Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Memory equations as reduced Markov processes

Artur Stephan[1], Holger Stephan [2]

submitted: March 22, 2018

[1] Insitut für Mathematik
Humboldt-Universität zu Berlin
Rudower Chaussee 25
12489 Berlin
Germany
E-mail: stephan@math.hu-berlin.de

[2] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany
E-Mail: holger.stephan@wias-berlin.de

No. 2496
Berlin 2018

2010 *Mathematics Subject Classification.* 39A06, 34K06, 60J27, 00A71, 34D05, 44A10, 26C15.

Key words and phrases. Markov generator, delay equation, Markov process without detailed balance, modeling memory equations, exponential kernel, reservoirs, quasiparticles, linear differential equations, Lagrange polynomial, Laplace transform, asymptotic behavior, simplex integrals, integro-differential equation, ordinary differential equations, meromorphic functions, non-autonomous, functional differential equation.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Memory equations as reduced Markov processes

Artur Stephan, Holger Stephan

Abstract

A large class of linear memory differential equations in one dimension, where the evolution depends on the whole history, can be equivalently described as a projection of a Markov process living in a higher dimensional space. Starting with such a memory equation, we give an explicit construction of the corresponding Markov process. From a physical point of view the Markov process can be understood as the change of the type of some quasiparticles along one-way loops. Typically, the arising Markov process does not have the detailed balance property. The method leads to a more realisitc modeling of memory equations. Moreover, it carries over the large number of investigation tools for Markov processes to memory equations, like the calculation of the equilibrium state, the asymptotic behavior and so on. The method can be used for an approximative solution of some degenerate memory equations like delay differential equations.

1 Introduction

Memory equations describe the time evolution of some quantity, considering the whole prehistory of the evolution: The past influences the future.

Markov processes, or more generally time evolutions with the Markov property, describe the problem under the assumption that the further evolution can be predicted, knowing only the current state: The present influences the future.

At first glance, by means of memory equations, it is possible to investigate a wider class of problems, since evolution equations with the Markov property can be regarded as degenerate memory problems, where the dependence of the past is concentrated in one moment.

But from a philosophical point of view, it seems to be natural that a complete description of a problem has to be a Markov one for the following reason: The Markov property means that the solution operator is a semigroup, i.e. it is invariant under a time shift. Due to Noether's theorem, this invariant corresponds to the conservation of some energy, the dual variable of time. Thus, the Markov property is the typical property of a model, where some energy is conserved.

Conversely, if the evolution is governed by a non-Markovian equation, it is not complete, some energy is lost. This requires finding more degrees of freedom unless the model is Markovian. In other words, it is to be expected that a non-Markovian description can be regarded as some part or restriction of a more-dimensional Markov process.

This theoretical thought can be confirmed in various practical situations:

- An arbitrary (nonlinear) dynamical system on a compact space $\mathcal{Z}$ can be equivalently formulated as a linear deterministic Markov process on the space of Radon measures on $\mathcal{Z}$ (see, e.g. [14]) via its Liouville equation.

DOI 10.20347/WIAS.PREPRINT.2496 Berlin 2018

- A general linear evolution equation that is nonlocal in space and time, including jumps and memory on some domain in $\mathbb{R}^n$, can be understood as a limit of a diffusion process (a special Markov process) on a complicated Riemannian manifold. (see [9])
- The projection of a general Brownian motion (a special Markov process in phase space) on the coordinate space is a diffusion process if the initial velocity is Maxwellian (see [13]).

Hence, the idea that a memory equation can be regarded as part of a higher dimensional Markov process, does not seem to be very surprising. Indeed, the main result in this paper is that we provide the construction of an easily analyzable Markov process for a more or less arbitrary given memory kernel.

Let us briefly revise the basic facts in modeling and analyzing Memory equations and Markov processes.

1.1 Memory Equations

Memory equations (ME) are differential equations where the evolution depends not only on the current state but also on the past. Memory equations are a special case of Functional Equations - an equation of unknown functions and their derivatives with different argument values. The mathematical theory of functional equations (or integro-differential equations) is treated in [10, 7].

From the viewpoint of modeling and analysis, Memory equations have attracted a lot of attention during the last decades. For example, they arise in modeling flows trough fissured media, [8, 11].

We consider Memory equations of a convolution type. Such equations arise as effective limits of homogenization problems, starting with the pioneering work of L. Tartar [16].

The object of interest is a linear memory equation of the form

$$\dot{u}(t) = -au + K * u = -au + \int_0^t K(t-s)u(s)\mathrm{d}s,\ u(0) = u_0, \tag{1}$$

where $u : [0, \infty[\to \mathbb{R}$ is a scalar state variable, $u_0 \in \mathbb{R}_{\geq 0}$ and $K : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ a positive real kernel. Please note, we focus on a scalar variable, but our considerations can be generalized to systems as well as to non-autonomous linear PDEs (like diffusion equations with time-dependent diffusion coefficients).

Let us briefly explain the ME (1). In contrast to $\dot{u} = -au$, where the decay is quite fast, in this equation the decay is damped due to the influence of former states. The ME can be interpreted as a reduction of the mass into unknown depots. Phenomenologically, this can be modeled by $a = a(t)$, which yields a non-autonomous equation. Another way to think about (1) is the following. Introducing the function A defined by $A' = -K$ and $A(0) = a$, we get

$$\dot{u}(t) = -A(0)u - \int_0^t A'(t-s)u(s)\mathrm{d}s = -\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t A(t-s)u(s)\mathrm{d}s.$$

Integrating the above equation, we get

$$u(t) = u(0) - \int_0^t A(t-s)u(s)\mathrm{d}s$$

that can be regarded as a continuous analogue of the time-discrete scheme

$$u_n = u_0 - a_1 u_{n-1} - a_2 u_{n-2} - \dots. \tag{2}$$

Equivalently, using partial integration we get

$$\dot{u}(t) = -A(t)u_0 - \int_0^t A(t-s)\dot{u}(s)\mathrm{d}s.$$

This form is often considered(e.g. in [11]). Subsequently, we use the form (1).

For solving a ME, the memory described by $K(t)$ or $A(t)$ has to be known for any time $t \geq 0$. This is often postulated, i.e. $K(t)$ is given by heuristic arguments.

A typical and simple example is $K_\alpha(t) = \alpha \mathrm{e}^{-\alpha t}$ for $\alpha > 0$. Then $K_\alpha(t) \geq 0$ and $\int_0^\infty K_\alpha(t)\mathrm{d}t = 1$.

In this case, for $\alpha \longrightarrow +\infty$, the integral on the right-hand side of (1) tends to $u(t)$ – the ME becomes an ordinary differential equation.

In the same sense, a sequence of some other integrals of convolution type can tend to a delay differential equation (DDE), that means $K(t) = \sum_j \alpha_j \delta(t-t_j)$ for large enough $t \geq 0$. So, the kernel K can be interpreted as a measure on the time line that can be approximated by the "simplest" measures: convex combinations of δ-measures. Note that DDEs with the above kernel of the form

$$\dot{u} = -au + \sum_j \alpha_j u(t-t_j),$$

are solved by virtue of an initial condition $\phi \in \mathrm{C}([-\max\{t_j\}, 0])$. That means the solution space is infinite dimensional. On the other hand regarding the modeling viewpoint, it is difficult to derive an initial value $\phi \in \mathrm{C}([0,T])$ for a DDE. Often the initial value ϕ is assumed to be constant or a simple given function. See e.g. [12] for more details, where the analysis and applications especially for modeling aftereffect phenomena are presented.

The ME needs the initial value only for one fixed value, say $t = 0$. On the other hand, if $t \geq \max\{t_j\}$, the DDE become a ME. This means, that the beginning of the evolution is also modeled in the ME. In this sense, MEs include many types of differential equations like ODEs and DDEs. We remark that also from the modeling viewpoint it is more natural to treat kernels located at smeared time values rather than precise time values.

Another important property is the asymptotic behavior. The ME is a non-autonomous differential equation. The equilibrium cannot be calculated setting $\dot{u} = 0$. Assuming $\int_0^\infty K(t)\mathrm{d}t = a$, any constant solution $u(t) = u_0$ satisfies

$$\lim_{t\to\infty} \left(-au(t) + \int_0^t K(s)u(t-s)\mathrm{d}s \right) = 0.$$

Assuming $\int_0^\infty K(t)\mathrm{d}t \neq a$, there is no non-trivial solution that makes the right-hand side zero, so that it is no equilibrium of the ME.

1.2 Markov Processes

There is a huge amount of literature on Markov Processes (see, e.g. [2, 3, 4]). Here we introduce our notation.

Let $\mathcal{Z}$ be a given state space, a compact topological space, $\mathcal{C} := \mathcal{C}(\mathcal{Z})$ the Banach space of continuous functions on $\mathcal{Z}$ and $\mathcal{P} := \mathcal{P}(\mathcal{Z})$ the set of probability measures, i.e. the subset of Radon measures p on $\mathcal{Z}$ with $p \geq 0$ and $p(\mathcal{Z}) = 1$.

A family $\mathbf{T}(t), t \geq 0$ of linear bounded operators in $\mathcal{C}$ is called a Markov semigroup if it is a semigroup, i.e. if it satisfies

$$\mathbf{T}(t_1 + t_2) = \mathbf{T}(t_1)\mathbf{T}(t_2),\ \mathbf{T}(0) = \mathbf{I},\ t_1, t_2 \geq 0\ ,$$

it is positive $\mathbf{T}(t) \geq 0$ in the cone sense of $\mathcal{C}$ and $\mathbb{1}$, the constant function is a fix-point of $\mathbf{T}(t)$ for all $t \geq 0$, $\mathbf{T}(t)\mathbb{1} = \mathbb{1}$. The semigroup property is often called Markov property and it is equivalent to the assumption that the trajectory depends only on the present time point and not on the past.

