UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ANALYSIS OF LOW-ORDER BOX MODELS FOR A CLASS OF OCEAN CIRCULATIONS

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

MASTER OF SCIENCE

By

TRUNG T. NGUYEN
Norman, Oklahoma
2015

ANALYSIS OF LOW-ORDER BOX MODELS
FOR A CLASS OF OCEAN CIRCULATIONS

A THESIS APPROVED FOR THE
SCHOOL OF COMPUTER SCIENCE

BY

Dr. S. Lakshmivarahan, Chair

Dr. K. Thulasiraman

Dr. Sudarshan Dhall

Acknowledgements

I would like to express my wholehearted appreciation to my advisor, Dr. Lakshmivarahan. I thank him for giving me the opportunity to do research with him and to learn from him. Without his guidance and encouragement, the completion of this thesis would nearly be impossible.

I extend my appreciation to the members of my thesis committee, Dr. Thulasiraman and Dr. Dhall. I thank them for offering their valuable time to serve on the committee.

My appreciation also goes to Dr. Abousleiman. I thank him for giving me opportunity to come to the United States for my Master's degree and to work at the PoroMechanics Institute (PMI), of which he is the director. I also want to thank my colleagues at PMI for their support. I especially am thankful to Mr. Son Hoang at PMI for his support and guidance at work as well as in life.

Finally, I am greatly indebted to my parents and family members for their unconditional support and caring. They always believe in me and that the reason why I always strive to be better and better.

Table of Contents

List of Tables

List of Figures

Abstract

Oceans occupy more than 70 percent of the planet earth. They play an important role in storing and transporting huge amount of heat. Understanding the physical behavior of the oceans is of interest to scientists in a variety of disciplines because ocean dynamics interacts with the atmospheric processes to govern climatic and biological processes. In this thesis, we investigate the dynamical behavior of the ocean models through a class of box models that capture the physics behind the thermohaline circulation and its changes as well. The goal is to gain some qualitative understanding of the stability and variability properties of thermohaline circulation. This steady-state analysis may provide a framework for a better understanding of climatic transitions between different stable regimes of the ocean system.

Chapter 1

Introduction

The key players that effect the unfolding of the real life drama called "climate dynamics" are the sun and the earth consisting of all of its oceans, its atmosphere and its chemical composition, land-surface and vegetation, the polar ice sheet, and the human activities. Of these, the oceans play a major role in the storage and redistribution of the thermal energy by evaporation and convection through a myriad of complex ocean circulations. In this thesis our goal is to provide a comprehensive review of a class of (spatially homogeneous) box-models that capture many of the qualitative features of one of the components of the ocean circulation known as the North Atlantic circulation (NAC).

To provide a broader perspective of the role by NAC on the climate, we start with some basic information on the earth, atmosphere and the ocean. The earth's surface area is $510 \times 10^6\ km^2$ (km – kilometer) of which nearly 71% ($362 \times 10^6\ km^2$) is covered the ocean. It is estimated that nearly all of earth's atmosphere is contained within the first $100\ km$ above the mean sea level and the outer space is said to begin at this level.

Our mother earth receives nearly $175\,PW$ (PW – petawatts and $1\,PW = 10^{15}$ watts) of incoming solar radiation in the upper reaches of our atmosphere. Roughly 30% of this input energy is reflected back into the outer space. The rest of the 70% ($112.5\,PW$) is used to warm the earth surface, the ocean waters and the atmosphere. Life on the earth cannot exist without this free supply of energy from the sun.

The sea water has a large thermal capacity and acts as a reservoir that stores and redistributes it to the cooler parts of the ocean from where it is radiated back to the atmosphere. This redistribution from the warmer and cooler regions is fundamental to maintaining the equilibrium condition known as the climate.

The primary mechanism for this redistribution of thermal energy is by circulation or convection. There are at least three different factors that affect this circulation: (1) surface wind and (2) the interaction of the gravity of the moon and the earth both of which create the visible ocean waves and (3) circulation induced by the pressure differences resulting from the changes in the density, $\rho(S,T)$ (measured as kg/km^3), of sea water with salinity, S (measured in terms of grams of salt in one kilogram of sea water and denoted by ‰) and the temperature, T (in degree ℃). The major focus of the box models is to concentrate on the dependence of NAC resulting from density differences.

The density $\rho(S,T)$ is a nonlinear function of S and T. It is well known that for a fixed T, ρ is on increasing function of S and for a given S, ρ is a decreasing function of T. Refer to Appendix A for a detailed discussion of the nonlinear empirical formula for $\rho(S,T)$ and many of its properties.

Atlantic waters in the tropical belt (covering $30°N$ to $30°S$ latitude) is warmer than those outside this belt, resulting in increased evaporation from the tropical belt. This has two effects. First, it increases the salinity and hence the density. Second, the warm water vapor raises and gets cooled and moved around by the atmospheric wind, which in turn enhances the cloud formation resulting in rain over the land and the ocean. While the pure rain water and the run-off from the land decreases the salinity and hence the density of the sea water, the latent heat release from condensation further fuels the wind, etc. Clearly, there is a tug of war between increased density resulting from evaporation and cooling and decreased density from increased fresh water flow and heating. The resulting circulation is called thermohaline circulation.

Under the prevailing climatic condition, Atlantic Ocean produces a thermohaline circulation where the cool dense water sinks in the northern hemisphere and is known to move south at depths of 2 – 4 km where it joins the other flows (Rahmstorf (1996)). To maintain the balance, there is a northward shallow warm water flow over the Atlantic.

The well-known Gulf Stream is an important component of this northward surface flux. It is estimated that the volume of water involved in this warm northward flux is about $17\,SV$, ($1\,SV = 10^6\,m^3/sec$). This convection is also responsible for the transfer of $1\,PW$ equivalent of thermal energy, which keeps the North Atlantic waters about 4℃ warmer than the corresponding Pacific waters at similar latitudes, and the air above the North Atlantic about 10℃ warmer.

To develop an appreciation for the amount of mass and heat transfer embedded in this thermohaline circulation, a few comparisons are in order. The combined rate of discharge from the American and the Canadian sides of the mighty Niagara Falls is only 750,000 US gallons/sec which is roughly $2.84 \times 10^3\,m^3/sec$, which is a miniscule compared to the flow of $17 \times 10^6\,m^3/sec$ (equivalent to 6×10^3 Niagara Falls) embedded in the northward flowing surface flux of the NAC. The total electric power consumed by the entire US population for the whole of the year 2012 was roughly $4.1\,PW$ which is about 4 times the energy level of $1\,PW$ transferred by NAC.

Much like the climate models, ocean models come in various shapes and sizes. The full models involve three spatial and the time dimension. These models are important to get a finer understanding of the nature of the ocean circulation. In this thesis we analyze the simplest possible spatially homogeneous model called the box – models.

There are essentially two ways to create spatially homogeneous models. The first is to expand the field of interest in a Fourier series where the spatial variations are captured by the chosen Fourier modes and the temporal variations by the dynamics of their amplitudes. The dynamics of variation of these amplitudes is then obtained by performing a Galerkin type projection onto a finite dimensional space corresponding to the dominant modes. This process helps to create a class of low-order models described by a system of nonlinear coupled ordinary differential equations. The second approach, pioneered by Stommel in 1961, is to capture the flow resulting from pressure difference resulting from the density differences using basic principles of flow in pipes.

While the full 4-D (3-space + time) models are useful to get information on the fine structure of the flow field, they are not useful to perform bifurcation analysis. The simple class of box models is amenable to easy analysis of the dependence of solutions on various parameters which in turn facilities bifurcation analysis.

In chapter 2, we provide a review and a comparison of the so-called lateral 2-box models. A summary of the vertical 2-box models is contained in chapter 3 where we describe the conditions for vertical oscillations again resulting from differential density of both sea and pure water. Chapters 4 and 5 contain a summary of the 3 and 4 box models, respectively. Chapter 6 contains the summary and problems for further research.

Chapter 2

Two-box models

2.1. A general two-box model

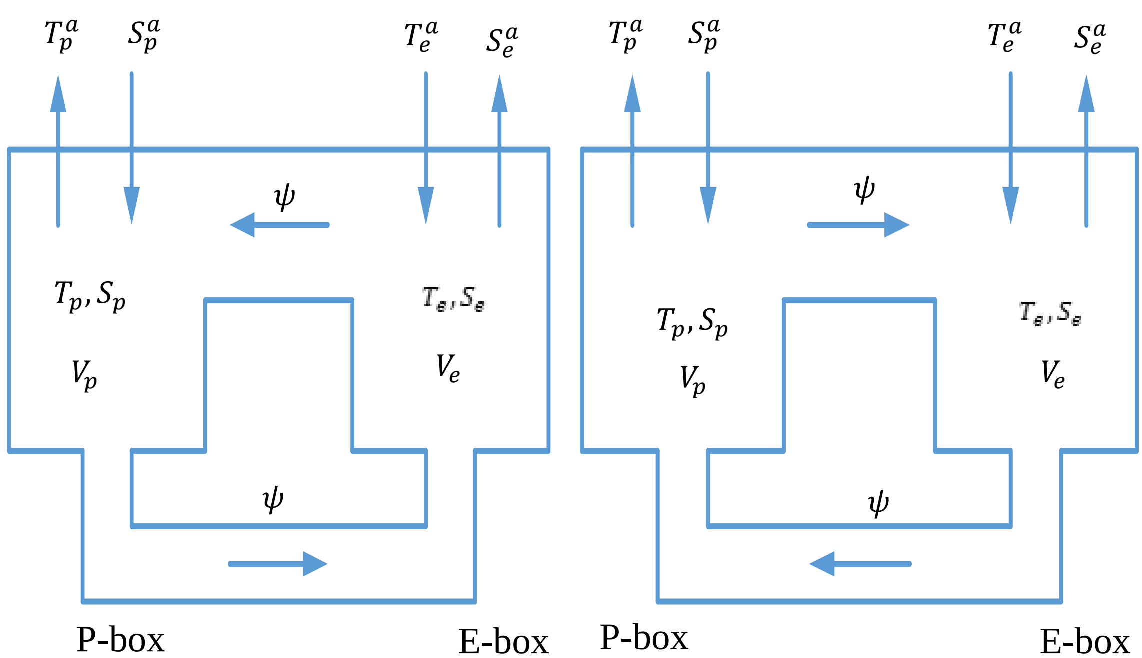
In this class of models the North Atlantic Ocean is represented by a system of two interconnected boxes as shown in Figure 2.1. The equatorial box (called E-box) is of volume V_e (measured in cubic meters – m^3) and polar box (called P-box) is of volume V_p. (Some authors specify that E-box extends northward from equator to $45°N$, and the P-box from $45°N$ to $75°N$ (Birchfield (1989), Roebber (1995)) and yet others assume $V_e = 2V_p$). It is assumed the sea water in both boxes are well mixed and the temperature (in °C) and salinity (grams of salt per kg of sea water, denoted by ‰) of the E-box are T_e and S_e, and those of the P-box are T_p and S_p, respectively.

It is also assumed that the two boxes are connected on the top by an over flow channel and at the bottom by a pipe as in Figure 2.1. These conduits allow for the exchange of heat and salt through convection/advection between the boxes.

Let T_e^a and T_p^a be the air temperature above the E and P-boxes where it is assumed that $T_e^a > T_p^a$ which reflects the current climate conditions. When $T_e^a > T_e$, the temperature of the E-box relaxes towards T_e^a by the standard Newtonian heat transfer process where $\frac{dT_e}{dt}$ is proportional to $(T_e^a - T_e)$. This increase in temperature immediately causes a reduction in the density $\rho(S,T)$ of the sea water in the E-box. As the sea water heats up, there is also evaporation (loss of fresh water) from the top of the E-box. Let $S_e^a > 0$ denote the increase in salinity resulting from the fresh water loss. Then, the salinity S_e increases where $\frac{dS_e}{dt}$ is proportional to $(S_e^a - S_e)$. This increases in salinity causes the density $\rho(S,T)$ of the E-box to increase.

From the point of view of modeling, it is useful to separate the temperature driven flow from the salinity driven flow. Recall from the above discussion, $T_e^a > T_e$ causes the density in E-box to decrease, but $T_p^a < T_p$ causes the density in the P-box to increase. This causes sinking in the P-box and upwelling in the E-box resulting in a temperature driven thermohaline circulation which is an anticlockwise flow ψ

(measured in cubic meters/sec – $m^3 sec^{-1}$) as shown in Figure 2.1 (a). On the other hand, evaporation denoted by S_e^a from the top of the E-box causes an increase in density in this box and freshwater in flow in the form of rain (denoted by S_p^a) results in a reduction of the density in the P-box. Consequently, there is sinking in the E-box and upwelling in the P-box resulting in the salinity driven thermohaline circulation which is a clockwise flow ψ as shown in Figure 2.1 (b). The primary goal of the two-box modeling effort is to understand the inherent tug of war between temperature driven and salinity driven flow and their dependence on the key parameters T_e^a, T_p^a, S_e^a, and S_p^a.



a) Anticlockwise (Positive) flow ψ (> 0) b) Clockwise (Negative) flow ψ (< 0)

Figure 2.1. A representation of the two-box model

Referring to Figure 2.1 (a) and (b), since the boxes are well mixed, it is immediate that flow ψ, independent of its direction, brings heat energy at the rate (ψT_p) into the E-box and takes away at the rate (ψT_e) from the E-box. Similarly, the flow ψ, independent of its direction, brings in salt at the rate (ψS_p) into the E-box and takes away at the rate (ψS_e) from the E-box. Consequently, the flow ψ contributes at the rate $|\psi|(T_p - T_e)$ towards increasing T_e and at the rate $|\psi|(S_p - S_e)$ towards increasing S_e, where $|\psi|$ denotes the magnitude of the flow ψ. A similar argument applies for P-box as well.

Combining these argument the dynamics of evolution of temperature and salinity is given by

$$V_e \frac{dT_e}{dt} = C_e^T (T_e^a - T_e) + |\psi|(T_p - T_e) \tag{2.1.1}$$

$$V_p \frac{dT_p}{dt} = C_p^T (T_p^a - T_p) + |\psi|(T_e - T_p) \tag{2.1.2}$$

$$V_e \frac{dS_e}{dt} = C_e^S (S_e^a - S_e) + |\psi|(S_p - S_e) \tag{2.1.3}$$

$$V_p \frac{dS_p}{dt} = C_p^S (S_p^a - S_p) + |\psi|(S_e - S_p) \tag{2.1.4}$$

where the units of the constants of proportionality C_e^T, C_p^T, C_e^S, and C_p^S are $m^3 sec^{-1}$. Dividing (2.1.1) and (2.1.3) by V_e, and (2.1.2) and (2.1.4) by V_p, we obtain the basis set of four equations given by

$$\frac{dT_e}{dt} = R_T (T_e^a - T_e) + \frac{|\psi|}{V_e}(T_p - T_e) \tag{2.1.5}$$

$$\frac{dT_p}{dt} = R_T (T_p^a - T_p) + \frac{|\psi|}{V_p}(T_e - T_p) \tag{2.1.6}$$

$$\frac{dS_e}{dt} = R_S (S_e^a - S_e) + \frac{|\psi|}{V_e}(S_p - S_e) \tag{2.1.7}$$

$$\frac{dS_p}{dt} = R_S (S_p^a - S_p) + \frac{|\psi|}{V_p}(S_e - S_p) \tag{2.1.8}$$

where

$$R_T = \frac{C_e^T}{V_e} = \frac{C_p^T}{V_p} \quad \text{and} \quad R_S = \frac{C_e^S}{V_e} = \frac{C_p^S}{V_p} \tag{2.1.9}$$

Thus, R_T and R_S have units sec^{-1}, where R_T^{-1} and R_S^{-1} are known as the temperature and salinity relaxation times or simply as time constants. It is well known that $R_T > R_S$.

Since the flow is induced by the temperature and salinity differences, it is useful to reduce the four equations to two by defining

$$\begin{aligned} T &= T_e - T_p, \text{ and } T^a = T_e^a - T_p^a \\ S &= S_e - S_p, \text{ and } S^a = S_e^a - S_p^a \end{aligned} \tag{2.1.10}$$

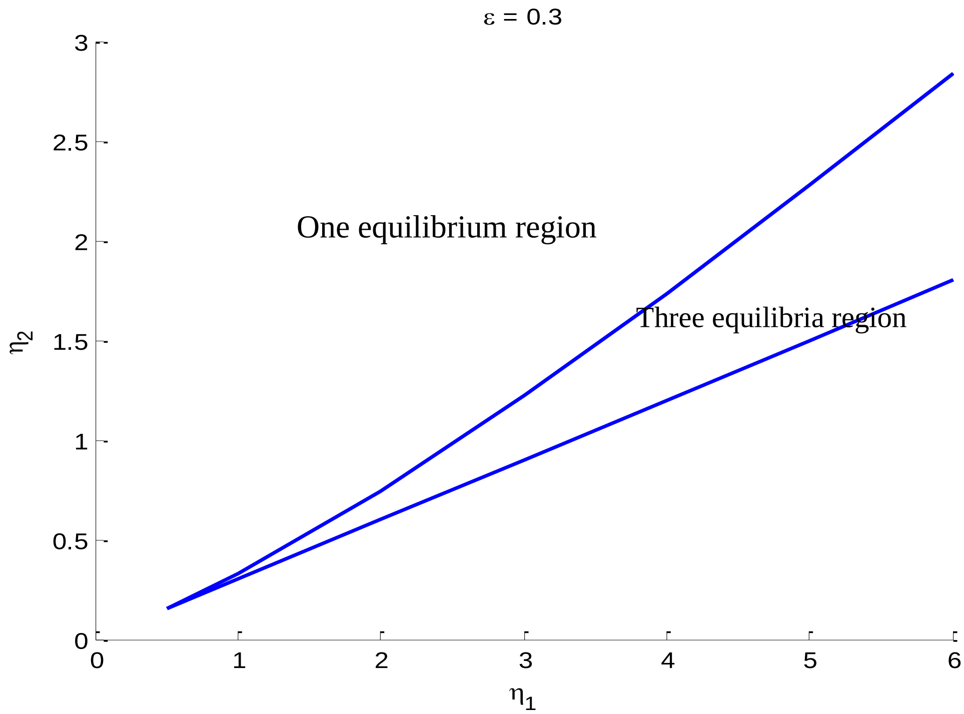


Figure 2.2. $(\boldsymbol{\eta_1}, \boldsymbol{\eta_2})$ parameter plane for one or three equilibria

Subtracting (2.1.6) from (2.1.5) and (2.1.8) from (2.1.7), we obtain

$$\frac{dT}{dt} = R_T(T^a - T) - \frac{|\psi|}{\bar{V}}T \tag{2.1.11}$$

$$\frac{dS}{dt} = R_S(S^a - S) - \frac{|\psi|}{\bar{V}}S \tag{2.1.12}$$

where

$$\bar{V} = \frac{V_e V_p}{(V_e + V_p)} \tag{2.1.13}$$

We now non-dimensionalize (2.1.11) – (2.1.12) by introducing a new set of scaled variables:

$$x = \frac{T}{\frac{\bar{V}R_T}{\gamma\alpha}}; \quad y = \frac{S}{\frac{\bar{V}R_T}{\gamma\beta}}$$
$$\bar{\psi} = \frac{\psi}{\bar{V}R_T}; \quad s = \frac{t}{(1/R_T)} \tag{2.1.14}$$

Table 2.1. The eigenvalues and eigenvectors at the equilibria when $\boldsymbol{\eta_1 = 3.0}$, $\boldsymbol{\eta_2 = 1.0}$, and $\boldsymbol{\varepsilon = 0}$.

	x	y	Eigenvalue	Eigenvector	Status
1	2.8778	2.9203	$-0.7136 - 1.3807 \times i$ $-0.7136 + 1.3807 \times i$	$(0.6194 - 0.3355 \times i, 0.7096)$ $(0.6194 + 0.3355 \times i, 0.7096)$	Stable
2	1.7035	0.9424	-0.6991 -2.8840	$(0.5244, 0.8514)$ $(0.9465, 0.3225)$	Stable
3	2.8251	2.7632	0.6991 -2.1848	$(0.5244, 0.8514)$ $(0.8565, 0.5160)$	Unstable

Table 2.2. The eigenvalue and eigenvector at the equilibrium when $\boldsymbol{\eta_1 = 1.0}$, $\boldsymbol{\eta_2 = 1.0}$, and $\boldsymbol{\varepsilon = 0.3}$

	x	y	Eigenvalue	Eigenvector	Status
1	0.6491	1.1896	$-1.4608 - 0.6693 \times i$ $-1.4608 + 0.6693 \times i$	$(0.3849 - 0.4525 \times i, 0.8043)$ $(0.3849 + 0.4525 \times i, 0.8043)$	Stable

Substituting (2.1.14) in (2.1.11) – (2.1.12) and simplifying, we obtain

$$\begin{aligned} \frac{dx}{ds} &= \eta_1 - x(1 + |x - y|) \\ \frac{dy}{ds} &= \eta_2 - y(\varepsilon + |x - y|) \end{aligned} \tag{2.1.15}$$

where

$$\begin{aligned} \varepsilon &= \frac{R_S}{R_T} \\ \eta_1 &= \frac{T^a \gamma \alpha}{\bar{V} R_T}; \quad \eta_2 = \varepsilon \frac{S^a \gamma \beta}{\bar{V} R_T} \end{aligned} \tag{2.1.16}$$

All the known two-box models are derived from (2.1.11) – (2.1.12), but they differ the way in which (i) the flow ψ and (ii) fresh water flux or equivalently S_p^* and S_e^* are modeled.

With fixed $\varepsilon = 0.3$, the regions in the (η_1, η_2) parameter plane for which the equations have one or three equilibria are shown in Figure 2.2.

With $\eta_1 = 3.0$, $\eta_2 = 1.0$, and $\varepsilon = 0.3$, these equations have three equilibria. The eigenvalues and eigenvectors at the equilibria are shown in Table 2.1. The values of x and y for steady states are shown in Figure 2.3.

