y},\mathbf{v} \in \mathbb{R}^n.$$

We postpone it to after we have discussed the actual necessities we have.

PI - "Exponential" Fractional Integrators

We start from the formula

$$u(t) = e_{\alpha,1}(t - t_0; \lambda)u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t - s; \lambda) f(s) \, \mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1} E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

and select a grid $\{t_i\}_{i=0}^{N}$, then

$$u(t_n) = e_{\alpha,1}(t_n - t_0; \lambda)u_0 + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} e_{\alpha,\alpha}(t_n - s; \lambda) f(s) \, \mathrm{d}s.$$

PI - "Exponential" Fractional Integrators

We start from the formula

$$u(t) = e_{\alpha,1}(t - t_0;\lambda)u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t-s;\lambda)f(s)\,\mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1}E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

and select a grid $\{t_i\}_{i=0}^{N}$, then

$$u(t_n) = e_{\alpha,1}(t_n - t_0;\lambda)u_0 + \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} e_{\alpha,\alpha}(t_n - s;\lambda)f(s)\,\mathrm{d}s.$$

- In general we have

$$e_{\alpha,\beta}(t;\lambda) = \tau^{\beta-1}e_{\alpha,\beta}(t/\tau;\tau^{\alpha}\lambda)$$

PI - “Exponential” Fractional Integrators

We start from the formula

$$u(t) = e_{\alpha,1}(t - t_0; \lambda) u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t - s; \lambda) f(s) \, \mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1} E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

and select a grid $\{t_i\}_{i=0}^{N}$, then

$$u(t_n) = e_{\alpha,1}(t_n - t_0; \lambda) u_0 + \sum_{j=0}^{n-1} \int_{t_j}^{t_{j+1}} e_{\alpha,\alpha}(t_n - s; \lambda) f(s) \, \mathrm{d}s.$$

- In general we have
$$e_{\alpha,\beta}(t; \lambda) = \tau^{\beta-1} e_{\alpha,\beta}(t/\tau; \tau^{\alpha}\lambda)$$
- For $s \in [t_j, t_{j+1}]$ let us consider the *change of variables* $s = t_j + r\tau$, $r \in [0, 1]$

PI - "Exponential" Fractional Integrators

We start from the formula

$$u(t) = e_{\alpha,1}(t - t_0; \lambda)u_0 + \int_{t_0}^{t} e_{\alpha,\alpha}(t - s; \lambda) f(s)\,\mathrm{d}s, \quad e_{\alpha,\beta} = t^{\beta-1} E_{\alpha,\beta}(-\lambda t^{\alpha}),$$

and select a grid $\{t_i\}_{i=0}^{N}$, then

$$u(t_n) = e_{\alpha,1}(t_n - t_0; \lambda)u_0 + \tau^{\alpha} \sum_{j=0}^{n-1} \int_0^1 e_{\alpha,\alpha}\left({}^{(t-t_j)}/_{\tau} - r; \tau^{\alpha}\lambda\right) f(t_j + r\tau)\,\mathrm{d}r.$$

- In general we have

$$e_{\alpha,\beta}(t; \lambda) = \tau^{\beta-1} e_{\alpha,\beta}({}^{t}/_{\tau}; \tau^{\alpha}\lambda)$$

- For $s \in [t_j, t_{j+1}]$ let us consider the *change of variables* $s = t_j + r\tau$, $r \in [0, 1]$

PI - “Exponential” Fractional Integrators

Then a PI rule for

$$u(t_n) = e_{\alpha,1}(t_n - t_0; \lambda)u_0 + \tau^\alpha \sum_{j=0}^{n-1} \int_0^1 e_{\alpha,\alpha}\left({}^{(t-t_j)}/{}_\tau - r; \tau^\alpha \lambda\right) f(t_j + r\tau)\, \mathrm{d}r.$$

is obtained by selecting $q+1$ *distinct* nodes $0 \leq c_0 < c_1 < \cdots < c_q \leq 1$ and replacing $f(t_j + r\tau)$ with

$$p_j^{[q]}(t_j{+}r\tau) = \sum_{\ell=0}^{q} L_\ell^{[q]}(r) f(t_j{+}c_\ell\tau), \quad r \in [0,1], \quad L_\ell^{[q]} \text{ Lagrange basis element of degree } q.$$

PI - “Exponential” Fractional Integrators

Then the PI rule is

$$u^{(n)} = e_{\alpha,1}(t_n - t_0; \lambda) y_0 + \tau^{\alpha} \sum_{j=0}^{n-1} \sum_{\ell=0}^{q} \omega_{\ell}^{[q;\alpha]}(n-j; \tau^{\alpha}\lambda) f(t_j + c_{\ell}\tau).$$

is obtained by selecting $q+1$ *distinct* nodes $0 \leq c_0 < c_1 < \cdots < c_q \leq 1$ and replacing $f(t_j + r\tau)$ with

$$p_j^{[q]}(t_j + r\tau) = \sum_{\ell=0}^{q} L_{\ell}^{[q]}(r) f(t_j + c_{\ell}\tau), \quad r \in [0,1], \quad L_{\ell}^{[q]} \text{ Lagrange basis element of degree } q.$$

And selecting the weights

$$\omega_{\ell}^{[q;\alpha]}(n, z) = \int_0^1 e_{\alpha,\alpha}(n - j - r; z) L_{\ell}^{[q]}(r)\, \mathrm{d}r.$$

PI - "Exponential" Fractional Integrators

Theorem (Garrappa and Popolizio 2011, Theorem 4.2)

Let $\alpha > 0$ and $f(t) \in \mathcal{C}^{q+2}([t_0, T])$. The error of a q-step exponential PI rule is given by

$$u(t_n) - u^{(n)} = \tau^{q+1} \frac{C_0^{[q]}}{(q+1)!} \int_{t_0}^{t_n} e_{\alpha,\alpha}(t_n - s; \lambda) f^{(q+1)}(s)\,\mathrm{d}s + O(\tau^{q+1+\alpha}),$$

where the constant $C_0^{[q]}$ depends only on the nodes c_ℓ.

- For $q = 2$, $c_0 = 0$, $c_1 = 1/2$ $c_2 = 1$, one finds $C_0^{[2]} = 0$, thus an interpolatory formula of order $O(\tau^{q+1+\alpha})$.
- The **general idea** is to select nodes c_ℓ in such way that

$$C_\nu^{[q]} = \int_0^1 \omega_q(r) \xi(1 - \nu, 1 - r)\,\mathrm{d}r, \quad \nu \in \mathbb{R},$$

for ξ the *Hurwitz zeta function*, are zeroed out in the error expansion for the method.

The MOL/Matrix case

Let us go back to the case that sparked our interest in going "exponential", that was the MOL problem

$$\begin{cases} {}_{CA}D^{\alpha}_{[0,t]}\mathbf{u}(t) + A\mathbf{u}(t) = \mathbf{g}(t), \quad t > 0, \\ \mathbf{u}(0) = \mathbf{u}_0. \end{cases}$$

By the variation of constant formula, we have seen that we can express the solution as

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1}E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

The MOL/Matrix case

Let us go back to the case that sparked our interest in going "exponential", that was the MOL problem

$$\begin{cases} {}_{CA}D^{\alpha}_{[0,t]}\mathbf{u}(t) + A\mathbf{u}(t) = \mathbf{g}(t), \quad t > 0, \\ \mathbf{u}(0) = \mathbf{u}_0. \end{cases}$$

By the variation of constant formula, we have seen that we can express the solution as

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

- In the general case we then have to apply one of the PI rules to compute the integral term,

The MOL/Matrix case

Let us go back to the case that sparked our interest in going “exponential”, that was the MOL problem

$$\begin{cases} {}_{CA}D^{\alpha}_{[0,t]}\mathbf{u}(t) + A\mathbf{u}(t) = \mathbf{g}(t), \quad t > 0, \\ \mathbf{u}(0) = \mathbf{u}_0. \end{cases}$$

By the variation of constant formula, we have seen that we can express the solution as

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1}E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

- In the general case we then have to apply one of the PI rules to compute the integral term,
- If $g(s) = \sum_{k=0}^{q} s^k \mathbf{v}_k$ for some vectors, we can compute the integral on the right-hand side in *closed form* and obtain

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{y}_0 + \sum_{k=0}^{q} \Gamma(k+1)t^{\alpha+k}E_{\alpha,\alpha+k+1}(-t^{\alpha}A)\mathbf{v}_k, \qquad t > 0.$$

Matrix functions: the normal case

If A is a normal matrix, and f is a function existing on the spectrum of A, then

$$f(A) = Uf(\Lambda)U^H, \quad U^H U = I, \quad \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n), \; A\mathbf{u}_i = \lambda_i \mathbf{u}_i, \quad U = [\mathbf{u}_1, \ldots, \mathbf{u}_n].$$

This is, e.g., sufficient for the cases in which

- A is the discretization of a self-adjoint operator,
- A is symmetric.

