BACHELOR

Mathematical analysis of generalized mode-coupling theory and numerical exploration of super-strong glass formation

Biezemans, Rutger A.

Award date:
2018

Technische Universiteit Eindhoven University of Technology

Department of Mathematics and Computer Science and
Department of Applied Physics

Mathematical analysis of generalized mode-coupling theory and numerical exploration of super-strong glass formation

Bachelor End Project

Rutger A. Biezemans
(Student number: 0957195)

Assessment committee M&CS:	Assessment committee AP:
Dr. B. Baumeier	Dr. L.M.C. Janssen
Dr. L. Iapichino	S. Ciarella
(Course code: 2WH40)	(Course code: 3FEX0)
Daily guidance M&CS:	
O. Caylak	

Eindhoven, July 2018

Abstract

Glass formation is one of the deepest challenges today in condensed matter physics. Only mild changes in structural properties in a supercooled liquid result in a dramatic change of dynamical properties, such as a material's viscosity. A better fundamental understanding of the glass transition would give us an insight in the underlying effects of this and could improve our material design and processing abilities. At present, mode-coupling theory (MCT) is the only fully microscopic theory to describe this. It is approximated as schematic MCT for analytical insight. In this project, an improvement on the theory, known as schematic generalized MCT (GMCT), is mathematically analyzed. It concerns a system of coupled, non-linear integro-differential equations that describe the intermediate scattering function of a material. We show, detailing and elaborating on the work of others, that a unique solution to the equations exists when the equations are obtained with the so-called exponential closure, and we can establish this also for one specific mean-field closure. For the over-damped approximation to the schematic GMCT equations, we find that a unique solution always exists. The earlier works we used either exploit linear aspects of the equation, or linearise the equations to show the result with a convergence property. We then aim to model super-strong glass formation observed in vitrimer simulations numerically. We find the Power Structure Factor (PSF) model, relating this glass formation to a static structure factor of the form $S(\Lambda) = 1 + \Lambda^{1/\alpha}$ where Λ is an inverse temperature, as one of the models that could be useful for the observed phenomena. In the numerical exploration, we find that the ratio of the frequency and damping coefficients in the schematic MCT equation is crucial. When looking at the correlation functions themselves, we find numerically a data collapse if we scale the time for each correlation function by the associated relaxation time. An analytic approach to explain this behaviour shows that this would hold universally if the relaxation time can be shown to depend on the coefficient ratio only.

Contents

1 Introduction

One of the major challenges in condensed matter physics is to understand the dynamics of supercooled liquids [16]. In particular, during their vitrification (or glass formation) the viscosity can grow orders of magnitude upon only mild variations in control parameters such as temperature or density, while also showing only weak changes in structural properties [13]. Phenomenologically, the glass transition, as it is also referred to, has extensively been studied and we will introduce the main concepts below. Theoretically, on the contrary, there is only one strictly first-principles, microscopically based theory that seeks to predict these phenomena: mode-coupling theory (MCT) [14], and there are still many phenomena that cannot be accounted for. Glasses are abundant and have among others applications in window panes, phase-change memory devices, pharmaceutical compounds and wearable electronics [13]. A better theoretical understanding of glass formation could therefore increase our capabilities in relevant material control and design [4].

In short, MCT studies the density correlation function in the Fourier domain and leads to an integro-differential equation with non-linear memory kernel which also depends on the unknown correlation function. To allow for analytical predictions, a wave vector independent approximation of the theory known as schematic MCT is often considered. In recent years, the theory has extended to include exact equations for the memory kernel, leading to a *system* of integro-differential equations. The system is in principle an infinite hierarchy, but for numerical results a so-called closure is used to cut off the system at a finite level. This system, in the schematic approximation, is the subject of an extensive mathematical analysis in this report, which will eventually bring us to existence and uniqueness results for the theory that is used in the physical practice. We can derive these from earlier work on the appropriate type of equations, whose proofs we provide in detail and aim to improve if necessary.

1.1 Glass formation: phenomenology

In this introduction, we first provide a discussion of the phenomenon of glass formation. We think this is necessary to give a mathematical audience some background as to what is studied here, but also to appropriately introduce the concepts that underlie the physical research question to be formulated below.

Microscopic structure
To study the properties of a glass, we start from a liquid. This can be turned into a glass by cooling down or compressing the material such that the viscosity exceeds a value of 10^{12} Pa.s (this is just a definition; the central concept is many orders-of-magnitude growth), but sufficiently quickly such that the liquid cannot crystallize. We thus supercool the liquid to below its crystallization temperature. A glass therefore is a solid material, lacking any long-range structure [13].

To describe the microscopic structure of a material, the radial distribution function $g(r)$ is typically used. This describes the probability of finding a particle in a thin, spherical shell a distance r away from some reference particle, with respect to this probability in an ideal gas. In an ideal gas, the particles are assumed to be evenly distributed over the whole occupied volume. For a typical crystalline solid, which has a well-defined and repetitive structure, $g(r)$ has sharp peaks at fixed intervals and vanishes in between. For a liquid, $g(r)$ is typically more smooth, but also displays peaks of decreasing height at distances corresponding to multiples of the particle diameter (referred to as the solvation shells). However, for large values of r, it becomes identical to an ideal gas, i.e. $g(r) \to 1$, since a liquid has no long-range structure. Now for an amorphous solid or a glass, the radial distribution function is very similar to the liquid state, although its dynamical properties clearly are not (leading to a giant growth of the viscosity). For a pictorial representation of $g(r)$, we refer to e.g. [13], [25].

In MCT, we study the microscopic structure in Fourier space, where the characteristics of $g(r)$ are captured by the static structure factor $S(k)$ as a function of the wave number k. It displays properties similar to those of $g(r)$ and in particular has a main peak at the wave number corresponding to the first solvation shell. It flattens out for greater wave numbers. We give a schematic representation of $S(k)$ in figure 1(a).

Relaxation
To study the dynamics of a glass-forming liquid, we use the density-density correlation function or intermediate scattering function F in the (spatial) Fourier domain over time. In words, it tells us how much the structure as given by g at some time t resembles the structure at some reference time 0. Once it has decayed to zero, the structure has become uncorrelated to the initial structure. At high temperatures, in the normal liquid state, this function will always decay to zero in a fashion similar to the representative curve in figure 1(b), within some time frame that is sufficient for the particles to move to new positions. At very short times, they are even free to move in any direction (this is before they bump into other particles and are influenced by the resulting interactions), which is called the ballistic regime. Thereafter, the microscopic regime follows and eventually, the function relaxes to zero [17].

At lower temperatures, when we have a supercooled liquid, the decay will look like the other curves in figure 1(b). After the microscopic regime, a plateau is attained. In this time regime, the particles are trapped by their neighbouring particles and can only move within a certain 'cage', this is the so-called cage effect. This time regime is referred to as the β-relaxation. Only for much larger times the particles will be able to escape from these cages such that the structure of the material is allowed to continue to change and F can decay to zero. This decay is called the α-relaxation regime. Based on experimental results, many features of the β- and α-relaxation can be described by appropriate fitting models. Although these are not microscopically motivated, we will see that MCT verifies some of these results [17].

When no relaxation to zero occurs any more, we have a liquid whose constituent particles can no longer move sufficiently to relax and we speak of an amorphous solid or a glass. In figure 1(b), this is illustrated by a curve for F staying in the β-regime plateau instead. This is called the (ideal) glass transition. Ideal, because it is not clear whether in reality F will never fully decay. We therefore speak of an experimental glass transition when the relaxation time increases over five orders of magnitude with respect to a typical high-temperature relaxation time, for this defines an experimental time scale on the order of about a week.

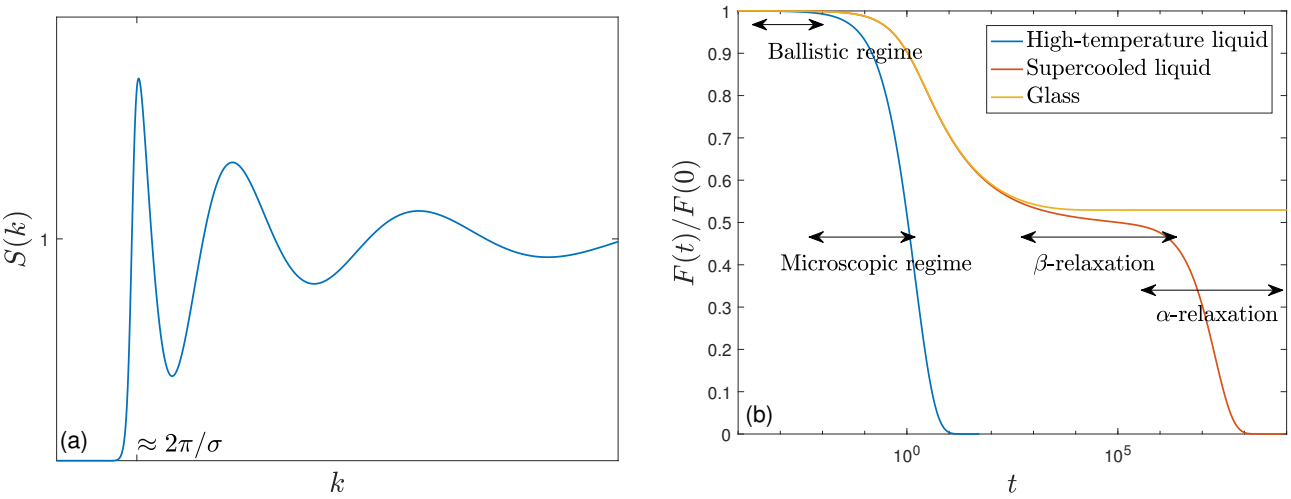

Figure 1: *Structural and dynamical properties of liquids. (a) Schematic representation of the structure factor S with the main peak at wave number $k \approx 2\pi/\sigma$, where σ is the first solvation shell. (b) Normalised intermediate scattering function F over time (note the logarithmic scale) for a high-temperature liquid, a supercooled liquid and a glass. For the latter, F cannot relax to 0 any more. The arrows indicates the ballistic, microscopic, β-relaxation and α-relaxation regimes. The last two can only be observed for supercooled liquids. The lines are representative numerical solutions of the MCT equations, which capture this phenomenological behaviour.*

Fragility

The relaxation time of the intermediate scattering function is the time needed to relax, or in the research practice the time needed to reach some typical small value. It is related to the viscosity with the shear modulus as a proportionality factor [13]. We can therefore use MCT to study F and obtain information about the viscosity. Of particular interest is then the slope of the logarithm of the viscosity or relaxation time of the material as a function of inverse temperature. For an example of what this might look like, the reader is invited to look ahead at various outcomes for the relaxation time in figure 9 of chapter 7. If such a graph is a straight line, this is called a strong glass-former. It corresponds to exponential growth of the relaxation time with inverse temperature and this growth-behaviour is called Arrhenius-type. Fragile glass formers have a smaller slope at higher temperatures and increasing slope at higher temperatures [17]. Formally, the fragility is described in terms of the slope at the glass transition. This is of practical importance, because the material might be less easily processable in temperature ranges with a high slope. Indeed, the viscosity would increase more and more strongly upon cooling and a very precise temperature control is necessary for processing, while a strong glass-former has a more gradual change of the viscosity [23].

As said, we will use MCT to predict the relaxation time. The only structural information that this theory needs is the static structure factor. Via mild changes of this structure factor with decreasing temperature, MCT will be able to predict a glass transition and accordingly a dramatic increase in the relaxation time over several orders of magnitude.

1.2 Research goals and structure of the report

This report has partly a mathematical focus and partly a physical one. Mathematically, we are interested in the existence and uniqueness of a solution to the schematic generalized MCT equations. We will study this system of integro-differential equations based on the work of Götze *et al.* [10] and Saal [29]. The research goal here is:

To obtain a proof for the existence and uniqueness of solutions to the schematic generalized MCT equations, detailing previous work such that the argument is understandable and convincing for a student at a third year Bachelor's level in mathematics.

We formally state the equations to be studied in chapter 3. We then start with preliminary mathematical results in chapter 4 (part of which is just a repetition and part of which is aimed at developing an understanding of new material) from real and complex analysis concerning inequalities, convergence, the Laplace transform and matrices in the context of differential equations. Also included are some notions and results from functional analysis, topology and variational calculus, in particular differentiation of normed operators. In chapter 5 we will then be able to develop the proofs for three existence and uniqueness results, building new arguments on top of the preliminary results where necessary. Then in chapter 6 we can slightly adapt the foregoing results such that they capture the exact structure of the equations in chapter 3 and provide the desired results for specific closures. Here we mainly show how certain inequalities are satisfied, but will for one case need some physical insight to allow for a mathematical trick before we can finalize the proof. We give a conclusion and outlook at the end in chapter 9. These chapters form the mathematical contents of this project.

The physical part of the project is motivated by vitrimers. This is a name for covalent polymer networks that can change their topology by thermally activated bond exchange reactions [3]. For vitrimers, these reactions are reversible such that material properties at high temperatures and at low temperatures are not lost on repetitive cooling and heating. Montarnal *et al.* have studied such a material and observed an Arrhenius behaviour of the fragility when approaching the glass transition, unlike for organic compounds or other polymers. At higher temperatures, the material is very processable. This way, the dimensional stability of thermosets and the processability and reparability of thermoplastics are combined. [23].

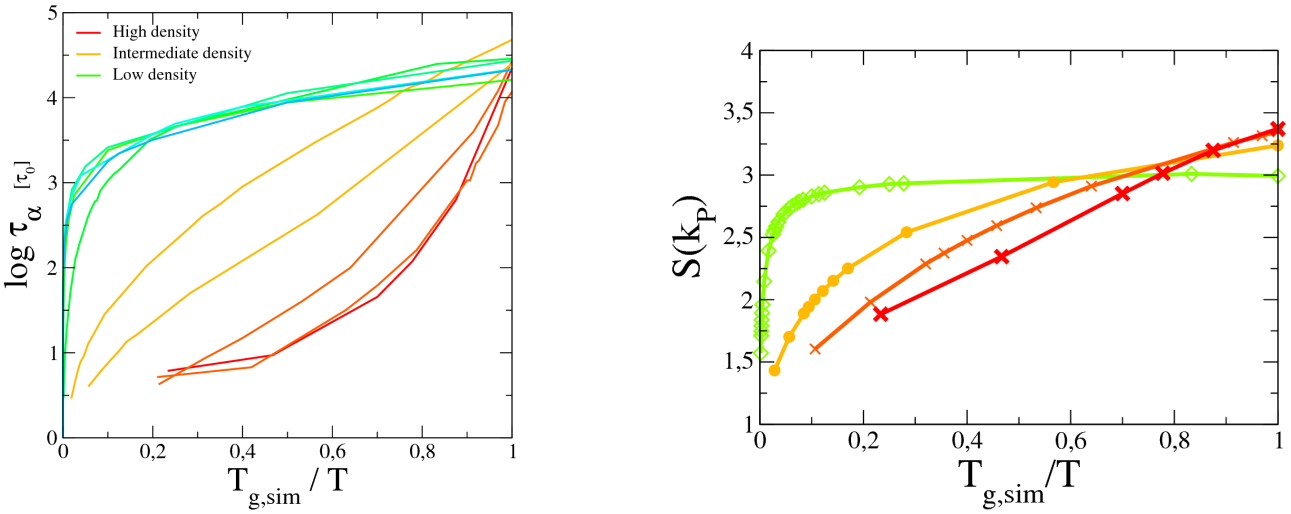

Figure 2: *Simulation results for glass-forming vitrimers. Left: logarithm of the (α-)relaxation time τ_α as a function of inverse temperature $1/T$ normalised by the glass transition temperature of the simulation $T_{\mathrm{g,sim}}$. For decreasing density, the fragility changes from weak to super-strong. Right: height of the main peak of the static structure factor S, at wave number k_p. Here the density also decreases from red to green. The qualitative change of shape for these graphs seems to be related to the qualitative change of glass-formation. The pictures were provided by S. Ciarella.*

In computer simulations done by S. Ciarella, a form of super-strong glass formation appeared where, near the glass transition, the growth of the logarithm of the viscosity becomes less strong with decreasing temperature. This same type of growth is also seen in the main peak of the static structure factor, as illustrated by figure 2. In this project, we try to establish a connection between the static structure factor and this growth of the relaxation time. We note that MCT always predicts a power-law divergence of the relaxation time at the glass transition, i.e., fragile behaviour. As a result, curves starting as those for the relaxation time in figure 2 will display a cross-over from a convex shape to a concave one. Therefore we try to obtain this cross-over at a value of the relaxation time on the order of 10^5, which we used to define an experimental glass and expresses

large practical time ranges. We will come back to precise definitions later in the report, but can now state the research question we seek to answer:

Can structural information in the form of a model for the static structure factor be used in schematic MCT to predict an increase of the relaxation time at the cross-over between super-strong and fragile behaviour beyond the experimental glass transition?

We start the physical aspect in chapter 2 with the theoretical background of MCT and its predictions. After the mathematical goal has been completed in chapter 6, We explore various models for schematic MCT and their results in chapter 7. In this chapter, we aim to answer the research question above by first looking at different models for the static structure factor and then building on these results to look for a way to shift the relaxation time to high time values. Then in chapter 8, we try to have a more physical look at the results of chapter 7 than the purely numerical view. We end with a conclusion and outlook for the whole report in chapter 9. These chapters might be of particular interest to the physically motivated reader.

2 Schematic generalized mode-coupling theory: physics and assumptions

This chapter introduces the theory behind the GMCT equations, which will be extensively studied mathematically in subsequent chapters. We do not aim to provide full technical details of their derivation, but rather wish to highlight the physics that underlie it and the assumptions that may limit their validity. Subsequently, we will give some attention to the predictions of the theory and the physical understanding that they bring.

Let us begin by defining the quantities needed to study the physical properties of interest. We look at a collection of N particles whose positions we describe by functions $\mathbf{r}_j(t)$ of time in our space ($j = 1, \ldots, N$). Then the (real-)space density ρ as measured with respect to a point $\mathbf{r}$ in space at time t is defined by

$$\rho(\mathbf{r}, t) = \sum_{j=1}^{N} \delta(\mathbf{r} - \mathbf{r}_j(t)). \tag{2.1}$$

We space Fourier transform this quantity to obtain the density modes $\rho(\mathbf{k}, t)$ as a function of wavevector $\mathbf{k}$, i.e.,

$$\begin{aligned} \rho(\mathbf{k}, t) &= \int_{\text{space}} e^{i\mathbf{k}\cdot\mathbf{r}} \rho(\mathbf{r}, t)\, \mathrm{d}\mathbf{r}, \\ &= \sum_{j=1}^{N} e^{i\mathbf{k}\mathbf{r}_j(t)} \end{aligned} \tag{2.2}$$

Now the dynamics of the system over time are described in this work (following e.g. [25]) by considering the intermediate scattering function $F(k, t)$, as a function of wave number and $k = |\mathbf{k}|$ and time t, defined by

$$F(k, t) = \frac{1}{N} \left\langle \rho(-\mathbf{k}, 0)\rho(\mathbf{k}, t) \right\rangle, \tag{2.3}$$

where the brackets denote a canonical ensemble average. The intermediate scattering function only depends on the wave number, instead of the actual wavevector in the right-hand side of (2.3), assuming isotropy of the material. This is the case for a powder or a 'simple' fluid, but not generally e.g. in the presence of an external field. [13]. This quantity, besides forming the backbone of our theory, is also of practical importance because it is information of the fluid that can be measured in inelastic neutron scattering experiments [17]. The static structure factor $S(k)$ is this structural information at time $t = 0$, i.e.,

$$S(k) \equiv F(k, 0) = \frac{1}{N} \left\langle \rho(-\mathbf{k}, 0)\rho(\mathbf{k}, 0) \right\rangle. \tag{2.4}$$

It is a Fourier transform-like variant of the radial distribution function $g(\mathbf{r})$, to which it is related by [25]

$$S(k) = 1 + \rho \int_{\text{space}} e^{-i\mathbf{k}\cdot\mathbf{r}} g(\mathbf{r})\, \mathrm{d}\mathbf{r}, \tag{2.5}$$

where $\rho = N/V$ is the bulk density of the system. We have briefly discussed the qualitative properties of these quantities in the introduction.

2.1 Mode-coupling theory

Mode-coupling theory (MCT) starts from two basic observations: firstly, it is used that the structural properties of a glass-forming liquid are very similar to those of the high-temperature liquid. (We have seen in the introduction that this is phenomenologically well-founded [17], as barely any structural change emerges when a fluid transitions to a glass.) Secondly, it builds on the observation that some dynamic variables evolve *fast* (such as vibrations or thermal noise), while others vary much *slower*, so there is a separation of time scales. The Mori-Zwanzig projection operator formalism allows one to derive exact equations of motion for the dynamical variables of interest.

This procedure can be sketched as follows: for a classical fluid, obeying Newton's laws of motion, it is known that the time evolution of any phase space variable $\underline{A}$ can be described by $\frac{d}{dt}\underline{A}(t) = i\mathcal{L}\underline{A}(t)$, whose solution is given by $\underline{A}(t) = e^{i\mathcal{L}t}\underline{A}(0)$, where $\mathcal{L}$ is the Liouville operator (but we do not go into its details here) [25]. In the Mori-Zwanzig formalism, we choose for $\underline{A}$ the slowly varying variables. Next one defines a scalar product for the phase space variables (note that $\underline{A}$ may actually be a vector of such variables) by means of the canonical ensemble and this is used to define a projection operator $\mathcal{P}_A$ onto the variable $\underline{A}$. By this method, each variable in the Hilbert space that they are described in, might have a component in our space of interest (slow dynamics), and a counterpart in the orthogonal complement.

Manipulating the exact equation of motion with this projection formalism and defining the correlation matrix $\underline{C}(t) = \langle \underline{A}^*(0)\underline{A}(t)\rangle$, one can derive that an exact equation of motion for these is

$$\frac{d\underline{C}(t)}{dt} = i\underline{\Omega}\cdot\underline{C}(t) - \int_0^t \underline{K}(\tau)\cdot\underline{C}(t-\tau)\,\mathrm{d}\tau. \tag{2.6}$$

Here, $\underline{\Omega}$ is some matrix called the frequency matrix and $\underline{K}$ a matrix called the memory function (their definitions can be found e.g. in [25]). $\underline{K}$ is the time-autocorrelation function of the 'fast' fluctuating force $\underline{f}$, which describes the fast components influencing the dynamics of $\underline{A}$. It is the difficulty of determining $\underline{K}(t)$ that still imposes a problem on finding $\underline{C}(t)$ from equation 2.6.

To apply this theory to molecular glass-forming liquids, the variables one considers are the density modes defined by equation (2.2) and the associated current modes

$$j(\mathbf{k},t) \equiv \frac{d}{dt}\rho(\mathbf{k},t) = i\sum_{l=1}^{N}\left(\mathbf{k}\cdot\dot{r}_l(t)\right)e^{i\mathbf{k}\cdot\mathbf{r}_l(t)}. \tag{2.7}$$

So we set $\underline{A}(t) = (\rho(\mathbf{k},t), j(\mathbf{k},t))^T$ for some wave vector.

In this setting, we are ultimately interested in the matrix element $C_{21}(t)$ in equation (2.6), because it gives us the time derivative of $F(k,t)$. Using that time dynamics are described by $\mathcal{L}$, which is a Hermitian operator (this means that its position in an inner product can be swapped without any modifications), we can see this as follows:

$$\begin{aligned}
\frac{d}{dt}F(k,t) &= \frac{1}{N}\frac{d}{dt}\left\langle \rho(\mathbf{k},0)^*\rho(\mathbf{k},t)\right\rangle, \\
&= \frac{1}{N}\left\langle \rho(\mathbf{k},0)^*\frac{d}{dt}e^{i\mathcal{L}t}\rho(\mathbf{k},0)\right\rangle, \\
&= \frac{1}{N}\left\langle \rho(\mathbf{k},0)^* i\mathcal{L}e^{i\mathcal{L}t}\rho(\mathbf{k},0)\right\rangle, \\
&= -\frac{1}{N}\left\langle i\mathcal{L}\rho(\mathbf{k},0)^* e^{i\mathcal{L}t}\rho(\mathbf{k},0)\right\rangle, \\
&= -\frac{1}{N}\left\langle \frac{d}{dt}\rho(\mathbf{k},0)^* e^{i\mathcal{L}t}\rho(\mathbf{k},0)\right\rangle, \\
&= -\frac{1}{N}\left\langle j(\mathbf{k},0)^* e^{i\mathcal{L}t}\rho(\mathbf{k},0)\right\rangle = -\frac{1}{N}C_{21}(t).
\end{aligned}$$

Some approximations are presented in [25] regarding the specific form of $\underline{f}$ to simplify $\underline{K}$ in the exact equation (2.6). Firstly, it is assumed that pair densities of the form $\rho_{\mathbf{k_1}}\rho_{\mathbf{k_2}}$ for all wave vectors $\mathbf{k_1}, \mathbf{k_2}$ dominate the slow mode in the time dependence of the fluctuating force $\underline{f}$. This turns its autocorrelation function $\underline{K}$ in the form of a sum over four-point density correlation functions $\left\langle \rho^*(\mathbf{k_1},0)\rho^*(\mathbf{k_2},0)e^{i[1-\mathcal{P}_A]\mathcal{L}t}\rho(\mathbf{k_3},t)\rho(\mathbf{k_4},t)\right\rangle$. Mathematically, this comes down to the first-order non-vanishing projection of $\underline{f}$ onto the space of our variables of interest [13].

Secondly, these four-point correlations are assumed to factorize into two-point correlations and the operator $\exp[i[1-\mathcal{P}_A]\mathcal{L}t]$ is replaced by $\exp[i\mathcal{L}t]$ to prevent that all terms vanish (due to projection by $1-\mathcal{P}_A$ on the fast space). This means that in further steps, we make the replacement $\langle\rho^*(\mathbf{k_1},0)\rho^*(\mathbf{k_2},0)\rho(\mathbf{k_3},t)\rho(\mathbf{k_4},t)\rangle \to \langle\rho^*(\mathbf{k_1},0)\rho(\mathbf{k_1},t)\rangle\langle\rho^*(\mathbf{k_2},0)\rho(\mathbf{k_2},t)\rangle = N^2F(\mathbf{k_1},t)F(\mathbf{k_2},t)$. This factorization would be exact for Gaussian variables, but density modes in general are not Gaussian. It is an ad hoc approximation, which only serves the purpose of simplification. Since we have, with the approximations, related (or coupled) the unknown correlations (or modes) to our variable F, we speak of mode-coupling theory at this point.

Working out the details of the approximations described above, the MCT equation of motion for $F(k,t)$ is obtained from (2.6) as

$$\begin{aligned}
0 &= \frac{d^2F(k,t)}{dt^2} + \frac{k^2k_BT}{mS(k)}F(k,t) + \int_0^t \underline{K}_{MCT}(k,t-\tau)\frac{dF(k,t)}{dt}\,\mathrm{d}\tau, \\
\underline{K}_{MCT}(k,t) &= \frac{\rho k_BT}{16\pi^3 m}\int |V_{\mathbf{q},\mathbf{k}-\mathbf{q}}|^2F(q,t)F(|\mathbf{k}-\mathbf{q}|,t)\,\mathrm{d}\mathbf{q},
\end{aligned} \tag{2.8}$$

where $V_{\mathbf{q},\mathbf{k}-\mathbf{q}}$ is a so-called vertex and depends on wave vectors $\mathbf{k}$ and $\mathbf{q}$, the bulk density ρ and the static structure factor evaluated at these wave vectors. Further k_B is the Boltzmann constant and m is the particle mass. We write these equations here to emphasize two things: the MCT approximations lead to an equation

where the intermediate scattering function at a specific wavevector depends on those at all other wave vectors via the integral in $\underline{K}_{MCT}$, and that apart from the system parameters ρ, m and T the only structural input for this theory is the liquid's structure factor $S(k)$ at all wavelengths.

Finally, instead of the function $F(k,t)$, we can also consider $\Phi(k,t) := F(k,t)/S(k)$, i.e. we consider a normalized intermediate scattering function that satisfies $\Phi(k,0) = 1$ by equation (2.4). Since $S(k)$ is constant with respect to the time t in equation (2.8), the equation for $\Phi(k,t)$ remains very similar. Apart from $\Phi(k,0) = 1$, the initial condition $\dot{\Phi}(k,0) = 0$ is imposed on Φ [15].

2.2 Interpretation

To interpret equation (2.8), we compare it to the equation of motion for a damped oscillator $x(t)$ [13],

$$\ddot{x}(t) + \omega^2 x(t) + 2\zeta\omega\dot{x}(t) = 0, \tag{2.9}$$

where dots denote time differentiation, ω is the frequency of the undamped oscillator and ζ is the damping coefficient. Comparison with (2.8) reveals that the role of ω^2 is fulfilled by the factor $k^2 k_B T/mS(k)$, which is the Ω_{21} component of the matrix Ω in equation (2.6). Hence the name frequency matrix. Further we see that the damping effect of ζ is similar to the time-dependent memory function $\underline{K}_{MCT}$ under the integral in equation (2.8). Indeed, the 'fast' fluctuating force $\underline{f}$ from which it ultimately stems can thus be interpreted as a dissipative force responsible for the evolution of our variables under the influence of fast variables.

2.3 Generalized MCT

To avoid the uncontrolled factorization of four-point density correlation functions, the Mori-Zwanzig formalism has been invoked again to find exact equations of motion for those, starting from a basis of two-point correlators $\rho(\mathbf{k_1}, t)\rho(\mathbf{k_2}, t)$, for wave vectors $\mathbf{k_1}, \mathbf{k_2}$, to project onto. Omitting the details, we state that the exact equations for these are similar in structure to equation (2.6), but now with a memory function which is in first order determined by six-point density correlations [15].

To be precise, a first derivative also appears outside the integral in the work of [15]. This is due to an approximation of the memory kernel in equation (2.8), where the fast dynamics are modelled by a delta function with a coefficient representing a friction coefficient (and the slow dynamics are added to this). This delta function takes one first-order derivative outside the integral and accounts for the friction coefficient in the equations in chapter 3. Often the so-called over-damped limit is used, where the friction coefficient outweighs the second derivative, and the second derivative is then dropped from the equations [15]. In any case, this does not influence the long-time behaviour as the second derivative always becomes negligible eventually.

For the four-point correlators, one can continue and find exact equations of motion depending on six-point correlators, etc. For each step, let us schematically denote by F_n a $2n$-point density correlator. (So F_1 is the function we have been interested in since equation (2.3) and which in the usual MCT approximations obeys equation (2.8).) This process gives a hierarchy of equations of motion for the correlators F_n, each of which then depends on the $2(n+1)$-point correlator F_{n+1} and is called generalised MCT (GMCT). This can in theory be continued indefinitely, obtaining infinite order GMCT. The same can be done for the normalised correlators $\Phi_n(\mathbf{k_1}, \ldots, \mathbf{k_n}, t) := F_n(\mathbf{k_1}, \ldots, \mathbf{k_n}, t)/F_n(\mathbf{k_1}, \ldots, \mathbf{k_n}, 0)$. For numerical tractability, however, this is impossible. Therefore an approximation is in practice eventually made for some correlator F_{N+1}, where we refer to N as the closure level and to the approximation made as the closure.

The so-called mean-field closure at level N uses the factorization approximation as before, but now for the equation of motion for the correlator F_N. Doing this for $N = 1$ corresponds to equation (2.8) and is referred to as standard MCT or first-order GMCT. One can choose between various factorizations in terms of all the $F_1, \ldots, F_N$ as long as the total number of point correlations sums up to $2(N+1)$. With the mean-field closure at level N, we obtain a finite-dimensional system for the correlators $F_1, \ldots, F_N$. Another possibility is the exponential closure, which sets $F_{N+1} = 0$. This is physically based on the observation that at higher levels, the correlators start to look like exponentially decaying functions. Since with the exponential closure, the kernel vanishes in the equation for F_N, this observation is made exact with the exponential closure as the solution for F_N then becomes a sum of exponentials (as we show in chapter 3).

2.4 Schematic GMCT

An analytical solution to the MCT equation (2.8) is not known, let alone to the infinite or closed system of GMCT equations. Simplifications are needed to obtain a better qualitative understanding of MCT. This can be done based on an observation by Bengtzelius *et al.*, namely that the main contribution to the memory function arises from the first peak of $S(k)$, located at a wavevector which we denote k_0 [1]. They make the simplification

$$S(k) = 1 + A\delta(k - k_0).$$

A is often defined to be the area under the first peak of $S(k)$. This modelling for $S(k)$ has the result that each equation of the GMCT hierarchy depends on k_0 only, which makes it essentially wavevector independent.

This approach is called schematic MCT in the case of first order GMCT, or for higher closure levels, schematic GMCT. Since all k-dependent factors in the GMCT equations also become constants, we label them μ_n for the frequency-like factor and λ_n for the factor in front of the integral, where n is the level of the GMCT hierarchy (note that we do allow for different constants at each level). The former is also sometimes referred to as ${\omega_n}^2$. We denote the normalised, wavevector independent intermediate scattering function for level n by ϕ_n. For the numerical investigation in later chapters, ϕ_1 as predicted by standard schematic MCT will be our variable of interest.

For decreasing temperature, it is seen that the structure factor becomes more peaked and increases and hence we should think of the parameter A as an increasing function of inverse temperature[1] [17]. At the level of the GMCT equations, this implies that the damping is strengthened and also the parameters λ_n in GMCT are to be thought of as an increasing function of inverse temperature.

We will precisely state the equations for schematic GMCT in the next chapter, together with a precise description of the closures one can use to close the hierarchy.

2.5 Predictions

We focus here on predictions of MCT (and GMCT or the schematic approximation) regarding fragility and the behaviour of the relaxation time up to the transition, since those are properties that this project has focussed on as well. The relaxation time can be defined in various ways, e.g. the time point where the correlator Φ_1 or ϕ_1 has decayed to a value below $1/e$ or 0.05, or the area under its graph as is done in [14] for analytical means. We stress once more that the predictions of MCT are based only on the static information contained in $S(k)$ (for e.g. various temperatures) and the bulk density and temperature one is interested in.

The first remarkable and most notable aspect is that MCT can indeed predict a transition in the sense that the time the correlation functions need to decay to zero, increases without bound at a finite temperature [17]. The transition might be indicative of a transition from a liquid to a solid state. However, since no information about an ordered state is used or imposed in the theory, Reichman and Charbonneau conclude in [25] that this can only be a disordered state and hence the prediction of a glass. Moreover, both for MCT and its schematic form, the time dependence of the correlators changes form from a simple decay in the high temperature regime to a relaxation exhibiting the characteristic β and α-relaxation in the super-cooled liquid regime, which we used to depict these phenomena in figure 1. So we see that MCT as well as its further approximation capture a physical mechanism that can be used to explain the glass transition [17].

Yet the temperature at which MCT predicts the transition is often not in good agreement with experiments or simulations [13]; the transition temperature of the theory is much higher than the true experimental temperature of the glass transition. We come back to this point below, in view of GMCT.

Another prediction is that the divergence of the relaxation time as a function of T always follows a power law close to the transition temperature. This implies that MCT can only account for fragile glass formation, and no Arrhenius T-dependence which is in fact one of the things people seek to describe by a macroscopic theory for glass forming systems. Still, for fragile glass formers, even when rescaling the transition temperature to the real glass forming temperature of the system, the dynamics are not accurately described by the theory close the the transition [13]. For further predictions on functional forms (which often correspond to phenomenological observations of glass formers) and other quantities, we refer to [17], [13] for an overview.

GMCT

Including exact equations for higher-order correlators, GMCT can make a number of predictions that improve on the above. Janssen *et al.* have identified parameter choices in the over-damped schematic GMCT equations (3.1) for λ_n and μ_n for the infinite order hierarchy which can be fitted to the standard $F2$ model (to be introduced in chapter 7), demonstrating that GMCT is more general than MCT and can indeed reproduce its predictions [14]. This fit was made only to a certain plateau height and relaxation time, yet reproduces all qualitative behaviour with respect to time as well as temperature. The authors further found a model that can, based on a single parameter, predict no divergence for any finite temperature, as well as divergence with different fragilities: from fragile to Arrhenius-type and even to super-strong glass formers. They expect these properties to be captured by the non-schematic GMCT as well. So GMCT really extends the prediction possibilities of MCT.

GMCT can also be compared to simulations. One then finds convergence towards the correlation functions obtained by simulations with an increasing closure level [16], whether one uses the exponential or mean-field closures at different levels. This uniform convergence can be rigorously demonstrated for a schematic model [22]. Further, it is found in the numerical studies of [16] that the mean-field closures give an upper bound to the dynamics (i.e., the correlation functions), while the exponential closures provides a lower bound.

[1] We can also think of it as a function of the bulk density or packing fraction, which are also parameters that can spur a glass transition. However, our interpretation will only be in terms of temperature

The convergence can be interpreted physically. Mean-field theories generally underestimate the effects of ergodicity-restoring fluctuations and hence overestimate the dynamic slowing down towards a non-ergodic state (the glass transition). In contrast, the exponential closure ignores higher order memory effects by setting $\Phi_{N+1}(k,t) \equiv 0$. This consequently leads to relaxation patterns that are too fast. So for both closures, a convergence notion exists in that we incorporate more memory effects in the solution. A higher closure level can also turn a non-relaxing correlation function into a relaxing one, removing the discrepancy between the MCT-predicted transition temperature and the simulated one [16].

We finally note that it is suggested that the MCT transition should be interpreted as a dynamical transition, where hopping effects or activated dynamics of particles 'hopping' out of their cages are responsible for ergodicity-restoring effects below the predicted transition temperature. An underlying argument is that the MCT-predicted transition does not adhere to experimentally observed divergence behaviour of the relaxation time [17]. In view of the above, GMCT improves the prediction of the transition of standard MCT without any need for fitting the results to simulations, from which it can be concluded that the inclusion of higher-order terms appears to describe all relevant relaxation patterns, at least in the weakly supercooled regime [16].

3 Mathematical statement of the equations

The equations that are studied in generalized mode coupling theory concern an infinite hierarchy of coupled integro-differential equations for the unknown functions ϕ_n, $n = 1, 2, \ldots$. In order for the equations below to make sense, we will assume that each of these functions is twice continuously differentiable, i.e. the functions of interest can be characterized as $\{\phi_n \in C^{(2)}(\mathbb{R}, \mathbb{R}) \mid n \in \mathbb{N}\}$. Physically these result from (real valued) density correlation functions in the Fourier domain and depend on a single variable, time.

3.1 Equations

In [15], a hierarchy of equations is derived in a schematic approximation, where wave-vector dependence of the density correlation functions is neglected (to produce ϕ_n) and brought into two characterizing (unknown) sets of constants $\mu_n \in \mathbb{R}^+$, $\lambda_n \in \mathbb{R}^+$, $n \in \mathbb{N}$. We briefly pointed towards this procedure in the introduction and more details are presented in the physical theory of chapter 2. In the over-damped limit, where a second derivative is ignored in the presence of a first derivative, this yields the following equations for $n \in \mathbb{N}$, $t > 0$:

$$\begin{cases} \dot{\phi}_n(t) + \mu_n \phi_n(t) + \lambda_n \int_0^t \phi_{n+1}(\tau)\dot{\phi}_n(t-\tau)d\tau = 0, \\ \phi_n(0) = 1. \end{cases} \tag{3.1}$$

where $\dot{\phi}_n(t) = \frac{d\phi_n}{dt}(t)$ denotes differentiation of ϕ_n.

Alternatively, one could decide not to make the assumption of the over-damped limit. In this case a second order derivative is not neglected and the equations contain an effective friction coefficient $\zeta > 0$. Moreover, an initial value is needed for the first derivative, which must equal 1 for all functions [15]. The equations that then result, for $n \in \mathbb{N}$ and $t > 0$, are given by:

$$\begin{cases} \ddot{\phi}_n(t) + \zeta\dot{\phi}_n(t) + \mu_n \phi_n(t) + \lambda_n \int_0^t \phi_{n+1}(\tau)\dot{\phi}_n(t-\tau)d\tau = 0, \\ \phi_n(0) = 1, \dot{\phi}_n(0) = 0. \end{cases} \tag{3.2}$$

The most eminent problem when discussing solutions to these equations is clear: the defining equation for each ϕ_n depends on another unknown function ϕ_{n+1}, which can by no means be found unless one knows the next unknown ϕ_{n+2} and this continuous for ever, not yielding a finite set of integro-differential equations to be solved. To this end, an assumption must be made at some finite value of n to end the infinite hierarchy. In [15], two ways are proposed which we will both investigate in the following section. These assumptions to end the hierarchy are called closures.

3.2 Closures

Let us denote the level at which we wish to close the hierarchy by N, by which we mean that the equations (3.1) and (3.2) will be considered for $n = 1, 2, \ldots, N$ and ϕ_{N+1} is expressed in terms of the functions $\phi_1, \phi_2, \ldots, \phi_N$ to yield a system of N equations for N unknown functions. We will denote such a closure function at level N by $\mathcal{C}_N$.

Definition 3.2.1. For $N \geq 1$ the so-called **mean-field closure** at level N is defined to be of the generic form [2]

$$\phi_{N+1} = \mathcal{C}_N(\phi_1, \ldots, \phi_N) = \prod_{n=1}^{N} (\phi_n)^{m_n} = \prod_{n=1}^{N} \phi_n^{m_n}, \tag{3.3}$$

where for $n = 1, 2, \ldots, N$ we have $m_n \in \mathbb{N} \cup \{0\}$ chosen such that

$$\sum_{n=1}^{N} n \cdot m_n = N + 1 \tag{3.4}$$

is satisfied. This closure is briefly written as MF-$N(1^{m_1}2^{m_2}\ldots n^{m_n}\ldots N^{m_N})$, omitting all n for which $m_n = 0$.