With $\eta_1 = 1.0$, $\eta_2 = 1.0$, and $\varepsilon = 0.3$, these equations have one equilibrium. The eigenvalue and eigenvector at the equilibrium are shown in Table 2.2. The values of x and y for steady state are shown in Figure 2.4.

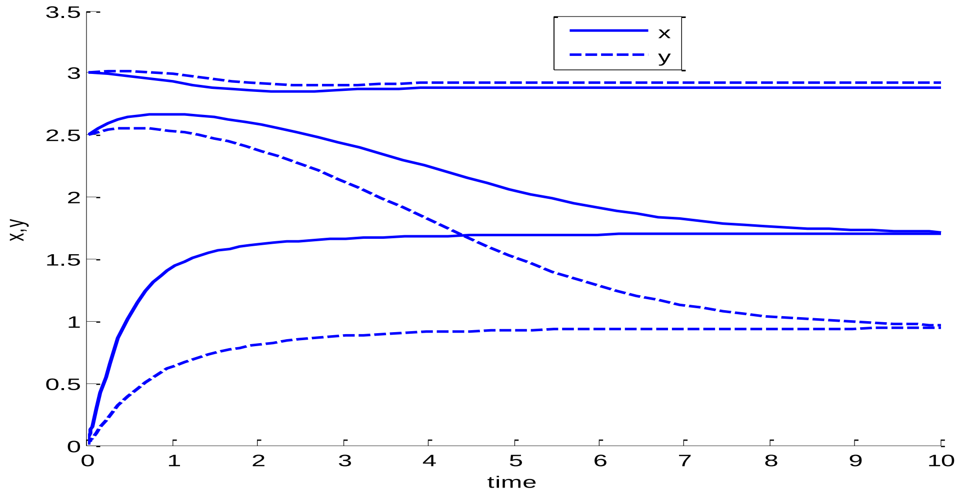


Figure 2.3. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\eta_1 = 3.0}$, $\boldsymbol{\eta_2 = 1.0}$, and $\boldsymbol{\varepsilon = 0.3}$

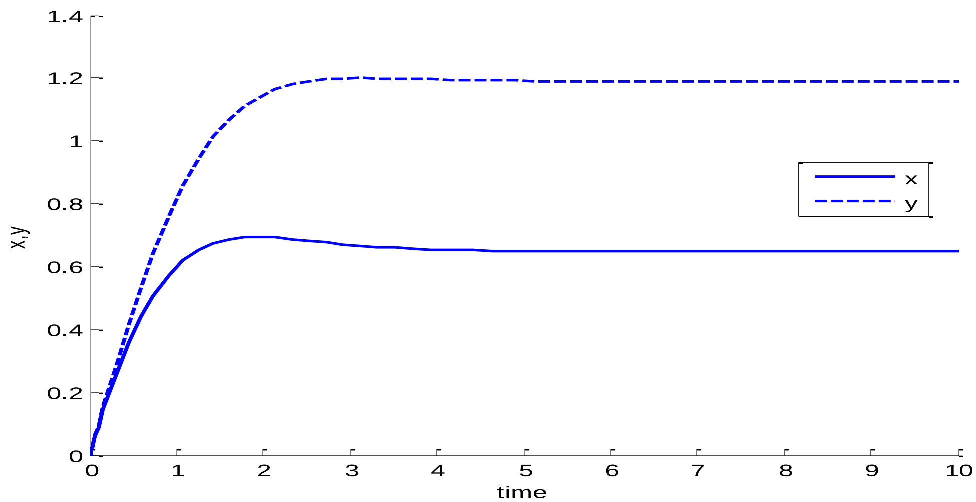


Figure 2.4. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\eta_1 = 1.0}$, $\boldsymbol{\eta_2 = 1.0}$, and $\boldsymbol{\varepsilon = 0.3}$

2.2. Stommel's model

Stommel in his 1961 classic paper considers as symmetric model by setting

$$\begin{aligned} &T_e^a = -T_p^a = T^a, S_e^a = -S_p^a = S^a \\ &T_e = -T_p = T, S_e = -S_p = S \\ &V_e = V_p = V, \bar{V} = \frac{V}{2}, and\ q = \frac{\psi}{V} \end{aligned} \tag{2.2.1}$$

where q is of dimension sec^{-1}. Substituting (2.2.1) into (2.1.11) – (2.1.12), we get

$$\begin{aligned} \frac{dT}{dt} &= R_T(T^a - T) - 2|q|T \\ \frac{dS}{dt} &= R_S(S^a - S) - 2|q|S \end{aligned} \tag{2.2.2}$$

To get a non-dimensional model, we define non-dimensional variables as

$$x = \frac{T}{T^a}, y = \frac{S}{S^a}, and\ r = R_T t \tag{2.2.3}$$

Substituting (2.2.3) into (2.2.2) and simplifying, we get the model equation of the form

$$\begin{aligned} \frac{dx}{dr} &= (1 - x) - |f|x \\ \frac{dy}{dr} &= \varepsilon_s(1 - y) - |f|y \end{aligned} \tag{2.2.4}$$

where the non-dimensional flow f is given by

$$|f| = \frac{2|q|}{R_T} \text{ and the Stommel epsilon } \varepsilon_s = \frac{R_S}{R_T} < 1 \tag{2.2.5}$$

To uncover the dependence of the flow variable f on the variables x and y, Stommel invokes the linearized version of the equation of state given by

$$kq = \rho_E - \rho_P = \rho_o\left[\alpha\left(T_e - T_p\right) - \beta\left(S_e - S_p\right)\right] \tag{2.2.6}$$

where k is called the resistance to the fluid flow. Using (2.2.1) and (2.2.3), it can be verified that

$$kq = 2\rho_o\alpha T^a(x - Ry) \tag{2.2.7}$$

where the ratio R is given by

$$R = \frac{\beta S^a}{\alpha T^a} \tag{2.2.8}$$

which decides the sign of the flow. Hence, it can be verified that

$$x - Ry = f\lambda \quad \text{and} \quad \lambda = \frac{R_T k}{4\rho_o \alpha T^a} \tag{2.2.9}$$

Substituting for f in (2.2.4), latter becomes

$$\begin{aligned} \frac{dx}{dr} &= (1-x) - \frac{x}{\lambda}|x - Ry| \\ \frac{dy}{dr} &= \varepsilon_s(1-y) - \frac{y}{\lambda}|x - Ry| \end{aligned} \tag{2.2.10}$$

Table 2.3. The eigenvalues and eigenvectors at the equilibria when $\varepsilon_s = 1/6$, $\lambda = 1/5$, and $R = 2$

	x	y	Eigenvalue	Eigenvector	Status
1	0.8202	0.4320	$-0.9119 - 1.8230 \times i$ $-0.9119 + 1.8230 \times i$	$(0.8896, 0.4115 + 0.1977 \times i)$ $(0.8896, 0.4115 - 0.1977 \times i)$	Stable
2	0.4835	0.1349	-0.7608 -3.6095	$(0.7922, 0.6102)$ $(0.9839, 0.1783)$	Stable
3	0.7650	0.3518	0.7608 -2.8486	$(0.7922, 0.6102)$ $(0.9582, 0.2860)$	Unstable

Table 2.4. The eigenvalue and eigenvector at the equilibrium when $\boldsymbol{\varepsilon_s = 1}$, $\boldsymbol{\lambda = 1/5}$, and $\boldsymbol{R = 2}$

	x	y	Eigenvalue	Eigenvector	Status
1	0.3582	0.3582	-2.7912 -4.5825	$(0.8944, 0.4472)$ $(0.7071, 0.7071)$	Stable

With $\varepsilon_s = 1/6$, $\lambda = 1/5$, and $R = 2$, these equations have three equilibria. The eigenvalues and eigenvectors at the equilibria are shown in Table 2.3. The values of x and y for steady states are shown in Figure 2.5.

With $\varepsilon_s = 1$, $\lambda = 1/5$, and $R = 2$, these equations have one equilibrium. The eigenvalue and eigenvector at the equilibrium are shown in Table 2.4. The values of x and y for steady state are shown in Figure 2.6.

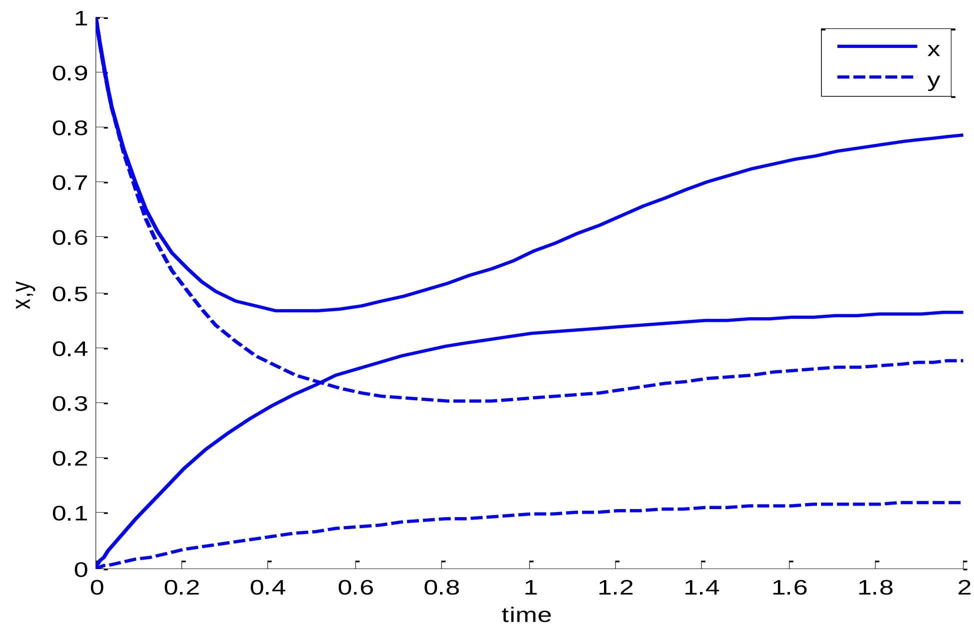


Figure 2.5. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\varepsilon_s = 1/6}$, $\boldsymbol{\lambda = 1/5}$, and $\boldsymbol{R = 2}$

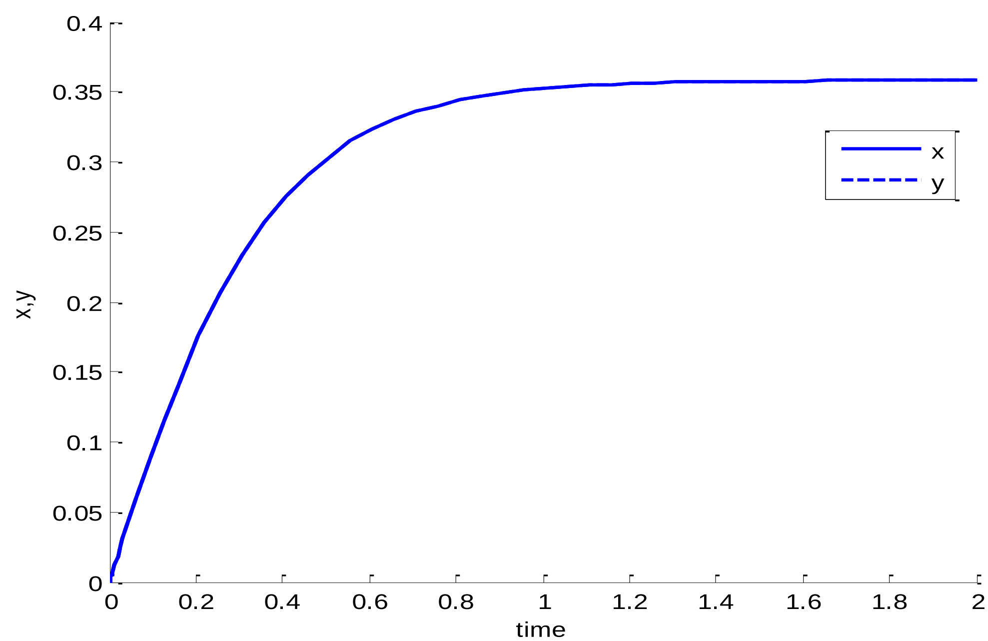


Figure 2.6. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\varepsilon_s = 1}$, $\boldsymbol{\lambda = 1/5}$, and $\boldsymbol{R = 2}$

2.3. Cessi's model

Like Stommel (1961), Cessi (1996) also considers a symmetric model but with a different parameterization for the fresh water flux and the flow. Refer to Figure 2.2 for the basic set up of the Cessi's model. The model dynamics can be either obtained from the first principle or by a special choice of terms in (2.1.11) – (2.1.12) such as

$$T_e^a = -T_p^a = \frac{\theta}{2} \qquad V_e = V_p = V \tag{2.3.1}$$

Instead of representing the fresh water loss due to evaporation from the top of the E-box and fresh water intake by way of rain into the top of the P-box by their salinity equivalent, Cessi represents these by explicit fresh water flux terms $\frac{F(t)}{2}$ from the E-box and $-\frac{F(t)}{2}$ into the P-box as given in Figure 2.2. The flow ψ is denoted by $\frac{\theta}{2}$ which is modelled differently compared to Stommel as will be shown below.

Using these substitutions, the Cessi model equations become

$$\frac{dT}{dr} = R_T(\theta - T) - Q(\Delta\rho)T \qquad \frac{dS}{dr} = \frac{F(t)}{H} S_o - Q(\Delta\rho)S \tag{2.3.2}$$

where T and S are as defined in (2.1.10), $S_0 = 35‰$ is the standard salinity, H is the height of the box (m). The dimensionless

$$\Delta\rho = \frac{\rho_E - \rho_P}{\rho_o} = -\alpha T + \beta S \tag{2.3.3}$$

where $\rho_o = 1029 kg.m^{-3}$ is the standard density of sea water at $T_o = 5$°C and $S_0 = 35‰$, $\alpha = 0.17 \times 10^{-3}$°C^{-1} and $\beta = 0.75 \times 10^{-3} psu^{-1}$.

Cessi postulates that Q can be modeled in one of three ways as

$Q_1 = R_d(sec^{-1})$, the diffusion time scale or time constant

$$Q_2 = R_d + \frac{q}{V}|\Delta\rho| \qquad Q_3 = R_d + \frac{q}{V}(\Delta\rho)^2 \tag{2.3.4}$$

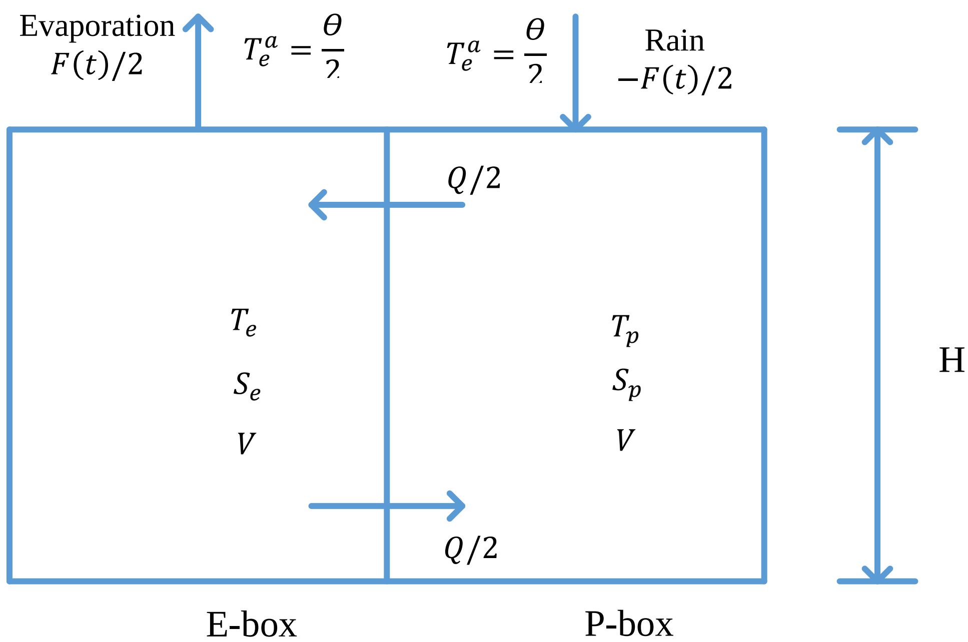


Figure 2.7. Cessi's model

The choice of $Q = Q_1$ renders the model linear, and analysis of a model with $Q = Q_2$ is pursued in Section 2.4. while Q_2 and Q_3 are qualitatively similar, Cessi and Young (1992) argued that under certain restriction on parameter ranges, Boussinesq convection leads to the choice Q_3.

By way non-dimensionalizing (2.3.2) define

$$x = \frac{T}{\theta}, y = \frac{\beta S}{\alpha \theta}, \text{ and the fast time } s = R_d t \tag{2.3.5}$$

First, using these in Q_3, the latter becomes

$$Q_3 = R_d[1 + \eta_c^2(x - y)^2] \tag{2.3.6}$$

where

$$\eta_c^2 = \frac{q}{V}\frac{(\alpha\theta)^2}{R_d} \tag{2.3.7}$$

Using (2.3.5) and (2.3.6) – (2.3.7) in (2.3.2), the latter becomes

$$\begin{aligned} \varepsilon \frac{dx}{ds} &= (1 - x) - \varepsilon x[1 + \eta_c^2(x - y)^2] \\ \frac{dy}{ds} &= \mu_c - y[1 + \eta_c^2(x - y)^2] \end{aligned} \tag{2.3.8}$$

where

$$\mu_c = \frac{\beta S_o}{H\alpha\Theta R_d} F(t)$$

and (2.3.9)

$$\varepsilon = \frac{R_d}{R_t} \ll 1$$

Table 2.5. The eigenvalues and eigenvectors at the equilibria when $\varepsilon = 0.01$, $\eta_c^2 = 7.5$, and $\mu_c = 1$

	x	y	Eigenvalue	Eigenvector	Status
1	0.9900	0.9993	-1.1397 -100.8628	$(-0.0013, 0.9999)$ $(0.9999, -0.0013)$	Stable
2	0.9491	0.1865	-3.4336 -116.0133	$(0.0958, 0.9953)$ $(0.9998, 0.0189)$	Stable
3	0.9878	0.8123	0.8544 -103.7785	$(0.0248, 0.9996)$ $(0.9997, 0.0204)$	Unstable

Table 2.6. The eigenvalue and eigenvector at the equilibrium when $\boldsymbol{\varepsilon = 0.01}$, $\boldsymbol{\eta_c^2 = 7.5}$, and $\boldsymbol{\mu_c = 1.5}$

	x	y	Eigenvalue	Eigenvector	Status
1	0.9874	1.1782	-4.7472 -98.3451	$(-0.0301, 0.9995)$ $(0.9993, -0.0359)$	Stable

With $\varepsilon = 0.01$, $\eta_c^2 = 7.5$, and $\mu_c = 1$, these equations have three equilibria. The eigenvalues and eigenvectors at the equilibria are shown in Table 2.5. The values of x and y for steady states are shown in Figure 2.9.

With $\varepsilon = 0.01$, $\eta_c^2 = 7.5$, and $\mu_c = 1.5$, these equations have one equilibrium. The eigenvalue and eigenvector at the equilibrium are shown in Table 2.6. The values of x and y for steady states are shown in Figure 2.10.

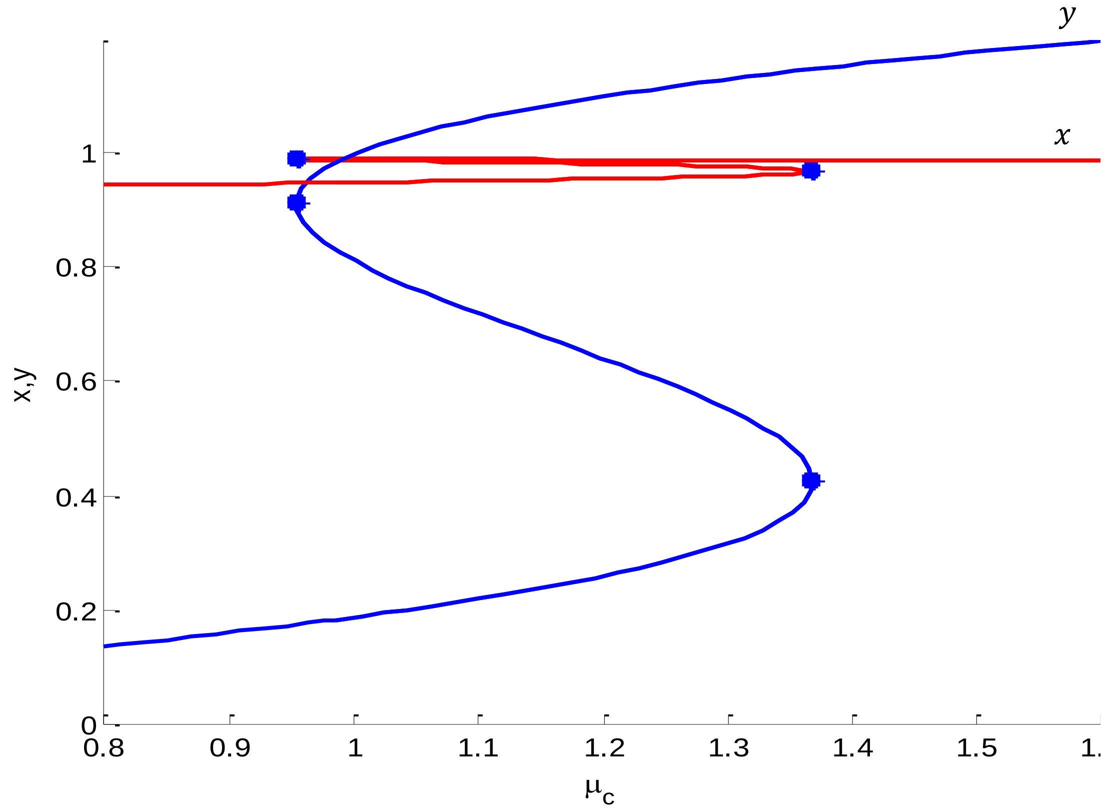


Figure 2.8. Bifurcation diagram for $\boldsymbol{\varepsilon = 0.01}$, $\boldsymbol{\eta_c^2 = 7.5}$, change $\boldsymbol{\mu_c}$

From Figure 2.8, we see that

- $\mu_c > 1.367681$, the system has one equilibrium with $x < y$.
- $\mu_c < 0.953247$, the system has one equilibrium with $x > y$,
- $0.953247 < \mu_c < 1.367681$, the system has three equilibria (two stable equilibria and one unstable equilibrium).

