



Utrecht University

METEOROLOGY,
PHYSICAL OCEANOGRAPHY
AND CLIMATE

MATHEMATICAL
SCIENCES



MATHEMATICAL
INSTITUTE (MI)

Justification of Delay Equations in Climate Models

Author:
S.K.J. Falkena

Supervisors:
prof. dr. ir. J.E. Frank (MI)
prof. dr. ir. H.A. Dijkstra (IMAU)

June 29, 2018

Abstract

A new technique to derive delay models from systems of partial differential equations is investigated. This technique is based on the Mori-Zwanzig formalism, which gives a formal rewriting of the system using a projection onto a set of resolved variables. The rewritten system contains a memory term. The computation of this memory term requires the solving of the so-called orthogonal dynamics equation, which represents the unresolved dynamics. Finding an accurate solution to this equation is crucial in the application of the formalism. Here, the new technique is applied to a two-strip model of the El Niño Southern Oscillation. A mathematical derivation of a delay differential model, using the Mori-Zwanzig formalism and an alternative (exact) method based on variation of constants, is given. The derived delay model contains an additional term compared to a previously proposed conceptual model. This new term leads to a higher period of the model, which is closer to that seen in data. The Mori-Zwanzig formalism turns out to be not necessary to arrive at the resulting equations. Furthermore, the technique is applied to a model of the Atlantic Multidecadal Oscillation. This results in a delay difference model for the phenomena. In addition to this result, which can also be obtained by integration along characteristics, error terms for a smoothing approximation of this delay difference system have been derived.

Contents

1	Introduction	2
2	Mori-Zwanzig Formalism	4
2.1	Linear Illustration	4
2.2	Derivation of the Formalism	5
2.3	Orthogonal Dynamics Equation	6
2.4	Memory Kernel	7
3	El Niño Southern Oscillation	8
3.1	Delay Model by Suarez and Schopf	9
3.2	Linear Two-Strip Model	10
3.2.1	Model Formulation	10
3.2.2	Mori-Zwanzig Formalism	11
3.2.3	Characteristics	14
3.2.4	Delay Model	16
3.3	Nonlinear Two-Strip Model	17
3.3.1	Model Formulation	17
3.3.2	Mori-Zwanzig Formalism	19
3.4	Delay Model Analysis	21
3.4.1	Bifurcation Analysis	22
3.4.2	Parameter Dependence	25
4	Atlantic Multidecadal Oscillation	27
4.1	AMO Model	28
4.1.1	Model Formulation	28
4.1.2	Background Overturning Circulation	31
4.2	Mori-Zwanzig formalism	32
4.2.1	Discretization	33
4.2.2	Projection	34
4.2.3	Noise and Memory Terms	35
4.2.4	Limit Behaviour	39
4.3	Delay Model	41
5	Summary, Discussion and Conclusion	43
A	ENSO: Discrete Mori-Zwanzig Formalism	48

1 Introduction

Modern climate models that are used to run global simulations are much more complex than two decades ago. An important reason for this is the increased available computing power. This allows for the models to be run at a higher spatial resolution, as well as for more processes, which happen at different timescales, to be included. These computations require a better understanding of the physics behind climate phenomena, to allow for a suitable parametrization in the models. To get a better understanding of the physical mechanisms and reduce parameter uncertainty, conceptual climate models are needed.

Conceptual climate models contain only the necessary features for a phenomenon to occur and thus represent the dominant physics. As a result these models are relatively simple and allow for mathematical analysis. A useful way in which these models can be analysed is through bifurcation analysis. This analysis allows for the distinction of different state regimes and possible other modes of variability. The dependence on the few parameters present can be investigated relatively easy compared to high complexity models. One type of a conceptual climate model is a delay model. This type of model is particularly useful as a conceptual climate model, since delay models are infinite-dimensional, meaning they can be quite complex, but only contain a limited number of variables. This allows for an easier mathematical treatment, while these models still represent the physics in a realistic manner.

Many processes in climate contain positive and negative feedbacks. Some of these feedbacks are delayed, for example by transport through an ocean basin. The presence of such a delayed feedback can sometimes be determined from data [33]. In delay models it is not necessary to resolve all of the processes involved in the feedback. A parametrization by the resulting delay time is sufficient. Therefore it is useful to be able to derive delay models from more complex climate models. This can result in a model which can be fully analysed, compared to models which can only be studied by numerical simulation.

In 2017 Keane et al. provided an overview of delay models used to describe climate processes [27]. The two main areas in the climate system on which delay models so far have focused are Energy Balance Models and models for the El Niño Southern Oscillation (ENSO). Especially for ENSO there is extensive literature on both the oscillation itself, as well as its timing during the year, using delay models. ENSO is a well understood climate phenomenon in the equatorial Pacific Ocean. During an El Niño event, the sea surface temperature in the eastern part of the basin is warmer than usual, resulting in severe weather disturbances in countries at both sides of the Pacific Ocean. Its counterpart is La Niña, when the sea surface temperature is colder than usual. These two events alternate with intermediate phases in between, resulting in an oscillation with a period of four to seven years.

One of the most successful models of ENSO is that by Zebiak and Cane (1987) [41]. The view of the behaviour of ENSO in terms of normal modes resulted in the so-called recharge-discharge oscillator mechanism of ENSO [9, 24, 25]. This mechanism provides insight in a delay mechanism present in the physics of the ENSO process. The delay mechanism here is related to the propagation of equatorial and off-equatorial waves. These waves need time to travel through the basin, resulting in a delayed arrival of a temperature signal. Already in 1988 Suarez and Schopf proposed a delay model to describe the dynamics of ENSO based on this delay mechanism [36].

As mentioned above, the underlying physical mechanism of ENSO is well established and there is a delay model describing its dynamics at a reasonably accurate level. This makes ENSO an ideal phenomenon to test a new method for deriving delay equations. In this method the Mori-Zwanzig formalism is used to derive delay equations starting from more complex models [5]. Both the initial complex model, as well as possible resulting delay models, are well known in the case of ENSO and have been studied extensively. The application of this method can provide a new justification for the use of delay models to describe the dynamics of ENSO.

In many processes in climate propagating waves play an important role. This can be an indication that delay models can be used in more instances than has been done so far [27], since in ENSO this wave propagation is the origin of the delay. Another climate process governed by propagating waves is the Atlantic Multidecadal Oscillation (AMO). This is an oscillation in sea surface temperature in the North Atlantic Ocean with a period of fifty to seventy years. The mechanism behind the AMO has been described by Te Raa and Dijkstra (2002) [39]. The sea

surface temperature mode of the AMO relies on the transport of temperature through thermal wind balance. Baroclinic instability plays an important role in the interaction of sea surface temperature modes in the North Atlantic Ocean [6, 11]. Low-order models of the AMO have been studied by, among others, Broer et al. [4], and recently by Sévellec and Huck [34]. In contrast to ENSO, there is no known physical mechanism for the AMO which would result in a delay. There is no localised feedback mechanism present in the AMO, while there is in the physics of ENSO.

Usually delay models for climate systems are derived from more complex models by making physical assumptions about the system [28]. This way, it is physically argued that there is a delay of a certain form, resulting in a simplified delay model. Battisti and Hirst have done so to arrive at a delay model for ENSO, corresponding on first order to the model stated by Suarez and Schopf [1, 2]. The Mori-Zwanzig formalism could provide a way to give this derivation of delay models a stronger mathematical foundation [5]. For the AMO no such physical justification for a possible delay model exists.

The Mori-Zwanzig formalism gives a formal rewriting of a system of ordinary differential equations with the goal of reducing the number of dependent variables [19]. The result is a system for only a select set of resolved variables. The unresolved variables are no longer present in the resulting system due to a projection onto the resolved variables. The formalism often is applied to systems with large scale differences [40], which in most cases have a Markovian parametrization [3]. Also for non-Markovian systems some results exist [8]. Hamiltonian systems are another class of systems to which the formalism has been applied [5]. The component of the rewritten system that is focused on in this thesis is the memory term. This term is the integral over a memory kernel, which contains the history of the system. This hints at the idea that, under some approximations, this memory term can be simplified to a term with a delay, since a delay also represents the history of the system.

In this thesis the Mori-Zwanzig formalism is applied to both an ENSO model and an AMO model. The ENSO model is used as a test case for the derivation of delay models using the formalism, since the physics here clearly indicate the presence of a delay mechanism. For the AMO model it is not yet clear whether a delay is present. In the process of deriving possible delay models, the necessity of using the Mori-Zwanzig formalism is always reviewed. Other mathematically based methods to arrive at the same result are considered.

In Section 2 the theory behind the Mori-Zwanzig formalism is discussed. In Section 3 and 4 the formalism is applied to models of both the ENSO and AMO. In these sections first the models are explained. For ENSO the model by Jin (1997) is used [24], while for the AMO the more recent model by Sévellec and Huck (2015) is considered [34]. After the discussion of the models, the Mori-Zwanzig formalism is applied to these systems of equations. The possible resulting delay models for the ENSO and AMO are then studied in some detail. In Section 5 the derived ENSO and AMO models and their behaviour are discussed and compared.

2 Mori-Zwanzig Formalism

The mathematical foundation used in this thesis for deriving delay equations is the Mori-Zwanzig formalism. Here the approach of Chorin et al. (2002) is followed [5]. They were the first to propose this way of applying the formalism. Their method is based on the work by Mori (1965) [30], and Zwanzig (1973) [42]. The Mori-Zwanzig formalism gives a formal rewriting of a set of ordinary differential equations to a reduced system for the resolved variables, which still captures all of the dynamics of the system. For this rewriting to be a valuable and realistic reduction of the original model, often approximations to some terms in the reduced system are needed.

This section starts by discussing a linear example to illustrate the Mori-Zwanzig formalism in Section 2.1. Then a derivation of the formalism is given in Section 2.2. After this derivation, in Section 2.3, an approximation to the orthogonal dynamics system for the unresolved variables is given. Lastly, in Section 2.4 an approximation for the memory kernel is derived.

2.1 Linear Illustration

To gain insight in the ideas and arguments for applying the Mori-Zwanzig formalism, a linear example is given. Here, the work by Gouasmi et al. is followed [20]. Consider the following linear system of ordinary differential equations:

$$\begin{aligned}\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2, \\ \frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2,\end{aligned}\tag{1}$$

with $x_i(t) : \mathbb{R} \rightarrow \mathbb{R}$ continuous and parameters $a_{ij} \in \mathbb{R}$ for $i, j = 1, 2$. The goal is to arrive at an equation for only a chosen set of resolved variables. Finding such an equation reduces the number of variables that have to be considered when solving the system, likely simplifying the computations. Here x_1 is chosen as the resolved variable, so the unresolved variable is x_2 . In the linear case of Equation (1) the formal solution of the equation for the unresolved variable x_2 is

$$x_2(t) = e^{a_{22}t}x_2(0) + \int_0^t a_{21}e^{a_{22}(t-s)}x_1(s)ds.\tag{2}$$

Here, the first term is the solution to the homogeneous equation for the unresolved variable. The second term is the particular solution obtained by the method of variation of constants [17]. Substituting this formal solution for x_2 into the equation for x_1 yields

$$\frac{dx_1}{dt} = a_{11}x_1(t) + a_{12}e^{a_{22}t}x_2(0) + a_{12} \int_0^t a_{21}e^{a_{22}(t-s)}x_1(s)ds.\tag{3}$$

Now the system in Equation (1) has been reduced to one equation for the resolved variable x_1 without losing the dynamical behaviour. The rewritten system is exact, since no approximations have been made in the derivation process.

The right-hand side of this reduced equation consists of three terms. The first term is the dependence on the resolved variable x_1 that was already present in the initial equation. It is referred to as the Markovian term, since it does not depend on the initial conditions or history of the system. The second term is the dependence on the initial conditions of the unresolved variable x_2 . Note that this term does not depend on x_1 , meaning its evolution is fixed by the homogeneous solution for x_2 . In statistical physics the unresolved dynamics often represents the fast (Brownian) motion, while the resolved dynamics give the main direction of the movement of particles [19]. The fast motion is often parametrized as noise in those cases. Therefore this term is called the noise term. The last term is an integral over the history of the resolved variable x_1 , weighted by some function dependent on the form of the homogeneous solution for the unresolved variable x_2 . Because of the dependence on the history of x_1 , this term is often referred to as the memory term and the component in front of x_1 in the integral as the memory kernel. In the derivation of delay equations using the Mori-Zwanzig formalism the focus is on this memory term.

2.2 Derivation of the Formalism

The methods employed in Section 2.1 to arrive at the final reduced equation work only for linear systems. In this section a similar result, which is also valid for nonlinear systems, is derived. The linear example discussed in the previous section gives an idea of the result that is expected. This anticipated result is an equation split up in a Markovian, noise, and memory term. The derivation given below is based on that by Chorin et al. [5].

Consider the following n -dimensional system of ordinary differential equations:

$$\frac{d}{dt}\phi(t) = R(\phi(t)), \quad \phi(0) = x_0, \quad (4)$$

where $\phi(t) \in \mathbb{R}^n$ is a continuous function of $t \in \mathbb{R}$, $x_0 \in \mathbb{R}^n$ is a variable representing the initial conditions and $R : \mathbb{R}^n \rightarrow \mathbb{R}^n$ has components R_i . For every initial condition x_0 there is a corresponding trajectory $\phi(t) = \phi(x_0, t)$, $\phi : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$. The flow map is $\Phi : x_0 \mapsto \phi(x_0, t)$. With this notion, the system of Equation (4) is equivalent to the Liouville equation for $u : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}^n$:

$$\frac{\partial}{\partial t}u(x_0, t) = \mathcal{L}u(x_0, t), \quad u(x_0, 0) = g(x_0), \quad (5)$$

for $g(x_0) = x_{0i}$. This way the solution is $u(x_0, t) = \phi_i(x_0, t)$, which is the i -th component of Equation (4). Here $\mathcal{L} = \sum_{i=1}^n R_i(x_0)\partial_{x_{0i}}$ is the Liouville operator. The proof of this equivalence can be found in the article by Chorin et al. [5]. The Liouville equation is mostly used in statistical and Hamiltonian mechanics [38]. The equation describes the evolution of the density in the system around any given point moving through phase-space. The Liouville operator is also known as the generator function. In statistical systems this can be related to infinitesimal generator matrix [31].

The goal, as for the linear system, is to find a system of equations for a select set of m resolved variables $\hat{\phi} \in \mathbb{R}^m$. The unresolved variables are denoted by $\tilde{\phi} \in \mathbb{R}^{n-m}$, such that $\phi = (\hat{\phi}, \tilde{\phi})$. To reduce the system from n components to the desired m components a projection operator $P : C(\mathbb{R}^n, \mathbb{R}) \rightarrow C(\mathbb{R}^m, \mathbb{R})$ is needed. An example of such a projection operator is the orthogonal projection:

$$(Pf)(\hat{x}_0) = \frac{\int f(x_0)\rho(x_0)d\tilde{x}_0}{\int \rho(x_0)d\tilde{x}_0}, \quad (6)$$

for $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and where $\rho : \mathbb{R}^n \rightarrow \mathbb{R}$ is a probability density function. Note that here integration is only over the unresolved variables. This projection operator is often applied to Hamiltonian systems [5, 20]. Another projection operator is the linear projection, defined by $P(f(\hat{x}_0, \tilde{x}_0)) = f(\hat{x}_0, 0) \equiv \hat{f}(\hat{x}_0)$, which sets all unresolved variables to zero and retains only the resolved components. The converse of P is $Q = I - P$. The choice of projection is important, since it determines the final result.

For a resolved variable one has $\phi_i(x_0, t) = e^{t\mathcal{L}}x_{0i}$ by Equation (5), where $e^{t\mathcal{L}}$ is the transfer operator corresponding to the Liouville equation. For this resolved variable Equation (5) can be written as

$$\frac{d}{dt}e^{t\mathcal{L}}x_{0i} = e^{t\mathcal{L}}\mathcal{L}x_{0i} = e^{t\mathcal{L}}P\mathcal{L}x_{0i} + e^{t\mathcal{L}}Q\mathcal{L}x_{0i}, \quad (7)$$

since \mathcal{L} and $e^{t\mathcal{L}}$ commute. Now the second term in the right-hand side of this equation is considered. This component gives the evolution of the unresolved variables. The Dyson formula is applied to find an expression for this term. This formula is [16]

$$e^{t(A+B)} = e^{tA} + \int_0^t e^{(t-s)(A+B)}Be^{sA}ds. \quad (8)$$

For $e^{t\mathcal{L}} = e^{t(P+Q)\mathcal{L}}$ this yields

$$e^{t\mathcal{L}} = e^{tQ\mathcal{L}} + \int_0^t e^{(t-s)\mathcal{L}}P\mathcal{L}e^{sQ\mathcal{L}}ds. \quad (9)$$

Now apply this operator to $Q\mathcal{L}x_{0i}$ to get

$$e^{t\mathcal{L}}Q\mathcal{L}x_{0i} = e^{tQ\mathcal{L}}Q\mathcal{L}x_{0i} + \int_0^t e^{(t-s)\mathcal{L}}P\mathcal{L}e^{sQ\mathcal{L}}Q\mathcal{L}x_{0i}ds. \quad (10)$$

Substituting this result into Equation (7) yields the generalized Langevin equation:

$$\begin{aligned} \frac{\partial}{\partial t}\phi_i(x_0, t) &= e^{t\mathcal{L}}P\mathcal{L}x_{0i} + F_i(x_0, t) + \int_0^t e^{(t-s)\mathcal{L}}K_i(\hat{x}_0, s)ds \\ &= R_i(\hat{\phi}(x_0, t)) + F_i(x_0, t) + \int_0^t K_i(\hat{\phi}(x_0, t-s), s)ds, \end{aligned} \quad (11)$$

where it is defined that $R_i(\hat{\phi}(x_0, t)) = (PR_i)(\phi(x_0, t))$ and with

$$F_i(x_0, t) = e^{tQ\mathcal{L}}Q\mathcal{L}x_{0i}, \quad K_i(\hat{x}_0, t) = P\mathcal{L}F_i(x_0, t). \quad (12)$$

Note that $F_i(x_0, t)$ is the solution to the so-called orthogonal dynamics equation:

$$\frac{d}{dt}F_i(x_0, t) = Q\mathcal{L}F_i(x_0, t), \quad F_i(x_0, 0) = Q\mathcal{L}x_{0i}. \quad (13)$$

Comparing the Langevin equation with the linear result of Equation (3), there is the Markovian term $R_i(\hat{\phi}(x_0, t))$, the noise term $F_i(x_0, t)$ and the memory term as integral over the memory integrand $K_i(\hat{\phi}(x_0, t-s), s)$. This memory integrand consists of a memory kernel applied to the resolved variables.

This rewriting of the system is exact, but that does not mean it is useful. If solving the orthogonal dynamics equation is just as difficult as solving the full system, there is no use in applying the formalism. The applicability thus depends on the system and whether a suitable projection exists. Such a projection would yield an orthogonal dynamics system which is relatively straightforward to solve or approximate in a good way. Applications to slow-fast [40], Markovian [3], non-Markovian [8], and Hamiltonian systems [5], as well as systems with an orthogonal basis of eigenfunctions [37], have been considered in literature [19]. For non-Hamiltonian systems of partial differential equations, which are studied in this thesis, less literature is available.

2.3 Orthogonal Dynamics Equation

The main difficulty in the applicability of the Mori-Zwanzig formalism is the solving of the orthogonal dynamics system of Equation (13). It is important that this equation is easier to solve than the original system. The choice of projection is an important factor in determining the form and complexity of the orthogonal dynamics equation. The projection needs to be chosen in such a way that the orthogonal dynamics system is stable and less involved than the original system. Alternatively, it has to be possible to accurately approximate the orthogonal dynamics equation by a stable and less involved system. Preferably, the orthogonal dynamics system decays at a faster rate than the full system. In that case it can be justified to neglect the noise term in the Langevin equation. For linear systems this means that the eigenvalues of the orthogonal dynamics system are more negative than those of the full system.

To simplify the issue of solving the orthogonal dynamics system, Gouasmi et al. derived the pseudo-orthogonal dynamics equation [20]. Under certain assumptions this is an exact rewriting of the orthogonal dynamics equation. In this rewritten form the orthogonal dynamics system can be more easily solved. The main assumption in the approach is the commutativity of $e^{tQ\mathcal{L}}$ and R :

$$e^{tQ\mathcal{L}}(R(x_0)) \approx R(e^{tQ\mathcal{L}}x_0) = R(\phi^Q(x_0, t)). \quad (14)$$

For linear systems this is always valid, since then all operators are linear. For nonlinear systems it mostly is not valid, but can be used as a first approximation. With this assumption, the orthogonal dynamics equation can be reformulated into the pseudo-orthogonal dynamics equation:

$$\frac{\partial}{\partial t}\phi^Q(x_0, t) = R(\phi^Q(x_0, t)) - R(\hat{\phi}^Q(x_0, t)). \quad (15)$$

Note that this equation can be implemented more straightforward in numerical codes compared to the original orthogonal dynamics equation. When this equation is solved, the noise term corresponding to the resolved component ϕ_i is given by

$$F_i(x_0, t) = R_i(\phi^Q(x_0, t)) - R_i(\hat{\phi}^Q(x_0, t)). \quad (16)$$

This is the part of the pseudo-orthogonal dynamics system corresponding to the respective resolved variable. The noise term thus can be retrieved directly when solving the pseudo-orthogonal dynamics equation.

2.4 Memory Kernel

Having a solution to the orthogonal dynamics system is not necessarily sufficient to also have a clear expression for the memory term. When the noise term is a complicated function of the solution to the orthogonal dynamics equation, it can be quite involved to get an expression for the memory kernel. Therefore, it is useful to look into approximations of the kernel. The following approximation has been derived by Gouasmi et al. [20]. The first step in the derivation of this approximation is to consider the n components of $\mathcal{L}F_i(x_0, t)$ as one sensitivity in the direction of

$$\bar{R}(x_0) = \frac{R(x_0)}{\|R(x_0)\|}, \quad (17)$$

instead of n separate sensitivities in directions $R_i(x_0)$. Here $\bar{R} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\|\cdot\|$ is the l^2 -norm, so the standard Euclidean norm as the space considered is \mathbb{R}^n . This yields

$$\begin{aligned} \mathcal{L}F_i(x_0, t) &= \sum_{j=1}^N R_j(x_0) \partial_{x_j} F_i(x_0, t) \\ &= \|R(x_0)\| \sum_{j=1}^N \bar{R}_j(x_0) \partial_{x_j} F_i(x_0, t). \end{aligned} \quad (18)$$

Now define $\nabla_{\bar{R}(x_0)} = \sum_{j=1}^N \bar{R}_j(x_0) \partial_{x_j}$, so

$$\begin{aligned} \mathcal{L}F_i(x_0, t) &= \|R(x_0)\| (\nabla_{\bar{R}(x_0)} F_i(x_0, t)) \\ &= \|R(x_0)\| \lim_{\epsilon \rightarrow 0} \frac{F_i(x_0 + \epsilon \bar{R}(x_0), t) - F_i(x_0, t)}{\epsilon}. \end{aligned} \quad (19)$$

Here $\nabla_{\bar{R}(x_0)} F_i(x_0, t)$ has been written as limit. The memory kernel is the projection of this equation onto the resolved variables. Since $F_i(x_0, t)$ is the solution to the orthogonal dynamics equation, it only depends on the unresolved variables. Therefore, this component disappears after projection onto the resolved variables, that is $PF_i(x_0, t) = F_i(\hat{x}_0, t) = 0$. Applying P to Equation (19) and using finite differences, results in the following approximation for the memory kernel:

$$K_i(\hat{x}_0, t) \approx \|R(\hat{x}_0)\| \frac{F_i(\hat{x}_0 + \epsilon \bar{R}(\hat{x}_0), t)}{\epsilon}. \quad (20)$$

The exact result can be recovered if the limit of $\epsilon \rightarrow 0$ exists. The memory integral can be approximated by the rectangle rule or another approximation method for integrals.