A linear operator $\mathbf{A}$ on $\mathcal{C}$ is called Markov generator if it is the generator of a Markov semigroup, i.e. if $g(t) = \mathbf{T}(t)g_0$, where $\mathbf{T}(t)$ is a Markov semigroup. Then $g(t) = \mathbf{T}(t)g_0$ is the solution of the equation

$$\dot{g}(t) = \mathbf{A}g(t),\ g(0) = g_0 \tag{3}$$

for an initial value g_0 from the domain of $\mathbf{A}$. This equation is called backward Chapman-Kolmogorov equation. A Markov process is the result of the action of the adjoint semigroup $\mathbf{T}^*(t)$ at a probability measure p_0, i.e. $p(t) = \mathbf{T}^*(t)p_0$. Any Markov process has at least one stationary probability measure $\mu \in \mathcal{P}$. It satisfies $\mathbf{T}^*(t)\mu = \mu$ for all $t \geq 0$. This is a consequence of the Markov-Kakutani Theorem. The stationary probability measure μ is an element of the null-space of $\mathbf{A}^*$.

In this paper we consider continuous-time Markov processes on discrete state spaces. The set $\mathcal{Z} = \{z_0, ..., z_N\}$ is a finite set of $N + 1$ states. In this case, we have $\mathcal{C} = \mathbb{R}^{N+1}$ and $\mathcal{P}$ is the simplex of probability vectors $\mathcal{P} := \mathrm{Prob}(\{z_0, \ldots, z_N\}) := \{p \in \mathbb{R}^{N+1} : p_i \geq 0, \sum_{i=0}^{N+1} p_i = 1\}$ and a subset of $\mathbb{R}^{N+1}$, too. A Markov semigroup is a real matrix family $\mathbf{T}(t)$ on $\mathbb{R}^{N+1}$ with positive entries and row sum 1. Its adjoint is the transposed matrix family $\mathbf{T}^*(t)$.

A Markov process is $p(t) = \mathbf{T}^*(t)p_0$, where p_0 is some given probability vector. It satisfies the set of equations

$$\dot{p}(t) = \mathbf{A}^*p(t),\ p(0) = p_0, \tag{4}$$

where $\mathbf{A}^*$ is the adjoint of the corresponding Markov generator. This equation is called forward Chapman-Kolmogorov equation. In contrast to equation (3) describing the evolution of moment functions, equation (4) describes the evolution of probability vectors. This means that one component of the vector $p(t)$ can be understood as the probability of the corresponding state, regardless of the probability of the other states.

It is well known that equation (4) has a unique solution $p(t) \in \mathcal{P}$ if and only if the off-diagonal elements are nonnegative and the columns of $\mathbf{A}^*$ sum up to 0. Thus, for $\mathbf{A} = (A_{ij})$ we have $A_{ij} \geq 0$ for $i \neq j$ and $A_{ii} = -\sum_{i\neq j=1}^{n} A_{ij}$.

For a generic Markov matrix the stationary probability μ is unique and all trajectories $\mathbf{T}^*(t)p_0$ for any initial state p_0 converge to μ. We only consider Markov processes with a unique stationary probability.

The eigenvalues of a Markov generator have always strongly negative real part, except one eigenvalue 0. The corresponding eigenvector is $\mathbb{1}$ for $\mathbf{A}$ and μ for $\mathbf{A}^*$. If the eigenvalues λ_i of $\mathbf{A}^*$ are all different, every component of the solution to (4), i.e. every component of $\mathbf{T}^*(t)p_0$ is a linear combination of $\mathbb{1}$ and exponential decaying functions $\mathrm{e}^{-\lambda_i t}$.

A Markov process in $\mathbb{R}^{N+1}$ allows for different physical interpretations. Apart from the canonical interpretations as a probability vector, it can be understood as some concentration or amount of $N + 1$ different materials. We will follow this interpretation and will assume that this amount of materials is represented by particles of different types. These particles can transform into each other, changing their type, which can be understood as a linear reaction. The entries of the Markov matrix A_{ij} describe the rates of transforming particles of type z_j into particles of type z_i. Therefore, if we are only

interested in the amount of material of one type, it is enough to consider the corresponding component of the vector $p(t)$ only. The initial amount of material is p_0. Since **A** is a Markov generator, positivity of the concentration and the whole mass is conserved.

If a Markov generator $\mathbf{A} = (A_{ij})$ and its stationary state $\mu = (\mu_i)$ satisfy $A_{ij}\mu_j = A_{ji}\mu_i$ for any $i, j \in \{1, \dots, n\}$, it is said that the corresponding Markov process has the detailed balance property. It is equivalent to the case that the matrix (A_{ij}) is symmetric in the L_2-Hilbert space over μ. Such a matrix has to have real eigenvalues. We remark that the opposite is not true in general: A Markov process without the detailed balance can have real eigenvalues, too. Moreover, there can be no Hilbert space at all, where it is symmetric. From a physical point of view, the condition $A_{ij}\mu_j = A_{ji}\mu_i$ means that any transition $z_i \Leftrightarrow z_j$ is in a local equilibrium. Thus, the detailed balance case is easier to analyze but it rarely appears in general. The systems that we consider do not have the detailed balance property in principle.

1.3 What our paper deals with

In this paper we connect the two concepts of Markovian dynamics and non-Markovian dynamics, which seem to be different at the first glance. Starting with a MP of a special form, we conclude a ME for the first coordinate. As already mentioned, the ME is a scalar differential equation, but our considerations can also be applied to PDEs. The resulting MP can be physically understood; the ME is governed by a kernel which is a sum of exponential functions. Then another path is taken: Starting with an ME with an exponential kernel, we find a MP where its first components yields again the ME. The other components can be understood as hidden degrees of freedom that have to be included in a complete description of the problem. This procedure is not unique and thus, it cannot be said that the hidden degrees of freedom are real physical variables. On the other hand, the construction of the MP out of the kernel is intuitive since the kernel is approximated by its moments. This method can be used to approximate a general positive kernel taking the enlarging of the MP into account. The simple case of two and three states is presented in chapter 2. In this case, all solutions and kernels can be calculated easily. In chapter 3 we consider the general case. The main Theorems are stated here.

The method has many physical and mathematical advantages – both for the theory of MPs and MEs. We want to highlight only two of them. Firstly, the modeling of a kernel for ME is usually done by heuristic arguments. The method presented here can be used to model kernel in a more convenient manner, since the MP has an underlying physical meaning. Moreover, the modeling of the beginning of the process is also done, as we already mentioned. Secondly, the asymptotic behavior of a non-autonomous differential equation can now be calculated easily from the Markovian dynamics.

The paper concludes with chapter 4. Here we remark the connection to delay differential equations, where the kernel is highly degenerate. This is also reflected in the setting of MP: The underlying Markov generator has a very special form. We observe that the solution of the ME converges to the equilibrium of the MP and also the spectral functions of ME and MP converge.

Summarizing, we have the following connection of model levels:

$$\text{MP} \subset \text{DDE} \subset \text{ME} \subset \text{MP'}.$$

Here MP' is a Markov process with a larger number of degrees of freedom.

It is well known that a linear delay equation with delay T in a state space $\mathcal{X}$ can be regarded as an autonomous equation in a much larger space $\mathrm{C}([-T, 0], \mathcal{X})$. There, the evolution of the delay equation is described by a semigroup of linear operators. This approach is not the aim in this paper.

Notion: In this paper, the Laplace transform is used frequently. Some properties are summarized in the appendix. ME of convolution type have the important property that the Laplace transform maps them into multiplication operators. The Laplace transform $\mathcal{L}(u)$ of a real valued function $t \mapsto u(t)$ is defined by $\mathcal{L}(u)(\lambda) = \hat{u}(\lambda) = \int_0^\infty \mathrm{e}^{-\lambda t} u(t)\mathrm{d}t$. If there is no confusion, we omit the 'hat' on $\hat{u}$ and just write u or $u(\lambda)$.

Some analytical tools concerning Lagrange polynomials and simplex integrals are moved to the appendix, too.

2 Some simple Markov processes and memory equations

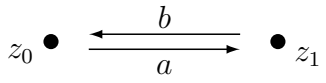
Before starting the general theory, we firstly present the basic ideas focusing on simple low dimensional examples – Markov processes with two and three states. Apart from the sake of simplicity nearly all phenomena of the general theory are eminent.

2.1 Two states

We consider a Markov process on a state space of two abstract states $\{z_0, z_1\}$, generated by the Markov generator

$$\mathbf{A} = \begin{pmatrix} -a & a \\ b & -b \end{pmatrix}, \text{ and its transpose } \mathbf{A}^* = \begin{pmatrix} -a & b \\ a & -b \end{pmatrix}. \tag{5}$$

The matrix $\mathbf{A}^*$ describes the switching between the two states with given rates $a \geq 0, b \geq 0$. We can think of an amount of matter, represented by particles, which can occur in two types. For some reason we are interested only in particles of the first type.



The equation describing the evolution of the vector $p = (u, v)$ reads $\dot{p} = \mathbf{A}^* p$ with $p(0) = p_0$. We assume that in the beginning the total mass is concentrated in the first variable, i.e. $p_0 = (u_0, 0)$. In other words, all particles have type z_0.

The eigenvalues of $\mathbf{A}^*$ are $\{0, -(a+b)\}$. The stationary solution is $\mu = \left(\frac{b}{a+b}u_0, \frac{a}{a+b}u_0\right)$. It is unique unless the non interesting case $a = b = 0$. Any Markov process with two states has the detailed balance property.