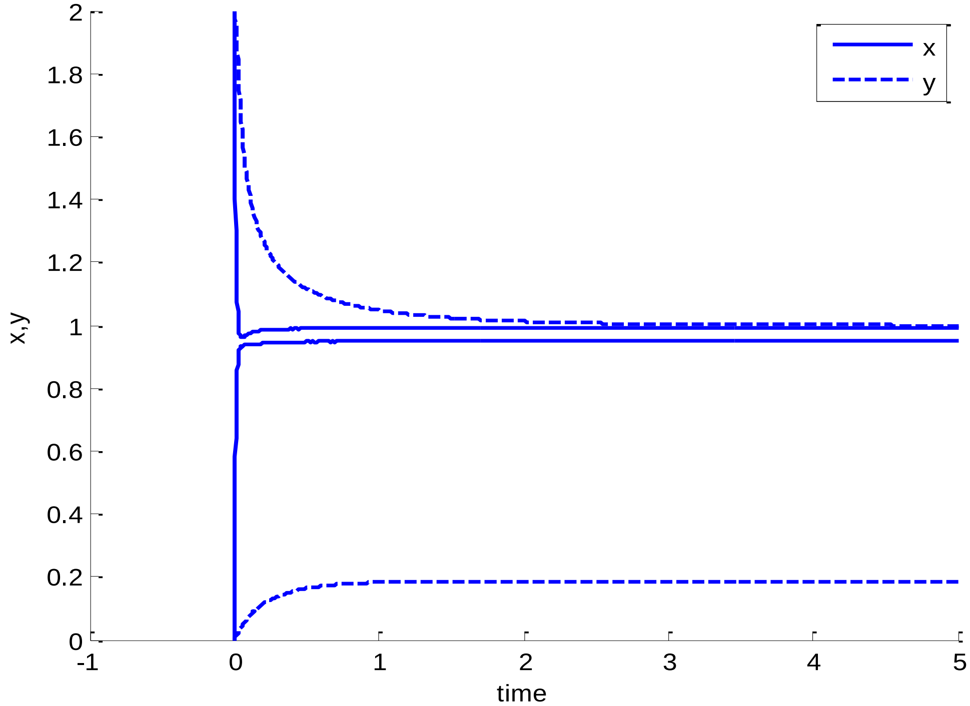


Figure 2.9. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\varepsilon} = \mathbf{0.01}$, $\boldsymbol{\eta_c^2} = \mathbf{7.5}$, and $\boldsymbol{\mu_c} = \mathbf{1}$

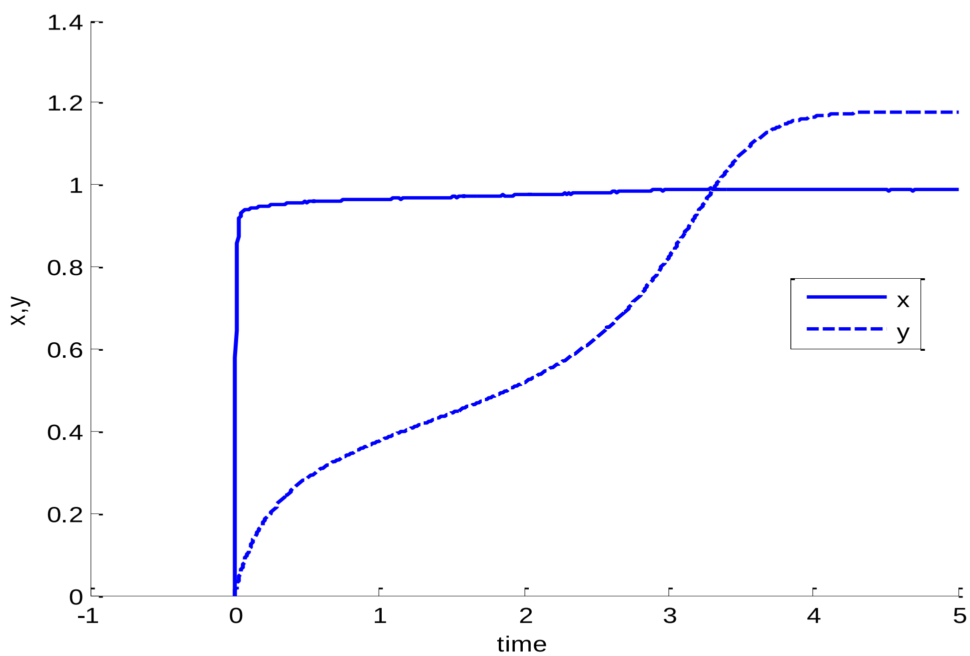


Figure 2.10. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\varepsilon} = \mathbf{0.01}$, $\boldsymbol{\eta_c^2} = \mathbf{7.5}$, and $\boldsymbol{\mu_c} = \mathbf{1.5}$

2.4. Van Veen's model

Van Veen (2001) considers a version the 2-box model which differs from Cessi's model in Section 2.3 in two respects. The precipitation is replaced by a constant flux and the flow Q is modeled using Q_2 in (2.3.4). With these changes, Van Veen's model equations are given by

$$\begin{aligned} \frac{dT}{dt} &= R_T(T^a - T) - (R_d + |f(S,T)|)T \\ \frac{dS}{dt} &= \delta - (R_d + |f(S,T)|)S \end{aligned} \tag{2.4.1}$$

where δ is the atmospheric water transport forcing

$$f(S,T) = \omega T - \xi S \tag{2.4.2}$$

and $f > 0$ if temperature driven and $f < 0$ if salinity driven.

Again, non-dimensionalize (2.4.1) using

$$x = \frac{T}{T^a}, \quad y = \frac{\xi S}{\omega T^a}, \quad \text{and the fast time} \quad s = R_d t \tag{2.4.3}$$

to obtain

$$\begin{aligned} \varepsilon\frac{dx}{ds} &= (1 - x) - \varepsilon x[1 + \eta_v|x - y|] \\ \frac{dy}{ds} &= \mu_v - y[1 + \eta_v|x - y|] \end{aligned} \tag{2.4.4}$$

where

$$\begin{aligned} \mu_v &= \frac{\xi R_T}{\omega R_d}\delta \\ \eta_v &= \frac{\omega}{R_d R_T} \end{aligned} \tag{2.4.5}$$

and

$$\varepsilon = \frac{R_d}{R_t} \ll 1$$

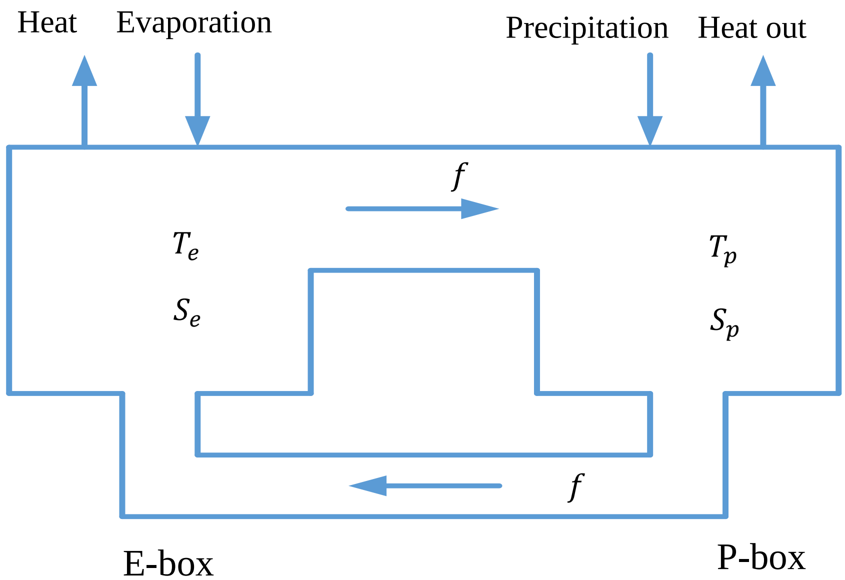


Figure 2.11. Van Veen's 2-box model

Table 2.7. The eigenvalues and eigenvectors at the equilibria when $\varepsilon = 0.1$, $\eta_v = 216.67$, and $\mu_v = 3$

	x	y	Eigenvalue	Eigenvector	Status
1	0.7060	0.7206	$-10.7441 - 38.9636 \times i$ $-10.7441 + 38.9636 \times i$	$(0.6807 - 0.1773 \times i, 0.7107)$ $(0.6807 + 0.1773 \times i, 0.7107)$	Stable
2	0.2371	0.0932	-27.7426 -77.7761	$(0.6153, 0.7882)$ $(0.9559, 0.2934)$	Stable
3	0.6929	0.6771	27.7426 -50.0335	$(0.6153, 0.7882)$ $(0.7950, 0.6065)$	Unstable

Table 2.8. The eigenvalue and eigenvector at the equilibrium when $\boldsymbol{\varepsilon = 0.1}$, $\boldsymbol{\eta_v = 216.67}$, and $\boldsymbol{\mu_v = 25}$

	x	y	Eigenvalue	Eigenvector	Status
1	0.1425	0.4155	-77.5531 -111.9222	$(0.6279, 0.3912)$ $(0.7782, 0.9202)$	Stable

With $\varepsilon = 0.1$, $\eta_v = 216.67$, and $\mu_v = 3$, these equations have three equilibria. The eigenvalues and eigenvectors at the equilibria are shown in Table 2.7. The values of x and y for steady states are shown in Figure 2.12.

With $\varepsilon = 0.1$, $\eta_v = 216.67$, and $\mu_v = 25$, these equations have one equilibrium. The eigenvalues and eigenvectors at the equilibrium are shown in Table 2.8. The values of x and y for steady state are shown in Figure 2.13.

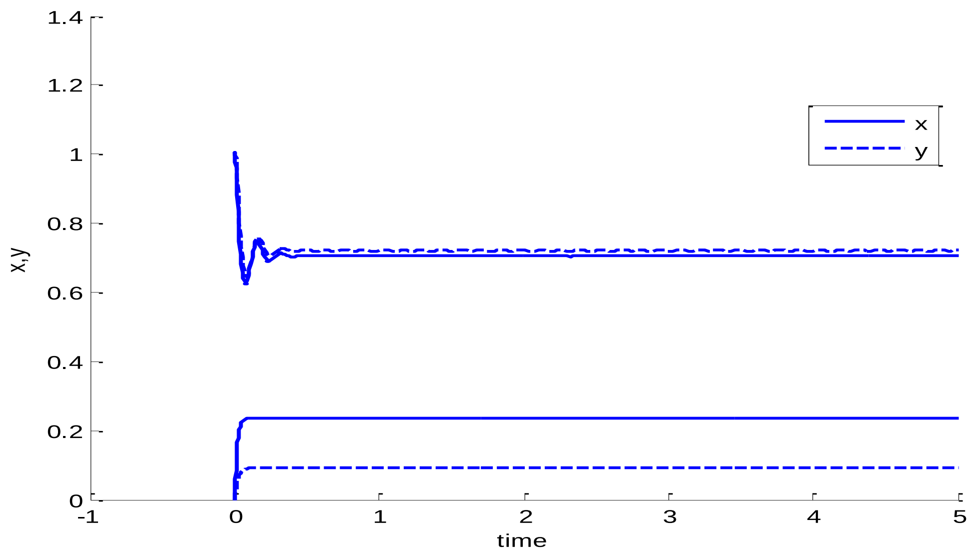


Figure 2.12. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\varepsilon = 0.1}$, $\boldsymbol{\eta_v = 216.67}$, and $\boldsymbol{\mu_v = 3}$

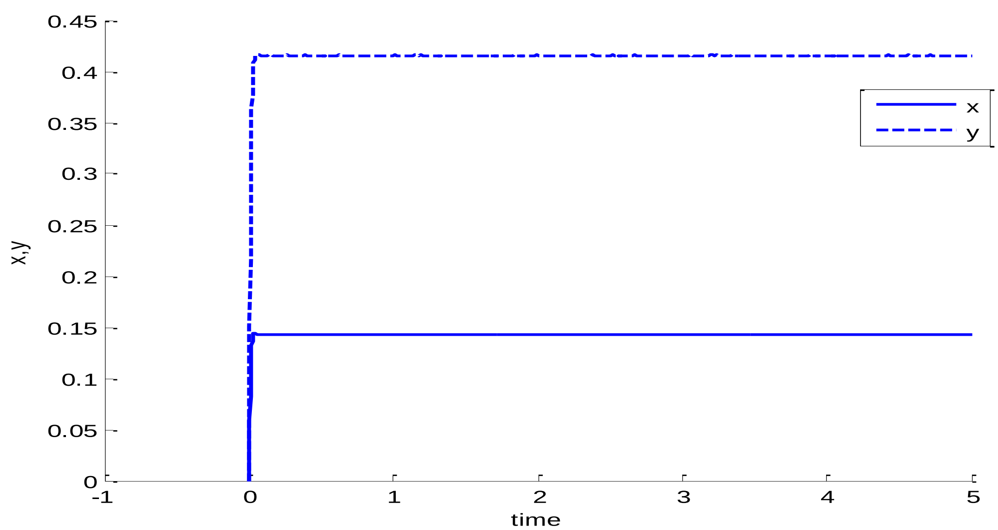


Figure 2.13. Plot of values of $\boldsymbol{x}$ and $\boldsymbol{y}$ for steady states with $\boldsymbol{\varepsilon = 0.1}$, $\boldsymbol{\eta_v = 216.67}$, and $\boldsymbol{\mu_v = 25}$

2.5. Marotzke's model

Marotzke (1990) contains, among others, a clear and concise description of a version of the 2-box model described in this subsection. The temperature difference $T = T_e - T_p$ is held fixed and the salinities evolve in response to evaporation/rain on one hand and to the flow on the other. Again the flow is modeled as

$$\psi = k(\alpha T - \beta S) \tag{2.5.1}$$

when $S = S_e - S_p$. Hence the dynamics of salinities in the boxes are given by

$$\begin{aligned} \frac{dS_e}{dt} &= F_s - |\psi|\left(S_e - S_p\right) \\ \frac{dS_p}{dt} &= -F_s + |\psi|\left(S_e - S_p\right) \end{aligned} \tag{2.5.2}$$

Subtracting, we get

$$\frac{dS}{dt} = 2F_s - 2|\psi|S \tag{2.5.3}$$

$$\frac{dS}{dt} = 2F_s - 2k|(\alpha T - \beta S)|S \tag{2.5.4}$$

Non-dimensionalize (2.5.4) using

$$S^* = \frac{\beta S}{\alpha T}, \quad F^* = \frac{\beta F_s}{k(\alpha T)^2},$$

then the dimensionless overturning rate $\psi^* = \frac{\psi}{k\alpha T} = 1 - S^*$, (2.5.5)

and the dimensionless time $t^* = 2k\alpha T t$

Equation (2.5.4) can be written as

$$\frac{dS^*}{dt^*} = F^* - |1 - S^*|S^* \tag{2.5.6}$$

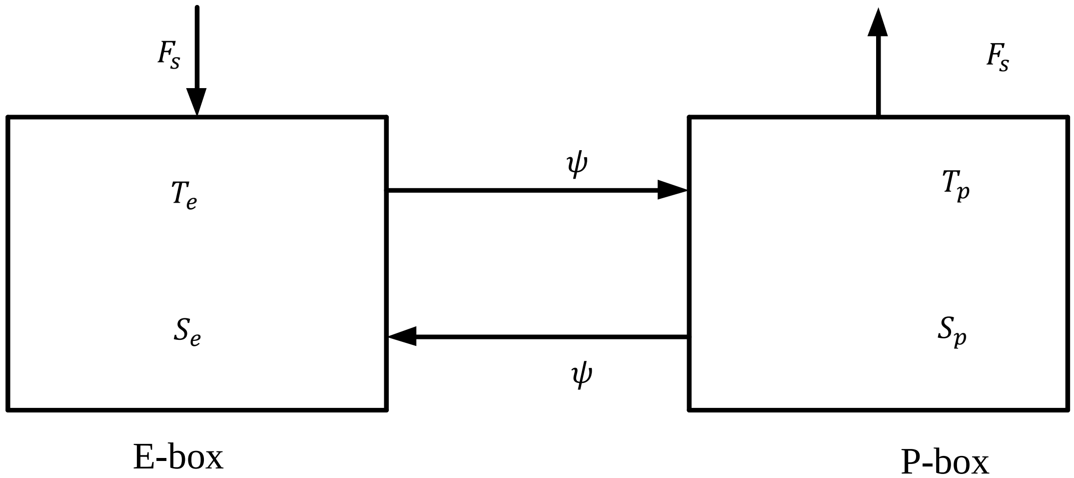


Figure 2.14. Marotzke's 2-box model set up where $T = T_e - T_p$ is held fixed.

which, given the dynamics (2.5.1), results in an equation for the derivatives of ψ^*

$$\frac{d\psi^*}{dt^*} = -F^* + |1 - \psi^*|\psi^* \tag{2.5.7}$$

We can derive stationary solution, which occur when the right-hand side of (2.5.7) equals zero.

$$-F^* + |1 - \psi^*|\psi^* = 0 \tag{2.5.8}$$

$$F^* = |1 - \psi^*|\psi^* \tag{2.5.9}$$

The solutions of (2.5.9) with $\psi^* > 0$, which are equivalent with $\psi > 0$, $\alpha T > \beta S$, represent the thermally driven thermohaline circulation. And the solutions of (2.5.9) with $\psi^* < 0$, which are equivalent with $\psi < 0$, $\alpha T < \beta S$, represent the salinity-driven thermohaline circulation. The plot of $f(\psi^*) = |1 - \psi^*|\psi^*$ is shown in Figure 2.15.

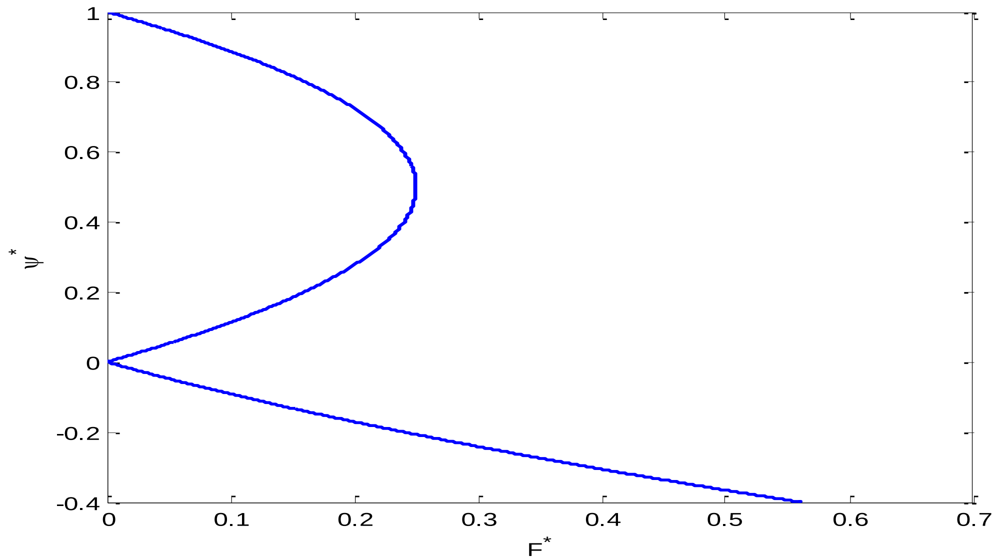


Figure 2.15. The plot of $\boldsymbol{f(\psi^*) = |1 - \psi^*|\psi^*}$

From Figure 2.16, we see that

- If $0 < F^* < 0.25$, the system has three equilibria. With $F^* = 0.1$, the equation has three equilibria (shown in Table 2.9).
- If $F^* > 0.25$, the system has only one equilibrium. With $F^* = 0.3$, the equation has one equilibrium (shown in Table 2.10).