The derived approximation is most useful when working numerically. To obtain analytical expressions using this method can be quite involved. However, if the result of the approximation remains tractable and the limit can be computed, it can result in an exact expression for the memory kernel.

3 El Niño Southern Oscillation

The El Niño Southern Oscillation (ENSO) is an oscillation in sea surface temperature in the Pacific Ocean with a period of four to seven years [10]. The normal situation above the Pacific Ocean consists of strong westward trade winds, which blow the warm surface water towards Indonesia and Australia. The east-west temperature differences in sea surface temperature interact with a convective circulation. At the western side of the ocean warm water heats the air, causing it to rise. High in the atmosphere the air flows eastward, after which it sinks near South America. At the surface it flows back towards the west, forming the trade winds. Since the water is blown away from the coast of South America, at that side of the ocean water rises from depths to the surface. This upwelling results in a shallow thermocline at the eastern boundary. The thermocline is the division layer between the upper-ocean warm mixed layer and the cooler deep waters below. In the thermocline temperature changes rapidly with depth.

During an El Niño event the atmospheric pressure differences between the east and west equatorial Pacific Ocean are smaller than usual, resulting in weaker trade winds. As a consequence, the warm surface water spreads further towards the east, reinforcing the small pressure difference by a reduced temperature gradient. During a La Niña event the opposite happens. The pressure difference is larger compared to the normal situation, resulting in stronger trade winds and even more upwelling at the coast of South America. In Figure 1 the sea surface temperature, wind patterns and thermocline depth of the three different situations are shown. These changes in sea surface temperature and thermocline depth, with El Niño as maximum and La Niña as minimum, follow an oscillation with a period of four to seven years. This oscillation is most pronounced in the eastern part of the Pacific Ocean.

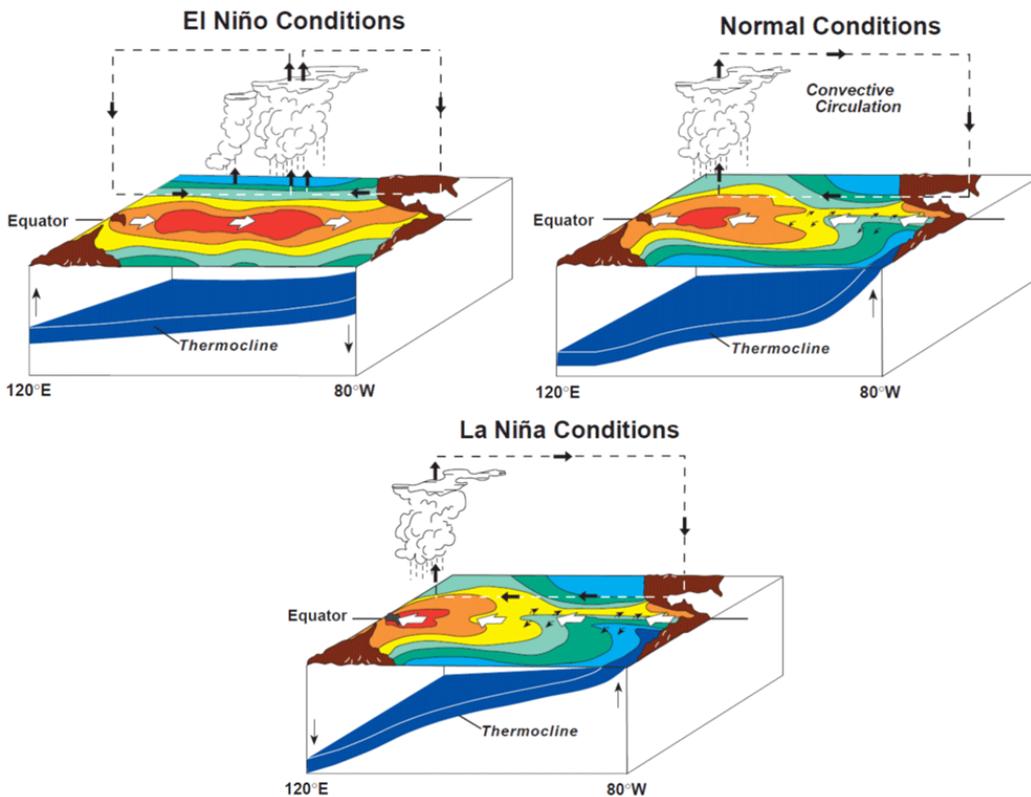


Figure 1: The situation in the Pacific Ocean during El Niño, normal and La Niña conditions. The colors show the sea surface temperature (red above 30°C, dark blue below 20°C) and the thermocline is shown in blue. The black arrows denote air movement, while the white arrows show ocean currents. Figure taken from PMEL/NOAA [29].

3.1 Delay Model by Suarez and Schopf

The processes behind ENSO, as explained before, are well understood. Many models have been developed to describe the dynamics of ENSO. An example of such an ENSO model is that by Zebiak and Cane (1987) [41]. This model is a system of partial differential equations for the flow in both the ocean and atmosphere, together with a ocean mixed layer temperature equation. In 1988 Suarez and Schopf proposed a much simpler model for ENSO. This delay model for only the sea surface temperature in the east of the Pacific Ocean is [36]

$$\frac{dT}{dt} = T(t) - T(t)^3 - \alpha T(t - \delta). \quad (21)$$

Here δ is the delay time and α a scaled parameter indicating the strength of the delayed feedback. Note that this model is scaled to contain as little parameters as possible. Battisti and Hirst gave a physical derivation of the linear terms in this model and some arguments for the form of the nonlinearity [2]. Their results are based on a sea surface temperature model by Battisti [1], which is similar to the model by Zebiak and Cane [41].

The delay in the processes of ENSO is caused by the physical mechanism depicted in Figure 2. A positive perturbation in the thermocline depth, meaning a deeper thermocline and a higher sea surface temperature, causes a weakening of the trade winds by a smaller zonal (east-west) temperature gradient. This weakening results in transport of warm surface water towards the equator. By this meridional (north-south) transport there is a positive temperature perturbation at the equator and a negative perturbation at higher latitudes. The positive perturbation travels eastward in the form of an equatorial Kelvin wave [7], causing the thermocline to deepen even further in the east of the basin. This positive feedback effect is considered to be immediate in the model by Suarez and Schopf. The negative perturbation at a higher latitude first travels westward in the form of a Rossby wave [7]. When it reaches the coast, the perturbation travels towards the equator as a coastal Kelvin wave and back to the eastern boundary as an equatorial Kelvin wave. Since the Rossby wave travels slow compared to the Kelvin waves, the travel time of this signal is much longer than that of the direct Kelvin wave.

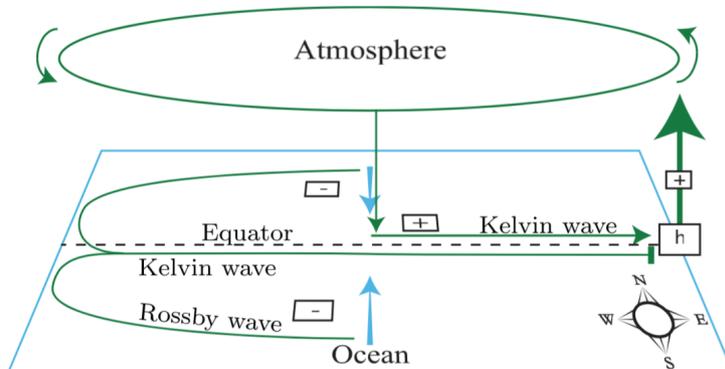


Figure 2: The mechanism behind the delay in ENSO. Figure taken from Keane et al. [27].

The delay model of Equation (21) gives oscillations with a period of two to three years. This is on the short side of the measured period of ENSO, indicating that there is an aspect missing in the model. The physics behind the delay is well understood, justifying the linear terms in Equation (21). However, the nonlinear term in the model by Suarez and Schopf is proposed ad hoc and no physical justification is given for it in their article. Battisti and Hirst provided some arguments for the form of the nonlinearity [2], but no thorough mathematical derivation has been given yet.

In the following sections a mathematical derivation of a delay model for ENSO is given. This derivation leads to an extended version of the delay model by Suarez and Schopf, resulting in an increase of the period. The derivation of a delay model also provides a proof of concept for the application of the Mori-Zwanzig formalism in deriving delay equations.

3.2 Linear Two-Strip Model

This section starts with the formulation of the (linear) model to which the Mori-Zwanzig formalism will be applied. After the formulation of the model, the application of the formalism is discussed, including the use of characteristics to arrive at a linear delay equation.

3.2.1 Model Formulation

The derivation of a delay model for ENSO starts from the two-strip model derived and studied by Jin in 1997 [24, 25]. This two-strip model is derived from the dimensionless shallow water equations [12]:

$$\begin{aligned}\partial_t u - yv + \partial_x h + \epsilon_0 u &= \tau, \\ yu + \partial_y h &= 0, \\ \partial_t h + \partial_x u + \partial_y v + \epsilon_0 h &= 0.\end{aligned}\tag{22}$$

Here u and v are the zonal and meridional velocities respectively and h is the thermocline depth. The wind forcing is represented by τ and ϵ_0 is a linear damping coefficient. The terms with y , the meridional coordinate, represent the β effect due to the earth's rotation. The model is scaled such that $x \in [0, 1]$ and perturbations in velocities and thermocline depth are of order one.

Reducing these equation to one equation for thermocline depth and assuming a parabolic dependence of the thermocline near the equator, yields a system of equations for the thermocline depth at the equator (h_e) and at some latitude y_n between 5° and 15° (h_n):

$$\begin{aligned}(\partial_t + \epsilon_0)(h_e - h_n) + \partial_x h_e &= \tau|_{y=0}, \\ (\partial_t + \epsilon_0)h_n - \frac{1}{y_n^2} \partial_x h_n &= \partial_y \left(\frac{\tau}{y} \right) \Big|_{y=y_n},\end{aligned}\tag{23}$$

with boundary conditions

$$h_e(0, t) = r_W h_n(0, t), \quad h_n(1, t) = r_E h_e(1, t).\tag{24}$$

Here r_W and r_E are a measure of the allowed mass flux at the western and eastern boundaries respectively. The ∂_x -terms give the advection of anomalies in the thermocline by Kelvin waves (h_e) and Rossby waves (h_n).

The wind forcing depends on the sea surface temperature at the equator (T_e), as discussed in Section 3.1. Following the Gill atmosphere model in a simplified version [12], it is given by

$$\tau = \mu A(T_e) e^{-\frac{(\alpha y)^2}{2}},\tag{25}$$

where μ is the coupling coefficient and α the fraction between meridional scales in the ocean and atmosphere. It is assumed that the wind stress depends on the temperature according to

$$A(T_e) = g(x) T_e(x_E, t).\tag{26}$$

Here x_E is a location in the east of the basin and $g(x)$ gives the pattern of the wind forcing in the zonal direction. This represents the physical mechanism as described in Section 3.1, where the temperature in the east affects the wind forcing. For simplicity the choice $x_E = 1$ is made.

An equation for the temperature perturbations T_e at the equator, derived by Dijkstra and Neelin [13], is added to the thermocline equations. This way the two-strip model describing the dynamics of ENSO becomes

$$\begin{aligned}(\partial_t + \epsilon_0)(h_e - h_n) + \partial_x h_e &= \mu g(x) T_e(x_E, t), \\ (\partial_t + \epsilon_0)h_n - \frac{1}{y_n^2} \partial_x h_n &= -\mu \frac{\theta}{y_n^2} g(x) T_e(x_E, t), \\ \partial_t T_e + c_T T_e - c_h h_e &= 0.\end{aligned}\tag{27}$$

Here θ is an order one coefficient representing the difference in wind forcing between the equator and higher latitudes. In the equation for temperature c_T represents local damping and c_h gives the effect of thermocline depth on temperature through upwelling.

For the formula determining c_T one can consult the article by Dijkstra and Neelin [13]. The formula for c_h is discussed here, since it will be used when a nonlinear variation of the model is considered. The expression for c_h is

$$c_h = f_h(x) \frac{\partial T_s}{\partial h}, \quad (28)$$

where f_h gives the background wind forcing (see Dijkstra and Neelin for the expression [13]) and T_s is the subsurface temperature as a function of thermocline depth at the equator. For the parametrization of the subsurface temperature the result by Hao et al. is used [22]:

$$T_s(h) = T_{s0} + (T_0 + T_{s0}) \tanh\left(\frac{h + h_0}{H^*}\right). \quad (29)$$

Here T_0 is the ocean equilibrium temperature in absence of dynamics, h_0 is an offset value for the thermocline and T_{s0} is the temperature at $h = -h_0$. H^* determines the steepness of the transition when h passes through $-h_0$. For the linear model c_h is considered at the reference state of the thermocline.

Considering the system of Equation (27), there is a way to rewrite the equations for the thermoclines such that decoupled equations for h_n and a new variable $h_c = h_e - \frac{1}{1+y_n^2} h_n$ are found. This new variable is dominated by the thermocline depth at the equator, but includes some influence of the higher latitudes as well. In these new variables the equations are

$$\begin{aligned} \partial_t h_c + \epsilon_0 h_c + \partial_x h_c &= \mu \left(1 - \frac{\theta}{1+y_n^2}\right) g(x) T_e(x_E, t), \\ \partial_t h_n + \epsilon_0 h_n - \frac{1}{y_n^2} \partial_x h_n &= -\mu \frac{\theta}{y_n^2} g(x) T_e(x_E, t), \\ \partial_t T_e + c_T T_e - c_h \left(h_c + \frac{1}{1+y_n^2} h_n\right) &= 0, \end{aligned} \quad (30)$$

with boundary conditions

$$h_c(0, t) = \left(r_W - \frac{1}{1+y_n^2}\right) h_n(0, t), \quad r_E h_c(1, t) = \left(1 - \frac{r_E}{1+y_n^2}\right) h_n(1, t). \quad (31)$$

If $r_E = 0$, then $h_n(1, t) = 0$, meaning no reflection occurs at the eastern boundary.

The homogeneous equations for thermocline depth in the rewritten system of Equation (30), without wind forcing, can be solved analytically. Doing so yields

$$\begin{aligned} h_c^0(x, t) &= H_c e^{\sigma_k t} e^{-(\sigma_k + \epsilon_0)x}, \\ h_n^0(x, t) &= H_n e^{\sigma_k t} e^{(\sigma_k + \epsilon_0) y_n^2 x}, \end{aligned} \quad (32)$$

with

$$\sigma_k = -\epsilon_0 + \frac{1}{1+y_n^2} \left(\ln \left(\frac{r_E r_W (1+y_n^2) - r_E}{(1+y_n^2) - r_E} \right) + 2\pi i k \right), \quad k \in \mathbb{N}. \quad (33)$$

Here by the boundary conditions it is required that $H_c = \left(r_W - \frac{1}{1+y_n^2}\right) H_n$. For $r_E = 0$, the only possible solution of the homogeneous system is the trivial one: $h_c^0 = h_n^0 = 0$. The solutions in Equation (32) are the eigensolutions of the two-strip model. Note that these eigensolutions are not orthogonal, meaning they are not convenient as a basis of the system on which can be projected.

3.2.2 Mori-Zwanzig Formalism

Starting from the rewritten version of the two-strip model in Equation (30), the goal is to derive a delay equation describing ENSO using the Mori-Zwanzig formalism. It is expected this model will be similar to that by Suarez and Schopf in Equation (21). The first issue to overcome, is that the Mori-Zwanzig formalism works on ordinary differential equations, while the two-strip model is a system of partial differential equations.

A route that is often taken in such situations is to use the eigensolutions as a basis for the general solutions [8, 19, 37]. Projecting onto these solutions yields a basis in which the system can be represented as a system of ordinary differential equations. However, for the two-strip model this does not yield a feasible result, since the eigensolutions are not orthogonal. Other projection methods, such as the use of Fourier exponentials or orthogonal polynomials as a basis [35], either do not work due to incompatible boundary conditions, or yield very tedious results which are not suitable for analytical computations.

Here, the solution used is to initially look at ∂_x as an operator and consider the system as ordinary differential equations in time. After going through the procedure of the Mori-Zwanzig formalism, the resulting terms that contain ∂_x are analysed. In this section the linear version of the two-strip model is considered, meaning c_T and c_h are allowed to depend on place (x) and possibly time (t), but not on any of the variables h_c , h_n or T_e . Here $c_T(x)$ and $c_h(x, \epsilon t)$ are taken.

Firstly the Liouville operator is identified:

$$\begin{aligned} \mathcal{L} = & \left(-(\epsilon_0 + \partial_x)h_c(x, 0) + \mu \left(1 - \frac{\theta}{1 + y_n^2} \right) g(x)T_e(x_E, 0) \right) \partial_{h_c} \\ & + \left(-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)h_n(x, 0) - \mu \frac{\theta}{y_n^2} g(x)T_e(x_E, 0) \right) \partial_{h_n} \\ & + \left(-c_T(x)T_e(x, 0) + c_h(x, \epsilon t) \left(h_c(x, 0) + \frac{1}{1 + y_n^2} h_n(x, 0) \right) \right) \partial_{T_e}. \end{aligned} \quad (34)$$

It gives the evolution of the system for arbitrary initial conditions. The model by Suarez and Schopf is a delay equation for the temperature at the equator in the east of the basin. Therefore, the equatorial temperature T_e is chosen as the resolved variable. A linear projection is used when applying the Mori-Zwanzig formalism:

$$P(f(T_e, h_c, h_n)) = f(T_e, 0, 0) \equiv \hat{f}(T_e), \quad (35)$$

with $P : C(\mathbb{R}^3, \mathbb{R}) \rightarrow C(\mathbb{R}, \mathbb{R})$, $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ and $\hat{f} : \mathbb{R} \rightarrow \mathbb{R}$, reducing the number of dependent variables from three to one.

Applying the formalism results in an equation, which is defined according to Equation (11), for just the resolved variable T_e . The different terms in this Langevin equation are computed and discussed separately. Firstly the Markovian term is computed:

$$\begin{aligned} e^{t\mathcal{L}} P\mathcal{L}(T_e(x, 0)) &= e^{t\mathcal{L}} P \left(-c_T(x)T_e(x, 0) + c_h(x, \epsilon t) \left(h_c(x, 0) + \frac{1}{1 + y_n^2} h_n(x, 0) \right) \right) \\ &= -c_T(x)e^{t\mathcal{L}}T_e(x, 0) \\ &= -c_T(x)T_e(x, t). \end{aligned} \quad (36)$$

As expected, this is the right-hand side dependence on the resolved variable T_e in the temperature component of the rewritten two-strip model.

To compute the noise and memory term the first step is to solve the orthogonal dynamics equation. Since here the system considered is linear, the pseudo-orthogonal dynamics approximation in Equation (15) is exact. This is used to compute the solutions of the orthogonal dynamics equation. The pseudo-orthogonal dynamics system for the thermocline equations of the two-strip model is

$$\begin{aligned} \frac{d}{dt} h_c^Q(x, t) &= -(\epsilon_0 + \partial_x)h_c^Q(x, t), \\ \frac{d}{dt} h_n^Q(x, t) &= -\left(\epsilon_0 - \frac{1}{y_n^2} \partial_x \right) h_n^Q(x, t). \end{aligned} \quad (37)$$

Here Q is used to denote the variables in the orthogonal dynamics equation. These equations are independent of each other and can be easily solved. The solutions are

$$\begin{aligned} h_c^Q(x, t) &= e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0), \\ h_n^Q(x, t) &= e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0). \end{aligned} \quad (38)$$

Note that this is another way of writing the solutions in Equation (32) of the homogeneous two-strip model. Having solved the orthogonal dynamics equation, the noise term can be computed. In the pseudo-orthogonal dynamics approximation it is defined by

$$F_{T_e}(x, t) = c_h(x, \epsilon t) \left(h_c^Q(x, t) + \frac{1}{1 + y_n^2} h_n^Q(x, t) \right). \quad (39)$$

This is the right-hand side of the equation that corresponds to the resolved variable T_e in the pseudo-orthogonal dynamics system. Substituting the solutions of the orthogonal dynamics system into this equation gives

$$F_{T_e}(x, t) = c_h(x, \epsilon t) \left(e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) + \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) \right). \quad (40)$$

The last component of the Langevin equation that needs to be computed is the memory term. To do this, first the memory kernel is computed:

$$\begin{aligned} K_{T_e}(T_e(x, 0), t) &= P\mathcal{L}F_{T_e}(x, t) \\ &= P\mathcal{L} \left(c_h(x, \epsilon t) \left(e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) + \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) \right) \right) \\ &= c_h(x, \epsilon t) \left(P\mathcal{L} \left(e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) \right) + \frac{1}{1 + y_n^2} P\mathcal{L} \left(e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) \right) \right). \end{aligned} \quad (41)$$

It is assumed that P and \mathcal{L} commute with $e^{-(\epsilon_0 + \partial_x)t}$ and $e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t}$. The validity of this assumption will be discussed in Section 3.3. This gives

$$\begin{aligned} K_{T_e}(T_e(x, 0), t) &= c_h(x, \epsilon t) \\ &\cdot P \left(e^{-(\epsilon_0 + \partial_x)t} \left(-(\epsilon_0 + \partial_x) h_c(x, 0) + \mu \left(1 - \frac{\theta}{1 + y_n^2} \right) g(x) T_e(x_E, 0) \right) \right. \\ &\quad \left. + \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} \left(-(\epsilon_0 - \frac{1}{y_n^2} \partial_x) h_n(x, 0) - \mu \frac{\theta}{y_n^2} g(x) T_e(x_E, 0) \right) \right) \\ &= c_h(x, \epsilon t) \left(\mu \left(1 - \frac{\theta}{1 + y_n^2} \right) e^{-(\epsilon_0 + \partial_x)t} - \mu \frac{\theta}{y_n^2} \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} \right) \\ &\quad \cdot g(x) T_e(x_E, 0). \end{aligned} \quad (42)$$

It is important to note that the memory kernel is the term in front of $T_e(x, 0)$, not the full expression given. Using the Langevin equation, the memory term is found by substituting the computed memory integrand into the integral. This yields

$$\begin{aligned} \int_0^t K_{T_e}(T_e(x, t-s), s) ds &= \int_0^t c_h(x, \epsilon(t-s)) \left(\mu \left(1 - \frac{\theta}{1 + y_n^2} \right) e^{-(\epsilon_0 + \partial_x)(t-s)} \right. \\ &\quad \left. - \mu \frac{\theta}{y_n^2} \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)(t-s)} \right) g(x) T_e(x_E, s) ds. \end{aligned} \quad (43)$$

Now all terms present in the Mori-Zwanzig formalism of Equation (11) are computed. The result, found by substituting the above computed expression into this Langevin equation, is the following equation for the temperature at the equator:

$$\begin{aligned} \frac{dT_e}{dt}(x, t) &= -c_T(x) T_e(x, t) \\ &\quad + c_h(x, \epsilon t) \left(e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) + \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) \right) \\ &\quad + \int_0^t c_h(x, \epsilon(t-s)) \left(\mu \left(1 - \frac{\theta}{1 + y_n^2} \right) e^{-(\epsilon_0 + \partial_x)(t-s)} \right. \\ &\quad \left. - \mu \frac{\theta}{y_n^2} \frac{1}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)(t-s)} \right) g(x) T_e(x_E, s) ds. \end{aligned} \quad (44)$$

The partial derivative to x still is present in the exponential terms. In the noise term this is not an issue, since the terms are exactly the solutions as given in Equation (32) to the homogeneous system:

$$e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) = h_c^0(x, t), \quad e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) = h_n^0(x, t). \quad (45)$$

In the memory term it is less clear what to do with these exponentials. In the next section a way of simplifying these terms using the method of characteristics is given.