For (u, v) the system reads as

$$\begin{cases} \dot{u} = -au + bv \\ \dot{v} = au - bv. \end{cases} \tag{6}$$

Using the Laplace transform and writing $u(\lambda) = \mathcal{L}(u(t))(\lambda)$ and $v(\lambda) = \mathcal{L}(v(t))(\lambda)$, we obtain a system of equations for (u, v) in the form

$$\begin{cases} (\lambda + a)u - u_0 = bv \\ (\lambda + b)v = au. \end{cases}$$

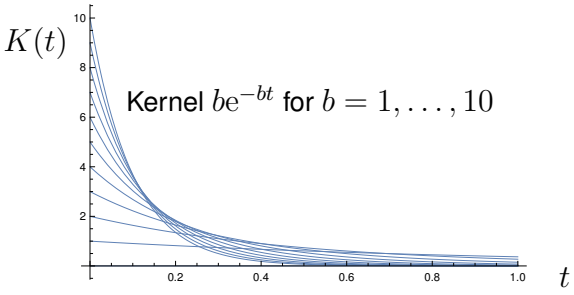
This yields an equation for u in the form

$$(\lambda + a)u - u_0 = \frac{ba}{\lambda + b}u \Rightarrow \lambda u - u_0 = -au + \frac{ba}{\lambda + b}u.$$

Using the inverse Laplace transform, we obtain a Memory Equation for u

$$\dot{u} = -au + ab\int_0^t \mathrm{e}^{-b(t-s)}u(s)\mathrm{d}s = -a\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{e}^{-b(t-s)}u(s)\mathrm{d}s. \tag{7}$$

The kernel $K(t) = b\mathrm{e}^{-bt}$ describe a dependence of the current state from previous time moments. For $b \longrightarrow \infty$, $K(t)$ tends to $\delta(t)$ and the equation becomes $\dot{u} = 0$. Thus, the right hand side of equation (7) consists of two terms, the first one, $-au$ describe an exponential decay, whereas the second one, the memory term describe an opposite effect: Particles that disappear, occur after a while. The time that passes between disappearing and reappearing, decreases with $1/b$. In the end, not all matter disappears like in a pure equation $\dot{u} = -au$ but an equilibrium between disappearance and reappearance arises.



The same effect is caused by the Markov process, changing the type of the particles. The particle changes the type from z_0 to z_1 with rate a, it seems to disappear, if we look only at type z_0. After a while it re-changes to type z_1 (it occurs) with rate b. This give the exponential time behavior e^{-bt} (corresponding to the memory kernel $K(t) = b\mathrm{e}^{-bt}$), characteristic for Markov processes.

The equation (7) – or equivalently the system (6) – can be solved explicitly. We obtain for the Laplace transform

$$u(\lambda) = \frac{\lambda + b}{\lambda(\lambda + a + b)}u_0 = \left(\frac{b}{a+b}\frac{1}{\lambda} + \frac{a}{a+b}\frac{1}{\lambda + a + b}\right)u_0$$

and for the solution itself

$$u(t) = \frac{b}{a+b}u_0 + \frac{a}{a+b}e^{-(a+b)t}u_0$$

The solution tends to an equilibrium state $u_\infty = \frac{b}{a+b}u_0$, the first component of the stationary solution μ.

It is not possible to calculate it from the memory equation (7), directly. Setting $\dot{u} = 0$, the equation

$$\dot{u} = -au + ab\int_0^t \mathrm{e}^{-b(t-s)}u(s)\mathrm{d}s = -a\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t \mathrm{e}^{-b(t-s)}u(s)\mathrm{d}s.$$

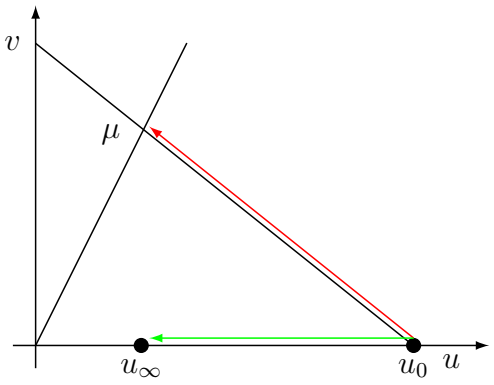
does not have any solution at all. Passing to the limit $t \longrightarrow \infty$ (rewriting at first $\int_0^t \mathrm{e}^{-b(t-s)}u(s)\mathrm{d}s = \int_0^t \mathrm{e}^{-bs}u(t-s)\mathrm{d}s$) we obtain

$$0 = -au_\infty + ab\int_0^\infty \mathrm{e}^{-bs}u_\infty \mathrm{d}s\ .$$

Any constant u_∞ solves this equation.

This strange behavior of the solution of memory equations is typical and can be illustrated in a picture,

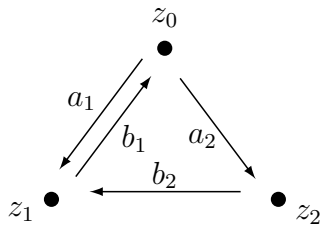
showing the time behavior of both, the solution of the Markov process and their first component – the solution of the memory equation.
Investigating only the solution of the memory equation, it is not clear why the trajectory $u(t)$ stops in u_∞. Whereas looking from above, the trajectory $(u(t), v(t))$ has to stop at the stationary state μ, the intersection of the subspace $u + v = 1$ with the null space of $\mathbf{A}^*$.



2.2 Three states

A general memory kernel must not be concentrated in $t = 0$. It can describe a transfer of mass from a very earlier time. It seems that this situation can be modeled by transitions between many quasiparticles before it appears at its starting type again. To understand the action of such a transition

loop, we investigate in detail a special case of three states, namely the transformation of a fixed particle (type z_0) in two different quasiparticles. One of them (type z_1) can be transformed back into type z_0 immediately, whereas the other (type z_2) can be transformed back into type z_0 only by two steps, changing at first to type z_1. This process is illustrated in the picture.



2.2.1 From Markov to Memory

The simple Markov process on a state space of three abstract states $\{z_0, z_1, z_2\}$ is described by the Markov generator

$$\mathbf{A} = \begin{pmatrix} -a_1 - a_2 & a_1 & a_2 \\ b_1 & -b_1 & 0 \\ 0 & b_2 & -b_2 \end{pmatrix}, \quad \mathbf{A}^* = \begin{pmatrix} -a_1 - a_2 & b_1 & 0 \\ a_1 & -b_1 & b_2 \\ a_2 & 0 & -b_2 \end{pmatrix} \tag{8}$$

with $a_1, a_2, b_1, b_2 \geq 0$. The equation, generating the Markov process is

$$\dot{p}(t) = \mathbf{A}^* p(t), \;\; p(0) = p_0. \tag{9}$$

Note, this is a Markov generator depending on four rates. A general Markov generator on $\mathbb{R}^3$ depend on six rates.

The stationary state μ is the solution to $\mathbf{A}^*\mu = 0$ and can be calculated easily as

$$\mu = \left(1 + \frac{a_1 + a_2}{b_1} + \frac{a_2}{b_2}\right)^{-1} \left(1, \frac{a_1 + a_2}{b_1}, \frac{a_2}{b_2}\right) u_0 = \frac{(b_1 b_2, a_1 b_2 + a_2 b_2, a_2 b_1)}{b_1 b_2 + a_1 b_2 + a_2 b_2 + a_2 b_1} u_0$$

The eigenvalues (they have always non-positive real part) of the matrix are $\lambda_0 = 0$ and

$$\lambda_{1,2} = -\frac{1}{2}\left(a_1 + a_2 + b_1 + b_2 \pm \sqrt{(a_1 + a_2 + b_1 + b_2)^2 - 4(a_1 b_2 + a_2 b_1 + a_2 b_2 + b_1 b_2)}\right)$$

Depending on a_1, a_2, b_1, b_2 the eigenvalues can be real (e.g. $\lambda_1 = -5$, $\lambda_2 = -11$ for $a_1 = 2, a_2 = 5, b_1 = 8, b_2 = 1$) or complex (e.g. for $\lambda_{1,2} = -9 \pm 2i$ for $a_1 = 2, a_2 = 5, b_1 = 8, b_2 = 3$). (By the way, these are suitable values for an explicite solution with rational terms, only.)

This Markov process has the detailed balance property, if $b_1 b_2 a_2 = 0$ what is not interesting, since the coupling chain is broken. Roughly speaking, the detailed balance property means that for any loop in one direction there is a loop backwards with the same product of the rates. But this is not the case in our model. Thus, the Markov process under consideration violate the detailed balance property, generically.

The stationary state is unique if and only if the real parts of $\lambda_{1,2}$ are strongly negative. Or, equivalently, $b_1 b_2 + a_1 b_2 + a_2 b_2 + a_2 b_1 = 0$. Since the a_i, b_i are non negative, this is a non interesting case that we exclude. Then, the stationary state is the equilibrium state for any initial value. Note, that nevertheless some of the a_i, b_i might be zero.

As in the case of two states, we are interested only in the state z_0 of the system and ask for an evolution equation of this state. To do this, we introduce the notion $p = (u, v_1, v_2)$ and look for the evolution of u with an initial state $p_0 = (u_0, 0, 0)$. This is naturally, since the states z_1 and z_2 are unknown, and there is no reason to assume something else than nothing in the beginning.

Equation (9) is now equivalent to the system

$$\left\{\begin{array}{rcrcrc} \dot{u}(t) &=& -(a_1+a_2)u(t) &+b_1 v_1(t) & & \\ \dot{v}_1(t) &=& a_1 u(t) &-b_1 v_1(t) &+b_2 v_2(t) & \\ \dot{v}_2(t) &=& a_2 u(t) & &-b_2 v_2(t) & \end{array}\right.$$

Passing to the Laplace transform, we obtain with $u = \mathcal{L}u$, $v_i = \mathcal{L}v_i$ the system

$$\left\{\begin{array}{rcrcrc} \lambda u &=& -(a_1+a_2)u &+b_1 v_1 & &+u_0 \\ \lambda v_1 &=& a_1 u &-b_1 v_1 &+b_2 v_2 & \\ \lambda v_2 &=& a_2 u & &-b_2 v_2 & \end{array}\right.$$

or equivalently, introducing $a = a_1 + a_2$, we get

$$\left\{\begin{array}{rcrc} (\lambda+a)u - u_0 &=& & b_1 v_1 \\ (\lambda+b_1)v_1 &=& a_1 u &+ b_2 v_2 \\ (\lambda+b_2)v_2 &=& & a_2 u. \end{array}\right.$$

Here, v_1 and v_2 can be eliminated as

$$v_2 = \frac{a_2}{\lambda+b_2}u\,, \quad v_1 = \frac{a_1}{\lambda+b_1}u + \frac{b_2}{\lambda+b_1}v_2 = \frac{a_1}{\lambda+b_1}u + \frac{a_2 b_2}{(\lambda+b_1)(\lambda+b_2)}u$$