[2] We slightly change notation here with respect to Janssen *et al.* in [15]. They adopt the definition that a closure at level N means that ϕ_N is expressed in terms of $\phi_1, \phi_2, \ldots, \phi_{N-1}$, but we think that taking N unknowns at level N enhances notational clarity for the explicit mathematical analysis that follows. The authors of [15] further write 3.3 as a product of $\phi_{n_i}^{m_i}$'s, with a new subindex, but this notation seems to be superfluous.

Example 3.2.2. With the above definition, we see that MF-4(14) means that we define $\phi_5 = \phi_1\phi_4$. More generally, MF-$N(1N)$ means that we define $\phi_{N+1} = \phi_1\phi_N$, and MF-$N(1^{N+1})$ defines $\phi_{N+1} = \phi_1^{N+1}$.

Remark 3.2.3. (Standard schematic MCT.) In both of the general examples above, closure at level 1 reduces to MF-1(1^2) such that $\phi_2 = \phi_1^2$. This is exactly the approximation made in standard schematic MCT, which is the schematic (wave vector independent) approximation of the theory discussed in 2.1.

Another approach suggested in [15] is to put $\phi_{N+1}(t) = 0$ for all $t \geq 0$. This leads to the following definition of the exponential closure.

Definition 3.2.4. Let $N \in \mathbb{N}$. The **exponential closure** of the system (3.1) or (3.2) at level N is defined by

$$\phi_{N+1} \equiv 0, \tag{3.5}$$

which turns ϕ_{N+1} into the zero function. This is denoted as the exp-N closure.

Remark 3.2.5. Looking at the system 3.1, which is studied in [15] for $\mu_n = n$, it follows readily that the exp-N closure yields $\phi_N(t) = e^{-\mu_n t}$. This explains the terminology used for the exponential closure.
The numerical studies in [15] suggest that for sufficiently large n, $\phi_n(t) \approx e^{-nt}$. This observation suggests the exponential closure, making the observation exact at level N.

Observation 3.2.6. For completeness, let us also analyse the effect of the exp-N closure on the system (3.2). For ϕ_N, ($N \geq 1$), the exp-N closure yields an initial value problem with an ordinary second-order differential equation for $t \geq 0$:

$$\begin{cases} \ddot{\phi}_N(t) + \zeta\dot{\phi}_N(t) + \mu_N\phi_N(t) = 0, \\ \phi_N(0) = \dot{\phi}_N(0) = 1, \end{cases}$$

having the constant coefficients ζ and μ_N in front of the unknown ϕ_N and its derivatives. Solving this initial value problem is straightforward [2, Chapter 3]: upon the substitution $\phi_N(t) = e^{\lambda t}$, we obtain the characteristic equation

$$\lambda^2 + \zeta\lambda + \mu_N = 0.$$

So when $\zeta^2 > 4\mu$ we obtain the real solutions $\lambda = \lambda_1 = -\frac{1}{2}\left(\zeta - \sqrt{\zeta^2 - 4\mu}\right)$ and $\lambda = \lambda_2 = -\frac{1}{2}\left(\zeta + \sqrt{\zeta^2 - 4\mu}\right)$ and the initial value problem has a unique solution, which is a linear combination of the form $C_1e^{\lambda_1 t} + C_2e^{\lambda_2 t}$, and hence again an exponential form. When $\zeta^2 < 4\mu$, we have a similar solution but with complex exponentials, or a superposition of sines and cosines if we consider real-valued solutions only. For the situation $\zeta^2 = 4\mu$, we have a solution of the form $(C_1 + tC_2)e^{-\frac{1}{2}\zeta t}$ and since $\zeta > 0$, the solution still displays decay dominated by an exponential in the long-time limit $t \to \infty$.

Remark 3.2.7. From a mathematical point of view, the exp-N closure drastically simplifies the coupled system, for one can solve the unknown functions $\phi_N, \ldots, \phi_1$ one by one, each time substituting the obtained solution in the equation for the function with the next lower index. However, the assumption of the exponential closure is in direct contradiction with the *infinite* hierarchy that is actually considered in the systems (3.1) and (3.2), since these demand that $\phi_{N+1}(t = 0)$ be equal to 1. This is avoided with the MF-N closures, which yield a self-consistent system for $\phi_1, \phi_2, \ldots, \phi_N$ that does not directly contradict the discarded equations of the infinite system.

These both closures finalize the definition of the equations of interest in generalized mode-coupling theory.

4 Preliminary mathematical results

In chapter 5, some existence and uniqueness results will be proved, extending the arguments provided by Götze *et al.* in [10] and Saal in [29]. To support the arguments they use, various results from mathematical theory will be invoked. In this chapter, we aim to organize these results and the underlying definitions to prepare for their application later. We give a reference for these results, or a proof for less general statements that we derive ourself for the purposes of this report. Alhough many results are interesting in their own right, the reader may wish to read the relevant sections of this chapter only when they are needed in chapter 5.

4.1 (In)equalities

We will need various inequalities to establish appropriate upper bounds in chapter 5. It turns out that for the convolution and double integral type expressions that we encounter, such estimates can be found. Also the Cauchy-Schwarz inequality for bounds on inner products and Gronwall's inequality for integral equations will be needed, and some more. However, we start with a very easy-to-prove inequality which will surprisingly often be used.

Inequality 4.1.1. Let $a, b \in \mathbb{R}$. Then the following holds:

$$a \cdot b \leq \frac{1}{2}\left(a^2 + b^2\right). \tag{4.1}$$

Proof:
From $0 \leq (a-b)^2 = a^2 - 2ab + b^2$, we immediately derive the desired result. □

Lemma 4.1.2. Let $T_0, T \in \mathbb{R}$, $T_0 < T$. Let $f : [T_0, T] \to \mathbb{R}$ $(T > T_0)$ be integrable and such that for all $x \in [T_0, T]$, $f(x)$ is non-negative. Then for all $t \in [T_0, T]$ the following inequality holds:

$$\int_{T_0}^{t}\int_{T_0}^{s} f(r)\,\mathrm{d}r\,\mathrm{d}s \leq (T - T_0)\int_{T_0}^{t} f(s)\,\mathrm{d}s. \tag{4.2}$$

Proof:
Reversing the order of integration, we can write

$$\begin{aligned}\int_{T_0}^{t}\int_{T_0}^{s} f(r)\,\mathrm{d}r\,\mathrm{d}s = \int_{T_0}^{t}\int_{r}^{t} f(r)\,\mathrm{d}s\,\mathrm{d}r &= \int_{T_0}^{t}(t-r)f(r)\,\mathrm{d}r,\\ &\leq \int_{T_0}^{t}(t-T_0)f(r)\,\mathrm{d}r \leq (T-T_0)\int_{T_0}^{t} f(r)\,\mathrm{d}r,\end{aligned}$$

where the inequality holds because $f(r) \geq 0$ for $r \in [T_0, T]$ by assumption. Note that r is a dummy variable in the last integral. □

Lemma 4.1.3. Let $f : [T_0, T] \to \mathbb{R}$ be non-negative and integrable on $[T_0, T]$. Then for all $t \in [T_0, T]$:

$$\int_{T_0}^{t}\int_{T_0}^{s} f(r)\cdot f(s)\,\mathrm{d}r\,\mathrm{d}s \leq (T - T_0)\int_{T_0}^{t} f(s)\,\mathrm{d}s. \tag{4.3}$$

Proof:
Splitting the sum in the integral after applying (4.1), next applying (4.2) to the integral with $f(r)$, we obtain

$$\begin{aligned}\int_{T_0}^{t}\int_{T_0}^{s} f(r)\cdot f(s)\,\mathrm{d}r\,\mathrm{d}s &\leq \frac{1}{2}\int_{T_0}^{t}\int_{T_0}^{s} f(r) + f(s)\,\mathrm{d}r\,\mathrm{d}s,\\ &\leq \frac{1}{2}\int_{T_0}^{t}\int_{T_0}^{s} f(r)\,\mathrm{d}r\,\mathrm{d}s + \frac{1}{2}\int_{T_0}^{t} f(s)\int_{T_0}^{s} 1\,\mathrm{d}r\,\mathrm{d}s,\\ &\leq \frac{1}{2}(T-T_0)\int_{T_0}^{t} f(s)\,\mathrm{d}s + \frac{1}{2}\int_{T_0}^{t}(s-T_0)f(s)\,\mathrm{d}s \leq (T-T_0)\int_{T_0}^{t} f(s)\,\mathrm{d}s.\end{aligned}$$

□

Lemma 4.1.4. Let $f : [0, 2T_0] \to \mathbb{R}$ $(T_0 > 0)$ be integrable and non-negative. For $t \in [T_0, 2T_0]$ it holds that

$$\int_{T_0}^{t} \int_{0}^{T_0} f(s-r)\,\mathrm{d}r\,\mathrm{d}s \leq T_0 \int_{0}^{t} f(s)\,\mathrm{d}s. \tag{4.4}$$

Proof:
Changing variables to $u = s - r$, we obtain

$$\int_{T_0}^{t} \int_{0}^{T_0} f(s-r)\,\mathrm{d}r\,\mathrm{d}s = \int_{T_0}^{t} \int_{s-T_0}^{s} f(u)\,\mathrm{d}u\,\mathrm{d}s.$$

In terms of u and s, the region of integration is represented by figure 3. It is a parallelogram with the two upper corners at $(u, s) = (t - T_0, t)$ and $(u, s) = (T_0, t)$, or the domain $D = \{(u, s) \in \mathbb{R}^2 \mid T_0 \leq s \leq t, s - T_0 \leq u \leq s\}$. Since we have $T_0 \leq t \leq 2T_0$, their relative positioning is in accordance with the configuration in figure 3 (except for the case when $t = 2T_0$, but that poses no complications below). Careful inspection of the figure allows us to describe the same domain D with the following boundaries, reversing the order of integration:

$$\begin{aligned}
\int_{T_0}^{t} \int_{0}^{T_0} f(s-r)\,\mathrm{d}r\,\mathrm{d}s &= \int_{0}^{t-T_0} \int_{T_0}^{u+T_0} f(u)\,\mathrm{d}s\,\mathrm{d}u + \int_{t-T_0}^{T_0} \int_{T_0}^{t} f(u)\,\mathrm{d}s\,\mathrm{d}u + \int_{T_0}^{t} \int_{u}^{t} f(u)\,\mathrm{d}s\,\mathrm{d}u, \\
&= \int_{0}^{t-T_0} u f(u)\,\mathrm{d}u + (t - T_0) \int_{t-T_0}^{T_0} f(u)\,\mathrm{d}u + \int_{T_0}^{t} (t-u) f(u)\,\mathrm{d}u, \\
&\leq (t - T_0) \int_{0}^{t} f(u)\,\mathrm{d}u \leq T_0 \int_{0}^{t} f(u)\,\mathrm{d}u.
\end{aligned}$$

We used that for $t \in [T_0, 2T_0]$ and $u \in [T_0, t]$, we have $0 \leq t - u \leq t - T_0$ and $t - T_0 \leq T_0$. □

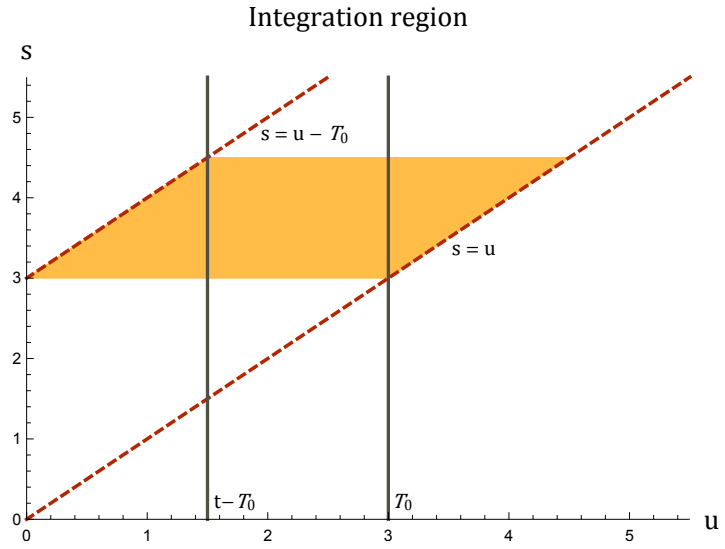

Figure 3: *Region of integration (orange) in the u, s-plane in the proof of Lemma 4.1.4, illustrated for $T_0 = 3$ and $t = 4.5$.*

Lemma 4.1.5. (Gronwall's inequality, [7, Section 1.5]).
Let $a < b$ and let $I = [a, b]$ denote an interval of the real line. Let $\alpha, \beta, u : I \to \mathbb{R}$ and assume these functions are continuous on I and that $\beta(x) \geq 0$ on I. Let u satisfy for all $t \in I$ the integral inequality

$$u(t) \leq \alpha(t) + \int_{a}^{t} \beta(s) u(s)\,\mathrm{d}s. \tag{4.5}$$

Then for all $t \in I$:

$$u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s) \exp\left(\int_s^t \beta(r)\,\mathrm{d}r\right)\mathrm{d}s. \tag{4.6}$$

If also α is non-decreasing[3], it holds for all $t \in I$ that

$$u(t) \leq \alpha(t) \exp\left(\int_a^t \beta(s)\,\mathrm{d}s\right). \tag{4.7}$$

Notation 4.1.6. Let U be an inner product space. We denote the inner product on U by $\cdot$, i.e. the infix operator $\cdot$ with the notation

$$\cdot : U \times U \to \mathbb{R}, (u, v) \mapsto u \cdot v.$$

For some vector object $u \in U$, we shall write $u \cdot u = u^2$. We will not make an explicit distinction in notation between inner products on different spaces.

Lemma 4.1.7. (Cauchy-Schwarz inequality, [19, Lemma 3.2-1]) Let U be an inner product space and let $u, v \in U$. Let $||.||$ denote the norm that is derived from the inner product on U, i.e. $||u|| = \sqrt{u^2}$. We then have

$$|u \cdot v| \leq ||u|| \cdot ||v||. \tag{4.8}$$

Observation 4.1.8. We use the same notation as above. In view of the Cauchy-Schwarz inequality and (4.1), we can also derive the following estimate of an inner product in terms of a sum of inner products:

$$|u \cdot v| \leq \frac{1}{2}(u^2 + v^2).$$

That is, (4.1) generalizes to inner products.

Lemma 4.1.9. Let $n \in \mathbb{N}$, $f \in C^{(1)}([a, b], \mathbb{R}^n)$ be a vector-valued function, continuously differentiable function for some interval $[a, b]$ in $\mathbb{R}$ and let $\cdot$ denote the standard inner product on$\mathbb{R}^n$. Then for all $t \in [a, b]$ we have

$$\frac{d}{dt}\frac{1}{2}(f \cdot f)(t) = f(t) \cdot \frac{d}{dt} f(t), \tag{4.9}$$

where $f \cdot f \in C^{(1)}([0, T], \mathbb{R})$ is defined by $f \cdot f(t) = f(t) \cdot f(t)$. This is an intuitive generalization of the chain rule in the case of an inner product.

Proof:

The proof can be found by straight-forward use of the definition of the standard inner product and application of the chain rule in the one-dimensional case for each component, because differentiation of a vector-valued function comes down to differentiating each component. Let us write for $t \in [a, b]$: $f(t) = (f_1(t), f_2(t), \ldots, f_n(t))$. We then have for all $t \in [a, b]$:

$$\begin{aligned} f(t) \cdot \frac{d}{dt} f(t) &= \sum_{i=1}^{n} f_i(t) \cdot \frac{d}{dt} f_i(t) = \sum_{i=1}^{n} \frac{d}{dt}\frac{1}{2} {f_i}^2(t), \\ &= \frac{d}{dt}\frac{1}{2}\sum_{i=1}^{n} {f_i}^2(t) = \frac{d}{dt}\frac{1}{2}(f \cdot f)(t). \end{aligned}$$

□

Theorem 4.1.10. (Leibniz rule for differentiation of integrals, [32, section 1.3]) Let $f = f(x, t)$ and $\frac{\partial f}{\partial t}$ be continuous functions in a domain of the x, t-plane that includes the rectangle $R = [a, b] \times [t_0, t_1]$. Further let $g, h : [a, b] \to [t_0, t_1]$ be differentiable on $[a, b]$. We then have for all $x \in [a, b]$:

$$\frac{d}{dx}\int_{g(x)}^{h(x)} f(x, t)\,\mathrm{d}t = f(x, h(x))\frac{dh(x)}{dx} - f(x, g(x))\frac{dg(x)}{dx} + \int_{g(x)}^{h(x)} \frac{\partial f(x, t)}{\partial x}\,\mathrm{d}t. \tag{4.10}$$

We finish these inequalities with the definition of Lipschitz continuity and show how this property can be derived from continuous differentiability.

[3] In [7] this assumption reads α is constant, but the proof can readily be extended to non-decreasing α since then for some $t \in I$, $\alpha(s)$ under the integral in (4.6) can be estimated by $\alpha(t)$ for $s \in [a, t]$.

Definition 4.1.11. Let U, V be metric spaces with metrics $d_U : U \times U \to \mathbb{R}$ and $d_V : V \times V \to \mathbb{R}$, respectively. Let $f : U \to V$ be a function from U to V. f is said to satisfy a **Lipschitz condition** with constant L, or to be **Lipschitz continuous** if a constant $L > 0$ exists such that for all $x_1, x_2 \in X$ we have

$$d_Y\left(f\left(x_1\right), f\left(x_2\right)\right) \leq L d_X\left(x_1, x_2\right).$$

Definition 4.1.12. Let $n \in \mathbb{N}$. We define the metric[4] $\delta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ for $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ by

$$\delta(x, y) = \sum_{i=1}^{n} |x_i - y_i|.$$

It will be useful in section 5.2 to work with this metric instead of the standard Euclidean metric on $\mathbb{R}^n$.

Lemma 4.1.13. (Lipschitz continuity for continuously differentiable functions) Let $[a, b]$ be a closed interval in $\mathbb{R}$, $n \in \mathbb{N}$ and let $f : [a, b]^n \to \mathbb{R}$ be continuous and continuously differentiable on $[a, b]^n$. Then f is Lipschitz continuous.
Proof:
Since f is continuous, it is also continuous in each of its n variables. Let us write for $x \in \mathbb{R}^n$, $x = (x_1, \ldots, x_n)$. Applying the Mean Value Theorem to the i^{th} variable yields for all $x, y \in [a, b]^n$ (with $x_i < y_i$) that some $\xi = \xi(x_i, y_i) \in (x_i, y_i)$ exists with

$$\frac{|f(x_1, \ldots, x_i, \ldots, x_d) - f(x_1, \ldots, x_{i-1}, y_i, x_{i+1}, \ldots, x_d)|}{|x_i - y_i|} \leq |\partial_i f(\xi)| \leq \max_{i=1,\ldots,n} \max_{t \in [a,b]} |\partial_i f(t)| := L,$$

where L is independent of x and y, and the component i under consideration. Therefore it also follows for all $x, y \in [a, b]^n$ that

$$\begin{aligned}
|f(x_1, \ldots, x_n) - f(y_1, \ldots, y_n)| \leq & |f(x_1, \ldots, x_n) - f(y_1, x_2, \ldots, x_n)| + |f(y_1, x_2, \ldots, x_n) - f(y_1, \ldots, y_n)|, \\
\leq & |f(x_1, \ldots, x_n) - f(y_1, x_2, \ldots, x_n)| + |f(y_1, x_2, \ldots, x_n) - f(y_1, y_2, x_3, \ldots, x_n)| \\
& + |f(y_1, y_2, x_3, \ldots, x_n) + f(y_1, \ldots, y_n)|, \\
\leq & |f(x_1, \ldots, x_n) - f(y_1, x_2, \ldots, x_n)| + |f(y_1, x_2, \ldots, x_n) - f(y_1, y_2, x_3, \ldots, x_n)| \\
& + \cdots + |f(y_1, \ldots, y_{n-1}, x_n) - f(y_1, \ldots, y_n)|, \\
\leq & L \sum_{i=1}^{n} |x_i - y_i| = L\delta(x, y).
\end{aligned}$$

□

4.2 Sequences of functions

Here we will state some results concerning uniform convergence of real-valued functions and consequent properties. These will be referred to in order to complete the technical arguments in chapter 5.

Theorem 4.2.1. (Cauchy criterion for sequences, [18, Theorem 8.2.8]) A sequence of functions $(f_n)_{n \in \mathbb{N}}$ converges uniformly on D if and only if for each $\epsilon > 0$ there exists an $n^* \in \mathbb{N}$ such that for all $x \in D$: $|f_n(x) - f_m(x)| < \epsilon$ whenever $n, m \geq n^*$.

Theorem 4.2.2. (Interchanging limit and integral) If $(f_n)_{n \in \mathbb{R}}$ is a sequence of continuous functions that converges uniformly to a function f on $[a, b]$, then

$$\lim_{n \to \infty} \int_a^b f_n(x)\, \mathrm{d}x = \int_a^b \left[\lim_{n \to \infty} f_n(x)\right] \mathrm{d}x = \int_a^b f(x)\, \mathrm{d}x.$$

I.e., if f_n converges uniformly on $[a, b]$, we can interchange taking the limit and taking the integral over the interval $[a, b]$.

[4] We will not explicitly check that this definition satisfies that of a metric. This is assumed to be familiar by experience.

Theorem 4.2.3. ([27, Theorem 7.17]) Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions, differentiable on $[a,b]$ $(a,b \in \mathbb{R}, a > b)$ and such that $(f_n(x_0))_{n\in\mathbb{N}}$ converges for some $x_0 \in [a,b]$. If $(f_n')_{n\in\mathbb{N}}$ converges uniformly on $[a,b]$, then $(f_n)_{n\in\mathbb{N}}$ converges uniformly on $[a,b]$, to a function f, and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \quad (a \le x \le b).$$

That is, under the conditions stated here, we can interchange differentiation and taking the limit of the sequence.

Lemma 4.2.4. Let $(f_n)_{n\in\mathbb{N}}$ and $(g_n)_{n\in\mathbb{N}}$ be sequences of continuous functions from $[0,T]$ to $\mathbb{R}^k$ for some $k \in \mathbb{N}, T > 0$. Suppose that $f_n \to f$ and $g_n \to g$, both uniformly, as $n \to \infty$ (in some norm $||.||_1$ on $\mathbb{R}^k$). Let $M : \mathbb{R}^k \to \mathbb{R}^m$ for some $m \in \mathbb{N}$ be uniformly continuous on a set D with respect to the norm $||.||_1$ on $\mathbb{R}^k$ and some norm $||.||_2$ on $\mathbb{R}^m$. Here $D = \{x \in \mathbb{R}^k \mid ||x||_1 < \max_{t\in[0,T]} ||f(t)||_1 + c\}$ for some $c > 0$. Then the following statements hold for $n \to \infty$, and each convergence is uniform on [0,T]:

(i) $f_n + g_n \to f + g$;

(ii) $M \circ f_n \to M \circ f$.

Proof:
Let $\epsilon > 0$ be given.

(i) For $t \in [0,T]$, we have

$$||f_n(t) + g_n(t) - f(t) - g(t)||_1 \le ||f_n(t) - f(t)||_1 + ||g_n(t) - g(t)||_1.$$

By uniform convergence, there exist $n_1 \in \mathbb{N}$ such that $n \ge n_1$ implies $||f_n(t) - f(t)||_1 < \epsilon/2$ and $n_2 \in \mathbb{N}$ such that $n \ge n_2$ implies $||g_n(t) - g(t)||_1 < \epsilon/2$, independent of t. Hence for $n \ge \max\{n_1, n_2\}$, we conclude that

$$||f_n(t) + g_n(t) - f(t) - g(t)||_1 < \epsilon.$$

(ii) By uniform continuity of M on D, for all $t \in [0,T]$ there exists a $\delta_\epsilon > 0$ (independent of t) such that for all $x \in D$, $||x - f(t)||_1 < \delta_\epsilon$ implies $||M(x) - M(f(t))||_2 < \epsilon$. By uniform convergence of (f_n), there exists an $N \in \mathbb{N}$ such that $n \ge N$ implies $||f_n(t) - f(t)||_1 < \min\{c, \delta_\epsilon\}$, for all $t \in [0,T]$. Let $n \ge N$. Then we see for all $t \in [0,T]$ that $||f_n(t)||_1 \le ||f(t)||_1 + ||f_n(t) - f(t)||_1 < \max_{s\in[0,T]} ||f(s)||_1 + c$, hence $f_n(t) \in D$, and

$$||f_n(t) - f(t)||_1 < \delta_\epsilon \Rightarrow ||M(f_n(t)) - M(f(t))||_2 < \epsilon.$$

□

Lemma 4.2.5. Let $(f_n)_{n\in\mathbb{N}}$ be a sequence of functions from $[0,T]$ to $\mathbb{R}^{k\times k}$ and let $(g_n)_{n\in\mathbb{N}}$ be a sequence of functions from $[0,T]$ to $\mathbb{R}^k$ for some $T > 0, k \in \mathbb{N}$. We work with some norm $||.||_1$ on $\mathbb{R}^k$ and the induced operator norm $||.||_2$ on $\mathbb{R}^{k\times k}$ (see remark 4.4.3). Suppose that (f_n) and (g_n) converge uniformly to functions f and g, respectively. Define $\Omega := \{(\tau,t) \in [0,T]^2 \mid t \le \tau\}$ and define $\bar{f}g : \Omega \to \mathbb{R}^k$ by $\bar{f}g(\tau,t) = f(\tau - t)g(t)$ and likewise for all f_n and g_n. Then $\bar{f}_n g_n \to \bar{f}g$ uniformly as $n \to \infty$.
Proof:
Let $\epsilon > 0$ be given. By $||.||_\infty$ we denote the appropriate supremum norms over the interval $[0,T]$. For all $n \in \mathbb{N}$ and all $(\tau,t) \in \Omega$, we have

$$||\bar{f}_n g_n(\tau,t) - \bar{f}g(\tau,t)||_1 \le ||g_n(t)||_1 ||f_n(\tau - t) - f(\tau - t)||_2 + ||f(\tau - t)||_2 ||g_n(t) - g(t)||_1.$$

By uniform convergence, an $n_1 \in \mathbb{N}$ exists such that for all $t \in [0,T]$, $n \ge n_1$ implies

$$||g_n(t) - g(t)||_1 < \frac{\epsilon/2}{1 + ||f||_\infty}.$$

Then we also obtain for $n > n_1$:

$$||g_n(t)||_1 \le ||g_n(t) - g(t)||_1 + ||g(t)||_1 < ||g(t)||_\infty + \frac{\epsilon/2}{1 + ||f||_\infty} \le ||g||_\infty + \frac{\epsilon}{2}.$$

Now also by uniform convergence, an $n_2 \in \mathbb{N}$ exists such that for all $(\tau,t) \in \Omega$, $n \ge n_2$ implies

$$||f_n(\tau - t) - f(\tau - t)||_2 < \frac{\epsilon/2}{||g||_\infty + \epsilon/2}.$$

Then for $n \geq \max\{n_1, n_2\}$, the following holds for all $(\tau, t) \in \Omega$:

$$\begin{aligned} ||\bar{f}_n g_n(\tau,t) - \bar{f}g(\tau,t)||_1 &< \frac{\left(||g||_\infty + \frac{\epsilon}{2}\right)\epsilon/2}{||g||_\infty + \epsilon/2} + \frac{||f(\tau - t)||_2 \epsilon/2}{1 + ||f||_\infty}, \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

□

Lemma 4.2.6. Let the definitions be as in Lemma 4.2.5. Let us further suppose that all f_n and g_n are continuous for $n \in \mathbb{N}$. (In particular, we know that f and g are continuous and hence all functions considered are also Riemann integrable.) We now work with a norm $||.||_1$ on $\mathbb{R}^k$ that respects the triangle inequality for integrals (e.g., the Euclidean norm). Define for all $n \in \mathbb{N}$ the function $F_n : [0, T] \to \mathbb{R}^k$ by $F_n(t) = \int_0^t f_n(t-s)g_n(s)\,\mathrm{d}s$. By Theorem 4.2.2 and Lemma 4.2.5 F_n converges pointwise to the function F defined by $F(t) = \int_0^t f(t-s)g(s)\,\mathrm{d}s$. This convergence is uniform on $[0, T]$.
Proof:
Let $\epsilon > 0$ be given. Due to the uniform convergence established in Lemma 4.2.5, an $n_1 \in \mathbb{N}$ exists such that $n \geq n_1$ implies for all $(t, s) \in \Omega$:

$$||f_n(t-s)g_n(s) - f(t-s)g(s)||_1 < \frac{\epsilon}{T}.$$

Then we obtain for $n \geq n_1$ and all $t \in [0, T]$:

$$\begin{aligned} \left|\left| \int_0^t f_n(t-s)g_n(s)\,\mathrm{d}s - \int_0^t f(t-s)g(s)\,\mathrm{d}s \right|\right|_1 &\leq \int_0^t ||f_n(t-s)g_n(s) - f(t-s)g(s)||_1\,\mathrm{d}s, \\ &< \int_0^t \frac{\epsilon}{T}\,\mathrm{d}s, \\ &= \frac{\epsilon t}{T} \leq \epsilon. \end{aligned}$$

□

4.3 The Laplace transform and monotonicity

To consider the so-called over damped case of the mode-coupling theory equations, we will need some results about the Laplace transform and monotonicity of functions. We define these notions and summarize the necessary results here. We will give an example, immediately illustrating the usefulness of the Laplace transform.

Definition 4.3.1. [30, section 1.1] Let f be a real or complex valued function of one variable. The **Laplace transform** of f, denoted $\mathcal{L}(f)$ or $\hat{f}$ is the function of the complex variable s defined by

$$\mathcal{L}(f)(s) = \hat{f}(s) = \int_0^\infty e^{-s\tau} f(\tau)\,\mathrm{d}\tau,$$

whenever the integral exists.

We can also refer to this integral as the **Laplace integral** at s, as is done in [8].

Theorem 4.3.2. [8, Theorem 3.1] A Laplace integral which converges absolutely at some point $s_0 \in \mathbb{C}$, converges absolutely in the right half-plane $\{s \in \mathbb{C} \mid \mathrm{Re}(s) > \mathrm{Re}(s_0)\}$.

The Laplace transform can be used to turn equations into simpler ones, with properties that we present now.

Theorem 4.3.3. (Integration Theorem, [8, Theorem 8.1]) Let f be some real or complex valued function defined on $[0, \infty)$. Define $\phi(t) = \int_0^t f(\tau)\,\mathrm{d}\tau$. If $\mathcal{L}(f)(s)$ converges for some real $s = x_0 > 0$, then $\mathcal{L}(\phi)(s)$ converges for $s = x_0$, and we have:

$$\hat{\phi}(s) = \mathcal{L}(s) = \frac{1}{s}\mathcal{L}(s) = \frac{1}{s}\hat{f}(s) \text{ for } s = x_0 \text{ and for } \mathrm{Re}(s) > x_0.$$

So when f has a convergent Laplace transform, the transform of the integral can be related to that one.

Theorem 4.3.4. (Differentiation Theorem, [8, Theorem 9.1]). If $f : [0, \infty) \to \mathbb{K}, t \mapsto f(t)$, where $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ is differentiable for $t > 0$, and $\mathcal{L}(f')(s)$ converges for some real $s = x_0 > 0$, then the limit $f(0^+)$ exists, and $\mathcal{L}(f)(s)$ too converges for $s = x_0$. We have the relation

$$\mathcal{L}(f')(s) = s\hat{f}(s) - f(0^+) \text{ for } s = x_0 \text{ and for } \operatorname{Re}(s) > x_0.$$

Definition 4.3.5. $\mathcal{T}_0$ is the set of functions from $[0, \infty)$ to $\mathbb{R}$ or $\mathbb{C}$ which are absolutely integrable and bounded on every finite interval that does not include the origin.

Theorem 4.3.6. (Convolution Theorem, [8, Theorem 10.1]) If $f_1, f_2 \in \mathcal{T}_0$ and $\mathcal{L}(f_1)(s)$ and $\mathcal{L}(f_2)(s)$ converge absolutely for $s = s_0$, then $\mathcal{L}(f_1 * f_2)(s)$ converges absolutely for $s = s_0$, and we have

$$\mathcal{L}(f_1 * f_2)(s) = \mathcal{L}(f_1)(s) \cdot \mathcal{L}(f_2)(s) \text{ for } \operatorname{Re}(s) \geq \operatorname{Re}(s_0).$$

Observation 4.3.7. A few more properties of the convolution integral are useful. They are briefly presented in [30, section 2.7] and follow quite readily from the definition and properties of the Riemann integral. Let f, g, h real or complex valued functions on some interval $[a, b]$ and let $c \in \mathbb{R}$. Then $c(f * g) = cf * g = f * cg$, $f * (g * h) = (f * g) * h$ and $f * (g + h) = (f * g) + (f * h)$.

It is also necessary to know if the Laplace transform is one-to-one and can be inverted after some manipulations, to go back to the original problem setting. This is ensured by the following theorem.

Theorem 4.3.8. (Uniqueness Theorem, [8]) Two functions, whose Laplace transforms are the same (in a right half-plane where both their Laplace integrals converge), differ at most by a null function.

From this, [30] infers the following theorem:

Theorem 4.3.9. (Lerch's Theorem, [30, Theorem 1.23]) Distinct continuous functions on $[0, \infty)$ have distinct Laplace transforms.

Hence if for some Laplace transform we encounter a continuous inverse transform, we can be sure it is the unique continuous inverse transform.

Example 4.3.10. (Ordinary differential equation) Consider the initial-value problem for a function $y : [0, \infty) \to \mathbb{R}$ given by

$$\frac{d}{dt}y(t) + y(t) = 1, \quad y(0) = 0.$$

Let us assume that the solution of this equation satisfies the conditions of Theorem 4.3.4. Applying the Laplace transform to both sides of the equality and using linearity, we then obtain

$$\mathcal{L}\left(\frac{dy}{dt}\right)(s) + \mathcal{L}(y)(s) = s\mathcal{L}(y)(s) - 0 + \mathcal{L}(y)(s) = \mathcal{L}(1)(s) = \frac{1}{s}.$$

The differential equation has thus been transformed to an algebraic equation for $\mathcal{L}(y)$, which can be solved to give

$$\mathcal{L}(y)(s) = \frac{1}{s(s+1)}.$$

To do an inverse Laplace transform in order to solve for the function y, one usually resorts to tables. For example, [30, p. 210] immediately gives for y:

$$y(t) = 1 - e^{-t}.$$

This procedure replaces the search for a solution to the homogeneous problem and a particular solution as is involved in the standard solution technique of ordinary differential equations. We will use this algebraic technique to formally solve an iterative process in section 5.2.

Definition 4.3.11. [33, Definition IV-2a] A function $f : [a, b] \to \mathbb{R}$ is **absolutely monotone** on the interval $[a, b]$ if it is continuous and has non-negative derivatives of all orders, satisfying

$$f^{(k)}(x) \geq 0 \quad (a < x < b; k = 0, 1, 2, \dots).$$

Definition 4.3.12. [33, Definition IV-2c] A function $f : [a, b] \to \mathbb{R}$ is **completely monotone** on the interval $[a, b]$ if and only if $f(-x)$ is absolutely monotone on $[-b, -a]$, i.e. f satisfies

$$(-1)^k f^{(k)}(x) \geq 0 \quad (a < x < b; k = 0, 1, 2, \dots).$$

Theorem 4.3.13. [33, Theorem IV-2b] If $f_1 : [a, b] \to \mathbb{R}$ is an absolutely monotone function on $[a, b]$, and if $f_2 : [a, b] \to [a, b]$ is completely monotone there, then $f_1 \circ f_2$ is completely monotone there.

We will generalise this result to a multivariate function.

Theorem 4.3.14. Let $n \in \mathbb{R}$. If $F : [a, b]^n \to \mathbb{R}$ is absolutely monotone in each of its n variables[5] and if $f_1, \dots, f_n : [a, b] \to [a, b]$ are completely monotone functions, then $M : [a, b] \to \mathbb{R}, t \mapsto F(f_1(t), \dots, f_n(t))$ is completely monotone on on $[a, b]$.
Proof:
It suffices to use the chain rule to obtain

$$\frac{d}{dt}M(t) = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i}(x_1, \dots, x_n)\bigg|_{(x_1,\dots,x_n)=(f_1(t),\dots,f_n(t))} \cdot \frac{df_i}{dt}(t).$$

This will be smaller than zero, because all terms in the summation are as for all x_i, $\frac{\partial F}{\partial x_i}$ satisfies absolute monotonicity and all f_i satisfy complete monotonicity. With the product rule, we obtain for the second derivative of M:

$$\begin{aligned}\frac{d^2}{dt^2}M(t) = &\sum_{i=1}^{n} \left(\sum_{j=1}^{n} \frac{\partial^2 F}{\partial x_j \partial x_i}(x_1, \dots, x_n)\bigg|_{(x_1,\dots,x_n)=(f_1(t),\dots,f_n(t))} \cdot \frac{df_j}{dt}(t)\frac{df_i}{dt}(t) \right) \\ &+ \frac{\partial F}{\partial x_i}(x_1, \dots, x_n)\bigg|_{(x_1,\dots,x_n)=(f_1(t),\dots,f_n(t))} \cdot \frac{d^2 f_i}{dt^2}(t).\end{aligned} \tag{4.11}$$

This will be non-negative, because the first derivatives of some f_i that occur are all non-positive, but always multiplied, and the second derivative of the f_i is non-negative by complete monotonicity. Further, all partial derivatives of F of first and second order are positive by the assumption of absolute monotonicity.
This line of thought continues: applying the product rule, for a next derivative all derivatives of some f_i are either multiplied by a new non-negative derivative of some f_j due to the chain rule, or differentiated once more. In either case, the contribution to the sum of derivatives will change sign and so will the derivative itself. □

We adapted the next result to a scalar valued function, because introducing the higher dimensional notions is not necessary for our purposes.

Theorem 4.3.15. [11, Theorem 5-2.6] The Laplace transform $\hat{a}(z)$ of a function $a : (0, \infty) \to \mathbb{C}$ that is locally integrable on $\mathbb{R}^+$ and completely monotone on $(0, \infty)$ has the following properties:

(i) $\hat{a}$ has an analytic extension to the region $\mathbb{C}\backslash\mathbb{R}^-$;

(ii) $\hat{a}(x) = \hat{a}^*(x)$ for $x \in (0, \infty)$;

(iii) $\lim_{x\to\infty} \hat{a}(x) = 0$ $(x \in \mathbb{R})$;

(iv) $\mathrm{Im}(\hat{a}(z)) \leq 0$ for $\mathrm{Im}(z) > 0$;

(v) $\mathrm{Im}(z\hat{a}(z)) \geq 0$ for $\mathrm{Im}(z) > 0$, and $\hat{a}(x) \geq 0$ for $x \in (0, \infty)$.

Conversely, every function $\hat{a}$ that satisfies (i)-(iii), together with (iv) or (v), is the Laplace transform of a function $a : (0, \infty) \to \mathbb{C}$ that is locally integrable on $\mathbb{R}^+$ and completely monotone on $(0, \infty)$.

4.4 Metric and normed spaces

We will be considering spaces of functions with the structure that they obtain from a metric. Also, for $\mathbb{R}^n$, a metric different from the Euclidean metric was already defined in definition 4.1.12. We will also consider some different metrics here. Using different ideas will allow us to exploit different properties of the functions considered in chapter 5. We will also summarize some notions about compactness in metric spaces from the literature.

[5] By this statement, we mean that partial differentiation with respect to all variables and for higher-order partial derivatives, all combinations of differentiation directions must result in non-negative derivatives on $[a, b]^n$. The reason for this demand relates directly to the occurrence of all possible combinations of partial derivatives in (4.11).

Definition 4.4.1. Let $a, b \in \mathbb{R}, n \in \mathbb{N}$. Also let $k \in \mathbb{N} \cup \{0\}$. Then $C^k([a,b], \mathbb{R}^n)$ denotes the vector space of all continuous functions from $[a,b]$ into $\mathbb{R}^n$ that are (at least) k times differentiable.

Definition 4.4.2. ($||.||_{(k)}$) Notation as in definition 4.4.1. Let us define the norm $||.||_{(k)}$ on $C^k([a,b], \mathbb{R}^n)$ which we define for $f \in C^k([a,b], \mathbb{R}^n)$ by

$$||.||_{(k)} : C^k([a,b], \mathbb{R}^n) \to \mathbb{R}, \quad ||f||_{(k)} = \sum_{i=0}^{k} \sup_{t \in [a,b]} ||f^{(i)}(t)||, \tag{4.12}$$

where $||.||$ is the Euclidean norm on $\mathbb{R}^n$. This norm can be defined, since we look at functions which are k times continuously differentiable on the closed and bounded interval $[a,b]$. Since continuous functions on closed and bounded intervals are bounded, all three suprema needed to define $||f||_{(k)}$ will exist whenever $f \in X$, for *finite* intervals $[a,b]$. The usefulness for working with this norm will become clear in 5.3.4. Moreover, this norm turns $C^k([a,b], \mathbb{R}^n)$ into a Banach space [26], i.e., all Cauchy sequences in $C^k([a,b], \mathbb{R}^n)$ will converge to a (unique) limit in $C^k([a,b], \mathbb{R}^n)$.

The particular advantage of this norm will become clear when working with the differentiation operator, which is linear and bounded when one uses this norm.