3.2.3 Characteristics

In this section the memory integral is considered, focusing on the exponential ∂_x -terms. Here the method of characteristics is used [17]. Both components in the memory kernel are of the form $e^{-(\epsilon_0 + c\partial_x)(t-s)} f(x, s)$, for either $c = 1$ or $c = -1/y_n^2$. This expression is the solution to the partial differential equation

$$\partial_t f + c\partial_x f = -\epsilon_0 f, \quad (46)$$

with initial conditions given at $t = s$. The characteristic equations for this partial differential equation are [17]

$$\frac{dt}{dy} = 1, \quad \frac{dx}{dy} = c. \quad (47)$$

The solutions to these equations are $t(y) = y + t_0$ and $x(y) = cy + x_0$. This results in characteristic curves $x - x_0 = c(t - t_0)$ along which $f(x, t)$ is conserved, meaning $f(x, t)$ is constant on such a curve apart from the damping caused by the ϵ_0 -term. This gives

$$e^{-(\epsilon_0 + c\partial_x)(t-s)} f(x, s) = e^{-\epsilon_0(t-s)} f(c(t-s) + x_s, s), \quad (48)$$

where x_s is the location at time s . This expression is valid as long as the argument of f lies in the domain. Otherwise boundary effects need to be considered.

In the two-strip model $x \in [0, 1]$, which means the reflection of characteristics at the boundary has to be discussed. In the memory integral there are two different exponentials, with $c = 1$ and $c = -1/y_n^2$. The corresponding characteristics represent the eastward traveling equatorial Kelvin waves for $c = 1$ and the westward traveling Rossby waves for $c = -1/y_n^2$. Note that the Rossby waves indeed take longer to cross the basin than the Kelvin waves as $y_n > 1$. In Figure 3 the characteristics of the memory term are shown. The red line shows what happens to a signal emitted from $x = 0.7$ at time zero until it arrives at the eastern boundary.

The characteristics of the system are used to get an expression for the memory term. The goal is to find a result for the temperature in the east of equatorial basin (x_E). This is where the model by Suarez and Schopf is valid and it is the location of the temperature on which the wind forcing depends. Looking at the signal at one location allows for the following of characteristics from a source given by $g(x)$ to that one location. Denoting $T_e(x_E, t) = T_e^E(t)$ the memory integral at $x = x_E$ is of the form

$$\int_0^t c_h(x_E, \epsilon(t-s)) \left(\mu \left(1 - \frac{\theta}{1 + y_n^2} \right) \cdot \left[e^{-(\epsilon_0 + \partial_x)(t-s)} g(x) \right]_{x_E} T_e^E(s) - \mu \frac{\theta}{y_n^2} \frac{1}{1 + y_n^2} \cdot \left[e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)(t-s)} g(x) \right]_{x_E} T_e^E(s) \right) ds. \quad (49)$$

The first term in the memory integral gives the waves that are traveling westward at $t = s$. Setting $x_E = 1$, these waves need a time $t = 1 - x$ to arrive at the eastern boundary. Assuming reflection takes place at the eastern and western boundary, the signal arrives a second time after $t = 1 - x + (y_n^2 + 1)$. This way the signal keeps reflecting through the basin, arriving at the eastern boundary after times $t_k = 1 - x + k(y_n^2 + 1)$ for $k = 0, 1, 2, \dots$. At each reflection the wave loses energy by a factor A_{rE} at the eastern boundary and a factor A_{rW} at the western boundary. These factors are determined by the amount of flux allowed through the boundaries and thus by the boundary conditions for h_n and h_c . This means

$$A_{rW} = r_W(1 + y_n^2) - 1, \quad A_{rE} = \left(\frac{1 + y_n^2}{r_E} - 1 \right)^{-1}, \quad (50)$$

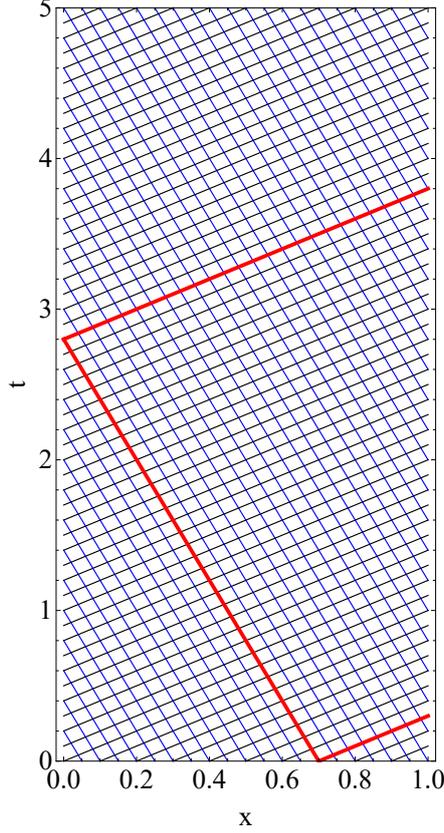


Figure 3: The characteristics of $\partial_t f + \partial_x f = -\epsilon_0 f$ (black) and $\partial_t f - \frac{1}{y_n^2} \partial_x f = -\epsilon_0 f$ (blue). In red the path of a signal following the characteristics is shown until it reaches the eastern boundary.

where the factor $1 + y_n^2$ relative to the thermocline boundary conditions in Equation (31) emerges due to the different contributions of h_c and h_n in the temperature equation of the two-strip model. The result of these reflections through the basin on the first part of the memory term at the eastern boundary is

$$\begin{aligned}
& \int_0^t c_h(1, \epsilon(t-s)) \mu \left(1 - \frac{\theta}{1 + y_n^2}\right) \left[e^{-(\epsilon_0 + \partial_x)(t-s)} g(x) \right]_{x_E} T_e^E(s) ds \\
&= \sum_{k=0}^{K_{max}(t)} \int_{t-(1+k(y_n^2+1))}^{t-k(y_n^2+1)} c_h(1, \epsilon(t-s)) \mu \left(1 - \frac{\theta}{1 + y_n^2}\right) e^{-\epsilon_0(t-s)} \\
&\quad \cdot (A_{rE} A_{rW})^k g(1 + k(y_n^2 + 1) - (t-s)) T_e^E(s) ds,
\end{aligned} \tag{51}$$

where $K_{max}(t) = \lfloor \frac{t-1}{y_n^2+1} \rfloor$ for $t \geq 1$ is the number of reflections that have occurred by time t . Note that there are time intervals for which this term has no effect at the eastern boundary, since it only represents half of the characteristics.

To get this part of the memory integral in a form which shows more of the delay behaviour, a change of coordinates is applied. Let $x = 1 + k(y_n^2 + 1) - (t-s)$, for which $\frac{dx}{ds} = 1$. Changing coordinates from s to x , the memory integral becomes

$$\begin{aligned}
& \sum_{k=0}^{K_{max}(t)} \int_0^1 c_h(1, \epsilon(1 + k(y_n^2 + 1) - x)) \mu \left(1 - \frac{\theta}{1 + y_n^2}\right) e^{-\epsilon_0(1+k(y_n^2+1)-x)} \\
&\quad \cdot (A_{rE} A_{rW})^k g(x) T_e^E(t - (1 + k(y_n^2 + 1) - x)) dx.
\end{aligned} \tag{52}$$

Writing the integral this way shows that the memory term contains a component that depends on past states of the resolved variable T_e^E . How strong the effect is of different times depends on the function $g(x)$, which gives the spatial distribution of the wind forcing.

The result for the waves that first travel towards the western boundary is achieved in a similar way. These waves need times $t_k = y_n^2 x + 1 + k(1 + y_n^2)$ for $k = 0, 1, 2, \dots$ to arrive at the eastern boundary. Now $x = -\frac{1}{y_n^2}(1 + k(y_n^2 + 1) - (t - s))$ is used for the change of variables. Going through the same steps as before, the total memory integral becomes

$$\begin{aligned} & \sum_{k=0}^{K_{max}(t)} (A_{rE} A_{rW})^k \mu \int_0^1 g(x) e^{-\epsilon_0 k(y_n^2 + 1)} \\ & \cdot \left(\left(1 - \frac{\theta}{1 + y_n^2}\right) c_h(1, \epsilon(1 + k(y_n^2 + 1) - x)) e^{-\epsilon_0(1-x)} T_e^E(t - (1 + k(y_n^2 + 1) - x)) \right. \\ & \left. - \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h(1, \epsilon(1 + k(y_n^2 + 1) + y_n^2 x)) e^{-\epsilon_0(1+y_n^2 x)} T_e^E(t - (1 + k(y_n^2 + 1) + y_n^2 x)) \right) dx. \end{aligned} \quad (53)$$

This expression shows there are multiple delays present in the two-strip model for ENSO. The exact form of the delay (distributed or discrete) and delay times are determined by the spatial pattern of the wind forcing $g(x)$.

3.2.4 Delay Model

Using the results from the two previous sections, an equation for the evolution of temperature at the eastern boundary of the domain can be obtained. Considering Equation (44) for $x = x_E = 1$ and using Equations (45) and (53) gives

$$\begin{aligned} \frac{dT_e^E}{dt} &= -c_T(1)T_e^E(t) + c_h(1, \epsilon t) \left(h_c^0(1, t) + \frac{1}{1 + y_n^2} h_n^0(1, t) \right) \\ &+ \sum_{k=0}^{K_{max}(t)} (A_{rE} A_{rW})^k \mu \int_0^1 g(x) e^{-\epsilon_0 k(y_n^2 + 1)} \\ &\cdot \left(\left(1 - \frac{\theta}{1 + y_n^2}\right) c_h(1, \epsilon(1 + k(y_n^2 + 1) - x)) e^{-\epsilon_0(1-x)} \right. \\ &\quad \cdot T_e^E(t - (1 + k(y_n^2 + 1) - x)) \\ &\quad \left. - \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h(1, \epsilon(1 + k(y_n^2 + 1) + y_n^2 x)) e^{-\epsilon_0(1+y_n^2 x)} \right. \\ &\quad \left. \cdot T_e^E(t - (1 + k(y_n^2 + 1) + y_n^2 x)) \right) dx. \end{aligned} \quad (54)$$

The effect of the components of the memory term decreases with k by the energy loss at reflection. To simplify this expression two assumptions are made.

First, it is assumed there is no reflection at the eastern boundary, so $r_E = 0$ and thus $A_{rE} = 0$. The only components in the sum of the memory term that remain with this assumption are the components for $k = 0$. Furthermore, as noted in Section 3.2, the homogeneous solution in that case is identically zero. This means the noise term vanishes in Equation (54). The equation for T_e^E then becomes

$$\begin{aligned} \frac{dT_e^E}{dt} &= -c_T(1)T_e^E(t) + \mu \int_0^1 g(x) \left(\left(1 - \frac{\theta}{1 + y_n^2}\right) c_h(1, \epsilon(1 - x)) e^{-\epsilon_0(1-x)} T_e^E(t - (1 - x)) \right. \\ &\quad \left. - \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h(1, \epsilon(1 + y_n^2 x)) e^{-\epsilon_0(1+y_n^2 x)} T_e^E(t - (1 + y_n^2 x)) \right) dx. \end{aligned} \quad (55)$$

This is already a strong simplification compared to Equation (54). However, the integral over the memory kernel still has to be computed. Unknown in this integral is the function $g(x)$, which determines the form of the memory kernel. This function gives the pattern of the wind forcing. More specifically, $g(x)$ indicates where the effect of the wind is strong and weak. From Section 3.1 it is known that the wind dominantly has an effect near the centre of the basin. Away from this location the effect is small, meaning the wind forcing acts quite locally. Here, this local effect of the wind forcing is approximated by $g(x) = A_0 \delta_{x_w}(x)$, a delta function of height A_0 at

$x = x_w$. This delta function leaves only the effect of that one location on the integral, meaning Equation (55) becomes

$$\begin{aligned} \frac{dT_e^E}{dt} = & -c_T(1)T_e^E(t) + \mu A_0 \left(\left(1 - \frac{\theta}{1 + y_n^2}\right) c_h(1, \epsilon(1 - x_w)) e^{-\epsilon_0(1 - x_w)} T_e^E(t - (1 - x_w)) \right. \\ & \left. - \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h(1, \epsilon(1 + y_n^2 x_w)) e^{-\epsilon_0(1 + y_n^2 x_w)} T_e^E(t - (1 + y_n^2 x_w)) \right). \end{aligned} \quad (56)$$

This way a linear discrete delay equation for temperature has been found.

The resulting equation does not yet resemble the model by Suarez and Schopf from Equation (21) in the linear terms. Instead of one, there are two delay times present. However, $1 - x_w \ll 1 + y_n^2 x_w$ indicating that this effect can be considered to be immediate compared to the other delay time. Thus it is assumed that $T_e^E(t - (1 - x_w)) \approx T_e^E(t)$. This is the same as assuming that Kelvin waves have a direct effect, as was discussed in Section 3.1. This approximation yields the final linear delay model:

$$\frac{dT_e^E}{dt} = c_S T_e^E(t) - c_L T_e^E(t - d), \quad (57)$$

where

$$\begin{aligned} c_S &= \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) c_h(1, \epsilon(1 - x_w)) e^{-\epsilon_0(1 - x_w)} - c_T(1), \\ c_L &= \mu A_0 \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h(1, \epsilon(1 + y_n^2 x_w)) e^{-\epsilon_0(1 + y_n^2 x_w)}, \\ d &= 1 + y_n^2 x_w. \end{aligned} \quad (58)$$

This model (after rescaling) gives the linear part of the model by Suarez and Schopf. The delay term is due to the emission of Rossby waves caused by a wind forcing which depends on the temperature near the eastern boundary. These waves travel to the western boundary, where they reflect in the form of Kelvin waves. The process described in Section 3.1 is well reflected throughout the derivation of Equation (57). This delay model does not yet account for the nonlinearity in the model by Suarez and Schopf. To be able to consider such a nonlinearity the model with which one starts already has to be nonlinear.

3.3 Nonlinear Two-Strip Model

In this section a nonlinear variation of the two-strip model is derived. Then this nonlinear model is studied using the Mori-Zwanzig formalism and a method based on variation of constants.

3.3.1 Model Formulation

To investigate how realistic the nonlinearity in the model by Suarez and Schopf is, a nonlinear version of the two-strip model is derived. To create such a nonlinearity in the model c_h is considered, meaning the upwelling feedback is assumed to be nonlinear. This is the nonlinearity that is found to be the most important in describing the behaviour of ENSO by Battisti and Hirst [2]. Recall that c_h , as defined by Equations (28) and (29), is proportional to the derivative of the subsurface temperature T_s with respect to the thermocline depth h . Computing this derivative gives

$$\frac{dT_s}{dh} = \frac{T_0 - T_{s0}}{H^*} \left(1 - \left(\frac{T_s - T_{s0}}{T_0 - T_{s0}}\right)^2\right). \quad (59)$$

The assumption made here is that T_e is proportional to $T_s - T_{s0}$, so perturbations in the equatorial sea surface temperature are assumed to be proportional to perturbations in the subsurface temperature. Using this assumption, Equation (28) becomes

$$c_h = f_h(x) \frac{T_0 - T_{s0}}{H^*} \left(1 - \left(\frac{c_{se} T_e}{T_0 - T_{s0}}\right)^2\right), \quad (60)$$

where c_{se} is the proportionality constant. This introduces a cubic nonlinearity in the temperature equation of the two-strip model.

To check the validity of the assumption made, buoy data of the equatorial Pacific Ocean is considered. There are ten locations in the equatorial Pacific where buoy data is available. Since the model only is concerned with the behaviour of the temperature in the east of the basin, it is sufficient to consider only buoys in the eastern part of the Pacific Ocean. To avoid coastal boundary layer effects the second most eastern buoy, which is located at 110 degrees west, is chosen. In Figure 4 the temperature data for this buoy is shown. The deviation from the mean for the sea surface temperature versus the offset subsurface temperature (see Equation 59) is depicted. The correlation between the two datasets is 0.87, indicating there is a strong relation, which is also visible in the figure. The slope between the two temperature perturbations is $c_{se} \approx 1$.

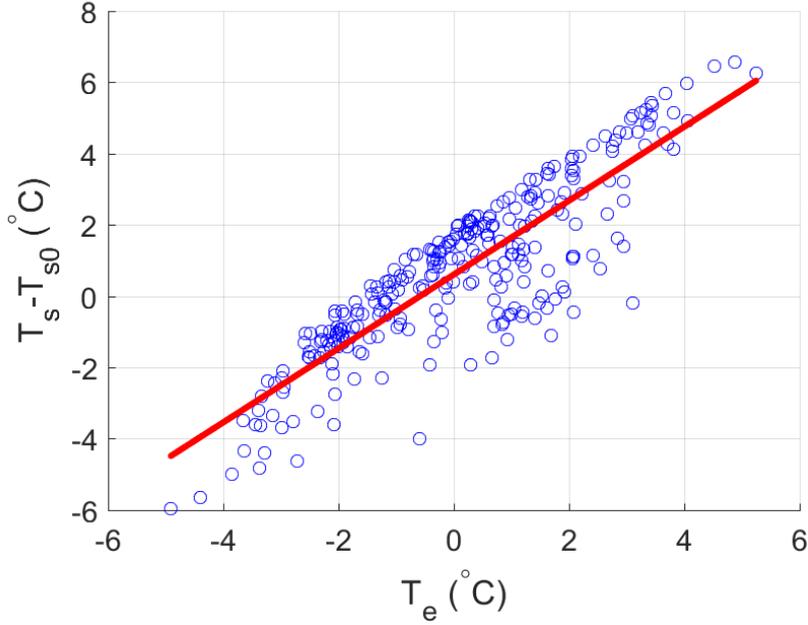


Figure 4: Temperature data from a buoy at the east side of the Pacific Ocean (110°W) for measurements at the surface and the subsurface (depth of 25m). Shown are the deviations from the average sea surface temperature versus the offset subsurface temperature. The red line is the best linear fit through the data and has a slope of 1.037 ± 0.035 . Data is taken from the Global Tropical Moored Buoy Array Project Office of NOAA/PMEL.

Using Equation (60) the nonlinear two-strip model (in the rewritten version) becomes

$$\begin{aligned}
 \partial_t h_c + \epsilon_0 h_c + \partial_x h_c &= \mu \left(1 - \frac{\theta}{1 + y_n^2} \right) g(x) T_e(x_E, t), \\
 \partial_t h_n + \epsilon_0 h_n - \frac{1}{y_n^2} \partial_x h_n &= -\mu \frac{\theta}{y_n^2} g(x) T_e(x_E, t), \\
 \partial_t T_e + c_T T_e - c_h^* (1 - \beta T_e^2) \left(h_c + \frac{1}{1 + y_n^2} h_n \right) &= 0.
 \end{aligned} \tag{61}$$

Here $c_h^* = f_h(x) \frac{T_0 - T_{s0}}{H^*}$ and $\beta = \left(\frac{1}{T_0 - T_{s0}} \right)^2$. It is important to realise that this result is in principle only valid in the eastern part of the basin, because at that location it has been shown there is a strong correlation and $c_{se} \approx 1$. At other locations the correlation turns out to be quite strong as well, but the proportionality constant between the surface and subsurface temperature is different.

Before continuing with the Mori-Zwanzig formalism for this nonlinear model, the model is studied in a more thorough way. Note that the equations for h_c and h_n are still linear, meaning

they can be rewritten in the form of integral equations, as was done in the linear illustration of Section 2.1. Using variation of constants, this gives

$$\begin{aligned} h_c(x, t) &= e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) + \int_0^t e^{-(\epsilon_0 + \partial_x)(t-s)} \mu \left(1 - \frac{\theta}{1 + y_n^2}\right) g(x) T_e(x_E, s) ds, \\ h_n(x, t) &= e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) - \int_0^t e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)(t-s)} \mu \frac{\theta}{y_n^2} g(x) T_e(x_E, s) ds. \end{aligned} \quad (62)$$

These expressions for h_c and h_n can be substituted into the equation for T_e . The resulting equation for T_e is

$$\begin{aligned} \frac{dT_e}{dt} &= -c_T(x) T_e(x, t) + c_h^*(x) (1 - \beta T_e^2(x, t)) \\ &\cdot \left(e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0) + \int_0^t e^{-(\epsilon_0 + \partial_x)(t-s)} \mu \left(1 - \frac{\theta}{1 + y_n^2}\right) g(x) T_e(x_E, s) ds \right. \\ &\left. + \frac{1}{1 + y_n^2} \left(e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} h_n(x, 0) - \int_0^t e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)(t-s)} \mu \frac{\theta}{y_n^2} g(x) T_e(x_E, s) ds \right) \right). \end{aligned} \quad (63)$$

The same procedure can be applied to the linear model, making the use of the Mori-Zwanzig formalism redundant. This way the commutativity assumption made in Section 3.2.2 can be verified. For fixed x the result is found to be exactly the same, providing a justification for the assumption.

Under the same assumptions as made in the previous section for the linear model (no reflection at the eastern boundary and a localized wind forcing), characteristics can be used to arrive at a delay equation. The delay model for the temperature at the eastern boundary becomes

$$\begin{aligned} \frac{dT_e^E}{dt} &= -c_T(1) T_e^E(t) + c_h^*(1) (1 - \beta T_e^E(t)^2) \mu A_0 \left(\left(1 - \frac{\theta}{1 + y_n^2}\right) e^{-\epsilon_0(1-x_w)} T_e^E(t - (1 - x_w)) \right. \\ &\left. - \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} e^{-\epsilon_0(1+y_n^2 x_w)} T_e^E(t - (1 + y_n^2 x_w)) \right). \end{aligned} \quad (64)$$

Assuming, as before, that the short delay is instantaneous, yields

$$\frac{dT_e^E}{dt} = (c_S^* - c_T(1)) T_e^E(t) - c_L^* T_e^E(t - d) - \beta c_S^* T_e^E(t)^3 + \beta c_L^* T_e^E(t)^2 T_e^E(t - d), \quad (65)$$

where

$$\begin{aligned} c_S^* &= \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) c_h^*(1) e^{-\epsilon_0(1-x_w)}, \\ c_L^* &= \mu A_0 \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h^*(1) e^{-\epsilon_0(1+y_n^2 x_w)}, \\ d &= 1 + y_n^2 x_w. \end{aligned} \quad (66)$$

This way a nonlinear discrete delay equation for the temperature at the equator, including two cubic terms, has been found. Compared to the model by Suarez and Schopf in Equation (21), a fourth term, which is proportional to $T_e^E(t)^2 T_e^E(t - d)$, emerges. The effect of this additional term on the dynamics of the delay model is discussed in Section 3.4. Before studying the behaviour of this extended delay model, first the application of the Mori-Zwanzig formalism to the nonlinear model in Equation (61) is discussed.

3.3.2 Mori-Zwanzig Formalism

The Mori-Zwanzig formalism is valid for both linear and nonlinear equations. The problem when considering nonlinear equations arises when solving the orthogonal dynamics equation. The first question that has to be answered to resolve this problem, is what the chosen projection does to nonlinear terms. Here the same linear projection as for the linear model is used (Equation (35)). The initial condition of the orthogonal dynamics system then is

$$Q\mathcal{L}(T_e(x, 0)) = c_h^*(x) (1 - \beta T_e^2(x, 0)) (h_c(x, 0) + \frac{1}{1 + y_n^2} h_n(x, 0)). \quad (67)$$

This expression is nonlinear, just as the initial system. Formally the noise term then can be written as

$$F_{T_e}(x, t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} (Q\mathcal{L})^k c_h^*(x) (1 - \beta T_e^2(x, 0)) (h_c(x, 0) + \frac{1}{1 + y_n^2} h_n(x, 0)), \quad (68)$$

by expressing the exponential as a sum. The components that are linear in $h_c(x, 0)$ and $h_n(x, 0)$ do not pose a problem. These terms both can be written in the form of an exponential sum. To see this consider

$$Q\mathcal{L}(h_c(x, 0)) = -(\epsilon_0 + \partial_x) h_c(x, 0), \quad (69)$$

which means

$$\sum_{k=0}^{\infty} \frac{t^k}{k!} (Q\mathcal{L})^k h_c(x, 0) = \sum_{k=0}^{\infty} \frac{t^k}{k!} (-\epsilon_0 + \partial_x)^k h_c(x, 0) = e^{-(\epsilon_0 + \partial_x)t} h_c(x, 0). \quad (70)$$

For h_n a similar result holds. This gives exactly the same result as found by using the pseudo-orthogonal dynamics approximation in the linear model.