We conclude the following equation for u

$$\lambda u - u_0 = \left(-a + a_1\frac{b_1}{\lambda+b_1} + a_2\frac{b_1}{\lambda+b_1}\frac{b_2}{\lambda+b_2}\right)u. \tag{10}$$

This is an equation for the first state, only. It can be solved explicitly with respect to u. But, at this moment, this is not our aim. We are looking for an equation for u. We write

$$\frac{b_1}{\lambda+b_1}\frac{b_2}{\lambda+b_2} = \frac{b_1 b_2}{b_2-b_1}\left(\frac{1}{\lambda+b_1} - \frac{1}{\lambda+b_2}\right),$$

and, after transforming inverse, we get an equation for the function $u(t)$, namely

$$\begin{aligned}\dot{u} &= -au + a_1 b_1 \int_0^t \mathrm{e}^{-b_1 s} u(t-s)\mathrm{d}s + a_2 \frac{b_1 b_2}{b_2 - b_1} \int_0^t (\mathrm{e}^{-b_1 s} - \mathrm{e}^{-b_2 s}) u(t-s)\mathrm{d}s \qquad (11)\\ &= -au + (K * u)(t),\end{aligned}$$

where

$$K(t) = b_1 a_1 \mathrm{e}^{-b_1 t} + a_2 \frac{b_1 b_2}{b_2 - b_1} \left(\mathrm{e}^{-b_1 t} - \mathrm{e}^{-b_2 t}\right) = \qquad (12)$$

$$= \left(b_1 a_1 + \frac{b_1 b_2 a_2}{b_2 - b_1}\right) \mathrm{e}^{-b_1 t} - \frac{b_1 b_2 a_2}{b_2 - b_1} \mathrm{e}^{-b_2 t} \qquad (13)$$

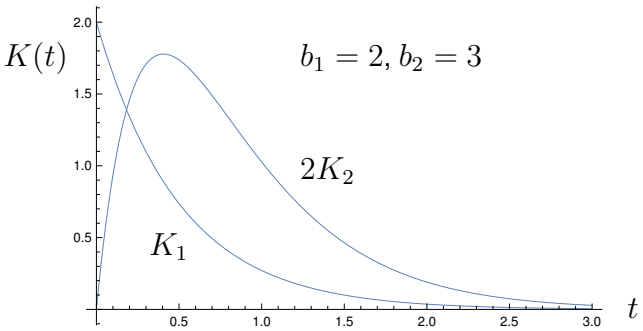
So, we obtain a memory equation with the kernel K. This equation describe the evolution of the first state of our physical system, depending on the whole past from 0 to time t. Obviously, this dependence is a result of the projection, since nothing else had be done. Thus, $u(t)$ is the solution of two equivalent equations, a memory equation and a component of a Markov system.

The kernel $K(t) = a_1 K_1(t) + a_2 K_2(t)$ is the sum of two parts

$$\begin{aligned} K_1(t) &= b_1 \mathrm{e}^{-b_1 t} \\ K_2(t) &= \frac{b_1 b_2}{b_2 - b_1} \left(\mathrm{e}^{-b_1 t} - \mathrm{e}^{-b_2 t}\right) \end{aligned}$$

each of them is obviously positive . If we denote $m_i = \int_0^\infty t K_i(t)\mathrm{d}t$ the mean time of a kernel, we have

$$m_1 = \frac{1}{b_1},\ m_2 = \frac{1}{b_1} + \frac{1}{b_2}$$



The first kernel K_1 describes a memory effect with small mean time and correspond to a small loop $z_0 \xrightarrow{a_1} z_1 \xrightarrow{b_1} z_0$ in the Markov process. The other kernel K_1 describes a memory effect with longer mean time and correspond to a longer loop $z_0 \xrightarrow{a_2} z_2 \xrightarrow{b_2} z_1 \xrightarrow{b_1} z_0$.

The relative coefficients a_i/a form a convex combination. The transitions $z_0 \xrightarrow{a_i} z_i$ split the whole number of particles in parts according to the loops.

Let us summarize some properties of the kernel $K(t)$.

- $K(t)$ is the sum of exponential decaying functions, where the exponents are the entries of the of diagonal elements of **A**.
- The arising memory equation is (11) with $a = \sum_i^N a_i$ or, equivalently, $k(\lambda = 0) = a$
- $K(t) \geq 0$ iff $k(\lambda) \geq 0$, since $a_i, b_i \geq 0$.

Equation (10) can be solved explicitely:

$$\begin{aligned} u\left(\lambda + a - \frac{a_1 b_1}{\lambda + b_1} - \frac{a_2 b_1 b_2}{(\lambda + b_1)(\lambda + b_2)}\right) &= u_0 \\ \lambda u \left(\frac{\lambda^2 + \lambda(a + b_1 + b_2) + a_2 b_1 + a_1 b_2 + a_2 b_2 + b_1 b_2}{(\lambda + b_1)(\lambda + b_2)}\right) &= u_0 \\ u = \frac{1}{\lambda} \frac{(\lambda + b_1)(\lambda + b_2)}{\lambda^2 + \lambda(a + b_1 + b_2) + a_2 b_1 + a_1 b_2 + a_2 b_2 + b_1 b_2} & \quad u_0. \end{aligned}$$

To get an explicite term for $u(t)$ we have to factorize the denominator what leads – of course – to the same time behavior as determined by the eigenvalues for the Markov process.

We compute the asymptotic behavior of the solution $u(t)$, using the asymptotic properties of the Laplace transform. We obtain for the equilibrium state

$$u_\infty = \lim_{\lambda \to 0} \lambda u = \frac{b_1 b_2}{a_2 b_1 + a_1 b_2 + a_2 b_2 + b_1 b_2} u_0.$$

For the other components we get in the same manner

$$\begin{aligned} v_1(t = \infty) &= \frac{a_1 b_2 + a_2 b_2}{a_2 b_1 + a_1 b_2 + a_2 b_2 + b_1 b_2} u_0 \\ v_2(t = \infty) &= \frac{a_2 b_1}{a_2 b_1 + a_1 b_2 + a_2 b_2 + b_1 b_2} u_0. \end{aligned}$$

These are the parts of the initial mass that remain in the states z_1 and z_2.

2.2.2 From Memory to Markov

Now, we go the opposite direction and start with a kernel that is the sum of two exponential decaying terms, i.e.

$$K(t) = c_1 \mathrm{e}^{-\alpha_1 t} + c_2 \mathrm{e}^{-\alpha_2 t} \tag{14}$$

with some real coefficients c_1, c_2. We assume $c_i \neq 0$, otherwise we are in the case of 2 states. For definiteness, we assume $\alpha_1 > \alpha_2 > 0$. The α_i has to be strongly positive, otherwise we have no decreasing of the time dependence of the past.

This kernel has to be written in the form (12) with positive coefficients. We have

$$\begin{aligned} K(t) &= c_1 \mathrm{e}^{-\alpha_1 t} + c_2 \mathrm{e}^{-\alpha_2 t} = \\ &= (c_1 + c_2)\mathrm{e}^{-\alpha_1 t} + c_2(\alpha_1 - \alpha_2)\frac{\mathrm{e}^{-\alpha_2 t} - \mathrm{e}^{-\alpha_1 t}}{\alpha_1 - \alpha_2} \end{aligned}$$

Thus, we have to demand $c_1 + c_2 \geq 0$ and $c_2 \geq 0$. Both are consequences of the positivity of $K(t)$, setting $t = 0$ and $t \longrightarrow \infty$.

Now, the Markov process is easily constructed. We set

$$\begin{aligned} b_1 &= \alpha_1 \\ b_2 &= \alpha_2 \\ a_2 &= \frac{c_2(\alpha_1 - \alpha_2)}{\alpha_1 \alpha_2} \\ a_1 &= \frac{c_1 + c_2}{\alpha_1}\,. \end{aligned}$$

The entries of the matrix b_1, b_2, a_2 are strongly positive, a_1 is non negative. This guarantees the uniqueness of the stationary solution. Moreover, this violates the detailed balance property.

The existence of a positive equilibrium is fulfilled, we have the equation

$$\dot{u} = -au + \int_0^t K(t-s)u(s)\mathrm{d}s,\ u(0) = u_0$$

and the property of consistency $k(0) = a$.

Summarizing, we get the following result:

Proposition 2.1. *The first component of the MP generated by* $\boldsymbol{A}^*$ *given by (8) is the solution to the ME (11).*

For a ME $\dot{u} = -au + (K * u)$ *with a kernel (14) with parameters* $c_1, c_2, \alpha_1, \alpha_2$ *satisfying* $\alpha_1 > \alpha_2 > 0$, $c_1 + c_2 \geq 0$ *and* $c_2 \geq 0$*, it can be constructed a three dimensional MP, where the first component coincides with the solution to the ME.*

3 General Memory Equations as Markov processes

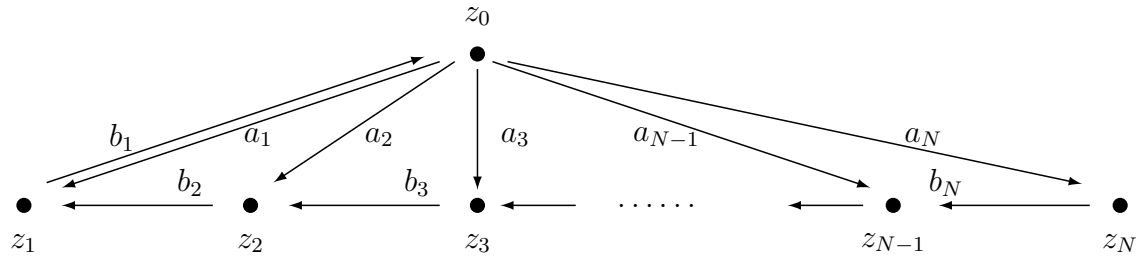
In this chapter, we generalize the ideas from the last chapter to an arbitrary finite dimensional Markov process. Firstly, we show that the first coordinate of a special Markov process, consisting of different transformation loops, satisfies a suitable memory equation with a more or less general kernel. Then, we go the opposite direction: We show that a ME with kernel of a special form yields the MP we started with. The construction of the Markov process is explicitly.