$E_{\alpha,\beta}(z)$ is an **analytic function**, and therefore we can compute it for every possible eigenvalue λ in the spectrum of A.

Matrix functions: the normal case

If A is a normal matrix, and f is a function existing on the spectrum of A, then

$$f(A) = Uf(\Lambda)U^H, \quad U^H U = I, \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n),\ A\mathbf{u}_i = \lambda_i \mathbf{u}_i, \quad U = [\mathbf{u}_1, \ldots, \mathbf{u}_n].$$

This is, e.g., sufficient for the cases in which

- A is the discretization of a self-adjoint operator,
- A is symmetric.

$E_{\alpha,\beta}(z)$ is an **analytic function**, and therefore we can compute it for every possible eigenvalue λ in the spectrum of A.

What about the *non-normal* and *nond-diagonalizable* case? For diagonalizable matrices, we can use the eigendecomposition at the same way.

Matrix functions: the Jordan Canonical Form

Jordan Canonical Form

We recall that any matrix $A \in \mathbb{C}^{n\times n}$ can be expressed in Jordan canonical form as

$$Z^{-1}AZ = J = \operatorname{diag}(J_1, \ldots, J_p), \quad \text{for } J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k \times m_k},$$

where Z is nonsingular and $m_1 + m_2 + \ldots + m_p = n$. If each block in which the eigenvalue λ_k appears is of size 1 then λ_k is said to be a *semisimple* eigenvalue.

- This is a *theoretical object*, it is useful to prove and define *things*, not to implement *things*.

Matrix functions: the Jordan Canonical Form

Jordan Canonical Form

We recall that any matrix $A \in \mathbb{C}^{n \times n}$ can be expressed in Jordan canonical form as

$$Z^{-1}AZ = J = \operatorname{diag}(J_1, \ldots, J_p), \qquad \text{for } J_k = J_k(\lambda_k) = \begin{bmatrix} \lambda_k & 1 & & \\ & \lambda_k & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix} \in \mathbb{C}^{m_k \times m_k},$$

where Z is nonsingular and $m_1 + m_2 + \ldots + m_p = n$. If each block in which the eigenvalue λ_k appears is of size 1 then λ_k is said to be a *semisimple* eigenvalue.

- This is a *theoretical object*, it is useful to prove and define *things*, not to implement *things*.
- Now that we have a decomposition of the matrix, we need to introduce a suitable definition of **being defined on the spectrum**.

Matrix functions: the general case

Let us denote by $\lambda_1, \dots, \lambda_s$ the distinct eigenvalues of A, and by n_i the order of the largest Jordan block in which the λ_i appears, i.e., the *index* of the eigenvalue λ_i.

Defined on the spectrum

The function f is *defined on the spectrum of* A if the values

$$f^{(j)}(\lambda_i), \qquad j = 0, 1, \dots, n_i - 1, \quad i = 1, \dots s,$$

exist, where $f^{(j)}$ denotes the jth derivative of f, with $f^{(0)} = f$.

Matrix functions: the general case

Let us denote by $\lambda_1, \dots, \lambda_s$ the distinct eigenvalues of A, and by n_i the order of the largest Jordan block in which the λ_i appears, i.e., the *index* of the eigenvalue λ_i.

Defined on the spectrum

The function f is *defined on the spectrum of* A if the values

$$f^{(j)}(\lambda_i), \qquad j = 0, 1, \dots, n_i - 1, \quad i = 1, \dots s,$$

exist, where $f^{(j)}$ denotes the jth derivative of f, with $f^{(0)} = f$.

⚠ Again for the ML function and $\alpha > 0$ we have no problem with this.

Matrix functions: the general case

Matrix function

Lef f be defined on the spectrum of $A \in \mathbb{C}^{n\times n}$, which is represented in Jordan canonical form as $Z^{-1}AZ = J$,

$$f(A) = Zf(J)Z^{-1} = Z\,\mathrm{diag}(f(J_1), \ldots, f(J_p))Z^{-1},$$

where

$$f(J_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \ldots & \frac{f^{(m_k-1)}(\lambda_k)}{(m_k-1)!} \\ & f(\lambda_k) & \ddots & \vdots \\ & & \ddots & f'(\lambda_k) \\ & & & f(\lambda_k) \end{bmatrix}.$$

Moreover, let f be a multivalued function and suppose some eigenvalues occur in more than one Jordan block. If the same choice of branch of f is made in each block, then we say that $f(A)$ is a *primary matrix function*.

Matrix functions: computing $f(A)$ and $f(A)\mathbf{v}$

To march our scheme for

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

we need to compute operations of the form $f(A)\mathbf{v}$, *nevertheless*, we will have to compute $f(\cdot)$ at least on some small matrix.

Matrix functions: computing $f(A)$ and $f(A)\mathbf{v}$

To march our scheme for

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

we need to compute operations of the form $f(A)\mathbf{v}$, *nevertheless*, we will have to compute $f(\cdot)$ at least on some small matrix.

Schur decomposition and matrix functions

Given a matrix A there exist always a matrix Q such that $Q^*Q = I$, and a upper triangular matrix T such that $A = QTQ^*$. Then, if f is defined on the spectrum of A we can compute $f(A)$ as $f(A) = Qf(T)Q^*$.

Matrix functions: computing $f(A)$ and $f(A)\mathbf{v}$

To march our scheme for

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

we need to compute operations of the form $f(A)\mathbf{v}$, *nevertheless*, we will have to compute $f(\cdot)$ at least on some small matrix.

Schur decomposition and matrix functions

Given a matrix A there exist always a matrix Q such that $Q^*Q = I$, and a upper triangular matrix T such that $A = QTQ^*$. Then, if f is defined on the spectrum of A we can compute $f(A)$ as $f(A) = Qf(T)Q^*$.

But how do we compute the matrix function of an upper triangular matrix?

Matrix functions: the upper triangular case

Assumption we assume that T is such that each block $T_{i,j}$ has clustered eigenvalues, and distinct diagonal blocks have *far enough* eigenvalues.

If the **assumption** doesn't hold we look for a block permutation.

$$\left[\begin{array}{cc|cc} (T_{1,1})_{1,1} & (T_{1,1})_{1,2} & \multicolumn{2}{c}{\multirow{2}{*}{$T_{1,2}$}} \\ 0 & (T_{1,1})_{2,2} & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$\mathbf{0}$}} & (T_{2,2})_{1,1} & (T_{2,2})_{1,2} \\ & & 0 & (T_{2,2})_{2,2} \end{array}\right]$$

Close eigenvalues may lead to severe *accuracy loss*, even far apert eigenvalues can produce more inaccurate answers than expected, see (Davies and Higham 2003).

Matrix functions: the upper triangular case

Assumption we assume that T is such that each block $T_{i,j}$ has clustered eigenvalues, and distinct diagonal blocks have *far enough* eigenvalues.

❗ If the **assumption** doesn't hold we look for a block permutation.

$$\left[\begin{array}{cc|cc} (T_{1,1})_{1,1} & (T_{1,1})_{1,2} & \multicolumn{2}{c}{\multirow{2}{*}{$T_{1,2}$}} \\ 0 & (T_{1,1})_{2,2} & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$\mathbf{0}$}} & (T_{2,2})_{1,1} & (T_{2,2})_{1,2} \\ & & 0 & (T_{2,2})_{2,2} \end{array}\right]$$

- To evaluate $f(T_{ii})$ we use the Taylor series in σ

$$f(T_{i,i}) = \sum_{k=0}^{+\infty} \frac{f^{(k)}}{k!} M^k,$$

for $\sigma = {}^{\mathrm{trace}(T_{i,i})}/_{m}$, $m = \dim(T_{i,i})$, and $M = T_{i,i} - \sigma I$.

Matrix functions: the upper triangular case

Assumption we assume that T is such that each block $T_{i,j}$ has clustered eigenvalues, and distinct diagonal blocks have *far enough* eigenvalues.

❗ If the **assumption** doesn't hold we look for a block permutation.

$$\left[\begin{array}{cc|cc} (T_{1,1})_{1,1} & (T_{1,1})_{1,2} & \multicolumn{2}{c}{\multirow{2}{*}{$T_{1,2}$}} \\ 0 & (T_{1,1})_{2,2} & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$\mathbf{0}$}} & (T_{2,2})_{1,1} & (T_{2,2})_{1,2} \\ & & 0 & (T_{2,2})_{2,2} \end{array}\right]$$

For the off-diagonal blocks we apply the block-Parlett recurrence

$F_{i,i} = f(T_{i,i}),\ i = 1, \ldots, n$;
for $j = 2, \ldots, n$ **do**
  for $i = j-1, j-2, \ldots, 1$ **do**
    Solve Sylvester equation for $F_{i,j}$:
    $T_{i,i}F_{j,j} - F_{i,j}T_{j,j} = F_{i,i}T_{i,j} - T_{i,j}F_{j,j} + \sum_{k=0}^{j-1}(F_{i,k} - T_{k,j} - T_{i,k}F_{k,j})$.
  end
end

Matrix functions: the upper triangular case

Assumption we assume that T is such that each block $T_{i,j}$ has clustered eigenvalues, and distinct diagonal blocks have *far enough* eigenvalues.