Now the nonlinear terms in Equation (68) are examined. Applying the operator $Q\mathcal{L}$ more times gives terms of increasing order in the initial conditions of T_e , h_c and h_n . This does not converge to an analytical computable solution. Therefore, an approximation has to be made. A possibility is to neglect terms above some order, but this does not yield a workable solution. Another option is to approximate the orthogonal dynamics equation by the pseudo-orthogonal dynamics equation as described in Section 2.3. The conditions for this approximation are not met for the nonlinear two-strip model, but it can be used as a first estimate. This last approximation will be applied in the remainder of this section.

The pseudo-orthogonal dynamics equations for the nonlinear two-strip model are

$$\begin{aligned} \frac{d}{dt} h_c^Q(x, t) &= -(\epsilon_0 + \partial_x) h_c^Q(x, t), \\ \frac{d}{dt} h_n^Q(x, t) &= -(\epsilon_0 - \frac{1}{y_n^2} \partial_x) h_n^Q(x, t), \\ \frac{d}{dt} T_e^Q(x, t) &= c_h^*(x) (1 - \beta T_e^Q(x, t)^2) (h_c^Q(x, t) + \frac{1}{1 + y_n^2} h_n^Q(x, t)). \end{aligned} \quad (71)$$

The first two equations also occur in the pseudo-orthogonal dynamics system for the linear model and have exponential functions as solutions. In addition, this time a solution for T_e^Q has to be found, since the noise term depends on it. Recall here that the noise term is given by the right-hand side of the equation for T_e^Q . Substituting the solutions for h_c^Q and h_n^Q into the equation for T_e^Q , the solution for T_e^Q with the condition that $(T_e^Q)^2 < \frac{1}{\beta}$ is

$$\begin{aligned} T_e^Q(x, t) &= \frac{1}{\beta} \tanh^2 \left(\operatorname{arctanh}(\sqrt{\beta} T_e(x, 0)) \right. \\ &\quad \left. + c_h^*(x) \sqrt{\beta} ((1 - e^{-(\epsilon_0 + \partial_x)t}) (\epsilon_0 + \partial_x)^{-1} h_c(x, 0) \right. \\ &\quad \left. + \frac{1}{1 + y_n^2} (1 - e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t}) (\epsilon_0 - \frac{1}{y_n^2} \partial_x)^{-1} h_n(x, 0)) \right). \end{aligned} \quad (72)$$

If $(T_e^Q)^2 > \frac{1}{\beta}$ the tanh has to be replaced by a coth, and when $(T_e^Q)^2 = \frac{1}{\beta}$ the result is a constant T_e^Q , since then $\frac{d}{dt} T_e^Q(x, t) = 0$. The initial conditions determine which of the solutions should be used. It is most likely that $(T_e^Q)^2 < \frac{1}{\beta}$, as β is small and T_e is of order one. Therefore in the following Equation (72) is used.

The noise term is given by the right-hand side of the equation for T_e^Q in Equation (71), for which now a closed expression is known. The next step is to compute the memory kernel. By the presence of the hyperbolic tangent and several nonlinearities, Equation (20) from Section 2.4 is used to approximate the memory kernel. Taking the limit $\epsilon \rightarrow 0$, yields

$$\begin{aligned} K_{T_e}(T_e(x, 0), t) &= c_h^*(x) (1 + \beta T_e^2(x, 0)) \left(\mu \left(1 - \frac{\theta}{1 + y_n^2} \right) e^{-(\epsilon_0 + \partial_x)t} g(x) T_e(x_E, 0) \right. \\ &\quad \left. - \mu \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} e^{-(\epsilon_0 - \frac{1}{y_n^2} \partial_x)t} g(x) T_e(x_E, 0) \right). \end{aligned} \quad (73)$$

All components of the equation for T_e using the Mori-Zwanzig formalism with the pseudo-orthogonal dynamics approximation are known. Using the above expression for the memory kernel and substituting the general solution for $T_e^Q(x, t)$ in the noise term, the approximate equation for T_e becomes

$$\begin{aligned} \frac{dT_e}{dt}(x, t) = & -c_T(x)T_e(x, t) + c_h^*(x)(e^{-(\epsilon_0 + \partial_x)t}h_c(x, 0) + \frac{1}{1 + y_n^2}e^{-(\epsilon_0 - \frac{1}{y_n^2}\partial_x)t}h_n(x, 0)) \\ & \cdot \left(1 - \tanh^2\left(\operatorname{arctanh}\left(\sqrt{\beta}T_e(x, 0)\right)\right.\right. \\ & \left.\left.+ c_h^*(x)\sqrt{\beta}\left((1 - e^{-(\epsilon_0 + \partial_x)t})\frac{h_c(x, 0)}{\epsilon_0 + \partial_x} + \frac{1}{1 + y_n^2}(1 - e^{-(\epsilon_0 - \frac{1}{y_n^2}\partial_x)t})\frac{h_n(x, 0)}{\epsilon_0 - \frac{1}{y_n^2}\partial_x}\right)\right)\right) \\ & + \int_0^t c_h^*(x)(1 + \beta T_e^2(x, s))\left(\mu\left(1 - \frac{\theta}{1 + y_n^2}\right)e^{-(\epsilon_0 + \partial_x)(t-s)}g(x)T_e(x_E, s)\right. \\ & \left.- \mu\frac{\theta}{y_n^2}\frac{A_{rW}}{1 + y_n^2}e^{-(\epsilon_0 - \frac{1}{y_n^2}\partial_x)(t-s)}g(x)T_e(x_E, s)\right)ds. \end{aligned} \quad (74)$$

Similar to the procedure followed for the linear model in Sections 3.2.3 and 3.2.4, this equation can be simplified. The desired result is an equation for the temperature in the east of the basin. Applying the method of characteristics one gets rid of the exponential ∂_x -terms. Assuming no reflection takes place at the eastern boundary, the noise term vanishes. The memory term further simplifies to two delay terms when a localized wind forcing is assumed. Considering the short delay as being instantaneous, the resulting nonlinear delay equation is

$$\frac{dT_e^E}{dt} = (c_S^* - c_T(1))T_e^E(t) - c_L^*T_e^E(t - d) - \beta c_S^*T_e^E(t)^3 + \beta c_L^*T_e^E(t - d)^3, \quad (75)$$

where

$$\begin{aligned} c_S^* &= \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) c_h^*(1) e^{-\epsilon_0(1-x_w)}, \\ c_L^* &= \mu A_0 \frac{\theta}{y_n^2} \frac{A_{rW}}{1 + y_n^2} c_h^*(1) e^{-\epsilon_0(1+y_n^2 x_w)}, \\ d &= 1 + y_n^2 x_w. \end{aligned} \quad (76)$$

This equation is almost the same as Equation (65) derived in the first part of Section 3.3 by applying variation of constants. The only difference, which arises due to the pseudo-orthogonal dynamics approximation, is the fourth term. Since Equation (65) is exact, this difference is due to errors in the pseudo-orthogonal dynamics approximation. Since the exact result is known, the accuracy of the approximation can be tested for this ENSO model.

3.4 Delay Model Analysis

In Sections 3.2 and 3.3 delay models for ENSO have been derived. The linear model in Equation (57) shows the presence of a delay in the two-strip model. Adding a nonlinearity to the model results in two different delay models given in Equations (65) and (75). The first of these is derived by applying variation of constants to the thermocline equations and is exact. The second delay model is derived using the pseudo-orthogonal dynamics approximation and contains some error. Both models contain an extra term compared to the delay model proposed by Suarez and Schopf as given in Equation (21).

In this section the behaviour of the three different nonlinear delay models is studied. In Figure 5 model simulations for the three different models are shown. The model derived using the Mori-Zwanzig formalism has a very different shape of the oscillation, which does not correspond to data of ENSO. Both derived models have a longer period than the model by Suarez and Schopf, as well as a smaller amplitude. These two aspects lead to a better representation of the measurements of ENSO. The main focus in this section is on the similarities and differences between the model by Suarez and Schopf and the exact delay model derived by applying variation of constants. Furthermore, the model derived using the approximated Mori-Zwanzig formalism is studied to get an idea of the error made in using the pseudo-orthogonal dynamics approximation.

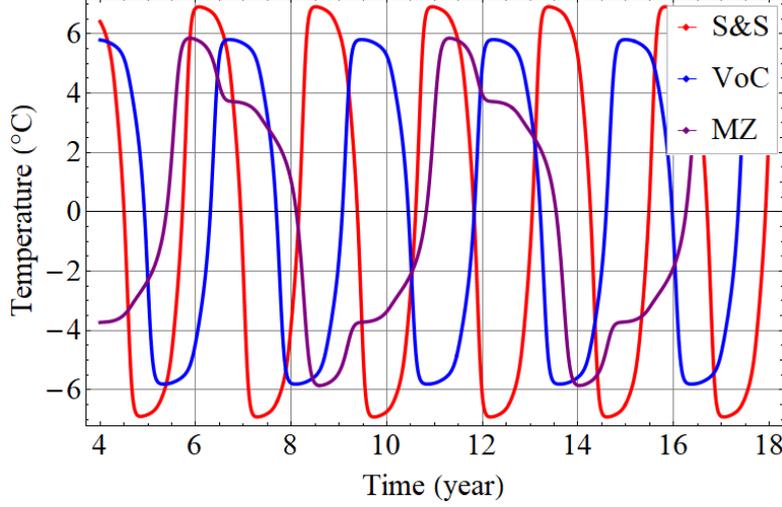


Figure 5: Model simulations of the different delay models for ENSO ($\alpha = 0.86$, $\gamma = 0.49$ and $\delta = 4.9$ after scaling). The model by Suarez and Schopf (red), the nonlinear model derived using variation of constants (blue) and the model derived using the Mori-Zwanzig approach (purple) are shown.

Before the models are studied in more detail, temperature and time are scaled to reduce the number of parameters. This allows for an easier treatment of the models in the following bifurcation analysis. On the other hand, the physical meaning of the results requires some translation in the end. In this section the sub- and superscripts of temperature are omitted, so T is written for T_e^E . Time is scaled by $t' = (c_S^* - c_T(1))t$ and temperature by $T' = \sqrt{\frac{\beta c_S^*}{c_S^* - c_T(1)}}T$. The scaled equation for the exact model of Equation (65), after dropping the primes, is

$$\frac{dT}{dt} = T(t) - T(t)^3 - \alpha T(t - \delta)(1 - \gamma T(t)^2), \quad (77)$$

where

$$\alpha = \frac{c_L^*}{c_S^* - c_T(1)}, \quad \gamma = \frac{c_S^* - c_T(1)}{c_S^*}, \quad \delta = (c_S^* - c_T(1))d. \quad (78)$$

Here the parameters are chosen such that for $\gamma = 0$ the model by Suarez and Schopf is recovered. Note that the scaled parameters do not depend on β . Only the temperature scale depends on this nonlinearity parameter. Since $c_S^* > c_T(1)$ all scaled parameters are positive and $\gamma < 1$. By the strong interdependence, it is difficult to relate values of the new parameters to values of the physical parameters. This issue is addressed at the end of the section. The scaling for the Mori-Zwanzig model of Equation (75) is the same.

3.4.1 Bifurcation Analysis

The exact delay model is analysed in the same way as was done by Suarez and Schopf [36]. First the steady states of Equation (77) are computed. Setting $\frac{dT}{dt} = 0$, three different equilibrium points are found:

$$T_{00} = 0, \quad T_{0\pm} = \pm \sqrt{\frac{1 - \alpha}{1 - \alpha\gamma}}. \quad (79)$$

To study the stability of these equilibrium points, a small perturbation T' from a steady state T_0 is considered. That is $T = T_0 + T'$. Substituting this into Equation (77) and considering only first order terms gives

$$\frac{dT'}{dt} = (1 - (3 - 2\alpha\gamma)T_0^2)T'(t) - \alpha(1 - \gamma T_0^2)T'(t - \delta). \quad (80)$$

For $T_0 = T_{00}$ the resulting linear delay model is not stable, indicating this is an unstable equilibrium. In the following $T_0 = T_{0\pm}$ is considered. Here, an exponential ansatz is made. This means $T' = \bar{T}e^{(\sigma_r + i\sigma_i)t}$ is substituted into the above equation. Splitting the resulting equation in a real and imaginary part yields

$$\begin{aligned}\sigma_r &= (1 - (3 - 2\alpha\gamma)T_0^2) - \alpha(1 - \gamma T_0^2) \cos(\sigma_i\delta)e^{-\sigma_r\delta}, \\ \sigma_i &= \alpha(1 - \gamma T_0^2) \sin(\sigma_i\delta)e^{-\sigma_r\delta}.\end{aligned}\tag{81}$$

To find where the equilibrium solutions change stability, the curves of neutral stability are computed by setting $\sigma_r = 0$. This way the frequency σ_i of the perturbation as a function of the parameters can be computed. The result is

$$\sigma_i = \sqrt{\alpha^2(1 - \gamma T_0^2)^2 - (1 - (3 - 2\alpha\gamma)T_0^2)^2},\tag{82}$$

and the corresponding two solutions for the delay time δ are

$$\begin{aligned}\delta_+ &= \frac{1}{\sigma_i} \left(2\pi k + \arccos \left(\frac{1 - (3 - 2\alpha\gamma)T_0^2}{\alpha(1 - \gamma T_0^2)} \right) \right), \\ \delta_- &= \frac{1}{\sigma_i} \left(2\pi k - \arccos \left(\frac{1 - (3 - 2\alpha\gamma)T_0^2}{\alpha(1 - \gamma T_0^2)} \right) \right),\end{aligned}\tag{83}$$

for a positive (or zero) integer k . Note that for $k = 0$ the two solutions are identical. In Figure 6 the neutral curves for $k = 0$ and $k = 1$, together with the dimensionless periods, are shown in α - δ -parameter space for two values of γ . The curve for $k = 0$ asymptotes to $\alpha = \frac{1}{\gamma}(1 - \sqrt{1 - \gamma})$ for large delay δ . Below the curve of $k = 0$, the non-zero steady states are stable and no oscillation occurs. Crossing the curve, a Hopf bifurcation occurs and a periodic solution emerges. At the curve with $k = 1$ another Hopf bifurcation occurs [23, 32], meaning a second periodic orbit emerges. This results in bistability of periodic orbits, which here has little effect on the period of the solutions.

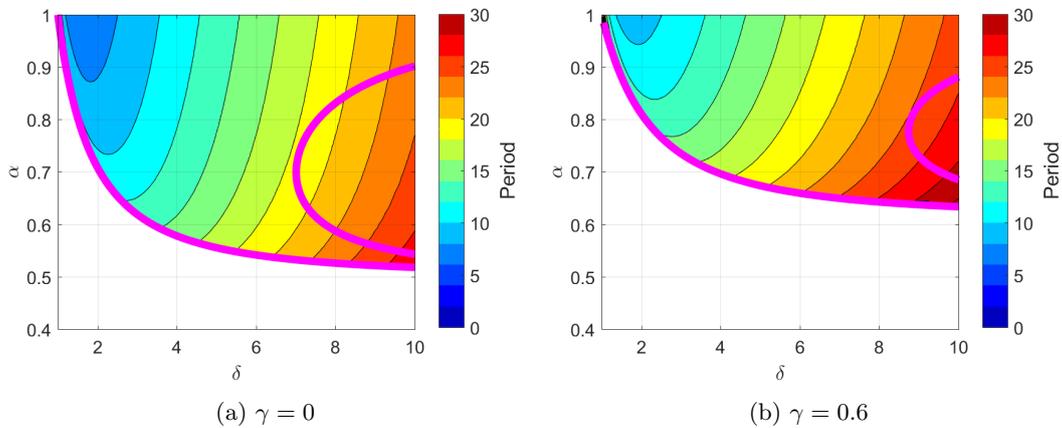


Figure 6: The (dimensionless) period of oscillation in α - δ -space for two different values of γ . The pink curves show the first and second Hopf bifurcation curves.

In Figure 7 the dimensionless period of the oscillations is shown in the full three dimensional α - δ - γ -parameter space. Note Figure 6 gives two slices of this plot for $\gamma = 0$ and $\gamma = 0.6$. The period increases as the delay time increases. Increasing the delay time means that the waves take longer to travel through the basin, and thus arrive later at the eastern side of the basin. Increasing γ leads to the shrinking of the region where oscillations occur. That means α has to be larger to get stable periodic solutions for the same delay time δ . Important to note is that for the same values of delay δ and parameter α , the period increases as γ increases. This shows that the derived delay model has a longer period compared to the original model by Suarez and Schopf, indicating it is closer to reality.

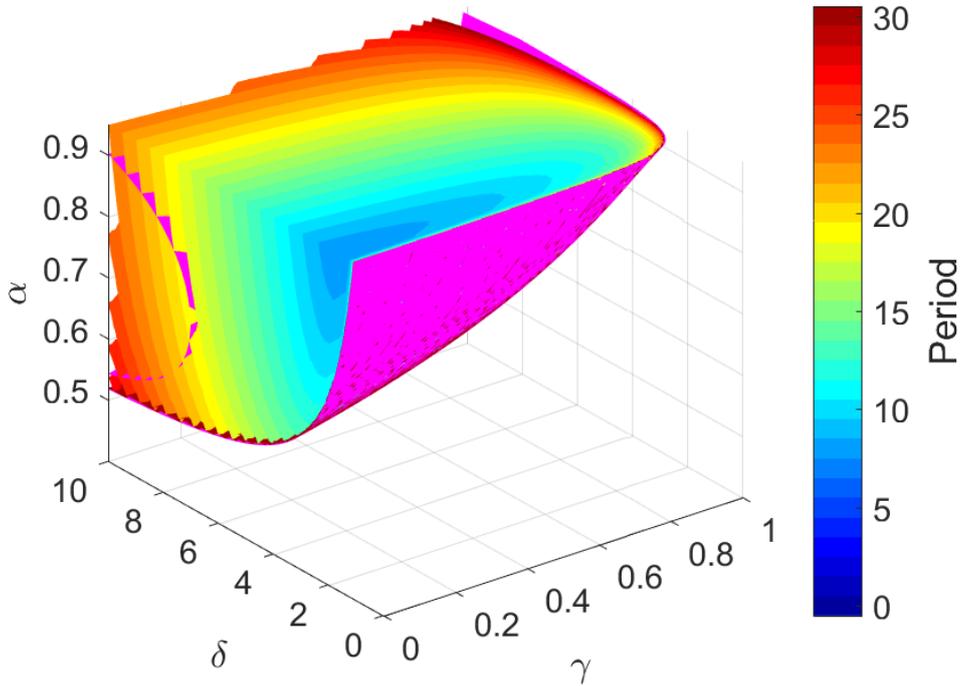


Figure 7: The (dimensionless) period of oscillation in α - δ - γ -space. The pink surfaces show the first and second Hopf bifurcation surfaces.

To investigate the errors made in applying the pseudo-orthogonal dynamics approximation to the Mori-Zwanzig formalism, a similar bifurcation analysis is performed on the scaled model of Equation (75). The result for $\gamma = 0.4$ is shown in Figure 8. Clearly the model derived using the approximate Mori-Zwanzig formalism exhibits a higher period than the exact model. Especially for large delays the difference is significant. Furthermore, the region in which oscillations are stable is smaller. As a consequence, it is less likely to find stable oscillations for realistic values of the parameters. Because of this last factor, this model is of little use in modeling the dynamics of ENSO.

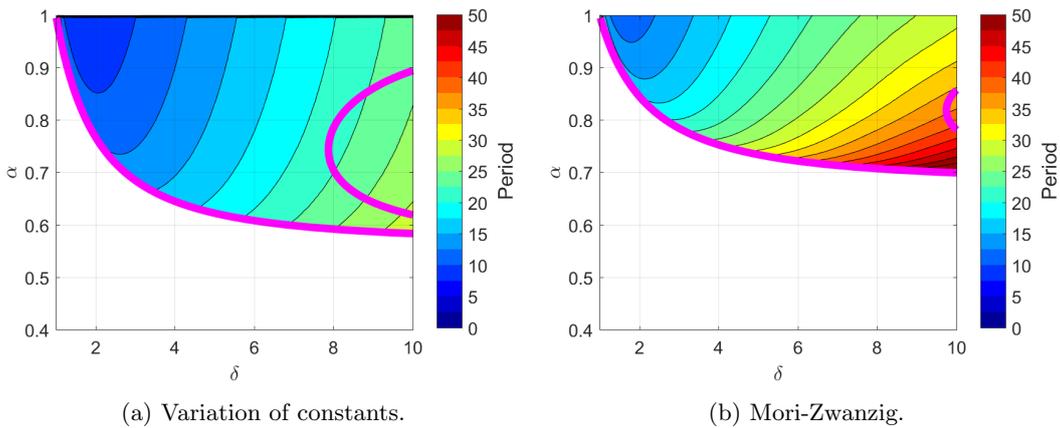


Figure 8: The (dimensionless) period of oscillation in α - δ -space for $\gamma = 0.4$ for the models derived using variation of constants and the approximated Mori-Zwanzig formalism. The pink curves show the first and second Hopf bifurcation curves.

3.4.2 Parameter Dependence

So far the discussion about the periodic behaviour has been given in abstract dimensionless scaled variables. To get an idea what realistic values are, the expressions for $c_T(x)$ and $c_h^*(x)$ from Dijkstra and Neelin are used [13]. They are

$$\begin{aligned} c_T(x) &= \epsilon_w + 0.5 \left((1 - \alpha_0) + (1 + \alpha_0) \tanh\left(\frac{\delta_F^1}{\epsilon} F(x)\right) \right) \delta_F^1 F(x), \\ c_h^*(x) &= 0.5 \left(\tanh\left(\frac{\delta_F^1}{\epsilon} F(x)\right) - 1 \right) \alpha_0 \delta_F^1 F(x) (T_0 - T_{s0}) \frac{H}{H^*}, \end{aligned} \quad (84)$$

with background wind forcing $F(x) = 0.6(0.12 - \cos(\frac{x-x_0}{2x_0}\pi)^2)$ for $x_0 = 0.57$ and parameters $\epsilon_w = \frac{\epsilon_T L}{c_0}$, $\alpha_0 = \frac{H_1}{\tilde{H}}$, $\delta_F^1 = \frac{\tau_0 L}{c_0} \frac{b_w}{H_1}$. The dimensional values are given in Table 1. The parameters are considered at $x = 0.9$ to avoid effects of the boundary layer in the background state.

Table 1: Parameter values used in determining c_T and c_h^* .

Damping scale Newtonian cooling	ϵ_T	$9.25 \cdot 10^{-8} \text{ s}^{-1}$
Basin length	L	$1.5 \cdot 10^7 \text{ m}$
Velocity first baroclinic Kelvin mode	c_0	2 m/s
Background wind forcing strength	τ_0	$2.667 \cdot 10^{-7} \text{ m/s}^2$
Parametrization constant	b_w	$1.026 \cdot 10^2 \text{ s}$
Depth surface layer	H_1	50 m
Depth top layer	H	200 m
Depth for temperature gradient	\tilde{H}	50 m
Steepness transition subsurface temperature	H^*	30 m
Temperature without dynamics	T_0	30 °C
Background subsurface temperature	T_{s0}	22 °C
Rayleigh friction coefficient	a_M	$1.3 \cdot 10^{-8} \text{ s}^{-1}$
Scaling parameter	ϵ	10^{-4}

The other dimensionless variables in the definitions of c_S^* and c_L^* are μ , ϵ_0 , A_{rW} , θ , y_n , A_0 and x_w . For the first three parameters the book *Nonlinear Physical Oceanography* by Dijkstra is followed [12]. Here μ is the coupling strength between the wind and the ocean for which a realistic value is $\mu = 1$. For the damping coefficient $\epsilon_0 = \frac{a_M L}{c_0}$ is taken. Recall that A_{rW} is defined by Equation (50), for which $r_W = 3/5$ is used. Lastly $x_w = 0.6$ is chosen as the location where the wind has its strongest effect. The other parameters are varied according to Table 2 to study the dependence of the period. Note that these values differ from the those in Jin [24] due to a different scaling.