3.1 From Markov to Memory

We consider a Markov process of $N+1$ abstract states $\{z_0, z_1, \ldots, z_N\}$ of the following form

$$\mathbf{A}^* = \begin{pmatrix} -a & b_1 & 0 & 0 & \ldots & 0 \\ a_1 & -b_1 & b_2 & 0 & \ldots & 0 \\ a_2 & 0 & -b_2 & b_3 & \ldots & 0 \\ a_3 & 0 & 0 & -b_3 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{N-1} & 0 & 0 & 0 & -b_{N-1} & b_N \\ a_N & 0 & 0 & 0 & 0 & -b_N \end{pmatrix}, \tag{15}$$

where $a_j \geq 0$ and $b_j > 0$ for $j = 1, \ldots, N$ are non negative rates and we set $a := \sum_{j=1}^N a_j$. The condition $b_j > 0$ is reasonable, since otherwise the loop is broken somewhere.



The process $p(t)$ is generated by the equation $\dot{p} = \mathbf{A}^* p$. We set $p = (u, v_1, \dots, v_N)$ and understand this quantity as the concentration of some particles. We assume that for $t = 0$ the total mass is concentrated in the first coordinate, i.e $p_0 = (u_0, 0, \dots, 0)$. The equation conserves positivity of p and the whole mass $u + v_1 + ... + v_N = u_0$. Thus, p is a vector on the positive simplex in $\mathbb{R}^{N+1}$, intersected by the hyperplane $u + v_1 + ... + v_N = u_0$. Of our interest is the first component, i.e. the amount of matter of particles of type z_0.

$\mathbf{A}^*$ is the generator of a special type of Markov processes. It describe the change of types in the following way: Particles of type z_0 can changes their type to type z_i with rates a_i. The back-changing of a particle of type z_i to type z_0 does not go in a direct way, but in i steps. Thus, we have an interaction between the $N + 1$ types in N loops (see the picture).

Easy calculations show that the stationary solution μ satisfying $A^*\mu = 0$ has the form

$$\mu = \frac{1}{Z}\left(1, \frac{a_1 + \dots + a_N}{b_1}, \frac{a_2 + \dots + a_N}{b_2}, \frac{a_3 + \dots + a_N}{b_3}, \dots, \frac{a_N}{b_N}\right) u_0,$$

where Z is the suitable normalization such that $\sum_{j=0}^{N} \mu_j = u_0$. Obviously,

$$Z = 1 + \sum_{i=1}^{N} \frac{1}{b_i} \sum_{j=i}^{N} a_j \ . \tag{16}$$

For the zeroth coordinate we have

$$u(\infty) = \frac{1}{Z}$$

Since any $b_j > 0$, this stationary solution is unique and is the equilibrium state for any initial condition.

Let us check, whether detail balance with respect to μ is satisfied. We have to check, that $A_{ij}\mu_j = A_{ji}\mu_i$. Since $A_{1j}\mu_j = A_{j1}\mu_1 = 0$ for $j \geq 2$, we obtain that $a_2 = a_3 = \dots a_N = 0$. Hence, the evolution of the states $z_2, \dots, z_N$ is not coupled to the evolution of z_0 and z_1. In this case, we get $N = 1$, the two dimensional case, where every MP has the detailed-balance property. That means, apart form this situation, the MP under consideration does not have the detailed-balance property.

The equation $\dot{p} = \mathbf{A}^* p$ is equivalent to the following system for $p = (u, v_1, \dots, v_N)$

$$\begin{cases} \dot{u} = -au + b_1 v_1 \\ \dot{v}_1 = a_1 u - b_1 u + b_2 v_2 \\ \dot{v}_2 = a_2 u - b_2 v_2 + b_3 v_3 \\ \dot{v}_3 = a_3 u - b_3 v_3 + b_4 v_4 \\ \dots\dots \\ \dot{v}_{N-1} = a_{N-1} u - b_{N-1} v_{N-1} + b_N v_N \\ \dot{v}_N = a_N u - b_N v_N. \end{cases}$$

Using the Laplace transform, we get the following equation for $(u, v_1, \dots, v_N)$

$$\begin{cases} (\lambda + a)u - u_0 = b_1 v_1 \\ (\lambda + b_1)v_1 = a_1 u + b_2 v_2 \\ (\lambda + b_2)v_2 = a_2 u + b_3 v_3 \\ (\lambda + b_3)v_3 = a_3 u + b_4 v_4 \\ \cdots\cdots \\ (\lambda + b_{N-1})v_{N-1} = a_{N-1} u + b_N v_N \\ (\lambda + b_1)v_N = a_N u \end{cases}$$

This yields for u

$$(\lambda + a)u - u_0 = \left(\frac{a_1 b_1}{\lambda + b_1} + \frac{a_2 b_1 b_2}{(\lambda + b_1)(\lambda + b_2)} + \frac{a_3 b_1 b_2 b_3}{(\lambda + b_1)(\lambda + b_2)(\lambda + b_3)} + \dots + \frac{a_N b_1 b_2 \cdots b_N}{(\lambda + b_1)(\lambda + b_2) \cdots (\lambda + b_N)} \right) u. \tag{17}$$

We define the kernel

$$\begin{aligned} k(\lambda) &= \sum_{j=1}^{N} a_j k_j(\lambda) \\ k_j(\lambda) &= \prod_{i=1}^{j} \frac{b_i}{\lambda + b_i} \end{aligned}$$

and hence the equation for the Laplace transform reads

$$\lambda u - u_0 = -au + k(\lambda)u. \tag{18}$$

Now, we formulate the memory equation in terms of $t \geq 0$ and some properties of the kernel. For this purpose, we introduce some quantities, connected with Lagrange polynomials (see the appendix for details) with different support points $b_1, ..., b_n$. Let

$$\psi_i^j = \prod_{k=1, k \neq i}^{j} \frac{b_k}{b_k - b_i},$$

assuming $b_i \neq b_k$ for $i \neq k$.

From the theory of Lagrange polynomials it is well known that

$$k_j(\lambda) = \prod_{i=1}^{j} \frac{b_i}{\lambda + b_i} = \sum_{i=1}^{j} \frac{b_i}{\lambda + b_i} \psi_i^j. \tag{19}$$

Using this, we can transform $k_i(\lambda)$ back and obtain

$$K(t) = \sum_{j=1}^{N} a_j K_j(t) \tag{20}$$

$$K_j(t) = \sum_{i=1}^{j} b_i \psi_i^j \mathrm{e}^{-b_i t} \tag{21}$$

(see the properties of the Laplace transform in the appendix).

The assumption $b_i \neq b_j$ for $i \neq j$ is not principial. If some or all b_i coincide, all formulae of the following can be obtained by some suitable limits. This is obviously done for the Laplace transform $k(\lambda)$. For $K(t)$ we get more complicated terms, involving not only exponential but also polynomials with degree, depending on the frequency of the b_i. We do not bore the reader with this technical complexity, since this is well known in the theory of Lagrange polynomials. Moreover, from a practical point of view, in a generic Markov matrix all entries can be chosen differently.

Surely, a different situation is, if the modeling requires equal b_i. This is the case for instance for DDE's. The case is considered in detail in chapter 4.

Now, we are ready for the following

Theorem 3.1. *Let $p = (u, v_1, \ldots, v_N)$ be the solution of $\dot{p} = \boldsymbol{A}^* p$ with $p_0 = (u_0, 0, \ldots, 0)$ where $\boldsymbol{A}^*$ is given via (15). Then $t \mapsto u(t)$ solves the memory equation*

$$\dot{u} = -au + \int_0^t K(t-s)u(s)\mathrm{d}s,\ u(0) = u_0, \tag{22}$$

where $K(t) = \sum_{j=1}^N a_j K_j(t)$ with $K_j(t) = \sum_{i=1}^j b_i \psi_i^j \mathrm{e}^{-b_i t}$ and $a = \sum_j a_j = k(0)$. Moreover, $K(t) \geq 0$ and $u_\infty = 1/Z$ where Z is given by (16).

Proof. From the definition of $k(\lambda)$ it is clear that $u(\lambda)$ defined by the MP is the solution to (18). If the inverse transformed function $t \mapsto u(t)$ is regular enough, it is solution to (22).

Rewriting (17) as

$$\lambda u(\lambda) = \frac{\lambda}{\lambda + a - \sum_{j=1}^N a_j k_j(\lambda)}\, u_0 \tag{23}$$

Since the $k_j(\lambda)$ are analytical functions and bounded on the right plane, so is $\lambda u(\lambda)$. Hence from the properties of the Laplace transform it follows that $u(t)$ is continuous differentiable. Thus, it solves (22).

To calculate u_∞ we use the representation (23) and investigate the behavior of $k_j(\lambda)$ for $\lambda \longrightarrow \infty$. We have

$$\begin{aligned}
k_j(\lambda) &= k_j(0) + \lambda k_j'(0) + o(\lambda) = \\
&= 1 + \lambda \left(\frac{b_1 b_2 \cdots b_j}{(\lambda + b_1)(\lambda + b_2) \cdots (\lambda + b_j)} \right)' \Bigg|_{\lambda=0} + o(\lambda) = \\
&= 1 - \lambda \left. \frac{b_1 b_2 \cdots b_j \cdot \left(b_1 b_2 \cdots b_j \sum_{i=1}^j \frac{1}{b_i} + o(\lambda) \right)}{\left[(\lambda + b_1)(\lambda + b_2) \cdots (\lambda + b_j) \right]^2} \right|_{\lambda=0} + o(\lambda) = \\
&= 1 - \lambda \sum_{i=1}^j \frac{1}{b_i} + o(\lambda)
\end{aligned}$$

By definition $a = \sum_{j=1}^N a_j$, and hence, it follows from (23)

$$\begin{aligned}
u(\infty) &= \lim_{\lambda \to \infty} \lambda u(\lambda) = \lim_{\lambda \to \infty} \frac{\lambda}{\lambda + a - \sum_{j=1}^N a_j \left[1 - \lambda \sum_{i=1}^j \frac{1}{b_i} + o(\lambda) \right]}\, u_0 = \\
&= \frac{1}{1 + \sum_{j=1}^N a_j \sum_{i=1}^j \frac{1}{b_i}}\, u_0 = \frac{1}{1 + \sum_{j=1}^N \frac{1}{b_j} \sum_{i=j}^N a_j}\, u_0 \,,
\end{aligned}$$

what is exactly the zeroth coordinate of μ, i.e. $1/Z$.