Remark 4.4.3. We will consider matrix-valued functions $m : \mathbb{R}^n \to \mathbb{R}^{n \times n}$ for some $n \in \mathbb{N}$. As matrices, the function values can again act as operators on vectors in $\mathbb{R}^n$. We will therefore consider $\mathbb{R}^{n \times n}$ as the vector space of (matrix representations of) bounded linear operators from $\mathbb{R}^n$ to $\mathbb{R}^n$, equipped with the induced operator norm from the Euclidean norm on $\mathbb{R}^n$. Then, for any $x \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$ it holds that $||Mx|| \leq ||M|| \cdot ||x||$ [19, 2.7-7].
The same result applies to arbitrary bounded linear operators between normed vector spaces, when no specific norm is mentioned. That is, if U and V are normed vector spaces and $A : U \to V$ is a bounded linear operator with norm $||A||$, we consider the induced operator norm on A and we have (denoting all norms by $||.||$) $\forall u \in U : ||Au|| \leq ||A|| \cdot ||u||$.

Observation 4.4.4. (Compactness, [19, Appendix A1.5]) Most generally, a topological space (or a subset thereof) is called compact if every open cover of this space or subset contains a finite sub cover (for a subset, one considers the induced subspace topology, defined by intersecting all open sets of the topological space with the set under consideration). A topological space or subset thereof is called sequentially compact if every sequence in the set contains a converging subsequence (thus with limit in the same set). In metric spaces, and hence also in normed spaces, the two notions are equivalent, and when compactness is mentioned we can give a proof by demonstrating sequential compactness. We further note that any compact subset of a metric space (and hence also of a normed space) is closed and bounded [19, Lemma 2.5-2].

Definition 4.4.5. ($\mathcal{C}(X,Y)$, [5, p. 348].) Let U, V be compact metric spaces. We define the metric space $\mathcal{C}(U,V)$ as the set consisting of all continuous functions from X to Y, equipped with the **supremum metric** $d_\infty : \mathcal{C}(U,V) \times \mathcal{C}(U,V) \to \mathbb{R}$ defined by

$$\forall f, g \in \mathcal{C}(U,V) : d_\infty(f,g) = \sup_{u \in U} d_V(f(u), g(u))),$$

where $d_V : V^2 \to \mathbb{R}$ is the metric defined for V. The compactness of U ensures that this metric is well-defined, since continuous functions map compact sets to compact sets [24, Theorem 26.5], which are bounded by observation 4.4.4.
Similarly, when V is a vector space with norm $||.||_V : V \to \mathbb{R}$, we can define the **supremum norm** $||.||_\infty : \mathcal{C}(U,V) \to \mathbb{R}$ by

$$\forall f \in \mathcal{C}(U,V) : d_\infty(f) = \sup_{u \in U} ||f(u)||_V.$$

Definition 4.4.6. (Compact operator, [5, Section 2.3].) An operator $\mathcal{A}$, mapping one normed vector space into another, is said to be **compact** if it maps the unit ball in its domain to a set whose closure is compact.

In this case, the image of the unit ball is also called relatively compact in the literature, e.g. in [19]. Here the closed unit ball is meant, i.e. for a normed space U with norm $||.||$ we consider the image of the set $\{u \in U \mid ||u|| \leq 1\}$.

4.5 Topology: the Arzelà-Ascoli theorem

Here we introduce some notions about equicontinuity in order to state the Arzelà-Ascoli theorem. Equicontinuity, lemma about closure, Arzelà-Ascoli. It characterizes compactness in spaces of continuous functions.

Definition 4.5.1. ([5, p. 348].) Let U, V be metric spaces and let $\mathcal{F}$ be a subset of $\mathcal{C}(U,V)$. Let d_U and d_V denote the metrics on U and V, respectively. Then $\mathcal{F}$ is said to be **equicontinuous** (or an equicontinuous family of functions) if the following holds:

$$\begin{gathered}\forall \epsilon > 0 \quad \exists \delta = \delta(\epsilon) \quad \forall f \in \mathcal{F} \quad \forall u, v \in U : \\ d_U(u,v) < \delta \Rightarrow d_V\left(f(u), f(v)\right) < \epsilon.\end{gathered} \tag{4.13}$$

That is, uniform continuity for each function in $\mathcal{F}$ can be established by use of the same number $\delta(\epsilon)$ once ϵ has been given.

Lemma 4.5.2. Let U, V be compact metric spaces with metrics d_U and d_V, respectively, and let $\mathcal{F} \subseteq \mathcal{C}(U,V)$ be equicontinuous. Assume that $f \in \mathcal{C}(U,V)$ and that there exists a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ such that $f_n \to f$ as $n \to \infty$ in $\mathcal{C}(U,V)$, i.e. in the supremum norm. (When for infinitely many $n \in \mathbb{N}$ we have that $f_n \neq f$, f is called an accumulation point of $\mathcal{F}$.) Then $\mathcal{F} \cup \{f\}$ is equicontinuous and for each $\epsilon > 0$, a choice of δ can be made that is independent of the specific function f or the sequence $(f_n)_{n\in\mathbb{N}}$, but only depends on $\mathcal{F}$ (and on ϵ, of course).[6]
Proof:
Due to convergence in the supremum norm, we have

$$\begin{gathered}\forall \tilde{\epsilon} \quad \exists N = N(\tilde{\epsilon}) \in \mathbb{N} \quad \forall n \geq N, u \in U : \\ d_V\left(f_n(u), f(u)\right) \leq \sup_{v\in U} d_V\left(f_n(v), f(v)\right) = d_\infty\left(f_n, f\right) < \tilde{\epsilon}.\end{gathered} \tag{4.14}$$

Let $\epsilon' > 0$ be arbitrary and let $u', v' \in U$. Then for all $n \in \mathbb{N}$, the following holds:

$$d_V\left(f(u'), f(v')\right) \leq d_V\left(f(u'), f_n(u')\right) + d_V\left(f_n(u'), f_n(v')\right) + d_V\left(f_n(v'), f(v')\right).$$

Now take $N \in \mathbb{N}$ such that (4.14) holds for $\tilde{\epsilon} = \frac{1}{2}\epsilon'$. Continuing with $n \geq N$, we obtain

$$d_V\left(f(u'), f(v')\right) < \tilde{\epsilon} + \tilde{\epsilon} + d_V\left(f_n(u'), f_n(v')\right) = \epsilon' + d_V\left(f_n(u'), f_n(v')\right). \tag{4.15}$$

Now let $\epsilon > 0$ be given. We will establish uniform continuity of f for this ϵ. We must prove that this can be done independently of the sequence $(f_n)_{n\in\mathbb{N}}$ or the function f. So let us take, in the notation of (4.13), $\delta = \delta(\epsilon/2)$ and let $u, v \in U$ with $d_U(u,v) < \delta$. Using in (4.15) that for such u and v we have for all $g \in \mathcal{F}$ that $d_V\left(g(u), g(v)\right) < \epsilon/2$, it holds that

$$d_V\left(f(u), f(v)\right) < \epsilon' + \frac{1}{2}\epsilon.$$

Since this inequality has been derived for arbitrary $\epsilon' > 0$ and ϵ is a fixed number, we obtain

$$d_V\left(f(u), f(v)\right) \leq \frac{1}{2}\epsilon < \epsilon.$$

Equicontinuity of $\mathcal{F} \cup \{f\}$ is thus established. Further, definition 4.5.1 ensures, that the equicontinuity can be proven only based on the fact that $\mathcal{F}$ is equicontinuous, regardless of the accumulation point of $\mathcal{F}$ that is considered. Hence the lemma has been proven. □

Corollary 4.5.3. (Closure of equicontinuous family of functions) Since the closure $\bar{\mathcal{F}}$ of $\mathcal{F}$ in $\mathcal{C}(U,V)$ can be characterized as the union of $\mathcal{F}$ and all its accumulation points, it follows from the foregoing lemma that uniform continuity of any $f \in \bar{\mathcal{F}}$ can be established for all $\epsilon > 0$ with one choice of δ, namely $\delta(\epsilon/2)$ in (4.13). Hence, the closure of an equicontinuous set in $\mathcal{C}(U,V)$ is itself again equicontinuous.

Theorem 4.5.4. (Arzelà-Ascoli, [5, Theorem 7.4-1].) Let U and V be compact metric spaces. A subset of $\mathcal{C}(U,V)$ is compact if and only if it is closed and equicontinuous.

[6] In the case that $f \in \mathcal{F}$, this fact is a trivial consequence of definition 4.5.1, but f need not be in $\mathcal{F}$ (and then it is certainly an accumulation point). Furthermore, we assume in the proof that U contains accumulation points (or equivalently, that the interior of U is non-empty), because otherwise the existence of $u, v \in U$ with $d_U(u,v) < \delta$ is not guaranteed for arbitrary $\delta > 0$. However, for isolated points in U, establishing uniform continuity is trivial.

4.6 Differentiation of operators in normed spaces

Some of the results needed for the existence and uniqueness rely on the differentiation of operators. In order to make sense of differentiation of an operator $\mathcal{A}$, we must use an extension from real (vector) valued functions to functions between arbitrary normed spaces. We adopt the presentation in [31, Section 9.2]. Let us start with an illustration of the usefulness of differentiation in a linearisation.

Example 4.6.1. (Stability of ordinary differential equations) Let us consider for a function $y : [0, T] \to \mathbb{R}^n$ for some $n \in \mathbb{N}$ the autonomous differential equation

$$\frac{d}{dt}y = F(y)$$

for a sufficiently smooth function $F : \mathbb{R}^n \to \mathbb{R}^n$. Suppose for $y_{eq} \in \mathbb{R}^n$ that $F(y_{eq}) = 0$, i.e., y_{eq} is an equilibrium point of the differential equation. Sufficiently close to y_{eq}, one can study properties of the linearised system

$$\frac{d}{dt}u = DF(y_{eq})u.$$

Here $DF(y_{eq})$ is the Jacobian matrix of F at y_{eq}, giving the total derivative. The eigenvalues of $DF(y_{eq})$ are used in [2, chapter 9] to describe the stability of the equilibrium y_{eq}. The linearisation is thus used to extract very useful properties of the system. We note that it is theoretically covered by the concept of differentiation. We now study a definition of differentiation that is more general than for functions of real variables, which will allow us to use the central theorem of section 5.3.

Definition 4.6.2. Let U, V be normed spaces and let $T : U \to V$ be an operator. Suppose $\Omega = \mathcal{D}(T) \subseteq U$ is an open set ($\mathcal{D}(T)$ denotes the domain of T), and let $u \in \Omega$. If a bounded[7] linear operator $T'(u) : U \to V$ exists such that

$$\lim_{||\Delta u|| \to 0} \frac{||T(u + \Delta u) - T(u) - (T'(u))(\Delta u)||}{||\Delta u||} = 0, \tag{4.16}$$

then $T'(u)$ is called the Fréchet derivative of the operator T at the vector u and T is said to be Fréchet-differentiable at u. The operator $T' : U \to \mathcal{B}(U, V)$, mapping $u \mapsto T'(u)$ is called the **Fréchet derivative** of T ($\mathcal{B}(U, V)$ denotes the set of all bounded linear operators from U to V). The domain $\mathcal{D}(T') \subseteq U$ consists of all vectors in U where the limit (4.16) exists. We say that T is Fréchet differentiable (on U) when $\mathcal{D}(T') = U$.

It follows readily from this definition that if two operators S and T are Fréchet-differentiable at some vector $u \in U$, then so is their sum and any scalar multiple. We have for all $\alpha \in \mathbb{R}$: $(\alpha T)'(u) = \alpha T'(u)$ and $(T + S)'(u) = T'(u) + S'(u)$. Also, when T is Fréchet-differentiable at u it is continuous at u [31, Theorem 9.2.4]. We shall from now on just write differentiable for Fréchet-differentiable. This can easily be done since for differentiable functions $f : \mathbb{R}^m \to \mathbb{R}^n$, the Fréchet derivative also exists and coincides with the linear operator represented by the $m \times n$ Jacobian matrix of f at all points where it is differentiable [31, Example 9.2.8]. Finally, if T has a Fréchet derivative, this bounded linear operator is unique [31, Theorem 9.2.5 and further].

Observation 4.6.3. (Bounded linear operators) Let U, V be normed spaces and let $T : U \to V$ be a bounded linear operator. We denote both the norm on U and on V by $||.||$. Then T is differentiable on U and T' maps u to T for all $u \in U$ (i.e., T' is a constant mapping from U to $\mathcal{B}(U, V)$). The proof follows immediately by noting that the linearity of T allows us to write for all $u, \Delta u \in U : T(u + \Delta u) = T(u) + T(\Delta u)$. Then (4.16) gives

$$\lim_{||\Delta u|| \to 0} \frac{||T(u + \Delta u) - T(u) - T(\Delta u)||}{||\Delta u||} = \lim_{||\Delta u|| \to 0} \frac{||T(u) + T(\Delta u) - T(u) - T(\Delta u)||}{||\Delta u||} = \lim_{||\Delta u|| \to 0} \frac{||0||}{||\Delta u||} = 0.$$

Since the Fréchet derivative is unique, we have shown that $\forall u \in U : T'(u) = T$.
Should we wish to differentiate this operator again after differentiation at a (fixed) vector u,[8] now differentiating at $v \in U$, we note the following: since the derivative gives the same operator regardless of the vector in U at which we compute the derivative, we have for each vector $u \in U$ in which we compute the first derivative (which we will denote as $T'(u) = T'_u$ here) and each vector $v \in U$ at which we wish to differentiate again, that

$$\lim_{||\Delta v|| \to 0} \frac{||T'_u(v) - T'_u(v + \Delta v) - T'_u(\Delta v)||}{||\Delta v||} = \lim_{||\Delta v|| \to 0} \frac{||T(v) - T(v + \Delta v) - T(\Delta v)||}{||\Delta v||} = 0,$$

[7] [31] also demands that his operator be continuous, but we know that for linear operators between normed spaces, boundedness and continuity are equivalent [19, Theorem 2.7-9].

[8] We did not explicitly define higher-order derivatives, but we just need to apply the definition to $T'(u)$, since this is again an operator from U to V. Note that this yields a mapping from $U \times U$ to $\mathcal{B}(U, V)$, so to specify the second derivative as an operator from U to V (i.e., in $\mathcal{B}(U, V)$), we must provide *two* arguments, and we even need three if we want to know its value in V at some vector in U.

because by linearity of T, the numerator yields zero for all $\Delta v \in V$. Hence, differentiating a bounded linear operator twice will always yield the same operator again, and continuing the differentiation process, we can differentiate as many times as we desire at all vectors in U and always know that T is the operator that results.

Definition 4.6.4. Let U, V be normed spaces, $D \subset U$ open and $T : U \to V$ differentiable on D. The space $\mathcal{B}(U, V)$ of bounded linear operators can also be turned into a normed vector space, with for instance the induced operator norm. We say that T is **continuously differentiable** if T' is continuous with the norms on U ($||.||_U$) and $\mathcal{B}(U, V)$ ($||.||_\mathcal{B}$), i.e.

$$\forall \epsilon > 0 \exists \delta > 0 \forall u, v \in D : ||u - v||_U < \delta \Rightarrow ||T'(u) - T'(v)||_\mathcal{B} < \epsilon.$$

Theorem 4.6.5. (Chain rule, [5, Theorem 3.2-1]) Let U, V, W be normed spaces and let $D \subseteq U$ and $D' \subseteq V$ be open sets. Let $T : D \to V$ and $S : D' \to W$ be operators. If T is differentiable at $u \in U$ with $T(u) \in D'$ and S is differentiable at $T(u)$, then $S \circ T$ is differentiable at u, and

$$(S \circ T)'(u) = S'(T(u)) \circ T'(u).$$

Theorem 4.6.6. (Product rule for convolution) Let $[a, b]$ be an interval in $\mathbb{R}$, $n \in \mathbb{N}$ and let $U = C^1([a, b], \mathbb{R}^n)$, $V = C([0, T], \mathbb{R}^{n\times n})$ and $W = C([0, T], \mathbb{R}^N)$. Let $A : U \to V$, $B : U \to W$ be bounded linear operators that are differentiable at $f \in U$ with derivatives denoted by A'_f and B'_f, respectively. Then also the convolution of the operators, $A * B : U \to W$, is differentiable at f and we have

$$(A * B)'_f = A'_f * B(f) + A(f) * B'_f. \tag{4.17}$$

Proof:
A proof is given in appendix A.

We are now well-equipped for the differentiation of operators to prove the respective part of lemma 5.3.4 in chapter 5. Another nice application of the Fréchet derivative is that the Mean Value Theorem can be generalized. We also derive some results from this theorem.

Theorem 4.6.7. (Mean Value Theorem, [5, Theorem 3.2-4]) Let U, V be normed vector spaces (for both of which we denote the norm by $||.||$), let $D \subset U$ be open and let f be a map from U to V. If for arbitrary $u, v \in U$ the line segment

$$S = \{tu + (1 - t)v \mid 0 \leq t \leq 1\}$$

lies in D and if $f'(x)$ exists for all $x \in S$, then

$$||f(u) - f(v)|| \leq ||u - v|| \sup_{x\in S} ||f'(x)||.$$

Corollary 4.6.8. Let U, V be normed vector spaces and let D be compact and convex subset of U. Let f be a continuously differentiable function from U to V, satisfying $f(u) = 0$ for $x \in U \backslash D$. Then f is bounded on U.
Proof:
If D is empty, we have $\forall x \in U : f(x) = 0$ by assumption and the statement is trivial.
Let us suppose the D is nonempty and let $a \in D$ be fixed. For arbitrary $u \in D$, we have by the triangle inequality: $||f(u)|| \leq ||f(u) - f(a)|| + ||f(a)||$. By the assumption that D is convex, the line segment S from a to u is contained in D and Theorem 4.6.7 can be applied. Let us use the boundedness of D (which is implied by its compactness) to write $\forall x \in D : ||x|| \leq k$. Finally, since f' is continuous on the compact set D by assumption, its image on D is also bounded, by K, say. Then we obtain

$$\begin{aligned} ||f(u)|| &\leq ||f(a)|| + ||u - a|| \sup_{x\in S} ||f'(x)||, \\ &\leq k + (||u|| + ||a||) \sup_{x\in D} ||f'(x)||, \\ &\leq k + 2kK. \end{aligned}$$

Since $u \in D$ was arbitrary and $f(u) = 0$ when $u \notin D$, we have $\forall x \in U : ||f(x)|| \leq k(1 + 2K)$. Hence, f is bounded on U. □

Lemma 4.6.9. Let U, V be normed vector spaces with norms denoted by $||.||$ and let $f : U \to V$ be continuously differentiable on some non-empty, compact and convex subset D of U. Then f is Lipschitz continuous on D, and hence, uniformly continuous on D.
Proof:
Let $u, v \in D$. Since D is convex, the line segment from u to v is contained in D and Theorem 4.6.7 can be applied. Since f' is assumed to be continuous on the compact set D, we can set $L := \max_{x\in D} ||f(x)||$. Then we obtain

$$\begin{aligned} ||f(u) - f(v)|| &\leq ||u - v|| \sup_{x\in S} ||f'(x)||, \\ &\leq ||u - v|| \sup_{x\in D} ||f'(x)||, \\ &= L||u - v||. \end{aligned}$$

Uniform continuity now follows easily: for arbitrary $\epsilon > 0$, let $u, v \in D$ with $||u - v|| < \epsilon/(1 + L)$. Then $||f(u) - f(v)|| < \epsilon$. □

We note that this Lemma actually also proves Lemma 4.1.13, where U is the one-dimensional real line and $V = \mathbb{R}^n$.

4.7 Matrices and differential equations

The theory of matrices in the context of ordinary differential equations will come in useful to analyse the differential parts of the system of generalised mode-coupling theory integro-differential equations. We first recall the definition of the matrix exponential. Its usefulness is at the end of this section illustrated by result 4.7.10.

Definition 4.7.1. For $n \in \mathbb{R}$, let $A \in \mathbb{C}^{n\times n}$ be any square matrix over the complex numbers. Then the **matrix exponential** e^A is defined as

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k.$$

Remark 4.7.2. For any matrix A, absolute convergence of the matrix exponential in the induced operator norm follows from

$$\left|\left| \sum_{k=0}^{\infty} \frac{1}{k!} A^k \right|\right| \leq \sum_{k=0}^{\infty} \frac{1}{k!} ||A^k|| \leq \sum_{k=0}^{\infty} \frac{1}{k!} ||A||^k = e^{||A||}.$$

Since the space $\mathbb{C}^{n\times n}$ is a finite-dimensional vector space over the complete field of the complex numbers, it also follows that $\mathbb{C}^{n\times n}$ is complete and hence absolute convergence implies convergence in $\mathbb{C}^{n\times n}$ for the matrix exponential.

Lemma 4.7.3. [12, Proposition 2.3-3, 2.3-6] Let A be an $n\times n$ matrix. Then e^A is invertible and $\left(e^A\right)^{-1} = e^{-A}$. Further, for any invertible $C \in \mathbb{C}^{n\times n}$, we have $e^{CXC^{-1}} = Ce^X C^{-1}$.

Next our aim is to find a bound on e^A if A is diagonalizable. We derive this bound with the help of a few intermediate calculations and results.

Observation 4.7.4. (Exponential of a diagonal matrix) Let $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ be a diagonal $n \times n$ matrix for some $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. We then have for any $k \in \mathbb{N}$ that $\Lambda^k = \text{diag}({\lambda_1}^k, \ldots, {\lambda_n}^k)$ and it follows that $e^\Lambda = \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n})$.

Observation 4.7.5. Let $A \in \mathbb{C}^{n\times n}$ be a matrix with n eigenvalues $\lambda_1, \ldots, \lambda_n$ and n linearly independent eigenvectors. Let $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$. Then an invertible matrix D exists (containing the eigenvectors A as columns) satisfying $\Lambda = D^{-1}AD$, or $A = D\Lambda D^{-1}$. Applying Lemma 4.7.3 we then find

$$e^A = e^{D\Lambda D^{-1}} = De^\Lambda D^{-1},$$

so e^A is seen to have the same eigenvectors as A (because it is diagonalized by the same matrix D) and as eigenvalues the diagonal values of e^Λ, i.e., in view of observation 4.7.4, $e^{\lambda_1}, \ldots, e^{\lambda_n}$.

Lemma 4.7.6. Let Λ be a symmetric $n \times n$ matrix for some $n \in \mathbb{N}$ (for instance it could be diagonal). Let $||.||$ denote some norm on $\mathbb{C}^n$ which can be derived from an inner product and let $||.||_M$ denote the induced operator norm on $\mathbb{C}^{n\times n}$. We have

$$||\Lambda||_M = \max_{k=1,\dots,n}\{|\lambda_1|,\dots,|\lambda_n|\} =: \lambda_{\max}.$$

Proof:
Since Λ is a symmetric matrix, there exists an orthonormal basis $\{e_1,\dots,e_n\}$ of eigenvectors of Λ. Let the corresponding eigenvalues be denoted by $\lambda_1,\dots,\lambda_n$. Let $v \in \mathbb{C}^n$ be any vector; we can decompose v with respect to the basis of eigenvectors as $v = v_1e_1 + \cdots + v_ne_n$. We then have

$$||\Lambda v|| = ||v_1\lambda_1e_1 + \cdots + v_n\lambda_ne_n|| = |\lambda_1|\cdot||v_1e_1|| + \cdots + |\lambda_n|\cdot||v_ne_n|| \leq \lambda_{\max}\left(||v_1e_1|| + \dots||v_ne_n||\right) = \lambda_{\max}||v||.$$

Hence $||\Lambda||_M \leq \lambda_{\max}$. Now let $m \in \{1,\dots,n\}$ be such that $|\lambda_m| = \lambda_{\max}$. We have for e_m:

$$||\Lambda e_m|| = ||\lambda_m e_m|| = |\lambda_m|\cdot||e_m|| = \lambda_{\max}||e_m||,$$

so also $||\Lambda||_M \geq \lambda_{\max}$. We conclude that $||\Lambda||_M = \lambda_{\max}$. □

Lemma 4.7.7. Let $A \in \mathbb{C}^{n\times n}$ be diagonalizable, i.e. A has n linearly independent eigenvectors). Let $\lambda_1,\dots,\lambda_n$ be the eigenvalues corresponding to the n eigenvectors of A and let $\lambda_{\max}$ be defined as in Lemma 4.7.6. Further let $||.||$ denote some norm on $\mathbb{C}^n$ which can be derived from an inner product and let $||.||_M$ denote the induced operator norm on $\mathbb{C}^{n\times n}$. Then some $b \geq 1$ exists such that

$$||A||_M \leq b\lambda_{\max}.$$

Proof:
Since A is diagonalizable, a matrix D exists such that $A = D\Lambda D^{-1}$. Applying remark 4.4.3 regarding the induced operator norm $||.||_M$ and Lemma 4.7.6 to $\Lambda := \mathrm{diag}(\lambda_1,\dots,\lambda_n)$, we have

$$||A||_M = ||D\Lambda D^{-1}|| \leq ||D||_M||\Lambda D^{-1}||_M \leq ||D||_M||D^{-1}||_M||\Lambda||_M = b\lambda_{\max},$$

with $b := ||D||_M||D^{-1}||_M$. Note that for D we have

$$1 = ||I||_M = ||DD^{-1}||_M \leq ||D||_M||D^{-1}||_M,$$

so in general we really need $b \geq 1$. □

Note that smaller values of b may be possible to bound A. However, we are mainly concerned with the form this estimate can take.

Observation 4.7.8. We can combine observation 4.7.5 with Lemma 4.7.7 to obtain the following result. Let $A \in \mathbb{C}^{n\times n}$ be some diagonalizable matrix with eigenvalues $\lambda_1,\dots,\lambda_n$, say. Then for the eigenvalues of e^A, we have $|e^{\lambda_k}| = e^{\mathrm{Re}(\lambda_k)}$ $(k = 1,\dots,n)$. So if we consider a norm $||.||_M$ on $\mathbb{C}^{n\times n}$ induced by a norm $||.||$ on $\mathbb{C}^n$ which can be derived from an inner product, there exists a $b \geq 1$ such that

$$||e^A||_M \leq be^c,$$

where $c = \max_{k=1,\dots,n}\{\mathrm{Re}(\lambda_1),\dots,\mathrm{Re}(\lambda_n)\}$.

Observation 4.7.9. Let us now extend this result to the matrix exponential function $f : \mathbb{R} \to \mathbb{C}^{n\times n}, t \mapsto e^{At}$. If A is diagonalizable with an invertible matrix D, we have $A = D(\Lambda t)D^{-1}$ and as above it follows that $e^{At} = De^{\Lambda t}D^{-1}$ for a diagonal matrix Λ containing the eigenvalues $\lambda_1,\dots,\lambda_n$ of A. Then the eigenvalues of Λt are $\lambda_1 t,\dots,\lambda_n t$. So we obtain the result of observation 4.7.8, namely that for each $t \in \mathbb{R}$:

$$||f(t)||_M = ||e^{At}||_M \leq be^{c(t)},$$

where for instance $b = ||D||_M||D^{-1}||_M$ can be taken and $c(t) = t\cdot\max_{k=1,\dots,n}\{\mathrm{Re}(\lambda_1),\dots,\mathrm{Re}(\lambda_n)\}$.

Result 4.7.10. We now turn our attention to the application of the matrix exponential to differential equations. Let A be a constant (coefficient) matrix in $\mathbb{R}^{n\times n}$ for some $n \in \mathbb{N}$. For $x \in C^1([0,\infty),\mathbb{R}^n)$ consider the linear system of first-order differential equations

$$\begin{cases} \frac{d}{dt}x(t) = Ax(t), \\ x(0) = x_0 \in \mathbb{R}^n \end{cases}$$

It is detailed in [2, section 7.7] that the unique solution to this initial value problem is

$$x(t) = e^{At}x_0.$$

4.8 Various

There are some more important statements and definitions needed before we can move on to the proof of results pertaining to the mode-coupling theory equations. We give them now at the end of the preliminary results.

Definition 4.8.1. [34, p. 472] Let U, V be Banach spaces with norms denoted by $||.||$. Let $A : U \to V$. Then A is called **weakly coercive** if

$$\lim_{||u||\to\infty} ||Au|| \to \infty.$$

Observation 4.8.2. As is noted by Saal [29], we see that we can equivalently say that an operator A is weakly coercive if and only if the image of an unbounded set is unbounded. This will be used to prove Lemma 5.3.6.

Definition 4.8.3. Let U, V be Banach spaces. A differentiable map $f : U \to V$ is called a **diffeomorphism** if it is a bijection and its inverse $f^{-1} : V \to U$ is also differentiable. For $k \in \mathbb{N}$, f is called a C^k**-diffeomorphism** if f is a diffeomorphism and, moreover, f and f^{-1} are k times continuously differentiable.

Theorem 4.8.4. [21, Theorem 3.11] Let $n \in \mathbb{N}$, $T > 0$, $g \in C^0([0,T], \mathbb{R}^n)$ and $k \in C^0([0,T]^2, \mathbb{R}^{n\times n})$ (meaning that all component functions of g and k are continuous). Then the system of integral equations

$$f(t) = g(t) + \int_0^t k(t,s)f(s)\,\mathrm{d}s \tag{4.18}$$

for $f \in C^0([0,T], \mathbb{R}^n)$ has a unique continuous solution for $t \in [0,T]$.

Remark 4.8.5. The equation (4.18) is referred to as a system of Volterra integral equations of the second kind, because the boundaries of the integral are not fixed (hence the Volterra-type) and the unknown function f appears under and outside the integral (hence the second kind).

5 Existence and uniqueness of solutions

One of the first questions of interest when confronted with some equation is whether a solution actually exists, whether it exists on the (complete) domain of interest and whether such a solution then is unique. When at least the answer to the first of these questions is affirmative, but preferably all, we can start looking for a (or the) solution to the equations. It is this topic about existence and uniqueness that is discussed in this chapter for the generalized mode-coupling theory equations that were presented in chapter 3. This chapter elaborates on the results of [10] and [29].

5.1 Position in literature

Integro-differential equations involve an integral as well as a differential operator of an unknown function for which the equation is to be solved. For integral equations, one usually makes the distinction between Fredholm and Volterra type. For the former, the boundaries of the integral operator are fixed, while for the latter, at least one of them is a variable. One further distinguishes between equations of first and second kind, which correspond to the unknown function occurring only inside or also outside the integral operator, respectively [32]. We have seen an example of a Volterra integral equation of the second kind in Theorem 4.8.4.

When a differential operator is also included, one can still make the characterizations above. We can then classify the schematic generalized mode-coupling theory equations as Volterra integro-differential equations of the second kind. Saal notes that in general, such equations are mostly studied when the unknown function depends in the integral only on the integration variable [29]. However, in equations (3.1) and (3.2), ϕ_n (or its derivative to be precise) and ϕ_{n+1} display dependence on the integration variable τ as well as on the independent variable t.

One usually captures the dependence on the independent variable under the integral in a function called the kernel of the integro-differential equation. We can therefore also say that the unusual property of the equations of chapter 3 is that they have a kernel that depends on the solution of the equation. This property of the kernel is generalized in the mathematical results studied in this chapter, allowing for more general kernels than those used specifically in schematic generalized mode-coupling theory.

5.2 Over-damped equations

In [10], a one-dimensional variant of the over-damped equations (3.1) is considered by Götze, and the existence of a unique solution on a bounded time interval is proved with a uniformly convergent iteration sequence. The extension to higher-dimensional cases is not given, but we will do so here.

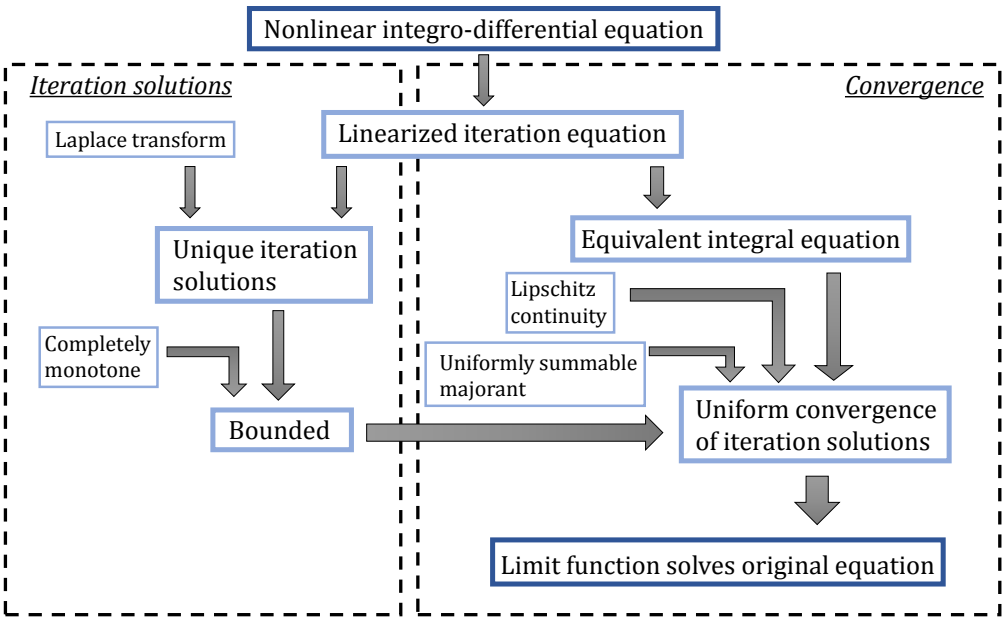

Figure 4: *Flow chart illustrating the existence proof in by Götze. We indicate the most important intermediate conclusions and methods used to obtain them. On the left are the steps to obtain bounded iteration solutions, on the right those that lead to their convergence properties.*

Theorem 5.2.1. (Existence and uniqueness for the over-damped equations) Let $T > 0$. Let us consider for N vector-valued functions $\phi = (\phi_1, \phi_2, \ldots, \phi_N)$ with $\phi_1, \ldots, \phi_N \in C^1([0,T], \mathbb{R}^N)$ the integro-differential equations

$$\begin{cases} \tau_q \dot{\phi}_q(t) = -\phi_q(t) - \int_0^t m_q(t-\tau)\dot{\phi}_q(\tau)\,\mathrm{d}\tau, \\ \phi_q(0) = 1, \end{cases} \qquad q = 1, \ldots, N, \quad 0 < t \leq T, \tag{5.1}$$

where for each q, $\tau_q > 0$ and $m_q(t) = F_q(\phi(t))$ and F_q is absolutely monotone in each of the N variables f_q restricted to $0 \leq f_q < 1 + \delta^*$ for some $\delta^* > 0$. Under these conditions, a unique solution ϕ to (5.1) exists.

Note that, since the value of m_q in (5.1) depends on ϕ, the functions F_q can be used to establish the closure that one chooses to use in the problem.

We illustrate the structure of Götze's proof for the existence part in figure 4. Based on a linearised, iterative variant of (5.1), we find unique iteration solutions via the Laplace transform, that are completely monotone. After integrating the equation, we can use these properties to derive some very useful estimates. These will allow us to identify a uniformly summable majorant sequence for the distances between successive iteration solutions. Finally then, the existence of a solution to (5.1) follows. Then we end this section with a proof for the uniqueness of such a solution.

Iteration sequence

We will consider the following iteration sequence to establish existence of a solution to (5.1): for $n = 1$, define $\phi_q^{(n)} : [0,T] \to \mathbb{R}$ by $\phi_q^{(n)}(t) = \exp(-\Gamma t)$ for some $\Gamma \geq 0$. For $n \geq 1$, define $\phi_q^{(n+1)}$ as the solution of the equation

$$\begin{cases} \tau_q \dot{\phi}_q^{(n+1)}(t) + \phi_q^{(n+1)}(t) + \int_0^t m_q^{(n)}(t-\tau)\dot{\phi}_q^{(n+1)}(\tau)\,\mathrm{d}\tau = 0, \\ \phi_q^{(n+1)}(0) = 1, \end{cases} \tag{5.2}$$

$$m_q^{(n)}(t) = F_q\left(\phi^{(n)}(t)\right) = F_q\left(\phi_1^{(n)}, \ldots, \phi_N^{(n)}\right).$$

Note that $m_q^{(n)}$ does not depend on any of the $\phi_q^{(n+1)}$, which it defines by means of a linearised variant of (5.1) because the ϕ-dependence of the kernel is suppressed.

Via the Laplace transform, we will show that these equations do yield a solution for $\phi_q^{(n+1)}$. (By Theorem 4.3.9, the Laplace transform yields an equivalent equation if we look at continuous functions, provided the Laplace transforms exit. This is justified as we find below that by Theorem 4.3.14, the Laplace transformed equations can be solved.) Using the results 4.3.4 and 4.3.6, we obtain from (5.2) the equation

$$\begin{aligned} s\tau_q \hat{\phi}_q^{(n+1)}(s) - \tau_q \phi_q^{(n+1)}(0) + \hat{\phi}_q^{(n+1)}(s) + \hat{m}_q^{(n)}(s)\hat{\dot{\phi}}_q^{(n+1)}(s) = 0, \\ s\tau_q \hat{\phi}_q^{(n+1)}(s) - \tau_q + \hat{\phi}_q^{(n+1)}(s) + \hat{m}_q^{(n)}(s)\left(s\hat{\phi}_q^{(n+1)}(s) - 1\right) = 0, \end{aligned}$$

leading to

$$\hat{\phi}_q^{(n+1)}(s) = \frac{\tau_q + \hat{m}_q^{(n)}(s)}{1 + s\left(\tau_q + \hat{m}_q^{(n)}(s)\right)}. \tag{5.3}$$

Complete monotonicity

We argue inductively that all iteration functions $\phi_q^{(n)}$ are completely monotone, $q = 1, \ldots, N$, $n \in \mathbb{N}$. This clearly is the case for $n = 1$. Let us assume that some $\phi_q^{(n)}$ is completely monotone. Applying Theorem 4.3.14 to $m_q^{(n)} = F_q \circ \phi^{(n)}$, we find that also $m_q^{(n)}$ is completely monotone with the assumption that F_q is completely monotone in all its N variables in the sense of the footnote in Theorem 4.3.14. Then its Laplace transform $\hat{m}_q^{(n)}$ fulfils the properties (i)-(v) of Theorem 4.3.14 and this is all we need to show that also $\hat{\phi}_q^{(n+1)}$ as given by (5.3) fulfils the properties (i)-(iii) and (v):

(i) Given that $\hat{m}_q^{(n)}$ has an analytic extension to the region $\mathbb{C}\backslash\mathbb{R}$, we can use this extension, which we shall also denote by $\hat{m}_q^{(n)}$ here, to define an analytic extension of $\hat{\phi}_q^{(n+1)}$ by plugging it in equation (5.3). This will then be an analytic function for all s in $\mathbb{C}$ that satisfy $1 = s(\tau_q + \hat{m}_q^{(n)}(s)) \neq 0$. Götze does not mention or explore this demand at all. At this point in time, we must also leave this subtlety unstudied.

(ii) We phrase this property of $\hat{m}_q^{(n)}(x)$ as being real-valued for $x \in \mathbb{R}^+$. We then see immediately that for such x, the value of $\hat{\phi}_q^{(n+1)}(x)$ as defined by equation (5.3) is real as well, for only real numbers appear in it.

(iii) Making use of basic limit properties and the assumed limit property of $\hat{m}_q^{(n)}$ in Theorem 4.3.14, we obtain for $x \in \mathbb{R}^+$:

$$\lim_{x\to\infty} \hat{\phi}_q^{(n+1)}(x) = \frac{\tau_q + \lim\limits_{x\to\infty} \hat{m}_q^{(n)}(x)}{1 + \left(\lim\limits_{x\to\infty} x\right)\left(\tau_q + \lim\limits_{x\to\infty} \hat{m}_q^{(n)}(x)\right)} = \frac{\tau_q + 0}{1 + \tau_q \lim\limits_{x\to\infty} x} = 0.$$

(v) To show that $\mathrm{Im}(\hat{\phi}_q^{(n+1)}(s)) \leq 0$ for $\mathrm{Im}(s) > 0$, we introduce the notation $s = x + iy$ with $x, y \in \mathbb{R}$ for arbitrary $s \in \mathbb{C}$ and $\mu_x := \mathrm{Re}(\hat{m}_q^{(n)}(s))$, $\mu_y := \mathrm{Im}(\hat{m}_q^{(n)}(s))$.
We first look at property (v) for $\hat{m}_q^{(n)}$, which yields for $y > 0$ that $\mathrm{Im}(s\hat{m}_q^{(n)}(s)) \geq 0$, or $x\mu_y + y\mu_x \geq 0$. To study the imaginary part of $s\hat{\phi}_q^{(n+1)}(s)$, we use expression (5.3) and write

$$s\hat{\phi}_q^{(n+1)}(s) = (x+iy)\frac{1 + (x-iy)(\tau_q + \mu_x - i\mu_y)}{1 + (x-iy)(\tau_q + \mu_x - i\mu_y)} \cdot \frac{\tau_q + \mu_x + i\mu_y}{1 + (x+iy)(\tau_q + \mu_x + i\mu_y)}.$$

Now in the denominator, we obtain a complex number multiplied by its complex conjugate, which is the modulus squared. Hence we can write this as some number $\alpha > 0$ (since we assume the denominator not to be 0) and we obtain

$$\begin{aligned} s\hat{\phi}_q^{(n+1)}(s) &= \frac{1}{\alpha}\left(x + iy + (x^2+y^2)(\tau_q + \mu_x - i\mu_y)\right)(\tau_q + \mu_x + i\mu_y), \\ &= \frac{1}{\alpha}\left\{\left(x + (\tau_q + \mu_x)(x^2+y^2)\right) + i\left(y - \mu_y(x^2+y^2)\right)\right\}(\tau_q + \mu_x + i\mu_y). \end{aligned}$$

A careful inspection then shows that for the imaginary part, we have

$$\begin{aligned} \alpha\,\mathrm{Im}\left(s\hat{\phi}_q^{(n+1)}(s)\right) &= \tau_q y + y\mu_x - \tau_q\mu_y(x^2+y^2) - \mu_x\mu_y(x^2+y^2) \\ &\quad + x\mu_y + \tau_q\mu_y(x^2+y^2) + \mu_x\mu_y(x^2+y^2) \Leftrightarrow \\ \mathrm{Im}\left(s\hat{\phi}_q^{(n+1)}(s)\right) &= \frac{1}{\alpha}(\tau_q y + x\mu_y + y\mu_x), \\ &\geq 0, \end{aligned}$$

where the final inequality results from $\alpha, y, \tau_q > 0$ and the inequality $x\mu_y + y\mu_x \geq 0$ derived above.