Table 2: The ranges in which the nondimensional parameters are varied.

Parameter		Dimensional	Dimensionless	Step
Wind forcing factor at y_n	θ	-	1.0 - 4.0	0.2
Wind forcing strength	A_0	$0.5 - 3.0 \cdot 10^{-2} \text{ Pa}$	0.1 - 0.6	0.05
Latitude Rossby waves	y_n	$5.0^\circ - 12.1^\circ$	1.4 - 3.4	0.2

Varying the parameters in the ranges given in Table 2, the period of oscillation is computed for both the exact derived model and the model by Suarez and Schopf. In Figure 9 the periods of oscillation for these models are shown versus one another. The period of the derived model is larger in every situation where still an oscillation is present. There are locations in parameter space where no oscillation occurs in the derived model, while they do occur in the model by Suarez and Schopf. This already was indicated by the smaller region of oscillation for larger γ in Figure 7.

The dependence of the period on the different parameters is shown in Figure 10. For increasing θ , that is, when the effect of the wind forcing at higher latitudes increases, the period slowly increases. There turns out to be a minimum value for θ around 1.7 below which no oscillations occur. In that case the signal at higher latitudes is too weak to have a significant effect at the

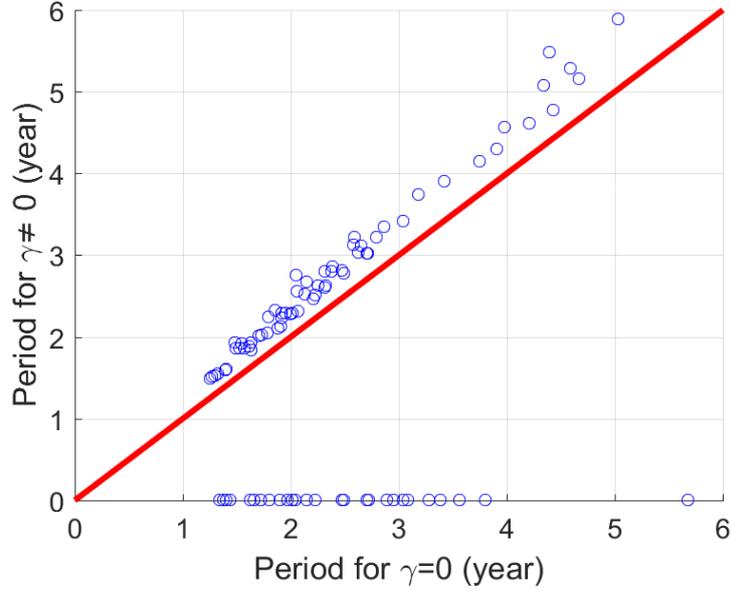


Figure 9: The period of the model derived using variation of constants versus the period according to the model by Suarez and Schopf. The line of equal period is shown in red.

eastern boundary. Considering the strength of the wind forcing A_0 , a stronger wind results in a shorter period. This can be due to the larger absolute difference between the effect at the equator and at higher latitudes, leading to a weaker effect of the latter. This decrease with increasing wind strength appears to be approximately exponential. For a realistic A_0 in the centre of the range, the period of the oscillation is approximately 2.5 to 3.5 years. This is still smaller than that of ENSO.

Looking at the latitude y_n at which the Rossby waves travel, instead of the latitude itself, $1/y_n^2$ is plotted, since this gives the velocity of the Rossby wave traveling at that latitude. For higher velocities, so lower latitudes, the oscillations have a smaller period. The faster the wave travels, the shorter the delay is, resulting in a shorter period. Similarly, slow waves result in longer periods.

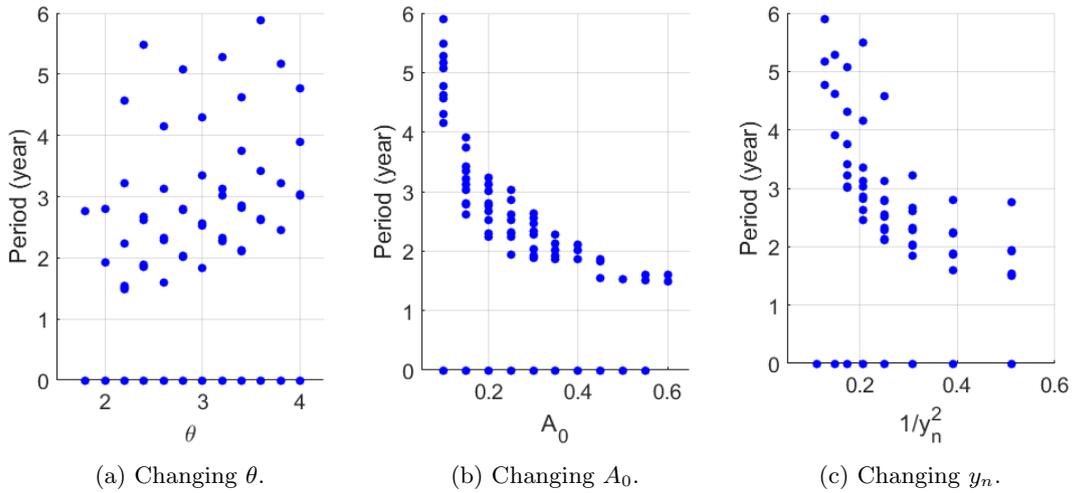


Figure 10: The dependence of the period of the oscillation on the parameters θ , A_0 and y_n for the exact derived model. The range for each value of one of the parameters is due to the variation of the other two parameters.

4 Atlantic Multidecadal Oscillation

The Atlantic Multidecadal Oscillation (AMO) is an oscillation in the sea surface temperature in the North-Atlantic Ocean with a period of fifty to seventy years [12]. In Figure 11 the sea surface temperature deviations in the North Atlantic Ocean are shown for the last 160 years. Even though there is not enough data to show multiple oscillations, there is evidence of a clear oscillation with a period of several decades. In addition, high resolution climate models show an oscillation on this same timescale. This phenomenon has not been studied as long or extensively as ENSO, because of its relatively recent discovery.

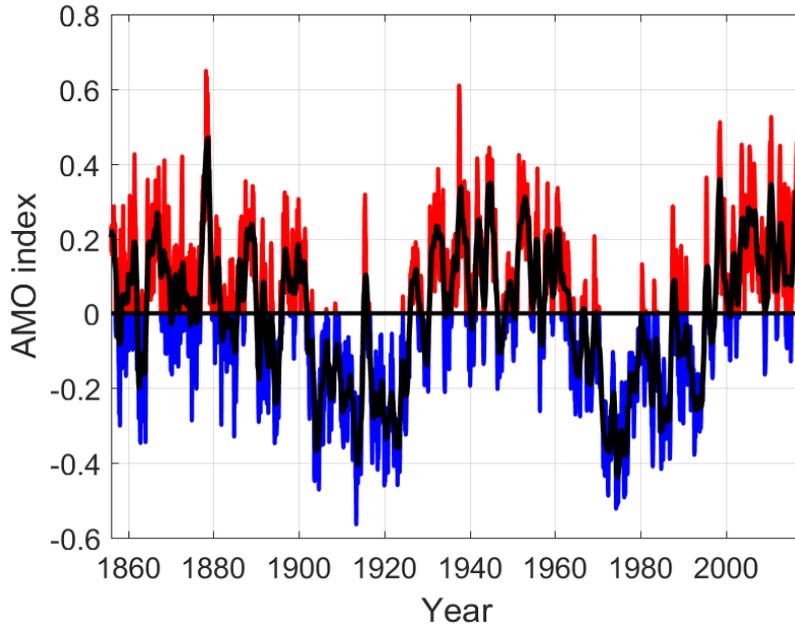


Figure 11: The deviations in sea surface temperature in the North Atlantic Ocean for the last 160 years. In black the 12-monthly running mean is shown. Data from the NCEP/NCAR reanalysis project [26].

The background state against which this oscillation occurs is that of the Atlantic Meridional Overturning Circulation (AMOC) [10]. In the south of the basin, near the equator, the water is heated by the sun. Via the Gulf Stream and other ocean currents this water is transported northward (and eastward), losing heat on its way. Near the poles the water is cooled down so much that it sinks. Near the bottom of the ocean it then flows southward. Also in the zonal direction such an overturning circulation occurs, but it is much weaker than the meridional flow.

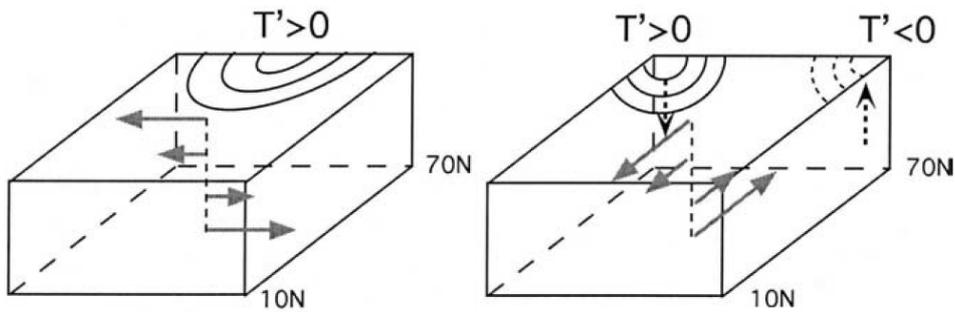


Figure 12: The physical mechanism responsible for the Atlantic Multidecadal Oscillation. Figure taken from Dijkstra (2006) [11].

The physical mechanism behind the AMO, as described in [39], depends on thermal wind balance, which states that flow is along temperature (and thus pressure) gradients. When there is a positive temperature perturbation in the northern-central part of the basin, there is a north-south temperature gradient. This meridional temperature gradient results in zonal overturning with westward surface flow through thermal wind balance. This is shown in the left part of Figure 12. The negative zonal flow transports the positive temperature anomaly towards the western boundary, creating a zonal temperature gradient. Again through thermal wind balance, this now leads to perturbations in the meridional overturning circulation (right part of Figure 12). This flow transports cold water from near the poles southward, reducing the meridional temperature gradient. This smaller north-south temperature gradient causes a positive (eastward) zonal flow, after which the same pattern as described above is followed with a sign change.

Similar to the procedure described in Section 3, it is investigated whether the AMO can be described by a delay model. However, here no clear traveling wave transporting temperature is part of the process. This could indicate that a delay equation may not be a feasible model for the AMO. No delay model for the AMO has yet been proposed.

4.1 AMO Model

In this section first the AMO model to which the Mori-Zwanzig formalism will be applied is formulated. The strengths and weaknesses of the model are discussed and a possible extension is given.

4.1.1 Model Formulation

To derive a delay model, the model for the AMO derived by Sévellec and Huck (2015) is taken as a starting point [34]. In the derivation of this three-layer model, they consider a background state with a zonal velocity \bar{u} and meridional and vertical temperature gradients ($\partial_y \bar{T}$ and $\partial_z \bar{T}$). No background meridional or vertical velocity is assumed, since the zonal flow, which is forced by the wind, is much stronger. Equations for the perturbations from this background state are considered. For temperature this yields

$$\partial_t T = -\bar{u} \partial_x T - v \partial_y \bar{T} - w \partial_z \bar{T} + \kappa \partial_{xx} T, \quad (85)$$

where κ is the horizontal eddy diffusivity coefficient.

For the velocities perturbations to the geostrophic and hydrostatic equations are considered. The geostrophic equation represents a balance between the Coriolis force and horizontal pressure gradients, while the hydrostatic equation gives an equilibrium between gravity and the vertical pressure gradient [7]. No inertia is taken into account. This results in the thermal wind balance:

$$f \partial_z v = \alpha_T g \partial_x T, \quad (86)$$

where f is the Coriolis parameter, α_T the thermal expansion coefficient and g the acceleration of gravity. Since it is assumed that $\partial_y T \approx 0$, this is the only part of the thermal wind balance that is present. This assumption also implies that $u = 0$. Vertical integration of the continuity equation ($\nabla \cdot \vec{u} = 0$), while assuming no vertical velocity at the surface and bottom, gives the baroclinic condition:

$$\int_{-H}^0 v dz = 0, \quad (87)$$

for an ocean of depth H . Discretization of the model over three vertical layers (of depth h_1 , h_2 and h_3), results in expressions for the meridional velocity in each layer via Equation (86). So these expressions are in terms of the x -derivatives of temperature in the different layers. Using the Sverdrup balance [7]

$$\beta v = f \partial_z w, \quad (88)$$

which employs the β -plane approximation of the Coriolis parameter $f = f_0 + \beta y$, results in expressions for the vertical velocities in the top two layers in a similar way.

Table 3: The values of the parameters in the AMO model by Sévellec and Huck [34].

Thickness layer one	h_1	600 m	Vertical temperature gradient
Thickness layer two	h_2	600 m	
Thickness layer three	h_3	3300 m	$\partial_z \bar{T} = -\frac{2c}{h_1+h_2} (\Delta T - \frac{\alpha_S}{\alpha_T} \Delta S)$
Total ocean depth	H	4500 m	
Zonal basin size	W	4.000 km	Control parameter c Standard $c = 1$
Meridional basin size	L	6500 km	
Horizontal diffusivity	κ	$2 \cdot 10^3 \text{ m}^2/\text{s}$	Meridional temperature gradient
Acceleration of gravity	g	9.8 m/s^2	
Coriolis parameter	f	10^{-4} s^{-1}	$\partial_y \bar{T} = \frac{2}{L} (\Delta T - \frac{\alpha_S}{\alpha_T} \Delta S)$
β effect	β	$1.5 \cdot 10^{-11} \text{ (ms)}^{-1}$	
Thermal expansion coefficient	α_T	$2 \cdot 10^{-4} \text{ K}^{-1}$	Meridional temperature gradient
Haline contraction coefficient	α_S	$7 \cdot 10^{-4} \text{ psu}^{-1}$	
Meridional temperature diff.	ΔT	-20 K	$\partial_y \bar{T} = \frac{2}{L} (\Delta T - \frac{\alpha_S}{\alpha_T} \Delta S)$
Meridional salinity diff.	ΔS	-1.5 psu	
Zonal velocity	\bar{u}	10^{-2} m/s	

Substituting the expressions for the velocities in the different layers into the temperature equation leads to a three-layer temperature model. It is assumed there is no background flow, nor a background temperature gradient in the bottom layer. The resulting model by Sévellec and Huck is

$$\begin{aligned}
 \partial_t T_1 &= a_1 \partial_x T_1 + b_1 \partial_x T_2 + c_1 \partial_x T_3 + \kappa \partial_{xx} T_1, \\
 \partial_t T_2 &= a_2 \partial_x T_1 + b_2 \partial_x T_2 + c_2 \partial_x T_3 + \kappa \partial_{xx} T_2, \\
 \partial_t T_3 &= \kappa \partial_{xx} T_3,
 \end{aligned} \tag{89}$$

with boundary conditions

$$T_i|_{West} = -T_i|_{East}, \quad i = 1, 2, 3. \tag{90}$$

The derivation of the boundary conditions can be found in the appendix to the article by Sévellec and Huck [34]. The constants in the model are all positive for physically realistic values and defined by

$$\begin{aligned}
 a_1 &= \frac{\alpha_T g}{2Hf} \left(-h_1(h_2 + h_3) \partial_y \bar{T} + \frac{\beta}{2f} h_1^2 (h_2 + h_3) \partial_z \bar{T} \right) - \bar{u}, \\
 b_1 &= \frac{\alpha_T g}{2Hf} \left(-h_2(h_2 + 2h_3) \partial_y \bar{T} + \frac{\beta}{2f} h_1 h_2 (h_2 + 2h_3) \partial_z \bar{T} \right), \\
 c_1 &= \frac{\alpha_T g}{2Hf} \left(-h_3^2 \partial_y \bar{T} + \frac{\beta}{2f} h_1 h_3^2 \partial_z \bar{T} \right), \\
 a_2 &= \frac{\alpha_T g}{2Hf} \left(h_1^2 \partial_y \bar{T} + \frac{\beta}{2f} h_1^2 (h_2 + 2h_3) \partial_z \bar{T} \right), \\
 b_2 &= \frac{\alpha_T g}{2Hf} \left(-h_2(h_3 - h_1) \partial_y \bar{T} + \frac{\beta}{2f} (4h_1 h_2 h_3 + h_2^2 (h_1 + h_3)) \partial_z \bar{T} \right) - \bar{u}, \\
 c_2 &= \frac{\alpha_T g}{2Hf} \left(-h_3^2 \partial_y \bar{T} + \frac{\beta}{2f} h_3^2 (2h_1 + h_2) \partial_z \bar{T} \right).
 \end{aligned} \tag{91}$$

The values of the parameters are given in Table 3.

The next step is to scale the equations. This is done by defining new variables $x = Wx'$ and $t = Yt'$ for space and time, with W as in Table 3 and Y a year (in seconds). The parameters then change by a factor $\frac{Y}{W}$ for a_i , b_i and c_i , e.g. $a'_i = \frac{Y}{W} a_i$. For the diffusion term the factor is $\frac{Y}{W^2}$. From here onwards the scaled model is considered, where the primes of the scaling are dropped.

Note in Equation (89) that the only term acting in the third layer is diffusion, since the bottom layer is assumed to be at rest. Therefore, any perturbation damps out. For this reason,

and to simplify the mathematical treatment of the system, it is assumed that there are no perturbations in the bottom layer ($T_3 = 0$). Furthermore, the diffusion terms are approximated by linear damping with damping coefficient α . Under these two approximations the system of Equation (89) becomes a two-layer system:

$$\begin{aligned}\partial_t T_1 &= a_1 \partial_x T_1 + b_1 \partial_x T_2 - \alpha T_1, \\ \partial_t T_2 &= a_2 \partial_x T_1 + b_2 \partial_x T_2 - \alpha T_2.\end{aligned}\tag{92}$$

This is the model for the AMO that is studied in this AMO section and to which the Mori-Zwanzig formalism is applied.

Before continuing the study of this model, first it is simulated without damping. This is done using an upwind discretization scheme for the x -derivatives and a forward Euler scheme in time. Note that this discretization includes numerical diffusion. The result for $dx = 0.0025$ (corresponding to 10 km), with an initial positive Gaussian temperature perturbation in the centre of the basin, is shown in Figure 13. Note the opposite sign of the temperature in the two layers, which is due to the baroclinic nature of the waves [7]. The model shows a combination of two oscillations with different periods. First, there is a long period of approximately sixty years, which corresponds to a thermal Rossby wave responsible for driving the AMO. Secondly, there is a high frequency oscillation with a period of around five years. This short period oscillation does not correspond to a planetary Rossby wave, as one might expect, since decreasing β does not result in a disappearance of these oscillations. It is a thermal Rossby wave, just as the one responsible for the AMO oscillation.

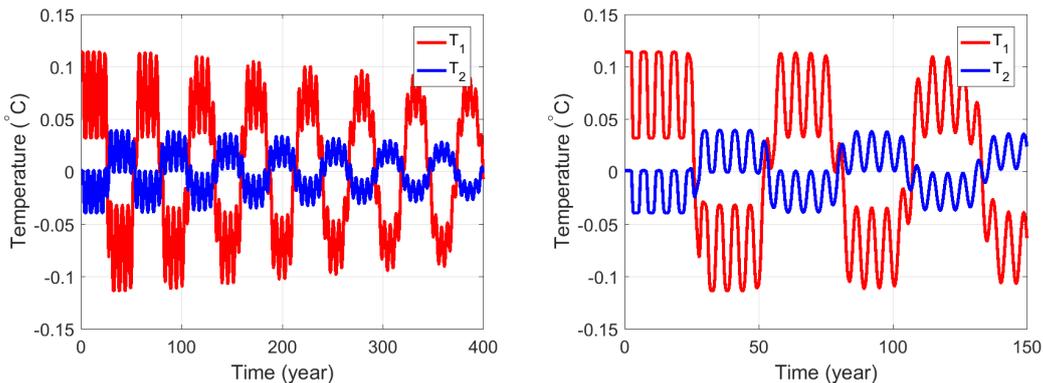


Figure 13: Model simulation of the AMO model in Equation (92) for an initial positive Gaussian temperature perturbation in the centre of the basin. In the right figure a zoom of the left figure for the first 150 years is shown.

The occurrence of this short period is surprising, as it is not known from literature. In Figure 14 the spectrum of the AMO is shown. This spectrum is computed by calculating the discrete Fourier transform of the data. The two clear peaks correspond to periods of one year, which is seasonality, and of approximately seventy years, which is the AMO. Between those two peaks there appears to be a small peak corresponding to a period of approximately eight years, which is on the same order has the short period in the model. However, this peak can also be due to noise in the data. The dominant appearance of this mode in the model simulations thus does not clearly correspond to the data of the AMO. It could be an artificial feature of the model.

This AMO model of temperature oscillations also provides information about oscillations in the zonal and meridional overturning circulation. By the thermal wind balance in Equation (86), the vertical shear in the meridional flow can be related to the zonal temperature gradient. This results in a quarter of a phase difference between the two. The vertical shear in the meridional flow indicates whether there is a positive or negative perturbation in the meridional overturning. A positive perturbation in $\partial_z v$ means more northward flow at the surface compared to the bottom, and thus a positive perturbation in the overturning circulation. Similarly a negative perturbation in $\partial_z v$ corresponds to a negative overturning perturbation.

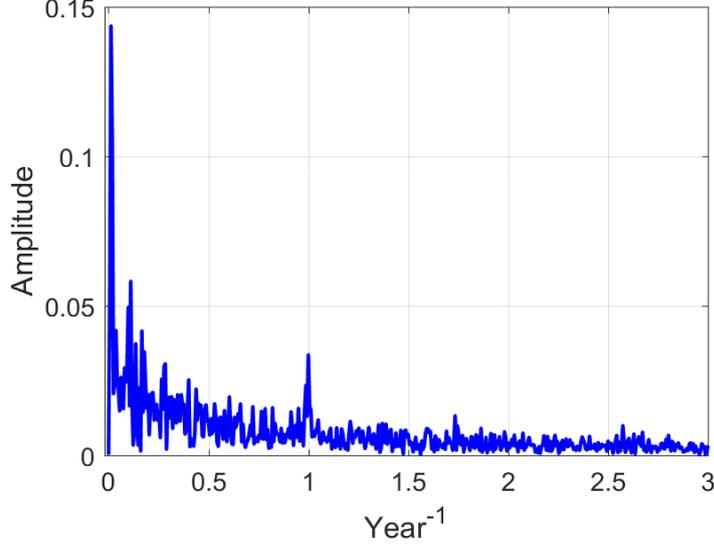


Figure 14: The spectrum of the AMO as computed by the discrete Fourier transform. Data from the NCEP/NCAR reanalysis project [26].

To get an idea of the behaviour of perturbations in the zonal overturning, the y -averaged continuity equation is considered. Here it is assumed that there is no flow through the boundaries of the basin, so $v|_{South} = v|_{North} = 0$. Using the Sverdrup balance from Equation (88), after taking the z -derivative of the integrated continuity equation, gives the following equation:

$$\partial_x(\partial_z u) = -\frac{\beta}{f}\partial_z v = -\frac{\beta}{f}\frac{\alpha_T g}{f}\partial_x T. \quad (93)$$

This indicates there is a difference of half a phase between zonal overturning perturbations and temperature perturbations. The phase difference between the zonal and meridional overturning perturbations is a quarter phase, as is expected from literature [12].

Computing the evolution of the meridional and zonal perturbations in the vertical shear from the temperature deviations, yields the results seen in Figure 15. For both the short and long period oscillations, the quarter and half phase difference between the temperature oscillations and the meridional and zonal overturning perturbations respectively can be seen. A positive peak in temperature coincides with a negative zonal overturning perturbation, which transports this perturbation westward. Physically a short delay between these two would be expected, but due to the assumption of immediate thermal wind balance this is not present in the model. The resulting zonal temperature gradient causes a negative meridional overturning perturbation with equatorward surface flow, lagging by a quarter phase. This is followed by a negative temperature perturbation and a positive zonal overturning, inducing a positive meridional overturning perturbation, after which the oscillation starts again. Apart from the timing of the peaks in temperature and zonal overturning perturbations, this is exactly the physical process of the AMO as described in the beginning of Section 4.