❗ If the **assumption** doesn't hold we look for a block permutation.

$$\left[\begin{array}{cc|cc} (T_{1,1})_{1,1} & (T_{1,1})_{1,2} & \multicolumn{2}{c}{\multirow{2}{*}{$T_{1,2}$}} \\ 0 & (T_{1,1})_{2,2} & & \\ \hline \multicolumn{2}{c|}{\multirow{2}{*}{$\mathbf{0}$}} & (T_{2,2})_{1,1} & (T_{2,2})_{1,2} \\ & & 0 & (T_{2,2})_{2,2} \end{array}\right]$$

What we need

To use the algorithm we have sketched out, we need to be able to compute the derivatives of the ML function sufficiently accurately.

Derivatives of the ML function

The key observation for this task is

$$\frac{d^k}{dz^k} E_{\alpha,\beta}(z) = \sum_{j=0}^{+\infty} \frac{(j+k)_k z^j}{\Gamma(\alpha j + \alpha k + \beta)} = \frac{k!}{\Gamma(k+1)} \sum_{j=0}^{+\infty} \frac{\Gamma(j+k+1) z^j}{j!\Gamma(\alpha j + \alpha k + \beta)} = k! E_{\alpha,\alpha k+\beta}^{k+1}(z),$$

where

$$E_{\alpha,\beta}^{\gamma}(z) = \frac{1}{\Gamma(\gamma)} \sum_{j=0}^{+\infty} \frac{\Gamma(1+\gamma) z^j}{j!\Gamma(\alpha j + \beta)},$$

is called the Prabhakar function.

Derivatives of the ML function

The key observation for this task is

$$\frac{d^k}{dz^k}E_{\alpha,\beta}(z)=\sum_{j=0}^{+\infty}\frac{(j+k)_k z^j}{\Gamma(\alpha j+\alpha k+\beta)}=\frac{k!}{\Gamma(k+1)}\sum_{j=0}^{+\infty}\frac{\Gamma(j+k+1)z^j}{j!\Gamma(\alpha j+\alpha k+\beta)}=k!E^{k+1}_{\alpha,\alpha k+\beta}(z),$$

where

$$E^{\gamma}_{\alpha,\beta}(z)=\frac{1}{\Gamma(\gamma)}\sum_{j=0}^{+\infty}\frac{\Gamma(1+\gamma)z^j}{j!\Gamma(\alpha j+\beta)},$$

is called the Prabhakar function.

Its **efficient computation** can be obtained, similarly to the ML function, by means of a *Laplace transform inversion*

$$\mathcal{L}\left\{t^{\beta-1}E^{\gamma}_{\alpha,\beta}(t^{\alpha}z)\right\}(s)=\frac{s^{\alpha\gamma-\beta}}{(s^{\alpha}-t^{\alpha}z)^{\gamma}},\quad \Re(s)>0,\quad |t^{\alpha}zs^{-\alpha}|<1.$$

Computing the Prabhakar function (Garrappa 2015)

We select $t = 1$ in

$$\mathcal{L}\left\{t^{\beta-1}E^{\gamma}_{\alpha,\beta}(t^{\alpha}z)\right\}(s) = \frac{s^{\alpha\gamma-\beta}}{(s^{\alpha}-t^{\alpha}z)^{\gamma}}, \quad \Re(s) > 0, \quad |t^{\alpha}zs^{-\alpha}| < 1.$$

Computing the Prabhakar function (Garrappa 2015)

Having selected $t = 1$ we have

$$\mathcal{L}\left\{E^{\gamma}_{\alpha,\beta}(z)\right\}(s) = \frac{s^{\alpha\gamma-\beta}}{(s^{\alpha}-z)^{\gamma}}, \quad \Re(s) > 0, \quad |zs^{-\alpha}| < 1.$$

Computing the Prabhakar function (Garrappa 2015)

Having selected $t = 1$ we have

$$\mathcal{L}\left\{E^{\gamma}_{\alpha,\beta}(z)\right\}(s) = \frac{s^{\alpha\gamma-\beta}}{(s^{\alpha}-z)^{\gamma}}, \quad \Re(s) > 0, \quad |zs^{-\alpha}| < 1,\ H_k(z;z) = \frac{s^{\alpha-\beta}}{(s^{\alpha}-z)^{k+1}}.$$

- Since

$$\frac{d^k}{dz^k}E_{\alpha,\beta}(z) = k!E^{k+1}_{\alpha,\alpha k+\beta}(z) = \frac{k!}{2\pi i}\int_{\mathcal{C}} e^{s}H_k(s;z)\mathrm{d}s \equiv I_k(z),$$

Computing the Prabhakar function (Garrappa 2015)

Having selected $t = 1$ we have

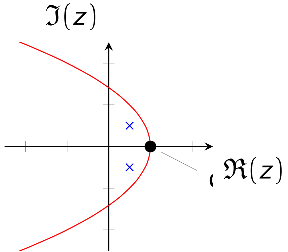
$$\mathcal{L}\left\{E^{\gamma}_{\alpha,\beta}(z)\right\}(s) = \frac{s^{\alpha\gamma-\beta}}{(s^{\alpha}-z)^{\gamma}}, \quad \Re(s) > 0, \quad |zs^{-\alpha}| < 1, \; H_k(z;z) = \frac{s^{\alpha-\beta}}{(s^{\alpha}-z)^{k+1}}.$$

- Since

$$\frac{d^k}{dz^k}E_{\alpha,\beta}(z) = k!E^{k+1}_{\alpha,\alpha k+\beta}(z) = \frac{k!}{2\pi i}\int_{\mathcal{C}} e^{s}H_k(s;z)\mathrm{d}s \equiv I_k(z),$$

- we use the *Optimal Parabolic Contour* we have already discussed in **Lecture 2** to determine the deformation of the Bromwich line to evaluate

$$I_k^{[N]} = \frac{k!h}{2\pi i}\sum_{j=-N}^{N} e^{\sigma(u_j)}H_k(\sigma(u_j);z)\sigma'(u_j).$$



$s = \mu(iu+1)^2$, $-\infty < u < \infty$

An alternative option (Higham and Liu 2021)

We needed the ML derivatives to apply Schur-Parlett to non-diagonalizable matrices.

An alternative option (Higham and Liu 2021)

We needed the ML derivatives to apply Schur-Parlett to non-diagonalizable matrices.

Diagonalization by perturbation

Let A be nonnormal

$$\tilde{A} = A + E$$

for E a suitable perturbation is *likely to be diagonalizable*. **Diagonalizable matrices are dense in** $\mathbb{C}^{n\times n}$, for a given A and machine precision ϵ then the best approximate diagonalization can be measured in terms of

$$\sigma(A,\epsilon) = \inf_{E,V} \sigma(A,V,E,\epsilon) = \inf_{E,V} \{\kappa_2(V)\epsilon + \|E\|_2\}.$$

An alternative option (Higham and Liu 2021)

We needed the ML derivatives to apply Schur-Parlett to non-diagonalizable matrices.

Diagonalization by perturbation

Let A be nonnormal

$$\tilde{A} = A + E$$

for E a suitable perturbation is *likely to be diagonalizable*. **Diagonalizable matrices are dense in** $\mathbb{C}^{n\times n}$, for a given A and machine precision ϵ then the best approximate diagonalization can be measured in terms of

$$\sigma(A,\epsilon) = \inf_{E,V} \sigma(A,V,E,\epsilon) = \inf_{E,V} \{\kappa_2(V)\epsilon + \|E\|_2\}.$$

We can expect to measure on $f(A)$ by estimating

$$\|f(A+E) - f(A)\| \lesssim \|L_f(A,E)\| \leq \|L_f(A)\|\|E\|,$$

for $L_f(A,E)$ the Fréchet derivative of f at A in direction E, $\|L_f(A)\| = \max_{\|E\|=1} \{\|L_f(A,E)\|\}$.

An alternative option (Higham and Liu 2021)

Fréchet derivative

The **Fréchet derivative** of a matrix function $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ at a point $X \in \mathbb{C}^{n\times n}$ is a linear mapping $L : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ $E \mapsto L_f(X, E)$ such that for all $E \in \mathbb{C}^{n\times n}$ we find

$$f(X + E) - f(X) - L(X, E) = o(\|E\|).$$

An alternative option (Higham and Liu 2021)

Fréchet derivative

The **Fréchet derivative** of a matrix function $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ at a point $X \in \mathbb{C}^{n\times n}$ is a linear mapping $L : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ $E \mapsto L_f(X, E)$ such that for all $E \in \mathbb{C}^{n\times n}$ we find

$$f(X + E) - f(X) - L(X, E) = o(\|E\|).$$

Thus, in our estimate we have

$$\|f(A + E) - f(A)\| \lesssim \|L_f(A, E)\| \leq \|L_f(A)\|\|E\|,$$

and therefore the change in f induced by E grows as $\|L_f(A)\|_2\|E\|_2$ and there are many cases in which $\|L_f(A)\|_2 \gg 1$.