Hence the second part of Theorem 4.3.14 yields that $\hat{\phi}_q^{(n+1)}$ is the Laplace transform of some completely monotone function, which we define to be $\phi_q^{(n+1)}$. Clearly this can be done for all q. This completes the induction argument. So the iteration sequences for all q are indeed well-defined by (5.2) and this is done uniquely in view of Theorem 4.3.9.

Now, due to complete monotonicity of all $\phi_q^{(n)}$, we know that $\dot{\phi}_q^{(n)}(t) \leq 0$ and therefore all $\phi_q^{(n)}$ are non increasing, such that $\phi_q^{(n)}(t) \leq \phi_q^{(n)}(0) = 1$ holds for all $q = 1, \ldots, N$ and all $n \in \mathbb{N}$. Also, due to complete monotonicity of all $m_q^{(n)}$, we know that these functions are all non-negative, such that $\phi_q^{(n+1)}(t) = 0$ for some $t > 0$ would lead to

$$\begin{aligned} \tau_q\dot{\phi}_q^{(n+1)}(t) &= -\int_0^t m_q^{(n)}(t-\tau)\dot{\phi}_q^{(n+1)}(\tau)\,\mathrm{d}\tau, \\ &> 0, \end{aligned}$$

by virtue of the definition (5.2) if $m_q^{(n)} \neq 0$, because $\dot{\phi}_q^{(n+1)}$ (which has a positive derivative) cannot be constantly zero on $[0,t]$ if $\phi_q^{(n+1)}$ goes from 1 to 0. This contradicts the complete monotonicity of $\phi_q^{(n)}$. If $m_q^{(n)}$ is the zero function, all solutions $\phi_q^{(n)}$ will be decreasing exponentials, hence they will also not intersect zero. We can thus conclude for all $t \in [0,T]$:

$$0 < \phi_q^{(n)} \leq 1, \quad q = 1, \ldots, N, \quad n \in \mathbb{N}. \tag{5.4}$$

($\phi_q^{(n)}$ can certainly not be smaller than 0 if it cannot be zero, as a consequence of the Intermediate Value Theorem that applies to continuous functions.)

Equivalent equation
To further analyse the iteration sequence, we will once integrate the definition (5.2). To simplify the expression,

we need the Leibniz rule for differentiation of integrals (4.10). Substituting u for t in (5.2) and reversing the integration direction in the convolution, this allows us to write for all $q = 1, \dots, N$:

$$\begin{aligned}
\int_0^u m_q^{(n)}(u-\tau)\dot{\phi}_q^{(n+1)}(\tau)\,\mathrm{d}\tau &= \int_0^u m_q^{(n)}(\tau)\dot{\phi}_q^{(n+1)}(u-\tau)\,\mathrm{d}\tau, \\
&= \int_0^u m_q^{(n)}(\tau)\frac{\partial}{\partial u}\phi_q^{(n+1)}(u-\tau)\,\mathrm{d}\tau, \\
&= \int_0^u \frac{\partial}{\partial u} m_q^{(n)}(\tau)\phi_q^{(n+1)}(u-\tau)\,\mathrm{d}\tau, \\
&= \frac{d}{du}\int_0^u m_q^{(n)}(\tau)\phi_q^{(n+1)}(u-\tau)\,\mathrm{d}\tau - m_q^{(n)}(u)\phi_q^{(n+1)}(0), \\
&= \frac{d}{du}\int_0^u m_q^{(n)}(u-\tau)\phi_q^{(n+1)}(\tau)\,\mathrm{d}\tau - m_q^{(n)}(u).
\end{aligned}$$

Still making the substitution u for t in (5.2), we see that integration over u from 0 to t yields

$$\begin{aligned}
\tau_q\phi_q^{(n+1)}(t) - \tau_q &= -\int_0^t \phi_q^{(n+1)}(\tau)\,\mathrm{d}\tau - \int_0^t \frac{d}{du}\int_0^u m_q^{(n)}(u-\tau)\phi_q^{(n+1)}(\tau)\,\mathrm{d}\tau\,\mathrm{d}u + \int_0^t m_q^{(n)}(\tau)\,\mathrm{d}\tau, \\
&= -\int_0^t \phi_q^{(n+1)}(\tau) - m_q^{(n)}(\tau)\,\mathrm{d}\tau - \int_0^t m_q^{(n)}(t-\tau)\phi_q^{(n+1)}(\tau)\,\mathrm{d}\tau \Leftrightarrow \\
\phi_q^{(n+1)}(t) &= 1 + \frac{1}{\tau_q}\int_0^t m_q^{(n)}(\tau) - \phi_q^{(n+1)}(\tau) - m_q^{(n)}(t-\tau)\phi_q^{(n+1)}(\tau)\,\mathrm{d}\tau, \\
&= 1 + \int_0^t K_q\left(\phi^{(n)}(\tau), \phi^{(n)}(t-\tau), \phi_q^{(n+1)}(\tau)\right)\mathrm{d}\tau, \qquad (5.5)
\end{aligned}$$

defining $K_q : \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R} \to \mathbb{R}$ as the function

$$K_q(x_1, \dots, x_N, y_1, \dots, y_N, z) = \frac{1}{\tau_q}\left(F_q(x_1, \dots, x_N) - F_q(y_1, \dots, y_N)z - z\right). \qquad (5.6)$$

Consequence of Lipschitz continuity

In view of (5.4), all K_q only need to be considered on the closed domain $[0,1]^{2N+1}$. Since all F_q are infinitely many times differentiable with respect to all their variables in $[0, \delta)$ for some $\delta > 1$ by assumption, it follows readily with the chain rule that all K_q have a continuous derivative on $[0,1]^{2N+1}$. So lemma 4.1.13 can be applied to K_q and it is Lipschitz continuous for all $q = 1, \dots, N$. Hence there exist constants $L_q > 0$ such that

$$\begin{aligned}
&\forall x, y, u, v \in [0,1]^N, z, w \in [0,1] : \\
|K(x_1, \dots, x_N, y_1, \dots, y_N, z) - K(u_1, \dots, u_N, v_1, \dots, v_N, w)| &\le L\left(|z-w| + \sum_{i=1}^N |x_i - u_i| + |y_i - v_i|\right), \\
&= L_q\left\{\delta(x,u) + \delta(u,v) + \delta(z,w)\right\},
\end{aligned}$$

(where $x = (x_1, \dots, x_N)$ etc.). Now define $L := \max_{q=1,\dots,N} L_q$, such that the inequality holds for all $q = 1, \dots, N$ with the same constant L.

Majorant

For the convergence of $\phi^{(n)}$, we will need a uniformly convergent majorant. We define this sequence here. Let $a_0 : [0,T] \to \mathbb{R}$ be defined by $a_0(t) = N$, and for $n \in \mathbb{N}$ let

$$a_n(t) = 3NL\int_0^t a_{n-1}(\tau)\,\mathrm{d}\tau,$$

so we readily see that for all $n \in \mathbb{N}$, $a_n(t) = N\frac{(3NL)^n}{n!}t^n$ and because $|a_n(t)| \leq a_n\left(\min\{1,T\}\right) = N\frac{(3NL\cdot\min\{1,T\})^n}{n!}$, we see that $a_n \to 0$ as $n \to \infty$ uniformly.

Estimates
We will use the sequence of functions $(a_n)_{n\in\mathbb{N}}$ defined above to demonstrate the uniform convergence of the iteration sequence $\phi_q^{(n)}$. Define

$$X^{(n)} : [0,T] \to \mathbb{R}, \text{ for } t \in [0,T] : X^{(n)}(t) = \delta\left(\phi^{(n+1)}(t), \phi^{(n)}(t)\right), \quad n \in \mathbb{N}.$$

Since all components of all $\phi^{(n)}(t)$ lie between 0 and 1 for all values of t, we have

$$\forall n \in \mathbb{N}, t \in [0,T] : X^{(n)}(t) = \sum_{q=1}^{N}\left|\phi_q^{(n+1)}(t) - \phi_q^{(n)}(t)\right| \leq N. \tag{5.7}$$

Using (5.5), obtained from the iterative definition of $\phi^{(n+1)}$, and the Lipschitz continuity of all K_q, we find for $X^{(n)}(t)$ $(n \in \mathbb{N})$:

$$\begin{aligned}
X^{(n)}(t) &= \sum_{q=1}^{N}\left|\int_0^t K_q\left(\phi^{(n)}(\tau), \phi^{(n)}(t-\tau), \phi_q^{(n+1)}(\tau)\right) - K_q\left(\phi^{(n-1)}(\tau), \phi^{(n-1)}(t-\tau), \phi_q^{(n)}(\tau)\right)\mathrm{d}\tau\right|,\\
&\leq \sum_{q=1}^{N}\int_0^t \left|K_q\left(\phi^{(n)}(\tau), \phi^{(n)}(t-\tau), \phi_q^{(n+1)}(\tau)\right) - K_q\left(\phi^{(n-1)}(\tau), \phi^{(n-1)}(t-\tau), \phi_q^{(n)}(\tau)\right)\right|\mathrm{d}\tau,\\
&\leq \sum_{q=1}^{N}\int_0^t L_q\left\{\delta\left(\phi^{(n)}(\tau), \phi^{(n-1)}(\tau)\right) + \delta\left(\phi^{(n)}(t-\tau), \phi^{(n-1)}(t\tau)\right) + \delta\left(\phi_q^{(n+1)}(\tau), \phi_q^{(n)}(\tau)\right)\right\}\mathrm{d}\tau,\\
&\leq \sum_{q=1}^{N}\int_0^t L\left\{X^{(n-1)}(\tau) + X^{(n-1)}(t-\tau) + X^{(n)}(\tau)\right\}\mathrm{d}\tau,\\
&= NL\int_0^t 2X^{(n-1)}(\tau) + X^{(n)}(\tau)\,\mathrm{d}\tau,
\end{aligned} \tag{5.8}$$

using $\int_0^t X^{(n)}(t-\tau)\,\mathrm{d}\tau = -\int_{u=t}^{u=0} X^{(n)}(u)\,\mathrm{d}u = \int_0^t X^{(n)}(\tau)\,\mathrm{d}\tau$ in the last equality.

Now we can identify the sequence $(a_n)_{n\in\mathbb{N}}$ as a majorant for the sequence $\left(X^{(n)}\right)_{n\in\mathbb{N}}$ with the help of mathematical induction. From (5.7), it immediately follows that $X^{(0)}(t) \leq a_0(t)$ for all t in $[0,T]$. We proceed by the principle of strong induction to show that

$$\forall n \in \mathbb{N} \quad \forall t \in [0,T] \quad \forall k \in \{1,\ldots,n\} : 0 \leq X^{(n)}(t) \leq a_k(t). \tag{5.9}$$

(The left inequality is already provided for by definition.) Let us assume for some $k \in \mathbb{N}$ that for all $t \in [0,T]$: $X^{(k)}(t) \leq a_n(t)$ for $n = 1, 2, \ldots, k$. Then we need to show that also $X^{(k+1)}(t) \leq a_{k+1}(t)$ holds. We have due to (5.8) and using (5.7):

$$\begin{aligned}
X^{(k+1)}(t) &\leq NL\int_0^t 2X^{(k)}(\tau) + X^{(k+1)}(\tau)\,\mathrm{d}\tau,\\
&\leq NL\int_0^t 2\cdot N + N\,\mathrm{d}\tau,\\
&= 3NL\int_0^t a_0(\tau)\,\mathrm{d}\tau = a_1(\tau).
\end{aligned}$$

Using this inequality and the induction hypothesis, and repeatedly making use of the newly found inequalities,

we obtain

$$\begin{aligned}
X^{(k+1)}(t) &\leq NL \int_0^t 2X^{(k)}(\tau) + X^{(k+1)}(\tau)\,\mathrm{d}\tau, \\
&\leq NL \int_0^t 2a_1(\tau) + a_1(\tau)\,\mathrm{d}\tau = a_2(t), \\
X^{(k+1)}(t) &\leq NL \int_0^t 2a_2(\tau) + a_2(\tau)\,\mathrm{d}\tau = a_3(t), \\
&\vdots \\
X^{(k+1)}(t) &\leq NL \int_0^t 2a_k(\tau) + a_k(\tau)\,\mathrm{d}\tau = a_{k+1}(t),
\end{aligned}$$

which concludes the proof of the desired property (5.9).

Convergence

With the property (5.9) at hand, we can show that all sequences $\left(\phi_q^{(n)}\right)_{n\in\mathbb{N}}$ $(q = 1, \ldots, N)$ converge uniformly on $[0, T]$. We will show that it satisfies the Cauchy criterion 4.2.1 for uniform convergences of sequences.

Let $\epsilon > 0$ be given. Let $t \in [0, T]$, $q \in \{1, \ldots, N\}$ and $m, n \in \mathbb{N}$. Without loss of generality, we assume $m > n$. Let $\beta = 3NL \min\{1, T\}$. Then we can derive the following inequalities:

$$\begin{aligned}
\left|\phi_q^{(n)}(t) - \phi_q^{(m)}(t)\right| &\leq \delta\left(\phi^{(n)}(t), \phi^{(m)}(t)\right), \\
&\leq \delta\left(\phi^{(n)}(t), \phi^{(n+1)}(t)\right) + \delta\left(\phi^{(n+1)}(t), \phi^{(n+2)}(t)\right) + \delta\left(\phi^{(m-1)}(t), \phi^{(m)}(t)\right), \\
&= X^{(n)}(t) + \cdots + X^{(m-1)}(t), \\
&\leq a_n(t) + \cdots + a_{m-1}(t), \\
&\leq a_n(\min\{1, T\}) + \ldots a_{m-1}(\min\{1, T\}), \\
&= N \sum_{k=n}^{m-1} \frac{\beta^k}{k!} < N \sum_{k=n}^{\infty} \frac{\beta^k}{k!}.
\end{aligned}$$

Now for the last series, we recognize that $\sum_{k=0}^{\infty} \beta^k/k! = e^\beta$, and in particular, the latter series converges. Hence $\sum_{k=n}^{\infty} \beta^k/k! \to 0$ as $n \to \infty$. So we can pick $n^* \in \mathbb{N}$ such that $\sum_{k=n}^{\infty} \beta^k/k! < \epsilon/N$ whenever $n, m \geq n^*$. This choice is clearly independent of t, and even q. So for all $q \in \{1, \ldots, N\}$ uniform convergence to a function $\tilde{\phi}_q : [0, T] \to \mathbb{R}$ follows according to Theorem 4.2.1 for $q = 1, \ldots, N$. This means that we can also define the vector-valued limit function $\tilde{\phi} : [0, T] \to \mathbb{R}^N$ as $\tilde{\phi} = (\tilde{\phi}_1, \ldots, \tilde{\phi}_N)$.

Properties of the limit function

At this point, Götze concludes in the one-dimensional proof that uniform convergence of the derivatives of $\phi_1^{(n)}$ also follows trivially in the same way. We accept this statement without further proof here. Since we have shown the same properties of each of our component functions $\phi_q^{(n)}$ $(1 = q, \ldots, N)$, we conclude that their derivatives also converge uniformly. The uniform convergence of the iteration functions and their derivatives allows us by Theorem 4.2.2 to take the limit inside the integral and the differentiation in the iteration equation (5.2) and we then see that $\tilde{\phi}$ solves equation (5.1). We have thus shown the existence part of Theorem 5.2.1.

Uniqueness

We finalize the proof of Theorem 5.2.1 by an argument based on an integral inequality, guided by the uniqueness proof in [2, Section 2.8] for ordinary differential equations.

Suppose two solutions $x, y : [0, T] \to \mathbb{R}^N$ to equation (5.1) exist. We denote their component functions by x_q and y_q, respectively, for $q = 1, \ldots, N$. The equivalent equation (5.5) for a solution to (5.1) with x and y

substituted for both iterations of $\phi^{(n)}$ yields that we have for each q:

$$\begin{aligned}|x_q(t) - y_q(t)| &= \left| \int_0^t K_q\left(x(\tau), x(t-\tau), x_q(\tau)\right) - K_q\left(y(\tau), y(t-\tau), y_q(\tau)\right) \mathrm{d}\tau \right|, \\ &\leq \int_0^t L\left\{\delta\left((x(\tau), y(\tau)\right) + \delta\left((x(t-\tau), y(t-\tau)\right) + \delta\left((x_q(\tau), y_q(\tau)\right)\right\}, \\ &\leq 3L \int_0^t \delta\left(x(\tau), y(\tau)\right) \mathrm{d}\tau. \end{aligned} \tag{5.10}$$

We also have by the definition of our δ-metric

$$\begin{aligned} 0 \leq |x_q(t) - y_q(t)| &\leq \sum_{k=1}^{N} |x_k(t) - y_k(t)|, \\ &= \delta\left(x(t), y(t)\right). \end{aligned} \tag{5.11}$$

Let us now define $U : [0,T] \to \mathbb{R}$ by $U(t) = \int_0^t \delta\left(x(\tau), y(\tau)\right) \mathrm{d}\tau$ for $t \in [0,T]$. We then see immediately that U satisfies (because x, y are differentiable and therefore continuous, such that also $\delta\left(x(t), y(t)\right)$ is a continuous function of t):

$$\frac{d}{dt} U(t) = \delta\left(x(t), y(t)\right), \quad U(0) = 0. \tag{5.12}$$

With the help of (5.10), we derive that

$$\begin{aligned} \delta\left(x(t), y(t)\right) &= \sum_{q=1}^{N} |x_q(t) - y_q(t)|, \\ &\leq \sum_{q=1}^{N} 3L \int_0^t \delta\left(x(\tau), y(\tau)\right) \mathrm{d}\tau, \\ &= 3LNU(t). \end{aligned} \tag{5.13}$$

Combining (5.12) and (5.13), and multiplying by $\exp(-3LNt)$ we find that

$$\begin{aligned} e^{-3LNt} \frac{d}{dt} U(t) - 3LNe^{-3LNt} U(t) &\leq 0, \\ \frac{d}{dt} e^{-3LNt} U(t) &\leq 0. \end{aligned}$$

Integrating finally gives that $e^{-3LNt} U(t) \leq U(0) = 0$ for $t \in [0,T]$. This implies $U(t) \leq 0$, but from its definition it is clear that $U(t) \geq 0$. So we conclude that $U(t) = 0 \quad \forall t \in [0,T]$. Consequently, we also have $\delta(x(t), y(t)) = \frac{d}{dt} U(t) = 0$ and according to (5.11) it now holds for all $t \in [0,T], q = 1, \ldots, N$ that

$$x_q(t) = y_q(t).$$

Sine x and y have now been shown to be equal component-wise, we have established that a solution to equation (5.1) is necessarily unique.

We can now end this section since the existence and uniqueness as asserted by Theorem 5.2.1 have been demonstrated.

5.3 Second-order system I

In the existing literature, a paper by Martin Saal provides several theorems concerning existence and uniqueness of solutions to the following equation [29]:

$$\begin{cases} \lambda \ddot{\psi}(t) + \dot{\psi}(t) + \psi(t) + \int_0^t m\left(\psi(t-\tau)\right) \dot{\psi}(\tau) d\tau = f(t), \\ \psi(0) = \Psi_0, \\ \dot{\psi}(0) = \Psi_1, \end{cases} \tag{5.14}$$

where the vector-valued function $\psi \in C^2([0,T],\mathbb{R}^N)$ for some $T > 0$ and some $N \in \mathbb{N}$ (which we identify with the closure level chosen), $\lambda > 0$ and $m : \mathbb{R}^N \to \mathbb{R}^{N\times N}$ is a matrix-valued function describing how the kernel depends on ψ and determines the coupling between the N component functions of ψ. Further, the initial values $\Psi_0, \Psi_1 \in \mathbb{R}^N$ and the source function $f : \mathbb{R} \to \mathbb{R}^N$ are assumed to be given.

To be able to speak of the different component functions of ψ, we introduce for these the notation $\psi = (\psi_1, \psi_2, \ldots, \psi_N)$, where for $n = 1, 2, \ldots, N : \psi_n \in C^2([0,T],\mathbb{R})$.

We will consider two conditions that lead to existence and uniqueness of a solution to (5.14) as provided by Saal [29]. He also proves these statements, but not all arguments are made explicit. In this section, we aim to extend the arguments such that the proof is understandable at a Bachelor's level, enriched by chapter 4. We state the theorem and introduce some suitable notation for the proof. Then we state a result which turns the proof into showing some properties of a non-linear operator by three lemmas. These are proved extensively afterwards.

Theorem 5.3.1. [29, Theorem 1]
Let $m \in C^1(\mathbb{R}^N, \mathbb{R}^{N\times N})$. Then for any given $\Psi_0, \Psi_1 \in \mathbb{R}^N$ and $f \in C^0([0,\infty),\mathbb{R}^N)$ there exists some $T > 0$ and a unique solution $\psi \in C^2([0,T],\mathbb{R}^N)$ of (5.14).
T is arbitrary if

$$\exists c > 0 \forall x, y \in \mathbb{R}^N : ||m(x)y|| \leq c(1 + ||x||)||y|| \tag{5.15}$$

holds.

Notation
Let $X := C^2([0,T],\mathbb{R}^N)$, where N is the closure level considered in equation (5.14) and T is some fixed real number defining the bounded interval on which the equation is considered. So X defines the vector space in which we wish to find a solution to (5.14). We will refer to the normed space $\left(X, ||.||_{(2)}\right)$ (see definition 4.12) simply by writing X.

Let us now define $Y := C^0([0,T],\mathbb{R}^N)\times\mathbb{R}^N\times\mathbb{R}^N$ with for $(f,x,y) \in Y$ the norm $||.||$ defined by $||(f,x,y)||_Y = \sup_{t\in[0,T]} |f(t)| + ||x|| + ||y||$. Here $||.||$ denotes the Euclidean norm on $\mathbb{R}^N$.[9] Then define the operator $\mathcal{A} : X \to Y$ as

$$\mathcal{A}(\psi) := \begin{pmatrix} \lambda\ddot{\psi} + \dot{\psi} + \psi + m(\psi) * \dot{\psi} \\ \psi(0) \\ \dot{\psi}(0) \end{pmatrix}.$$

Then the system (5.14) can be rewritten as

$$\mathcal{A}(\psi) = \begin{pmatrix} f \\ \Psi_0 \\ \Psi_1 \end{pmatrix}$$

and Theorem 5.3.1 is proven if we show that $\mathcal{A}$ is invertible for arbitrary $T > 0$.

Saal shows that $\mathcal{A}$ is a C^k-diffeomorphism for $1 \leq k \leq \infty$ by splitting $\mathcal{A} = \mathcal{L} + \mathcal{N}$, where $\mathcal{L}$ is a linear operator defined by

$$\mathcal{L}(\psi) := \begin{pmatrix} \lambda\ddot{\psi} + \dot{\psi} + \psi \\ \psi(0) \\ \dot{\psi}(0) \end{pmatrix}$$

and $\mathcal{N}$ is the non-linear part of $\mathcal{A}$ defined by

$$\mathcal{N}(\psi) := \begin{pmatrix} m(\psi) * \dot{\psi} \\ 0 \\ 0 \end{pmatrix}. \tag{5.16}$$

The argument used by Saal is a proof that the result below can be applied to $\mathcal{A}$. No explicit reference is given, so we refer to Saal.

[9] [29] uses X for this space and does not abbreviate what we called X, but introducing notation for both seems most convenient to me.

Theorem 5.3.2. [29, p. 3] Let the notation be as introduced above. If $\mathcal{L}$ (a linear operator) is a C^∞-diffeomorphism from X onto Y, then $\mathcal{A} = \mathcal{L} + \mathcal{N}$ is a C^k-diffeomorphism ($1 \leq k \leq \infty$) if the following conditions are satisfied:

(i) $\mathcal{N}$ is a compact C^k-mapping;

(ii) The derivative $\mathcal{A}'(\psi)$ is injective for all $\psi \in X$;

(iii) $\mathcal{A}$ is weakly coercive.

Now the proof under condition (5.15) is constructed by showing that all conditions are met by the operator $\mathcal{A}$ as we encounter it to capture equation (5.14). This is done in four separate lemmas that we prove in this chapter. We summarize this approach in figure 5, where we make explicit that the linear growth condition on m, (5.15), is only needed to show weak coerciveness. Local existence is then derived by cutting off the kernel at hand if it does not satisfy condition (5.15), and uniqueness in that situation is proved in a last lemma.

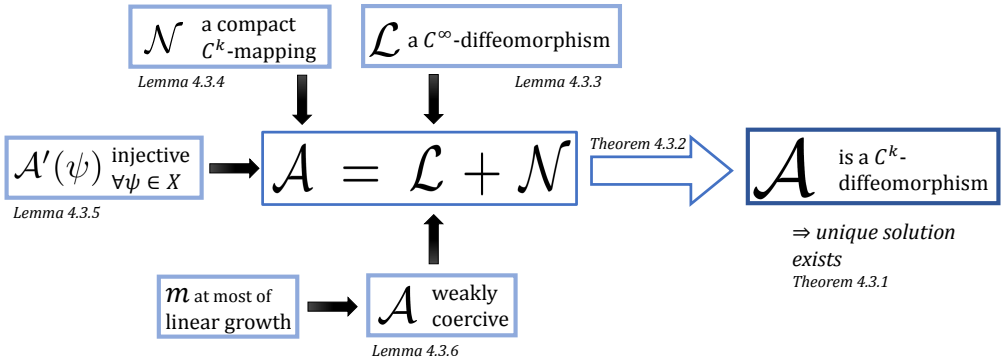

Figure 5: *Flow chart interconnecting the Lemmas that we prove in this section to prove Theorem 5.3.1, global existence and uniqueness of a solution for (5.14). The linear growth condition (5.15) is only needed in the steps to show weak coerciveness.*

Lemma 5.3.3. It is asserted by [29] that $\mathcal{L}$ is a C^∞-diffeomorphism from X onto Y.
This can be seen as follows: $\mathcal{L}$ is surjective due to the existence theorem for systems of ordinary differential equations. For instance, since f is continuous by assumption, after a transformation of $\mathcal{L}(\psi) = f$ to a first order system, [2, Theorem 7.1.2] can be applied and this yields that a unique solution ψ exists globally on $[0, T]$. Injectivity follows by the same theorem, which also ensures uniqueness of the solution. Hence $\mathcal{L}$ is an invertible linear operator from X to Y. Observation 4.6.3 yields that $\mathcal{L}$ is infinitely many times differentiable. Since the inverse of a linear operator is itself a linear operator (if it exists) [19, Theorem 2.6-10], it is also infinitely many times differentiable. Hence $\mathcal{L}$ is a C^∞-diffeomorphism. □

Lemma 5.3.4. [29, Lemma 2] Let $m \in C^k(\mathbb{R}^N, \mathbb{R}^{N\times N})$ ($1 \leq k \leq \infty$). Then $\mathcal{N}$ is a compact C^k-operator with the derivative given by

$$\mathcal{N}'(\psi) : X \to Y, \quad \mathcal{N}'(\psi)h = \begin{pmatrix} (m'(\psi)(h)) * \dot{\psi} + m(\psi) * \dot{h} \\ 0 \\ 0 \end{pmatrix}. \tag{5.17}$$

Proof:
To prove the lemma, we have to establish differentiability and compactness for $\mathcal{N}$. For the latter, we will show that the image of the closed unit ball in X is compact and hence closed. Then its closure coincides with the image itself because a set is closed if and only if it equals its closure [24, Corollary 17.7].
1 Differentiability of $\mathcal{N}$.
Let us denote the differentiation operator by $\mathcal{D} : X \to C^0([0, T], \mathbb{R}^N)$, and rewrite the effect of $\mathcal{N}$ on some $\psi \in X$ as

$$\mathcal{N}(\psi) = \begin{pmatrix} m(\psi) * \mathcal{D}(\psi) \\ 0 \\ 0 \end{pmatrix},$$

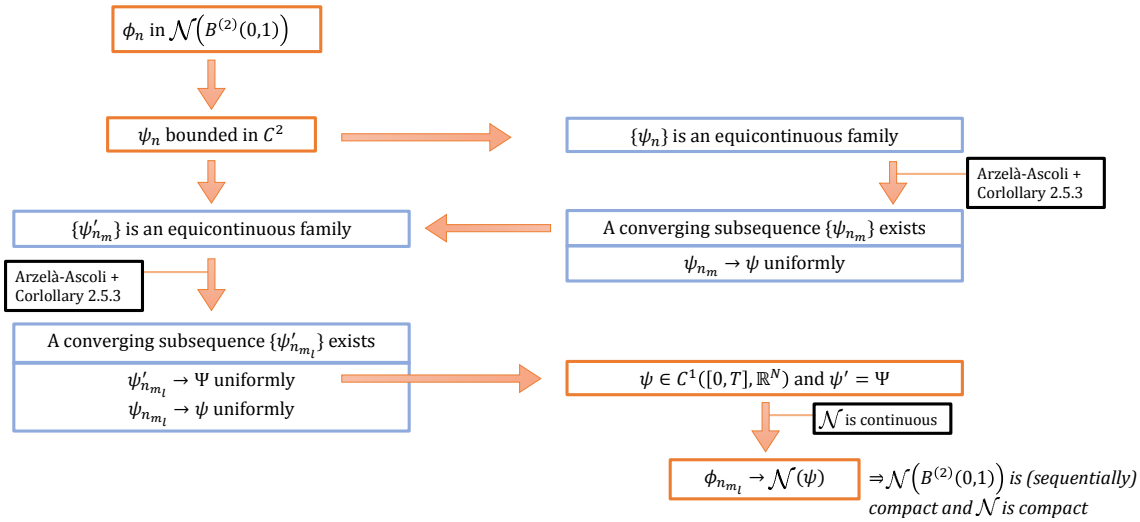

Figure 6: *Overview of the steps taken in the compactness proof of Lemma 5.3.4 and the arguments that justify the implications we make.*

or omitting the presence of ψ in the notation:

$$\mathcal{N} = \begin{pmatrix} m * \mathcal{D} \\ 0 \\ 0 \end{pmatrix}.$$

So $\mathcal{N}$ is the convolution of two operators ("followed" by twice the zero mapping). Now m is k times differentiable by assumption. With respect to the norms used on X and Y, $\mathcal{D}$ is a bounded operator, for if $\psi \in X$ with norm $||\psi||_{(2)}$, then

$$\sup_{t\in[0,T]} |\mathcal{D}(\psi)(t)| = \sup_{t\in[0,T]} |\dot{\psi}(t)| \leq ||\psi||_{(2)}. \tag{5.18}$$

Furthermore, differentiation is linear, so $\mathcal{D}$ is a bounded linear operator and in view of observation 4.6.3, it is differentiable as many times as we wish. Hence the chain rule for the convolution product (stated as (4.17)) yields (after repeated application) that $\mathcal{N}$ is k times continuously differentiable and more explicitly, we have for the first derivative that[10]

$$\mathcal{N}' : X \to \mathcal{B}(X,Y), \quad \mathcal{N}' = \begin{pmatrix} m' * \mathcal{D} + m * \mathcal{D}' \\ 0 \\ 0 \end{pmatrix},$$

which yields after evaluation at some $\psi \in X$ an operator from X to Y. Remembering that the fact that $\mathcal{D}$ is a linear operator implies that $\mathcal{D}'(\psi) = \mathcal{D}$ for all $\psi \in X$, we find:

$$\forall \psi \in X : \mathcal{N}'(\psi) : X \to Y \text{ is for } h \in X \text{ given by}$$
$$\mathcal{N}(\psi)(h) = \begin{pmatrix} m'(\psi) * \mathcal{D}(\psi) + m(\psi) * \mathcal{D} \\ 0 \\ 0 \end{pmatrix}(h) = \begin{pmatrix} (m'(\psi)(h)) * \mathcal{D}(\psi) + m(\psi) * \mathcal{D}(h) \\ 0 \\ 0 \end{pmatrix},$$

as is the claim in (5.17).

2 Compactness of $\mathcal{N}$.

Let $\bar{B}^{(2)}(0,1)$ denote the closed unit ball in X. In order to prove compactness of $\mathcal{N}$, we start with a sequence $(\phi_n)_{n\in\mathbb{N}}$ in $\mathcal{N}(\bar{B}^{(2)}(0,1))$. Then there is by definition a sequence $(\psi_n)_{n\in\mathbb{N}}$ in $\bar{B}^{(2)}(0,1)$ such that for all $n \in \mathbb{N}$: $\mathcal{N}(\psi_n) = \phi_n$.

We first show that $(\psi_n)_{n\in\mathbb{N}}$ contains a subsequence which converges to a limit with a differentiable derivative. This will yield sequential compactness, hence compactness of $\mathcal{N}(\bar{B}^{(2)}(0,1))$. The steps are summarized by figure 6. Let us denote the compact interval $[0,T]$ by I. For all $\psi \in \bar{B}^{(2)}(0,1)$, we know that $||\psi||_{(2)} \leq 1$,

[10] We are a little sloppy with respect to the vector-valued character of the values that $\mathcal{N}$ takes, since we only consider the first of the three spaces $C^0([0,T],\mathbb{R}^N)$, $\mathbb{R}^N$ and $\mathbb{R}^N$. However, since we add the norms to obtain Y, it is easily seen that related results (such as differentiation of the operators considered here) can equivalently be treated in each space separately. Since the quantities in $\mathbb{R}^N$ associated with $\mathcal{N}$ are constant mappings, their differentiation yields zero.

which implies for the supremum norm that $||\psi||_\infty = \sup_{t\in I} ||\psi(t)|| \leq 1$[11]. So the image of I under some ψ lies in $\tilde{R} := \{y \in \mathbb{R}^N \mid ||y|| \leq 1\}$ and we can write $\bar{B}^{(2)}(0,1) \subseteq C^2(I,\tilde{R})$. Both I and $\tilde{R}$ are compact sets (closed and bounded subsets of a finite dimensional normed space), so a sequence in $C^2(I,\tilde{R})$ is a sequence in $\mathcal{C}(I,\tilde{R})$, in which we work with the supremum norm induced by $C^0(I,\mathbb{R}^N)$. Below, we will change between norms $(||.||, ||.||_{(1)}, ||.||_{(2)})$ to obtain the information we need. The details of the steps in figure 6 now follow.

- Let $\mathcal{C}^{(1)}(I,\tilde{R}) := \{f \in \mathcal{C}(I,\tilde{R}) \mid f' \text{ exists and } f' \in \mathcal{C}(I,\tilde{R})\}$. As an immediate consequence of the definition of the norm $||.||_{(2)}$, for all n in $\mathbb{N}$, $\psi_n \in \mathcal{C}^{(1)}(I,\tilde{R})$. We show that $\mathcal{C}^{(1)}(I,\tilde{R})$ is equicontinuous. Using the Mean Value Theorem [5, Theorem 3.2-3] and the fact that for $f \in \mathcal{C}^{(1)}(I,\tilde{R})$, we know that $||f'|| \leq 1$, we obtain

$$\begin{aligned} \forall f \in &\mathcal{C}^{(1)}(I,\tilde{R}) \quad \forall t,t' \in I \text{ with } t < t' : \\ &||f(t') - f(t)||_\infty \leq 1 \cdot |t' - t|. \end{aligned}$$

  Hence, for all $\epsilon > 0$, letting $t, t' \in I$ with $|t - t'| < \epsilon$ ensures for all $f \in \mathcal{C}^{(1)}(I,\tilde{R})$ that $||f(t') - f(t)||_\infty < \epsilon$ and the set $\mathcal{C}^{(1)}(I,\tilde{R})$ is thus equicontinuous.

- To apply the Arzelà-Ascoli Theorem 4.5.4, we need to look at the closure $\bar{\mathcal{C}}^{(1)}(I,\tilde{R})$ of $\mathcal{C}^{(1)}(I,\tilde{R})$. Since $\mathcal{C}^{(1)}(I,\tilde{R})$ is equicontinuous, so is $\bar{\mathcal{C}}^{(1)}(I,\tilde{R})$ by corollary 4.5.3.

- In summary, I, $\tilde{R}$ are compact metric spaces and $\bar{\mathcal{C}}^{(1)}(I,\tilde{R})$ is a closed and equicontinuous subset of $\mathcal{C}(I,\tilde{R})$. Hence it is compact and also sequentially compact, so there exists a converging subsequence $(\psi_{n_m})_{m\in\mathbb{N}}$ of $(\psi_n)_{n\in\mathbb{N}}$ with limit $\psi \in \bar{\mathcal{C}}^{(1)}(I,\tilde{R})$. This convergence is in the supremum norm. However, we do not yet know whether $\psi \in \mathcal{C}^{(1)}(I,\tilde{R})$, i.e., whether ψ is also continuously differentiable.

- We now look at the sequence $\left(\psi'_{n_m}\right)_{m\in\mathbb{N}}$ of derivatives of ψ_{n_m}. Because $\forall m \in \mathbb{N} : \psi_{n_m} \in \bar{B}^{(2)}(0,1)$ (which ensures uniform boundedness by 1 of the second derivatives), it now also holds that $\forall m \in \mathbb{N} : \psi'_{n_m} \in \mathcal{C}^{(1)}(I,\tilde{R})$. Applying the same argumentation as above, there exists a converging subsequence $\left(\psi'_{n_{m_l}}\right)_{l\in\mathbb{N}}$ of $\left(\psi'_{n_m}\right)_{m\in\mathbb{N}}$ with limit $\Psi \in \bar{\mathcal{C}}^{(1)}(I,\tilde{R}) \subset \mathcal{C}(I,\tilde{R})$. So this limit is continuous. The convergence is again in the supremum norm. Clearly, $\left(\psi_{n_{m_l}}\right)_{l\in\mathbb{N}}$ still converges to ψ.

- For the sequences $\left(\psi_{n_{m_l}}\right)_{l\in\mathbb{N}}$ and $\left(\psi'_{n_{m_l}}\right)_{l\in\mathbb{N}}$, all conditions of Theorem 4.2.3 are satisfied, since convergence in the supremum norm is (by definition) uniform convergence.[12] Hence it follows that ψ is differentiable on I and that $\psi' = \lim_{l\to\infty} \psi'_{n_{m_l}} = \Psi$. Indeed, since Ψ is continuous, we have established that ψ is continuously differentiable.

- Since $\psi_{n_{m_l}} \to \psi$ and $\psi'_{n_{m_l}} \to \psi'$ uniformly as $l \to \infty$, we see that $\psi_{n_{m_l}} \to \psi$ in the $||.||_{(1)}$ norm as well. Hence we can say that any sequence in $\bar{B}^{(2)}(0,1)$ contains a subsequence that converges in $C^{(1)}(I,\mathbb{R}^N)$.