4.1.2 Background Overturning Circulation

The AMO model by Sévellec and Huck does not contain a background overturning circulation, as the meridional and vertical flow is neglected in the background state. This results in the presence of a negative overturning circulation in the model. Adding background meridional and vertical velocities (\bar{v} and \bar{w}) yields the following perturbation equation for temperature:

$$\partial_t T = -\bar{u}\partial_x T - \bar{v}\partial_y T - \bar{w}\partial_z T - v\partial_y \bar{T} - w\partial_z \bar{T} + \kappa\partial_{xx} T. \quad (94)$$

If $\partial_y T$ is taken into account, the two equations for thermal wind balance are

$$f\partial_z v = \alpha_T g\partial_x T, \quad f\partial_z u = -\alpha_T g\partial_y T. \quad (95)$$

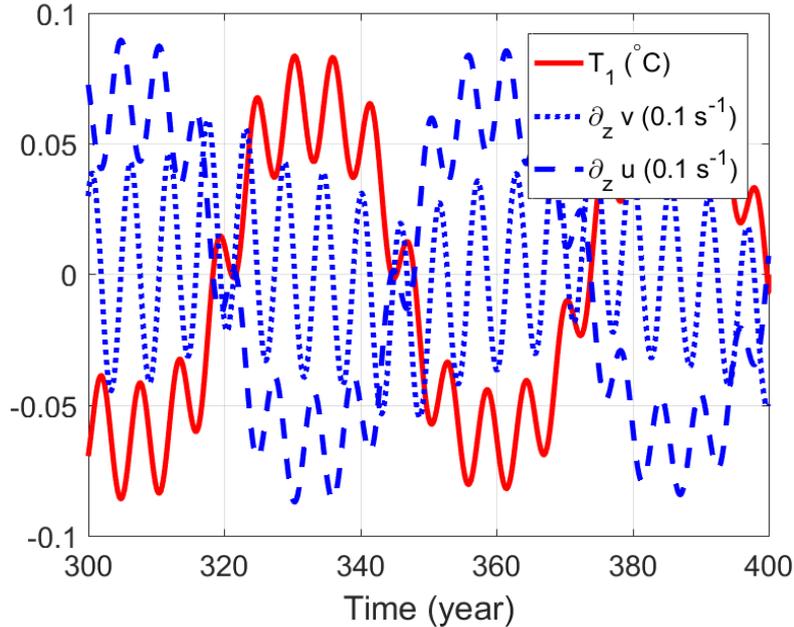


Figure 15: The evolution of the vertical shear in the meridional and zonal direction (blue) together with the temperature oscillations (red) in the top layer.

This means the partial derivative of temperature to y can be expressed in terms of $\partial_z u$. In addition to the previous equations, also expressions for $\partial_z T$ and $\partial_z u$ in the three different layers are needed. For u this is done by assuming a baroclinic condition, similar to Equation (87). For T the vertical derivatives in the different layers can be computed if it is assumed there are no temperature perturbations at the bottom ($T|_{-H} = 0$).

If, as before, it is assumed that no perturbations in the third layer are present, the model reduces again to a two-layer model. A model simulation, with $\bar{v} = 0.5 \cdot 10^{-2}$ m/s and $\bar{w} = -0.17 \cdot 10^{-6}$ m/s, is shown in Figure 16. The result of adding the background overturning circulation, is a damping of the high frequency oscillation. The short period oscillation is absorbed by the background overturning circulation, while the long period oscillation persists. The amplitude of the oscillation corresponding to the AMO appears to be barely affected by the damping. This can be due to the presence of an amplifying effect of the background overturning in some parts of the equations.

In the following sections the AMO model without background overturning is considered, to simplify the computations. However, one has to keep in mind that the high frequency oscillation present in the model, does not sustain in the presence of a background overturning.

4.2 Mori-Zwanzig formalism

For the ENSO model, as described in Section 3.2, the procedure of projecting onto the temperature at one location presented itself in a quite straight-forward manner. For the AMO model from Section 4.1 it is more difficult to see how this projection is done.

The model is linear, meaning it is possible to formally write down a solution for T_2 using variation of constants and substitute this into the equation for T_1 . This yields

$$\partial_t T_1 = a_1 \partial_x T_1 - \alpha T_1 + b_1 \partial_x \left(e^{(b_2 \partial_x - \alpha)t} T_2(0) + \int_0^t e^{(b_2 \partial_x - \alpha)(t-s)} a_2 \partial_x T_1(s) ds \right). \quad (96)$$

Here the application of integration along characteristics to get rid of the exponentials is not sufficient, because the exponential terms act on the x -derivative of T_1 instead of T_1 itself. By the nature of the equations, there are waves traveling in the system, but by rewriting the system this way it is difficult to see whether a delay is present. Note that because of the x -derivative to T_1 in the equation, it is also difficult to arrive at an equation for temperature at one location.

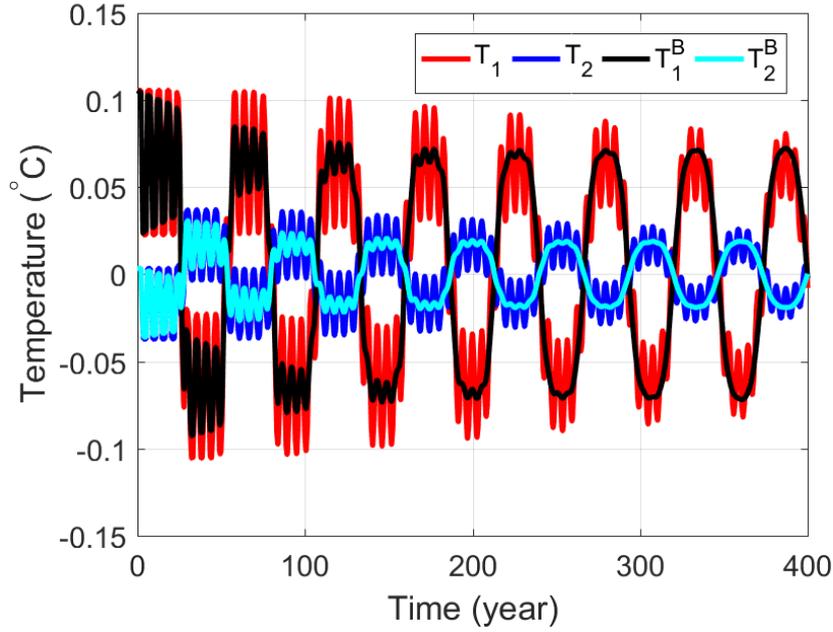


Figure 16: Simulations of the temperature in the two layers without (red, blue) and with (black, cyan) a background overturning circulation.

The way in which is proceeded here, is to first change the partial differential equations into ordinary differential equations by discretization. On the resulting large system of ordinary differential equations the Mori-Zwanzig formalism is applied. In the following sections this procedure is described. In the end this results in a delay model for the AMO. The nature of this model is quite different from the model found for ENSO. An alternative method is to use a change of variables. This way a system with two separated variables can be found, after which integration along characteristics can be applied.

4.2.1 Discretization

The first step is to find a stable discretization of the AMO model in Equation (92). Since all parameters in the model are positive, the direction in which the waves travel is known. This indicates the best way to discretize the model. A grid of $N + 1$ -points in space with distance $dx = \frac{1}{N}$ is used. Since all waves travel westward, an upwind discretization scheme is used. The discretized equations are

$$\begin{aligned}\partial_t T_1^n &= \frac{a_1}{dx}(T_1^{n+1} - T_1^n) + \frac{b_1}{dx}(T_2^{n+1} - T_2^n) - \alpha T_1^n, \\ \partial_t T_2^n &= \frac{a_2}{dx}(T_1^{n+1} - T_1^n) + \frac{b_2}{dx}(T_2^{n+1} - T_2^n) - \alpha T_2^n,\end{aligned}\tag{97}$$

for $n = 0, \dots, N$, with boundary conditions

$$\begin{aligned}T_1^N &= -T_1^0, \\ T_2^N &= -T_2^0.\end{aligned}\tag{98}$$

Because of the circular nature of the boundary conditions, this is a $2N$ -dimensional system. Let $\vec{T} = (T_1^0, T_2^0, \dots, T_1^{N-1}, T_2^{N-1})$. The system from Equation (97) can be written as a matrix equation:

$$\partial_t \vec{T} = M \vec{T},\tag{99}$$

To solve this system, the first step is to compute the eigenvalues and (generalized) eigenvectors. The eigenvalues of M_Q are

$$\lambda_{\pm} = -\alpha - \frac{l_{\pm}}{dx}, \quad (103)$$

where

$$l_{\pm} = \frac{1}{2} \left(a_1 + b_2 \pm \sqrt{a_1^2 + b_2^2 - 2a_1b_2 + 4a_2b_1} \right). \quad (104)$$

Each of them has multiplicity $N - 1$. The corresponding generalized eigenvectors of M_Q for $i = 1, \dots, N - 1$ are

$$\vec{v}_{\pm}^i = \left(\frac{dx}{l_{\pm}} \right)^{i-1} \cdot (0, \dots, 0, w_{\pm}, 1, 0, \dots, 0), \quad (105)$$

where the non-zero values are located on the coordinates corresponding to location i , and with

$$w_{\pm} = \frac{1}{2a_2} \left(a_1 - b_2 \pm \sqrt{a_1^2 + b_2^2 - 2a_1b_2 + 4a_2b_1} \right). \quad (106)$$

Using the eigenvalues and generalized eigenvectors, the solution to the orthogonal dynamics system can be found [14]. It is

$$\begin{aligned} \vec{T}_Q(t) = & e^{\lambda_+ t} \left(c_+^1 \vec{v}_+^1 + c_+^2 (t\vec{v}_+^1 + \vec{v}_+^2) + \dots + c_+^i \left(\frac{t^{i-1}}{(i-1)!} \vec{v}_+^1 + \frac{t^{i-2}}{(i-2)!} \vec{v}_+^2 + \dots + \vec{v}_+^i \right) \right. \\ & \left. + \dots + c_+^{N-1} \left(\frac{t^{N-2}}{(N-2)!} \vec{v}_+^1 + \dots + \vec{v}_+^{N-1} \right) \right) \\ & + e^{\lambda_- t} \left(c_-^1 \vec{v}_-^1 + c_-^2 (t\vec{v}_-^1 + \vec{v}_-^2) + \dots + c_-^i \left(\frac{t^{i-1}}{(i-1)!} \vec{v}_-^1 + \frac{t^{i-2}}{(i-2)!} \vec{v}_-^2 + \dots + \vec{v}_-^i \right) \right. \\ & \left. + \dots + c_-^{N-1} \left(\frac{t^{N-2}}{(N-2)!} \vec{v}_-^1 + \dots + \vec{v}_-^{N-1} \right) \right). \end{aligned} \quad (107)$$

Here the constants c_{\pm}^i are determined by the initial conditions. Each generalized eigenvector has only components in the directions of T_1^i and T_2^i . This means that, to find expressions for the constants, the following system has to be solved for each i :

$$\begin{aligned} c_+^i w_+ \left(\frac{dx}{l_+} \right)^{i-1} + c_-^i w_- \left(\frac{dx}{l_-} \right)^{i-1} &= T_1^i(0), \\ c_+^i \left(\frac{dx}{l_+} \right)^{i-1} + c_-^i \left(\frac{dx}{l_-} \right)^{i-1} &= T_2^i(0). \end{aligned} \quad (108)$$

The solution is

$$\begin{aligned} c_+^i &= \left(\frac{l_+}{dx} \right)^{i-1} \cdot \frac{T_1^i(0) - w_- T_2^i(0)}{w_+ - w_-}, \\ c_-^i &= - \left(\frac{l_-}{dx} \right)^{i-1} \cdot \frac{T_1^i(0) - w_+ T_2^i(0)}{w_+ - w_-}. \end{aligned} \quad (109)$$

This way an analytical solution to the orthogonal dynamics equation for general initial conditions has been found.

This general solution is used to compute first the noise terms, and subsequently the memory terms. The noise terms for the AMO model are defined by

$$\begin{aligned} F_{T_1^0}(t) &= \frac{a_1}{dx} T_{1Q}^1(t) + \frac{b_1}{dx} T_{2Q}^1(t), \\ F_{T_2^0}(t) &= \frac{a_2}{dx} T_{1Q}^1(t) + \frac{b_2}{dx} T_{2Q}^1(t). \end{aligned} \quad (110)$$

Only the generalized eigenvectors v_{\pm}^1 in the solution of the orthogonal dynamics equation contribute to the components of the noise terms in the above equations. The resulting equations are

$$\begin{aligned}
F_{T_1^0}(t) &= a_1 N \left(w_+ e^{\lambda+t} \sum_{i=1}^{N-1} c_+^i \frac{t^{i-1}}{(i-1)!} + w_- e^{\lambda-t} \sum_{i=1}^{N-1} c_-^i \frac{t^{i-1}}{(i-1)!} \right) \\
&\quad + b_1 N \left(e^{\lambda+t} \sum_{i=1}^{N-1} c_+^i \frac{t^{i-1}}{(i-1)!} + e^{\lambda-t} \sum_{i=1}^{N-1} c_-^i \frac{t^{i-1}}{(i-1)!} \right), \\
F_{T_2^0}(t) &= a_2 N \left(w_+ e^{\lambda+t} \sum_{i=1}^{N-1} c_+^i \frac{t^{i-1}}{(i-1)!} + w_- e^{\lambda-t} \sum_{i=1}^{N-1} c_-^i \frac{t^{i-1}}{(i-1)!} \right) \\
&\quad + b_2 N \left(e^{\lambda+t} \sum_{i=1}^{N-1} c_+^i \frac{t^{i-1}}{(i-1)!} + e^{\lambda-t} \sum_{i=1}^{N-1} c_-^i \frac{t^{i-1}}{(i-1)!} \right),
\end{aligned} \tag{111}$$

where dx has been replaced by $\frac{1}{N}$.

To be able to compute the memory terms, it has to be investigated what the operator $P\mathcal{L}$ does when it acts on the noise terms. Since the noise terms are linear in the initial conditions, it is sufficient to consider the effect of this operator on each of these conditions separately. Acting on each of the initial conditions gives

$$\begin{aligned}
&P\mathcal{L}(T_1^1(0), T_2^1(0), \dots, T_1^i(0), T_2^i(0), \dots, T_1^{N-1}(0), T_2^{N-1}(0)) \\
&= P(\dots, \frac{a_1}{dx}(T_1^{i+1}(0) - T_1^i(0)) + \frac{b_1}{dx}(T_2^{i+1}(0) - T_2^i(0)) - \alpha T_1^i(0), \\
&\quad \frac{a_2}{dx}(T_1^{i+1}(0) - T_1^i(0)) + \frac{b_2}{dx}(T_2^{i+1}(0) - T_2^i(0)) - \alpha T_2^i(0), \dots) \\
&= (0, \dots, 0, -\frac{a_1}{dx}T_1^0(0) - \frac{b_1}{dx}T_2^0(0), -\frac{a_2}{dx}T_1^0(0) - \frac{b_2}{dx}T_2^0(0)).
\end{aligned} \tag{112}$$

This means only the terms that, before the application of the operator $P\mathcal{L}$, depend on $T_1^{N-1}(0)$ and $T_2^{N-1}(0)$ remain in the memory kernel. Those are the terms from the general solution which contain c_{\pm}^{N-1} . Replacing dx by $\frac{1}{N}$, the memory integrands are

$$\begin{aligned}
K_{T_1^0}(T_1^0(0), T_2^0(0), t) &= P\mathcal{L} \left(N(a_1 w_+ + b_1) e^{\lambda+t} c_+^{N-1} \frac{t^{N-2}}{(N-2)!} \right. \\
&\quad \left. + N(a_1 w_- + b_1) e^{\lambda-t} c_-^{N-1} \frac{t^{N-2}}{(N-2)!} \right), \\
K_{T_2^0}(T_1^0(0), T_2^0(0), t) &= P\mathcal{L} \left(N(a_2 w_+ + b_2) e^{\lambda+t} c_+^{N-1} \frac{t^{N-2}}{(N-2)!} \right. \\
&\quad \left. + N(a_2 w_- + b_2) e^{\lambda-t} c_-^{N-1} \frac{t^{N-2}}{(N-2)!} \right).
\end{aligned} \tag{113}$$

Substituting the values of λ_{\pm} and c_{\pm}^{N-1} , from Equations (104) and (109) respectively, yields

$$\begin{aligned}
K_{T_1^0}(T_1^0(0), T_2^0(0), t) &= N \frac{t^{N-2}}{(N-2)!} e^{-\alpha t} \\
&\quad \cdot \left(e^{-l+Nt} (a_1 w_+ + b_1) P\mathcal{L} \left((l_+ N)^{N-2} \frac{T_1^{N-1}(0) - w_- T_2^{N-1}(0)}{w_+ - w_-} \right) \right. \\
&\quad \left. + e^{-l-Nt} (a_1 w_- + b_1) P\mathcal{L} \left(-(l_- N)^{N-2} \frac{T_1^{N-1}(0) - w_+ T_2^{N-1}(0)}{w_+ - w_-} \right) \right) \\
&= N^2 \frac{t^{N-2}}{(N-2)!} e^{-\alpha t} \left((l_+ N)^{N-2} e^{-l_+ Nt} (A_{1+} T_1^0(0) + B_{1+} T_2^0(0)) \right. \\
&\quad \left. + (l_- N)^{N-2} e^{-l_- Nt} (A_{1-} T_1^0(0) + B_{1-} T_2^0(0)) \right),
\end{aligned} \tag{114}$$

and similarly

$$K_{T_2^0}((T_1^0(0), T_2^0(0), t) = N^2 \frac{t^{N-2}}{(N-2)!} e^{-\alpha t} \left((l_+ N)^{N-2} e^{-l_+ N t} (A_{2+} T_1^0(0) + B_{2+} T_2^0(0)) \right. \\ \left. + (l_- N)^{N-2} e^{-l_- N t} (A_{2-} T_1^0(0) + B_{2-} T_2^0(0)) \right), \quad (115)$$

with

$$A_{1+} = \frac{(a_1 w_+ + b_1)(-a_1 + w_- a_2)}{w_+ - w_-}, \quad A_{1-} = \frac{-(a_1 w_- + b_1)(-a_1 + w_+ a_2)}{w_+ - w_-}, \\ B_{1+} = \frac{(a_1 w_+ + b_1)(-b_1 + w_- b_2)}{w_+ - w_-}, \quad B_{1-} = \frac{-(a_1 w_- + b_1)(-b_1 + w_+ b_2)}{w_+ - w_-}, \\ A_{2+} = \frac{(a_2 w_+ + b_2)(-a_1 + w_- a_2)}{w_+ - w_-}, \quad A_{2-} = \frac{-(a_2 w_- + b_2)(-a_1 + w_+ a_2)}{w_+ - w_-}, \\ B_{2+} = \frac{(a_2 w_+ + b_2)(-b_1 + w_- b_2)}{w_+ - w_-}, \quad B_{2-} = \frac{-(a_2 w_- + b_2)(-b_1 + w_+ b_2)}{w_+ - w_-}. \quad (116)$$

All components of Equation (11), the system resulting from the application of the Mori-Zwanzig formalism, are now known. The resulting equations for T_1^0 and T_2^0 , found by substituting the noise and memory terms into the Langevin equation, are

$$\partial_t T_1^0 = -a_1 N T_1^0 - b_1 N T_2^0 - \alpha T_1^0 \\ + N e^{-\alpha t} \sum_{i=1}^{N-1} \left((a_1 w_+ + b_1) e^{-l_+ N t} c_+^i + (a_1 w_- + b_1) e^{-l_- N t} c_-^i \right) \frac{t^{i-1}}{(i-1)!} \\ + \int_0^t N^2 \frac{(t-s)^{N-2}}{(N-2)!} e^{-\alpha(t-s)} \left((l_+ N)^{N-2} e^{-l_+ N(t-s)} (A_{1+} T_1^0(s) + B_{1+} T_2^0(s)) \right. \\ \left. + (l_- N)^{N-2} e^{-l_- N(t-s)} (A_{1-} T_1^0(s) + B_{1-} T_2^0(s)) \right) ds, \\ \partial_t T_2^0 = -a_2 N T_1^0 - b_2 N T_2^0 - \alpha T_2^0 \\ + N e^{-\alpha t} \sum_{i=1}^{N-1} \left((a_2 w_+ + b_2) e^{-l_+ N t} c_+^i + (a_2 w_- + b_2) e^{-l_- N t} c_-^i \right) \frac{t^{i-1}}{(i-1)!} \\ + \int_0^t N^2 \frac{(t-s)^{N-2}}{(N-2)!} e^{-\alpha(t-s)} \left((l_+ N)^{N-2} e^{-l_+ N(t-s)} (A_{2+} T_1^0(s) + B_{2+} T_2^0(s)) \right. \\ \left. + (l_- N)^{N-2} e^{-l_- N(t-s)} (A_{2-} T_1^0(s) + B_{2-} T_2^0(s)) \right) ds. \quad (117)$$

This system still depends on the discretization, or more precisely, on the number of points N . The discretization affects both the noise and memory terms. The effect of this dependence on the noise terms is considered in the next section. The effect on the memory terms is briefly discussed here. To see what this effect of N is, consider the function

$$f_K(t) = N^2 \frac{t^{N-2}}{(N-2)!} e^{-\alpha t} (\mu N)^{N-2} e^{-\mu N t}, \quad (118)$$

which contains the dependence of the memory kernel on t and N for $\mu = l_{\pm}$. This function provides information about the behaviour of the memory kernel for increasing N . In Figure 19 this function is plotted for several N and fixed μ . For increasing N the function approaches a peak of increasing height at $1/\mu$. This blow up at one point is due to the projection onto one location. This prevents the waves from traveling through the basin in the orthogonal dynamics system, resulting in an accumulation of energy at the location of the resolved variables. This blow up of the memory kernel has to be dealt with. Here it is important to note that the memory kernel occurs in an integral, meaning the surface below f_K is the determining factor, not the peak height. In the next section the limit behaviour of the memory integral is investigated by applying Laplace's approximation.

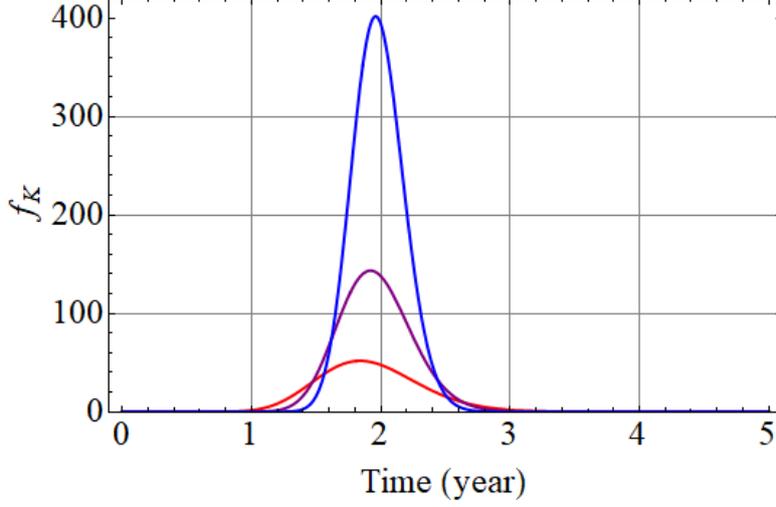


Figure 19: The function $f_K(t)$ with $\mu = 0.5$ for $N = 25$ (red), $N = 50$ (purple) and $N = 100$ (blue).