The positivity of the $K_j(t)$, $t \geq 0$ follows from their representation with simplex integrals (see the appendix). We have

$$K_j(t) \;=\; \sum_{i=1}^{j} b_i \psi_i^j \mathrm{e}^{-b_i t} = \int_{S_j} (-1)^{j-1} f^{(j-1)}\big(\langle \alpha, s\rangle t\big)\Big|_{s_j = 1-s_1-s_2-s_{j-1}} ds_{j-1}\cdots ds_1$$

with $f(x) = \mathrm{e}^{-xt}$ and $\langle \alpha, s\rangle = \alpha_1 s_1 + \alpha_2 s_2 + \ldots + \alpha_j s_j$. $(-1)^{j-1} f^{(j-1)}\big(\langle \alpha, s\rangle t\big) = t^{j-1}\mathrm{e}^{-\langle \alpha, s\rangle t} \geq 0$ proves the positivity of $K_j(t)$ and therefore we have $K(t) \geq 0$, since the coefficients a_j in (19) are positive, too. This completes the proof of the theorem. □

3.2 From Memory to Markov

We consider memory equations of the form

$$\dot{u}(t) = -au + K * u = -au + \int_0^t K(t-s)u(s)\mathrm{d}s,$$

where $a > 0$ is a real parameter and K is a positive kernel. The aim is to embed the evolution of u into a Markov process introducing new variables.

Our main assumptions are $K(t) \geq 0$ and $\int_0^\infty K(t)\mathrm{d}t = a$. Clearly, starting with some given $K(t)$ we want to end up with a kernel of the shape (20-21). Then going forward to a kernel like in (17), the entries of the Markov generator matrix can be taken immediately.

The kernels (21) are positive although this are linear combinations of exponential with – maybe – negative coefficients.

It may seem that any nonnegative kernel $K(t)$ can be presented in such a form. But this is not the case. We show this in a

Counterexample: Let

$$K(t) = 3\mathrm{e}^{-t} - 8\mathrm{e}^{-2t} + 6\mathrm{e}^{-3t}$$

and

$$f(t) = \mathrm{e}^{4t}K(t) = 3\mathrm{e}^{3t} - 8\mathrm{e}^{2t} + 6\mathrm{e}^{t}$$

$f(t)$ has a unique minimum $f(0.215315...) = 0.8590718....$ Thus $K(t) \geq 0$.

Seeking for coefficients A, B, C, D, E, F, G (this is the representation (21)) with

$$\begin{aligned} K(t) \;&=\; A\mathrm{e}^{-3t} + B\mathrm{e}^{-2t} + C\mathrm{e}^{-t} + D\frac{\mathrm{e}^{-t}-\mathrm{e}^{-2t}}{1} + E\frac{\mathrm{e}^{-t}-\mathrm{e}^{-3t}}{2} + F\frac{\mathrm{e}^{-2t}-\mathrm{e}^{-3t}}{1} + \\ &+\; G\left(\frac{\mathrm{e}^{-t}}{1\cdot 2} + \frac{\mathrm{e}^{-2t}}{(-1)\cdot 1} + \frac{\mathrm{e}^{-3t}}{1\cdot 2}\right) \end{aligned}$$

the resulting system for the coefficients leads to

$$0 = 2 + D + E + F + B + C$$

that does not have nonnegative solutions.

We think, there is no hope to find a corresponding MP for an arbitrary nonnegative kernel. Therefore we go another way and try to derive a class of sensible kernels starting from physical considerations. Furthermore, the following reasoning shows how the time interval of the memory effect is connected with rates of the loops of the MP.

First of all we have to ask: How one can model a meaningful kernel for a ME. We can assume that the dependence on the past is concentrated at some time point before the present, say $t - t_1$, ... $t - t_N$ where t_j are ordered time values, i.e. $0 < t_1 < t_2 < \cdots < t_N$, with some coefficients $\gamma_1, ..., \gamma_N$ with $\gamma_i \geq 0$ and $\sum \gamma_i = 1$ that gives the relative proportion of each time point. The corresponding memory kernel of such an ansatz is

$$\tilde{K}(t) = \sum_{j=1}^{N} \gamma_j \delta(t - t_j)$$

(here δ means the "δ-function", the "density" of the Dirac measure). The kernel $\tilde{K}$ occurs when starting from a discrete time model, like equation (2). Clearly, this is a first guess. A real memory kernel seems to be more smeared. Therefore, we can try to find kernels $\tilde{K}_j(t)$ with mean time at t_j, i.e

$$\int_0^\infty \tilde{K}_j(t)\mathrm{d}t = a, \quad \int_0^\infty t\tilde{K}_j(t)\mathrm{d}t = \int_0^\infty t\delta(t - t_j)\mathrm{d}t = t_j,$$

We will show that such kernels $\tilde{K}_j(t)$ can be found and it is possible to find a suitable MP for them. Note, that this does not determine the kernels $\tilde{K}_j$ uniquely, of course.

We show that our kernels of shape (20) are suitable for this.

Proposition 3.2. *Let a sequence $0 < t_1 < t_2 < \cdots < t_N < \infty$ be given. There are kernels $K(t) = \sum_{j=1}^{N} a_i K_i(t)$ such that $K \geq 0$ and $\int_0^\infty K(t)\mathrm{d}t = a$ and $\int_0^\infty tK_i(t)\mathrm{d}t = t_i$.*

Proof. We define $b_j \in \mathbb{R}$ via $t_i = \sum_{j=1}^{i} \frac{1}{b_j}$. Since the t_i are ordered, we get $b_j > 0$. We define

$$K(t) = \sum_{j=1}^{N} a_j K_j(t), \text{ where } K_j(t) = \sum_{i=1}^{j} b_i \psi_i^j \mathrm{e}^{-b_i t}.$$

We prove that K satisfies the desired properties. Using the Laplace transform, we get

$$\mathcal{L}(K_j(t))(\lambda) = \sum_{i=1}^{j} b_i \psi_i^j \frac{1}{\lambda + b_i} = \prod_{i=1}^{j} \frac{b_i}{\lambda + b_i} =: k_j(\lambda).$$

This yields $\int_0^\infty K_j(t)\mathrm{d}t = k_j(\lambda = 0) = 1$. Moreover, $\int_0^\infty tK_j(t)\mathrm{d}t = -k_j'(\lambda = 0)$. We have

$$k_j'(\lambda) = \sum_{i=1}^{j} \frac{b_1}{\lambda + b_1} \cdot \frac{b_2}{\lambda + b_2} \cdots \frac{-b_i}{(\lambda + b_i)^2} \cdots \frac{b_{j-1}}{\lambda + b_{j-1}} \cdot \frac{b_j}{\lambda + b_j}$$

This yields $-k_j'(\lambda = 0) = \sum_{i=1}^{j} \frac{1}{b_i} = t_j$, i.e. $\int_0^\infty tK_j(t)\mathrm{d}t = t_j$. □

Theorem 3.3. *Let $K(t)$ be a memory kernel of the form*

$$K(t) = \sum_{j=1}^{N} \alpha_i K_i(t), \text{ where } K_i(t) = \sum_{j=1}^{i} b_j \psi_j \mathrm{e}^{-b_j t}.$$

and $\alpha = \sum_j \alpha_j$. Let u be the solution to the equation $\dot{u}(t) = -\alpha u + K * u$ with $u(0) = u_0$. Then, there is a Markov process $\dot{p} = \mathbf{A}^* p$ in $\mathbb{R}^{N+1}$ generated by a Markov matrix $\mathbf{A}$ and an initial condition $p(0)$ such that $u(t) = p_0(t)$.

Proof. Define the Markov generator matrix via $a = \alpha$, $a_i = \alpha_i$, $b_i = \beta_i$. The initial condition for the Markov process is $p_0 = (u_0, 0, \dots, 0)$. The claim follows. □

For the asymptotic behavior of the ME, we immediately get the following statement.

Corollary 3.4. *Let $K(t) = \sum_{j=1}^{N} a_i K_i(t)$, where $K_i(t) = \sum_{j=1}^{i} b_j \psi_j \mathrm{e}^{-b_j t}$ and $a = \sum_j a_j$. Let u be the solution to the equation $\dot{u}(t) = -au + K * u$ with $u(0) = u_0$. Then $u(t) \to u_\infty$ as $t \to \infty$, where $u_\infty = \frac{1}{Z} u_0$ and Z is given by (16).*

3.3 Remarks

1 Kernels like $k_j(\lambda) = \prod_{i=1}^{j} \left(\frac{b_i}{\lambda + b_i}\right)^{m_i}$ with suitable chosen $m_i \in \mathbb{N}$ may approximates a δ-kernel better. Especially it allows to take into account more moments then only the first one. This requires to allow the b_i to be equal. This is possible without any principial problems (see the note above Theorem 3.1). A special case is treated in the next chapter, where one delay is approximated arbitrary precise. To prove positivity of the corresponding functions Lemma 5.1 from the appendix can be used.

Kernels like in (17) are rational functions of degree N, having poles on the left plane. They approximate meromorphic functions. This makes one able to consider more general kernels then linear combinations of exponents – at least approximately.