So there is a subsequence $(\psi_{n_l})_{l\in\mathbb{N}}$ that converges in $C^{(1)}(I,\mathbb{R}^N)$. We note from (5.18) that $\mathcal{D}$ is also a bounded linear operator from $C^{(1)}(I,\mathbb{R}^N)$ to Y with the $||.||_{(1)}$-norm on the former, hence it is continuous as an operator between these spaces. Then we know that $\mathcal{N}$ is continuous from $C^{(1)}(I,\mathbb{R}^N)$ to Y. So because $(\psi_{n_l})_{l\in\mathbb{N}}$ is a subsequence that converges in $C^{(1)}(I,\mathbb{R}^N)$, $\mathcal{N}$ provides a subsequence $(\mathcal{N}(\psi_{n_l}))_{l\in\mathbb{N}}$ of $(\phi_n)_{n\in\mathbb{N}}$ that converges in Y (due to continuity of $\mathcal{N}$, it converges to $\mathcal{N}(\psi)$. So $\mathcal{N}(\bar{B}^{(2)}(0,1))$ is a sequentially compact subset of the metric space Y, and therefore a compact subset. Its closure is thus equal to the set itself and still compact. Hence, $\mathcal{N}$ is a compact operator from X to Y. □

Lemma 5.3.5. [28, Lemma 3] Let $m \in C^{(k)}\left(\mathbb{R}^N, \mathbb{R}^{N\times N}\right)$ $(1 \leq k \leq \infty)$ and $\psi \in X$. Then $\mathcal{A}'(\psi)$ is injective.
Proof:
Let $m \in C^{(k)}\left(\mathbb{R}^N, \mathbb{R}^{N\times N}\right)$ $(1 \leq k \leq \infty)$ and $\psi \in X$. To show injectivity of $\mathcal{A}'(\psi)$, let $h \in X$. Since $\mathcal{A}'(\psi)$ is a linear operator from X to Y, we need to show that $\mathcal{A}(\psi)h = 0$ implies $h = 0$.
We have previously defined $\mathcal{A} = \mathcal{L} + \mathcal{N}$. Now $\mathcal{L}$ is a bounded linear operator, such that in view of observation 4.6.3, we have $\mathcal{L}'(\psi) = \mathcal{L}$. We have already shown that the derivative $\mathcal{N}'(\psi)$ is given by (5.17). Since derivatives add, we see that $\mathcal{A}(\psi)h = 0$ means that

$$\begin{pmatrix} \lambda\ddot{h} + \dot{h} + h + (m'(\psi)h) * \psi + m(\psi) * \dot{h} \\ h(0) \\ \dot{h}(0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

[11] $||.||$ denotes the Euclidean norm on $\mathbb{R}^N$ here.
[12] All ψ_n are vector-valued functions, so we actually apply Theorem 4.2.3 to each of the N component functions and use that uniform convergence in the Euclidean norm is equivalent to uniform convergence of each of the component functions.

so in particular $h(0) = \dot{h}(0) = 0$.
Our further investigation concerns for $s \in [0, T]$ the equation

$$\lambda \ddot{h}(s) + \dot{h}(s) + h(s) + \int_0^s (m'(\psi)h(s-r))\,\psi(r)\,\mathrm{d}r + \int_0^s m(\psi(s-r))\dot{h}(r)\,\mathrm{d}r = 0. \tag{5.19}$$

Taking the inner product with $\dot{h}(s)$ (since a real inner product is symmetric in its two arguments, we can take the inner product left or right for each term as suits best) and integration over s from 0 to t yields with (4.9):

$$\begin{aligned}
\int_0^t \lambda \ddot{h}(s) \cdot \dot{h}(s)\,\mathrm{d}s + \int_0^t \dot{h}(s) \cdot \dot{h}(s)\,\mathrm{d}s + \int_0^t \dot{h}(s) \cdot h(s)\,\mathrm{d}s &= -\int_0^t \dot{h}(s) \cdot \int_0^s (m'(\psi)h(s-r))\,\psi(r)\,\mathrm{d}r\,\mathrm{d}s \\
&\quad - \int_0^t \dot{h}(s) \cdot \int_0^s m(\psi(s-r))\dot{h}(r)\,\mathrm{d}r\,\mathrm{d}s, \\
\int_0^t \frac{1}{2}\lambda \frac{d}{ds}\dot{h} \cdot \dot{h}(s)\,\mathrm{d}s + \int_0^t \dot{h}(s) \cdot \dot{h}(s)\,\mathrm{d}s + \int_0^t \frac{1}{2}\frac{d}{ds}h \cdot h(s)\,\mathrm{d}s &= -\int_0^t \int_0^s (m'(\psi)h(s-r))\,\psi(r) \cdot \dot{h}(s)\,\mathrm{d}r\,\mathrm{d}s \\
&\quad - \int_0^t \int_0^s m(\psi(s-r))\dot{h}(r) \cdot \dot{h}(s)\,\mathrm{d}r\,\mathrm{d}s.
\end{aligned}$$

We will now write $h(s) \cdot h(s) = h(s)^2$ and use that we already found $h(0) = \dot{h}(0) = 0$:

$$\begin{aligned}
\frac{1}{2}\lambda \dot{h}(t)^2 + \int_0^t \dot{h}(s)^2\,\mathrm{d}s + \frac{1}{2}h(t)^2 &\le \left| \int_0^t \int_0^s (m'(\psi)h(s-r))\,\psi(r) \cdot \dot{h}(s)\,\mathrm{d}r\,\mathrm{d}s \right| \\
&\quad + \left| \int_0^t \int_0^s m(\psi(s-r))\dot{h}(r) \cdot \dot{h}(s)\,\mathrm{d}r\,\mathrm{d}s \right|, \\
&\le \int_0^t \int_0^s \left| (m'(\psi)h(s-r))\,\psi(r) \cdot \dot{h}(s) \right| \mathrm{d}r\,\mathrm{d}s \\
&\quad + \int_0^t \int_0^s \left| m(\psi(s-r))\dot{h}(r) \cdot \dot{h}(s) \right| \mathrm{d}r\,\mathrm{d}s, \\
&\le \int_0^t \int_0^s ||m'(\psi)|| \cdot ||h(s-r)|| \cdot ||\psi(r)|| \cdot ||\dot{h}(s)||\,\mathrm{d}r\,\mathrm{d}s \\
&\quad + \int_0^t \int_0^s ||m(\psi(s-r))|| \cdot ||\dot{h}(r)|| \cdot ||\dot{h}(s)||\,\mathrm{d}r\,\mathrm{d}s.
\end{aligned}$$

We used the Cauchy-Schwarz inequality (4.8) and remark 4.4.3 for the last inequality.
Since m is continuous and ψ is continuous, the image of the compact interval $[0, T]$ under $m \circ \psi : [0, T] \to \mathbb{R}^{N \times N}$ is bounded, such that a constant $c_1 > 0$ exists with $\forall \tau \in [0, T] : ||m(\psi(\tau))|| \le c_1$. Also due to continuity of ψ, a constant $c_2 > 0$ exists such that $\forall \tau \in [0, T] : ||\psi(\tau)|| \le c_2$. Since m' denotes differentiation of m in the operator sense, $m'(\psi)$ is by definition a bounded linear operator for fixed ψ, so $||m'(\psi)||$ must be an existing constant, fundamentally independent of any variable in the integration. Now let $c := \max\{1, c_1, c_2, ||m'(\psi)||\}$. Then

$$\begin{aligned}
\frac{1}{2}\lambda \dot{h}(t)^2 + \frac{1}{2}h(t)^2 &\le -\int_0^t \dot{h}(s)^2\,\mathrm{d}s + c^2 \left(\int_0^t \int_0^s ||h(s-r)|| \cdot ||\dot{h}(s)||\,\mathrm{d}r\,\mathrm{d}s + \int_0^t \int_0^s ||\dot{h}(r)|| \cdot ||\dot{h}(s)||\,\mathrm{d}r\,\mathrm{d}s \right), \\
&\le -\int_0^t \dot{h}(s)^2\,\mathrm{d}s + c^2 \left(\int_0^t \int_0^s ||h(r)|| \cdot ||\dot{h}(s)||\,\mathrm{d}r\,\mathrm{d}s + \int_0^t \int_0^s ||\dot{h}(r)|| \cdot ||\dot{h}(s)||\,\mathrm{d}r\,\mathrm{d}s \right).
\end{aligned}$$

Inequality (4.1) can be applied to the above integrals to give

$$\frac{1}{2}\lambda\dot{h}(t)^2+\frac{1}{2}h(t)^2\leq-\int_0^t\dot{h}(s)^2\,\mathrm{d}s+\frac{1}{2}c^2\left(\int_0^t\int_0^s||h(r)||^2+||\dot{h}(s)||^2\,\mathrm{d}r\,\mathrm{d}s+\int_0^t\int_0^s||\dot{h}(r)||^2+||\dot{h}(s)||^2\,\mathrm{d}r\,\mathrm{d}s\right).$$

Now applying (4.2) leads to

$$\begin{aligned}\frac{1}{2}\lambda\dot{h}(t)^2+\frac{1}{2}h(t)^2&\leq-\int_0^t\dot{h}(s)^2\,\mathrm{d}s\\&\quad+\frac{1}{2}c^2T\left(\int_0^t||h(s)||^2\,\mathrm{d}s+\int_0^t||\dot{h}(s)||^2\,\mathrm{d}s+\int_0^t||\dot{h}(s)||^2\,\mathrm{d}s+\int_0^t||\dot{h}(s)||^2\,\mathrm{d}s\right),\qquad(5.20)\\&\leq-\int_0^t\dot{h}(s)^2\,\mathrm{d}s+3c^2T\max\left\{\frac{1}{\lambda},1\right\}\left(\int_0^t\frac{1}{2}\lambda\dot{h}(s)^2\,\mathrm{d}s+\int_0^t\frac{1}{2}h(s)^2\,\mathrm{d}s\right).\end{aligned}$$

We can now use Gronwall's inequality, Lemma 4.1.5,, identifying

$$\begin{aligned}u(t)&=\frac{1}{2}\lambda\dot{h}(t)^2+\frac{1}{2}h(t)^2,\\\alpha(t)&=-\int_0^t\dot{h}(s)^2,\\\beta(t)&=3c^2T\max\left\{\frac{1}{\lambda},1\right\},\end{aligned}$$

such that (4.5) is satisfied. With these functions for α and β, since α is a non-positive function, (4.6) yields

$$u(t)=\frac{1}{2}\lambda\dot{h}(t)^2+\frac{1}{2}h(t)^2\leq 0.$$

Since this is the sum of two inner products of a vector with itself, which are by definition non-negative, we can conclude that $h=\dot{h}=0$, i.e., $\mathcal{A}$ is injective. □

We now turn to the last lemma needed to prove Theorem 5.3.1 under condition (5.15).

Lemma 5.3.6. [29, Lemma 4] Let $m\in C^k\left(\mathbb{R}^N,\mathbb{R}^{N\times N}\right)$ $(1\leq k\leq\infty)$ and assume that condition (5.15) holds. Then $\mathcal{A}$ is weakly coercive.

Proof:

For $B>0$ fixed, let $Z=\{\psi\in X\mid||\mathcal{A}(\psi)||_Y\leq B\}$, i.e. a subset of X with functions whose images under $\mathcal{A}$ are bounded in Y by a constant B. Without loss of generality, we pick $B\geq 1$, because sets that are bounded by a smaller constant are also bounded by 1. In view of observation 4.8.2, we can show that the image of an unbounded set is unbounded. To prove this statement, it suffices to show that the fact that B is bounded imposes a bound on $||\psi||_{(2)}$ for $\psi\in Z$, i.e. a constant that bounds $||\psi(t)||$, $||\dot{\psi}(t)||$ and $\ddot{\psi}(t)$ for $t\in[0,T]$. This will be a constant in the sense that it depends on B, λ, T, m and some arbitrarily chosen $\epsilon>0$. We will critically check that this constant does not depend on any specific $\psi\in Z$.

Let $\psi\in Z$ and define

$$\begin{pmatrix}f\\\psi_0\\\psi_1\end{pmatrix}:=\mathcal{A}(\psi)=\begin{pmatrix}\lambda\ddot{\psi}+\dot{\psi}+\psi+m(\psi)*\dot{\psi}\\\psi(0)\\\dot{\psi}(0)\end{pmatrix}.\qquad(5.21)$$

We will find an upper bound for ψ and $\dot{\psi}$ by iteratively bounding them on successive intervals.

Let $\epsilon>0$ arbitrary. Because ψ and m are continuous, so is their composition $m\circ\psi$ and this guarantees the existence of some $T'>0$ with $|m(\psi(t))|\leq|m(\psi_0)|+\epsilon$ for $t\in[0,T']$. Now, because $\psi\in Z$, it holds that $||\psi_0||\leq||\mathcal{A}(\psi)||_Y\leq B$ in the Euclidean norm for ψ_0. Again with continuity of m, it maps compact sets (in $\mathbb{R}^N$, where compactness is equivalent to being closed and bounded) to compact sets (in $\mathbb{R}^{N\times N}$), so a constant $\kappa>0$ exists such that for all $\phi\in Z:||m(\phi(0))||\leq\kappa$. In conclusion, for $t\in[0,T']$, we have $||m(\psi(t))||\leq\kappa+\epsilon=:k_1$, where k_1 depends only on some chosen ϵ, B and m.

Let $T_0:=\min\{T',\frac{1}{2k_1},T\}$.[13] Note that T_0 does depend on ψ because T' does. However, $T_0\leq\frac{1}{2k_1}$ and $T_0<T$,

[13] We wish to ensure that $T_0\leq T$ to be able to integrate up to T_0. However, in the following argument, we will mostly assume that $T_0<T$ and even $2T_0<T$.

so it has an upper bound that does not depend on the specific function $\psi \in Z$. Now for $t \in [0,T]$, starting from $f(s) = \lambda\ddot{\psi}(s) + \dot{\psi}(s) + \psi(s) + (m(\psi) * \dot{\psi})(s)$ and taking the inner product with $\dot{\psi}(s)$, followed by integration over s from 0 to t results in (making use of great similarities with the proof of Lemma 5.3.5):

$$\frac{1}{2}\lambda\left(\dot{\psi}(t)^2 - \psi_1^2\right) + \int_0^t \dot{\psi}(s)^2\,\mathrm{d}s + \frac{1}{2}\left(\psi(t)^2 - \psi_0\right) + \int_0^t \int_0^s m(\psi(s-r))\dot{\psi(r)}\cdot\dot{\psi}(s)\,\mathrm{d}r\,\mathrm{d}s =$$

$$\int_0^t f(s)\cdot\psi(s)\,\mathrm{d}s \leq \int_0^t \left|f(s)\cdot\dot{\psi}(s)\right|\mathrm{d}s \leq \int_0^t ||f(s)||\cdot||\dot{\psi}(s)||\,\mathrm{d}s \leq \frac{1}{2}\int_0^t f(s)^2 + \dot{\psi}(s)^2\,\mathrm{d}s \Rightarrow$$

$$\psi(t)^2 + \lambda\dot{\psi}(t)^2 \leq {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^t f(s)^2\,\mathrm{d}s - \int_0^t \dot{\psi}(s)^2\,\mathrm{d}s + 2\int_0^t\int_0^s \left|m\left(\psi(s-r)\right)\dot{\psi}(r)\dot{\psi}(s)\right|\mathrm{d}r\,\mathrm{d}s. \tag{5.22}$$

We used Cauchy-Schwarz and again inequality (4.1). Continuing with $t \in [0,T_0]$, we can make the following estimates:

$$\begin{aligned}
\psi(t)^2 + \lambda\dot{\psi}(t)^2 &\leq {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^t f(s)^2\,\mathrm{d}s - \int_0^t \dot{\psi}(s)^2\,\mathrm{d}s + 2k_1\int_0^t\int_0^s ||\dot{\psi}(s)||\cdot||\dot{\psi}(r)||\,\mathrm{d}r\,\mathrm{d}s,\\
&\overset{(4.3)}{\leq} {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^t f(s)^2\,\mathrm{d}s - \int_0^t \dot{\psi}(s)^2\,\mathrm{d}s + 2k_1T_0\int_0^t \dot{\psi}(s)^2\,\mathrm{d}s,\\
&= {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^t f(s)^2\,\mathrm{d}s + (2k_1T_0 - 1)\int_0^t \dot{\psi}(s)^2\,\mathrm{d}s.
\end{aligned}$$

Because $T_0 \leq 1/(2k_1)$, we have that $2k_1T_0 - 1 \leq 0$, so we have

$$\begin{aligned}
\psi(t)^2 + \lambda\dot{\psi}(t)^2 &\leq {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^t f(s)^2\,\mathrm{d}s,\\
&\leq {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^{T_0} f(s)^2\,\mathrm{d}s,\\
&\leq {\psi_0}^2 + \lambda{\psi_1}^2 + \int_0^{T} f(s)^2\,\mathrm{d}s,\\
\psi(t)^2 + \dot{\psi}(t)^2 &\leq \max\left\{1, \frac{1}{\lambda}\right\}\left(\psi(t)^2 + \lambda\dot{\psi}(t)^2\right),\\
&\leq \max\left\{1, \frac{1}{\lambda}\right\}\left(B^2 + \lambda B^2 + \int_0^T B^2\,\mathrm{d}s\right) =: {c_1}^2
\end{aligned}$$

Thus we have identified an upper bound c_1 for both $\psi(t)$ and $\dot{\psi}(t)$ for $t \in [0,T_0]$. This bound only depends on B, T and λ and is thus valid for all $\psi \in B$.
If $T_0 = T$, we can proceed to showing that also $\ddot{\psi}(t)$ is bounded here. However, we proceed first for the general case that $T_0 < T$. For ease of notation, we will also assume that $2T_0 < T$. Should this not be the case, the same bounds hold, but one should be aware that the integrals can only be taken up to T, strictly speaking.
For $t \in [T_0, 2T_0]$, we will first examine the integral $\int_0^t\int_0^s \left|m(\psi(s-r)) * \dot{\psi(r)}\cdot\psi(s)\right|\mathrm{d}r\,\mathrm{d}s$. We will split the integral over different regions. To see how, let us symbolically write

$$\int_0^t\int_0^s = \int_0^{T_0}\int_0^s + \int_{T_0}^t\int_0^s = \int_0^{T_0}\int_0^s + \int_{T_0}^t\int_0^{T_0} + \int_{T_0}^t\int_{T_0}^s.$$

This allows us to also use the upper bounds previously identified on the interval $[0, T_0]$. We find

$$\begin{aligned}
&\int_0^t \int_0^s \left| m(\psi(s-r)) * \dot{\psi}(r) \cdot \dot{\psi}(s) \right| \mathrm{d}r \, \mathrm{d}s, \\
&\quad \leq \int_0^{T_0} \int_0^s ||m(\psi(s-r))|| \cdot ||\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s + \int_{T_0}^t \int_0^s \left| m(\psi(s-r))\dot{\psi}(r) \cdot \dot{\psi}(s) \right| \mathrm{d}r \, \mathrm{d}s, \\
&\quad \leq {k_1 c_1}^2 \int_0^{T_0} \int_0^s 1 \, \mathrm{d}r \, \mathrm{d}s + \int_{T_0}^t \int_0^{T_0} ||m(\psi(s-r))\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s + \\
&\qquad \int_{T_0}^t \int_{T_0}^s ||m(\psi(s-r))|| \cdot ||\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s.
\end{aligned} \tag{5.23}$$

We will consider the three integrals that are left separately. The first integral can be explicitly computed:

$$k_1 {c_1}^2 \int_0^{T_0} \int_0^s 1 \, \mathrm{d}r \, \mathrm{d}s = \frac{1}{2} k_1 {c_1}^2 {T_0}^2.$$

For the second integral, using condition (5.15), we have

$$\int_{T_0}^t \int_0^{T_0} ||m(\psi(s-r))\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s \leq \int_{T_0}^t \int_0^{T_0} c(1 + ||\psi(s-r)||)||\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s.$$

In this integral, $r \in [0, T_0]$, so we can bound $||\dot{\psi}(r)||$. Further, we have for all $a, b \in \mathbb{R} : (1+a)b = 1 \cdot b + ab \leq \frac{1}{2}(1^2 + b^2 + a^2 + b^2) < b^2 + 1 + a^2$. We also need inequality 4.1.4. This leads us to

$$\begin{aligned}
\int_{T_0}^t \int_0^{T_0} ||m(\psi(s-r))\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s &\leq cc_1 \int_{T_0}^t \int_0^{T_0} \dot{\psi}(s)^2 \, \mathrm{d}r \, \mathrm{d}s + cc_1 \int_{T_0}^t \int_0^{T_0} 1 + \psi(s-r)^2 \, \mathrm{d}s \, \mathrm{d}r, \\
&\overset{(4.4)}{\leq} cc_1(t - T_0)T_0 + cc_1 T_0 \int_{T_0}^t \dot{\psi}(s)^2 \, \mathrm{d}s + cc_1 T_0 \int_0^t \psi(s)^2 \, \mathrm{d}s, \\
&\leq cc_1 {T_0}^2 + cc_1 T_0 \int_{T_0}^t \dot{\psi}(s)^2 \, \mathrm{d}s + cc_1 T_0 \int_0^t \psi(s)^2 \, \mathrm{d}s.
\end{aligned} \tag{5.24}$$

For the third integral, we notice that $s \in [T_0, 2T_0]$, $r \in [T_0, 2T_0]$, so $s - r \in [0, T_0]$ and $s - T_0 \leq T_0$. We then obtain

$$\begin{aligned}
\int_{T_0}^t \int_{T_0}^s ||m(\psi(s-r))|| \cdot ||\dot{\psi}(r)|| \cdot ||\dot{\psi}(s)|| \, \mathrm{d}r \, \mathrm{d}s &\leq \frac{1}{2} k_1 \int_{T_0}^t \int_{T_0}^s \dot{\psi}(r)^2 + \dot{\psi}(s)^2 \, \mathrm{d}r \, \mathrm{d}s, \\
&\overset{(4.2)}{\leq} \frac{1}{2} k_1 T_0 \int_{T_0}^t \dot{\psi}(s)^2 \, \mathrm{d}s + \frac{1}{2} k_1 \int_{T_0}^t (s - T_0)\dot{\psi}(s)^2 \, \mathrm{d}s, \\
&\leq k_1 T_0 \int_{T_0}^t \dot{\psi}(s)^2 \, \mathrm{d}s.
\end{aligned} \tag{5.25}$$

Combining these inequalities with (5.23) gives

$$\begin{aligned}
\int_0^t \int_0^s \left| m(\psi(s-r)) * \dot{\psi}(r) \cdot \dot{\psi}(s) \right| \mathrm{d}r \, \mathrm{d}s &\leq \frac{1}{2} k_1 {c_1}^2 {T_0}^2 + cc_1 {T_0}^2 + cc_1 T_0 \int_0^t \psi(s)^2 \, \mathrm{d}s + (k_1 + cc_1) T_0 \int_{T_0}^t \dot{\psi}(s)^2 \, \mathrm{d}s, \\
&\leq \frac{1}{2} k_1 {c_1}^2 {T_0}^2 + cc_1 {T_0}^2 + (k_1 + cc_1) T_0 \max\left\{1, \frac{1}{\lambda}\right\} \int_0^t \psi(s)^2 + \lambda \dot{\psi}(s)^2 \, \mathrm{d}s.
\end{aligned}$$

Putting this in (5.22), which was obtained for all $t \in [0,T]$, we find[14]

$$\begin{aligned}\psi(t)^2 + \lambda\dot{\psi}(t)^2 \leq & {\psi_0}^2 + \lambda {\psi_1}^2 + \int_0^T f(s)^2 \,\mathrm{d}s + k_1 {c_1}^2 {T_0}^2 + 2cc_1 {T_0}^2 \\ & + 2\left(k_1 + cc_1\right) T_0 \max\left\{1, \frac{1}{\lambda}\right\} \int_0^t \psi(s)^2 + \lambda\dot{\psi}(s)^2 \,\mathrm{d}s, \\ & \leq \alpha_2 + \beta_2 \int_0^t \psi(s)^2 + \lambda\dot{\psi}(s)^2 \,\mathrm{d}s,\end{aligned} \tag{5.26}$$

with

$$\begin{aligned}\alpha_2 &= B^2(1+\lambda+T) + T^2 c_1(k_1 c_1 + 2c), \\ \beta_2 &= 2(k_1 + cc_1)T \max\left\{1, \frac{1}{\lambda}\right\}.\end{aligned}$$

These constants depend only on B, ϵ, λ, T and the nature of m. Furthermore, they are certainly positive. Hence Gronwall's inequality (4.7) can be applied to give, for all $t \in [0, 2T_0]$:

$$\begin{aligned}\psi(t)^2 + \dot{\psi}(t)^2 &\leq \max\left\{1, \frac{1}{\lambda}\right\}\left(\psi(t)^2 + \lambda\dot{\psi}(t)^2\right), \\ &\leq \max\left\{1, \frac{1}{\lambda}\right\} \alpha_2 \exp(\beta_2 t), \\ &\leq \max\left\{1, \frac{1}{\lambda}\right\} \alpha_2 \exp(\beta_2 T) =: {c_2}^2.\end{aligned}$$

So we have found a bound c_2 for $\psi(t)$ and $\dot{\psi}(t)$ on $t \in [0, 2T_0]$. This bound holds for all $\psi \in Z$ because it only depends on B, ϵ, λ, T and the nature of m. This bound on $\psi(t)$ for all $t \in [0, 2T_0]$ again gives a constant k_2 such that for all $t \in [0, 2T_0]$: $m(\psi(t)) \leq k_2$ due to continuity of m.
These are the ingredients needed to find a bound on $\psi(t)$ for $t \in [0, 4T_0]$. We can follow the same procedure as above, now splitting the integral for $t \in [2T_0, 4T_0]$ as

$$\int_0^t \int_0^s = \int_0^{2T_0} \int_0^s + \int_{2T_0}^t \int_0^s = \int_0^{2T_0} \int_0^s + \int_{2T_0}^t \int_0^{2T_0} + \int_{2T_0}^t \int_{2T_0}^s .$$

Inequalities analogous to (5.23), (5.24) and (5.25) can be obtained (with the appropriate bounds) with the use of k_2 and c_2 instead of k_1 and c_1, respectively. In inequality (5.26), we could add the constant $(1+\lambda){c_2}^2$ to ensure that the inequality also holds for $t \in [0, 2T_0]$. Then Gronwall's inequality can again be invoked to bound $\psi(t)$ and $\dot{\psi}(t)$ by a constant c_3 for all $t \in [0, 4T_0]$. This constant can be found such that it only depends on B, ϵ, λ, T and m, so the estimate is again valid for all $\psi \in Z$. We find a new bound k_3 on $m(\phi(t))$ for all $\phi \in Z$ and all $t \in [0, 4T_0]$ and continue to find a bound on $t \in [0, 8T_0]$, etcetera.
This results in an upper bound K, say, for $t \in [0,T]$ on $\psi(t)$ and $\dot{\psi}(t)$ for all $\psi \in Z$, where the bound K still only depends on B, ϵ, λ, T and m. Consequently, by continuity of m, $||psi(t)|| \leq K$ also ensures the some $L > 0$ exists with $m(\psi(t)) \leq L$ for $t \in [0,T]$.
Now we have to bound the second derivate to obtain a bound of ψ in the sense of the $||.||_{(2)}$ norm. This is now immediately derived from (5.21), since for all $t \in [0,T]$, we have:

$$\begin{aligned}\lambda||\ddot{\psi}(t)|| &\leq ||f(t)|| + ||\dot{\psi}(t)|| + ||\psi(t)|| + ||m(\psi) * \dot{\psi}(t)||, \\ &\leq B + 2K + TLK.\end{aligned}$$

So, if the image under $\mathcal{A}$ of a set of functions in X is bounded, so is the set of functions itself. Then observation 4.8.2 ensures that $\mathcal{A}$ is weakly coercive. □

Now we have shown with Lemmas 5.3.4, 5.3.5 and 5.3.6, together with Theorem 5.3.2, that $\mathcal{A}$ is a C^k-diffeomorphism from X onto Y for any $T > 0$, provided that m satisfies condition (5.15).

[14] The newly found inequalities cannot be used on $[0, T_0]$, since they were explicitly derived for $t \in [0, 2T_0]$. However, we see that the first three terms on the next line already provide a bound on $\psi(t) + \lambda\dot{\psi}(t)$ there.

Local existence
When the kernel m does not satisfy condition (5.15), a local existence result can be derived from the global existence result if we replace the kernel by a bounded one, which equals m on some restricted domain of $\mathbb{R}^N$. To this end, let $k_1 > ||\Psi_0||$, $k_2 > k_1$ arbitrary but fixed, and define $\tilde{m} \in C^1\left(\mathbb{R}^N, \mathbb{R}^{N\times N}\right)$ by

$$\tilde{m}(x) := \begin{cases} m(x) & ||x|| \leq k_1, \\ 0 & ||x|| \geq k_2, \end{cases}$$

and continuously differentiable extended on $\{x \in \mathbb{R}^N \mid k_1 \leq ||x|| \leq k_2\}$. So by construction $\tilde{m}$ is indeed continuously differentiable and with $D = \{x \in \mathbb{R}^N \mid ||x|| \leq k_2\}$, we can employ Corollary 4.6.8 to conclude that $\tilde{m}$ is bounded on $\mathbb{R}^N$, say $\forall x \in \mathbb{R}^N : ||\tilde{m}(x)|| \leq K$. Hence we can verify condition (5.15) for $\tilde{m}$. Let $x, y \in \mathbb{R}^N$. By remark 4.4.3, we have

$$||\tilde{m}(x)y|| \leq ||\tilde{m}(x)|| \cdot ||y|| \leq K||y|| \leq K(1 + ||x||)||y||,$$

which is indeed (5.15). Consequently, with the kernel $\tilde{m}$, we know that the system (5.14) admits a unique solution ψ on $[0, T]$ for any $T \in \mathbb{R}$. We have ensured that $||\psi(0)|| = ||\Psi_0|| < k_1$ and since the solution ψ is continuous, there exists some $\tau > 0$ with $||\psi(t)|| \leq k_1$ for $t \in [0, \tau]$. Then on $[0, \tau]$, it holds that $m(\psi(t)) = \tilde{m}(\psi(t))$, so ψ is a (local) solution to (5.14) with the kernel m on the interval $[0, \tau]$.

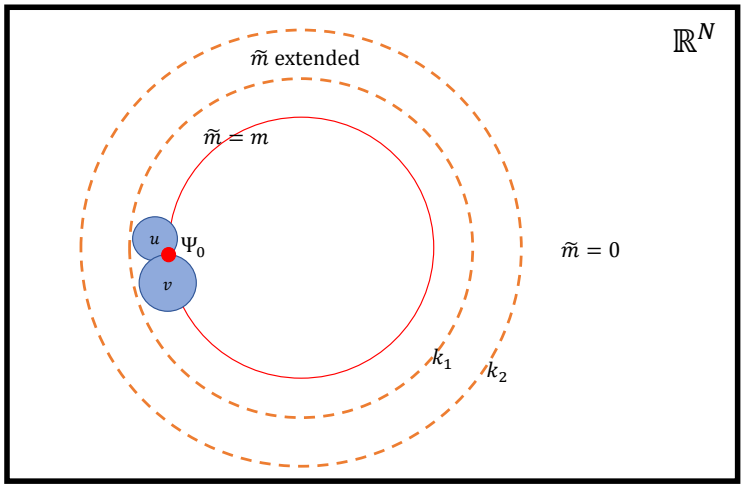

Figure 7: *Illustration of the cut-off kernel $\tilde{m}$. The kernel is left unchanged in some domain around Ψ_0 (up to k_1), such that a solution (u or v in the figure) with $\tilde{m}$ can locally be taken as a solution with the original kernel m.*

To show uniqueness of any local solution, we can weaken the conditions on the kernel m to local Lipschitz continuity such that m is Lipschitz continuous on the closed ball with $||x|| \leq k_1$ in $\mathbb{R}^N$, which contains the local solutions to the problem.

Lemma 5.3.7. (Uniqueness, [29, Lemma 5]) Let $T > 0$ arbitrary and let $m \in C^0(\mathbb{R}^N, \mathbb{R}^{N\times N})$ be locally Lipschitz continuous, in the sense that it is Lipschitz continuous when restricted to any closed ball with arbitrary finite radius[15]. Then for any given $\Psi_0, \Psi_1 \in \mathbb{R}^N$ and $f \in C^0([0, T], \mathbb{R}^N)$ there exists at most one solution $\psi \in C^2([0, T], \mathbb{R}^N)$ of (5.14). Therefore, we demand this property for m restricted to arbitrary closed balls.
Proof:
Suppose two solutions u, v of (5.14) exist. Let $w = u - v$. Then we have $w(0) = 0 = \dot{w}(0)$ since both solutions satisfy the same initial conditions. Further equating the inhomogeneous equation in (5.14) for both solutions u and v yields

$$0 = \lambda\ddot{w}(t) + \dot{w}(t) + w(t) + \int_0^t m(u(t-s))\dot{u}(s) - m(v(t-s))\dot{v}(s)\,\mathrm{d}s.$$

[15] Saal is not clear about what he means with local Lipschitz continuity in [29]. However, the definition that for any point in $x \in \mathbb{R}^N$, there should be an open region U_x of x such that m constrained to U_x is Lipschitz continuous, is insufficient to support the argument for arbitrary solutions of (5.14).

The presence of w here is due to the linearity of $\mathcal{L}$ in $\mathcal{A} = \mathcal{L} + \mathcal{N}$. Taking the inner product with $\dot{w}(t)$, this leads to

$$\frac{\lambda}{2}\frac{d}{dt}(\dot{w}(t))^2 + \frac{1}{2}\frac{d}{dt}(w(t))^2 = -\dot{w}(t)^2 - \dot{w}(t)\int_0^t m(u(t-s))\dot{u}(s) - m(v(t-s))\dot{v}(s)\,\mathrm{d}s.$$

Since u, v and m are continuous (and hence the composition $m \circ u$ as well), there exist constants $c_1, c_2, c_3 > 0$ with $||u||_\infty \leq c_1$, $||v||_\infty \leq c_2$ and $||m \circ u||_\infty|| \leq c_3$, where all supremum norms are taken over $[0,T]$. All bounds hold for the constant $c := \max\{c_1, c_2, c_3\}$. In terms of figure 7, we can interpret c as k_1 for some cut-off kernel if taken larger than $||\Psi_0||$ such that local existence of a solution is obtained on $[0,T]$.
For the integrand, we can find the following estimates:

$$\begin{aligned}
||m(u(t-s))\dot{u}(s) - m(v(t-s))\dot{v}(s)|| &\leq ||m(u(t-s))\dot{u}(s) - m(u(t-s)\dot{v}(s) - m(u(t-s)\dot{v}(s) - m(v(t-s))\dot{v}(s)||,\\
&\leq ||m(u(t-s))\,[\dot{u}(s) - \dot{v}(s)]\,|| + ||\,[m(u(t-s) - m(v(t-s))]\,\dot{v}(s)||,\\
&\leq ||m(u(t-s))|| \cdot ||\dot{u}(s) - \dot{v}(s)|| + ||m(u(t-s) - m(v(t-s))|| \cdot ||\dot{v}(s)||,\\
&\leq c||\dot{w}(t-s)|| + L||u(t-s) - v(t-s)||c,\\
&= c||\dot{w}(t-s)|| + cL||w(t-s)||,
\end{aligned}$$

using a suitable estimate for matrix multiplication in view of remark 4.4.3 and the fact that the local Lipschitz continuity of m allows us to write for some constant $L > 0$ that $||m(u(t)) - m(v(t))|| \leq L||u(t) - v(t)||$ for all $t \in [0,T]$, as u and v lie in the compact set $\{f \in C^2([0,T], \mathbb{R}^N) \mid ||f||_\infty \leq c\}$.
Now, since $\int_0^t f(t-s)\,\mathrm{d}s = \int_0^t f(s)\,\mathrm{d}s$ for any integrable function f, integrating over t from 0 to, say $r \in [0,T]$, we obtain with the triangle inequality for integrals (recall the initial values $w(0)$ and $\dot{w}(0)$ we found for w) and the Cauchy-Schwarz inequality:

$$\begin{aligned}
\frac{\lambda}{2}\dot{w}(r)^2 + \frac{1}{2}w(r)^2 &\leq -\int_0^r \dot{w}(t)^2\,\mathrm{d}t + \left|\int_0^r\int_0^t \dot{w}(t)m(u(t-s))\dot{u}(s) - m(v(t-s))\dot{v}(s)\,\mathrm{d}s\,\mathrm{d}t\right|,\\
&\leq \int_0^r ||\dot{w}(t)|| \cdot \int_0^t ||m(u(t-s))\dot{u}(s) - m(v(t-s))\dot{v}(s)||\,\mathrm{d}s\,\mathrm{d}t,\\
&\leq \int_0^r ||\dot{w}(t)|| \cdot \int_0^t ||\dot{w}(s)|| + cL||w(s)||\,\mathrm{d}s\,\mathrm{d}t,\\
&\overset{(4.1)}{\leq} \int_0^r\int_0^t \frac{1}{2}(1+cL)\dot{w}(t)^2 + \frac{1}{2}\dot{w}(s)^2 + \frac{1}{2}cLw(s)\,\mathrm{d}s\,\mathrm{d}t,\\
&\overset{(4.2)}{\leq} \int_0^r t\frac{1}{2}(1+cL)\dot{w}(t)^2\,\mathrm{d}t + T\int_0^r \frac{1}{2}\dot{w}(s)^2\,\mathrm{d}s + T\int_0^r \frac{1}{2}cLw(s)^2\,\mathrm{d}s,\\
&\overset{*}{\leq} T\int_0^r \frac{1}{2}(1+cL)\dot{w}(s)^2\,\mathrm{d}s + T\int_0^r \frac{1}{2}\dot{w}(s)^2\,\mathrm{d}s + T\int_0^r \frac{1}{2}cLw(s)^2\,\mathrm{d}s,\\
&\leq \left(\frac{1}{2}L + 1\right)cT\int_0^r \dot{w}(s)^2\,\mathrm{d}s + \frac{1}{2}cLT\int_0^r w(s)^2\,\mathrm{d}s,\\
&\leq \left(\frac{1}{2}L + 1\right)cT\int_0^r \dot{w}(s)^2 + w(s)^2\,\mathrm{d}s,\\
&\leq (1L+2)\,cT\max\left\{1, \frac{1}{\lambda}\right\}\int_0^r \frac{\lambda}{2}\dot{w}(s)^2 + \frac{1}{2}w(s)^2\,\mathrm{d}s.
\end{aligned}$$

At * we could estimate t in the first integral by T since $\dot{w}(t)^2 \geq 0$ for all t between 0 and T. So we conclude with Gronwall's inequality that for all $t \in [0,T]$:

$$\dot{w}(t)^2 + w(t)^2 \leq 0,$$

since in (4.5), we identify $\alpha = 0$. So for all $t \in [0,T] : u(t) = v(t)$ and the solution is indeed unique. □

This concludes the proof of Theorem 5.3.1.

5.4 Second order system II

To extend the class of kernels giving rise to a global solution to (5.14) (i.e., existing for all $T > 0$ on the interval $[0, T]$), Saal uses an iterative approach with a linearised equation. This is in fact reminiscent of the approach of Götze *et al.* in [10] as discussed in section 5.2. We first behold the linearisation in a slightly more general context.

For $l \in \mathbb{N}$, $x_0 \in \mathbb{R}^l$, $\Phi, M \in C^0([0,\infty), \mathbb{R}^{l\times l})$ and $f \in C^0([0,\infty), \mathbb{R}^l)$, let $x \in C^0([0,\infty), \mathbb{R}^l)$ be a solution to

$$x(t) = \Phi(t)x_0 + \Phi(t)\int_0^t \Phi(-s)\int_0^s M(s-r)x(r)\,\mathrm{d}r\,\mathrm{d}s + \Phi(t)\int_0^t \Phi(-s)f(s)\,\mathrm{d}s. \tag{5.27}$$

Reversing the order of integration in the double integral, we rewrite this as

$$x(t) = \Phi(t)\int_0^t \left(\int_r^t \Phi(-s)M(s-r)\,\mathrm{d}s\right)x(r)\,\mathrm{d}r + \Phi(t)\left(x_0 + \int_0^t \Phi(-s)f(s)\,\mathrm{d}s\right),$$

or

$$x(t) = \int_0^t \tilde{M}(t,r)x(r)\,\mathrm{d}r + \tilde{f}(t),$$

where

$$\tilde{M}(t,r) := \Phi(t)\int_r^t \Phi(-s)M(s-r)\,\mathrm{d}s,$$
$$\tilde{f}(t) := \Phi(t)\left(x_0 + \int_0^t \Phi(-s)f(s)\,\mathrm{d}s\right).$$

Due to the assumptions made on Φ, M and f and the continuity of multiplication and the integral in the above definitions, $\tilde{M}$ and $\tilde{f}$ are seen to be continuous functions as well. Here we recognize the Volterra equation of the second kind of Theorem 4.8.4, hence a unique solution $x \in C^0([0,T], \mathbb{R}^l)$ exists to equation (5.27).
If we further assume that $\Phi \in C^1([0,\infty), \mathbb{R}^{l\times l})$ (so it is differentiable), (5.27) can be differentiated and we find

$$\begin{aligned}\dot{x}(t) =& \Phi'(t)\left[x_0 + \int_0^t \Phi(-s)\int_0^s M(s-r)x(r)\,\mathrm{d}r\,\mathrm{d}s + \int_0^t \Phi(-s)f(s)\,\mathrm{d}s\right] \\ &+ \Phi(t)\left[\Phi(-t)\int_0^s M(t-r)x(r)\,\mathrm{d}r + \Phi(-t)f(t)\right],\end{aligned} \tag{5.28}$$

so the solution x is also continuously differentiable: $x \in C^1([0,T], \mathbb{R}^l)$.

We proceed to the next theorem, which bounds the solution of (5.27).

Theorem 5.4.1. [29, Theorem 6] In the notation introduced above, for certain $b, c_0, c_1, k \in \mathbb{R}$ with $c_1 > c_0$, let $\forall y \in \mathbb{R}^l : ||\Phi(t)y|| \le be^{-c_0 t}||y||$ for $t \in [0,\infty)$, $||M(s-r)x(r)|| \le ke^{-c_1(s-r)}||x(r)||$ for $s, r \in [0,\infty)$[16], $s \ge r$ and let $\left|\left|\int_0^t \Phi(-s)f(s)\,\mathrm{d}s\right|\right| \le k_1$ for some $k_1 > 0$ independent of t. Then

$$||x(t)|| \le (||x_0|| + k_1)e^{\left(-c_0 - \frac{kb^2}{c_1 - c_0}\right)t} \tag{5.29}$$

holds for the solution x of (5.27). Especially in the case $c_0 - \frac{k}{c_1 - c_0} > 0$, we have an exponentially decaying solution.
Proof:

[16] In [29], Saal omits the constant b (i.e., $b = 1$). However, in view of observation 4.7.8, we think that it is in general necessary to account for another constant. We show in appendix B that for the matrix we wish to apply the theorem to, that is indeed the case when working with the Euclidean norm and the matrix norm it induces.