4.2.4 Limit Behaviour

The components within the memory integrals satisfy the conditions for Laplace's approximation [15]. This is used to expand the memory integrals in $\epsilon = \frac{1}{N}$. Laplace's approximation says that for a large number N , a smooth function $h(x)$, and a twice differentiable function $g(x)$, the following integral can be approximated by

$$\int_a^b h(x)e^{Ng(x)} dx = \sqrt{\frac{2\pi}{N|g''(x_0)|}} e^{Ng(x_0)} \cdot \left(h(x_0) + \frac{1}{N} \left(-\frac{h''(x_0)}{2g''(x_0)} + \frac{h(x_0)g''''(x_0)}{8(g''(x_0))^2} \right. \right. \\ \left. \left. + \frac{h'(x_0)g'''(x_0)}{2(g''(x_0))^2} - \frac{5h(x_0)(g'''(x_0))^2}{24(g''(x_0))^3} \right) + \mathcal{O}\left(\frac{1}{N^2}\right) \right), \quad (119)$$

provided that there is an $x_0 \in (a, b)$ such that $g(x)$ is only close to $g(x_0)$ if x is close to x_0 . At that point $g(x)$ is required to have a maximum, so $g''(x_0) < 0$.

For the AMO model considered, all components in the memory integrals can be written in the form $h_{ij\pm}(s)e^{Ng_{\pm}(s)}$, for $i = 1, 2$ and $j = 1, 2$, with

$$g_{\pm}(s) = \log(l_{\pm}(t-s)) - l_{\pm}(t-s), \\ h_{ij\pm}(s) = \frac{N^N}{(N-2)!} (l_{\pm}(t-s))^{-2} e^{-\alpha(t-s)} C_{j\pm} T_i(s), \quad (120)$$

where $C_{j\pm}$ is either $A_{j\pm}$ or $B_{j\pm}$ as defined in Equation (116). To check that the conditions for the Laplace approximation are satisfied, the first two derivatives of $g_{\pm}(s)$ are computed:

$$g'_{\pm}(s) = -\frac{1}{t-s} + l_{\pm}, \\ g''_{\pm}(s) = -\frac{1}{(t-s)^2}. \quad (121)$$

The first derivative indicates there is one extreme at $s_0 = t - \frac{1}{l_{\pm}}$. This means the first condition is satisfied. Furthermore, $g_{\pm}(s_0) = -l_{\pm}^2 < 0$, indicating the second condition is satisfied. This means that Laplace's approximation can be applied. To be able to compute the first error term of Laplace's approximation, also the third and fourth derivative of $g_{\pm}(s)$ are computed:

$$g'''_{\pm}(s) = -\frac{2}{(t-s)^3}, \\ g''''_{\pm}(s) = -\frac{6}{(t-s)^4}, \quad (122)$$

as well as the first two derivatives of $h_{ij\pm}$:

$$\begin{aligned} h'_{ij\pm}(s) &= \frac{N^N}{(N-2)!} (l_{\pm}(t-s))^{-2} e^{-\alpha(t-s)} C_{j\pm} \left(\left(\alpha + \frac{2}{t-s} \right) T_i(s) + T'_i(s) \right), \\ h''_{ij\pm}(s) &= \frac{N^N}{(N-2)!} (l_{\pm}(t-s))^{-2} e^{-\alpha(t-s)} C_{j\pm} \left(\left(\left(\alpha + \frac{2}{t-s} \right)^2 + \frac{2}{(t-s)^2} \right) T_i(s) \right. \\ &\quad \left. + 2 \left(\alpha + \frac{2}{t-s} \right) T'_i(s) + T''_i(s) \right). \end{aligned} \quad (123)$$

Substituting all the above computed terms into the Laplace approximation of Equation (119) gives

$$\begin{aligned} \int_0^t h_{ij\pm}(s) e^{Ng_{\pm}(s)} ds &= \sqrt{\frac{2\pi}{Nl_{\pm}^2}} e^{-N} \frac{N^N}{(N-2)!} e^{-\frac{\alpha}{l_{\pm}}} C_{j\pm} \left(T_i \left(t - \frac{1}{l_{\pm}} \right) \right. \\ &\quad \left. + \frac{1}{N} \frac{1}{2l_{\pm}^2} \left((2l_{\pm}(l_{\pm} + \alpha) + \alpha^2) T_i \left(t - \frac{1}{l_{\pm}} \right) + 2(l_{\pm} + \alpha) T'_i \left(t - \frac{1}{l_{\pm}} \right) \right. \right. \\ &\quad \left. \left. + T''_i \left(t - \frac{1}{l_{\pm}} \right) \right) + \mathcal{O} \left(\frac{1}{N^2} \right) \right). \end{aligned} \quad (124)$$

This gives an approximation for every part of the memory integrals of the AMO model. In the above equation delay terms have emerged, meaning that also the AMO can be modeled by some type of delay equation.

In Equation (124) there still is N -dependence present. To investigate the behaviour of this dependence one factor of N is taken out, after which a Taylor expansion in $1/N$ around zero is applied. The result is

$$\sqrt{\frac{2\pi}{Nl_{\pm}^2}} e^{-N} \frac{N^N}{(N-2)!} = \frac{N}{l_{\pm}} \left(1 - \frac{13}{12} \frac{1}{N} + \mathcal{O} \left(\frac{1}{N} \right) \right). \quad (125)$$

This means all terms in the memory integrals can be expanded in terms of $\epsilon = \frac{1}{N}$. That is

$$\begin{aligned} \int_0^t h_{ij\pm}(s) e^{Ng_{\pm}(s)} ds &= \frac{N}{l_{\pm}} e^{-\frac{\alpha}{l_{\pm}}} C_{j\pm} \left(T_i \left(t - \frac{1}{l_{\pm}} \right) + \epsilon \frac{1}{2l_{\pm}^2} \left(((l_{\pm} + \alpha)^2 - \frac{7}{6} l_{\pm}^2) T_i \left(t - \frac{1}{l_{\pm}} \right) \right. \right. \\ &\quad \left. \left. + 2(l_{\pm} + \alpha) T'_i \left(t - \frac{1}{l_{\pm}} \right) + T''_i \left(t - \frac{1}{l_{\pm}} \right) \right) + \mathcal{O}(\epsilon^2) \right). \end{aligned} \quad (126)$$

So far, only the memory terms of the Mori-Zwanzig formalism have been considered. The behaviour of the noise terms for increasing N also needs to be investigated. Each component of the sum in the noise terms in Equation (111) is proportional to

$$N e^{-l_{\pm} N t} \frac{(l_{\pm} N t)^{N-k}}{(N-k)!}, \quad (127)$$

for some $k = 2, \dots, N$. Dividing by N the resulting function peaks at time $t = \frac{1}{l_{\pm}}$, just as the memory kernels do. If N is increased, the contribution of all other times goes to zero faster than ϵ^2 . At the time where the noise term peaks, it can be expanded in ϵ , just as has been done for the memory terms. The result is

$$N e^{-N} \frac{(N)^{N-k}}{(N-k)!} = \frac{1}{\sqrt{2\pi}} \frac{N}{\sqrt{\epsilon}} \left(1 - \frac{13}{12} \epsilon + \mathcal{O}(\epsilon^2) \right). \quad (128)$$

This shows that the noise terms indeed decay faster than the memory terms at dominant order. The decay is proportional to $\sqrt{\epsilon}$, which is slower than the decay of the first perturbation to the memory terms. However, this is an effect that happens at one time only, as a remnant of the initial conditions. At later times it does not have any effect, and therefore it is neglected when the resulting equations are considered in the next section. But when considering the possible predictability of noise, this can be an interesting term to consider in more detail.

4.3 Delay Model

The limiting behaviour of the noise and memory terms has been discussed in the previous section. Dividing by N in Equation (117), the resulting system can be written as an expansion in ϵ . Here, as noted in the previous section, the noise term is neglected. Up to first order the resulting equations are

$$\begin{aligned}\epsilon \frac{dT_1}{dt} &= -a_1 T_1(t) - b_1 T_2(t) + \frac{A_{1+}}{l_+} e^{-\frac{\alpha}{l_+}} T_1\left(t - \frac{1}{l_+}\right) + \frac{B_{1+}}{l_+} e^{-\frac{\alpha}{l_+}} T_2\left(t - \frac{1}{l_+}\right) \\ &\quad + \frac{A_{1-}}{l_-} e^{-\frac{\alpha}{l_-}} T_1\left(t - \frac{1}{l_-}\right) + \frac{B_{1-}}{l_-} e^{-\frac{\alpha}{l_-}} T_2\left(t - \frac{1}{l_-}\right) + \epsilon f_{\epsilon 1}(t) + \mathcal{O}(\epsilon^2), \\ \epsilon \frac{dT_2}{dt} &= -a_2 T_1(t) - b_2 T_2(t) + \frac{A_{2+}}{l_+} e^{-\frac{\alpha}{l_+}} T_1\left(t - \frac{1}{l_+}\right) + \frac{B_{2+}}{l_+} e^{-\frac{\alpha}{l_+}} T_2\left(t - \frac{1}{l_+}\right) \\ &\quad + \frac{A_{2-}}{l_-} e^{-\frac{\alpha}{l_-}} T_1\left(t - \frac{1}{l_-}\right) + \frac{B_{2-}}{l_-} e^{-\frac{\alpha}{l_-}} T_2\left(t - \frac{1}{l_-}\right) + \epsilon f_{\epsilon 2}(t) + \mathcal{O}(\epsilon^2),\end{aligned}\tag{129}$$

where

$$\begin{aligned}f_{\epsilon 1}(t) &= -\alpha T_1(t) + \frac{A_{1+}}{l_+} e^{-\frac{\alpha}{l_+}} g_{\epsilon+}(T_1) + \frac{B_{1+}}{l_+} e^{-\frac{\alpha}{l_+}} g_{\epsilon+}(T_2) \\ &\quad + \frac{A_{1-}}{l_-} e^{-\frac{\alpha}{l_-}} g_{\epsilon-}(T_1) + \frac{B_{1-}}{l_-} e^{-\frac{\alpha}{l_-}} g_{\epsilon-}(T_2), \\ f_{\epsilon 2}(t) &= -\alpha T_2(t) + \frac{A_{2+}}{l_+} e^{-\frac{\alpha}{l_+}} g_{\epsilon+}(T_1) + \frac{B_{2+}}{l_+} e^{-\frac{\alpha}{l_+}} g_{\epsilon+}(T_2) \\ &\quad + \frac{A_{2-}}{l_-} e^{-\frac{\alpha}{l_-}} g_{\epsilon-}(T_1) + \frac{B_{2-}}{l_-} e^{-\frac{\alpha}{l_-}} g_{\epsilon-}(T_2),\end{aligned}\tag{130}$$

for

$$g_{\epsilon\pm}(T) = \frac{1}{2l_{\pm}^2} \left((l_{\pm} + \alpha)^2 - \frac{7}{6} l_{\pm}^2 \right) T\left(t - \frac{1}{l_{\pm}}\right) + 2(l_{\pm} + \alpha) T'\left(t - \frac{1}{l_{\pm}}\right) + T''\left(t - \frac{1}{l_{\pm}}\right).\tag{131}$$

This equation is the main result achieved by applying the Mori-Zwanzig formalism to the discretized system of the AMO in Equation (92).

Letting $\epsilon \rightarrow 0$, the equations simplify to a set of delay difference equations. These delay difference equations can also be derived by integration along characteristics of the original system [17]. Such a delay difference equation shows the transport of temperature through the basin with (possibly) different velocities. For more theory on delay difference and delay differential equations one can consult the book by Hale and Verduyn Lunel [21]. Delay difference equations can exhibit an increasing frequency of switching between different states if the delay is irrational [18]. To prevent this possibility of infinite switching, in numerical studies an $\epsilon \frac{d}{dt}$ -term is added. This term smooths the solution by preventing the infinite switching. In applying the Mori-Zwanzig formalism an error term for the addition of this $\epsilon \frac{d}{dt}$ -term, has been computed. This error term corresponds to the error made in the discretization of the x -derivative, indicating it reflects (numerical) diffusion.

The method employed here to derive a delay model for the AMO, including error terms, can be generalized to other systems. Also for the ENSO model this method results in a delay equation. The derivation of this is given in Appendix A. Note that both the AMO and ENSO model are diagonalizable in their ∂_x -dependence. It is expected that other diagonalizable systems of wave equations also have a corresponding delay (difference) equation. In the ENSO model the result of a delay differential equation is (partly) due to absence of a ∂_x -terms in the temperature equation. To see whether this can be generalized more research is needed.

In Figure 20 a simulation of the delay model of the AMO is shown for $\epsilon = 10^{-2}$, together with the resulting error term. The history taken here is from a simulation of the initial AMO model. For this ϵ the model is stable and shows the transport of the history with two different velocities. The error is present dominantly for the high frequency oscillations and, as expected, an order ϵ smaller than the original oscillation. The fast velocity corresponds to a delay of about 3 years, while the second delay is approximately 27 years. These delay times correspond to the

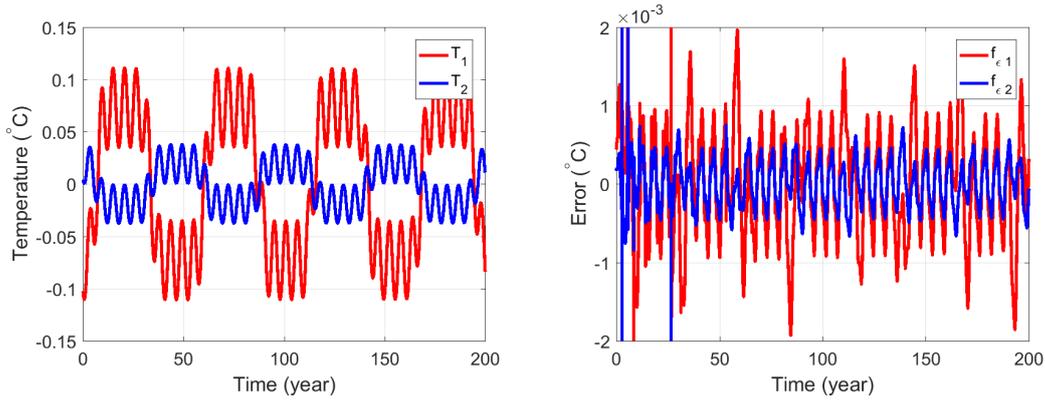


Figure 20: A model simulation of the AMO delay model in Equation (129) with $\epsilon = 10^{-2}$ where a simulation of the original model is used as history. In the right figure the corresponding first order error term is shown.

time the waves need to cross the basin. The long delay time corresponds to the AMO and the short delay time represents another thermal Rossby wave present in the model. By the inversion in the boundary conditions, the period doubles compared to the delay.

The behaviour of this delay model for the AMO is similar to that of the full model by Sévellec and Huck. The periods of oscillation are the same, meaning both the five and sixty year period are recovered. An interesting aspect of this delay model is that it can be initiated with real data to predict the evolution of the AMO, including possibly the behaviour of noise. The dominance of the high frequency in the error terms, indicates that this component might be less stable, as already was shown in Section 4.1.2 by adding a background overturning circulation. In summary, this delay (difference) model with additional $\epsilon \frac{d}{dt}$ -term is just as good a model for the AMO as the full model by Sévellec and Huck when initialized with a realistic history.

5 Summary, Discussion and Conclusion

Delay models are useful as conceptual climate models by their infinite dimensional nature and limited number of parameters. They are suited for mathematical analysis and this way can add to the physical understanding of the processes involved. In this thesis the Mori-Zwanzig formalism has been investigated as a method to derive delay equations. The two-strip model of the El Niño Southern Oscillation (ENSO) was used as a test case for the application of the technique. The reason for this was that delay models had already been proposed for this phenomenon. The formalism then was applied to a model of the Atlantic Multidecadal Oscillation (AMO). No delay model for this phenomenon has been proposed up to now.

The Mori-Zwanzig formalism gives an exact rewriting of a system of ordinary differential equations [5]. The rewritten equation contains a Markovian, noise, and memory term. Here the focus was on the memory term, since this is an integral over the history of the system, just as the delay term represents this history. By making approximations of the memory integral, delay terms can be derived. This was achieved for models of both the ENSO and AMO.

For the model of ENSO, in addition to the use of the Mori-Zwanzig formalism, also a generalized version of the method of variation of constants has been employed. Starting from the linear two-strip model [24], the linear part of the delay model by Suarez and Schopf has been derived [36]. The two methods used here yield the same result. Furthermore, a nonlinear version of the two-strip model was derived by assuming that the sea surface temperature is proportional to the subsurface temperature. This assumption is supported by (NOAA) buoy data from the Pacific Ocean. Also for the nonlinear two-strip model (nonlinear) delay models have been derived. Here the two methods do not yield the same result. This is due to the approximations needed to obtain a closed-form equation from the Mori-Zwanzig formalism. The method of variation of constants is exact.

Both derived nonlinear delay models contain an extra cubic delay term compared to the model by Suarez and Schopf [36]. In both cases this additional term results in an increased period of the model oscillation. Another consequence of the additional term is the decrease of the area in parameter space where stable oscillations occur. For the model derived using the approximated Mori-Zwanzig formalism, this decrease is so large that no stable periodic behaviour occurs for realistic values of the parameters. The exact derived model does show oscillations for these parameter values. The period of the model derived using variation of constants is closer to the real period of ENSO than the model proposed by Suarez and Schopf. However, its period is still smaller than what is seen in data. An option to improve the match between model period and data, is to no longer assume a delta-function for the spatial pattern of the wind forcing, but rather take a more realistic pattern. As a consequence, the resulting delay model will no longer contain a discrete delay, but a distributed delay. Another option for improvement is to add additional nonlinearities, for example in the thermocline equations.

Also for the AMO a delay model has been derived from a three-layer model by Sévellec and Huck [34]. In contrast to the delay differential model for ENSO, the derived model for the AMO is a delay difference model at first order. This means the current state is fully determined by past states. This type of model can exhibit an increasing switching frequency between states [18], making it physically unrealistic. Often an $\epsilon \frac{d}{dt}$ -term is added to prevent this behaviour and allow for better numerical treatment. Starting from the discretized AMO model, an error term for this approximation was derived. This error term corresponds to the upwind discretization scheme used.

The AMO model by Sévellec and Huck does not contain a background overturning circulation. This results in a high frequency model oscillation which is not clearly seen in data. As discussed in Section 4.1.2, adding the meridional overturning circulation to the background state of the model, results in a damping of this high frequency oscillation. It would be interesting to see what the effect of this background overturning is on the resulting delay equations. The addition of the overturning results in a change of the eigenvalues of the discretized system. These eigenvalues then show a stronger decay for the high frequency mode and a weaker decay for the low frequency mode. It is expected that this will be expressed in the parameters in front of the different delay terms in the resulting delay model. There is not expected to be an effect on the delay times, but this has to be verified. The application of the Mori-Zwanzig formalism to this extended AMO model remains for future research.

The method of deriving delay equations applied to the AMO model can be generalized. It is expected that every diagonalizable linear system of wave equations can be rewritten in the form of a delay difference equation. Integration along characteristics already yields the dominant terms. In addition, the Mori-Zwanzig formalism gives error terms to a smoothening approximation of the resulting delay difference system. For non-diagonalizable systems a similar statement might be valid, but the computations become more involved. The necessity of the diagonalizability remains for future work.

For the ENSO model it is not necessary to use the Mori-Zwanzig formalism to arrive at a delay equation. It does not give additional understanding compared to the method of variation of constants. This last method works for both the linear and nonlinear model, since the nonlinear model is still linear in the equations for the unresolved variables (thermocline depth). Therefore, the variation of constants method yields a delay equation for the resolved variable (temperature) in both cases. The Mori-Zwanzig formalism gives the same result for the linear system, but the computation is more involved. For nonlinear systems the formalism does not give accurate results, since approximations are needed.

For the derivation of the delay model of the AMO, the Mori-Zwanzig formalism is not strictly necessary, just as is the case for the ENSO model. Since the model is linear, it can be diagonalized in the ∂_x -terms. When linear terms, so terms that do not contain an x -derivative, are not present, there is a change of variables to a system in which the equations for all different variables are separated. Integration along characteristics will give the delay difference equations as found by the Mori-Zwanzig formalism in this thesis. However, when linear terms are present the Mori-Zwanzig formalism still gives a resulting delay difference equation, in contrast to the alternative method. The Mori-Zwanzig formalism is needed as well to find the error terms of the smoothening approximation for the delay difference model.

When the equations of the model considered are also nonlinear in the unresolved variables, the Mori-Zwanzig formalism is the only method that will give a result. In such a nonlinear case the orthogonal dynamics system has to be approximated. This approximation needs to be an improvement on the pseudo-orthogonal dynamics approximation [20], since this approximation was shown to be not accurate for the ENSO model. Only if its accuracy can be shown for a specific model, it can be justified to apply this approximation method. The derivation of improved approximations is a first step that needs to be taken to apply the Mori-Zwanzig formalism accurately to nonlinear models. This is a necessary step to be able to reliably derive nonlinear delay models for climate systems.

Since many climate models are wave equations in one form or another, it is expected that they can also be represented by a delay system. In the models studied in this thesis, it was shown that projecting a system of wave equations onto one location yields a delay model. This would imply that there is an abundance of processes in climate that can be described by a delay equation. For linear models, this thesis provides the handles to derive such a delay model. Here, the Mori-Zwanzig formalism is not always necessary. For nonlinear models it should be possible to arrive at a delay model as well. To do so, the Mori-Zwanzig formalism is needed, unless the unresolved dynamics still is linear. When the formalism is applied, the orthogonal dynamics equation has to be approximated. To do this accurately, new approximation methods need to be developed. Only then it is possible to derive accurate nonlinear delay models for climate systems.

Acknowledgements

I would like to thank both Henk Dijkstra and Jason Frank for their supervision and the discussions during my research. Furthermore, I am grateful to Jan Sieber for joining in on the project and supervising me during my time in Exeter. The discussions together greatly improved my knowledge of delay equations and improved my mathematical writing. Also a very special thanks to Courtney Quinn for working together in Exeter (and the Netherlands). I really appreciated the discussions about the physics of the AMO and the help with a thorough understanding of the Mori-Zwanzig formalism.

During my time at the University of Exeter I have been warmly welcomed by everyone. Here I want to name, next to Jan and Courtney, Lamees Felemban, Paul Ritchie, Hassan Alkhayoun and Damian (and Bogna) Smug, since they were great office mates and showed me some nice parts around Exeter. Also a special thanks to Nicci Rodie for having me as a lodger, joining me in watching *Wie is de Mol* and introducing me to *The Crown* and English habits.

In the Netherlands I want to thank everyone who worked in the masterroom and with whom I spend many lunch and coffee breaks. Furthermore, I am grateful to Rineke Falkena, Monique Dorresteyn, Elsa van Roessel, Thijs van Gogh and Sibrand Rinzema for reading through my thesis and giving useful comments and constructive criticism.

Lastly I am really lucky to have such supportive parents and a great boyfriend. Especially a really big thank you to my mom for bringing me to Exeter and back by car. This way I could use my bike there and continue my Dutch cycling way of life. I like to thank both my parents for visiting me in Exeter, so that I could show them around and explore even more of the English countryside. Last but not least, I am very grateful for having such a sweet boyfriend, Sibrand Rinzema, who accepted my absence for almost 5 months and came to visit me twice, even though the snow stopped him at first for the second visit. I want to thank him for exploring Exeter (and surroundings) and Dartmoor together, trusting me in driving on the left side of the road with the steering wheel at the wrong side of the car and off course for the runs up and down hill, the skype conversations and advice on no matter what.