2 There are other (similar) MP that lead to a ME and vice versa. For example the MP with the generator

$$\mathbf{A}^* = \begin{pmatrix} -a & c_1 & c_2 & c_3 & \dots & c_N \\ a & -c_1 - b_1 & 0 & 0 & \dots & 0 \\ 0 & b_1 & -c_2 - b_2 & 0 & \dots & 0 \\ 0 & 0 & b_2 & -c_3 - b_3 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & b_{N-1} & -c_N \end{pmatrix},$$

can also be used for embedding the presented exponential kernels. Such MP can be understood in the same manner like at the picture on page 12 but with reversed arrows. Although this approach is more difficulty from a technical point of view.

3 The presented results can be applied in various manner. We focus on ordinary differential equations to present the general idea. Linear Memory equations in infinite dimensional space like diffusion equations with time depending diffusion coefficients are also possible.

Moreover, the well known tools for investigating MP, like inequalities for Lyapunov functions (see [14]) can now be carried over to explore ME.

4 Special Markov process leads to a Delay Differential equation

In this section we consider a special form of the MP. We define $a_j = 0$ for $j = 1, 2, \ldots, N-1$ and put $a_N = a$ and $b_j = b \in \mathbb{R}$. Using the observation from the last section we consider a general cyclic MP with one single but long loop. The Markov process in $\mathbb{R}^{N+1}$ is generated by the matrix

$$\mathbf{A}^* = \begin{pmatrix} -a & b & 0 & \cdots & 0 \\ 0 & -b & b & \cdots & 0 \\ 0 & 0 & -b & \cdots & 0 \\ \vdots & & & \ddots & b \\ a & \cdots & 0 & \cdots & -b \end{pmatrix}.$$

We assume the initial mass is concentrated in the first reservoir. Then, the equation reads

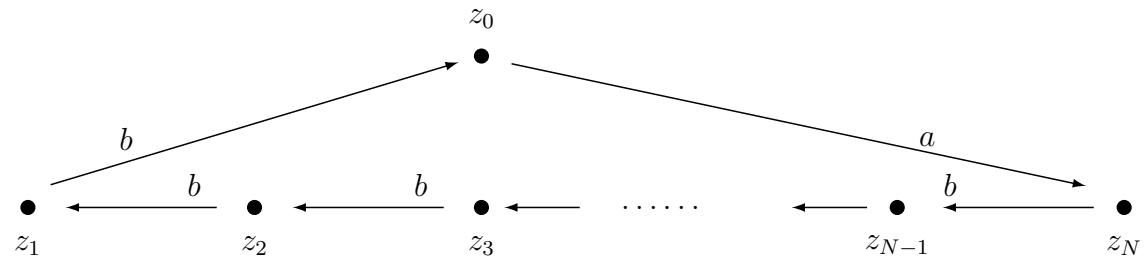
$$\dot{p}(t) = \mathbf{A}^* p(t),\ p(0) = p_0,$$

where $p = (u, v_1, v_2, \ldots, v_n)^T$ and $p_0 = (u_0, 0, \ldots, 0)^T$.

The stationary solution is

$$\mu = \frac{1}{Z}\left(\frac{1}{a}, \frac{1}{b}, \frac{1}{b}, ..., \frac{1}{b}\right)^T u_0 \in \mathbb{R}^{N+1},$$

where $Z = \frac{1}{a} + \frac{N}{b} = \frac{b+aN}{ab}$. Note, the system does not have the detailed balance property.



We get

$$(\lambda + a)\hat{u} - u_0 = a\left(\frac{b}{\lambda + b}\right)^N u.$$

It holds

$$\left(\frac{b}{\lambda + b}\right)^N = \mathcal{L}\left(\frac{b^N}{(N-1)!} t^{N-1} \mathrm{e}^{-bt}\right)(\lambda).$$

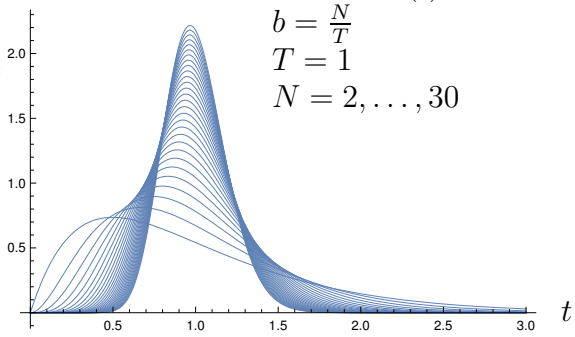
Hence, we get

$$\dot{u}(t) = -au(t) + \frac{ab^N}{(N-1)!}\int_0^t s^{N-1}\mathrm{e}^{-bs}u(t-s)\mathrm{d}s = \\ = -a\left(u(t) - \int_0^t K_N(s)u(t-s)\mathrm{d}s\right),$$

where we introduced the kernel

$$K_N(t) = \frac{b^N}{(N-1)!}t^{N-1}\mathrm{e}^{-bt}.$$

Kernel $K_N(t)$ for:
$b = \frac{N}{T}$
$T = 1$
$N = 2, \ldots, 30$



A delay equation can be understood as a memory equation with a δ-kernel. To do this, we fix $T > 0$ and introduce $\delta_T(t) = \delta(t - T)$. We get

$$\int_0^\infty \delta_T(t)\mathrm{e}^{-\lambda t}\mathrm{d}t = \int_0^\infty \delta(t-T)\mathrm{e}^{-\lambda t}\mathrm{d}t = \mathrm{e}^{-\lambda T}. \tag{24}$$

Moreover, for $t > T$ we have

$$u(t-T) = \int_0^\infty u(s)\delta(t-T-s)\mathrm{d}s = \int_0^\infty u(s)\delta_T(t-s)\mathrm{d}s = \tag{25}$$

$$= \int_0^t u(s)\delta_T(t-s)\mathrm{d}s = u(t) * \delta_T(t). \tag{26}$$

Hence,

$$\mathcal{L}(u(t-T))(\lambda) = \hat{u}(\lambda)\mathrm{e}^{-\lambda T}. \tag{27}$$

Putting $b = \frac{N}{T}$, we approximate the Laplace transform of the kernel δ_T, i.e.

$$\mathcal{L}(\delta_T)(\lambda) = e^{-\lambda T} \approx \left(1 + \frac{\lambda T}{N}\right)^{-N} = \left(\frac{\frac{N}{T}}{\frac{N}{T} + \lambda}\right)^N = \mathcal{L}(K_N(t))(\lambda). \tag{28}$$

Hence, we conclude

$$\mathcal{L}(K_n(t))(\lambda) \xrightarrow{n\to\infty} e^{-\lambda T} = \mathcal{L}(\delta(t-T))(\lambda), \tag{29}$$

and the limiting (DDE) reads as

$$\dot{u} = \begin{cases} -au(t), & \text{if } 0 \le t \le T \\ -au(t) + au(t-T), & \text{if } t \ge T, \end{cases} \tag{30}$$

or equivalently

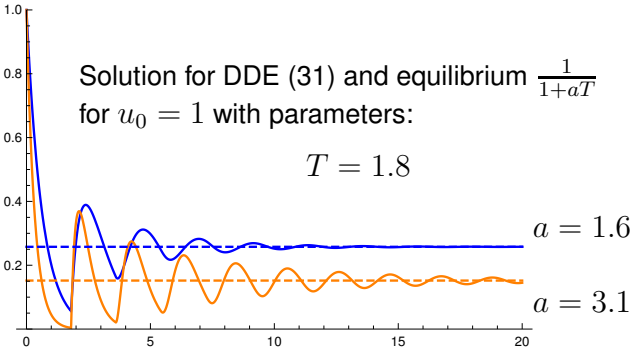
$$\dot{u} = -au(t) + au(t-T), \text{ for } t \ge T, \quad \text{and} \quad u|_{[0,T]}(t) = \mathrm{e}^{-at}u_0. \tag{31}$$

Let us note that the initial condition $u|_{[0,T]}(t) = \mathrm{e}^{-at}u_0$ results from the modeling ansatz. No other initial condition is possible. Here, the initial condition was calculated in contrast to the usual way, to guess it.

Let us compute the limiting stationary solution for $N \to \infty$ of the first coordinate of the MP. This means the MP has long loops, but mass is transferred with a high rate. We have $Z = \frac{Na+b}{ab}$. Putting $b = \frac{N}{T}$, we conclude for the zeroth coordinate of the stationary solution.

$$\mu_0 = \frac{1}{Za} = \frac{b}{Na+b} = \frac{\frac{N}{T}}{Na + \frac{N}{T}} = \frac{1}{1+aT}.$$

The solution of the DDE and the stationary solution μ_0 of the MP can be seen in the picture. The solution of the DDE converges nicely to μ_0.



Finally, we remark some properties of the spectrum. The spectrum of the DDE is given by inserting $\mathrm{e}^{\lambda t}$ for $\lambda \in \mathbb{C}$ into the equation (see e.g. [12]). This yields for given $a, T \geq 0$ the equation

$$\lambda = -a + a\mathrm{e}^{-\lambda T}. \tag{32}$$

This transcendental equation (in $\lambda \in \mathbb{C}$) has in general an infinite discrete amount of solutions.

The eigenvalues of $\mathbf{A}^*$ for fixed $N \in \mathbb{N}$ are given by the characteristic equation

$$\phi(\lambda) = -ab^{N-1} + (\lambda + b)^{N-1}(\lambda + a) = 0,$$

that can be computed easily. Hence, setting $b = \frac{N}{T}$ we get $\phi(\lambda) = 0$ if and only if

$$\frac{a}{a+\lambda} = \left(\frac{\lambda + b}{b}\right)^{N-1} = \left(1 + \frac{\lambda T}{N}\right)^{N-1}.$$

For $N \to \infty$, right hand side converges to $\mathrm{e}^{\lambda T}$. So, in the limit $\lambda \in \mathbb{C}$ satisfies the equation

$$\frac{a}{a+\lambda} = \mathrm{e}^{\lambda T},$$

i.e. the same equation as (32). In this sense, one can say that not only the solution converges but also the spectrum of the MP and of the ME converges to each other. Note, that the convergence of the spectrum is very slow, as the convergence of the exponential function is.