Using the assumed estimate for multiplication with $\Phi(t)$ and the triangle inequality - and timely making use of the other estimates - we can start from (5.27) with

$$||x(t)|| \leq ||\Phi(t)x_0|| + \left|\left|\Phi(t)\int_0^t\int_0^s \Phi(-s)M(s-r)x(r)\,\mathrm{d}r\,\mathrm{d}s\right|\right| + \left|\left|\Phi(t)\int_0^t \Phi(-s)f(s)\,\mathrm{d}s\right|\right|,$$
$$\leq be^{-c_0t}||x_0|| + be^{-c_0t}\int_0^t\int_0^s ||\Phi(-s)M(s-r)x(r)||\,\mathrm{d}r\,\mathrm{d}s + be^{-c_0t}\int_0^t ||\Phi(-s)f(s)||\,\mathrm{d}s \Rightarrow$$

$$\frac{1}{b}e^{c_0t}||x(t)|| \leq ||x_0|| + \int_0^t\int_0^s be^{c_0s}||M(s-r)x(r)||\,\mathrm{d}r\,\mathrm{d}s + k_1,$$
$$\leq ||x_0|| + k_1 + bk\int_0^t\int_0^s e^{c_0s}e^{-c1(s-r)}||x(r)||\,\mathrm{d}r\,\mathrm{d}s,$$
$$= ||x_0|| + k_1 + bk\int_0^t\int_0^s e^{c_0r}e^{-(c1-c_0)(s-r)}||x(r)||\,\mathrm{d}r\,\mathrm{d}s,$$
$$= ||x_0|| + k_1 + bk\int_0^t e^{c_0r}||x(r)||\int_r^t e^{-(c1-c_0)(s-r)}\,\mathrm{d}s\,\mathrm{d}r.$$

Now the integral over s can be directly computed:

$$\int_r^t e^{-(c1-c_0)(s-r)}\,\mathrm{d}s = \left[\frac{-1}{c_1-c_0}e^{-(c1-c_0)(s-r)}\right]_{s=r}^{s=t} = \frac{e^{(c_0-c_1)} - e^{(c_0-c_1)(t-r)}}{c_1-c_0}.$$

Since $c_1 > c_0$, the term with the minus sign can be discarded to obtain the bounds

$$\frac{e^{(c_0-c_1)} - e^{(c_0-c_1)(t-r)}}{c_1-c_0} \leq \frac{e^{-(c_1-c_0)}}{c_1-c_0} \leq \frac{e^0}{c_1-c_0} = \frac{1}{c_1-c_0}.$$

Continuing the estimate for $||x(t)||$ with this result gives

$$e^{c_0t}||x(t)|| \leq b\left(||x_0|| + k_1\right) + \int_0^t \frac{kb^2}{c_1-c_0}e^{c_0r}||x(r)||\,\mathrm{d}r.$$

As a result, we can apply Gronwall's inequality (4.7) with

$$\alpha(t) = b\left(||x_0|| + k_1\right),$$
$$\beta(t) = \frac{kb^2}{c_1-c_0},$$

and bringing e^{c_0t} to the other side gives

$$||x(t)|| \leq be^{-c_0t}(||x_0|| + k_1)\exp\left(\int_0^t \frac{kb^2}{c_1-c_0}\right) = b(||x_0|| + k_1)e^{-\left(c_0 - \frac{kb^2}{c_1-c_0}\right)t}.$$

□

First order system

The foregoing result will be applied in the iteration process with a linearised, first-order system. So Saal transforms the system (5.14) to a first-order system by means of the following definitions:

- We look at the variable $x = (x_1, x_2) := (\psi, \dot{\psi}) \in C^1([0,\infty), \mathbb{R}^{2N})$;
- The source function is transformed to $F := \frac{1}{\lambda}(0, f) \in C^0([0,\infty), \mathbb{R}^{2N})$;

- Saal (implicitly) extends the interpretation of matrix multiplication: for a general matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

and a vector $y = (y_1, y_2) \in \mathbb{R}^{2N}$, we define their product $Ay = ((Ay)_1, (Ay)_2) \in \mathbb{R}^{2N}$ by

$$(Ay)_i := \sum_{j=1}^{2} a_{ij} y_j \in \mathbb{R}^N, \quad i = 1, 2;$$

- For our transformation, define

$$A := \begin{pmatrix} 0 & 1 \\ -\frac{1}{\lambda} & -\frac{1}{\lambda} \end{pmatrix};$$

- For the kernel, we keep using the usual matrix multiplication. We define for arbitrary $g = (g_1, g_2) \in C^0([0,\infty), \mathbb{R}^N \times \mathbb{R}^N)$ and $t \in [0,\infty)$:

$$M(g(t)) := \begin{pmatrix} 0 & 0 \\ 0 & m(g_1(t-s)) \end{pmatrix} \in \mathbb{R}^{2N \times 2N},$$

where each block is to be read as a block of size $N \times N$.

The equation for x is now first order and reads for $t \in [0,\infty)$:

$$\begin{cases} \dot{x}_1(t) = \dot{\psi}(t) & = x_2(t) - \frac{1}{\lambda} \int_0^t 0 \cdot x_1(s)\,\mathrm{d}s, \\ \dot{x}_2(t) = \ddot{\psi}(t) & = -\frac{1}{\lambda}\psi(t) - \frac{1}{\lambda}\dot{\psi}(t) - \frac{1}{\lambda} \int_0^t m(\psi(t-s))\dot{\psi}(s)\,\mathrm{d}s + f(t), \\ & = -\frac{1}{\lambda}x_1(t) - \frac{1}{\lambda}x_2(t) - \frac{1}{\lambda} \int_0^t m(x_1(t-s))x_2(s)\,\mathrm{d}s + f(t). \end{cases}$$

and initial value $x(0) = (\Psi_0, \Psi_1) := x_0$. With the notations and definitions above we see that this can compactly be rewritten as

$$\begin{cases} \dot{x}(t) = Ax(t) - \frac{1}{\lambda} \int_0^t M(x(t-s))x(s)\,\mathrm{d}s + F(t), \\ x(0) = x_0. \end{cases} \tag{5.30}$$

The linear part, without the integral and F, explicitly reads

$$\begin{cases} (\dot{x}_1)_1 = (x_2)_1, \\ (\dot{x}_2)_1 = -\frac{1}{\lambda}(x_1)_1 - \frac{1}{\lambda}(x_2)_1, \\ \quad \vdots \\ (\dot{x}_1)_N = (\dot{x}_2)_N, \\ (\dot{x}_2)_N = -\frac{1}{\lambda}(x_1)_N - \frac{1}{\lambda}(x_2)_N, \end{cases}$$

so we see that we effectively have N linear first-order systems each of which can be solved as described in result 4.7.10 to obtain (with the correct initial conditions)

$$\begin{cases} \begin{pmatrix} (x_1)_1 \\ (x_2)_1 \end{pmatrix} = e^{At} \begin{pmatrix} (\Psi_0)_1 \\ (\Psi_1)_1 \end{pmatrix}, \\ \quad \vdots \\ \begin{pmatrix} (x_1)_N \\ (x_2)_N \end{pmatrix} = e^{At} \begin{pmatrix} (\Psi_0)_N \\ (\Psi_1)_N \end{pmatrix}. \end{cases}$$

With the extended interpretation of matrix multiplication above, we can now write the solution to the linear part as

$$x(t) = e^{At}x_0 =: \Phi(t)x_0$$

Variation of constants

With the solution for the linear part of (5.30) at hand, we employ the technique of variation of constants to find a modified equation to solve (5.30). To this end, we suppose the solution x is of the form $x(t) = \Phi(t)y(t)$ for

some function $y \in C^1([0,\infty), \mathbb{R}^{2N})$ with $y(0) = x_0$ to satisfy the initial conditions. Substituting this supposition yields for $t \in [0,\infty)$ (with $\frac{d}{dt}e^{At} = Ae^{At} = A\Phi(t)$):

$$A\Phi(t)y(t) + \Phi(t)\dot{y}(t) = A\Phi(t)y(t) - \frac{1}{\lambda}\int_0^t M(\Phi(t-s)y(t-s))\Phi(s)y(s)\,\mathrm{d}s + F(t).$$

Since the matrix inverse of e^{At} is e^{-At} according to Lemma 4.7.3, multiplying by $\Phi(-t)$ [17] and integrating yields the integral equation

$$y(t) = x_0 - \frac{1}{\lambda}\int_0^t \Phi(-r)\int_0^r M(\Phi(r-s)y(r-s))\Phi(s)y(s)\,\mathrm{d}s\,\mathrm{d}r + \int_0^t \Phi(-r)F(r)\,\mathrm{d}r.$$

With $x(t) = \Phi(t)y(t)$, this leads us for x to the integral equation

$$x(t) = \Phi(t)x_0 - \frac{1}{\lambda}\Phi(t)\int_0^t \Phi(-s)\int_0^s M(x(s-r))x(r)\,\mathrm{d}r\,\mathrm{d}s + \Phi(t)\int_0^t \Phi(-s)F(s)\,\mathrm{d}s. \tag{5.31}$$

We changed the role of the dummies r and s for maximum resemblance to (5.27).

Iteration sequence
Saal now defines a sequence $(x^{(n)})_{n\in\mathbb{N}}$ of functions $x^{(n)} \in C^1([0,\infty), \mathbb{R}^N \times \mathbb{R}^N)$ for $n \in \mathbb{N}$ by

$$x^{(n)}(t) = \Phi(t)x_0 - \frac{1}{\lambda}\Phi(t)\int_0^t \Phi(-s)\int_0^s M(x^{(n-1)}(s-r))x^{(n)}(r)\,\mathrm{d}r\,\mathrm{d}s + \Phi(t)\int_0^t \Phi(-s)F(s)\,\mathrm{d}s \tag{5.32}$$

with $x_0(t) := x_0 e^{-c_1 t}$, where $c_1 > 0$ will be chosen in the proof of Theorem 5.4.2. Due to the discussion at the beginning of this section, we know that these equations yield a unique solution for all $x^{(n)}$ which is continuously differentiable (with its derivative given by (5.28), for f we read F and for M we should read $\frac{1}{\lambda}M \circ x^{(n-1)}$).

Boundedness for Φ
Before we can bound the solution of (5.30), we need to bound $\Phi(t) = e^{At}$ and we do this in the manner of observation 4.7.9. Computing the eigenvalues λ_1, λ_2 of A results in

$$\begin{cases} \lambda_{1,2} = -\frac{1\pm\sqrt{1-4\lambda}}{2\lambda} & \text{if } \lambda \le \frac{1}{4}, \\ \lambda_{1,2} = -\frac{1\pm i\sqrt{4\lambda-1}}{2\lambda} & \text{if } \lambda > \frac{1}{4}. \end{cases}$$

Now using the bound of observation 4.7.9, we thus find that a $b \ge 1$ exists such that for all $y \in \mathbb{R}^{2N}$

$$||\Phi(t)y|| \le be^{-c_0 t}||y||$$

with $c_0 = \mathrm{Re}\left(\frac{1-\sqrt{1-4\lambda}}{2\lambda}\right)$ (where we read $\sqrt{1-4\lambda} = i\sqrt{4\lambda-1}$ when $\lambda > \frac{1}{4}$), since this is the largest (real part) of the two possible eigenvalues. We see that c_0 is positive, and hence for all $t \in [0,\infty)$, we conclude that

$$||\Phi(t)y|| \le be^{-c_0 t}||y||. \tag{5.33}$$

To be precise, we need to check the inequalities in section 4.7 again for our renewed interpretation of multiplication in equation 5.31. This only needs to be done for Lemma 4.7.7, since then observations 4.7.8 and 4.7.9 follow from (ordinary) scalar and matrix multiplications and the property of the norm described in remark 4.4.3. The result of Lemma 4.7.7 can be verified by rewriting the expression at hand with usual matrix

[17] We should actually check that the definition on page 47 is associative, such that multiplication by the inverse really removes $\Phi(t)$. This boils down to associativity of scalar addition and scalar multiplication with vectors as follows: let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ and $y = (y_1, y_2) \in \mathbb{R}^N \times \mathbb{R}^N$. We then have $(B(Ay))_i = \sum_{j=1}^2 b_{ij}(Ay)_j = \sum_{j=1}^2 b_{ij}\left(\sum_{k=1}^2 a_{jk}y_k\right) = \sum_{j=1}^2\sum_{k=1}^2 b_{ij}a_{jk}y_k = \sum_{k=1}^2\left(\sum_{j=1}^2 b_{ij}a_{jk}\right)y_k = \sum_{k=1}^2 (BA)_{ik}y_k = ((BA)y)_i$.

multiplications as follows (on $\mathbb{R}^n \times \mathbb{R}^n$ we use the Euclidean norm):

$$\begin{aligned}
||Ay||^2 =& \left|\left| \begin{pmatrix} a_{11}y_1 + a_{12}y_2 \\ a_{21}y_1 + a_{22}y_2 \end{pmatrix} \right|\right|^2, \\
=& \left|\left| \begin{pmatrix} a_{11}(y_1)_1 + a_{12}(y_2)_1 \\ \vdots \\ a_{11}(y_1)_N + a_{12}(y_2)_N \\ a_{21}(y_1)_1 + a_{22}(y_2)_1 \\ \vdots \\ a_{21}(y_1)_N + a_{22}(y_2)_N \end{pmatrix} \right|\right|^2, \\
=&(a_{11}(y_1)_1 + a_{12}(y_2)_1)^2 + \cdots + (a_{11}(y_1)_N + a_{12}(y_2)_N)^2 + \\
&(a_{21}(y_1)_1 + a_{22}(y_2)_1)^2 + \cdots + (a_{21}(y_1)_N + a_{22}(y_2)_N)^2, \\
=& \left|\left| A \begin{pmatrix} (y_1)_1 \\ (y_2)_1 \end{pmatrix} \right|\right|^2 + \cdots + \left|\left| A \begin{pmatrix} (y_1)_N \\ (y_2)_N \end{pmatrix} \right|\right|^2, \\
\leq& b^2 {\lambda_{\max}}^2 \left({(y_1)_1}^2 + {(y_2)_1}^2 + \cdots + {(y_1)_N}^2 + {(y_2)_N}^2 \right), \\
=& b^2 {\lambda_{\max}}^2 \left({(y_1)_1}^2 + \cdots + {(y_1)_N}^2 + {(y_2)_1}^2 + \cdots + {(y_2)_N}^2 \right), \\
=& b^2 {\lambda_{\max}}^2 ||y||^2.
\end{aligned}$$

Taking the square root yields the result of Lemma 4.7.7.

Theorem 5.4.2. [29, Theorem 7][18] Assume that

$$\exists \alpha > 1, v_1 > 0, k_1 > 0 \text{ with } \forall y, z \in \mathbb{R}^N, ||z|| \leq ||x_0|| + k_1 : ||m(z)x|| \leq v_1 ||z||^\alpha ||x||$$
$$\text{and } \forall t \in [0, \infty) : \left|\left| \int_0^t \Phi(-s) F(s) \, \mathrm{d}s \right|\right| \leq k_1$$

holds. Then there exists a constant $k > 0$ dependent on b, α and c_0 such that if $v_1(||x_0|| + k_1)^\alpha \leq k$ is fulfilled[19], we can choose a $c_1 > 0$ such that

$$||x_n(t)|| \leq b(||x_0|| + k_1) e^{-c_1 t} \tag{5.34}$$

holds for all $n \in \mathbb{N}$ and all $t \in [0, \infty)$.
Proof:
Set $k := \frac{\lambda(\alpha-1)^2}{4\alpha b^3} {c_0}^2$ and $c_1 := \frac{\alpha+1}{2\alpha} c_0$. Then for x_0, the statement of this theorem holds by definition.
So we proceed by induction. For some $n \in \mathbb{N}$, assume that $||x^{(n-1)}(t)|| \leq b(||x_0|| + k_1) e^{-c_1 t}$ for all $t \in [0, \infty)$ (Induction Hypothesis). We wish to apply Theorem 5.4.1[20] and the inequality (5.29) it provides. For $s, r \in \mathbb{R}, s \geq r \geq 0$, we obtain for the kernel in equation (5.32), making use of the assumptions for this theorem:

$$\begin{aligned}
\left|\left| \frac{1}{\lambda} M(x^{(n-1)}(s-r)) x^{(n)}(r) \right|\right| &= \frac{1}{\lambda} ||m\left(x_1^{(n-1)}(s-r)\right) x_2^{(n)}(r)||, \\
&\leq \frac{v_1}{\lambda} ||x_1^{(n-1)}(s-r)||^\alpha ||x_2^{(n)}(r)||, \\
&\overset{I.H.}{\leq} \frac{b v_1}{\lambda} (||x_0|| + k_1)^\alpha e^{-\alpha c_1 (s-r)} ||x^{(n)}(r)||, \\
&\leq \frac{kb}{\lambda} e^{-\alpha c_1 (s-r)} ||x^{(n)}(r)||.
\end{aligned} \tag{5.35}$$

We have $\alpha c_1 = \frac{\alpha+1}{2} c_0 > c_0$, so together with the assumption on F above, the assumptions of Theorem 5.4.1 are satisfied and this yields for $x^{(n)}$:

$$||x^{(n)}(t)|| \leq b(k_1 + ||x_0||) e^{-\left(c_0 - \frac{kb^3}{\lambda} \frac{1}{\alpha c_1 - c_0}\right)},$$

[18] The constants in this theorem are slightly modified with respect to Saal [29], for the same reason as explained in footnote 16.
[19] Since the constant k is specifically chosen as a function of b, α and c_0, and these cannot take any values (c_0 is fixed, α must satisfy the first assumption of this theorem and b must be at least large enough to satisfy the bound (5.33)), this is in fact another constraint on the kernels to which this theorem applies.
[20] The object Φ in Theorem 5.4.1 is used with the ordinary matrix multiplication, unlike in equation (5.32). However, all we need in Theorem 5.4.1 is the bound for the multiplication with Φ in the assumptions, and this was verified just above the current theorem. So if we verify the other conditions as well, the inequality (5.29) can still be applied.

for all $t \in [0, \infty)$, where with the definitions of c_1 and k it follows that

$$c_0 - \frac{kb^3}{\lambda}\frac{1}{\alpha c_1 - c_0} = c_0 - \frac{kb^3}{\lambda}\frac{2}{(\alpha+1)c_0 - 2c_0} = c_0 - \frac{2(\alpha-1)^2 {c_0}^2}{4\alpha(\alpha-1)c_0} = c_0 - \frac{(\alpha-1)c_0}{2\alpha} = \frac{(\alpha+1)c_0}{2\alpha} = c_1.$$

□

Corollary 5.4.3. [29, Corollary 8] Let $m \in C^1(\mathbb{R}^N, \mathbb{R}^{N\times N})$. Then for any given $\Psi_0, \Psi_1 \in \mathbb{R}^N$ and $f \in C^0([0,\infty), \mathbb{R}^N)$ satisfying the conditions of Theorem 5.4.2[21] (with $x_0 = (\Psi_0, \Psi_1)$), there is a unique global solution $\psi \in C^2([0,\infty), \mathbb{R}^N)$ of (5.14). Additionally

$$||\dot{\psi}(t)|| + ||\psi(t)|| \leq (||x_0|| + k_1)e^{-c_1 t}$$

holds for $t \in [0, \infty)$.
Proof:
1 Existence.
We show existence of a solution on each bounded interval $[0, T]$ for some $T > 0$. Constructing an iteration sequence as has been done above, we can apply Theorem 5.4.2 and we know that each iteration function $x^{(n)} = (x_1^{(n)}, x_2^{(n)}) \in \mathbb{R}^{2N}$ is uniformly bounded by $(||x_0|| + k_1)e^{-c_1}$ on $[0, T]$. Differentiating (5.32), we obtain

$$\begin{aligned}\dot{x}^{(n)}(t) &= Ax^{(n)}(t) - \frac{1}{\lambda}\Phi(t)\Phi(-t)\int_0^t M(x^{(n-1)}(t-r))x^{(n)}(r)\,\mathrm{d}r + \Phi(t)\Phi(-t)F(t),\\ &= Ax^{(n)}(t) - \frac{1}{\lambda}\int_0^t M(x^{(n-1)}(t-r))x^{(n)}(r)\,\mathrm{d}r + F(t). \end{aligned} \tag{5.36}$$

We used that $\dot{\Phi}(t) = A\Phi(t)$ (it is the matrix exponential of A, which yields the unique solution to the initial value problem of result 4.7.10) and the fact that $\Phi(t)^{-1} = \Phi(-t)$ according to Lemma 4.7.3. We recognize a linearised version of (5.30, which is not unexpected since (5.32) was obtained from integrating that very equation. With this equality, we can also derive the following bound for $t \in [0, T]$:

$$\begin{aligned}||\dot{x}^{(n)}(t)|| &\leq ||A|| \cdot ||x^{(n)}(t)|| + \frac{1}{\lambda}\int_0^t ||M(x^{(n-1)}(t-r))x^{(n)}(r)||\,\mathrm{d}r + ||F(t)||,\\ &\overset{(5.35)}{\leq} ||A|| \cdot ||x^{(n)}(t)|| + \frac{1}{\lambda}\sup_{t\in[0,T]}||f(t)|| + \frac{kb}{\lambda}\int_0^t e^{-\alpha c_1(t-r)}||x^{(n)}(r)||\,\mathrm{d}r,\\ &\overset{(5.34)}{\leq} ||A|| \cdot ||x^{(n)}(t)|| + \frac{1}{\lambda}\sup_{t\in[0,T]}||f(t)|| + \frac{kb^2}{\lambda}(||x_0|| + k_1)\int_0^t e^{-\alpha c_1}e^{-c_1 r}\,\mathrm{d}r,\\ &= ||A|| \cdot ||x^{(n)}(t)|| + \frac{1}{\lambda}\sup_{t\in[0,T]}||f(t)|| + \frac{kb^2}{c_1\lambda}(||x_0|| + k_1)e^{-\alpha c_1}\left(1 - e^{-c_1 t}\right),\\ &\leq ||A||(||x_0|| + k_1)e^{-c_1} + \frac{1}{\lambda}\sup_{t\in[0,T]}||f(t)|| + \frac{kb^2}{c_1\lambda}(||x_0|| + k_1)e^{-\alpha c_1}. \end{aligned}$$

So we have a sequence of uniformly bounded, differentiable vector valued functions $x^{(n)}$, whose derivatives are also uniformly bounded. Moreover, it follows that $x^{(n)}$ and $\dot{x}^{(n)}$ can for all $n \in \mathbb{N}$ be bounded by a single constant. This brings us to the same situation as in the second part of the proof of Theorem 5.3.4. Following the line of argument shown there, it follows with the Arzelà-Ascoli Theorem that there is a convergent subsequence $\left(x^{(n_k)}\right)_{k\in\mathbb{N}}$. Let $x = (x_1, x_2) \in C([0, T], \mathbb{R}^{2N})$ denote the limit. Since the convergence with the Arzelà-Ascoli Theorem is in the supremum norm over the interval $[0, T]$, it is uniform.
We must make some observations to obtain convergence properties of $\dot{x}^{(n)}$ now. Since m is assumed to be continuously differentiable, the same holds for M. Since all iteration functions are uniformly bounded by the same constant, the same holds for the limit function x. Then by Lemma 4.6.9, M is Lipschitz continuous and uniformly continuous on a ball in $\mathbb{R}^{2N}$ with any radius larger than this constant. Then we can apply Lemma 4.2.4 and we see that the continuous function defined by the product $Ax^{(n)}$ converges uniformly, as does the

[21] Saal only mentions the assumptions of the bounds with α, v_1 and k_1, but as we remarked in footnote 19, also the demanded inequality with k poses a constraint.

composition $M \circ x^{(n-1)}$. Hence by Lemma 4.2.6, the integral itself converges uniformly and the uniform limit of (5.36) becomes

$$\begin{aligned}\lim_{n\to\infty} \dot{x}^{(n)}(t) &= Ax(t) - \frac{1}{\lambda}\int_0^t M(x(t-r))x(r)\,\mathrm{d}r + F(t),\\ &= \frac{d}{dt}x(t)\end{aligned}$$

The second equality holds in view of Theorem 4.2.3 (extended to vector valued functions by uniform convergence of the derivatives of all components functions) and shows that the limit function x is continuously differentiable, so $x \in C^1([0,T],\mathbb{R}^{2N})$ and x solves (5.30). Since this is the first-order transformation of (5.14), we see that x provides a solution ψ to (5.14) by setting $\psi(t) = x_1(t)$ for $t \in [0,T]$. Since $\ddot{\psi} = \dot{x_2}$ and we have seen $\dot{x}$ to be continuous, we obtain $\psi \in C^2([0,T],\mathbb{R}^N$, as desired. Finally note that x indeed takes the correct initial values because all iteration functions do so by definition according to equation (5.32), such that $\psi(0) = x_1(0) = \Psi_0$ and $\dot{\psi}(0) = x_2(0) = \Psi_1$.

2 Uniqueness.

Suppose two solutions $u, v \in C^2([0,T],\mathbb{R}^N)$ to (5.14) exist. By continuity of u, v and m, we know that $u(t)$, $v(t)$ and $m(u(t))$ can be bounded by some constant c for $t \in [0,T]$. Since m is continuously differentiable on $\mathbb{R}^N$, so is M, and by Lemma 4.6.9, M is Lipschitz continuous on every closed ball centred around the origin in $\mathbb{R}^N$. This is enough to obtain the local Lipschitz continuity required for the proof of Lemma 5.3.7. Hence we apply the argument given there and see that indeed we must have $u = v$ on $[0,T]$.

Note 5.4.4. We have established in both Theorem 5.3.1 and 5.4.3 existence and uniqueness on bounded intervals, of the form $[0,T]$. This suffices to define a global solution, for all $t \in \mathbb{R}^+$, because for any $t \in \mathbb{R}^+$ some $T > t$ can be taken to define the solution up to there. Since any such solution will also be a solution on all smaller intervals, it is also the unique solution there, and indeed this argument allows us to really extend the solution to as large intervals of definition as we wish.

We have now studied in-depth the existence and uniqueness proofs presented in [10] and [29] and have seen how their conclusions arise. In the next chapter, we will finalize our study of the schematic GMCT equations of chapter 3 by comparing the results of this chapter to the form of those equations. The detailed study of the theorems in this chapter allows us to make adaptations such that the theorems exactly match the equations for schematic GMCT.

6 Theorems applied to mode-coupling theory

In this chapter we check whether the theorems of chapter 5 can be applied to yield existence and uniqueness of a solution on $[0,\infty)$ (recall note 5.4.4) to the schematic GMCT equations (3.1) and (3.2) when a suitable closure $\mathcal{C}_N$ is applied. First we identify for both closures the appropriate kernel function. For (3.2), we will also have to take another look at the proofs of the theorems in [29], because some more parameters occur in GMCT than are treated by Saal.

6.1 Over-damped equations

To identify (3.1) with (5.1), for which the proof of section 5.2 is constructed, we immediately see that we can set (using n to indicate the levels $1,\dots,N$ of the system)

$$\begin{aligned}\tau_n &= \frac{1}{\mu_n},\\ m_n(t) = F_n(\phi(t)) &= \frac{\lambda_n}{\mu_n}\phi_{n+1}(t), \quad n = 1,\dots,N-1,\\ m_N(t) &= \mathcal{C}_N(\phi(t)).\end{aligned}$$

So Theorem 5.2.1 applies to our situation and we need to check whether each F_n defined above is completely monotone in each argument. This leads to the following lemma.

Lemma 6.1.1. For each closure function $\mathcal{C}_N$ that is absolutely monotone in each of its N variables on the interval $[0,\delta)$ for some $\delta > 1$, a unique solution to (3.1) on $[0,\infty)$ exists.
Proof:
Using the notation $x_1,\dots,x_N$ for the variables of F_n ($n \in \{1,\dots,N-1\}$), we have that $\frac{\partial F_n}{\partial x_m} = \frac{\lambda_n}{\mu_n}\delta_{n,m-1}$ (the Kronecker delta). It further follows that any second or higher order derivative is identically zero. So for $F_1,\dots,F_{N-1}$, we see that absolute monotonicity holds on the whole of $\mathbb{R}^N$. Now demanding the same of $\mathcal{C}_N$ on the interval of interest, we see that all requirements for Theorem 5.2.1 are met and this yields the desired conclusion. □

It is consequently clear that Theorem 5.2.1 can indeed be applied to the over-damped equations, since non-negativity of all combinations of partial derivatives follows readily for both the exp-N and all MF-N closures on the region $(\phi_1,\dots,\phi_N) \in [0,\delta)^N$.

6.2 Second-order system

In equation (5.14), we are short of parameters to describe the system (3.2) that is studied in GMCT. We divide (3.2) by ζ to get rid of the parameter in front of the first derivative and obtain the system

$$\begin{cases}\frac{1}{\zeta}\ddot{\phi}_n(t) + \dot{\phi}_n(t) + \frac{\mu_n}{\zeta}\phi_n(t) + \int_0^t \frac{\lambda_n}{\zeta}\phi_{n+1}(t)\dot{\phi}_n(t-\tau)d\tau = 0,\\ \phi_n(0) = 1, \dot{\phi}_n(0) = 0.\end{cases} \tag{6.1}$$

So we can write $\mathcal{A} = \mathcal{L} + \mathcal{N}$ as in section 5.3 to describe this system by $\mathcal{A}(\phi) = (0,1,0) \in \mathbb{R}^{N\times N\times N}$ by defining

$$\mathcal{L}(\phi) := \begin{pmatrix}\lambda\ddot{\phi} + \dot{\phi} + M\phi\\ 1\\ 0\end{pmatrix}$$

with $\lambda = \frac{1}{\zeta}$, the 1 and 0 denoting $(1,\dots,1) \in \mathbb{R}^N$ and $(0,\dots,0) \in \mathbb{R}^N$, respectively, and

$$M := \frac{1}{\zeta}\begin{pmatrix}\mu_1 & & \varnothing\\ & \ddots & \\ \varnothing & & \mu_N\end{pmatrix},$$

and $\mathcal{N}$ defined as in equation (5.16) with

$$m(\phi_1,\dots,\phi_N) := \frac{1}{\zeta}\begin{pmatrix}\lambda_1\phi_2 & & \varnothing\\ & \ddots & \\ \varnothing & & \lambda_N\mathcal{C}_N(\phi)\end{pmatrix}.$$

Our goal now is to use the conclusion of Theorem 5.3.1. We check that with the presence of M in the definition of $\mathcal{L}$, the same steps to prove it can be used. We outline here the modifications that are needed for each of the lemmas that are established to construct the proof.

1. Lemma 5.3.3.
   With M in the definition of $\mathcal{L}$, it still holds that $\mathcal{L}$ is a linear differential operator with constant coefficients (the matrix M only contains some constants). Therefore the same line of argument holds.

2. Lemma 5.3.4.
   The operator $\mathcal{N}$ is still exactly of the form that is assumed in the proof. So we need to check that m is k times continuously differentiable. For this we can demand that all components of m are, which is to say that we demand that $\mathcal{C}_N$ is k times continuously differentiable. We elaborate on the computation of the derivative of m in appendix C. We see there that a sufficient condition for m to be Fréchet differentiable on $\mathbb{R}^N$ is that all its component functions be continuously partial differentiable in all its variables on all of $\mathbb{R}^N$. In the GMCT situation, we see that we then need to demand this only of $\mathcal{C}_N$.

3. Lemma 5.3.5.
   In the inequalities that are derived in the proof, the presence of M now introduces some trouble, since estimates for h were obtained by taking the inner product with $\dot{h}$. We try to remedy this. Starting from taking the inner product of the appropriate version of equation (5.19) with $\dot{h}$ and integrating, we see that we need to consider what to do with
   $$\int_0^t Mh(s)\cdot\dot{h}(s)\,\mathrm{d}s.$$
   Let $\mu_{\min} := \frac{1}{\zeta}\min\{\mu_1,\dots,\mu_N\}$. Since M is diagonal, we can work out the integral to obtain
   $$\begin{aligned}\int_0^t Mh(s)\cdot\dot{h}(s)\,\mathrm{d}s &= \frac{1}{\zeta}\int_0^t\sum_{i=1}^N \mu_i h_i(s)\dot{h}_i(s)\,\mathrm{d}s,\\ &= \frac{1}{\zeta}\sum_{i=1}^N\mu_i\int_0^t\frac{1}{2}\frac{dh_i(s)^2}{ds}\,\mathrm{d}s,\\ &= \frac{1}{\zeta}\sum_{i=1}^N\frac{1}{2}\mu_i h_i(t)^2,\\ &\geq \frac{1}{2}\mu_{\min}\sum_{i=1}^N h_i(t)^2,\\ &= \frac{1}{2}\mu_{\min}h(t)^2.\end{aligned}$$
   (Note that $h(0)=0$ in this lemma.) The appropriate inequalities that can be derived corresponding to equation (5.20) become
   $$\begin{aligned}\frac{1}{2}\lambda\dot{h}(t)^2+\frac{1}{2}\mu_{\min}h(t)^2 &\leq \frac{1}{2}\lambda\dot{h}(t)^2+\frac{1}{\zeta}\sum_{i=1}^N\frac{1}{2}\mu_i h_i(t)^2,\\ &\leq -\int_0^t\dot{h}(s)^2\,\mathrm{d}s+\frac{1}{2}c^2T\left(\int_0^t||h(s)||^2\,\mathrm{d}s+\int_0^t||\dot{h}(s)||^2\,\mathrm{d}s+\int_0^t||\dot{h}(s)||^2\,\mathrm{d}s+\int_0^t||\dot{h}(s)||^2\,\mathrm{d}s\right),\\ &\leq -\int_0^t\dot{h}(s)^2\,\mathrm{d}s+3c^2T\max\left\{\frac{1}{\lambda},\frac{1}{\mu_{\min}}\right\}\left(\int_0^t\frac{1}{2}\lambda\dot{h}(s)^2\,\mathrm{d}s+\int_0^t\frac{1}{2}\mu_{\min}h(s)^2\,\mathrm{d}s\right).\end{aligned}$$
   This is again the suitable form to use Gronwall's inequality. We used that for arbitrary $\lambda,\mu>0, a,b\geq 0$, we can write
   $$a+b=\frac{1}{\lambda}\lambda a+\frac{1}{\mu}\mu b\leq\max\left\{\frac{1}{\lambda},\frac{1}{\mu}\right\}\lambda a+\max\left\{\frac{1}{\lambda},\frac{1}{\mu}\right\}\mu b=\max\left\{\frac{1}{\lambda},\frac{1}{\mu}\right\}(\lambda a+\mu b).$$
   So the conclusion of the lemma still holds when M appears in the equation.

4. Lemma 5.3.6
   Here we need again to adapt the estimates to the situation with M in the equation. The procedure to obtain the estimates is largely the same as above: taking the inner product with $\dot{h}(t)$ and integrating, after which the integrals in which M does not appear are estimated. The fact that this is done on successive intervals is no problem, since we can adapt the estimate on each interval. An extra subtlety is that the initial values in equation (5.22) are non-zero. Hence ${\psi_0}^2$ should be replaced by $\frac{1}{\zeta}\sum_{i=1}^{N}\mu_i\psi_i(0)^2$ (the ψ_i are the vector components of ψ). In terms of the constants we use in Lemma 5.3.6, this can be estimated as follows:
   $$\frac{1}{\zeta}\sum_{i=1}^{N}\mu_i\psi_i(0)^2 \leq \mu_{\max}\sum_{i=1}^{N}\psi_i(0)^2 = \mu_{\max}{\psi_0}^2 \leq \mu_{\max}B^2,$$
   with $\mu_{\max} := \frac{1}{\zeta}\max\{\mu_1,\ldots,\mu_N\}$. Then applying the method above for the lower sides of the inequalities obtained in the proof, we see that on each interval a bound on ψ and $\dot{\psi}$ can be identified with Gronwall's inequality as is done in the original proof. Hence also the conclusion of this lemma holds.

In conclusion, we see that also in the situation with M that is encountered in GMCT, we can conclude that Theorem 5.3.1 holds. So if m satisfies condition (5.15) and $\mathcal{C}_N$ is continuously partial differentiable in all its variables on all of $\mathbb{R}^N$, we know that a global solution exists. We formalize this below in Lemma 6.3.1.

6.3 Existence for different closures

Here we check explicitly for the exponential and mean-field closures if we can establish global existence and uniqueness of a solution to the GMCT equations.

Lemma 6.3.1. For any closure function $\mathcal{C}_N$ that is continuously partial differentiable with respect to all its arguments, there exists some $T > 0$ such that a unique solution of (5.14) exists on $[0, T]$.
Proof:
Cf. Theorem 5.3.1 and the discussion of section 6.2. Since in this situation $f \equiv 0 \in C^0([0,\infty), \mathbb{R}^N)$, we have to check that $m \in C^1[\mathbb{R}^N, \mathbb{R}^{N\times N}]$. Let us denote the N arguments of m by $x = (x_2, x_3, \ldots, x_N)$ here. For $i = 1, \ldots, N$ we see that
$$\left(\frac{\partial m}{\partial x_i}\right)_{jj}(x) = \begin{cases} 0 & \text{if } i \neq j+1 \wedge j < N, \\ \frac{\lambda_{i-1}}{\zeta} & \text{if } i = j+1, \\ \frac{\lambda_N}{\zeta}\frac{\partial \mathcal{C}_N}{\partial x_i}(x) & \text{if } j = N. \end{cases}$$
All off-diagonal elements are also 0. Now a vector-valued function is continuous if and only if all its components are continuous functions and for any partial derivative of m but for the last component, these will be constants (and thus continuous). Since the last component equals $\lambda_N/\zeta \cdot \frac{\partial \mathcal{C}_N}{\partial x_i}(x)$, this function is continuous for all $i = 1, \ldots, N$ if $\frac{\partial \mathcal{C}_N}{\partial x_i}$ is a continuous function on $\mathbb{R}^N$ for all $i = 1, \ldots, N$, which is the assumption for this lemma. Continuity of all partial derivatives on all of $\mathbb{R}^N$ now ensures that $m \in C^1(\mathbb{R}^N, \mathbb{R}^{N\times N})$, such that Theorem (5.3.1) can be applied to yield the desired conclusion. □

Remark 6.3.2. For the exp-N closure, $\mathcal{C}_N = 0$ so this is certainly continuously differentiable. For any MF-N closure, $\mathcal{C}_N$ is a monomial in the variables $x_1, \ldots, x_N$ and is therefore also continuously differentiable on $\mathbb{R}^N$. So for both closure types, some interval $[0, T]$ for $T > 0$ exists such that a local, unique solution for the mode-coupling hierarchy exists.

Lemma 6.3.3. (Existence for exponential closure.) For the exp-N closure, the condition (5.15) holds. Consequently, a unique global solution on $[0,\infty)$ of (5.14) exists.
Proof:
Let $x = (x_1, \ldots, x_N) \in \mathbb{R}^N$ and $y = (y_1, \ldots, y_N) \in \mathbb{R}^N$. Also define $\Lambda = \frac{1}{\zeta}\max_{i=1,\ldots,N-1}\{\lambda_i\}$. For the exp-$N$ closure, we have
$$m(x) = \frac{1}{\zeta}\begin{pmatrix} \lambda_1 x_2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \lambda_{N-1}x_N & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

We now find the following (in)equalities:

$$\begin{aligned}
||m(x)y||^2 &= \left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}\left(m(x)\right)_{ij} y_j\right)^2\right), \\
&= \sum_{i=1}^{N-1}\left(\left(\frac{\lambda_i}{\zeta}\right)^2 {x_{i+1}}^2 {y_i}^2\right), \\
&\leq \Lambda^2 \sum_{i=1}^{N-1}\left({x_{i+1}}^2 {y_i}^2\right), \\
&\leq \Lambda^2 \sum_{i=0}^{N-1}\left({x_{i+1}}^2 \sum_{j=1}^{N} {y_j}^2\right) = \Lambda^2 \sum_{i=1}^{N} {x_i}^2 \sum_{j=1}^{N} {y_j}^2, \\
&= \Lambda^2 ||x||^2 ||y||^2.
\end{aligned}$$

Hence $||m(x)y|| \leq \Lambda ||x|| \cdot ||y||$ and also $||m(x)y|| \leq \Lambda(1+||x||)||y||$ (as $\Lambda, |y| \geq 0$), so we see that condition (5.15) is indeed satisfied with e.g. $c = \Lambda$. □

Lemma 6.3.4. (Condition for mean-field closures.) Let $\mathcal{C}_N$ be determined by some MF-N closure. Then condition (5.15) does not hold.
Proof:
We will construct a vector $y \in \mathbb{R}^N$ and a sequence of vectors $\left(x^k\right)_{k\in\mathbb{N}}$ such that $||m(x^k)y||$ becomes arbitrarily much larger than $(1 + ||x^k||)||y||$ as $k \to \infty$.
Let $y = e_N \in \mathbb{R}^N$ be the N^{th} unit vector in $\mathbb{R}^N$. Then $m(x^k)y = \lambda_N/\zeta \cdot \mathcal{C}_N(x^k) \cdot e_N$ and $||y|| = 1$.
In the notation of definition 3.2.1, we can write $\mathcal{C}_N(x^k) = \prod_{n=1}^{N} \left(x_n^k\right)^{m_n}$. Now there is a minimal $i \in \{1, 2, \ldots, N\}$ such that m_i is nonzero, since all m_i must add to $N + 1$ according to (3.4).
Let us first suppose that there is also some $j \in \{i + 1, i + 2, \ldots, N\}$ such that m_j is nonzero (and $i \neq j$). In particular, this can only hold for $N > 2$. Now define for $k \in \mathbb{N}$ and with (yet) arbitrary scalars $\alpha_k \in \mathbb{R}$ the components $(x^k)_n$ of the vector x^k by

$$\left(x^k\right)_n = \begin{cases} \alpha_k & \text{if } n \in \{i, j\}, \\ 1 & \text{otherwise,} \end{cases} \quad n = 1, \ldots, N.$$

It follows for x^k that $\mathcal{C}_N(x^k) = {\alpha_k}^{m_i+m_j}$ and $||x^k||^2 = (N - 2) + 2{\alpha_k}^2$. We restrict α_k to values greater than or equal to 1 such that we have $||x||^2 \geq 1$. We now have the following inequality:

$$\begin{aligned}
\frac{||m(x^k)y||^2}{(1 + ||x^k||)^2||y||^2} &\geq \frac{||m(x^k)y||^2}{(2||x^k||)^2||y||^2}, \\
&= \frac{(\lambda_N/\zeta)^2}{4} \frac{{\alpha_k}^{2m_i+2m_j}}{(N - 2) + 2{\alpha_k}^2}.
\end{aligned}$$

We now choose $\alpha_k = k$ for all $k \in \mathbb{N}$ and let $k > N$ such that also $k^2 > N$. Then we proceed:

$$\begin{aligned}
\frac{||m(x^k)y||^2}{(1 + ||x^k||)^2||y||^2} &> \frac{(\lambda_N/\zeta)^2}{4} \frac{k^{2m_i+2m_j}}{N + 2k^2}, \\
&> \frac{(\lambda_N/\zeta)^2}{4} \frac{k^{2m_i+2m_j}}{k^2 + 2k^2}, \\
&\geq \frac{(\lambda_N/\zeta)^2}{12} k^2,
\end{aligned}$$

since $m_i, m_j \geq 1$. Hence the expression on the left cannot be bounded by any constant $c > 0$.