An alternative option (Higham and Liu 2021)

Fréchet derivative

The **Fréchet derivative** of a matrix function $f : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ at a point $X \in \mathbb{C}^{n\times n}$ is a linear mapping $L : \mathbb{C}^{n\times n} \to \mathbb{C}^{n\times n}$ $E \mapsto L_f(X, E)$ such that for all $E \in \mathbb{C}^{n\times n}$ we find

$$f(X + E) - f(X) - L(X, E) = o(\|E\|).$$

Thus, in our estimate we have

$$\|f(A + E) - f(A)\| \lesssim \|L_f(A, E)\| \leq \|L_f(A)\|\|E\|,$$

and therefore the change in f induced by E grows as $\|L_f(A)\|_2\|E\|_2$ and there are many cases in which $\|L_f(A)\|_2 \gg 1$.

- The idea from (Higham and Liu 2021) is to use a *structured perturbation*:
  "take E to be upper triangular standard Gaussian matrix."

An alternative option (Higham and Liu 2021)

The idea in few steps

1. Compute the Schur decomposition $A = QTQ^*$,

An alternative option (Higham and Liu 2021)

The idea in few steps

1. Compute the Schur decomposition $A = QTQ^*$,
2. Consider the perturbed matrices $\tilde{T} = T + E$

An alternative option (Higham and Liu 2021)

The idea in few steps

1. Compute the Schur decomposition $A = QTQ^*$,
2. Consider the perturbed matrices $\tilde{T} = T + E$
   - $\tilde{T}$ is still upper triangular,
   - Eigenvectors can be compute by back-substitution: $(\tilde{T} - \tilde{t}_{i,i} I)\mathbf{v}_i = 0$, $i = 1, \ldots, m$,

An alternative option (Higham and Liu 2021)

The idea in few steps

1. Compute the Schur decomposition $A = QTQ^*$,
2. Consider the perturbed matrices $\tilde{T} = T + E$
   - $\tilde{T}$ is still upper triangular,
   - Eigenvectors can be compute by back-substitution: $(\tilde{T} - \tilde{t}_{i,i} I)\mathbf{v}_i = 0$, $i = 1, \ldots, m$,
3. Compute in precision u_h the diagonalization

$$\tilde{T} = VDV^{-1}, \quad D = \mathrm{diag}(\lambda_i),$$

with **distinct** λ_i,

An alternative option (Higham and Liu 2021)

The idea in few steps

1. Compute the Schur decomposition $A = QTQ^*$,
2. Consider the perturbed matrices $\tilde{T} = T + E$
   - $\tilde{T}$ is still upper triangular,
   - Eigenvectors can be compute by back-substitution: $(\tilde{T} - \tilde{t}_{i,i} I)\mathbf{v}_i = 0$, $i = 1, \dots, m$,
3. Compute in precision u_h the diagonalization

$$\tilde{T} = VDV^{-1}, \quad D = \mathrm{diag}(\lambda_i),$$

   with **distinct** λ_i,
4. Form $f(\tilde{T}) = Vf(D)V^{-1}$ in precision u_h

An alternative option (Higham and Liu 2021)

The idea in few steps

1. Compute the Schur decomposition $A = QTQ^*$,
2. Consider the perturbed matrices $\tilde{T} = T + E$
   - $\tilde{T}$ is still upper triangular,
   - Eigenvectors can be compute by back-substitution: $(\tilde{T} - \tilde{t}_{i,i}I)\mathbf{v}_i = 0$, $i = 1, \ldots, m$,
3. Compute in precision u_h the diagonalization

$$\tilde{T} = VDV^{-1}, \quad D = \mathrm{diag}(\lambda_i),$$

   with **distinct** λ_i,
4. Form $f(\tilde{T}) = Vf(D)V^{-1}$ in precision u_h

What precision do we need?

To have $\kappa_1(V)u_h \lesssim u$ we select for $c_m u \approx \min_i |\mathrm{diag}(\tilde{t}_{1,1}I - \tilde{T}_{2,2})|$

$$u_h \lesssim \frac{c_m u^2}{\max_{i<j} |\tilde{t}_{i,j}| \left(\max_{i<j} |\tilde{t}_{i,j}|/c_m u + 1\right)^{k-2}}, \quad k = \text{“size of the Jordan block”} \geq 2.$$

From small to large matrices

We now know how to compute $E_{\alpha,\beta}(A)$ for a *small matrix A*, either with

- Classical Schur-Parlett algorithm with Laplace inversion technique for the needed derivative of the ML function (Garrappa and Popolizio 2018),
  - https://it.mathworks.com/matlabcentral/fileexchange/66272-mittag-leffler-function-with-matrix-arguments
- Multiprecision derivative-free Schur-Parlett algorithm (Higham and Liu 2021),
  - https://github.com/Xiaobo-Liu/mp-spalg

From small to large matrices

We now know how to compute $E_{\alpha,\beta}(A)$ for a *small matrix* A, either with

- Classical Schur-Parlett algorithm with Laplace inversion technique for the needed derivative of the ML function (Garrappa and Popolizio 2018),
  - https://it.mathworks.com/matlabcentral/fileexchange/66272-mittag-leffler-function-with-matrix-arguments
- Multiprecision derivative-free Schur-Parlett algorithm (Higham and Liu 2021),
  - https://github.com/Xiaobo-Liu/mp-spalg

What about *large matrices*?

From small to large matrices

We now know how to compute $E_{\alpha,\beta}(A)$ for a *small matrix* A, either with

- Classical Schur-Parlett algorithm with Laplace inversion technique for the needed derivative of the ML function (Garrappa and Popolizio 2018),
  - https://it.mathworks.com/matlabcentral/fileexchange/66272-mittag-leffler-function-with-matrix-arguments
- Multiprecision derivative-free Schur-Parlett algorithm (Higham and Liu 2021),
  - https://github.com/Xiaobo-Liu/mp-spalg

What about *large matrices*?

Projection methods for matrix functions

We can exploit the *subspace projection* idea, take $V \in \mathbb{R}^{n\times k}$ spanning a given subspace $\mathcal{W}_k$

$$f(A)\mathbf{v} \approx Vf(V^TAV)V^T\mathbf{v} \qquad V^TAV \in \mathbb{R}^{k\times k}, \quad k \ll n.$$

Krylov Projection Methods

Different methods are obtained for **different** choices of the **projection spaces** $\mathcal{W}_k(A, \mathbf{v})$.

Krylov Projection Methods

Different methods are obtained for **different** choices of the **projection spaces** $\mathcal{W}_k(A, \mathbf{v})$.

A general framework

Given a set of scalars $\{\sigma_1, \ldots, \sigma_{k-1}\} \subset \overline{\mathbb{C}}$ (the extended complex plane), that are not eigenvalues of A, let

$$q_{k-1}(z) = \prod_{j=1}^{k-1} (\sigma_j - z).$$

The rational Krylov subspace of order k associated with A, $\mathbf{v}$ and q_{k-1} is defined by

$$\mathcal{Q}_k(A, \mathbf{v}) = [q_{k-1}(A)]^{-1}\, \mathcal{K}_k(A, \mathbf{v}), \qquad \mathcal{K}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, A\mathbf{v}, \ldots, A^{k-1}\mathbf{v}\}.$$

Krylov Projection Methods

Different methods are obtained for **different** choices of the **projection spaces** $\mathcal{W}_k(A, \mathbf{v})$.