Table 2.9. The eigenvalues and eigenvectors at the equilibria when $F^* = 0.1$

x	ψ^*	Eigenvalue	Status
1	0.8872	-0.7745	Stable
2	0.1127	0.7745	Unstable
3	-0.0916	-1.1832	Stable

Table 2.10. The eigenvalue and eigenvector at the equilibrium when $\boldsymbol{F^* = 0.3}$

	ψ^*	Eigenvalue	Status
1	-0.2416	-1.4832	Stable

2.6. Sensitivity analysis

a) Cessi's model

The equations of the model

$$\varepsilon\frac{dx}{ds} = (1-x) - \varepsilon x[1+\eta_c^2(x-y)^2]$$
$$\frac{dy}{ds} = \mu_c - y[1+\eta_c^2(x-y)^2] \tag{2.6.1}$$

is equivalent with

$$\dot{x} = \frac{1-x}{\varepsilon} - x[1+\eta_c^2(x-y)^2]$$
$$\dot{y} = \mu_c - y[1+\eta_c^2(x-y)^2] \tag{2.6.2}$$

Let $a = \frac{1}{\varepsilon}$, parameters $\alpha = (a, \mu_c, \eta_c)$, and variables $X = (x, y)$

$$\dot{x} = a(1-x) - x[1+\eta_c^2(x-y)^2] = f_1(x,y,\alpha)$$
$$\dot{y} = \mu_c - y[1+\eta_c^2(x-y)^2] = f_2(x,y,\alpha) \tag{2.6.3}$$

- **Sensitivity matrix for initial conditions** $\boldsymbol{X} = \begin{pmatrix}\boldsymbol{x(t, x(0), y(0), \alpha)},\\ \boldsymbol{y(t, x(0), y(0), \alpha)}\end{pmatrix}$

$$\frac{d}{dt}\begin{bmatrix}\frac{\partial x(t)}{\partial x(0)} & \frac{\partial x(t)}{\partial y(0)}\\ \frac{\partial y(t)}{\partial x(0)} & \frac{\partial y(t)}{\partial y(0)}\end{bmatrix} = \begin{bmatrix}\frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y}\\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y}\end{bmatrix} \times \begin{bmatrix}\frac{\partial x(t)}{\partial x(0)} & \frac{\partial x(t)}{\partial y(0)}\\ \frac{\partial y(t)}{\partial x(0)} & \frac{\partial y(t)}{\partial y(0)}\end{bmatrix} \tag{2.6.4}$$

$$\frac{d}{dt}D_{X(0)}(X) = D_X(t) * D_{X(0)}(X) \tag{2.6.5}$$

where

$$D_{X(0)}\big(X(0)\big) = \begin{bmatrix}\frac{\partial x(0)}{\partial x(0)} & \frac{\partial x(0)}{\partial y(0)}\\ \frac{\partial y(0)}{\partial x(0)} & \frac{\partial y(0)}{\partial y(0)}\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} = I \tag{2.6.6}$$

$D_X(t) =$

$$\begin{bmatrix}-1-a-\eta_c^2(x-y)^2-\eta_c^2x(2x-2y) & \eta_c^2x(2x-2y)\\ -\eta_c^2y(2x-2y) & \eta_c^2y(2x-2y)-\eta_c^2(x-y)^2-1\end{bmatrix} \tag{2.6.7}$$

With $\varepsilon = 0.01$, $\mu_c = 1$ and $\eta_c^2 \simeq 7.5$, $x(0) = 0.9378$, $y(0) = 0.8123$, Figure 2.16 and 2.17 show the sensitivity to initial conditions of the model.

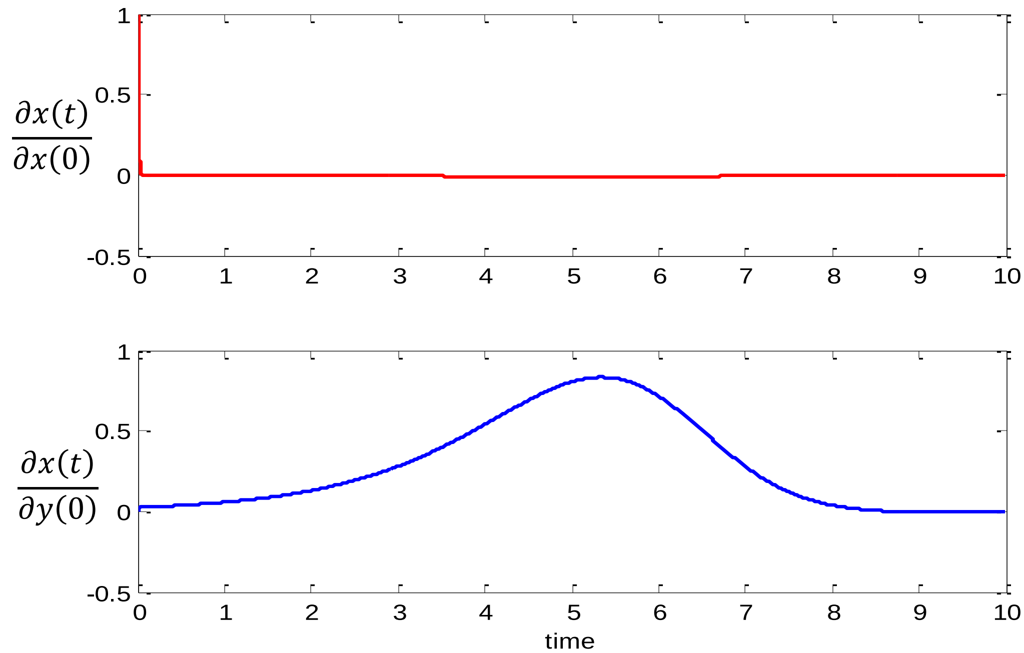


Figure 2.16. The sensitivity to initial conditions of $\boldsymbol{x}$

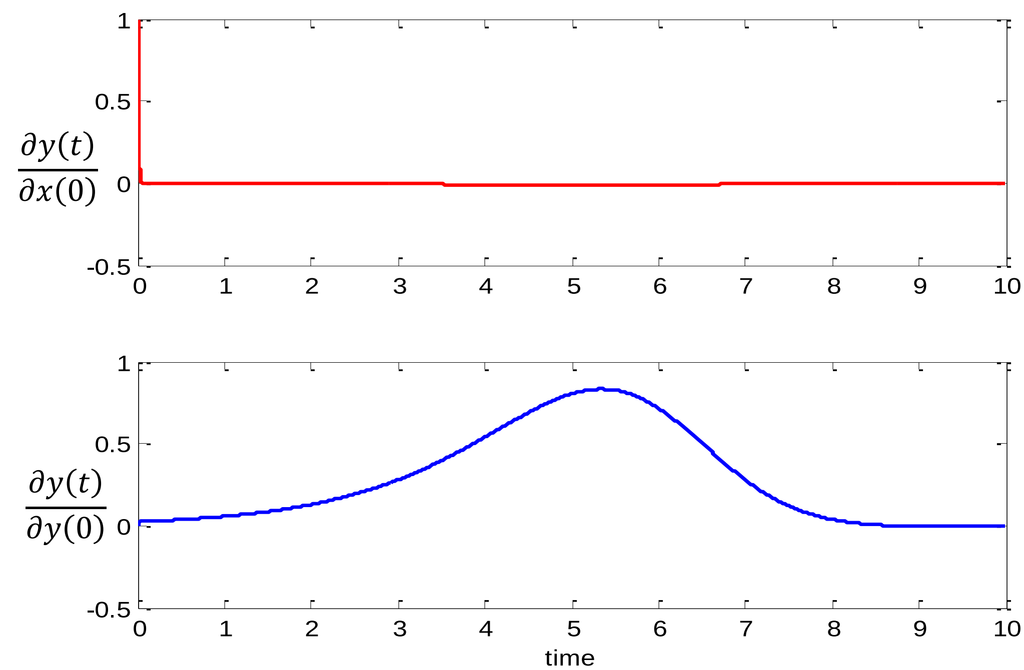


Figure 2.17. The sensitivity to initial conditions of $\boldsymbol{y}$

- **Sensitivity matrix for parameters $X = \begin{pmatrix} \boldsymbol{x(t, x(0), y(0), \alpha)}, \\ \boldsymbol{y(t, x(0), y(0), \alpha)} \end{pmatrix}$**

$$\frac{d}{dt}\begin{bmatrix} \frac{\partial x(t)}{\partial a} & \frac{\partial x(t)}{\partial \mu_c} & \frac{\partial x(t)}{\partial \eta_c} \\ \frac{\partial y(t)}{\partial a} & \frac{\partial y(t)}{\partial \mu_c} & \frac{\partial y(t)}{\partial \mu_c} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} \times \begin{bmatrix} \frac{\partial x(t)}{\partial a} & \frac{\partial x(t)}{\partial \mu_c} & \frac{\partial x(t)}{\partial \eta_c} \\ \frac{\partial y(t)}{\partial a} & \frac{\partial y(t)}{\partial \mu_c} & \frac{\partial y(t)}{\partial \mu_c} \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial a} & \frac{\partial f_1}{\partial \mu_c} & \frac{\partial f_1}{\partial \eta_c} \\ \frac{\partial f_2}{\partial a} & \frac{\partial f_2}{\partial \mu_c} & \frac{\partial y f_2}{\partial \mu_c} \end{bmatrix} \tag{2.6.8}$$

$$\frac{d}{dt} D_\alpha(X) = D_X(t) * D_\alpha(X) + D_\alpha(t) \tag{2.6.9}$$

where

$$D_\alpha\big(X(0)\big) = 0 \tag{2.6.10}$$

$$D_X(t) =$$

$$\begin{bmatrix} -1 - a - \eta_c^2(x-y)^2 - \eta_c^2 x(2x-2y) & \eta_c^2 x(2x-2y) \\ -\eta_c^2 y(2x-2y) & \eta_c^2 y(2x-2y) - \eta_c^2(x-y)^2 - 1 \end{bmatrix} \tag{2.6.11}$$

$$D_\alpha(t) = \begin{bmatrix} 1-x & 0 & -2\eta_c x(x-y)^2 \\ 0 & 1 & -2\eta_c y(x-y)^2 \end{bmatrix} \tag{2.6.12}$$

With $\varepsilon = 0.01$, $\mu_c = 1$ and $\eta_c^2 \simeq 7.5$, $x(0) = 0.9378$, $y(0) = 0.8123$, Figure 2.18 and 2.19 show the sensitivity to parameters of the model.

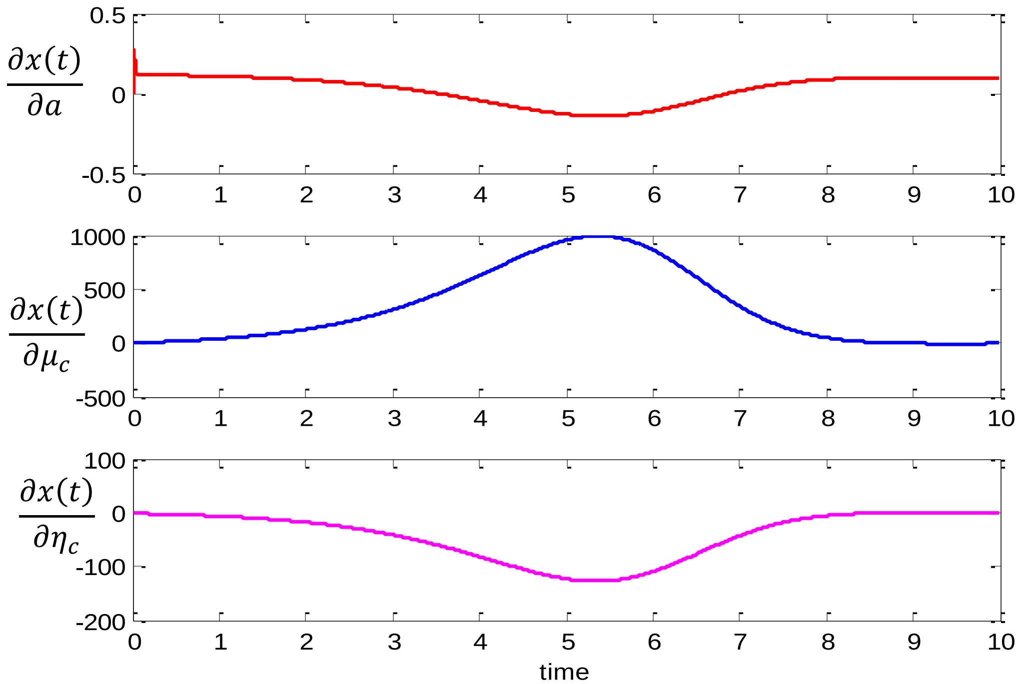


Figure 2.18. The sensitivity to parameters of $\boldsymbol{x}$

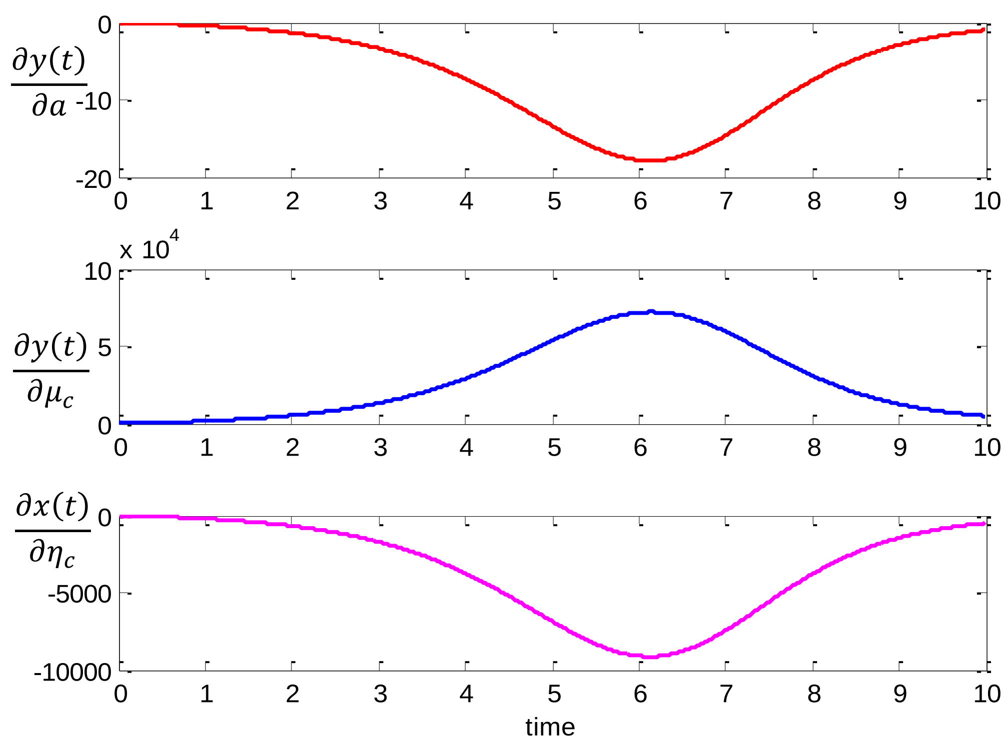


Figure 2.19. The sensitivity to parameters of $\boldsymbol{y}$

b) Van Veen's model

The equations of the model

$$\varepsilon\dot{x} = (1-x) - \varepsilon(1+\eta_v|x-y|)x$$
$$\dot{y} = \mu_v - (1+\eta_v|x-y|)y \tag{2.6.13}$$

is equivalent with

$$\dot{x} = \frac{1-x}{\varepsilon} - (1+\eta_v|x-y|)x$$
$$\dot{y} = \mu_v - (1+\eta_v|x-y|)y \tag{2.6.14}$$

Let $a = \frac{1}{\varepsilon}$, parameters $\alpha = (a, \mu_v, \eta_v)$, and variables $X = (x, y)$

$$\dot{x} = a(1-x) - (1+\eta_v|x-y|)x = f_1(x,y,\alpha)$$
$$\dot{y} = \mu_v - (1+\eta_v|x-y|)y = f_2(x,y,\alpha) \tag{2.6.15}$$

- **Sensitivity matrix for initial conditions** $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x(t, x(0), y(0), \alpha)}, \\ \boldsymbol{y(t, x(0), y(0), \alpha)} \end{pmatrix}$

$$\frac{d}{dt}\begin{bmatrix} \frac{\partial x(t)}{\partial x(0)} & \frac{\partial x(t)}{\partial y(0)} \\ \frac{\partial y(t)}{\partial x(0)} & \frac{\partial y(t)}{\partial y(0)} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} \times \begin{bmatrix} \frac{\partial x(t)}{\partial x(0)} & \frac{\partial x(t)}{\partial y(0)} \\ \frac{\partial y(t)}{\partial x(0)} & \frac{\partial y(t)}{\partial y(0)} \end{bmatrix} \tag{2.6.16}$$

$$\frac{d}{dt}D_{X(0)}(X) = D_X(t) * D_{X(0)}(X) \tag{2.6.17}$$

where

$$D_{X(0)}\big(X(0)\big) = \begin{bmatrix} \frac{\partial x(0)}{\partial x(0)} & \frac{\partial x(0)}{\partial y(0)} \\ \frac{\partial y(0)}{\partial x(0)} & \frac{\partial y(0)}{\partial y(0)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \tag{2.6.18}$$

$$D_X(t) = \begin{bmatrix} -1 - a - \eta_v|x-y| - \eta_v * x * sign(x-y) & \eta_v * x * sign(x-y) \\ -\eta_v * y * sign(x-y) & \eta_v * y * sign(x-y) - \eta_v|x-y| - 1 \end{bmatrix} \tag{2.6.19}$$

With $\varepsilon = 0.1,\ \eta_v = 216.67,\ and\ \mu_v = 3,\ x(0) = 0.6929,\ y(0) = 0.6771,$ Figure 2.20 and 2.21 show the sensitivity to initial conditions of the model.

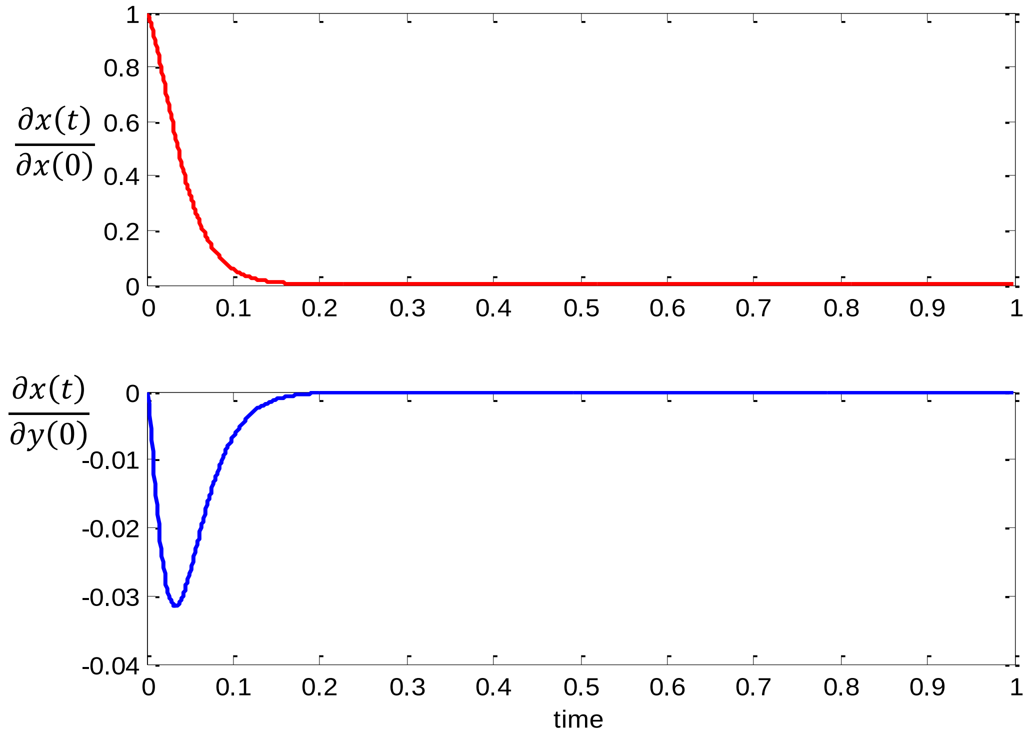


Figure 2.20. The sensitivity to initial conditions of $\boldsymbol{x}$

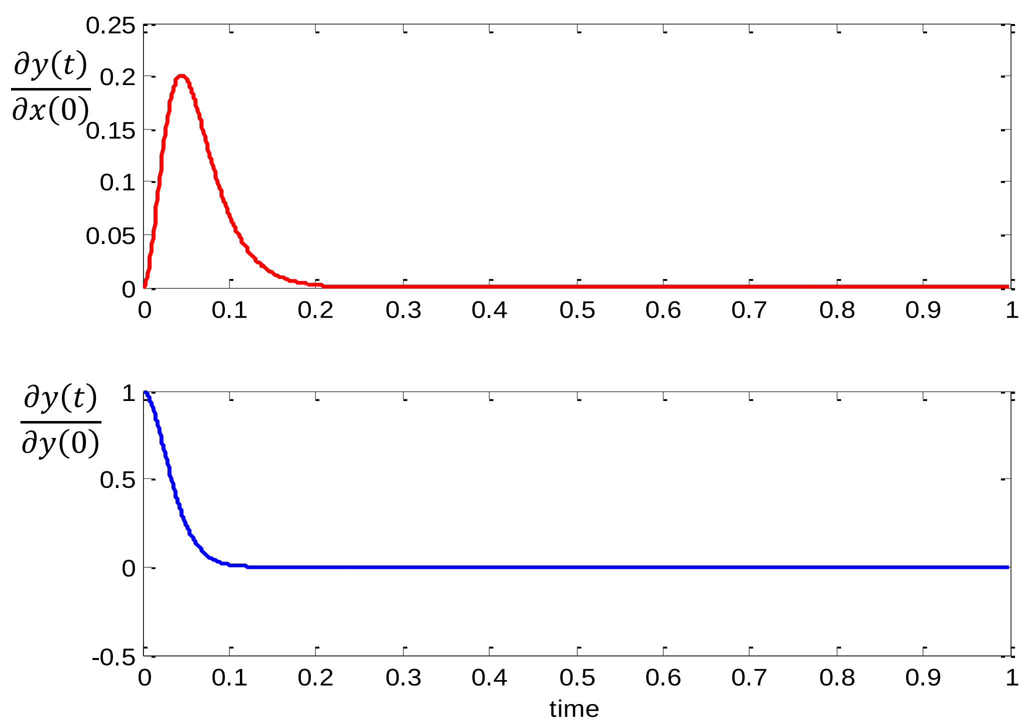


Figure 2.21. The sensitivity to initial conditions of $\boldsymbol{y}$

- **Sensitivity matrix for parameters $X = \begin{pmatrix} \boldsymbol{x(t, x(0), y(0), \alpha)}, \\ \boldsymbol{y(t, x(0), y(0), \alpha)} \end{pmatrix}$**

$$\frac{d}{dt}\begin{bmatrix} \frac{\partial x(t)}{\partial a} & \frac{\partial x(t)}{\partial \mu_v} & \frac{\partial x(t)}{\partial \eta_v} \\ \frac{\partial y(t)}{\partial a} & \frac{\partial y(t)}{\partial \mu_v} & \frac{\partial y(t)}{\partial \mu_v} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} \times \begin{bmatrix} \frac{\partial x(t)}{\partial a} & \frac{\partial x(t)}{\partial \mu_v} & \frac{\partial x(t)}{\partial \eta_v} \\ \frac{\partial y(t)}{\partial a} & \frac{\partial y(t)}{\partial \mu_v} & \frac{\partial y(t)}{\partial \mu_v} \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial a} & \frac{\partial f_1}{\partial \mu_v} & \frac{\partial f_1}{\partial \eta_v} \\ \frac{\partial f_2}{\partial a} & \frac{\partial f_2}{\partial \mu_v} & \frac{\partial y f_2}{\partial \mu_v} \end{bmatrix} \tag{2.6.20}$$

$$\frac{d}{dt} D_\alpha(X) = D_X(t) * D_\alpha(X) + D_\alpha(t) \tag{2.6.21}$$

where

$$D_\alpha\big(X(0)\big) = 0 \tag{2.6.22}$$

$D_X(t) =$

$$\begin{bmatrix} -1 - a - \eta_v|x - y| - \eta_v * x * sign(x - y) & \eta_v * x * sign(x - y) \\ -\eta_v * y * sign(x - y) & \eta_v * y * sign(x - y) - \eta_v|x - y| - 1 \end{bmatrix} \tag{2.6.23}$$

$$D_\alpha(t) = \begin{bmatrix} 1 - x & 0 & -x * |x - y| \\ 0 & 1 & -y * |x - y| \end{bmatrix} \tag{2.6.24}$$

With $\varepsilon = 0.1,\ \eta_v = 216.67,\ and\ \mu_v = 3,\ x(0) = 0.6929,\ y(0) = 0.6771,$ Figure 2.22 and 2.23 show the sensitivity to parameters of the model.