Let us now turn to the situation where $\mathcal{C}_N$ can be written as a function of just one of the components of the vector x^k, say $\mathcal{C}_N(x^k) = \mathcal{C}_N(x_i^k)$ for some $i \in \{1, \ldots, N\}$ fixed. Note that this also includes the standard mode-coupling theory MF-1 closure discussed in remark 3.2.3. For this case, we define the components $\left(x^k\right)_n$ by

$$\left(x^k\right)_n = \begin{cases} k & \text{if } n = i, \\ 1 & \text{otherwise,} \end{cases} \quad n = 1, \ldots, N.$$

We then have $||x||^2 = (N-1)^2 + k^2$ and $\mathcal{C}_N(x^k) = k^{m_i}$, where due to (3.4), $m_i = \frac{N+1}{i}$. In particular, since $i \in \{1, \ldots, N\}$, we see that $m_i > 1$. We still let $y = e_N$. Now for $k > N-1$, we obtain the following inequalities:

$$\begin{aligned}
\frac{||m(x^k)y||^2}{(1+||x^k||)^2||y||^2} &\geq \frac{||(\lambda_N/\zeta)\cdot k^{m_i} e_N||^2}{(2||x||)^2||y||^2}, \\
&= \frac{(\lambda_N/\zeta)^2 k^{2m_i}}{4((N-1)^2+k^2)}, \\
&> \frac{(\lambda_N/\zeta)^2 k^{2m_i}}{8k^2}, \\
&\geq \frac{(\lambda_N/\zeta)^2}{8} k^2.
\end{aligned}$$

We can conclude that also in this situation, the quantity on the left cannot be bounded by any constant $c > 0$. So any MF-N closures never satisfy condition (5.15). □

Remark 6.3.5. Lemma 6.3.4 shows that global existence for a solution to a system for which the mean-field closure is used, cannot be proved based on Theorem 5.3.1 provided by [29]. This is in contrast to Lemma 6.3.3 for the exponential closure. We can solve this problem for the MF-$N(1N)$ closure with some physical insight that allows us to redefine m outside the region $[-1,1] \times \mathbb{R}^{N-1}$. A proof provided by S. Ciarella shows that for physical solutions to (6.1), it holds that $|\phi_1(t)| \leq 1$ for all $t \in \mathbb{R}$. For any MF-N closure, we could show that condition (5.15) holds for the corresponding coupling function m for input vectors x restricted to the condition $||x|| < \sqrt{2}$. So redefining m for inputs with $|x_1| > 1$ will not change the physics and the definition outside this region can be made such that m does satisfy (5.15) on all of $\mathbb{R}^N$ in a way that preserves also the differentiability condition. We will make all this rigorous now with Lemma 6.3.6 and construction 6.3.7.

Lemma 6.3.6. Let $x \in \mathbb{R}^N$ satisfy $|x_1| \leq 1$. Then for the MF-$N(1N)$ closure, m satisfies the following:

$$\exists c > 0 \quad \forall y \in \mathbb{R}^N : ||m(x)y|| \leq c(1+||x||)||y||.$$

Proof:
Let $y = (y_1, \ldots, y_N) \in \mathbb{R}^N$ and define $\Lambda = \frac{1}{\zeta} \cdot \max_{i=1,\ldots,N}\{\lambda_i\}$. For the MF-$N(1N)$ closure, we have

$$m(x) = \frac{1}{\zeta}\begin{pmatrix} \lambda_1 x_2 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \lambda_{N-1}x_N & 0 \\ 0 & \cdots & 0 & \lambda_N x_1 x_N \end{pmatrix}.$$

We now find the following (in)equalities:

$$\begin{aligned}
||m(x)y||^2 &= \left(\sum_{i=1}^{N}\left(\sum_{j=1}^{N}(m(x))_{ij}\, y_j\right)^2\right), \\
&= \sum_{i=1}^{N-1}\left(\left(\frac{\lambda_i}{\zeta}\right)^2 {x_{i+1}}^2 {y_i}^2\right) + \left(\frac{\lambda_N}{\zeta}\right)^2 {x_1}^2 {x_N}^2 {y_N}^2, \\
&\leq \Lambda^2 \sum_{i=1}^{N-1}\left({x_{i+1}}^2 {y_i}^2\right) + \Lambda^2 \cdot 1 \cdot {x_N}^2 {y_N}^2, \\
&\leq \Lambda^2 \sum_{i=1}^{N-1} {x_{i+1}}^2 \sum_{i=1}^{N-1} {y_i}^2 + \Lambda^2 {y_N}^2 \sum_{i=1}^{N-1} {x_{i+1}}^2, \\
&= \Lambda^2 \sum_{i=1}^{N-1} {x_{i+1}}^2 \sum_{i=1}^{N} {y_i}^2, \\
&\leq \Lambda^2 \sum_{i=1}^{N} {x_i}^2 \sum_{i=1}^{N} {y_i}^2, \\
&= \Lambda^2 ||x||^2 ||y||^2.
\end{aligned}$$

Hence $||m(x)y|| \leq \Lambda ||x|| \cdot ||y|| \leq \Lambda^2(1+||x||)||y||$. □

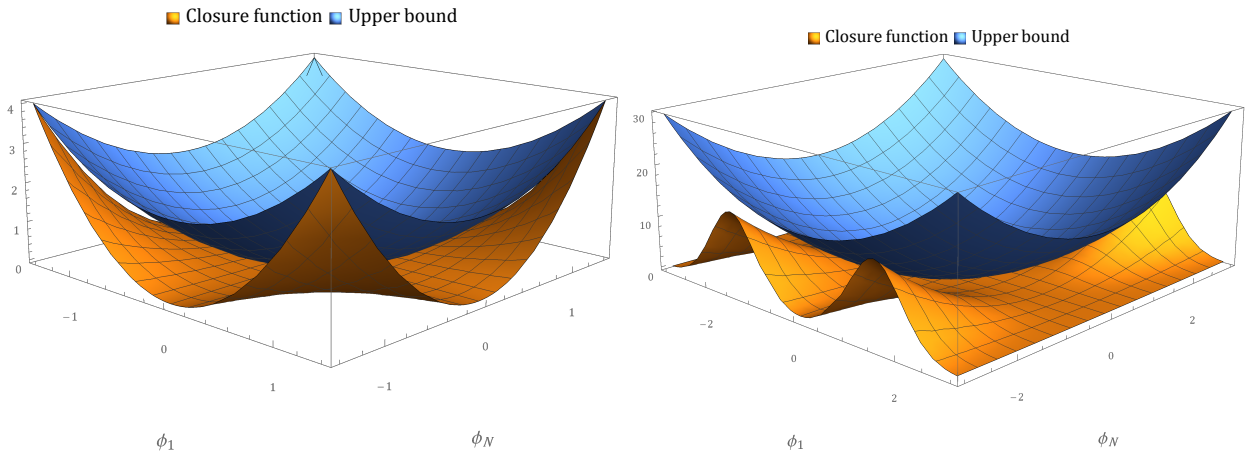

Figure 8: *Graphs illustrating the construction and boundedness of the coupling function $\mathcal{C}_N$ as a function of ϕ_1 and ϕ_N. Left: The original MF-$N(1N)$ closure function (orange shades), intersecting its allowed bound (blue shades) for $(\phi_1, \phi_N) = (\pm\sqrt{2}, \pm\sqrt{2})$. Right: the MF-$N(1N)$ closure function, identical to its physical definition for $\phi_1 \in [-1, 1]$ and continuously differentiable extended on the rest of $\mathbb{R}^2$ to satisfy stay below a bound $\alpha({x_1}^2 + {x_2}^2)$ for some $\alpha > \sqrt{e}$, which is the bound used in construction 6.3.7.*

Construction 6.3.7. We now construct a new function $m : \mathbb{R}^N \to \mathbb{R}^N$ which still defines the MF-$N(1N)$ closure for $x_1 \in [-1, 1]$, but also fulfils the growth condition (5.15) on all of $\mathbb{R}^N$ and is continuously differentiable. From the foregoing proof of Lemma 6.3.6, because m is diagonal in matrix notation, it is actually clear that only the closure function $\mathcal{C}_N$ is of importance and that its value squared should be bounded by some expression that is at most quadratic in the components $x_1, \ldots, x_N$. To this end, let us define

$$\mathcal{C}_N(x_1, \ldots, x_N) = \mathcal{C}_N(x_1, x_N) = \begin{cases} x_1 \cdot x_N, & x_1 \in [-1, 1], \\ -x_N \cdot e^{\frac{1}{4}-(x_1+\frac{3}{2})^2}, & x_1 < -1, \\ x_N \cdot e^{\frac{1}{4}-(x_1-\frac{3}{2})^2}, & x_1 > 1. \end{cases}$$

For a visual illustration of this function, see figure 8. In words, while $\mathcal{C}_N$ is allowed to continue to grow linearly in x_N, the linear modulation, say, by x_1 is replaced by an exponentially decaying modulation when x_1 is no longer in the interval $[-1, 1]$. We will continue by establishing that $\mathcal{C}_N$ is continuously differentiable and satisfies the imposed bound, such that the same can be said about the closure function m. We will only do this explicitly for the region $x_1 \geq 1$. On $[-1, 1]$, $\mathcal{C}_N$ is simply a polynomial that satisfies the growth condition as a result of Lemma 6.3.6 and the case $x_1 \leq -1$ is very similar.

- **Continuity** on $x_1 = 1$.
  We have $\lim_{x_1 \uparrow 1} \mathcal{C}_N(x_1, x_N) = \lim_{x_1 \uparrow 1} x_1 \cdot x_N = x_N = \mathcal{C}_N(1, x_N)$, and $\lim_{x_1 \downarrow 1} \mathcal{C}_N(x_1, x_N) = \lim_{x_1 \downarrow 1} x_N \cdot e^{\frac{1}{4}-(x_1-\frac{3}{2})^2} = x_N \cdot e^{\frac{1}{4}-(1-\frac{3}{2})^2} = x_N$. This holds for all $x_N \in \mathbb{R}$, so continuity follows.

- **Continuity of** $\frac{\partial}{\partial x_1}\mathcal{C}_N$.
  From the left ($x_1 \leq 1$), we have $\lim_{x_1 \uparrow 1} \frac{\partial}{\partial x_1}\mathcal{C}_N(x_1, x_N) = \lim_{x_1 \uparrow 1} x_N = x_N$. From the right ($x_1 > 1$), we find $\lim_{x_1 \downarrow 1} \frac{\partial}{\partial x_1}\mathcal{C}_N(x_1, x_N) = \lim_{x_1 \downarrow 1} -2(x_1 - \frac{3}{2})x_N e^{\frac{1}{4}-(x_1-\frac{3}{2})^2} = x_N$. This yields continuity of the x_1-partial derivative on $x_1 = 1$.

- **Continuity of** $\frac{\partial}{\partial x_N}\mathcal{C}_N$.
  From the left, we obtain $\lim_{x_1 \uparrow 1} \frac{\partial}{\partial x_N}\mathcal{C}_N(x_1, x_N) = \lim_{x_1 \uparrow 1} x_1 = 1$ and from the right, $\lim_{x_1 \downarrow 1} \frac{\partial}{\partial x_N}\mathcal{C}_N(x_1, x_N) = \lim_{x_1 \downarrow 1} e^{\frac{1}{4}-(x_1-\frac{3}{2})^2} = e^0 = 1$. So both partial derivatives are continuous at $x_1 = 1$ for all x_N, hence C_N is (totally) differentiable.

- **Correct boundedness**
  Let $d : \mathbb{R}^2 \to \mathbb{R}$ be a function defined by $d(x_1, x_N) = \alpha({x_1}^2 + {x_N}^2) - \sqrt{e}{x_N}^2 e^{-2(x_1-\frac{3}{2})^2}$ for some number

$\alpha > \sqrt{e}$ (e.g. $\alpha = 2$ would suffice). We then have that

$$\begin{aligned}\frac{\partial}{\partial x_N} d(x_1, x_N) &= 2\alpha x_N - 2\sqrt{e} x_N e^{-2(x_1 - \frac{3}{2})^2}, \\ &\geq 2x_N \left(\alpha - \sqrt{e}\right) \Rightarrow\end{aligned}$$

$$\begin{cases} \frac{\partial}{\partial x_N} d(x_1, x_N) \geq 0 & \text{if } x_N \geq 0, \\ \frac{\partial}{\partial x_N} d(x_1, x_N) < 0 & \text{otherwise.} \end{cases}$$

So in the x_N direction, $d(x_1, x_N)$ is always non-decreasing if we move away from the $x_N = 0$ axis. Furthermore, we have on this axis,

$$\forall x_1 \in \mathbb{R} : d(x_1, 0) = \alpha {x_1}^2 \geq 0.$$

Combining these facts, we conclude that $d(x_1, x_N) \geq 0$ for all x_1 and x_N, or equivalently, recognizing that for $x_1 > 1$, we have $d(x, y) = \alpha({x_1}^2 + {x_N}^2) - \mathcal{C}_N(x_1, x_N)^2$,

$$\forall x_1 > 1, x_N \in \mathbb{R} : \mathcal{C}_N(x_1, x_N)^2 \leq \alpha({x_1}^2 + {x_N}^2).$$

Putting this in the proof of Lemma 6.3.6, we find that the following holds for the MF-$N(1N)$ closure, redefined for $|x_1| > 1$:

$$\begin{aligned}||m(x)y||^2 &= \left(\sum_{i=1}^{N} \left(\sum_{j=1}^{N} \left(m_i(x)\right)_j y_j \right)^2 \right), \\ &= \sum_{i=1}^{N-1} \left(\left(\frac{\lambda_i}{\zeta} \right)^2 {x_{i+1}}^2 {y_i}^2 \right) + \left(\frac{\lambda_N}{\zeta} \right)^2 \mathcal{C}_N(x_1, x_N)^2 {y_N}^2, \\ &\leq \Lambda^2 \sum_{i=1}^{N-1} \left({x_{i+1}}^2 {y_i}^2 \right) + \alpha\Lambda^2({x_1}^2 + {x_N}^2){y_N}^2, \\ &\leq \Lambda^2 \sum_{i=1}^{N} {x_i}^2 \sum_{i=1}^{N-1} {y_i}^2 + \alpha\Lambda^2 {y_N}^2 \sum_{i=1}^{N} {x_i}^2, \\ &\leq \alpha\Lambda^2 \sum_{i=1}^{N} {x_i}^2 \sum_{i=1}^{N} {y_i}^2, \\ &= \alpha\Lambda^2 ||x||^2 ||y||^2.\end{aligned}$$

It follows that also $||m(x)y||^2 \leq \Lambda^2 \max\{\alpha, 1\} ||x||^2 ||y||^2$, so the MF-$N(1N)$ closure (for $|x_1| \leq 1$) satisfies condition (5.15) and by Theorem 5.3.1, a unique solution to to the GMCT equations that is twice continuously differentiable exists for this closure.

We do not formally elaborate on the application of the existence result of section 5.4. Without further details, we state that have seen that also in this case it is applicable to the exponential closure, but not to the MF-closures. However, we can now conclude this section with the existence and uniqueness of solutions to the over-damped and full schematic GMCT equations based on two of the results of chapter 5, as we investigated their applicability with various lemmas. In the next chapter, we start to explore numerical results produced by the standard schematic MCT equation, i.e., GMCT closed at first order by the mean-field closure.

7 Modelling different relaxation time behaviours

We would now like to see if standard schematic MCT (GMCT closed at first order with the MF-$N(1^2)$ closure) is capable of predicting super-strong glass formation over a range of large relaxation times that we can manipulate ourselves. We look at the over-damped limit. Denoting by τ the relaxation time (which we define below), we would particularly like $\log(\tau)$ to have a decreasing slope as a function of inverse temperature, such as is found in the simulation results of figure 2. Since all structural input to MCT of the liquid is contained in the static structure factor S, our tool will be the choice of S as a function of temperature. We have seen in the introduction that under some conditions for vitrimers, its main peak (which is the actual input for the schematic approximation of MCT), grows in a way that also displays a decreasing slope as a function of inverse temperature (figure 2). Further this seems to impact the relaxation time in the same way, leading to super-strong glass formation until the power-law divergence of τ arises that is always predicted by MCT, but not necessarily physically correct. In accordance with the research goal stated in the introduction, we aim to place the cross-over from super-strong glass formation to power-law divergence to high relaxation times and as a measure for this we set the goal of an increase of 10^5.

Let us first make some definitions that we will use in the presentation of the results. Looking at standard schematic MCT in the over-damped limit, we denote by μ_1 and λ_1 the frequency and damping coefficients, respectively, appearing in the equations (3.1) for the case $N = 1$. As we do not consider GMCT now, we drop the indices and write μ and λ. We are interested in the relaxation time τ of the normalized density correlator $\phi_1(t)$, which we define by

$$\phi_1(\tau) = \frac{1}{e}, \tag{7.1}$$

i.e., the time needed for the normalized intermediate scattering function to reach below a typical decay value. Since we are interested in τ as a function of temperature, and this temperature decreases when we approach the transition, we look at the inverse temperature Λ,

$$\Lambda = T^{-1}.$$

We define Λ_c to be the value of Λ for which $\tau(\Lambda)$ diverges, when we approach the divergence from below (that is, from the high-temperature regime). In [20], the equations for ϕ without dropping the second derivative are studied, and he shows that divergence occurs at Λ_c which satisfies $\lambda_1/\mu_1 = 4$. Since the second derivative does not change the long-time dynamics (which are relevant for a divergent relaxation time), we can use this in our numerical work. The outcomes are always consistent with this result. Finally, we have already seen that the predicted transition does not correspond to the physical glass transition. We therefore scale the variable Λ in our results, making use of the variable

$$\epsilon = \frac{\Lambda - \Lambda_c}{\Lambda_c} \tag{7.2}$$

We noted in section 2.4 that the structure factor S is usually seen to become more peaked at lower temperatures. We therefore model S as an increasing function of Λ. Based on the observed possibly logarithmic growth of the first peak height in S in simulations for vitrimers (figure 2), we will try to establish the desired effect also by such functional forms. This will result in the functional form we choose for λ and μ, which will thus depend on Λ as a control parameter by which we approach the transition.

7.1 On the numerics used

To analyse the relaxation time as a function of Λ, we will integrate the schematic MCT equations, using an implementation by L.M.C. Janssen of the algorithm provided by M. Fuchs [9], up to time values of 10^{20} or a value of $\phi_1(t) < 2^{-52} \cdot 10^4$ at the end time of the integration. This last number is 10 000 times the floating point accuracy of Matlab. To implement the definition for τ in (7.1), we average the two time values for which $\phi_1(t)$ is just above and just below $0.367879 \approx 1/e$. The use of a time grid introduces a numerical error here, as does the rounding of $1/e$. In the data computed, it is therefore seen that the curves for τ as a function of Λ display step-wise growth for a fined grid of Λ. This means that increasing Λ in smaller steps does not result in more smooth data. An analysis of these errors is beyond the scope of this work. The resulting data is smooth enough for our purposes and error analysis will be needed when more quantitative results are required.

The data we consider will typically look as shown in figure 9 at the curve labelled with the number 2. Our interest is to see at what value of τ the curve changes from a convex shape to a concave one, i.e. from super-strong to fragile glass formation. We call this the cross-over point. To be precise, we do not wish to determine an exact value from the data, but aim to see if we can influence its height by changing the structure factor in

the theory. To this end, it suffices to identify a rough but systematic estimate of its height. Less quantitatively, the desired effect can of course be observed from comparing the various graphs of τ as a function of Λ.

We summarize our procedure to determine the estimated cross-over point. We use the forward-difference for a numerical approximation of the first derivative. We do so over a grid of 80 to 120 values of Λ. The cross-over corresponds to the minimum of this function, but the data displays a bumpy plateau instead of a smooth valley (this can be ascribed to the small numerical errors introduced above). Our choice is to take the average value of the lowest ten percent of the forward difference as a threshold, then consider the forward differences from the highest value of Λ to below and put the cross-over at the last value of Λ before crossing the threshold. This procedure results in the circles in figure 9. A more elaborate and quantitative analysis of the cross-over could be executed with e.g. a polynomial fit to the data which allows for the determination of the cross-over as the inflection point.

Physically, the important aspect of the cross-over is its height as a function of Λ, since a large relaxation time relates to a large viscosity upon cooling. To compare the magnitude of the relaxation time across various models, we divide them all by one typical high-temperature value. This will be the time scale on which we wish to reach the experimental glass transition as a growth over five orders of magnitude. In order to do so, we first introduce the standard $F2$ model by Leutheusser [20].

1 The standard $F2$ model is characterized by the parameters

$$\begin{aligned} \mu(\Lambda) &= 1, \\ \lambda(\Lambda) &= \Lambda. \end{aligned} \tag{7.3}$$

This is a simple and one of the first models for which the MCT equations have been studied and its name arises from the fact that the closure is obtained with the function F^2 in the memory kernel of the integral. The dependence of τ on Λ for this model is shown in figure 9 labelled by the number 1. It shows the power-law divergence near the critical value $\Lambda_c = 4$ (which solves $\lambda(\Lambda_c)/\mu(\Lambda_c) = 4$) and shows fragile behaviour over the rest of the temperature domain. As a typical high-temperature value, we choose $\Lambda = 0.02$. At this temperature, we find

$$\tau_{\mathrm{F2}} = 1.024 \tag{7.4}$$

for the relaxation time. This is our reference value for all the results that we present hereafter.

7.2 Various modelling approaches

In this section, we explore numerical results for various functional forms of λ and μ. In [6], the following is derived, relating these parameters to the structure factor S:

$$\begin{aligned} \mu(\Lambda) &= \frac{k_B T}{m} \frac{{k_0}^2}{S(\Lambda)}, \\ \lambda(\Lambda) &\approx 0.0082 \frac{k_B T}{\rho m} {k_0}^3 (S(\Lambda) - 1)^2, \end{aligned} \tag{7.5}$$

where S is a model for the structure factor at the fixed wave number k_0. Since we are only interested in the mathematical flexibility of the MCT equation, we drop all constant factors in (7.5) that we do not aim to control directly, and replace the definitions by (recall that $T = 1/\Lambda$)

$$\begin{aligned} \mu(\Lambda) &= \frac{1}{\Lambda S(\Lambda)}, \\ \lambda(\Lambda) &= \frac{(S(\Lambda) - 1)^2}{\Lambda}. \end{aligned} \tag{7.6}$$

All results for the choices below are plotted in figure 9, which also includes the standard $F2$ model. The numbers we assign to each choice correspond to the labels of all curves in the figure.

2 Being interested in the influence of a logarithmic form for the structure factor as an intuitive guess from figure 2, let us first behold the choice $S(\Lambda) = \ln(\Lambda)$ (the natural logarithm, i.e. with base e). Now the logarithm is not a perfectly sensible model for the structure factor, as it diverges to $-\infty$ for high temperatures, while it should be 1, resembling an ideal gas. We set $S(\Lambda) = 1 + \ln(\Lambda)$ for $\lambda(\Lambda)$ to make it somewhat more reasonable for values of $\Lambda > 1$, and our approximation is not to add the 1 to μ to still simplify the expressions slightly and

obtain a factor $1/\Lambda$ in front. All in all, we set

$$\begin{aligned} \mu(\Lambda) &= \frac{1}{\Lambda \ln(\Lambda)}, \\ \lambda(\Lambda) &= \frac{\ln(\Lambda)^2}{\Lambda}. \end{aligned} \tag{7.7}$$

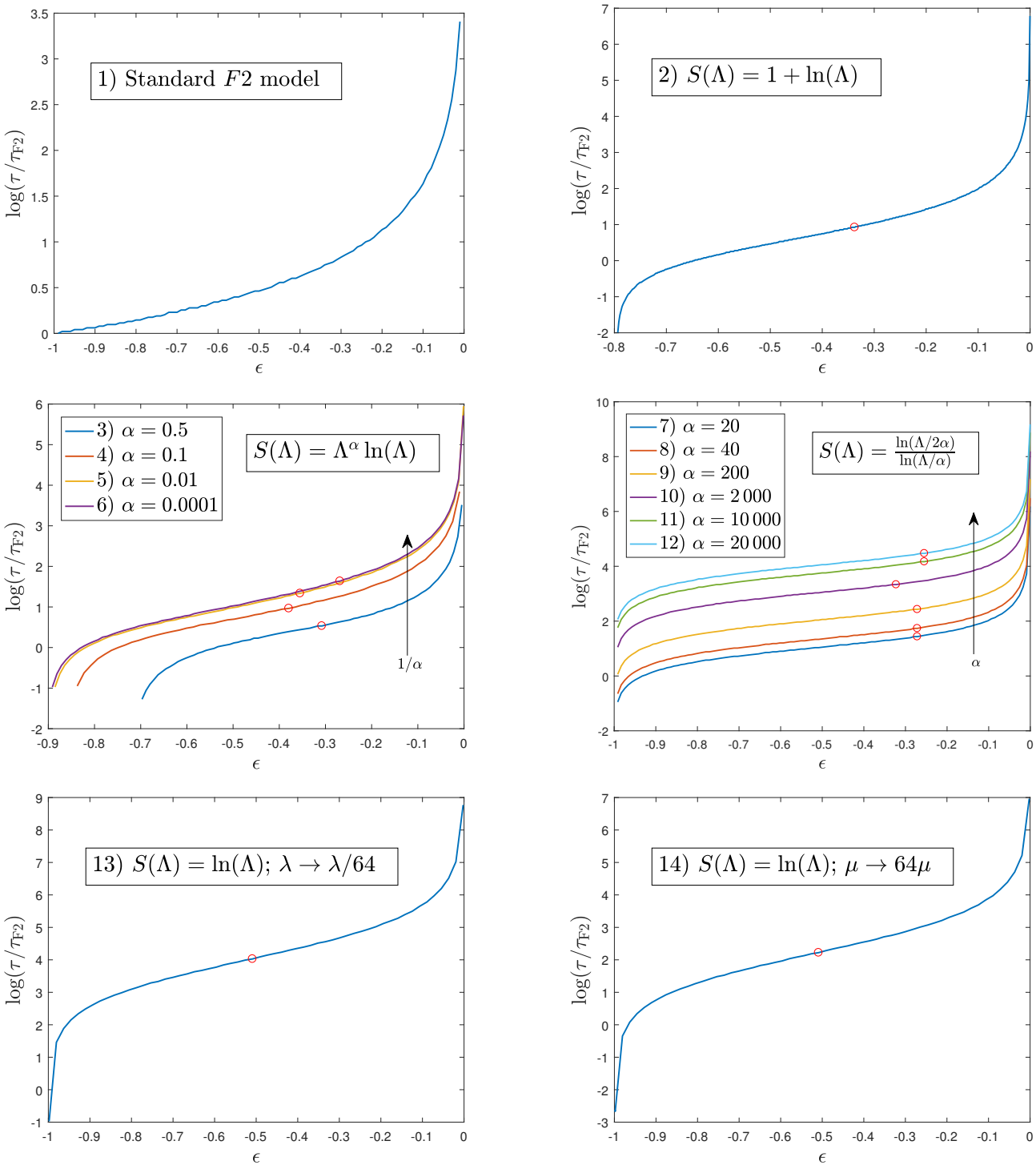

Figure 9: *The relaxation time τ as a function of ϵ, which is a scaled inverse temperature, for different functional forms of the structure factor S. The red circles represent our estimates for the cross-over from super-strong glass formation to power-law divergence, following the procedure of section 7.1, which is absent in the linear standard F2 model. We see various ways to increase τ at the cross-over by tuning the functional forms used. Detailed discussions for each curve can be found in the main text, where the numerical labels correspond to the numbered boxes on pages 60-63.*

From figure 9, we see that a strong glass former is manifest at higher temperatures, while the power-law develops for ϵ closer to 0, i.e., when Λ approaches Λ_c. In particular, a characteristic cross-over point can be located when $\log(\tau)$ changes from a strong-glass former to power-law behaviour as described in section 7.1. We observe that the non-linear dependence on Λ of the parameters in (7.7) spurs quite different behaviour of the relaxation time as a function of temperature than is seen in the standard $F2$-model.

Solving $\lambda(\Lambda_c)/\mu(\Lambda_c) = 4$ for Λ_c, we find $\Lambda_c = \exp(\sqrt[3]{4}) \approx 4.89$. We use this analytical value to scale our variable to ϵ according to formula (7.2). Further, we can use this knowledge to set the right frame for our numerical computation of $\phi_1(t)$ for varying values of Λ and confirm the procedure's implementation[22].
3 To investigate the character of $\tau(\Lambda)$ (or $\tau(\epsilon)$) further and have a look at the possible shifting of the relaxation time (in particular, moving the cross-over point to high values of the relaxation time), we now propose models of the form $S(\Lambda) = \Lambda^\alpha \ln(\Lambda)$ where α is some fixed number for a particular model. To obtain a glass-forming system at some temperature, we must choose α such that λ/μ can attain the value 4, where $\tau(\Lambda)$ diverges. For these computations, we did not add 1 to the structure factor (although it would result in a more physical model, but we will include these concerns in the next section when we search for a really successful model). So here we have for arbitrary α:

$$\begin{aligned} \mu(\Lambda) &= \frac{1}{\Lambda^{1+\alpha}\ln(\Lambda)}, \\ \lambda(\Lambda) &= \frac{(\Lambda^\alpha \ln(\Lambda) - 1)^2}{\Lambda}. \end{aligned} \tag{7.8}$$

We start with $\alpha = 0.5$. Numerically solving $\lambda(\Lambda_c)/\mu(\Lambda_c) = 4$, we find $\Lambda_c \approx 3.467$. In figure 9, we see again the transition from a strong glass-former to a power-law with cross-over point, albeit at a lower value of τ than the previous case.

4 The next model uses the parameters of equations (7.8) with $\alpha = 0.1$. This yields $\Lambda_c \approx 6.765$. We see the same qualitative behaviour as for $\alpha = 0.5$ in figure 9, with the cross-over point lying at a slightly higher value of τ.

5 Hoping to further raise the τ-value at the cross-over point, we set $\alpha = 0.01$ in (7.8), for which $\Lambda_c \approx 9.610$. We still see the same qualitative behaviour as for $\alpha = 0.5$, with the cross-over point lying again at a higher value of τ than for the previous values $\alpha = 0.1$.

6 As a final test of the tunability of this model, we consider $\alpha = 0.0001$, corresponding to $\Lambda_c \approx 10.115$. We see that τ barely increases any more for this extreme decrease of our tuning parameter with respect to the situation $\alpha = 0.01$, so this model does not seem to reach the flexibility we seek.

7 As another tunable model, we consider a structure factor of the form $S(\Lambda) = \ln(\Lambda/(2\alpha))/\ln(\Lambda/(\alpha))$. Since this formula shows a diverging structure factor at $\Lambda = \alpha$, this is only sensible in the range $0 \leq \Lambda \leq \alpha$. In this range, S is also an increasing function of Λ. Further note that

$$\lim_{\Lambda\to 0} \frac{\ln(\Lambda/(2\alpha))}{\ln(\Lambda/\alpha)} = 1,$$

such that there is no reason to add any constant to the structure factor here. With equation (7.6), this model corresponds to the parameters

$$\begin{aligned} \mu(\Lambda) &= \frac{\ln(\Lambda/\alpha)}{\Lambda\ln(\Lambda/2\alpha)}, \\ \lambda(\Lambda) &= \frac{1}{\Lambda}\left(\frac{\ln(\Lambda/2\alpha)}{\ln(\Lambda/\alpha)} - 1\right)^2. \end{aligned} \tag{7.9}$$

We call this the LD model as an abbreviation for logarithm division for later reference. For a reasonable temperature range, we start with $\alpha = 20$, which corresponds to a divergence of τ at $\Lambda_c \approx 11.804$. Again for this model, the results are depicted in figure 9.

8 Next we use $\alpha = 40$ in equation (7.9) for which we have $\Lambda_c \approx 23.609$. We see that the cross-over point increases. We thus pursue increasing α for some more values.

[22] The knowledge of the divergence value of Λ is indeed a perfect means to make a grid of about 100 values for Λ. However, at the time of most computations, we did not yet realise this and therefore the exact grids may change for each choice of parameters. Since we are mostly qualitatively interested in the behaviour of the cross-over, we do not provide further details on the exact computations.

9 This model uses $\alpha = 200$, which yields $\Lambda_c \approx 118.043$.

10 We proceed for $\alpha = 2\,000$, corresponding to $\Lambda_c \approx 1\,180.43$.

11 Our next model is for $\alpha = 10\,000$ and corresponds to $\Lambda_c \approx 5\,902.14$.

12 As the last model based on the parameters given by (7.9), we use $\alpha = 20\,000$. In this case, divergence occurs at $\Lambda_c \approx 11\,804.3$.
It may at first be surprising that Λ_c seems to scale linearly with α. This can in fact be seen analytically, since $\mu(\Lambda)/\lambda(\Lambda)$ is a function of Λ and α such that only the ratio Λ/α appears in the expression. Hence the critical value Λ_c is found for a fixed value of Λ_c/α, which implies the linear scaling of Λ_c with α.

Further it is notable that with increasing α, we keep increasing the relaxation time at the cross-over point. In figure 9, we see an increase of τ of over approximately 3 orders of magnitude for α also varying in this range. We will look into the relation between the parameter α and the value of τ at the cross-over in the next section.

13 We consider two more models in this section, inspired by the apparent importance of the ratio λ/μ (at least for the critical value Λ_c). We ask ourselves what the effect would be of altering this ratio. We test this for the model described by (7.7). Specifically, we decrease[23] the relative value of λ such that

$$\begin{aligned} \mu(\Lambda) &= \frac{1}{\Lambda \ln(\Lambda)}, \\ \lambda(\Lambda) &= \frac{\ln(\Lambda)^2}{64\Lambda}. \end{aligned} \tag{7.10}$$

This leads to a value for the divergence temperature $\Lambda_c = \exp(\sqrt[3]{64 \cdot 4}) \approx 572.27$. The result for τ as a function of Λ is again shown in figure 9, at the label 13. We see that this model increases the relaxation time at the cross-over with respect to the model given by (7.7).

14 We finally investigate this rescaling of the parameters by adapting μ, setting

$$\begin{aligned} \mu(\Lambda) &= \frac{64}{\Lambda \ln(\Lambda)}, \\ \lambda(\Lambda) &= \frac{\ln(\Lambda)^2}{\Lambda}, \end{aligned} \tag{7.11}$$

such that the critical temperature is still found for $\Lambda_c = \exp(\sqrt[3]{64 \cdot 4}) \approx 572.27$. This rescaling leads again to a higher relaxation time at the cross-over, as can be seen in figure 9 at the label 14.

7.3 Scaling-inspired structure factors

In the previous section, tuning the ratio λ/μ seemed to be a powerful tool to influence the value of τ at the cross-over. In particular, we saw that decreasing this ratio for all Λ can make τ increase. In this section, we look at some more results of this method and use this idea to guide us to other models.

We first continue with the scaling of the model described by (7.7), now investigating μ and λ of the form

$$\begin{aligned} \mu(\Lambda) &= \frac{\alpha}{\Lambda \ln(\Lambda)}, \\ \lambda(\Lambda) &= \frac{\ln(\Lambda)^2}{\Lambda} \end{aligned} \tag{7.12}$$

for variable α. We call this the log model, since it makes use of the (natural) logarithm for the structure factor. In figure 10, we show the relaxation time as a function of ϵ for different values of α, resulting in curves that lay at higher τ values for bigger values of α. We also show the estimate of τ at the cross-over, which we denote by τ_{x}, versus α and see that this estimate also increases as a function of the parameter α. As a universal normalisation parameter of the relaxation time, we still use the value τ_{F2} of equation (7.4). The numerical results show that this approach to scaling the parameters λ and μ allows one to obtain values of τ_{x} ranging over 3 orders of magnitude for α between 1 and 200, which is what we are looking for. In this range, the critical value Λ_c varies from 4.9 to $1.1 \cdot 10^4$.

[23] We have observed that increasing the relative value of λ has the opposite effect. Since we wish to increase τ at the cross-over, we only show the case for which it does actually increase.

We note that the numerical values may not look exactly smooth, which is due to the estimates we make numerically as explained in section 7.1. However, the aim of the present work is not to improve on the predicted values; we believe that figure 10(a) sufficiently depicts the increase of τ_x with increasing α, wherever it may be located exactly.

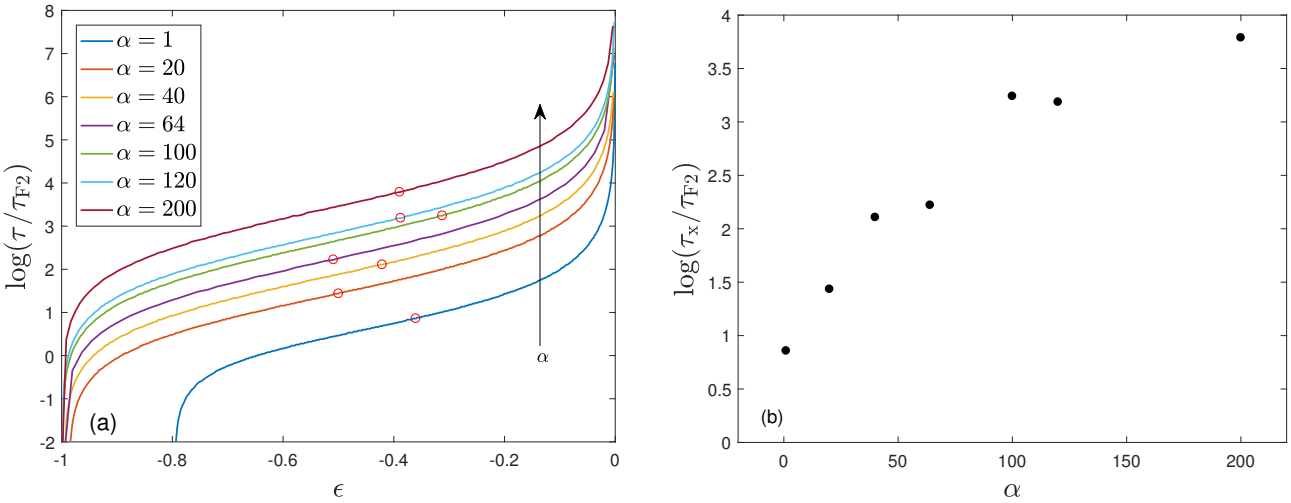

Figure 10: *Numerical results for the relaxation time τ predicted by schematic MCT for the log model with parameters given by (7.12), controlled by the parameter α. (a) τ as a function of scaled inverse temperature ϵ for varying α, including circles at the numerical estimates for the cross-over from super-strong behaviour to power-law divergence. (b) The value of τ at the cross-over (τ_x) versus α. We see that these various models have the ability to predict values of τ_x over a large time range. We do not elaborate on precise numerical values and errors, but only wish to establish a shift of τ_x to higher values.*

Our next aim is to see if we can relate this scaling to the structure factor. To this end, it is most useful to look at μ in (7.12) and note that we can alternatively write this as

$$\mu(\Lambda) = 1/(\Lambda \ln(\Lambda^{1/\alpha})).$$

This inspires us to choose for S some function which is purely controlled by a power of Λ, where the constant α seems an appropriate tool to influence the height of τ_x.

We first investigate this for a model that resembles the standard $F2$ model, setting

$$\begin{aligned} \lambda(\Lambda) &= \Lambda^{1/\alpha}, \\ \mu(\Lambda) &= 1, \end{aligned} \tag{7.13}$$

again for varying values of α. Due to its resemblance to the $F2$ model, we call it the Power $F2$ model. We display the results for τ as a function of ϵ and for τ_x as a function of α in figure 11. From (7.13), we see that the critical temperatures used for normalization of the data are given by ${\Lambda_c}^{1/\alpha} = 4$, or $\Lambda_c = 4^\alpha$. So for the results of figure 11, these vary from 8 to $1.2 \cdot 10^{24}$. If this is used in a real physical setting with a temperature in Kelvin, note that this value is very close to a temperature of absolute zero. We emphasize that the determination of τ_x is influenced by numerical errors that we do not analyse here, so the position of the circles in figure 11 is only indicative of an upward shift of the relaxation time at the cross-over point and should not be taken as an exact numerical result for these. We do believe, however, that the increase of τ_x with increasing α is clear from the graph. So we can conclude that the model (7.13) can predict the desired effect.