A general framework

Given a set of scalars $\{\sigma_1, \ldots, \sigma_{k-1}\} \subset \overline{\mathbb{C}}$ (the extended complex plane), that are not eigenvalues of A, let

$$q_{k-1}(z) = \prod_{j=1}^{k-1} (\sigma_j - z).$$

The rational Krylov subspace of order k associated with A, $\mathbf{v}$ and q_{k-1} is defined by

$$\mathcal{Q}_k(A, \mathbf{v}) = [q_{k-1}(A)]^{-1} \mathcal{K}_k(A, \mathbf{v}), \qquad \mathcal{K}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, A\mathbf{v}, \ldots, A^{k-1}\mathbf{v}\}.$$

A matrix expression

Given $\{\mu_1, \ldots, \mu_{k-1}\} \subset \overline{\mathbb{C}}$ such that $\sigma_j \neq \mu_j^{-2}$, we define the matrices

$$C_j = (\mu_j \sigma_j A - I)\,(\sigma_j I - A)^{-1}, \text{ and } \mathcal{Q}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, C_1\mathbf{v}, \ldots, C_{k-1} \cdots C_2 C_1 \mathbf{v}\}.$$

Krylov Projection Methods: special cases

A matrix expression

Given $\{\mu_1, \dots, \mu_{k-1}\} \subset \overline{\mathbb{C}}$ such that $\sigma_j \neq \mu_j^{-2}$, we define the matrices

$$C_j = (\mu_j \sigma_j A - I)(\sigma_j I - A)^{-1}, \text{ and } \mathcal{Q}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, C_1\mathbf{v}, \dots, C_{k-1} \cdots C_2 C_1 \mathbf{v}\}.$$

Polynomial Krylov $\mathcal{W}_k(A, \mathbf{v}) = \mathcal{K}_k(A, \mathbf{v})$ set $\mu_j = 1$ and $\sigma_j = \infty$ for each j,

Krylov Projection Methods: special cases

A matrix expression

Given $\{\mu_1, \ldots, \mu_{k-1}\} \subset \overline{\mathbb{C}}$ such that $\sigma_j \neq \mu_j^{-2}$, we define the matrices

$$C_j = (\mu_j \sigma_j A - I)\,(\sigma_j I - A)^{-1}, \text{ and } \mathcal{Q}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, C_1\mathbf{v}, \ldots, C_{k-1} \cdots C_2 C_1 \mathbf{v}\}.$$

Polynomial Krylov $\mathcal{W}_k(A, \mathbf{v}) = \mathcal{K}_k(A, \mathbf{v})$ set $\mu_j = 1$ and $\sigma_j = \infty$ for each j,

Extended Krylov $\mathcal{W}_{2k-1}(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, A^{-1}\mathbf{v}, A\mathbf{v}, \ldots, A^{-(k-1)}\mathbf{v}, A^{k-1}\mathbf{v}\}$, set

$$(\mu_j, \sigma_j) = \begin{cases} (1, \infty), & \text{for } j \text{ even}, \\ (0, 0), & \text{for } j \text{ odd}. \end{cases}$$

Krylov Projection Methods: special cases

A matrix expression

Given $\{\mu_1, \ldots, \mu_{k-1}\} \subset \overline{\mathbb{C}}$ such that $\sigma_j \neq \mu_j^{-2}$, we define the matrices

$$C_j = (\mu_j \sigma_j A - I)\,(\sigma_j I - A)^{-1}, \text{ and } \mathcal{Q}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, C_1\mathbf{v}, \ldots, C_{k-1}\cdots C_2 C_1 \mathbf{v}\}.$$

Polynomial Krylov $\mathcal{W}_k(A, \mathbf{v}) = \mathcal{K}_k(A, \mathbf{v})$ set $\mu_j = 1$ and $\sigma_j = \infty$ for each j,

Extended Krylov $\mathcal{W}_{2k-1}(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, A^{-1}\mathbf{v}, A\mathbf{v}, \ldots, A^{-(k-1)}\mathbf{v}, A^{k-1}\mathbf{v}\}$, set

$$(\mu_j, \sigma_j) = \begin{cases} (1, \infty), & \text{for } j \text{ even}, \\ (0, 0), & \text{for } j \text{ odd}. \end{cases}$$

Shift-And-Invert $\mathcal{W}_k(A, \mathbf{v}) = \mathrm{Span}\{\mathbf{v}, (\sigma I - A)^{-1}\mathbf{v}, \ldots, (\sigma I - A)^{-(k-1)}\mathbf{v}\}$, take $\mu_j = 0$ and $\sigma_j = \sigma$ for each j,

The ML function (Moret and Novati 2011)

To estimate the convergence behavior of general projection methods in the non-normal we need the concept of **field of values** (or *numerical range*.)

.

The ML function (Moret and Novati 2011)

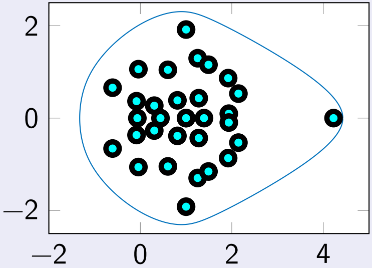
To estimate the convergence behavior of general projection methods in the non-normal we need the concept of **field of values** (or *numerical range*.)

Field of Values/Numerical Range

Given $A \in \mathbb{C}^{N \times N}$ we denote its **field of values** as

$$W(A) = \left\{ \frac{\langle \mathbf{x}, A\mathbf{x} \rangle}{\langle \mathbf{x}, \mathbf{x} \rangle}, \quad \mathbf{0} \neq \mathbf{x} \in \mathbb{C}^N \right\},$$

where $\langle \cdot, \cdot \rangle$ represents the Euclidean inner product.



The ML function (Moret and Novati 2011)

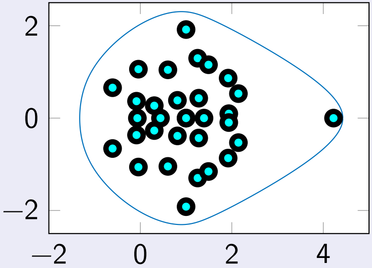
To estimate the convergence behavior of general projection methods in the non-normal we need the concept of **field of values** (or *numerical range*.)

Field of Values/Numerical Range

Given $A \in \mathbb{C}^{N\times N}$ we denote its **field of values** as

$$W(A) = \left\{ \frac{\langle \mathbf{x}, A\mathbf{x}\rangle}{\langle \mathbf{x}, \mathbf{x}\rangle}, \quad \mathbf{0} \neq \mathbf{x} \in \mathbb{C}^N \right\},$$

where $\langle \cdot, \cdot \rangle$ represents the Euclidean inner product.



It has many **properties**, *e.g.*, $W(A) \subseteq D(0, \|A\|)$ (disk centered on 0 with radius $\|A\|$), is *compact*, sub-additive $W(A+B) \subseteq W(A) + W(B)$, unitarily invariant $W(UAU^H) = UW(A)U^H$, *etc.* see (Benzi 2021).

The ML function (Moret and Novati 2011)

Assumptions:

(A1) We assume that $\exists a > 0$, $\theta \in [0, \pi/2)$ such that

$$W(A) \subset \Sigma_{\theta,a} = \{\lambda \in \mathbb{C} \, : \, |\arg(\lambda) - a| \leq \theta\}.$$

(A2) $\beta > 0$, $\alpha \in (0, 2)$ be such that $\alpha\pi/2 < \pi - \theta$, $\varepsilon > 0$ and

$$\frac{\alpha\pi}{2} < \mu \leq \min\{\pi, \alpha\pi\}, \quad \mu < \pi - \theta.$$

Method of choice: we use polynomial Krylov method $\mathcal{K}_m(A, \mathbf{v})$:

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} \mathbf{e}_m^T, \quad \mathrm{Span}\, V_m = \mathrm{Span}\{\mathbf{v}_i\}_{i=1}^m = \mathcal{K}_m(A, \mathbf{v}), \quad H_m = V_m^H A V_m.$$

We want to bound:

$$R_m = E_{\alpha,\beta}(-A)\mathbf{v} - V_m E_{\alpha,\beta}(-H_m)\mathbf{e}_1, \quad m \geq 1.$$

The ML function (Moret and Novati 2011)

We first express the error in *integral form*, starting from (Podlubny 1999, Theorem 1.1)

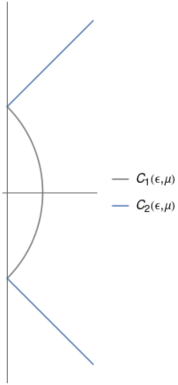
$$E_{\alpha,\beta}(z) = \frac{1}{2\alpha\pi i}\int_{C(\varepsilon,\mu)} \frac{\exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}}{\lambda - z}\,\mathrm{d}\lambda, \quad z \in G^{-}(\varepsilon,\mu),$$

where

- $\forall\, \varepsilon > 0,\ 0 < \mu < \pi$

$$C(\varepsilon,\mu) = \bigcup \begin{cases} C_1(\varepsilon,\mu) = \{\lambda \,:\, \lambda = \varepsilon \exp(i\varphi), \quad -\mu \le \varphi \le \mu\}, \\ C_2(\varepsilon,\mu) = \{\lambda \,:\, \lambda = r \exp(\pm i\mu), \quad r \ge \varepsilon\}. \end{cases}$$

- The contour $C(\varepsilon,\mu)$ divides the complex plane into two domains, $G^{-}(\varepsilon,\mu)$ and $G^{+}(\varepsilon,\mu)$ lying respectively on the left and on the right of $C(\varepsilon,\mu)$.