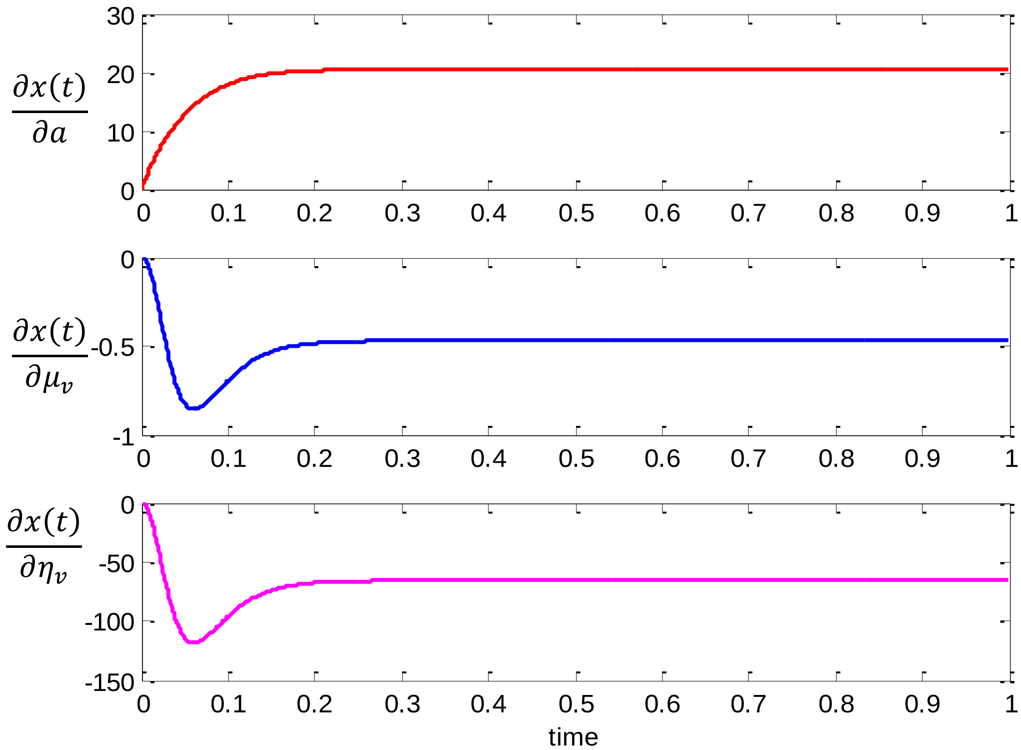


Figure 2.22. The sensitivity to parameters of $\boldsymbol{x}$

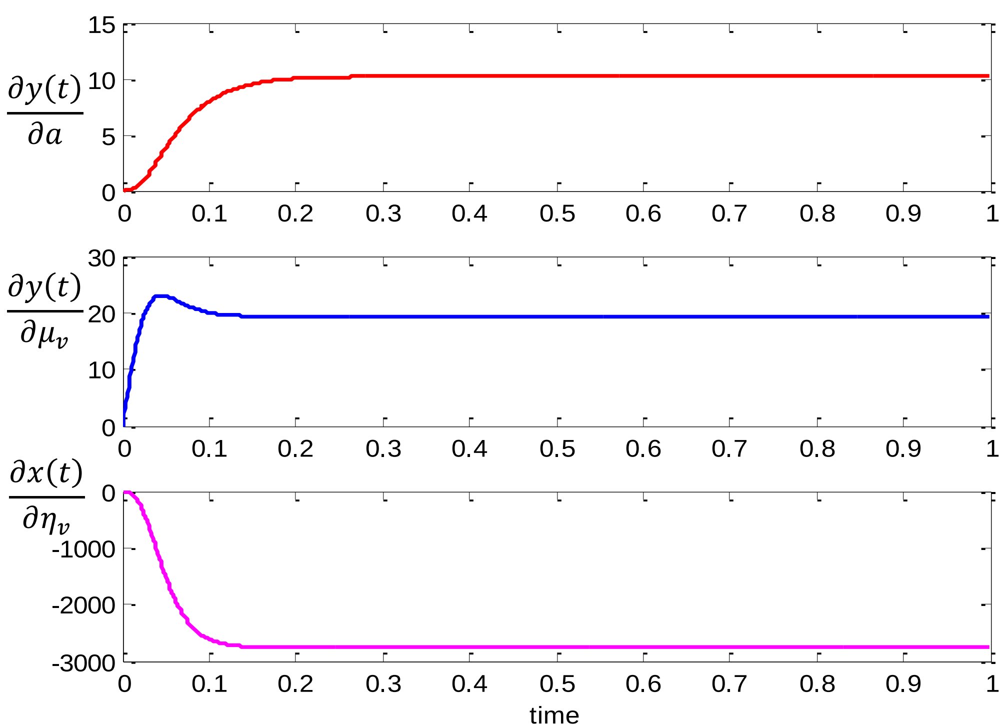


Figure 2.23. The sensitivity to parameters of $\boldsymbol{y}$

2.7. Conclusions

In this chapter, we have studied two-box model of thermohaline circulation in several ways. First, we consider the equilibria and their stability. In most models, there exists one stable equilibrium or three equilibria, of which one is unstable, one is thermo-driven stable, and the last one is haline-driven stable. We then examine the bifurcation diagram in which we see the possible values of equilibria of a system as a function of a bifurcation parameters in the syste. Last, we look at the sensitivity analysis of Cessi and Van Veen' model with respect to initial conditions and parameters.

Chapter 3

Two-box oscillation

3.1. A simple heat-salt oscillator

Welander (1982) described an interesting variant of a 2-box model arranged vertically one on the top of the other instead of the two boxes arranged side by side as in Stommel's model, and refer to Figure 3.1.

In this model the ocean is vertically divided into two layers where the top layer is assumed to be well mixed that interacts both with the atmosphere above and the well stratified bottom deep layer blow. It is assumed that the temperature T_o and salinity S_o of the deep layer are fixed constants. The temperature T of the top layer reacts to the temperature difference between the air temperature T_A and T and the salinity of the top layer changes with respect to the salinity equivalent S_A resulting from fresh water input. The transfer of heat and salt between the top shallow layer and the bottom deep layer is primarily induced by the density difference $\Delta\rho = \rho - \rho_o$ where ρ is the density of the top layer and ρ_o is that of the bottom layer. The first order approximation is

$$\Delta\rho = -\alpha(T - T_o) + \beta(S - S_o) \tag{3.1.1}$$

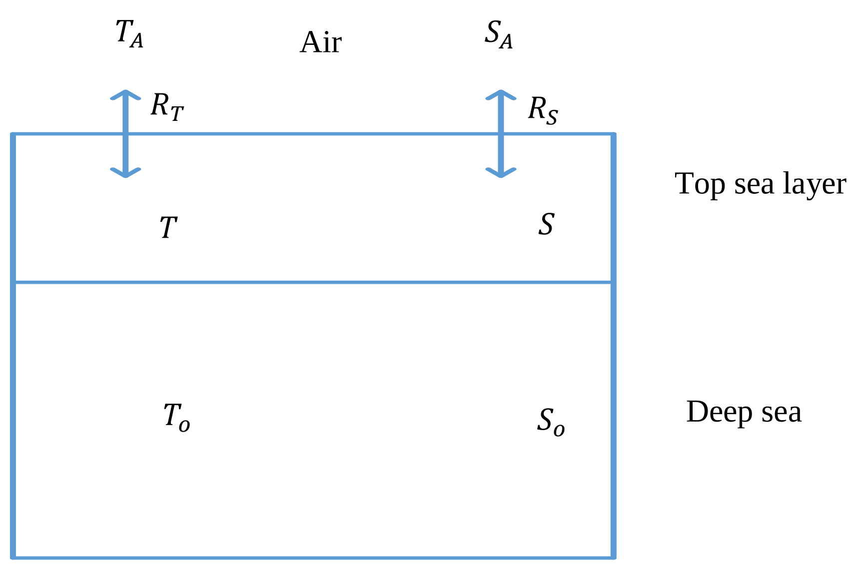


Figure 3.1. Welander's heat-salt oscillator

The model equations are then given by

$$\frac{dT}{dt} = R_T(T_A - T) - k(\Delta\rho)T$$
$$\frac{dS}{dt} = R_S(S_A - S) - k(\Delta\rho)S \tag{3.1.2}$$

where $k(\Delta\rho)$ is an increasing function of $\Delta\rho$. In the so-called flip model, $k(\Delta\rho)$ is chosen as

$$k(\Delta\rho) = \begin{cases} k_o \; if \; \Delta\rho \leq -\varepsilon \\ k_1 \; if \; \Delta\rho > -\varepsilon \end{cases} \tag{3.1.3}$$

where $k_o \ll k_1$.

This flip-flop model can in principle exhibit two steady states:

$$\bar{T_1} = \frac{R_T}{R_T + k_o} T_A \approx T_A$$
$$\bar{S_1} = \frac{R_T}{R_T + k_o} S_A \approx S_A \tag{3.1.4}$$

when $\Delta\rho \leq -\varepsilon$, since k_o is small and

$$\bar{T_2} = \frac{R_T}{R_T + k_1} T_A \approx \frac{R_T}{k_1} T_A$$
$$\bar{S_2} = \frac{R_T}{R_T + k_1} S_A \approx \frac{R_T}{k_1} S_A \tag{3.1.5}$$

when $\Delta\rho > -\varepsilon$. However, if the parameters are such that

$$-\alpha\bar{T_2} + \beta\bar{S_2} < -\varepsilon < -\alpha\bar{T_1} + \beta\bar{S_1}$$

then the two steady states can never be reached.

To demonstrate the oscillation (see Figure 3.2), we consider a numerical example. The parameter value chosen are $T_o = S_o = \rho_o = 0$, $\frac{\alpha T_A}{\beta S_A} = 0.2$, $\frac{R_T}{R_S} = 10$, $\frac{k_o}{R_T} = 0$, $\frac{k_1}{R_T} = 5$, $\frac{\varepsilon}{\beta S_A} = -0.01$. Using non-dimensional variables $T^* = \frac{T}{T_A}$, $S^* = \frac{S}{S_A}$, $\Delta\rho^* = \frac{\Delta\rho}{\beta S_A}$, and $t^* = R_T t$, equations (3.1.1) and (3.1.2) become

$$\Delta\rho = -0.2T^* + S^* \tag{3.1.6}$$

$$\frac{dT^*}{dt} = 1 - T^* - \binom{0}{5} T^*$$
$$\frac{dS^*}{dt} = 0.1(1 - S^*) - \binom{0}{5} S^* \tag{3.1.7}$$

where $k^*(\rho^*) = \begin{cases} 0 \; if \; \rho^* \leq -0.01 \\ 5 \; if \; \rho^* > -0.01 \end{cases}$

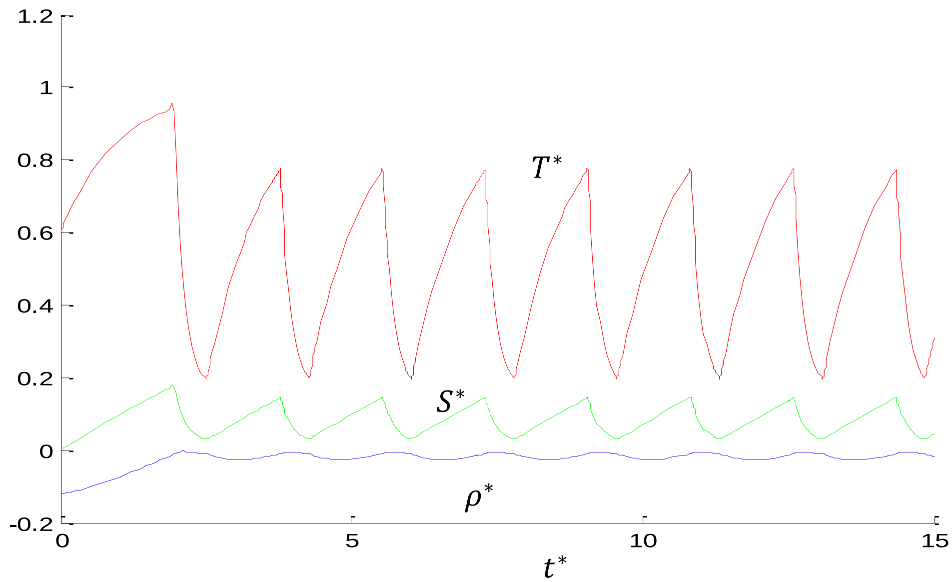


Figure 3.2. The solution $\boldsymbol{T}^*(\boldsymbol{t}^*)$, $\boldsymbol{S}^*(\boldsymbol{t}^*)$, and $\boldsymbol{\rho}^*(\boldsymbol{t}^*)$ in the numerical example

3.2. A simple pure water oscillator

A fresh water lake can be conveniently modeled by a system of two well mixed boxes arranged vertically as shown in the Figure 3.3. The shallow surface box of volume V_s (m^3) is at a temperature T_s (°C). T_s varies due to the thermal interaction – Newtonian heating/cooling, between the atmosphere above at temperature $T_a(\lesseqgtr T_s)$ and due to turbulent mixing with the deep box resulting from the static instability due to increased density of water in the surface box arising from the heat exchange with the atmosphere. The deep box is of volume V_d (m^3) $(\gg V_s)$ and is assumed to be at a constant temperature T_d. The potential for the vertical oscillation in this system is largely a result of the nonlinear dependence of density $\rho(T)$ of the pure water on the temperature T as described in the Appendix C.

Accordingly the dynamics of evolution of T_s can be captured by the following ordinary differential equation

$$V_s \frac{d(T_s - T_d)}{dt} = C_T(T_a - T_s) - \gamma f(\Delta\rho)(T_s - T_d) \qquad (3.2.1)$$

where (it is assumed for definiteness that $T_a > T_s$) the first term on the right hand side accounts for the heat exchange with atmosphere by Newtonian heating process and the constant C_T has units $m^3 sec^{-1}$. The second term on the right hand side of (3.2.1) represents the heat exchange resulting from the flow where $\gamma f(\Delta\rho)$ is the measure of the flow in $m^3 sec^{-1}$, the normalized density difference $\Delta\rho = \Delta\rho(T_s)$ for simplicity in notation, the constant γ $(m^3 sec^{-1})$ is known as the hydraulic constant and $f(\Delta\rho)$ is a monotonic non-decreasing function of $\Delta\rho$. Refer to Appendix C for characterization of $\Delta\rho$.

Dividing both sides by V_s, we get

$$\frac{d(T_s - T_d)}{dt} = k_T(T_a - T_s) - \eta f(\Delta\rho)(T_s - T_d) \tag{3.2.2}$$

where $k_T = \frac{C_T}{V_s}$ (sec^{-1}) is the reciprocal of the temperature relaxation time for the surface box and $\eta = \frac{\gamma}{V_s}$ (sec^{-1}).

To capture the generic behavior of this model, we now non-dimensionalize the temperature and time as follows

$$x = \frac{T_s - T_d}{T_a - T_d} \qquad \text{and} \qquad s = \frac{t}{(1/k_T)} \tag{3.2.3}$$

Substituting (3.2.3) into (3.2.2), since T_d is a constant, the latter becomes

$$\frac{dx}{ds} = (1 - x) - \bar{\eta} f\big(\Delta\bar{\rho}(x)\big)x \tag{3.2.4}$$

where

$$\bar{\eta} = \frac{\eta}{k_T} \tag{3.2.5}$$

is a non-dimensional parameter, and $\Delta\bar{\rho}(x)$ is the fractional density difference expressed in term of the nondimensional temperature x . Refer to Appendix C for the transformation of the dependence of $\Delta\rho$ from T_s to x. By a clever choice of the function $\eta f(\Delta\rho)$, we can obtain quite a variety of models. In this note, following Welander (1982), we consider the so called the flip or threshold model.

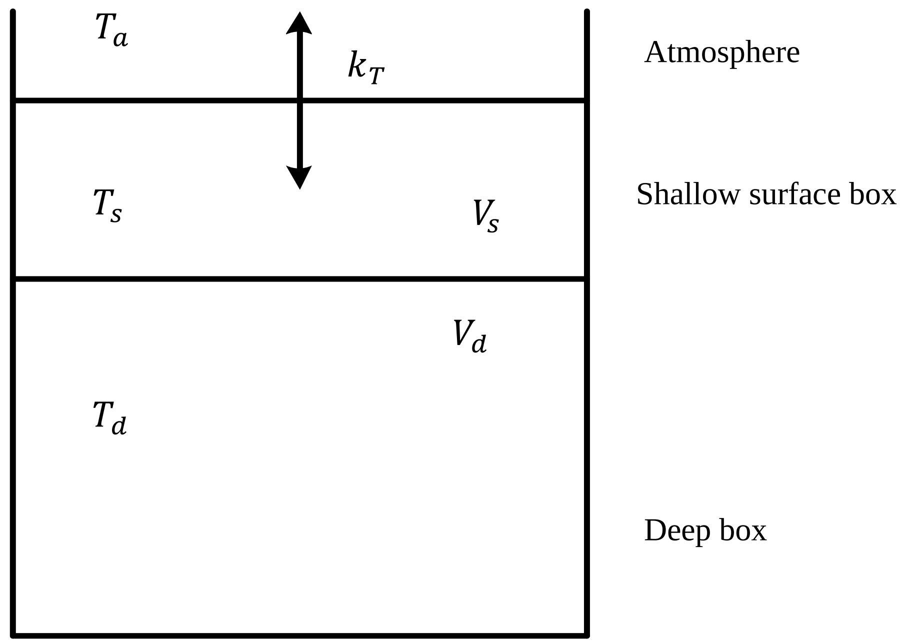


Figure 3.3. A 2-box model for the fresh water lake. The shallow surface box has volume $\boldsymbol{V_s}$ $(\boldsymbol{m^3})$ and deep box has volume $\boldsymbol{V_d}$ $(\boldsymbol{m^3})$ with $\boldsymbol{V_d} \gg \boldsymbol{V_s}$. It is assumed that $\boldsymbol{T_d = 2.0}$℃.

The flip model: In this model, we define

$$\bar{\eta} f\big(\Delta\bar{\rho}(x)\big) = \begin{cases} k_0 & if \quad \Delta\bar{\rho}(x) \le \varepsilon \\ k_1 & if \quad \Delta\bar{\rho}(x) > \varepsilon \end{cases} \tag{3.2.6}$$

where $\varepsilon \ge 0$ is a small pre-specified threshold parameter and $k_0 \ll k_1$ are two constants that reflect the vertical turbulent heat-transfer resulting from the static instability induced by increasing density of the surface layer with temperature. An example of the range of temperature x where $\Delta\bar{\rho}(x) - \varepsilon > 0$ for $\varepsilon = 10^{-5}$ is given in Figure 3.4.

Define an indicator function

$$I_\varepsilon\big(\Delta\bar{\rho}(x)\big) = \begin{cases} 0 & if \quad \Delta\bar{\rho}(x) \le \varepsilon \\ 1 & if \quad \Delta\bar{\rho}(x) > \varepsilon \end{cases} \tag{3.2.7}$$

Combining (3.2.6) and (3.2.7), the model equation (3.2.4) can be written as

$$\frac{dx}{ds} = (1 - x) - k_0 x - (k_1 - k_0) x I_\varepsilon\big(\Delta\bar{\rho}(x)\big) \tag{3.2.8}$$

or equivalently as

$$\frac{dx}{ds} = \begin{cases} f_1(x) = 1 - (1 + k_0)x & if \quad \Delta\bar{\rho}(x) - \varepsilon \le 0 \\ f_2(x) = 1 - (1 + k_1)x & if \quad \Delta\bar{\rho}(x) - \varepsilon > 0 \end{cases} \tag{3.2.9}$$

Clearly, this model equation has discontinuous right hand side (DRHS) and the conditions for the existence and uniqueness of the solution are not satisfied. Filippov (1988) presents an interesting extension of the notion of the solution of the ODE with DRHS, refer to Appendix D for a summary of Fillipov's framework. There are essentially two ways to proceed. First is to apply Filippov's theory to quantify the flow field near the points of discontinuity which guarantees the existence and uniqueness of the solution. A second and simpler approach is to invoke to the regularization method where we replace DRHS with a continuous approximation. In this note, while we follow the Fillipov's approach, for completeness, we also indicate a simple regularization method.