References

- [1] D.S. Battisti. “Dynamics and Thermodynamics of a Warming Event in a Coupled Tropical Atmosphere-Ocean Model”. In: *Journal of the Atmospheric Sciences* 45.20 (1988), pp. 2889–2919.
- [2] D.S. Battisti and A.C. Hirst. “Interannual Variability in a Tropical Atmosphere-Ocean Model: Influence of the Basic State, Ocean Geometry and Nonlinearity”. In: *Journal of the Atmospheric Sciences* 46.12 (1988), pp. 1687–1712.
- [3] C.L. Beck et al. “Model reduction, optimal prediction, and the Mori-Zwanzig representation of Markov chains”. In: *Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*. 2009.
- [4] H.W. Broer et al. “The Dynamics of a Low-Order Model for the Atlantic Multidecadal Oscillation”. In: *Discrete and Continuous Dynamical Systems* 16 (2011), pp. 73–107.
- [5] A.J. Chorin, O.H. Hald, and R. Kupferman. “Optimal prediction with memory”. In: *Physica D* 166 (2002).
- [6] C. Colin de Verdière and T. Huck. “Baroclinic Instability: An Oceanic Wavemaker for Interdecadal Variability”. In: *Journal of Physical Oceanography* 29 (1999), pp. 893–910.
- [7] B. Cushman-Roisin and J. Beckers. *Introduction to Geophysical Fluid Dynamics*. Vol. 101. Academic Press, 2011.
- [8] E. Darve, J. Solomon, and A. Kia. “Computing generalized Langevin equations and generalized Fokker-Planck equations”. In: *Proceedings of the National Academy of Sciences* 106.27 (2009), pp. 10884–10889.
- [9] H.A. Dijkstra. “A Normal Mode Perspective of Intrinsic Ocean-Climate Variability”. In: *Annual Review of Fluid Mechanics* 48 (2016), pp. 341–363.
- [10] H.A. Dijkstra. *Dynamical Oceanography*. Springer, 2008.
- [11] H.A. Dijkstra. “Interaction of SST Modes in the North Atlantic Ocean”. In: *Journal of Physical Oceanography* 36 (2006), pp. 286–299.
- [12] H.A. Dijkstra. *Nonlinear Physical Oceanography. A Dynamical Systems Approach to the Large Scale Ocean Circulation and El Niño*. 2nd Revised and Enlarged Edition. Vol. 28. Springer, 2005.
- [13] H.A. Dijkstra and J.D. Neelin. “On the attractors of an intermediate coupled ocean-atmosphere model”. In: *Dynamics of Atmosphere and Ocean*. 22nd ser. 19–48 (1995).
- [14] J.J. Duistermaat and W. Echhaus. *Analyse van Gewone Differentiaalvergelijkingen*. Epsilon Uitgaven, 1995.
- [15] A. Erdélyi. *Asymptotic Expansions*. Dover Publications, 1956.
- [16] D.J. Evans and G.P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. 2nd. ANU E Press, 2007, p. 61.
- [17] L.C. Evans. *Partial Differential Equations*. 2nd. Vol. 19. American Mathematical Society, 2010.
- [18] M. Ghil, I. Zaliapin, and B. Coluzzi. “Boolean delay equations: A simple way of looking at complex systems”. In: *Physica D* 237 (2008), pp. 2967–2986.
- [19] D. Givon, R. Kupferman, and A. Stuart. “Extracting macroscopic dynamics: model problems and algorithms”. In: *Nonlinearity* 17 (2004), pp. 55–127.
- [20] A. Gouasmi, E.J. Parish, and K. Duraisamy. “A priori estimation of memory effects in reduced-order models of nonlinear systems using the Mori-Zwanzig formalism”. In: *Proc. R. Soc. A*. 473rd ser. 20170385 (2017).
- [21] J.K. Hale and Verduyn Lunel S.M. *Introduction to Functional Differential Equations*. Springer, 1993.
- [22] Z. Hao, J.D. Neelin, and F. Jin. “Nonlinear Air-Sea Interaction in the Fast-Wave Limit”. In: *Journal of Climate* 6 (1993), pp. 1523–1544.

- [23] A.R. Humphries et al. “Dynamics of a delay differential equation with multiple state-dependent delays”. In: *Discrete and Continuous Dynamical Systems* 32.8 (2012), pp. 2701–2727.
- [24] F. Jin. “An Equatorial Ocean Recharge Paradigm for ENSO. Part I: Conceptual Model”. In: *Journal of Atmospheric Sciences* 54 (1997), pp. 811–829.
- [25] F. Jin. “An Equatorial Ocean Recharge Paradigm for ENSO. Part II: A Stripped-Down Coupled Model”. In: *Journal of Atmospheric Sciences* 54 (1997), pp. 830–847.
- [26] E. Kalnay and Coauthors. “The NCEP/NCAR Reanalysis 40-year Project”. In: *Bulletin of the American Meteorological Society* 77 (1996), pp. 437–471.
- [27] A. Keane, B. Krauskopf, and C.M. Postlethwaite. “Climate models with delay differential equations”. In: *Chaos* 27.114309 (2017).
- [28] B. Krauskopf and J. Sieber. “Bifurcation analysis of delay-induced resonances of the El Niño Southern Oscillation”. In: *Proceedings of the Royal Society A* 470.2169 (2014).
- [29] Pacific Marine Environmental Laboratory. *Explaining El Niño*. NOAA. May 23, 2018. URL: <https://www.pmel.noaa.gov/elnino/schematic-diagrams>.
- [30] H. Mori. “Transport, Collective Motion and Brownian Motion”. In: *Progress of Theoretical Physics* 33.3 (1965), pp. 423–455.
- [31] G.A. Pavliotis. *Stochastic Processes and Applications. Diffusion Processes, the Fokker-Planck and Langevin Equations*. 2014.
- [32] D. Roose and R. Szalai. “Continuation and Bifurcation Analysis of Delay Differential Equations”. In: *Numerical Continuation Methods for Dynamical Systems*. Ed. by B. Krauskopf, Osinga H.M., and J. Galán-Vioque. Springer, Dordrecht, 2007.
- [33] J. Runge, V. Petoukhov, and J. Kurths. “Quantifying the Strength and Delay of Climate Interactions: The Ambiguities of Cross Correlation and Novel Measure Based on Graphical Models”. In: *Journal of Climate* 27 (2014), pp. 720–739.
- [34] F. Sévellec and T. Huck. “Theoretical Investigation of the Atlantic Multidecadal Oscillation”. In: *Journal of Physical Oceanography* 45 (2015).
- [35] J. Shen, T. Tang, and L. Wang. *Spectral Methods. Algorithms, Analysis and Applications*. Springer, 2011.
- [36] M.J. Suarez and P.S. Schopf. “A Delayed Action Oscillator for ENSO”. In: *Journal of the Atmospheric Sciences* 45.21 (1988), pp. 3283–3287.
- [37] R. Szalai. “Modelling elastic structures with strong nonlinearities with application to stick-slip friction”. In: *Proceedings of the Royal Society A* 470, 20130593 (2014).
- [38] J.R. Taylor. *Classical Mechanics*. University Science Books, 2005.
- [39] L.A. Te Raa and H.A. Dijkstra. “Instability of the Thermohaline Ocean Circulation on Interdecadal Timescales”. In: *Journal of Physical Oceanography* 32 (2002), pp. 138–160.
- [40] J. Wouters et al. “Parametrization of stochastic multiscale triads”. In: *Nonlinear Processes in Geophysics* 23 (2016), pp. 435–445.
- [41] S.E. Zebiak and M.A. Cane. “A Model of El Niño-Southern Oscillation”. In: *Monthly Weather Review* 115 (1987), pp. 2262–2278.
- [42] R. Zwanzig. “Nonlinear Generalized Langevin Equations”. In: *Journal of Statistical Physics* 9.3 (1973), pp. 215–220.

A ENSO: Discrete Mori-Zwanzig Formalism

The discretization method applied to the AMO model of Equation (92) in Section 4.2, can also be applied to the ENSO model in Equation (30). This gives the exact same result as found in Section 3.2.2 for a correct parametrization of the delta-function. Just as for the AMO model, the basin is divided into $N + 1$ points and discretized according to an upwind scheme following the waves. This discretization is shown in Figure 21. The resulting discretized model is

$$\begin{aligned} \partial_t h_c^n &= -\epsilon_0 h_c^n - \frac{1}{dx}(h_c^n - h_c^{n-1}) + \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) N \delta_{n_w}(n) T_e^N, \quad n = 1 \dots N, \\ \partial_t h_n^n &= -\epsilon_0 h_n^n + \frac{1}{y_n^2} \frac{1}{dx}(h_n^{n+1} - h_n^n) - \mu A_0 \frac{\theta}{y_n^2} N \delta_{n_w}(n) T_e^N, \quad n = 0 \dots N - 1, \\ \partial_t T_e^n &= -c_T T_e^n + c_h \left(h_c^n + \frac{1}{1 + y_n^2} h_n^n\right), \quad n = 0 \dots N, \end{aligned} \quad (132)$$

with boundary conditions

$$\begin{aligned} h_c^0 &= \frac{A_{rW}}{1 + y_n^2} h_n^0, \\ h_n^N &= 0. \end{aligned} \quad (133)$$

Here already the assumptions of a localized wind effect and no reflection at the eastern boundary from Section 3.2.4 have been applied. The temperature equation only couples back into the equations for the thermocline depth for $n = N$, which therefore is the only n for which the temperature equation is considered. Note that in the discretization of the continuous delta-function a factor N is incorporated to have the same energy in the system for every N . This system can be written in matrix form, just as the discretized AMO model.

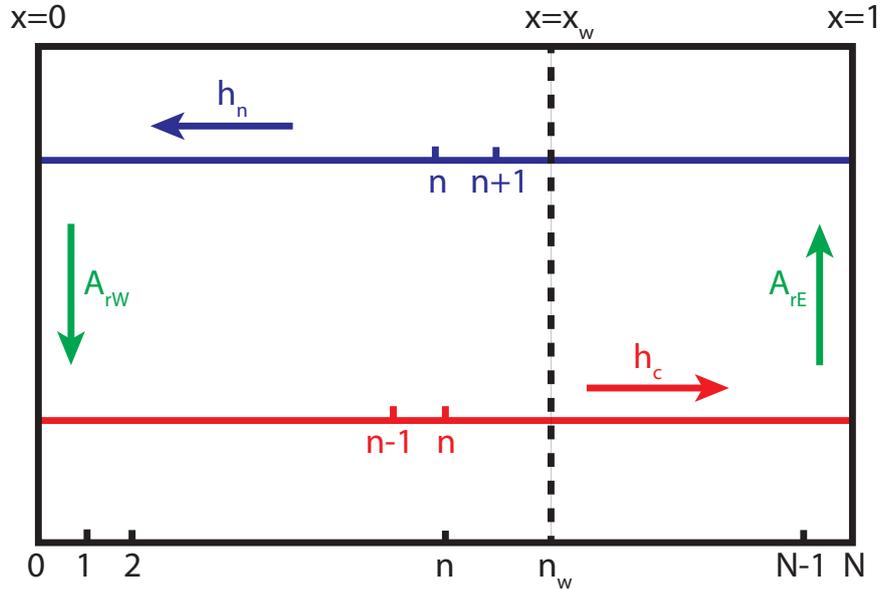


Figure 21: The basin and discretization of the ENSO model. The equator is shown in red, the higher latitude in blue and the boundary conditions in green.

The projection in the AMO model was on one location in x . In that case the memory kernel blew up because the wave was prevented from traveling through the basin. This resulted in an accumulation of wave energy at one location in the orthogonal dynamics model. To achieve a similar result for the ENSO model, it is not sufficient to only project onto T_e^N , as then the waves can still propagate fully in the orthogonal dynamics system. A projection onto both T_e^N and h_n^0

conditions is checked. This gives

$$\begin{aligned}
& P\mathcal{L}(\dots, h_c^i(0), \dots, h_n^i(0), \dots, T_e^N) \\
&= P(\dots, -\epsilon_0 h_c^i(0) - \frac{1}{dx}(h_c^i(0) - h_c^{i-1}(0)) + \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) N \delta_{n_w}(i) T_e^N(0), \\
&\quad \dots, -\epsilon_0 h_n^i(0) + \frac{1}{y_n^2} \frac{1}{dx}(h_n^{i+1}(0) - h_n^i(0)) - \mu A_0 \frac{\theta}{y_n^2} N \delta_{n_w}(i) T_e^N(0), \dots, \\
&\quad - c_T T_e^N(0) + c_h h_c^N(0)) \\
&= \left(\frac{1}{dx} \frac{A_r W}{1 + y_n^2} h_n^0(0), \dots, \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) N \delta_{n_w}(n_w) T_e^N(0), \dots, \right. \\
&\quad \left. \dots, -\mu A_0 \frac{\theta}{y_n^2} N \delta_{n_w}(n_w) T_e^N(0), \dots, -c_T T_e^N(0)\right).
\end{aligned} \tag{140}$$

So only the terms that before application of $P\mathcal{L}$ depend on $h_c^1(0)$, $h_c^{n_w}(0)$, $h_n^{n_w}(0)$ and $T_e^N(0)$ are left in the memory kernels. The corresponding constants are c_c^N , $c_c^{N+1-n_w}$ and $c_n^{n_w}$. This means

$$\begin{aligned}
K_{h_n^0}(t) &= -\mu A_0 \frac{\theta}{y_n^2} e^{-\epsilon_0 t} e^{-\frac{N}{y_n^2} t} N \left(\frac{N}{y_n^2}\right)^{n_w} \frac{t^{n_w-1}}{(n_w-1)!} T_e^N(0). \\
K_{T_e^N}(t) &= c_h e^{-\epsilon_0 t} \left(\frac{A_r W}{1 + y_n^2} e^{-Nt} N^N \frac{t^{N-1}}{(N-1)!} h_n^0(0) \right. \\
&\quad \left. + \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) e^{-Nt} N^{N+1-n_w} \frac{t^{N-N_w}}{(N-n_w)!} T_e^N(0)\right).
\end{aligned} \tag{141}$$

These memory kernels show peaks at certain times, just as in the AMO model. However, here only the kernel for h_n^0 blows up, while the one for T_e^N remains bounded. This is because the original temperature equation does not contain any x -derivatives.

To compute the contribution of the memory term, Laplace's approximation of Equation (119) is used. Here, there are three different components to which the approximation has to be applied. With

$$g_1(s) = \frac{n_w}{N} \log\left(\frac{t-s}{y_n^2}\right) - \frac{t-s}{y_n^2}, \quad h_1(s) = \frac{N^{n_w+1}}{(n_w-1)!} \frac{e^{-\epsilon_0(t-s)}}{t-s} T_e^N(s), \tag{142}$$

$$g_2(s) = \log(t-s) - (t-s), \quad h_2(s) = \frac{N^N}{(N-1)!} \frac{e^{-\epsilon_0(t-s)}}{t-s} h_n^0(s), \tag{143}$$

$$g_3(s) = \frac{N-n_w}{N} \log(t-s) - (t-s), \quad h_3(s) = \frac{N^{N+1-n_w}}{(N-n_w)!} e^{-\epsilon_0(t-s)} T_e^N(s), \tag{144}$$

the contributions to the memory terms are

$$-\mu A_0 \frac{\theta}{y_n^2} \int_0^t h_1(s) e^{N g_1(s)} ds, \tag{145}$$

and

$$c_h \frac{A_r W}{1 + y_n^2} \int_0^t h_2(s) e^{N g_2(s)} ds + c_h \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2}\right) \int_0^t h_3(s) e^{N g_3(s)} ds, \tag{146}$$

in the equations for h_n^0 and T_e^N respectively. Applying Laplace's approximation to the above memory terms yields

$$\begin{aligned}
& -\mu A_0 \frac{\theta}{y_n^2} \frac{n_w^{n_w+1}}{(n_w-1)!} \sqrt{\frac{2\pi}{n_w}} y_n^2 e^{-n_w} \cdot \left(e^{-\epsilon_0 \frac{n_w}{N} y_n^2} \frac{N}{n_w} \frac{1}{y_n^2} T_e^N\left(t - \frac{n_w}{N} y_n^2\right) \right. \\
& \quad + \frac{1}{N} \frac{1}{2} e^{-\epsilon_0 \frac{n_w}{N} y_n^2} \left(\left(\frac{1}{6 y_n^2} \left(\frac{N}{n_w}\right)^2 + \epsilon_0^2 y_n^2\right) T_e^N\left(t - \frac{n_w}{N} y_n^2\right) + 2\epsilon_0 y_n^2 T_e^{N(1)}\left(t - \frac{n_w}{N} y_n^2\right) \right. \\
& \quad \left. \left. + y_n^2 T_e^{N(2)}\left(t - \frac{n_w}{N} y_n^2\right) + \mathcal{O}\left(\frac{1}{N^2}\right)\right)\right),
\end{aligned} \tag{147}$$

for the h_n^0 memory term and

$$\begin{aligned}
& c_h \frac{A_{rW}}{1+y_n^2} \frac{N^N}{(N-1)!} \sqrt{\frac{2\pi}{N}} e^{-N} \cdot \left(e^{-\epsilon_0} h_n^0(t-1) \right. \\
& \quad \left. + \frac{1}{N} \frac{1}{2} e^{-\epsilon_0} \left(\left(\frac{1}{6} + \epsilon_0^2 \right) h_n^0(t-1) + 2\epsilon_0 h_n^{0(1)}(t-1) + h_n^{0(2)}(t-1) \right) + \mathcal{O}\left(\frac{1}{N^2}\right) \right) \\
& + c_h \mu A_0 \left(1 - \frac{\theta}{1+y_n^2} \right) \frac{(N-n_w)^{N+1-n_w}}{(N-n_w)!} \sqrt{\frac{2\pi}{N-n_w}} e^{-(N-n_w)} \\
& \quad \cdot \left(e^{-\epsilon_0(1-\frac{n_w}{N})} T_e^N \left(t - \left(1 - \frac{n_w}{N} \right) \right) + \frac{1}{N} \frac{1}{2} e^{-\epsilon_0(1-\frac{n_w}{N})} \right. \\
& \quad \cdot \left(\left(\frac{1}{6} \frac{N}{N-n_w} - 2\epsilon_0 + \epsilon_0^2 \frac{N-n_w}{N} \right) T_e^N \left(t - \left(1 - \frac{n_w}{N} \right) \right) \right. \\
& \quad \left. \left. - \left(2 - \epsilon_0 \frac{N-n_w}{N} \right) T_e^{N(1)} \left(t - \left(1 - \frac{n_w}{N} \right) \right) + \frac{N-n_w}{N} T_e^{N(2)} \left(t - \left(1 - \frac{n_w}{N} \right) \right) \right) \right) \\
& \quad \left. + \mathcal{O}\left(\frac{1}{N^2}\right) \right), \tag{148}
\end{aligned}$$

for the memory term of T_e^N . In the limit $N \rightarrow \infty$ it holds that $\frac{n_w}{N} \rightarrow x_w$. Writing $\epsilon = \frac{1}{N}$, just as in the AMO model, results in

$$\begin{aligned}
\partial_t h_n^0(t) &= -\epsilon_0 h_n^0(t) - \frac{N}{y_n^2} h_n^0(t) - \mu A_0 \frac{\theta}{y_n^2} N \frac{n_w^{n_w}}{(n_w-1)!} \sqrt{\frac{2\pi}{n_w}} y_n^2 e^{-n_w} \\
& \quad \cdot \left(e^{-\epsilon_0 x_w y_n^2} \frac{1}{y_n^2} T_e^N \left(t - x_w y_n^2 \right) + \epsilon \frac{1}{2} e^{-\epsilon_0 x_w y_n^2} \right. \\
& \quad \cdot \left(\left(\frac{1}{6 x_w y_n^2} + \epsilon_0^2 x_w y_n^2 \right) T_e^N \left(t - x_w y_n^2 \right) + 2\epsilon_0 x_w y_n^2 T_e^{N(1)} \left(t - x_w y_n^2 \right) \right. \\
& \quad \left. \left. + x_w y_n^2 T_e^{N(2)} \left(t - x_w y_n^2 \right) \right) + \mathcal{O}(\epsilon^2) \right), \\
\partial_t T_e^N(t) &= -c_T T_e^N(t) + c_h \frac{A_{rW}}{1+y_n^2} \frac{N^N}{(N-1)!} \sqrt{\frac{2\pi}{N}} e^{-N} \cdot \left(e^{-\epsilon_0} h_n^0(t-1) \right. \\
& \quad \left. + \epsilon \frac{1}{2} e^{-\epsilon_0} \left(\left(\frac{1}{6} + \epsilon_0^2 \right) h_n^0(t-1) + 2\epsilon_0 h_n^{0(1)}(t-1) + h_n^{0(2)}(t-1) \right) + \mathcal{O}(\epsilon^2) \right) \tag{149} \\
& + c_h \mu A_0 \left(1 - \frac{\theta}{1+y_n^2} \right) \frac{(N-n_w)^{N+1-n_w}}{(N-n_w)!} \sqrt{\frac{2\pi}{N-n_w}} e^{-(N-n_w)} \\
& \quad \cdot \left(e^{-\epsilon_0(1-x_w)} T_e^N \left(t - (1-x_w) \right) + \epsilon \frac{1}{2} e^{-\epsilon_0(1-x_w)} \right. \\
& \quad \cdot \left(\left(\frac{1}{6(1-x_w)} - 2\epsilon_0 + \epsilon_0^2(1-x_w) \right) T_e^N \left(t - (1-x_w) \right) \right. \\
& \quad \left. \left. - \left(2 - \epsilon_0(1-x_w) \right) T_e^{N(1)} \left(t - (1-x_w) \right) + (1-x_w) T_e^{N(2)} \left(t - (1-x_w) \right) \right) \right) \\
& \quad \left. + \mathcal{O}(\epsilon^2) \right).
\end{aligned}$$

Taking the limit $N \rightarrow \infty$ while dividing the equation for h_n^0 by N , yields the final delay system:

$$\begin{aligned}
\epsilon \partial_t h_n^0(t) &= -\epsilon \epsilon_0 h_n^0(t) - \frac{1}{y_n^2} h_n^0(t) - \mu A_0 \frac{\theta}{y_n^2} \frac{1}{y_n^2} e^{-\epsilon_0 x_w y_n^2} T_e^N(t - x_w y_n^2) \\
&\quad - \epsilon \mu A_0 \frac{\theta}{y_n^2} e^{-\epsilon_0 x_w y_n^2} \left(\frac{1}{2} \epsilon_0^2 x_w y_n^2 - \frac{1}{x_w y_n^2} \right) T_e^N(t - x_w y_n^2) + \epsilon_0 x_w y_n^2 T_e^{N(1)}(t - x_w y_n^2) \\
&\quad + \frac{1}{2} x_w y_n^2 T_e^{N(2)}(t - x_w y_n^2) + \mathcal{O}(\epsilon^2), \\
\partial_t T_e^N(t) &= -c_T T_e^N(t) + c_h \frac{A_{rW}}{1 + y_n^2} e^{-\epsilon_0} h_n^0(t - 1) \\
&\quad + c_h \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2} \right) e^{-\epsilon_0(1-x_w)} T_e^N(t - (1 - x_w)) \\
&\quad + \epsilon \left(c_h \frac{A_{rW}}{1 + y_n^2} e^{-\epsilon_0} \left(\left(\frac{1}{2} \epsilon_0^2 - 1 \right) h_n^0(t - 1) + \epsilon_0 h_n^{0(1)}(t - 1) + \frac{1}{2} h_n^{0(2)}(t - 1) \right) \right. \\
&\quad + c_h \mu A_0 \left(1 - \frac{\theta}{1 + y_n^2} \right) e^{-\epsilon_0(1-x_w)} \left(\left(\frac{1}{2} \epsilon_0^2 (1 - x_w) - \epsilon_0 - \frac{1}{1 - x_w} \right) T_e^N(t - (1 - x_w)) \right. \\
&\quad \left. \left. - \left(1 - \frac{1}{2} \epsilon_0 (1 - x_w) \right) T_e^{N(1)}(t - (1 - x_w)) + \frac{1}{2} (1 - x_w) T_e^{N(2)}(t - (1 - x_w)) \right) \right) \\
&\quad + \mathcal{O}(\epsilon^2).
\end{aligned} \tag{150}$$

Note that by letting $\epsilon \rightarrow 0$ and considering the equation for h_n^0 at time $t - 1$ the derived delay model for ENSO from Equation (57) is found.