An Expression for the Error

From the previous we find

$$E_{\alpha,\beta}(-A) = \frac{1}{2\alpha\pi i}\int_{C(\varepsilon,\mu)} \exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}(\lambda I + A)^{-1}\,\mathrm{d}\lambda, \quad \sigma(-A) \in G^{-}(\varepsilon,\mu),$$

and together with

$$R_m = E_{\alpha,\beta}(-A)\mathbf{v} - V_m E_{\alpha,\beta}(-H_m)\mathbf{e}_1, \quad m \geq 1,$$

we write

$$R_m = \frac{1}{2\alpha\pi i}\int_{C(\varepsilon,\mu)} \exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}\delta_m(\lambda), \mathrm{d}\lambda,$$

for

$$\begin{aligned}\delta_m(\lambda) &= (\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1}\mathbf{e}_1\\ &= (\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1}V_m^H\mathbf{v}.\end{aligned}$$

An Expression for the Error

Observe now that

$$\delta_m(\lambda) = (\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1}V_m^H\mathbf{v} = \Delta_m\mathbf{v},$$

An Expression for the Error

Observe now that

$$\delta_m(\lambda) = (\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1} V_m^H \mathbf{v} = \Delta_m \mathbf{v},$$

By using the Arnoldi relation, since $\mathbf{v}_{m+1} \perp V_m$:

$$V_m^H(\lambda I + A) V_m = \lambda I + H_m,$$

An Expression for the Error

Observe now that

$$\delta_m(\lambda) = (\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1} V_m^H \mathbf{v} = \Delta_m \mathbf{v},$$

By using the Arnoldi relation, since $\mathbf{v}_{m+1} \perp V_m$:

$$V_m^H (\lambda I + A) V_m = \lambda I + H_m,$$

Therefore we have

$$\Delta_m (\lambda I + A) V_m = 0.$$

An Expression for the Error

Observe now that

$$\delta_m(\lambda) = (\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1} V_m^H \mathbf{v} = \Delta_m \mathbf{v},$$

By using the Arnoldi relation, since $\mathbf{v}_{m+1} \perp V_m$:

$$V_m^H (\lambda I + A) V_m = \lambda I + H_m,$$

Therefore we have

$$\Delta_m (\lambda I + A) V_m = 0.$$

For an arbitrary $\mathbf{y} \in \mathbb{C}^m$ we have then

$$(\lambda I + A)^{-1}\mathbf{v} - V_m(\lambda I + H_m)^{-1} V_m^H \mathbf{v} = \Delta_m(\mathbf{v} - (\lambda I + A) V_m \mathbf{y}) = \Delta_m p_m(A)\mathbf{v},$$

where $p_m(z)$ is a polynomial of degree $\leq m$ with $p_m(-\lambda) = 1$.

An Expression for the Error

We have therefore proved that

$$\|\delta_m(A)\| \leq \|(\lambda I + A)^{-1} - V_m(\lambda I + H_m)^{-1} V_m^H\| \|p_m(A)\mathbf{v}\|, \forall p_m \in \mathbb{P}_{\leq m}[z] \text{ with } p_m(-\lambda) = 1.$$

By using (Diele, Moret, and Ragni 2008/09, Lemma 2) we also have the following expression

$$\|\delta_m(\lambda)\| = \frac{\prod_{j=1}^{m} h_{j+1,j}}{|\det(\lambda I + H_m)|} \|(\lambda I + A)^{-1}\mathbf{v}_{m+1}\|.$$

An Expression for the Error

We have therefore proved that

$$\|\delta_m(A)\| \leq \|(\lambda I + A)^{-1} - V_m(\lambda I + H_m)^{-1} V_m^H\| \|p_m(A)\mathbf{v}\|, \forall p_m \in \mathbb{P}_{\leq m}[z] \text{ with } p_m(-\lambda) = 1.$$

By using (Diele, Moret, and Ragni 2008/09, Lemma 2) we also have the following expression

$$\|\delta_m(\lambda)\| = \frac{\prod_{j=1}^{m} h_{j+1,j}}{|\det(\lambda I + H_m)|} \|(\lambda I + A)^{-1}\mathbf{v}_{m+1}\|.$$

To obtain the first bound we call then

$$D(\lambda) = \operatorname{dist}(\lambda, W(-A)) \quad \forall \lambda \in C(\varepsilon, \mu).$$

An Expression for the Error

We have therefore proved that

$$\|\delta_m(A)\| \leq \|(\lambda I + A)^{-1} - V_m(\lambda I + H_m)^{-1} V_m^H\| \|p_m(A)\mathbf{v}\|, \forall p_m \in \mathbb{P}_{\leq m}[z] \text{ with } p_m(-\lambda) = 1.$$

By using (Diele, Moret, and Ragni 2008/09, Lemma 2) we also have the following expression

$$\|\delta_m(\lambda)\| = \frac{\prod_{j=1}^{m} h_{j+1,j}}{|\det(\lambda I + H_m)|} \|(\lambda I + A)^{-1}\mathbf{v}_{m+1}\|.$$

To obtain the first bound we call then

$$D(\lambda) = \operatorname{dist}(\lambda, W(-A)) \quad \forall \lambda \in C(\varepsilon, \mu).$$

Representation function

Using (A1) and (A2) we can find a function $\nu(\varphi)$ such that

$$\forall \lambda = |\lambda| \exp(\pm i\varphi) \in C(\varepsilon, \mu) \quad D(\lambda) \geq \nu(\varphi)|\lambda|, \quad \nu(\varphi) \geq \nu > 0.$$

A First Error Bound

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M) \prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos(\mu/\alpha)| + 1))}{m\alpha - 1 + \beta} \right).$$

A First Error Bound

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos(\mu/\alpha)| + 1))}{m\alpha - 1 + \beta} \right).$$

Proof. We use $\|(\lambda I + A)^{-1}\| \leq D(\lambda)^{-1}$ and $W(H_m) \subseteq W(A)$ in the error expression R_m

$$\begin{aligned}\|R_m\| &= \left\| \frac{1}{2\alpha\pi i} \int_{C(\varepsilon,\mu)} \exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}\delta_m(\lambda), \mathrm{d}\lambda \right\| \\ &\leq \frac{\prod_{j=1}^{m} h_{j+1,j}}{2\pi\alpha} \int_{C(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}\right|}{D(\lambda)^{m+1}} |\mathrm{d}\lambda|.\end{aligned}$$

A First Error Bound

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos(\mu/\alpha)|+1))}{m\alpha - 1 + \beta} \right).$$

Proof. We use $\|(\lambda I + A)^{-1}\| \leq D(\lambda)^{-1}$ and $W(H_m) \subseteq W(A)$ in the error expression R_m

$$\|R_m\| \leq \frac{\prod_{j=1}^{m} h_{j+1,j}}{2\pi\alpha}(I_1 + I_2),$$

with

$$I_1 = \int_{C_1(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}\right|}{D(\lambda)^{m+1}} |\mathrm{d}\lambda| \leq 2\varepsilon^{\frac{1-\beta}{\alpha}-m} \int_0^{\mu} \frac{\exp(\varepsilon^{1/\alpha}\cos(\varphi/\alpha))}{\nu(\varphi)^{m+1}} \,\mathrm{d}\varphi,$$

A First Error Bound

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos(\mu/\alpha)|+1))}{m\alpha - 1 + \beta} \right).$$

Proof. We use $\|(\lambda I + A)^{-1}\| \leq D(\lambda)^{-1}$ and $W(H_m) \subseteq W(A)$ in the error expression R_m

$$\|R_m\| \leq \frac{\prod_{j=1}^{m} h_{j+1,j}}{2\pi\alpha} \left(2\varepsilon^{\frac{1-\beta}{\alpha}-m} \int_0^{\mu} \frac{\exp(\varepsilon^{1/\alpha}\cos(\varphi/\alpha))}{\nu(\varphi)^{m+1}}\,\mathrm{d}\varphi + I_2 \right),$$

with

$$I_2 = \int_{C_2(\varepsilon,\mu)} \frac{\left|\exp(\lambda^{1/\alpha})\lambda^{1-\beta/\alpha}\right|}{D(\lambda)^{m+1}}\,|\mathrm{d}\lambda| \leq \frac{2}{\nu^{m+1}} \int_{\varepsilon}^{+\infty} \frac{r^{\frac{1-\beta}{\alpha}} \exp(-r^{\frac{1}{\alpha}}|\cos(\mu/\alpha)|)}{r^{m+1}}\,\mathrm{d}r$$

$$= \frac{2}{\nu^{m+1}} \int_{\varepsilon^{1/\alpha}}^{+\infty} \frac{\exp(-s|\cos(\mu/\alpha)|)}{s^{m\alpha+\beta}}\,\mathrm{d}s \leq \frac{2\alpha\exp(-\varepsilon^{1/\alpha}|\cos(\mu/\alpha)|)}{(m\alpha+\beta-1)\nu^{m+1}\varepsilon^{\frac{m\alpha+\beta-1}{\alpha}}}.$$

A First Error Bound

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos(\mu/\alpha)| + 1))}{m\alpha - 1 + \beta} \right).$$