To this end, recall, we can approximate a unit step function

$$I(y) = \begin{cases} 0 & if \quad y \le 0 \\ 1 & if \quad y > 0 \end{cases} \tag{3.2.10}$$

by a continuous function of the form

$$H_\beta(y) = \frac{1}{2}[1 + tanh(\beta y)] \tag{3.2.11}$$

where

$$tanh(y) = \frac{e^y - e^{-y}}{e^y + e^{-y}} \tag{3.2.12}$$

It is well known that as β increases, the quality of approximation of the discontinuous function $I(y)$ in (3.2.10) by continuous function $H_\beta(y)$ in (3.2.11) steadily improves. Since $\Delta\bar{\rho}(x) - \varepsilon$ is of the order 10^{-4}, in the following, for better accuracy, we set $\beta = 10^6$.

Consequently, we can regularize the indicator function $I_\varepsilon\big(\Delta\bar{\rho}(x)\big)$ in (3.2.9) by replacing it by $H_\beta(\Delta\bar{\rho}(x) - \varepsilon)$. Thus, the regularized version of the flip model is given by

$$\frac{dx}{ds} = f(x) \tag{3.2.13}$$

where

$$f(x) = (1 - x) - k_0 x - (k_1 - k_0) x H_\beta(\Delta\bar{\rho}(x) - \varepsilon) \tag{3.2.14}$$

describes the flow field of this model dynamics.

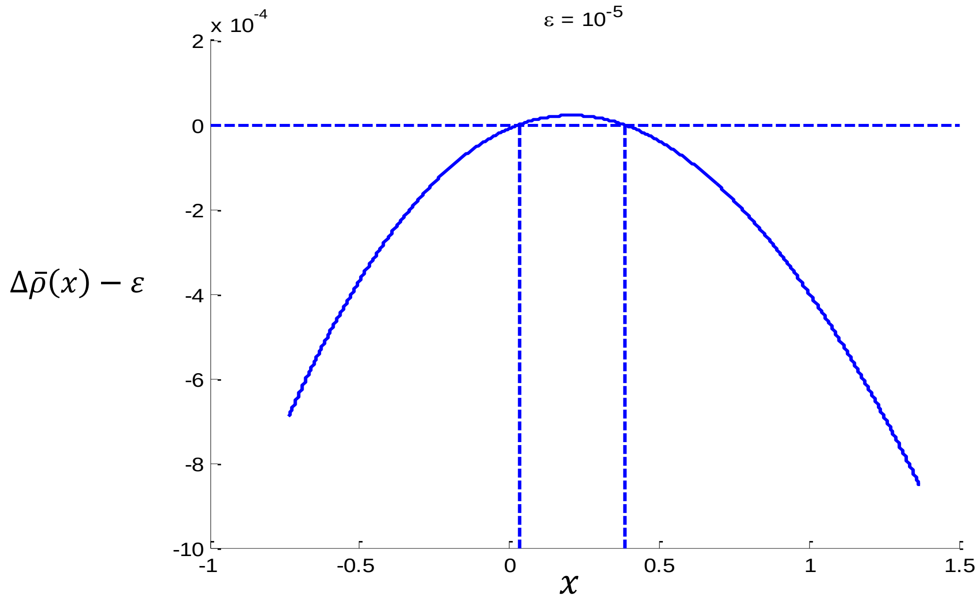


Figure 3.4. Plot of $\Delta\bar{\boldsymbol{\rho}}(\boldsymbol{x}) - \boldsymbol{\varepsilon}$ for $\boldsymbol{\varepsilon} = \mathbf{10^{-5}}$

$\Delta\bar{\rho}(x) - \varepsilon > 0$ when $x_1 \approx 0.0352 < x < x_2 \approx 0.3850$

In this section we provide a complete catalog of the behavior of the model solution with respect to the choice of values of three parameters: ε, k_0 and k_1 in (3.2.9) and (3.2.14).

a) Properties of $\Delta\bar{\boldsymbol{\rho}}(\boldsymbol{x})$ and effect of $\boldsymbol{\varepsilon}$

From the definition of $\Delta\bar{\rho}(x)$ in (C.9) in the Appendix C, the following properties readily follow:

P1) $\Delta\bar{\rho}(x)$ attains its maximum value of 3.2087×10^{-5} at $x = x_0 = 0.2086$

P2) $\Delta\bar{\rho}(0) = 0$ and $\Delta\bar{\rho}(1) = -3.8778 \times 10^{-4}$.

P3) Setting $\varepsilon = 10^{-5}$, and solving the polynomial equation

$$\Delta\bar{\rho}(x) = \varepsilon \tag{3.2.15}$$

with real coefficients, we find that there are two real solutions at $x_1 \approx 0.0352$ and $x_2 \approx 0.3850$. Refer to Figure 3.4 for an illustration.

P4) $\Delta\bar{\rho}(x) - \varepsilon \geq 0$ for all $x_1 \leq x \leq x_2$, and $\Delta\bar{\rho}(x) - \varepsilon < 0$, otherwise.

P5) The separation between x_1 and x_2 measured by the distance $|x_2 - x_1|$ is an decreasing function of ε.

In the subsequent analysis, we fix $\varepsilon = 10^{-5}$. Stated in other words, the points of discontinuity in the flow field $f(x)$ in (3.2.9) occurs at x_1 and x_2 whose numerical value depends on the threshold parameter $\varepsilon \geq 0$. Further, since (3.2.15) is a polynomial with empirically determined real coefficients, the solution x_1 and x_2 of (3.2.15) in general are not rational number, hence cannot be represented exactly by a terminating decimal expansion. Hence, for the flow field $f(x)$ of the model (3.2.9) is given by $x_1 \approx 0.0352$ and $x_2 \approx 0.3850$.

$$f(x) = \begin{cases} f_1(x) = 1 - (1 + k_0)x & for \quad x \leq x_1 \; and \; x \geq x_2 \\ f_2(x) = 1 - (1 + k_1)x & for \qquad x_1 < x < x_2 \end{cases} \tag{3.2.16}$$

Against this background we now move on to analyzing the impact of the variation of parameters k_0 and k_1 on the flow field of the model dynamics.

b) Analysis of equilibria and their stability

Analysis of the impact of the two parameters k_0 and k_1 on the model dynamics is largely a consequence of two facts: (F1) $0 \leq k_0 < k_1$ and (F2) both $f_1(x)$ and $f_2(x)$ in (3.2.16) are affine functions with negative slopes $-(1 + k_0)$ and $-(1 + k_1)$ respectively. To fix the bearings, it is useful to refer to the plot of $f(x)$ in (3.2.16) with $k_0 = 0$ and $k_1 = 35$ in Figure 3.5. From the two facts listed above and a glance at the Figure 3.5, we can easily verify the following relations between the values of $f_1(x)$ and $f_2(x)$ at the points of discontinuities, $x_1 \approx 0.0352$ and $x_2 \approx 0.3850$.

Since $0 \le k_0 < k_1$, it follows that (R1) $f_1(x_1) > f_2(x_1)$ and (R2) $f_1(x_2) > f_2(x_2)$. Since $f_1(x)$ and $f_2(x)$ are affine with negative slopes, from $x_1 < x_2$, it also follows that (R3) $f_1(x_1) > f_1(x_2)$ and (R4) $f_2(x_1) > f_2(x_2)$.

Given these four relations, there are eight other possibilities for the values of $f_i(x_1)$ and $f_i(x_2)$ for $i = 1, 2$. A complete listing of these eight possibilities and the conditions for their existence are given in Table 3.1.

Combining the four natural relations in (R1) – (R4) and eight conditions C1 – C8 in Table 3.1, we now enlist a set of all six mutually exclusive combinations that provide a complete characterization of the model dynamics corresponding to various combinations parameter ranges in Table 3.2. A partition of the 2-D parameter space for $k_1 \ge k_0 \ge 0$ induced by the six cases in Table 3.2 is given in Figure 3.6. The completeness of the six cases listed in Table 3.2 follows immediately from the fact that the union of the six partitions P_1 through P_6 is equal to the upper half $(k_1 \ge k_0)$ of the positive quadrant of the 2D plane.

For completeness, we present analogous results for the regularized version of $f(x)$ in (3.2.14) with $\beta = 10^6$ are given in Table 3.4. It can be verified that entries in Table 3.4 converge to the corresponding entries in Table 3.3 as β increases.

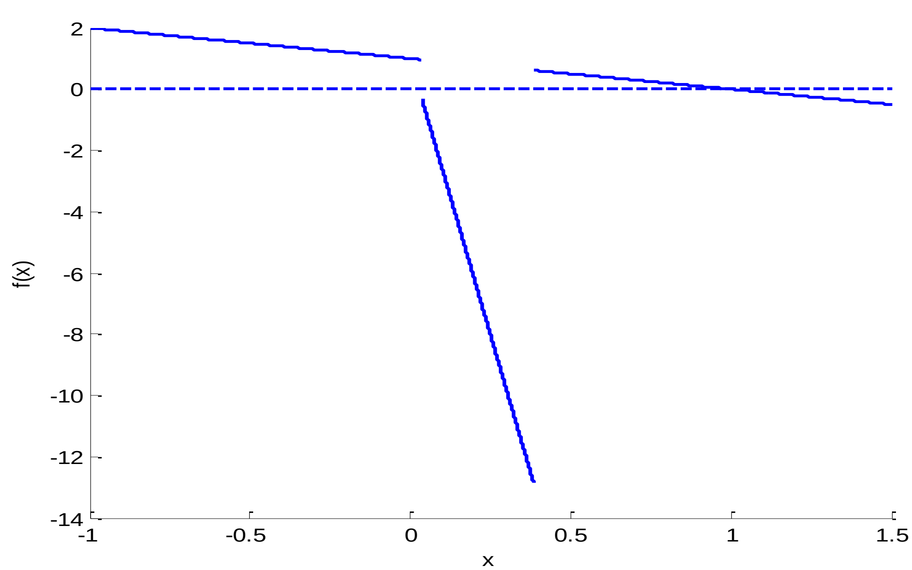


Figure 3.5. Plot of the $f(x)$ in (3.2.16) with $k_0 = 0$ and $k_1 = 35$.

Table 3.1. Conditions on the parameters k_0 and k_1 for the sign definiteness of $f_i(x_j)$ for $i = 1, 2$ and $j = 1, 2$ where $0 < x_1 < x_2$. Let $b = \frac{1-x_1}{x_1}$ and $a = \frac{1-x_2}{x_2}$. Given that $x_1 \approx 0.0352$ and $x_2 \approx 0.3850$, hence $a \approx 1.597$ and $b \approx 27.409$

Number	Property	Condition
C1	$f_1(x_1) > 0$	$k_0 < b$
C2	$f_1(x_1) < 0$	$k_0 > b$
C3	$f_2(x_1) > 0$	$k_1 < b$
C4	$f_2(x_1) < 0$	$k_1 > b$
C5	$f_1(x_2) > 0$	$k_0 < a$
C6	$f_1(x_2) < 0$	$k_0 > a$
C7	$f_2(x_2) > 0$	$k_1 < a$
C8	$f_2(x_2) < 0$	$k_1 > a$

Table 3.2. A listing of the set of all six mutually exclusive cases: $a = \frac{1-x_2}{x_2} \approx 1.597$, $b = \frac{1-x_1}{x_1} \approx 27.409$; $f_1 = f_1(x)$, $f_2 = f_2(x)$.

Case number	Condition at x_1	Condition at x_2	Condition on k_0 and k_1 defining the partition P_i
1	$f_1 > 0 > f_2$	$f_1 > 0 > f_2$	$P_1: 0 \leq k_0 < a < b < k_1$
2	$f_1 > f_2 > 0$	$f_1 > 0 > f_2$	$P_2: 0 \leq k_0 < a < k_1 < b$
3	$f_1 > f_2 > 0$	$f_1 > f_2 > 0$	$P_3: 0 \leq k_0 < k_1 < a < b$
4	$0 > f_1 > f_2$	$0 > f_1 > f_2$	$P_4: a < b < k_0 < k_1$
5	$f_1 > f_2 > 0$	$0 > f_1 > f_2$	$P_5: a < k_0 < k_1 < b$
6	$f_1 > 0 > f_2$	$0 > f_1 > f_2$	$P_6: a < k_0 < b < k_1$

A graphical view of the phase plot – plot of $f(x)$ in (3.2.16) vs. x along with the display of equilibria for the six cases in Table 3.3 are given in Figures 3.7 through 3.12 respectively.

It turns out that the set of all equilibria E_1 through E_{10} in Table 3.3 can be classified into two groups. First is the set of all six stable equilibria $S_1 =$

$\{E_3,\ E_4,\ E_6,\ E_7,\ E_8,\ E_9\}$ that lie at a point of continuity of either $f_1(x)$ or $f_2(x)$ in (3.2.16). The second consisting of the rest of the four $S_2 = \{E_1,\ E_2,\ E_5,\ E_{10}\}$ that lie at the points of discontinuities x_1 and x_2 of $f(x)$ in (3.2.16).

According to the classical theory of ODE, a stable equilibrium of a smooth vector field is a local attractor in the sense that the solutions starting from its basin of attraction will eventually converge to this stable equilibrium point. The last column of Table 3.3 contains the basin of attraction for all the six stable equlibria in the set S_1. As an illustration, for any initial condition satisfying $x_2 < x(0) < \infty$, the solution $x(t)$ of $\dot{x} = f_1(x) = 1 - (1 + k_0)x$ is such that $\lim_{t\to\infty} x(t) = \frac{1}{1+k_0}$ which is the equilibrium E_3 in Table 3.3. Similar convergence results hold for the rest of the five stable equilibria in S_1.

Consequently we only need to apply Fillipov's theory in Appendix D only to the four cases in the set S_2. We consider two cases.

A. Behavior of (3.2.16) around x_1:

Referring to Appendix D, let $h(x) = x - x_1$, and hence $h_x(x) = 1$. Then (3.2.16) becomes

$$\dot{x} = \begin{cases} f_1(x) = 1 - (1 + k_0)x & if \quad h(x) < 0 \\ f_2(x) = 1 - (1 + k_1)x & if \quad h(x) > 0 \end{cases} \tag{3.2.17}$$

where $f_1(x_1) \neq f_2(x_1)$ since $k_0 < k_1$.

From (D.9) – (D.10), since $h_x(x) = 1$, it follows that

$$\beta(x_1) = \frac{f_1(x_1) + f_2(x_1)}{f_1(x_1) - f_2(x_1)} \tag{3.2.18}$$

and

$$f_U(x_1) = 0 \tag{3.2.19}$$

Similarly, from (D.11) – (D.12), it follows that

$$\alpha(x_1) = \frac{f_1(x_1)}{f_1(x_1) - f_2(x_1)} \tag{3.2.20}$$

and

$$f_F(x_1) = 0 \tag{3.2.21}$$

Table 3.3. Distribution and stability of equilibria for $\boldsymbol{f}(\boldsymbol{x})$ in (3.2.16)

Recall that $x_1 \approx 0.0352$ and $x_2 \approx 0.3850$, a_i^* denotes a large positive real number.

Case number	Parameter values	Equilibria	Eigenvalue of the Jacobian of $f(x)$ at the equilibria	Stability	Basin at attraction
1	$k_0 = 0$ $k_1 = 35$	$E_1 = x_1$	$-a_1^*$	E_1 – Stable – chatter	
		$E_2 = x_2$	a_2^*	E_2 – Unstable	
		$E_3 = \dfrac{1}{1+k_0}$	-1.0	E_3 – Stable	(x_2, ∞)
2	$k_0 = 0$ $k_1 = 10$	$E_4 = \dfrac{1}{1+k_1}$	-11.0	E_1 – Stable	(x_1, x_2)
		$E_5 = x_2$	a_3^*	E_2 – Unstable	
		$E_6 = \dfrac{1}{1+k_0}$	-1.0	E_3 – Stable	(x_2, ∞)
3	$k_0 = 0$ $k_1 = 1$	$E_7 = \dfrac{1}{1+k_0}$	-1.0	E_1 – Stable	(x_2, ∞)
4	$k_0 = 30$ $k_1 = 35$	$E_8 = \dfrac{1}{1+k_0}$	-41.12	E_1 – Stable	$(-\infty, x_1)$
5	$k_0 = 5$ $k_1 = 20$	$E_9 = \dfrac{1}{1+k_1}$	-21.67	E_1 – Stable	(x_1, x_2)
6	$k_0 = 5$ $k_1 = 35$	$E_{10} = x_1$	$-a_4^*$	E_1 – Stable – chatter	

Table 3.4. Distribution and stability of equilibria for $f(x)$ in (3.2.16).

Recall that $x_1 \approx 0.0352$ and $x_2 \approx 0.3850$

Case number	Parameter values	Equilibria	Eigenvalue of the Jacobian of $f(x)$ at the equilibria	Stability
1	$k_0 = 0$ $k_1 = 35$	$E_1 = 0.0373$	-154.76	E_1 – Stable – chatter
		$E_2 = 0.3911$	295.98	E_2 – Unstable
		$E_3 = 1.0$	-1.0	E_3 – Stable
2	$k_0 = 0$ $k_1 = 10$	$E_4 = 0.0909$	-11.0	E_1 – Stable
		$E_5 = 0.3883$	258.03	E_2 – Unstable
		$E_6 = 1.0$	-1.0	E_3 – Stable
3	$k_0 = 0$ $k_1 = 1$	$E_7 = 1.0$	-1.0	E_1 – Stable
4	$k_0 = 30$ $k_1 = 35$	$E_8 = 0.0316$	-41.12	E_1 – Stable
5	$k_0 = 5$ $k_1 = 20$	$E_9 = 0.0477$	-21.67	E_1 – Stable
6	$k_0 = 5$ $k_1 = 35$	$E_{10} = 0.0369$	-144.70	E_1 – Stable – chatter

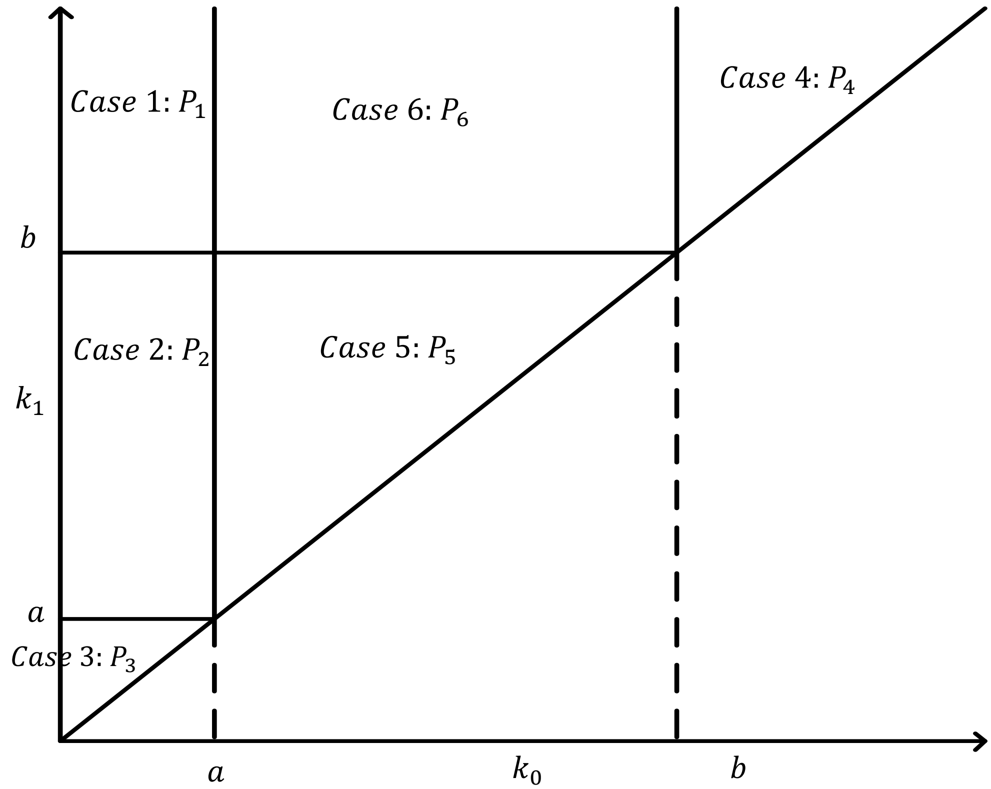


Figure 3.6. A partition of the 2-D parameter space with $\boldsymbol{k_0}$ as x-axis and $\boldsymbol{k_1}$ as y-axis induced by the six cases listed in Table 3.2. $\boldsymbol{a = \frac{1-x_2}{x_2} \approx 1.597}$, $\boldsymbol{b = \frac{1-x_1}{x_1} \approx 27.409}$. Since $\boldsymbol{k_0 < k_1}$ we only need to consider the space above the $\boldsymbol{45°}$ line where $\boldsymbol{k_0 = k_1}$ completeness of the six cases in Table 3.2 follows from the fact that the union of the partitions $\boldsymbol{P_1}$ through $\boldsymbol{P_6}$ is equal to the upper half $(\boldsymbol{k_1 \geq k_0})$ of the positive quadrant of the 2-D plane.

In other words, by Fillipov – Utkin theory implies that $f(x_1) = 0$ which in turn means that the trajectory $x(t)$ starting from $x(0)$ where $-\infty < x(0) < x_1$ or $x_1 < x(0) < x_2$, should ever reach the point x_1. From (3.2.19) and (3.2.21), it is immediate that it should remain there forever. A little reflection however reveals that x_1 is a point of discontinuity of the field $f(x)$ and x_1 by virtue of being a root of a 5th degree polynomial equation is, in general, an irrational number. Hence, the probability that the trajectory $x(t)$ will ever settle on x_1 is zero.