Next we try to combine the idea $S(\Lambda) = \Lambda^{1/\alpha}$ with the parameters λ and μ from equation (7.6) by looking at the functional forms

$$\begin{aligned} \lambda(\Lambda) &= \Lambda^{\frac{2}{\alpha}-1}, \\ \mu(\Lambda) &= \frac{1}{\Lambda^{1+\frac{1}{\alpha}}}, \end{aligned} \tag{7.14}$$

and we call this the PSF model as an abbreviation for Power Structure Factor model (we use $\alpha > 0$). For the expression for λ, we have again used $1 + S(\Lambda)$, because the substituted structure factor then approaches 1 as Λ approaches 0. For different values of α, it results in the τ behaviour of figure 12. We can again observe a

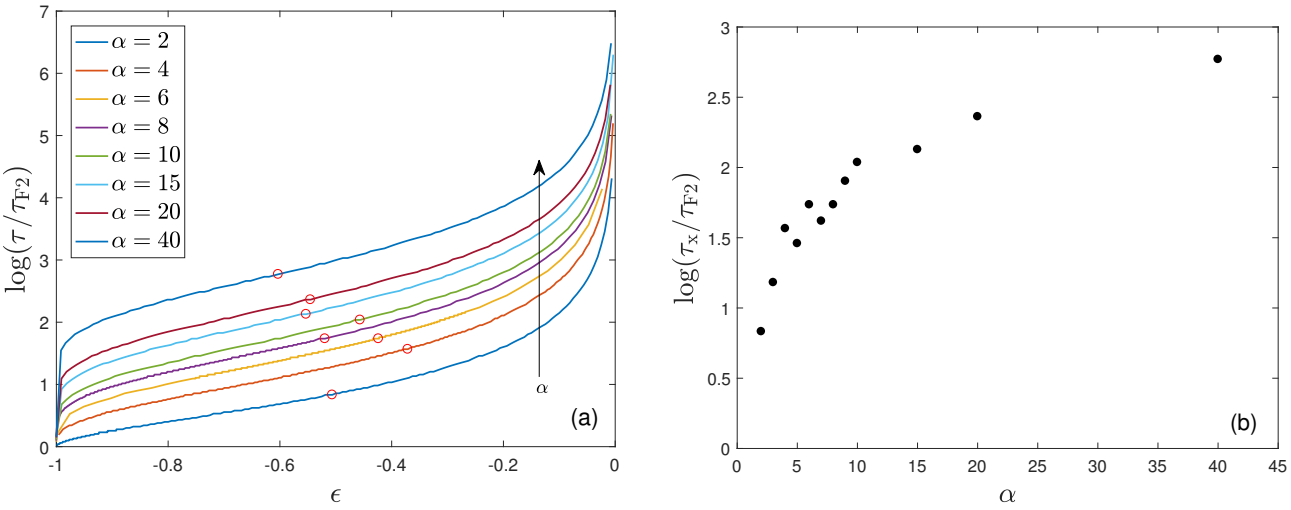

Figure 11: *Numerical results for the relaxation time τ predicted by schematic MCT for the power $F2$ model with parameters given by (7.13), controlled by the parameter α. (a) τ as a function of scaled inverse temperature ϵ for varying α, including circles at the numerical estimates for the cross-over from super-strong behaviour to power-law divergence. (b) The value of τ at the cross-over (τ_x) versus α. We see that these various models have the ability to predict values of τ_x over a large time range. The results indicate that τ_x is shifted upwards for increasing values of the parameter α.*

transition from super-strong to fragile glass formation and see that at the cross-over, τ_x again increases with increasing α. For the critical temperature in this model, we see that $\lambda(\Lambda_c)/\mu(\Lambda_c) = 4$ is satisfied for $\Lambda_c = 4^{(3/\alpha)}$. The results of figure 12 have values of Λ_c varying from 4 to $1.0 \cdot 10^4$. For the values of α here, we again observe the increasing trend of τ_x with increasing α. Although the numerical data should undergo an error analysis before a quantitative analysis can be made, our results do suggest that for the PSF model, $\log(\tau_x)$ might grow linearly as a function of α. This could be an interesting observation to motivate a more accurate numerical analysis. Perhaps most importantly, this growth has finally allowed us to obtain a a relaxation time that has grown 5 orders of magnitude with respect to the $F2$ model at $\Lambda = 0.02$. We expect that more orders can be reached when we further increase α.

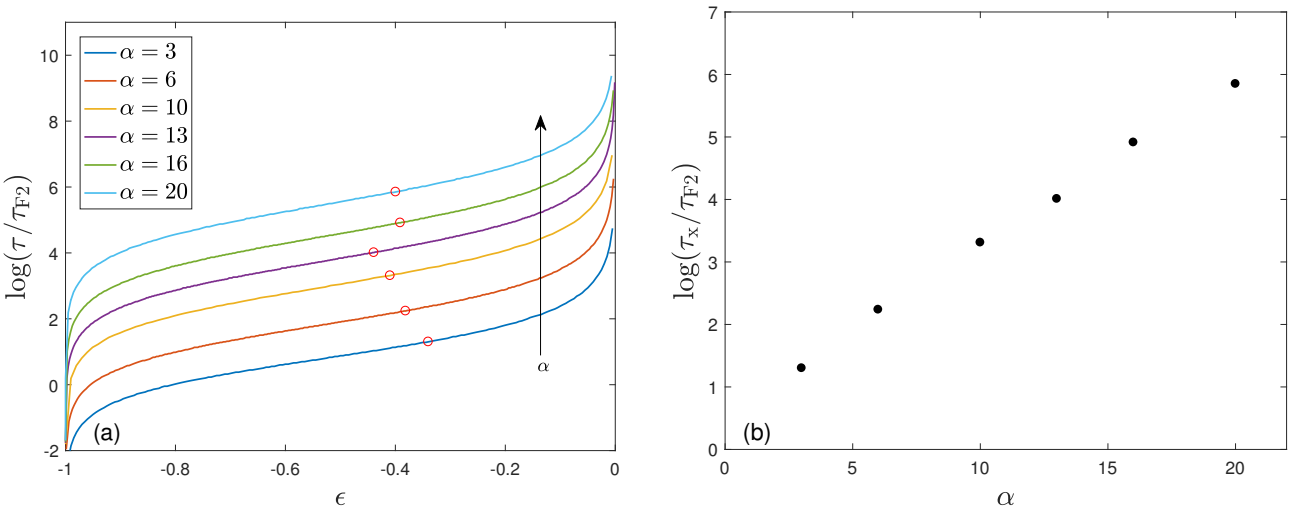

Figure 12: *Numerical results for the relaxation time τ predicted by schematic MCT for the PSF model with parameters given by (7.12), controlled by the parameter α. (a) τ as a function of scaled inverse temperature ϵ for varying α, including circles at the numerical estimates for the cross-over from super-strong behaviour to power-law divergence. (b) The value of τ at the cross-over (τ_x) versus α. The results indicate that τ_x is shifted upwards for increasing values of the parameter α.*

As a last consideration, we come back to the rather successful LD model of the previous section, given by equations (7.9). Looking at the structure factor as a function of Λ in figure 13(a) for $\alpha = 20$, we see that it

has a convex shape like $\Lambda^{1/\alpha}$ for small Λ, but for the greatest temperature range where it does not diverge, it is concave. As for the value of τ_x in this model, we see from figure 13(b) that $\log(\tau_x)$ seems to grow linearly with $\log(\alpha)$, or τ_x linearly with α. So this is another model that easily allows us to explore multiple orders of magnitude for the relaxation time at the cross-over point. We note however, that the divergence of this model is not physically consistent with the observation that the structure factor changes mildly upon temperature variations. Still it might be a useful model up to the point where divergence occurs, or to be corrected to become more physical.

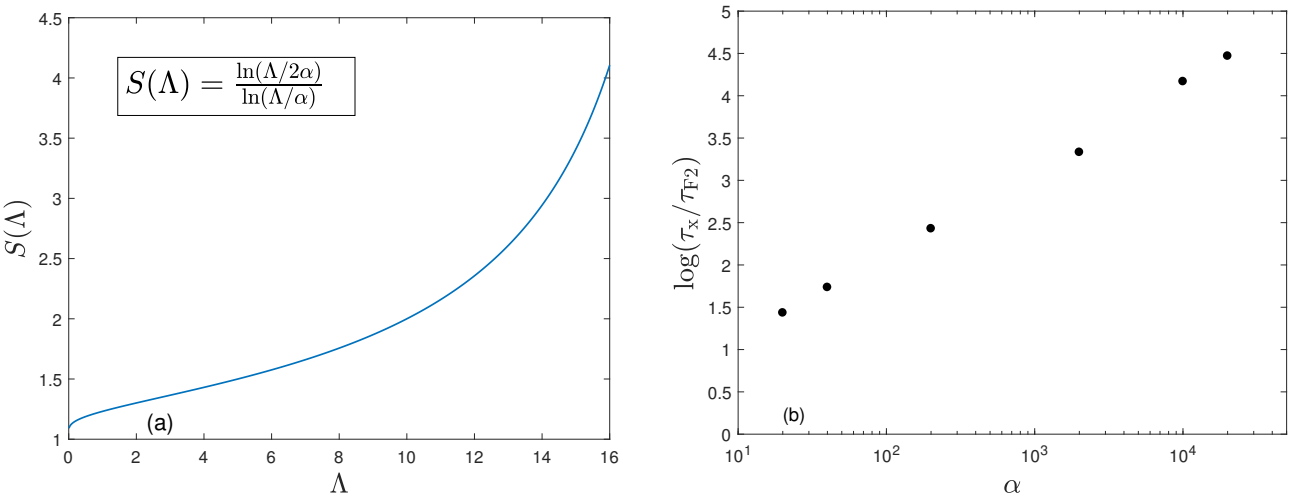

Figure 13: *A second look at the LD model, whose structure factor is indicated and plotted in panel (a) for $\alpha = 20$. With the parameters of equation (7.9), MCT predicts the values of τ_x of panel (b). We did no numerical error analysis, but the graph suggests a linear relationship between τ_x and the parameter α that is used to model different structure factors.*

7.4 A closer look at $F2$

In this section we aim to provide some perspective to the behaviour of the relaxation time that we have seen above. We would like to ensure that the super-strong growth of the relaxation time that arises for non-linear dependence of λ and μ on Λ is not a result of the temperature range we consider, for the results would be similar to the standard $F2$ model if we started at higher values of Λ. Therefore, we consider the model for $\mu = 100$ and $\lambda = \Lambda$, such that $\Lambda_c = 200$ and we increase the temperature range up to the divergence by two orders of magnitude while maintaining the $F2$ dependence on Λ.

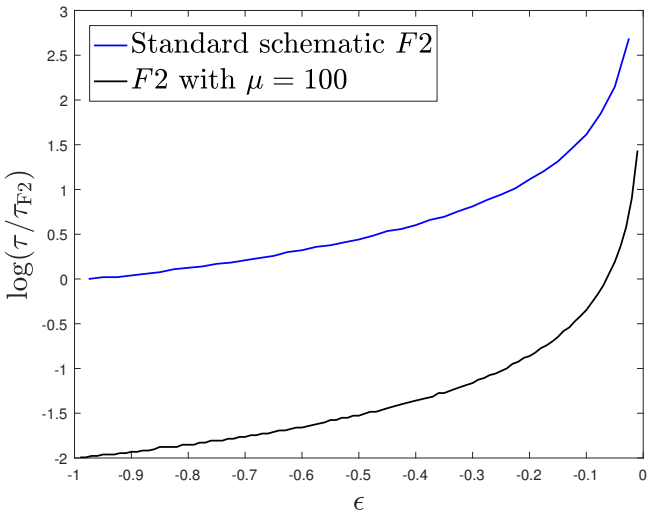

Figure 14: *Comparison between the $F2$ model and the same Λ dependence, but with $\mu = 100$. Since the shapes of τ as a function of ϵ are similar, we conclude that the super-strong glass prediction is not just a temperature window effect.*

In figure 14, we compare the results for this model with the standard $F2$ model (for which we have $\mu = 1$). We see the same form in the graph for $\log(\tau)$ as a function of ϵ and conclude that the super-strong behaviour of the relaxation time is not just a temperature window effect.

8 Physical interpretation

In this chapter, we first try to find a physical intuition to explain the results obtained in chapter 7 based on the role of the ratio λ/μ. Then we take a look at the intermediate scattering functions calculated with schematic MCT to see if there are also effects of the ratio λ/μ on this quantity.

It may be verified that all our modelling choices in chapter 7 come down to decreasing the ratio of $\lambda(\Lambda)/\mu(\Lambda)$ (at least from a certain value of Λ on) with increasing α. This is of course explicit in the choice of equations (7.12). In section 2.2 we have seen that the memory kernel represents a damping effect that acts on the intermediate scattering function. In schematic MCT, the strength of this effect is captured by the parameter λ. Similarly, μ comes in the place of the component of the frequency matrix, which one can compare to an oscillation frequency in the system. Therefore we can interpret a larger value of the ratio λ/μ at the same temperature as greater damping on the intermediate scattering function. Hence, the relaxation to 0 takes place on a shorter time scale and divergence of the relaxation time can be expected to happen when further lowering temperatures, as is consistent with the results of the previous chapter.

8.1 Intermediate scattering functions and time scaling

Here we try to understand the role of the ratio λ/μ and of scaling by looking at the normalized intermediate scattering function $\phi(t)$ directly. In figure 15, we compare such functions for different models, but the same ratio of λ and μ.

We choose the ratio to be fixed at $\lambda(\Lambda)/\mu(\Lambda) = 3.415$. For the standard $F2$ model, this means that we consider the solution for $\Lambda = 3.415$.

For the PSF model (equations (7.14), we set $\alpha = 10$ and look at the solution for $\Lambda = 3.415^{10/3}$ to obtain the correct ratio. We also compute the solution for $\Lambda = 13$ to have a comparison solution with a ratio of 2.16.

We also compare these to a few solutions of the log model (equations 7.12). Solving for the desired ratio yields the relation $\alpha = (\log(\Lambda))^3/3.415$. We therefore look at the solutions for $\Lambda = 3, 10$ and 100 and use the corresponding α value to ensure we can make a comparison at a fixed ratio of λ/μ. We note that it does not really matter from which models the ratios are derived. The resulting graphs are shown in figure 15.

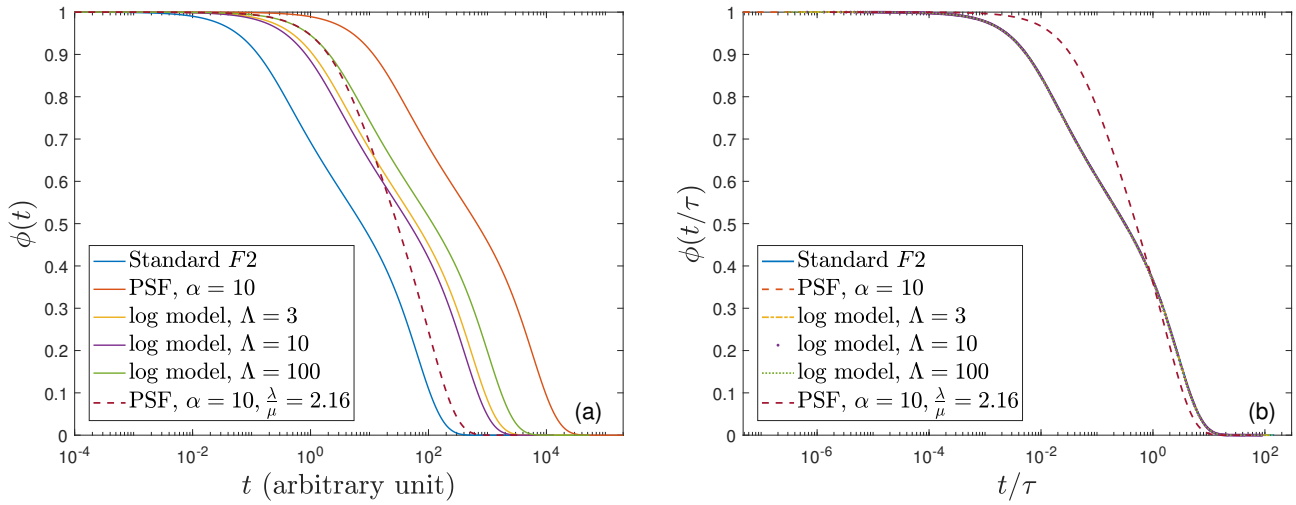

Figure 15: *Solutions $\phi(t)$ as a function of time t of the MCT equation for different parameter values but with a fixed ratio. (a) $\phi(t)$ on their real time scale, where it can be seen that different ratios result in similarly shaped graphs. Note that the deviating graph was obtained at a different ratio for extra comparison. (b) The solutions of panel (a) with the time scaled by the relaxation time for each. For a fixed ratio, we see that a master curve results that represents all solutions.*

For different parameter values, it is no surprise that different numerical outcomes of $\phi(t)$ are obtained. We see that the shape depends on the ratio used, but seems not to depend on the precise values of λ and μ. If we make use of a new time scale, where we divide by the relaxation time of each solution, this is confirmed. The curves for the same ratio display a data collapse. However, a definite observation that the rescaled solutions coincide can only be made analytically. If this is the case, this would imply that a single master curve is predicted by the schematic MCT equation for a fixed ratio of λ/μ and that the MCT predictions only differ in the associated time scales.

We can make a start based on a time scale invariance which relates to a friction prefactor for the first derivative in the MCT equation (3.1) (for the $N = 1$ case). We will derive this result in the context of our fixed ratio. We consider the MCT equation with a fixed ratio $r = \lambda/\mu$, i.e.

$$\dot{\phi}(t) + \mu\phi(t) + r\mu \int_0^t \phi(t-\tau)^2 \dot{\phi}(\tau)\,\mathrm{d}\tau = 0, \tag{8.1}$$

for some $\mu > 0$. This is equivalent to

$$\frac{1}{\mu}\dot{\phi}(t) + \phi(t) + r \int_0^t \phi(t-\tau)^2 \dot{\phi}(\tau)\,\mathrm{d}\tau = 0. \tag{8.2}$$

We are now interested in the solutions as a function of a rescaled time: $\psi(t) = \phi(t/t_\infty)$ for some time scale t_∞, or $\phi(t) = \psi(t \cdot t_\infty)$. Then we know, by the chain rule, that

$$\dot{\phi}(t) = \frac{d}{dt}\phi(t) = \frac{d}{dt}\psi(tt_\infty) = t_\infty \dot{\psi}(tt_\infty). \tag{8.3}$$

This allows us to transform (8.2) into an equation for ψ (we write $\tau = u$ in the integral for later convenience):

$$\frac{1}{\mu}t_\infty\dot{\psi}(tt_\infty) + \psi(tt_\infty) + r \int_0^t \psi((t-u)t_\infty)^2 \dot{\psi}(ut_\infty)t_\infty\,\mathrm{d}u = 0. \tag{8.4}$$

In the integral, we can now make the change of variables $\tau = t_\infty u$, corresponding to $\mathrm{d}u = t_\infty\,\mathrm{d}\tau$ and a change of the integration bounds to 0 and $t_\infty \cdot t$:

$$\frac{1}{\mu}t_\infty\dot{\psi}(tt_\infty) + \psi(tt_\infty) + r \int_0^{tt_\infty} \psi(tt_\infty - \tau)^2 \dot{\psi}(\tau)\,\mathrm{d}\tau = 0. \tag{8.5}$$

Hence we see that if $\phi(t)$ solves (8.1), then $\psi(t) = \phi(t/t_\infty)$ solves (replacing tt_∞ simply by t in equation (8.5))

$$\frac{t_\infty}{\mu}\dot{\psi}(t) + \psi(t) + r \int_0^t \psi(t-\tau)^2 \dot{\psi}(\tau)\,\mathrm{d}\tau = 0. \tag{8.6}$$

If now t_∞ is the relaxation time as was used for the scaling in figure 15(b) and the relaxation time and μ are related in such a way that $t_\infty/\mu = C(r)$, some constant only depending on the fixed ratio r, then we would indeed have that the rescaled curves all satisfy the same ratio-universal equation

$$C(r)\dot{\psi}(t) + \psi(t) + r \int_0^t \psi(t-\tau)^2 \dot{\psi}(\tau)\,\mathrm{d}\tau = 0. \tag{8.7}$$

However, a relation of the form $t_\infty/\mu = C(r)$ is non-trivial. Therefore, the result of figure 15 is useful in that it shows that a search to this relation would directly give the relaxation time based on the parameters μ and r, but would also enable us to predict the full relaxation based on only the one parameter r. One specific role of r is already known to us, namely the divergence of the relaxation time at the fixed value $r = 4$.

9 Conclusion

Mathematical research

In this report, we have taken a detailed look at existence and uniqueness of solutions to the coupled integro-differential equations of schematic GMCT, both in the over-damped limit and for the actual second-order system that has been derived (stated in chapter 3). For the over-damped limit, we have written the higher-dimensional extension to a system of equations of the proof given by Götze [10] and have elaborated almost all details of the proof that were omitted by Götze. The proof makes use of a linearised iteration variant of the equation of interest and then derives suitable estimates to prove uniform convergence of the iteration sequence one obtains to show (global) existence. The uniqueness part follows a standard argument from ordinary differential equations.

For the second-order system, we were able to provide full details of the proofs given by Saal [29] with the help of the preliminary theory. This allowed us to work towards the application of Theorem 5.3.2 to the operator $\mathcal{A}$ that describes the system, such that the main steps of the proof were based on Fréchet differentiation and the derivation of several integral inequalities. Invertibility of $\mathcal{A}$ followed, which yields the existence and uniqueness. We had to impose a growth condition on the kernel function m to obtain existence for arbitrary intervals of definition, which limits the applicability to practical (physical) systems. For other kernels, a local existence result is derived and uniqueness of such a solution follows for kernels m satisfying a Lipschitz condition. A second proof was also studied which again employed a linearised iteration equation after transforming the system into a first-order one. Using the integro-differential equation to obtain bounds on the iteration sequence, convergence was shown which established existence. Uniqueness was established in a similar manner as before.

Subsequently, we saw that the theorem in [10] was directly applicable to schematic GMCT and found existence of a unique solution for the exponential and mean-field closures. We then modified the equation studied in Theorem 5.3.1 by Saal in [29] and highlighted the modifications needed in the proof to account for all parameters that are used in schematic GMCT. This yields existence of a unique solution for the exponential closure. For the MF-$N(1N)$ closure, we could establish this with a kernel function that we redefined outside the region where physical solutions exist. This was necessary in order to respect the linear growth condition needed in the proof.

Research outlook

Regarding the type of integro-differential equations studied here, a result without the linear growth condition on the kernel used by Saal would be useful to be searched for in future research. For example, if we want to employ it directly to all types of mean-field closures that can be used in GMCT, any type of polynomial growth should be allowed. To this end, new estimates would be needed to show that the operator $\mathcal{A}$ is weakly coercive if one wishes to fit this into the proof of Theorem 5.3.1. It would further be interesting to look at uniqueness and existence results for the full GMCT equations, without the schematic approximation. In this situation, each correlation function ϕ_n becomes also a function of n wave numbers and the kernel depends on the correlation functions via an integral over wave vectors. However, the time-dependence is fully captured by the equations studied here, so the proofs of this report might already give a good starting point.

Of further interest are the solutions themselves (both for GMCT and for the schematic case). These are often found numerically in the physical practice and the work of Janssen *et al.* suggests that for increasing closure levels, the solution converges to the solution of the infinite hierarchy [15]. It would be powerful to have theorem that rigorously confirms this for the numeric schemes used, e.g. in [15] based on the algorithm of [9]. For some specific parameter choices in schematic GMCT, an analytic solution to the infinite hierarchy of equations is known [15]. It would also be very interesting to derive such results for other hierarchies in order to better understand the predictions that GMCT can make.

Physical research

We have studied a multitude of models for the parameters μ and λ for standard schematic MCT with which MCT predicts super-strong behaviour of the relaxation time with decreasing temperature before the always predicted power-law divergence. Relating the models to the static structure factor, we have found the PSF (power structure factor) model of the form $S(\Lambda) = 1 + \Lambda^{1/\alpha}$ ($\alpha > 0$). We obtained an increase of the relaxation time at the cross-over from super-strong to power-law behaviour over 5 orders of magnitude. Other successful models could be the log model or the LD (logarithm division) model, but they are physically less apt to describe a static structure factor.

The search for physically realistic models was started from the apparent importance of the ratio λ/μ that seems to control the divergence. A bigger ratio can be physically interpreted as stronger damping. This interpretation is consistent with the numerical results we found as we have seen that a change of the structure factor that results in a bigger ratio yields the prediction of faster relaxation. Further our numerical results suggest that for a fixed ratio λ/μ, a master curve for the solutions to schematic MCT is obtained when scaling

the time by the relaxation time. The scaling invariance of the schematic MCT equation with a replaced friction coefficient hints towards an explanation that describes for a fixed ratio r, a universal expression for the relaxation time $t_\infty = \mu C(r)$ for all μ. Such a dependence would enable us to strongly characterize MCT predictions.

Research outlook

In this work, the value of the relaxation time at the cross-over from super-strong to fragile glass formation was only estimated. For a good quantitative description of the glass formation, these methods could be improved and should be analysed to allow for error estimation and fits through the numerical results. Only then can the MCT predictions be interpreted quantitatively. In particular, the possible linear dependencies observed for (the logarithm of) the relaxation time at the cross-over and the model parameter for the PSF and LD model might then be investigated. Moreover, the results of the PSF model can be compared to vitrimer simulations to see if the desired effect of super-strong glass formation can quantitatively be fitted to those practical results.

Theoretically, the investigated models can be further explored by the methods of higher order GMCT. It is interesting to see if the convergence notions seen in previous work will also hold in these models with new qualitative behaviour, and what the effect on the cross-over point will be (would it disappear, leaving super-strong glass formation at the glass transition temperature?). In line with earlier research, one could e.g. as the simplest case set $\mu_n = n \cdot \mu$, weighing all wave numbers in GMCT on an equal footing, and keep all coefficients λ_n equal to the instances of λ (λ_1 in GMCT) considered here [15].

Finally the importance of the λ/μ ratio can be further pursued in numerical and theoretical work. The data collapse of the intermediate scattering function, depending on this ratio only, could be explored for more values of the ratio than has been done here. This might be done with the purpose of providing more convincing evidence for the data collapse, or to establish specific regimes for the ratio where the data collapse might hold. Further, numerical observations could be used to steer a search for the dependence of τ_∞ on this ratio. When this is pursued, future research should not overlook the physical relevance of the master curve. If there is a feasible explanation, it would be interesting to see if this can be taken to GMCT as well.

References

[1] U. Bengtzelius, W. Götze, and A. Sjölander. Dynamics of supercooled liquids and the glass transition. *Journal of Physics C: Solid State Physics*, 17:5915–5934, 1984.

[2] W. E. Boyce and R.C. DiPrima. *Elementary Differential Equations and Boundary Value Problems: International Student Version*. John Wiley & Sons, 10th edition, 2013.

[3] M. Capelot, M. Unterlass, F. Tournilhac, and L. Leibler. Catalytic Control of the Vitrimer Glass Transition. *ACS Macro Letters*, 1(7):789–792, 7 2012.

[4] P. Charbonneau, Y. Jin, G. Parisi, and F. Zamponi. Hopping and the Stokes-Einstein relation breakdown in simple glass formers. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42):15025–15030, 2014.

[5] W. Cheney. *Analysis for applied mathematics*, volume 208 of *Graduate Texts in Mathematics*. 2001.

[6] S. Ciarella and L.M.C. Janssen. Appendix: our F2. 2018.

[7] C. Corduneaunu. *Principles of Differential and Integral Equations*. Allyn and Bacon, Inc., Boston, Massachusetts, 1971.

[8] G. Doetsch and L. Debnath. *Introduction to the Theory and Application of the Laplace Transformation*. Springer-Verlag, Berlin, 1974.

[9] M. Fuchs, W. Götze, I. Hofacker, and A. Latz. Comments on the alpha -peak shapes for relaxation in supercooled liquids. *Journal of Physics: Condensed Matter*, 3(26), 1991.

[10] W. Götze and L. Sjögren. General Properties of Certain Non-linear Integro-Differential Equations. *Journal of Mathematical Analysis and Applications*, 195:230–250, 1995.

[11] G. Grippenberg, S.O. Londen, and O. Staffans. *Volterra Integral and Functional Equations*. Camridge University Press, Cambridge, 1990.

[12] B.C. Hall. *Lie Groups, Lie Algebras, and Representations*, volume 222 of *Graduate Texts in Mathematics*. Springer International Publishing, Cham, 2015.

[13] L.M.C. Janssen. Mode-Coupling Theory of the Glass Transition: A Primer. *Under review*, 2018.

[14] L.M.C. Janssen, P. Mayer, and D.R. Reichman. Relaxation patterns in supercooled liquids from generalized mode-coupling theory. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 90(5):1–8, 2014.

[15] L.M.C. Janssen, P. Mayer, and D.R. Reichman. Generalized mode-coupling theory of the glass transition: Schematic results at finite and infinite order. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(5), 2016.

[16] L.M.C. Janssen and D.R. Reichman. Microscopic Dynamics of Supercooled Liquids from First Principles. *Physical Review Letters*, 115(20):1–5, 2015.

[17] W. Kob. *Supercooled liquids, the glass transition, and computer simulations*. Les Houches 2002 Summer Sch. - Sess. LXXVII Slow Relaxations Nonequilibrium Dyn. Condens. Matter. 2002.

[18] W.A.J. Kosmala. *A Friendly Introduction to Analaysis*. Pearson Education, Inc., 2nd edition, 2004.

[19] E. Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley & Sons. Inc., 1989.

[20] E. Leutheusser. Dynamical model of the liquid-glass transition. *Physical Review A*, 29(5):2765–2773, 1984.

[21] P. Linz. *Analytical and Numerical Methods for Volterra Equations*. Studies in Applied and Numerical Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 1985.

[22] P. Mayer, K. Miyazaki, and D.R. Reichman. Cooperativity beyond caging: Generalized mode-coupling theory. *Physical Review Letters*, 97(9):1–4, 2006.

[23] D. Montarnal, M. Capelot, F. Tournilhac, and L. Leibler. Silica-Like Malleable Materials from Permanent Organic Networks. *Science*, 334(6058):965–968, 2011.

[24] J.R. Munkres. *Topology*. Prentice Hall, Inc., 2nd edition, 2000.

[25] D.R. Reichman and P. Charbonneau. Mode-coupling theory. *Journal of Statistical Mechanics: Theory and Experiment*, (5):267–289, 2005.

[26] T. Rowland. Cˆk function, 2015.

[27] W. Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, Inc., 3rd edition, 1976.

[28] M. Saal and S.O. Londen. Global existence and blow-ups for certain ordinary integro-differential equations. *Journal of Integral Equations and Applications*, 27(4):573–602, 2015.

[29] Martin Saal. Well-posedness and asymptotics of some nonlinear integro-differential equations. 2012.

[30] Joel L. Schiff. *The Laplace Transform*. Springer Science + Business Media , LLC, 2011.

[31] E. Suhubi. *Functional Analysis*. Springer, 2003.

[32] A.M. Wazwaz. *Linear and Nonlinear Integral Equations*. Higher Education Press and Springer-Verlag, 2011.

[33] D.V. Widder. *The Laplace Transform*. Princeton University Press, 1946.

[34] E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume II/B: Nonl. Springer Science + Business Media, New York, 1990.

Appendices

A Convolution and product rule

Here we elaborate a little on the convolution integral and prove the product rule for operator differentiation (Theorem 4.17).

Definition A.1. Let $n \in \mathbb{N}$ and let $F : [0,T] \to \mathbb{R}^{n\times n}$ and $v : [0,T] \to \mathbb{R}^n$ be continuous functions, for some $T > 0$. Then the **convolution** $F * v$ is defined as the continuous function $F * v : [0,T] \to \mathbb{R}^n$ by

$$F * v(t) = \int_0^t F(t-s)v(s)\,\mathrm{d}s.$$

Observation A.2. Due to linearity of the multiplication of matrices and vectors and of computing integrals, we readily see that for all $F, G : [0,T] \to \mathbb{R}^{n\times n}$ and $v, w : [0,T] \to \mathbb{R}^n$ and all $\lambda, \mu \in \mathbb{R}$ we have $(F+G) * v = F * v + G * v$, $F * (v+w) = F * v + F * w$ and $(\lambda F) * (\mu)v = (\lambda\mu)(F * v)$.

Observation A.3. With the Euclidean norm on $\mathbb{R}^n$ ($||.||$), the induced operator norm on $\mathbb{R}^{n\times n}$, and the supremum norm over $[0,T]$ denoted by $||.||_\infty$ for both, and the triangle inequality for integrals, we see that

$$||F * v(t)|| = \left|\left| \int_0^t F(t-s)v(s)\,\mathrm{d}s \right|\right| \leq \int_0^t ||F(t-s)|| \cdot ||v(s)||\,\mathrm{d}s = \int_0^t ||F||_\infty ||v||_\infty\,\mathrm{d}s = ||F||_\infty * ||v||_\infty.$$

The last step basically uses a convolution of two constant functions. This holds for all $t \in [0,T]$, so for brevity we write

$$||F * v|| \leq ||F||_\infty * ||v||_\infty.$$

Theorem A.4. (Product rule for convolution) Let $[a,b]$ be an interval in $\mathbb{R}$, $n \in \mathbb{N}$ and let $U = C^1([a,b], \mathbb{R}^n)$, $V = C([0,T], \mathbb{R}^{n\times n})$ and $W = C([0,T], \mathbb{R}^N)$. Let $A : U \to V$, $B : U \to W$ be bounded linear operators that are differentiable at $f \in U$ with derivatives denoted by A'_f and B'_f, respectively. Then also the convolution of the operators, $A * B : U \to W$, is differentiable at f and we have

$$(A * B)'_f = A'_f * B(f) + A(f) * B'_f. \tag{A.1}$$

Proof:
To be explicit, we use the $||.||_{(2)}$ norm on U, the Euclidean norm on the convolution integral and the induced operator norms to find estimates involving A'_f and B'_f when needed. We have for all $\Delta f \in U$:

$$\begin{aligned}
&||A(f+\Delta f) * B(f+\Delta f) - A(f) * B(f) - A'_f(\Delta f) * B(f) - A(f) * B'_f(\Delta f)|| = \\
&||\left\{A(f+\Delta f) - A(f) - A'_f(\Delta f)\right\} * B(f+\Delta f) + A(f) * B(f+\Delta f) + A'_f(\Delta f) * B(f+\Delta f) \\
&\qquad - A(f) * B(f) - A'_f(\Delta f) * B(f) - A(f) * B'_f(\Delta f)|| = \\
&||\left\{A(f+\Delta f) - A(f) - A'_f(\Delta f)\right\} * B(f+\Delta f) + A(f) * \left\{B(f+\Delta f) - B(f) - B'_f(\Delta f)\right\} \\
&\qquad + A'_f(\Delta f) * \{B(f+\Delta f) - B(f)\}||.
\end{aligned}$$

Hence with the observations above (where it is certainly allowed to replace the norm $||.||_\infty$ by the norm $||.||_{(2)}$, preserving the inequality) and the triangle inequality, we have for all $\Delta f \in U \backslash \{0\}$:

$$\begin{aligned}
&\frac{||A(f+\Delta f) * B(f+\Delta f) - A(f) * B(f) - A'_f(\Delta f) * B(f) - A(f) * B'_f(\Delta f)||}{||\Delta f||} \leq \\
&\frac{||A(f+\Delta f) - A(f) - A'_f(\Delta f)||}{||\Delta f||} * ||B(f+\Delta f)|| + ||A(f)|| * \frac{||B(f+\Delta f) - B(f) - B'_f(\Delta f)||}{||\Delta f||} \\
&\qquad ||A'_f|| \cdot \frac{||\Delta f||}{||\Delta f||} * ||B(f+\Delta f) - B(f)||.
\end{aligned}$$

By the definition of the Fréchet derivative, we can make use of the fact that

$$\lim_{||\Delta f|| \to 0} \frac{||A(f+\Delta f) - A(f) - A'_f(\Delta f)||}{||\Delta f||} = 0,$$

and likewise for B. The differentiability at f further ensures that B is continuous at f, so

$$\lim_{||\Delta f||\to 0} B(f+\Delta f) = B(f).$$

We note that $||A(f)||$, $||B(f)||$ and $||A'_f||$ are fixed numbers (for fixed f) and letting $||\Delta f|| \to 0$, the continuity of the norm and the convolution allows us to take all limits above separately and it follows that

$$\lim_{||\Delta f||\to 0} \frac{||A(f+\Delta f) * B(f+\Delta f) - A(f) * B(f) - A'_f(\Delta f) * B(f) - A(f) * B'_f(\Delta f)||}{||\Delta f||} = 0.$$

This proves (A.1). □

B Appropriate bound for matrices

Here we show why for the matrix A defined on page 47, a bound of the form

$$||e^{At}|| \leq e^{-c_0 t}$$

as defined in Lemma 4.7.6 is insufficient if we adhere to the norm induced by the Euclidean norm on $\mathbb{R}^2$, but should rather be replaced by one of the form

$$||A|| \leq be^{-c_0 t},$$

with some $b > 1$. In the case of a diagonal matrix we could do with $b = 1$ in observation 4.7.8 since then $b = ||D|| \cdot ||D^{-1}|| = ||I|| \cdot ||I|| = 1$.
In this appendix, we give a numerical example for the case $t = 1$.

Let $\lambda = 0.2$, so

$$A = \begin{pmatrix} 0 & 1 \\ -5 & -5 \end{pmatrix}.$$

This gives for A approximately the eigenvalues $\lambda_1 = -3.32$ and $\lambda_2 = -1.38$, and for e^A the eigenvalues $e^{\lambda_1} = 0.03$ and $e^{\lambda_2} = 0.25$, so $\lambda_{\max} = 0.25$. Further let us behold the vector $v = (2, 3)$. We then have with the Euclidean norm

$$\frac{||e^A v||}{||v||} = 0.4.$$

This clearly violates the inequality obtained in Lemma 4.7.6. Hence a scaling constant $b > 1$ is indeed needed here to bound the induced operator norm, as introduced in Lemmas 4.7.7 and 4.7.8.

C Fréchet derivative of matrix-valued functions

After definition 4.6.2, we have seen that the Fréchet derivative of vector-valued functions coincides with the Fréchet derivative. To see how Fréchet differentiation generalizes to matrix-valued functions, let us first look at the Jacobian matrix for a slightly different point of view than that of a matrix.

Let $n, m \in \mathbb{N}$ and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function on all of $\mathbb{R}^n$. We denote partial derivatives of the kth component function with respect to the lth coordinate in a point $y \in \mathbb{R}^n$ as $\partial_l f_k(y)$, which are the elements of the Jacobian matrix J_y. Then the value of the Fréchet derivative Df_y of f in y (which is the linear operator corresponding to the Jacobian matrix) evaluated at some vector $u = (u_1, \ldots, u_n) \in \mathbb{R}^N$ is given by

$$Df_y(u) = J_y u = \sum_{i=1}^{n} u_i Df_y(e_i) = \sum_{i=1}^{n} u_i \begin{pmatrix} \partial_i f_1(y) \\ \vdots \\ \partial_i f_m(y) \end{pmatrix}$$

if we denote by $e_1, \ldots, e_n$ the unit vectors of $\mathbb{R}^n$ that correspond to the coordinate system in which f is defined. The linearity of the Fréchet derivative thus manifests itself in taking a linear combination of partial derivatives of the function f at y.

We can use the same linearity for the Fréchet derivative of a matrix-valued function. Let $m : \mathbb{R}^n \to \mathbb{R}^{m\times l}$ be such a function with component functions $m_{ij} : \mathbb{R}^n \to \mathbb{R}$ which are all continuously differentiable ($i = 1, \ldots, m$,

$j = 1, \ldots, l)$. Then we claim that the Fréchet derivative $Dm_y(u)$ of m at a vector $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ evaluated at a vector $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ is given by

$$Dm_y(u) = \sum_{k=1}^{N} u_k \begin{pmatrix} \partial_k m_{11}(y) & \ldots & \partial_k m_{1l}(y) \\ \vdots & \ddots & \vdots \\ \partial_k m_{m1}(y) & \ldots & \partial_k m_{ml}(y) \end{pmatrix}, \tag{C.1}$$

which is a linear combination of partial derivatives of matrices. To prove this, we need to show that

$$\frac{||m(y+u) - m(y) - Dm_y(u)||}{||u||} \to 0 \text{ as } ||u|| \to 0, \tag{C.2}$$

for the appropriate norms that one has chosen on the vector spaces $\mathbb{R}^n$ and $\mathbb{R}^{m \times l}$. Certainly, if each component in the expression $\frac{1}{||u||}(m(y+u) - m(y) - Dm_y(u)) \to 0$, the same will hold true for the norm in (C.2). Applying (C.1) he (i, j)th component here reads

$$\frac{1}{||u||} \left(m_{ij}(y+u) - m(y) - u \cdot \nabla m_{ij}(y) \right),$$

where ∇ denotes the gradient operator in Cartesian coordinates. From the directional derivative from basic multivariate calculus, we know that this expression goes to 0 as $||u|| \to 0$ for all i, j. So we have shown that (C.1) indeed gives the Fréchet derivative of the matrix-valued function m at y evaluated at the vector u and all we need for m to be Fréchet differentiable is thus continuous partial differentiability of its component functions.