Proof. We use $\|(\lambda I + A)^{-1}\| \leq D(\lambda)^{-1}$ and $W(H_m) \subseteq W(A)$ in the error expression R_m

$$\|R_m\| \leq \frac{\prod_{j=1}^{m} h_{j+1,j}}{2\pi\alpha} \left(2\varepsilon^{\frac{1-\beta}{\alpha}-m} \int_0^{\mu} \frac{\exp(\varepsilon^{1/\alpha}\cos(\varphi/\alpha))}{\nu(\varphi)^{m+1}} \,\mathrm{d}\varphi + \frac{2\alpha \exp(-\varepsilon^{1/\alpha}|\cos(\mu/\alpha)|)}{(m\alpha + \beta - 1)\nu^{m+1}\varepsilon^{\frac{m\alpha+\beta-1}{\alpha}}} \right)$$

The result follows then by setting $\varepsilon = M^{\alpha}$ and simplifying the expression. □

A First Error Bound

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos({}^{\mu}\!/\!{}_{\alpha})|+1))}{m\alpha - 1 + \beta} \right).$$

Proof. We use $\|(\lambda I + A)^{-1}\| \leq D(\lambda)^{-1}$ and $W(H_m) \subseteq W(A)$ in the error expression R_m

$$\|R_m\| \leq \frac{\prod_{j=1}^{m} h_{j+1,j}}{2\pi\alpha} \left(2\varepsilon^{\frac{1-\beta}{\alpha}-m} \int_0^{\mu} \frac{\exp(\varepsilon^{1/\alpha}\cos({}^{\varphi}\!/\!{}_{\alpha}))}{\nu(\varphi)^{m+1}}\, d\varphi + \frac{2\alpha \exp(-\varepsilon^{1/\alpha}|\cos({}^{\mu}\!/\!{}_{\alpha})|)}{(m\alpha+\beta-1)\nu^{m+1}\varepsilon^{\frac{m\alpha+\beta-1}{\alpha}}} \right)$$

The result follows then by setting $\varepsilon = M^{\alpha}$ and simplifying the expression. □

⚠ With the same proof another bound for the case of small α can be obtained.

A First Error Bound: small αs

Theorem (Moret and Novati 2011, Theorem 3.2)

Let assumptions (A1) and (A2) hold, then for $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{\pi \nu^{m+1} M^{m\alpha+\beta-1}} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(|\cos(\mu/\alpha)| + 1))}{m\alpha - 1 + \beta} \right).$$

Corollary (Moret and Novati 2011, Corollary 3.3)

Let assumptions (A1) and (A2) hold. Let $m \geq 1$ be such that $m\alpha + \beta > 0$, then for every $M > 0$, we have

$$\|R_m\| \leq \frac{\exp(M)\prod_{j=1}^{m} h_{j+1,j}}{4\nu^{m+1} M^{m\alpha}} \frac{4M^{1-\beta}}{\pi} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(1 + |\cos(\mu/\alpha)|))}{M|\cos(\mu/\alpha)|} \right).$$

A First Error Bound: some observations

- The ML function is entire for $\alpha > 0 \Rightarrow$ superlinear convergence for large enough m:

$$M = m\alpha + \beta - 1 \;\Rightarrow\; \|R_m\| \propto \left(\frac{\exp(1)}{M}\right)^{M} \nu^{-(m+1)} \prod_{j=1}^{m} h_{j+1,j}.$$

A First Error Bound: some observations

- The ML function is entire for $\alpha > 0 \Rightarrow$ superlinear convergence for large enough m:

$$M = m\alpha + \beta - 1 \;\Rightarrow\; \|R_m\| \propto \left(\frac{\exp(1)}{M}\right)^M \nu^{-(m+1)} \prod_{j=1}^{m} h_{j+1,j}.$$

- To better understand this, we use that for every monic polynomial of degree m we find

$$\prod_{j=1}^{m} h_{j+1,j} \leq \|q_m(A)v\|,$$

  Therefore, if we take q_m as the **monic Faber polynomial** associated to a closed convex subset $\Omega \supset W(-A)$ we get the bound in terms of the **logarithmic capacity** γ of Ω.

A First Error Bound: some observations

- The ML function is entire for $\alpha > 0 \Rightarrow$ superlinear convergence for large enough m:

$$M = m\alpha + \beta - 1 \;\Rightarrow\; \|R_m\| \propto \left(\frac{\exp(1)}{M}\right)^M \nu^{-(m+1)} \prod_{j=1}^{m} h_{j+1,j}.$$

- To better understand this, we use that for every monic polynomial of degree m we find

$$\prod_{j=1}^{m} h_{j+1,j} \leq 2\gamma^m,$$

  Therefore, if we take q_m as the **monic Faber polynomial** associated to a closed convex subset $\Omega \supset W(-A)$ we get the bound in terms of the **logarithmic capacity** γ of Ω.

$\Rightarrow$ we have discovered:

$$\|R_m\| \propto \left(\frac{\exp(1)}{m\alpha}\right)^{m\alpha} \left(\frac{\gamma}{\nu}\right)^m.$$

Specialized bounds

The bound can be refined under stricter hypotheses.

Specialized bounds

The bound can be refined under stricter hypotheses.

Theorem (Moret and Novati 2011, Theorem 3.5)

Assume that A is Hermitian with $\sigma(A) \subseteq [a, b] \subset [0, +\infty)$. Assume that $0 < \alpha < 1$, $\beta \geq \alpha$. Let $\mu \leq \pi/2$, $\alpha\pi/2 < \mu < \alpha\pi$. Then for every index $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{4M^{1-\beta}}{\pi} \left(\frac{\mu}{\alpha} + \frac{\exp(-M(1 + |\cos(\mu/\alpha)|))}{M|\cos(\mu/\alpha)|} \right) \exp(M)\Phi(u(M^{\alpha} \exp(i\mu)))^{-m}.$$

for $\Phi(u) = u + \sqrt{u^2 - 1}$, $u(z) = (|b+z|+|a+z|)/b-a$.

Specialized bounds

The bound can be refined under stricter hypotheses.

Theorem (Moret and Novati 2011, Theorem 3.5)

Assume that A is Hermitian with $\sigma(A) \subseteq [a, b] \subset [0, +\infty)$. Assume that $0 < \alpha < 1$, $\beta \geq \alpha$. Let $\mu \leq \pi/2$, $\alpha\pi/2 < \mu < \alpha\pi$. Then for every index $m \geq 1$ and for every $M > 0$ we have

$$\|R_m\| \leq \frac{4M^{1-\beta}}{\pi}\left(\frac{\mu}{\alpha} + \frac{\exp(-M(1+|\cos(\mu/\alpha)|))}{M|\cos(\mu/\alpha)|}\right)\exp(M)\Phi(u(M^{\alpha}\exp(i\mu)))^{-m}.$$

for $\Phi(u) = u + \sqrt{u^2 - 1}$, $u(z) = (|b+z|+|a+z|)/b-a$.

Limiting relation

If $\alpha \to 0$, $\beta = 1$, we have $E_{0,1}(-z) = (1+z)^{-1}$, $|z| < 1$. Then setting $\mu = \alpha\pi$ and letting $M = 1$, we find

$$\|R_m\| \leq \frac{4(\pi\exp(1) - \exp(-1))}{\pi\Phi(u(1))^m}$$

The Shift-and-Invert Method (Moret and Novati 2011)

We remain under the assumptions (A1) and (A2) and consider the matrix

$$Z = (I + hA)^{-1}, \qquad h > 0,$$

together with the space $\mathcal{K}_m(Z, \mathbf{v})$.

The Shift-and-Invert Method (Moret and Novati 2011)

We remain under the assumptions (A1) and (A2) and consider the matrix

$$Z = (I + hA)^{-1}, \qquad h > 0,$$

together with the space $\mathcal{K}_m(Z, \mathbf{v})$.
We can write the **analogous Arnoldi relation** for $U_m = [\mathbf{u}_1, \ldots, \mathbf{u}_m]$ spanning $\mathcal{K}_m(Z, \mathbf{v})$:

$$ZU_m = U_m S_m + s_{m+1,m} u_{m+1} \mathbf{e}_m^T, \qquad S_m = U_m^H Z U_m.$$

The Shift-and-Invert Method (Moret and Novati 2011)

We remain under the assumptions (A1) and (A2) and consider the matrix

$$Z = (I + hA)^{-1}, \qquad h > 0,$$

together with the space $\mathcal{K}_m(Z, \mathbf{v})$.
We can write the **analogous Arnoldi relation** for $U_m = [\mathbf{u}_1, \dots, \mathbf{u}_m]$ spanning $\mathcal{K}_m(Z, \mathbf{v})$:

$$ZU_m = U_m S_m + s_{m+1,m} u_{m+1} \mathbf{e}_m^T, \qquad S_m = U_m^H Z U_m.$$

The **approximation** is then given by

$$\mathbf{y} = f(A)\mathbf{v} \approx \mathbf{y}_m = V_m f(B_m)\mathbf{e}_1 \text{ where } (I + hB_m)S_m = I.$$

The Shift-and-Invert Method (Moret and Novati 2011)

We remain under the assumptions (A1) and (A2) and consider the matrix

$$Z = (I + hA)^{-1}, \qquad h > 0,$$

together with the space $\mathcal{K}_m(Z, \mathbf{v})$.
We can write the **analogous Arnoldi relation** for $U_m = [\mathbf{u}_1, \ldots, \mathbf{u}_m]$ spanning $\mathcal{K}_m(Z, \mathbf{v})$:

$$ZU_m = U_m S_m + s_{m+1,m} u_{m+1} \mathbf{e}_m^T, \qquad S_m = U_m^H Z U_m.$$

The **approximation** is then given by

$$\mathbf{y} = f(A)\mathbf{v} \approx \mathbf{y}_m = V_m f(B_m)\mathbf{e}_1 \text{ where } (I + hB_m)S_m = I.$$

We can repeat the general error analysis using

$$R_m = E_{\alpha,\beta}(-A)\mathbf{v} - U_m E_{\alpha,\beta}(-B_m)\mathbf{e}_1 = \frac{1}{2\pi\alpha i} \int_{C(\varepsilon,\mu)} \exp(\lambda^{1/\alpha})\lambda^{(1-\beta)/\alpha} b_m(\lambda)\, \mathrm{d}\lambda,$$

for $b_m(\lambda) = (\lambda I + A)^{-1}\mathbf{v} - U_m(\lambda I + B_m)^{-1}\mathbf{e}_1$.