It can be shown (refer to Appendix E) that the trajectory $x(t)$ starting from $(-\infty, x_1)$ or (x_1, x_2) will enter a δ-neighborhood of x_1 in finite time, say k^*. Since [*] $f_1(x_1^-) > 0$ and $f_2(x_1^+) > 0$, it is immediate that the trajectory once it reaches the δ-neighborhood of x_1 in time k^*, then for all $k > k^*$, it will oscillate between the two sides of x_1. That is x_1 corresponds to attracting sliding regime in Appendix D.

The above analysis applies to equilibria E_1 and E_{10}. Refer to Figures 3.13 and 3.14 for an illustration of case 1 in Table 3.3.

B. Behavior of (3.2.16) around x_2:

By a similar analysis, using $h(x) = x - x_1$, it also follows that $f_U(x_2) = f_F(x_2) = 0$. However, since $f_1(x_2^+) > 0$ and $f_2(x_2^-) > 0$, x_2 is basically repulsive and the solution starting from (x_1, x_2) or (x_2, ∞) can never approach x_2. Hence, the equilibria E_2 and E_5 can never be reached at all.

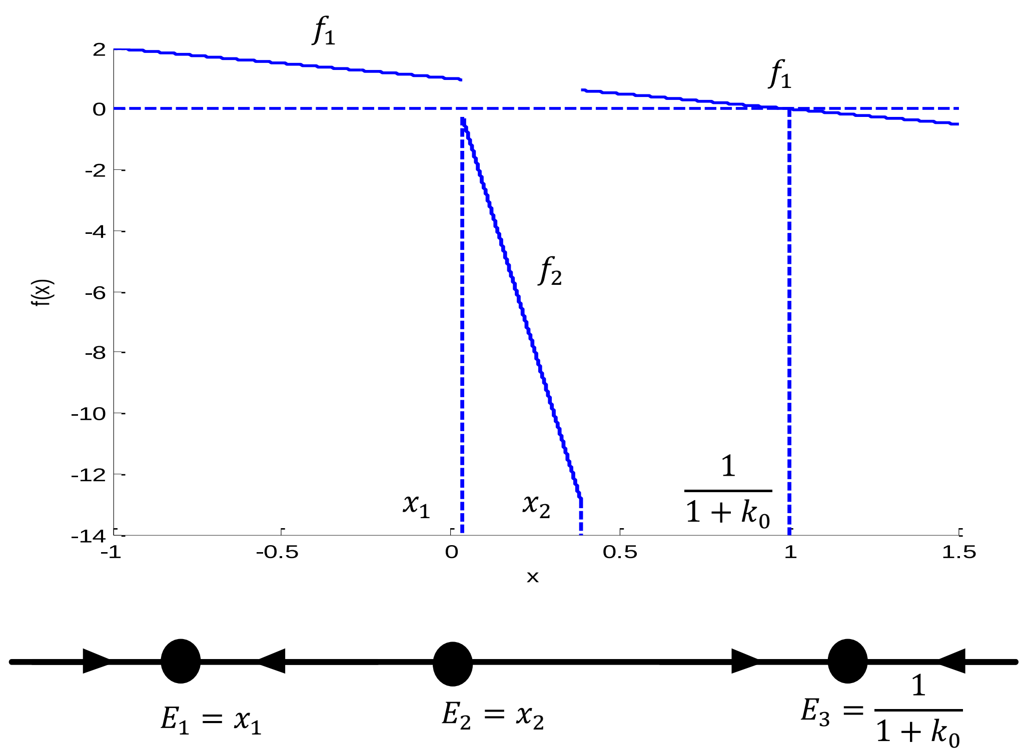


Figure 3.7. Flow field for case 1 ($\boldsymbol{k_0 = 0}$ and $\boldsymbol{k_1 = 35}$)

(*) $\boldsymbol{x^-}$ refers to points left of $\boldsymbol{x}$ and $\boldsymbol{x^+}$ refers to points right of $\boldsymbol{x}$

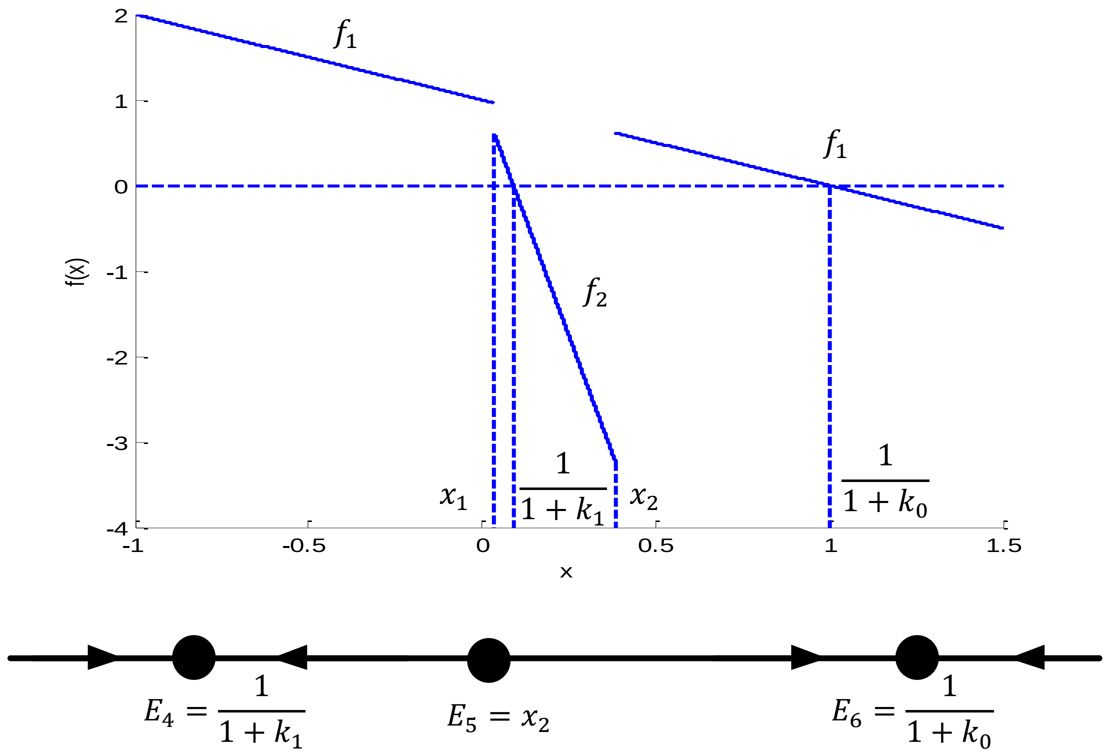


Figure 3.8. Flow field for case 2 ($\boldsymbol{k_0 = 0}$ and $\boldsymbol{k_1 = 10}$)

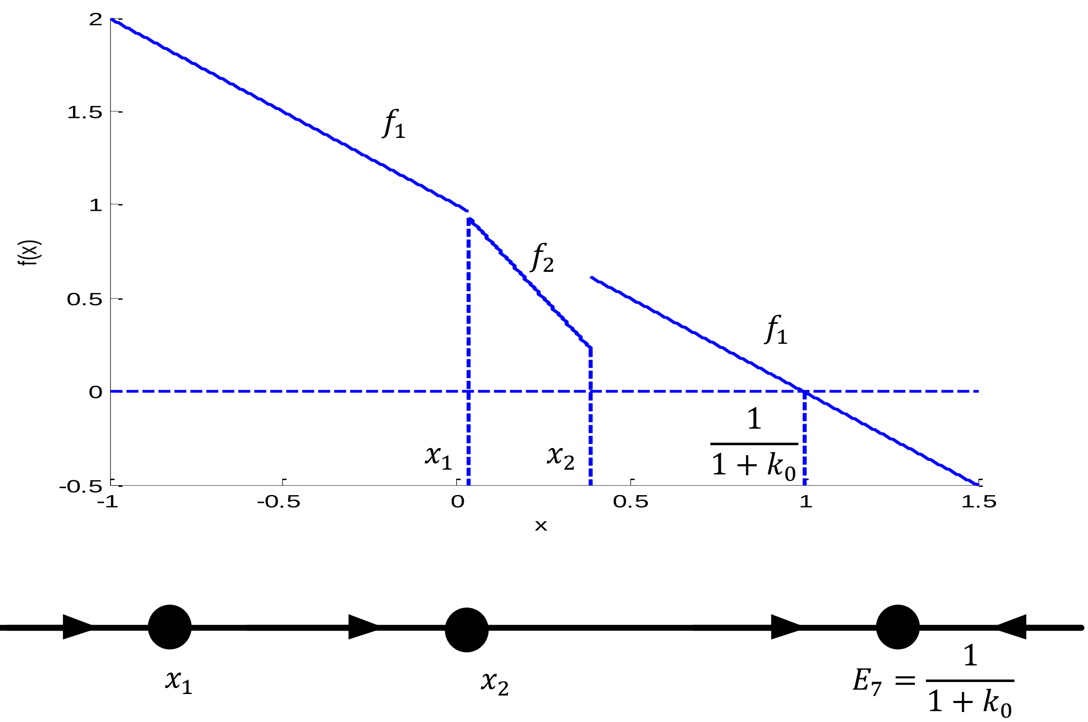


Figure 3.9. Flow field for case 3 ($\boldsymbol{k_0 = 0}$ and $\boldsymbol{k_1 = 1}$)

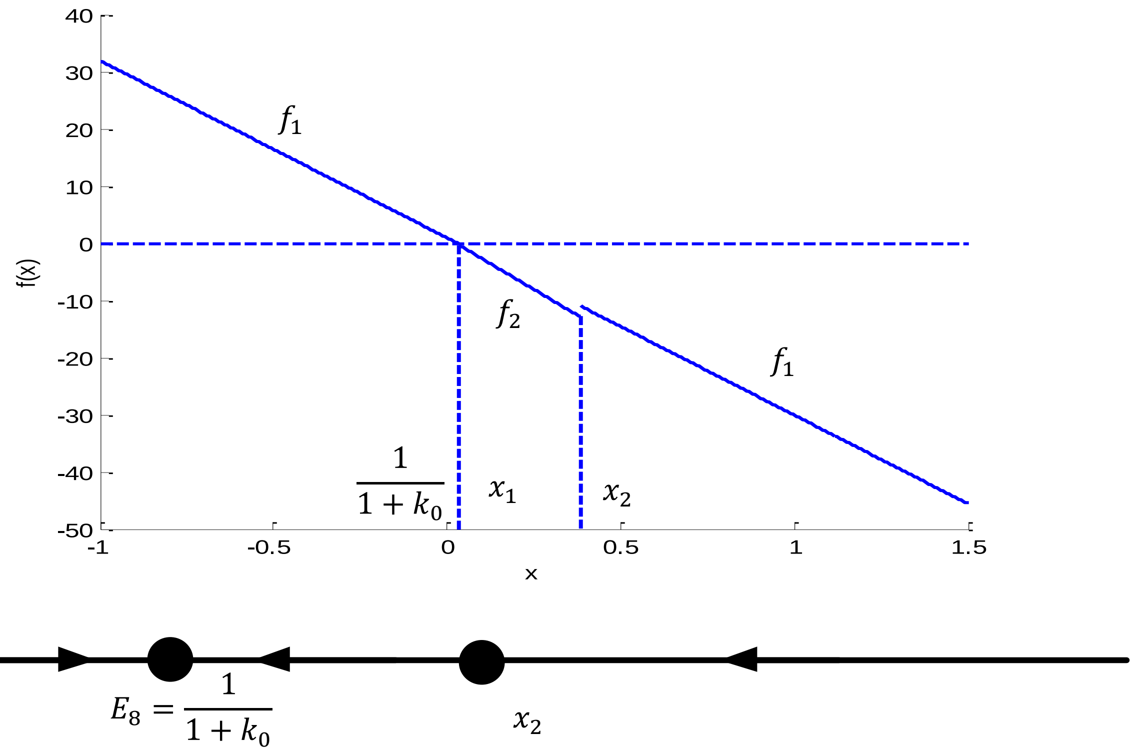


Figure 3.10. Flow field for case 4 ($\boldsymbol{k_0 = 30}$ and $\boldsymbol{k_1 = 35}$)

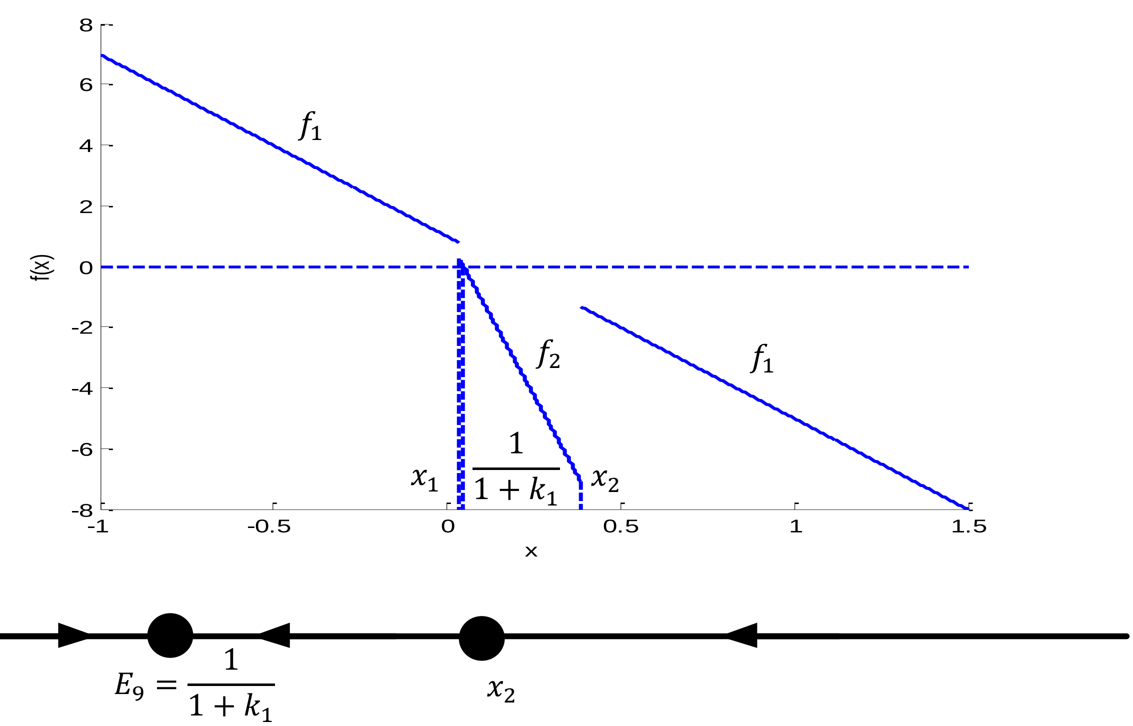


Figure 3.11. Flow field for case 5 ($\boldsymbol{k_0 = 5}$ and $\boldsymbol{k_1 = 20}$)

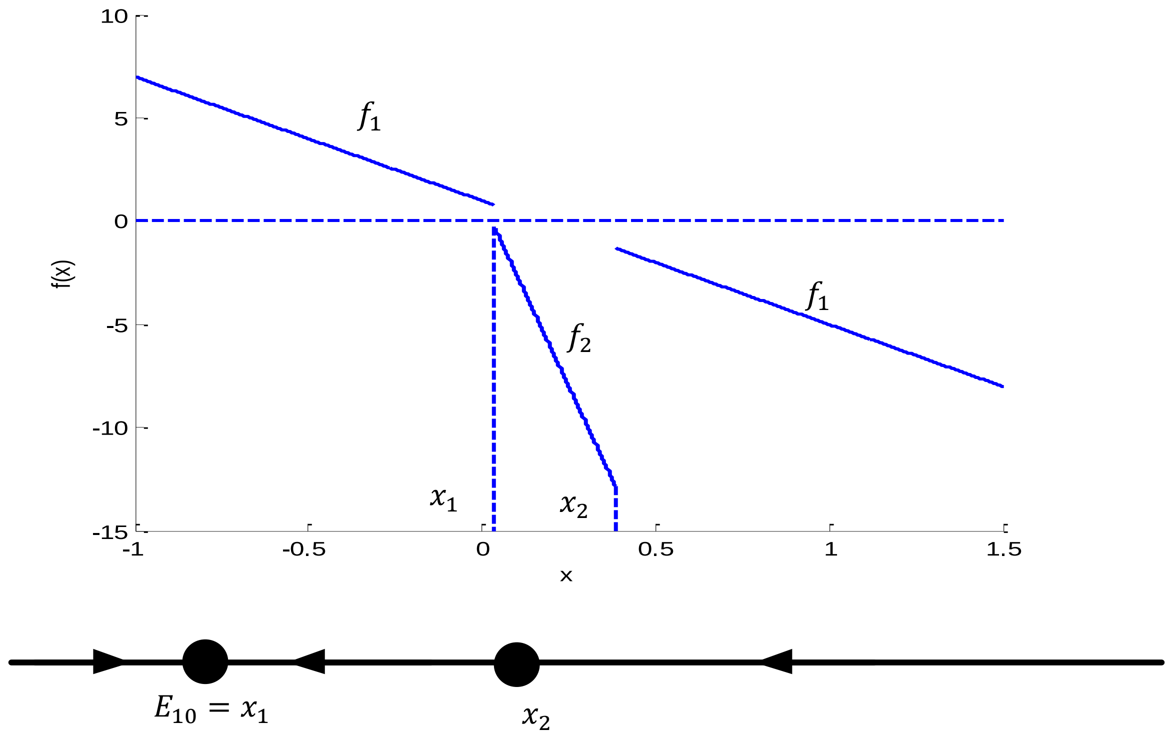


Figure 3.12. Flow field for case 2 ($\boldsymbol{k_0 = 0}$ and $\boldsymbol{k_1 = 10}$)

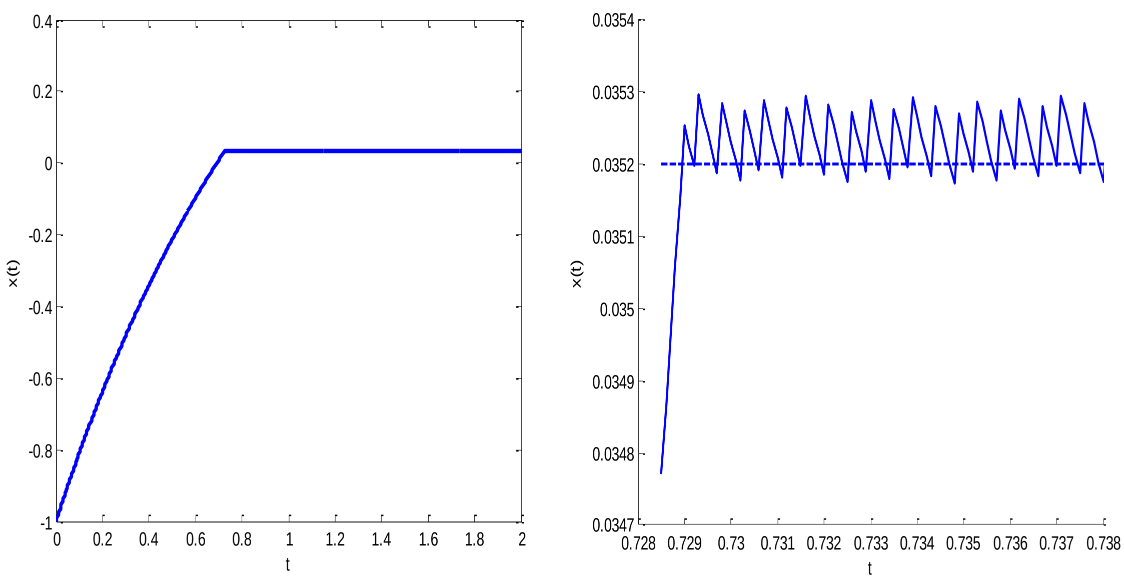


Figure 3.13. Plot of the solution $\boldsymbol{x(t)}$ vs. $\boldsymbol{t}$ for $\boldsymbol{x(0) = -1}$

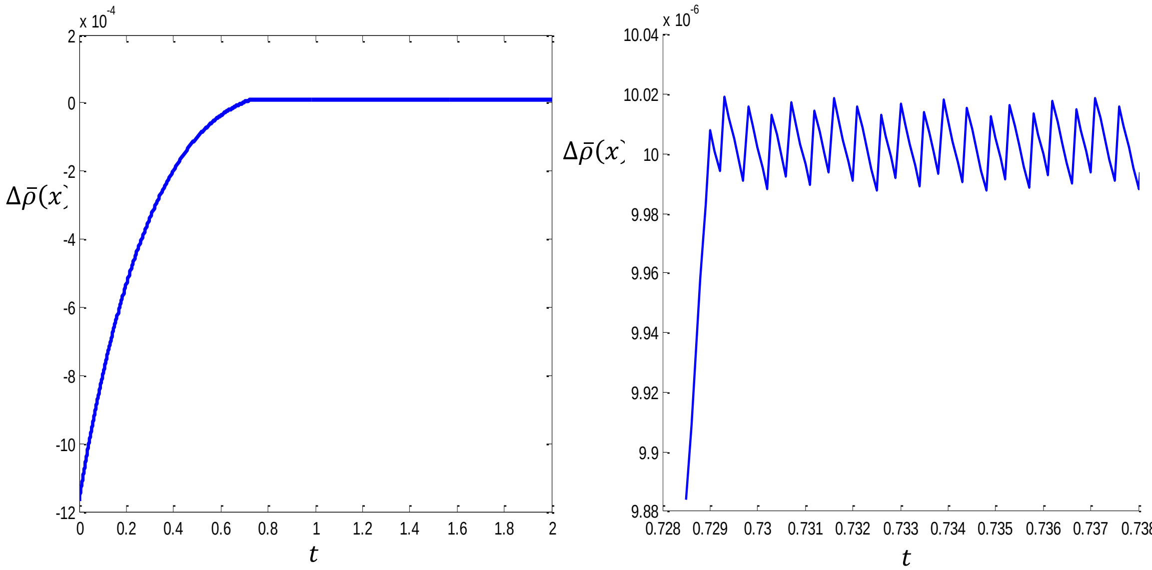


Figure 3.14. Plot of the variation of $\Delta\bar{\boldsymbol{\rho}}(\boldsymbol{x})$ vs. $\boldsymbol{t}$ for $\boldsymbol{x(0) = -1}$

c) A look at bifurcation

Referring to Figure 3.6, consider case 1 when $0 \le k_0 < a$, and $b < k_1$ with $f_1(x_1) > 0 > f_2(x_1)$, and $f_1(x_2) > 0 > f_2(x_2)$. For definiteness, let us fix k_0 to some value in the range $0 \le k_0 < a$. This means that the shape of $f_1(x)$ is fixed. Now, consider the effect of decreasing k_1 on the shape of $f_2(x)$ from its current value $b < k_1$. When $k_1 = b$, it follows that $f_2(x_1) = 0$. That is, as the value of k_1 is moved down across b, $f_2(x_1)$ increases from being negative to positive which in turn means that case 1 morphs to case 2 in Figure 3.6. Also refer to Tables 3.2 and 3.3 for changes in the equilibria and their properties, when k_1 reaches down to the level $k_1 = a$, it can be verified that $f_2(x_2) = 0$. That is, as k_1 decreases across a, $f_2(x_2)$ increases from negative to positive . Thus, case 2 morphs to case 3 in Figure 3.6. Again refer to Tables 3.2 and 3.3 for changes in the number and properties of equilibria. A similar bifurcation can be demonstrated when (a) one keeps k_1 fixed and increases k_0 while keeping $k_0 < k_1$ and (b) by varying both k_0 and k_1 across a diagonal parallel to the 45° line in $k_0 - k_1$ plane in Figure 3.6.