5 Appendix

5.1 Laplace transform

Here, we summarize some facts of the Laplace transform. More details can be found in [15]. For a given function $u : [0, \infty) \in t \mapsto u(t) \in \mathbb{R}$ that does not grow faster than an exponential function in time, the Laplace transform is defined by

$$\hat{u}(\lambda) = (\mathcal{L}u)(\lambda) = \int_0^\infty e^{-\lambda t} u(t) \mathrm{d}t$$

We will use the following formulas that can be checked easily:

$$\begin{aligned}
\mathcal{L}(\dot{u})(\lambda) &= \lambda\hat{u}(\lambda) - u_0 \\
\mathcal{L}(K * u) &= (\mathcal{L}K) \cdot (\mathcal{L}u) \\
\mathcal{L}(\mathrm{e}^{-a\cdot})(\lambda) &= \frac{1}{\lambda + a} \\
\mathcal{L}\left(\frac{1}{(n-1)!} t^{n-1}\mathrm{e}^{-at}\right)(\lambda) &= \frac{1}{(\lambda + a)^n}
\end{aligned}$$

The Laplace transform has an interesting asymptotic behavior. The limit for large times $u(t) \overset{t\to\infty}{\longrightarrow} u_\infty$ can be calculated with the Laplace transform. It holds $\lambda\hat{u}(\lambda) \overset{\lambda\to 0}{\longrightarrow} u_\infty$. Thus, there is no need to know the whole solution $u(t)$ if one is interested only in the equilibrium case. This is important since, in general for non-autonomous equations, the equilibrium case can not be calculated by setting $\dot{u} = 0$.

Let us note that the uniform convergence on compact sets of $t \in \mathbb{R}_+$ carries over to uniform convergence on compact sets of λ in the domain of analyticity.

To carry over positivity properties between the original and the transformation the following Lemma is usefull:

Lemma 5.1. *Let* $K(t) = \sum_{j=1}^N \gamma_j \mathrm{e}^{-\alpha_j t}$ *with its Laplace transform* $k(\lambda) = \sum_{j=1}^N \gamma_j \frac{1}{\lambda+\alpha_j}$. *Then* $K(t) \geq 0$ *if and only if* $\sum_{j=1}^N \frac{\gamma_j}{(\lambda+\alpha_j)^m} \geq 0$ *for any* $m \in \mathbb{N}$.

Proof. Let $K(t) \geq 0$. Since $K(0) \geq 0$, we get $\sum_{j=1}^N \gamma_j \geq 0$, i.e. the claim holds for $m = 0$. For $m \geq 0$, we get $0 \leq \int_0^\infty t^m K(t) \mathrm{e}^{-\lambda t} \mathrm{d}t = (-1)^m k^{(m)}(\lambda) = \sum_{j=1}^N \frac{\gamma_j}{(\lambda+\alpha_j)^{m+1}}$ what proves the claim in one direction.

For the other direction, we put $\lambda = \frac{n}{t}$ and $m + 1 = n$. Then

$$0 \leq \sum_{j=1}^N \frac{\gamma_j (\frac{n}{t})^n}{(\frac{n}{t} + \alpha_j)^n} = \sum_{j=1}^N \frac{\gamma_j}{(1 + \frac{\alpha_j n}{t})^n} = \sum_{j=1}^N \gamma_j \left(1 + \frac{\alpha_j t}{n}\right)^{-n} \to \sum_{j=1}^N \gamma_j \mathrm{e}^{-\alpha_j t}, \text{ as } n \to \infty,$$

which proves the claim of the Lemma. □

5.2 Simplex integrals

In Theorem 3.1, we proved the positivity of the kernel $K(t)$ using an integral over a simplex. This is based on the following observation.

Let $S_{n-1} \subset \mathbb{R}^n$ be the simplex, defined as

$$S_{n-1} = \{s \in \mathbb{R}^n \mid s_i \geq 0,\ s_1 + ... + s_n = 1\}$$

We consider functions $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ and their integrals over S_{n-1}. We have

$$\begin{aligned} \int\limits_{\mathcal{S}_{n-1}} g(s) \mathrm{d}\sigma(s) &= \frac{1}{\sqrt{n}} \int\limits_{S_{n-1}} g(s_1, s_2, ..., s_{n-1}, 1 - s_1 - \ldots - s_{n-1}) \mathrm{d}s_1 \cdots \mathrm{d}s_{n-1} = \\ &= (n-1)! \int_0^1 \mathrm{d}s_1 \int_0^{1-s_1} \mathrm{d}s_2 \int_0^{1-s_1-s_2} \mathrm{d}s_3 \cdots \int_0^{1-s_n-\ldots-s_{n-2}} \mathrm{d}s_{n-1}\, g(s_1, s_2, ..., s_n)\Big|_{s_n = 1-s_1-\ldots-s_{n-1}}, \end{aligned}$$

where $\sigma(\mathrm{d}s)$ is the Lebesgue measure on S_{n-1} and $\sqrt{n}$ is the volume of S_{n-1}.

Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a smooth enough function, $f^{(k)}$ its k- derivative and $x_1, ..., x_n$ given different reals. Set $g(s) = f(\langle x, s \rangle)$, where $\langle x, s \rangle = x_1 s_1 + x_2 s_2 + \ldots + x_n s_n$ is the scalar product in $\mathbb{R}^n$.

Now, using induction one can prove that

$$\sum_{i=1}^n f(x_i) \prod_{j \neq i}^n \frac{1}{x_i - x_j} = \int\limits_{S_{n-1}} f^{(n-1)}(\langle x, s \rangle) \sigma(\mathrm{d}s)\,.$$

This formula gives a powerfull tool to switch between expessions connected with Lagrange polynomials and expessions connected with simplex integrals. In Theorem 3.1, we used this formula with $f(x) = \mathrm{e}^{-xt}$.

5.3 Lagrange polynomials

Here we summarize basic facts from the theory of Lagrange polynomials. Let

$$L_i^j(x) = \prod_{k=1,k\neq i}^{j} \frac{x - x_k}{x_i - x_k} \,,$$

assuming $x_i \neq x_k$ for $i \neq k$. Obviously $L_i^j(x)$ is a polynomial of degree $j-1$ and we have $L_i^j(x_k) = \delta_{ik}$ with δ_{ik} the Kronecker symbol. Hence, the polynomial

$$P(x) = \sum_{i=1}^{j} p_i L_i^j(x)$$

of degree $j-1$ satisfy $P(x_i) = p_i$.

Seeking for a polynomial $P(x) = q_0 + q_1 x + ... + q_{j-1}x^{j-1}$ with the condition $P(x_i) = p_i = \frac{x_i}{z+x_i}$ we get as the result coefficients q_i with $q_0 = \prod_{i=1}^{j} \frac{x_i}{z+x_i}$ among them. Hence, we have on the one hand

$$P(0) = q_0 = \prod_{i=1}^{j} \frac{x_i}{z + x_i}$$

and on the other hand

$$P(0) = \sum_{i=1}^{j} p_i L_i^j(0) = \sum_{i=1}^{j} \frac{x_i}{z + x_i} \prod_{k=1,k\neq i}^{j} \frac{(-x_k)}{x_i - x_k} = \sum_{i=1}^{j} \frac{x_i}{z + x_i} \prod_{k=1,k\neq i}^{j} \frac{x_k}{x_k - x_i} \,.$$

It follows

$$\prod_{i=1}^{j} \frac{x_i}{z + x_i} = \sum_{i=1}^{j} \frac{x_i}{z + x_i} \prod_{k=1,k\neq i}^{j} \frac{x_k}{x_k - x_i} \,.$$

Note, in our explanation we used $\psi_i^j = (-1)^{j-1} L_i^j(0)$.

References

[1] W. ARENDT, ET AL., *One-parameter Semigroups of Positive Operators*, Springer, 1986.

[2] A. BOBROWSKI, *Functional Analysis for Probability and Stochastic Processes*, Cambridge University Press, 2005.

[3] R. DURRETT, *Probability: Theory and Examples*, Cambridge University Press, 2010.

[4] E.B. DYNKIN, *Markov Processes*, Springer, 1965.

[5] K. ENGEL, R. NAGEL, *One-Parameter Semigroups for Linear Evolution Equations*, Graduate texts in mathematics, 194. Springer-Verlag, New York, Berlin, Heidelberg, 1999.

[6] M.E. GURTIN, A.C. PIPKIN, *A general Theory of Heat Conduction with Finite Wave Speeds*, Arch. Rational Mech. Anal. 31, p. 113–126, 1968.

[7] J.K. HALE, S.M.V. LUNEL, *Introduction to functional differential equations*, Springer, 1993.

[8] U. HORNUNG, R.E. SHOWALTER, *Diffusion Models to Fractured Media*, Journal of Mathematical Analysis and Applications, 147, p. 69–80, 1990.

[9] A. KHRABUSTOVSKYI, H. STEPHAN, *Positivity and time behavior of a general linear evolution system, non-local in space and time*, Math. Methods Appl. Sci., p. 1809–1834, 2008.

[10] V. KOLMANOVSKII, A. MYSHKIS, *Introduction to the Theory and Applications of Functional Differential Equations*, Springer, 2013.

[11] M. PESZYNSKA, *Analysis of an integro-differential equation arising form modelling of flows with fading memory through fissured media*, Journal of Partial Differential Equations 8., 1991.

[12] H. SMITH, *An Introduction to Delay Differential Equations with Applications to the Life Sciences*, Springer, 2010.

[13] H. STEPHAN, *A dimension-reduced description of general Brownian motion by non-autonomous diffusion-like equations*, J. Mathematical Physics Analysis Geometry (MAG), 12, p. 187–202, 2005.

[14] H. STEPHAN, *Lyapunov functions for positive linear evolution problems*, ZAMM Z. Angew. Math. Mech., 85, p. 766–777, 2005.

[15] B.YU. STERNIN, V.E. SHATALOV, *Borel-Laplace Transform and Asymptotic Theory: Introduction to Resurgent Analysis*, CRC Press, 288 pages, 1995.

[16] L. TARTAR, *Memory Effects and Homogenization*, in: Mechanics and Thermodynamics of Continua, Springer, p. 537–549, 1991.