Error bound (Moret and Novati 2011)

Theorem (Moret and Novati 2011, Theorem 4.3)

For every matrix A satisfying (A1) and (A2), assume $0 < \alpha < 1$ and $\beta \geq \alpha$. Then, there exists a function $g(h)$, continuous in any bounded interval $0 < h_1 \leq h \leq h_2$, such that for $m \geq 2$,

$$\|R_m\| \leq \frac{g(h)}{m-1}.$$

Error bound (Moret and Novati 2011)

Theorem (Moret and Novati 2011, Theorem 4.3)

For every matrix A satisfying (A1) and (A2), assume $0 < \alpha < 1$ and $\beta \geq \alpha$. Then, there exists a function $g(h)$, continuous in any bounded interval $0 < h_1 \leq h \leq h_2$, such that for $m \geq 2$,

$$\|R_m\| \leq \frac{g(h)}{m-1}.$$

Theorem (Moret and Novati 2011, Theorem 4.5)

Assume that A is Hermitian with $\sigma(A) \subseteq [a, +\infty)$, $a \geq 0$. Assume $0 < \alpha \leq 2/3$ and $\beta \geq \alpha$. Then, for every $m \geq 1$ we have

$$\|R_m\| \leq \frac{K_1 Q_m h^{\frac{\beta-1}{\alpha}}}{(1+\sqrt{2})^{m-1}} + \frac{K_2 h^{\beta/\alpha}}{(m-1)^2} \exp\left(-\frac{h^{-1/\alpha}}{\sqrt{2}}\right),$$

where $Q_m = \max_{0 \leq |\varphi| \leq 3\alpha\pi/4} \exp\left(h^{-1/\alpha} \cos\varphi/\alpha\right)(1-\cos\varphi)^{\frac{m-1}{2}}$, with K_1, K_2 constants.

ML function, what have we found?

- The *polynomial method* suffers both for **small** α **values** and for **large field of values**.

ML function, what have we found?

- The *polynomial method* suffers both for **small** α **values** and for **large field of values**.
- For the *shift-and-invert* method the convergence doesn't deteriorate with the size of $W(A)$, its uniform with respect to the h parameter.

ML function, what have we found?

- The *polynomial method* suffers both for **small α values** and for **large field of values**.
- For the *shift-and-invert* method the convergence doesn't deteriorate with the size of $W(A)$, its uniform with respect to the h parameter.
- To obtain a complete method one still has to find a way to repeatedly compute the matrix functions in

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

ML function, what have we found?

- The *polynomial method* suffers both for **small** α **values** and for **large field of values**.
- For the *shift-and-invert* method the convergence doesn't deteriorate with the size of $W(A)$, its uniform with respect to the h parameter.
- To obtain a complete method one still has to find a way to repeatedly compute the matrix functions in

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

- **Research ideas:** finding better rational approximations/poles/expansions together with error analysis for the ML function.

ML function, what have we found?

- The *polynomial method* suffers both for **small** α **values** and for **large field of values**.
- For the *shift-and-invert* method the convergence doesn't deteriorate with the size of $W(A)$, its uniform with respect to the h parameter.
- To obtain a complete method one still has to find a way to repeatedly compute the matrix functions in

$$\mathbf{u}(t) = E_{\alpha,1}(-t^{\alpha}A)\mathbf{u}_0 + \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-A(t-s)^{\alpha})\mathbf{g}(s)\,\mathrm{d}s.$$

- **Research ideas:** finding better rational approximations/poles/expansions together with error analysis for the ML function.

Other extensions

A variant with *restart* is discussed in (Moret and Popolizio 2014), the combination with other matrix-functions in (Moret and Novati 2019).

Bibliography I

Benzi, M. (2021). "Some uses of the field of values in numerical analysis". In: *Boll. Unione Mat. Ital.* 14.1, pp. 159–177. ISSN: 1972-6724. DOI: 10.1007/s40574-020-00249-2. URL: https://doi.org/10.1007/s40574-020-00249-2.

Davies, P. I. and N. J. Higham (2003). "A Schur-Parlett algorithm for computing matrix functions". In: *SIAM J. Matrix Anal. Appl.* 25.2, pp. 464–485. ISSN: 0895-4798. DOI: 10.1137/S0895479802410815. URL: https://doi.org/10.1137/S0895479802410815.

Diele, F., I. Moret, and S. Ragni (2008/09). "Error estimates for polynomial Krylov approximations to matrix functions". In: *SIAM J. Matrix Anal. Appl.* 30.4, pp. 1546–1565. ISSN: 0895-4798. DOI: 10.1137/070688924. URL: https://doi.org/10.1137/070688924.

Garrappa, R. (2015). "Numerical evaluation of two and three parameter Mittag-Leffler functions". In: *SIAM J. Numer. Anal.* 53.3, pp. 1350–1369. ISSN: 0036-1429. DOI: 10.1137/140971191. URL: https://doi.org/10.1137/140971191.

Bibliography II

Garrappa, R. and M. Popolizio (2011). "On accurate product integration rules for linear fractional differential equations". In: *J. Comput. Appl. Math.* 235.5, pp. 1085–1097. ISSN: 0377-0427. DOI: 10.1016/j.cam.2010.07.008. URL: https://doi.org/10.1016/j.cam.2010.07.008.

— (2018). "Computing the matrix Mittag-Leffler function with applications to fractional calculus". In: *J. Sci. Comput.* 77.1, pp. 129–153. ISSN: 0885-7474. DOI: 10.1007/s10915-018-0699-5. URL: https://doi.org/10.1007/s10915-018-0699-5.

Higham, N. J. and X. Liu (2021). "A multiprecision derivative-free Schur-Parlett algorithm for computing matrix functions". In: *SIAM J. Matrix Anal. Appl.* 42.3, pp. 1401–1422. ISSN: 0895-4798. DOI: 10.1137/20M1365326. URL: https://doi.org/10.1137/20M1365326.

Metzler, R. and J. Klafter (2000). "The random walk's guide to anomalous diffusion: a fractional dynamics approach". In: *Phys. Rep.* 339.1, p. 77. ISSN: 0370-1573. DOI: 10.1016/S0370-1573(00)00070-3. URL: https://doi.org/10.1016/S0370-1573(00)00070-3.

Bibliography III

- Moret, I. and P. Novati (2011). "On the convergence of Krylov subspace methods for matrix Mittag-Leffler functions". In: *SIAM J. Numer. Anal.* 49.5, pp. 2144–2164. ISSN: 0036-1429. DOI: 10.1137/080738374. URL: https://doi.org/10.1137/080738374.
- — (2019). "Krylov subspace methods for functions of fractional differential operators". In: *Math. Comp.* 88.315, pp. 293–312. ISSN: 0025-5718. DOI: 10.1090/mcom/3332. URL: https://doi.org/10.1090/mcom/3332.
- Moret, I. and M. Popolizio (2014). "The restarted shift-and-invert Krylov method for matrix functions". In: *Numer. Linear Algebra Appl.* 21.1, pp. 68–80. ISSN: 1070-5325. DOI: 10.1002/nla.1862. URL: https://doi.org/10.1002/nla.1862.
- Podlubny, I. (1999). *Fractional differential equations*. Vol. 198. Mathematics in Science and Engineering. An introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications. Academic Press, Inc., San Diego, CA, pp. xxiv+340. ISBN: 0-12-558840-2.

Bibliography IV

Sokolov, I. M. and J. Klafter (2005). "From diffusion to anomalous diffusion: a century after Einstein's Brownian motion". In: *Chaos* 15.2, pp. 026103, 7. ISSN: 1054-1500. DOI: 10.1063/1.1860472. URL: https://doi.org/10.1063/1.1860472.