3.3. Conclusions

In this chapter, a theoretical study is made of simple mixed-layer models, one with salt water and the other with pure water. A simple limiting variant of the models described is obtained by taking $k(\Delta\rho)$ (in heat-salt model) and $\bar{\eta}f\big(\Delta\bar{\rho}(x)\big)$ (in pure water model) as a step function. Therefore, the ultimate use of the analysis rests on the quantifying the values of k_0 and k_1 in practical cases. We then demonstrate the oscillation in both cases by numerical examples.

Chapter 4

Three-box models

4.1. Rooth's model

While all the known two-box models (described in Section 2) cover only the North Atlantic basin, Rooth (1982) was the first to consider a symmetric system of well mixed three-boxes, labelled N, S and E-boxes covering the entire North, South and equatorial Atlantic Ocean. Refer to Figure 4.1 for an illustration. Two polar boxes exist with volume 1 and V_S, where surface cooling (Q_N and $Q_N > 0$) and an upward virtual salt flux (F_{SN} and F_{SS}) due to a precipitation excess maintain a relatively low temperature and salinity. At the levels in and above the thermohaline these boxes are connected with a sub-tropical and tropical box (volume V_E) that extends across the equator in both hemispheres. The equatorial box is warmed by air-sea interaction ($Q_E <$ 0). Because of salt conservation in the ocean, the virtual mass flux equal to $-(F_{SN} + F_{SS})$ due to a net evaporation excess tends to increase the equatorial salinity.

The thermohaline circulation in this model maintains a transport rate ψ between the boxes.

$$\psi = k[\alpha(T_S - T_N) - \beta(S_S - S_N)] \tag{4.1.1}$$

where k is a hydraulic constant, α and β are thermal and haline expansion coefficients.

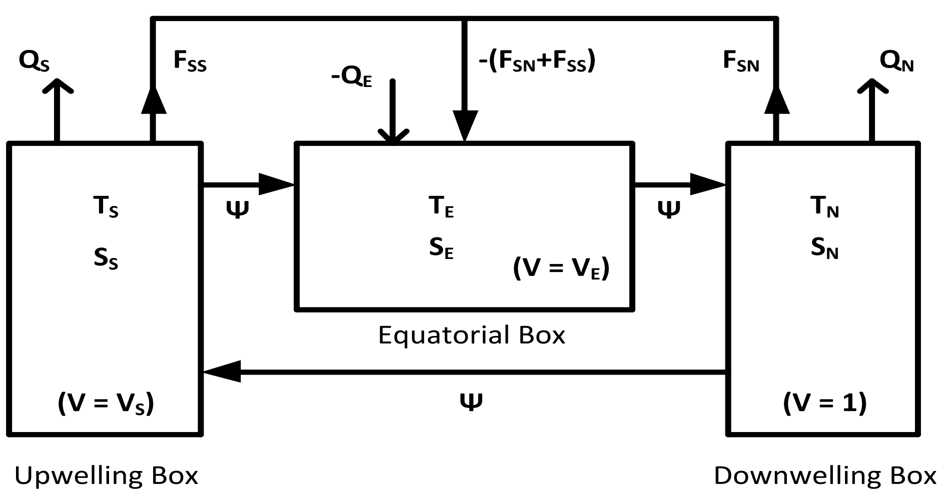


Figure 4.1. The scheme for the interhemispheric box model, introduced by Rooth.

The equations that describe the development of temperature and salinity of each box when, as in the present thermohaline circulation, $\psi \geq 0$ are

$$\frac{dT_N}{dt} = -Q_N + \psi(T_E - T_N) \tag{4.1.2}$$

$$\frac{dT_E}{dt} = -\frac{Q_E}{V_E} + \frac{\psi(T_S - T_E)}{V_E} \tag{4.1.3}$$

$$\frac{dT_S}{dt} = -\frac{Q_S}{V_S} + \frac{\psi(T_N - T_S)}{V_S} \tag{4.1.4}$$

$$\frac{dS_N}{dt} = -F_{SN} + \psi(S_E - S_N) \tag{4.1.5}$$

$$\frac{dS_E}{dt} = \frac{F_{SN} + F_{SS}}{V_E} + \frac{\psi(S_S - S_E)}{V_E} \tag{4.1.6}$$

$$\frac{dS_S}{dt} = -\frac{F_{SS}}{V_S} + \frac{\psi(S_N - S_S)}{V_S} \tag{4.1.7}$$

where Q_N and Q_S are the integrated heat losses to the atmosphere of the downwelling and the upwelling polar box (both > 0), while $-Q_E$ is the heat gain in the cross-equatorial box ($Q_E < 0$); F_{SN} and F_{SS} are the virtual salinity fluxes for, respectively, the northern and southern polar box (both > 0).

Equilibrium analysis:

$$-Q_N + \psi(T_E - T_N) = 0 \tag{4.1.8}$$

$$-\frac{Q_E}{V_E} + \frac{\psi(T_S - T_E)}{V_E} = 0 \tag{4.1.9}$$

$$-\frac{Q_S}{V_S} + \frac{\psi(T_N - T_S)}{V_S} = 0 \tag{4.1.10}$$

$$-F_{SN} + \psi(S_E - S_N) = 0 \tag{4.1.11}$$

$$\frac{F_{SN} + F_{SS}}{V_E} + \frac{\psi(S_S - S_E)}{V_E} = 0 \tag{4.1.12}$$

$$-\frac{F_{SS}}{V_S} + \frac{\psi(S_N - S_S)}{V_S} = 0 \tag{4.1.13}$$

From (4.1.10) and (4.1.13), we get

$$T_N - T_S = \frac{Q_S}{\psi} \tag{4.1.14}$$

$$S_N - S_S = -\frac{F_{SS}}{\psi} \tag{4.1.15}$$

Substituting (4.1.14) and (4.1.15) into (4.1.1), we obtain the stationary overturning $\bar{\psi}$

$$\bar{\psi}^2 = k(-\alpha Q_S + \beta F_{SS}) \tag{4.1.16}$$

This implies that the stationary overturning rate in this model is proportional to the square root of the net virtual mass flux due to air-sea interaction over the upwelling box only.

According to (4.1.14), (4.1.16) also can be written as

$$\bar{\psi}^2 = k[-\alpha(T_N - T_S)\bar{\psi} + \beta F_{SS}] \tag{4.1.17}$$

This quadratic equation of the stationary overturning rate $\bar{\psi}$ has solutions

$$\bar{\psi} = -\frac{\alpha k(T_N - T_S)}{2} \pm \frac{1}{2}\sqrt{k^2\alpha^2(T_N - T_S)^2 - 4k\beta F_{SS}} \tag{4.1.18}$$

4.2. Welander's model

An extension of Rooth (1982) box model has been studied by Welander (1986) and Thual and McWilliams (1992). Refer to Figure 4.2 for an illustration. This expanded symmetric arrangement extend the intrinsic power of the two-box models by capturing not only the equator to pole-ward (E – S – E and E – N – E) circulation but also the pole to pole-ward (S – E – N – E – S and N – E – S – E – N) circulations as well.

Referring to the Figure 4.2, the three triplets (T_n, S_n, V_n), (T_e, S_e, V_e) and (T_s, S_s, V_s) refer respectively to the temperature, salinity and the volume of the N, E and

S-boxes. Similarly, let the three pairs (T_n^a, S_n^a), (T_e^a, S_e^a), and (T_s^a, S_s^a) refer air temperature and the salinity equivalent of the fresh water input to N, E, and S-boxes respectively.

Following the first principle stated at the beginning of Section 2.1, the dynamics of evolution of temperatures and salinities in the three boxes are given in (4.2.1) – (4.2.6):

$$V_n \frac{dT_n}{dt} = C_n^T (T_n^a - T_n) + |\psi_n| (T_e - T_n) \tag{4.2.1}$$

$$V_e \frac{dT_e}{dt} = C_e^T (T_e^a - T_e) + |\psi_s| (T_s - T_e) + |\psi_n| (T_n - T_e) \tag{4.2.2}$$

$$V_s \frac{dT_s}{dt} = C_s^T (T_s^a - T_s) + |\psi_s| (T_e - T_s) \tag{4.2.3}$$

$$V_n \frac{dS_n}{dt} = C_n^S (S_n^a - S_n) + |\psi_n| (S_e - S_n) \tag{4.2.4}$$

$$V_e \frac{dS_e}{dt} = C_e^S (S_e^a - S_e) + |\psi_s| (S_s - S_e) + |\psi_n| (S_n - S_e) \tag{4.2.5}$$

$$V_s \frac{dS_s}{dt} = C_s^S (S_s^a - S_s) + |\psi_s| (S_e - S_s) \tag{4.2.6}$$

where the flows are given by

$$\begin{aligned} \psi_s &= \gamma[-\alpha(T_e - T_s) + \beta(S_e - S_s)] \\ \psi_n &= \gamma[-\alpha(T_e - T_n) + \beta(S_e - S_n)] \end{aligned} \tag{4.2.7}$$

Define

$$\begin{aligned} \frac{C_s^T}{V_s} = \frac{C_e^T}{V_e} = \frac{C_n^T}{V_n} = R_T \\ \frac{C_s^S}{V_s} = \frac{C_e^S}{V_e} = \frac{C_n^S}{V_n} = R_S \end{aligned} \tag{4.2.8}$$

where $R_S \ll R_T$. We can reduce this system of six equations in (4.2.1) – (4.2.6) to a system of four equations. Let

$$\begin{aligned} \xi_s &= T_e - T_s, \xi_n = T_e - T_n \\ \eta_s &= S_e - S_s, \eta_n = S_e - S_n \\ \xi_s^a &= T_e^a - T_s^a, \xi_n^a = T_e^a - T_n^a \end{aligned} \tag{4.2.9}$$

$$\eta_s^a = S_e^a - S_s^a, \eta_n^a = S_e^a - S_n^a$$

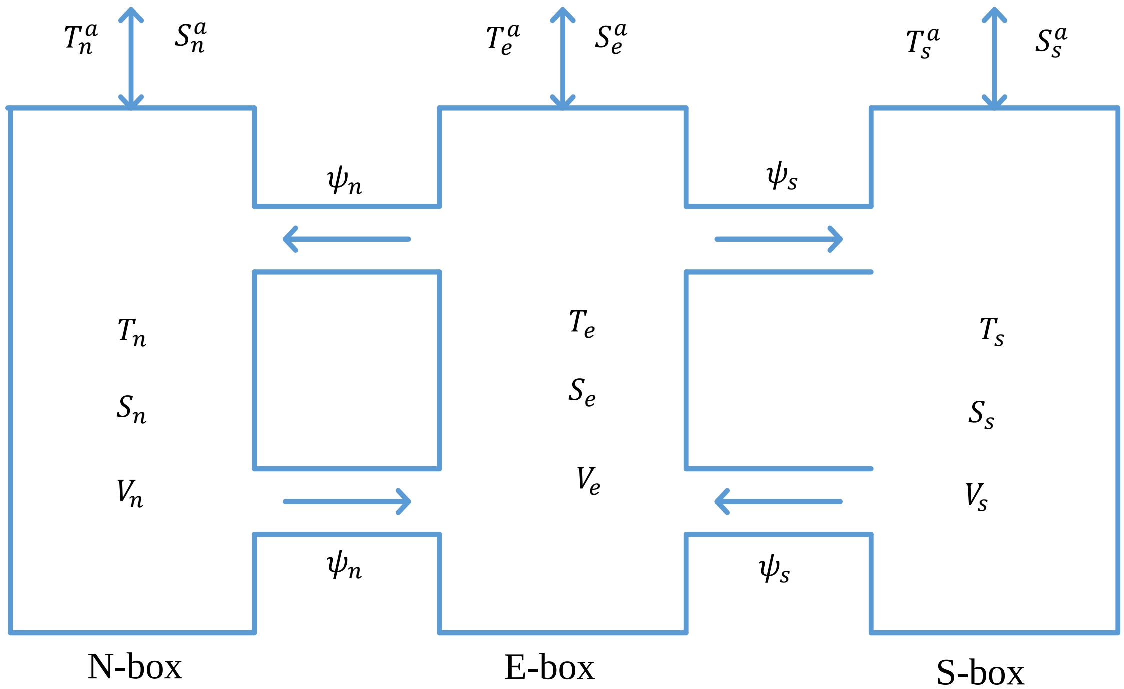


Figure 4.2. A three-box model covering the entire Atlantic Ocean where in each box the sea water is assumed to be well mixed.

Dividing (4.2.2) by V_e, (4.2.1) by V_n and subtracting the latter from the former, using (4.2.8), we obtain

$$\frac{d\xi_n}{dt} = R_T(\xi_n^a - \xi_n) - \left(\frac{|\psi_n|}{V_e} + \frac{|\psi_n|}{V_n}\right)\xi_n - \frac{|\psi_s|}{V_e}\xi_s \tag{4.2.10}$$

Similarly, dividing (4.2.2) by V_e, (4.2.3) by V_s, and subtracting the latter from the former, we obtain

$$\frac{d\xi_s}{dt} = R_T(\xi_s^a - \xi_s) - \left(\frac{|\psi_s|}{V_e} + \frac{|\psi_s|}{V_s}\right)\xi_s - \frac{|\psi_s|}{V_e}\xi_n \tag{4.2.11}$$

By similar operations on (4.2.4) – (4.2.6), we get

$$\frac{d\eta_n}{dt} = R_S(\eta_n^a - \eta_n) - \left(\frac{|\psi_n|}{V_e} + \frac{|\psi_n|}{V_n}\right)\eta_n - \frac{|\psi_s|}{V_e}\eta_s \tag{4.2.12}$$

$$\frac{d\eta_s}{dt} = R_S(\eta_s^a - \eta_s) - \left(\frac{|\psi_s|}{V_e} + \frac{|\psi_s|}{V_s}\right)\eta_s - \frac{|\psi_s|}{V_e}\eta_n \tag{4.2.13}$$

Now, define a new set of dimensionless temperature, salinity, flow and time variables as

$$x_n = \frac{\xi_n}{\left(\frac{V_S R_T}{2\gamma\alpha}\right)}, \; x_s = \frac{\xi_s}{\left(\frac{V_S R_T}{2\gamma\alpha}\right)}$$

$$y_n = \frac{\eta_n}{\left(\frac{V_S R_T}{2\gamma\beta}\right)}, \; y_s = \frac{\eta_s}{\left(\frac{V_S R_T}{2\gamma\beta}\right)} \tag{4.2.14}$$

$$\bar{\psi} = \frac{\psi}{\left(\frac{V_S R_T}{2}\right)}, \text{ and } s = \frac{t}{(1/R_T)}$$

To simplify the expression, first set $V = V_n = V_s = \frac{V_e}{2}$, then dividing (4.2.10) by R_T and substituting (4.2.14) and simplifying, (4.2.10) takes the dimensionless form as

$$\frac{dx_n}{ds} = x_n^a - x_n\left(1 + \frac{3}{4}\left|\overline{\psi_n}\right|\right) - \frac{\left|\overline{\psi_s}\right|}{4}x_s \tag{4.2.15}$$

where the forcing term is given by

$$x_n^a = \frac{2\gamma\alpha}{VR_T}(T_e^a - T_n^a) \tag{4.2.16}$$

The dimensionless flows from (4.2.7), (4.2.9) and (4.2.14) are

$$\overline{\psi_n} = \frac{2\psi_n}{VR_T} = \frac{2\gamma}{VR_T}(-\alpha\xi_n + \beta\eta_n) = -x_n + y_n \tag{4.2.17}$$

and

$$\overline{\psi_s} = \frac{2\psi_s}{VR_T} = \frac{2\gamma}{VR_T}(-\alpha\xi_s + \beta\eta_s) = -x_s + y_s \tag{4.2.18}$$

Substituting (4.2.17) – (4.2.18) into (4.2.15), the latter becomes

$$\frac{dx_n}{ds} = x_n^a - x_n\left(1 + \frac{3}{4}|y_n - x_n|\right) - \frac{x_s}{4}|y_s - x_s| \tag{4.2.19}$$

By similar sequence of operations, the dimensionless form of the rest of the three equations (4.2.4) – (4.2.6)

$$\frac{dx_s}{ds} = x_s^a - x_s\left(1 + \frac{3}{4}|y_s - x_s|\right) - \frac{x_n}{4}|y_n - x_n| \tag{4.2.20}$$

$$\frac{dy_n}{ds} = y_n^a - y_n\left(\varepsilon + \frac{3}{4}|y_n - x_n|\right) - \frac{y_s}{4}|y_s - x_s| \tag{4.2.21}$$

$$\frac{dy_s}{ds} = y_s^a - y_s\left(\varepsilon + \frac{3}{4}|y_s - x_s|\right) - \frac{y_n}{4}|y_n - x_n| \tag{4.2.22}$$

where the dimensionless forcing term are

$$x_s^a = \frac{2\gamma\alpha}{VR_T}(T_e^a - T_s^a)$$

$$y_n^a = \frac{R_S}{R_T}\frac{2\gamma\alpha}{VR_T}(S_e^a - S_n^a) \tag{4.2.23}$$

$$y_s^a = \frac{R_S}{R_T}\frac{2\gamma\alpha}{VR_T}(S_e^a - S_s^a)$$

and

$$\varepsilon = \frac{R_S}{R_T} \tag{4.2.24}$$

4.3. Roebber's model

The basic geometry of the Roebber's model is shown in Figure 4.3. The North Atlantic, which is considered to be isolated from the rest of the world ocean, is represented by a deep ocean box and two near-surface boxes, the latter boxes extending from the equator to $45°N$, and then from $45°N$ to $70°N$.

The evolution of ocean temperature (T, in Kelvin) and salinity (S, in practical salinity unit or psu) is calculated in the following set of equations:

$$V_1\frac{dT_1}{dt} = \frac{1}{2}q(T_2 - T_3) + K_T(T_{A_1} - T_1) - K_{Z_1}(T_1 - T_3) \tag{4.3.1}$$

$$V_2\frac{dT_2}{dt} = \frac{1}{2}q(T_3 - T_1) + K_T(T_{A_2} - T_2) - K_{Z_2}(T_2 - T_3) \tag{4.3.2}$$

$$V_3\frac{dT_3}{dt} = \frac{1}{2}q(T_1 - T_2) + K_{Z_1}(T_1 - T_3) + K_{Z_2}(T_2 - T_3) \tag{4.3.3}$$

$$V_1\frac{dS_1}{dt} = \frac{1}{2}q(S_2 - S_3) - K_{Z_1}(S_1 - S_3) - Q_s \tag{4.3.4}$$

$$V_2\frac{dS_2}{dt} = \frac{1}{2}q(S_3 - S_1) - K_{Z_2}(S_2 - S_3) + Q_s \tag{4.3.5}$$

$$V_3 \frac{dS_3}{dt} = \frac{1}{2} q(S_1 - S_2) + K_{Z_1}(S_1 - S_3) + K_{Z_2}(S_2 - S_3) \tag{4.3.6}$$

where q is the thermohaline circulation with the flow as indicated in Figure 2, V_j is the box volume, T_j and S_j denote the volume average temperature and salinity, K_{Z_j} is the vertical eddy mixing coefficient of box j, K_T is the temperature restoring constant, and Q_s is the volume averaged equivalent salt flux, proportional to the differential net surface evaporation.

The thermohaline circulation is defined by

$$q = \frac{k(\rho_1 - \rho_2)}{\rho_0} = k[\alpha(T_2 - T_1) - \beta(S_2 - S_1)] \tag{4.3.7}$$

where k is a hydraulic constant linking volume transport m to the density difference, ρ_0 is the reference density, α and β are thermal and haline expansion coefficients.

We now non-dimensionalize equations (4.3.1) – (4.3.6) by introducing a new set of scaled variables.

$$\begin{gathered} T_i' = \frac{T_i}{T_0},\ i = 1, 2, 3 \\ S_i' = \frac{\beta S_i}{\alpha T_0},\ i = 1, 2, 3 \\ q' = \frac{q}{q^*} \\ t' = \frac{q^* t}{V_1 + V_2 + V_3} \end{gathered} \tag{4.3.8}$$

where α and β are thermal and haline expansion coefficients, and T_0 and q^* are a reference temperature and thermohaline circulation.

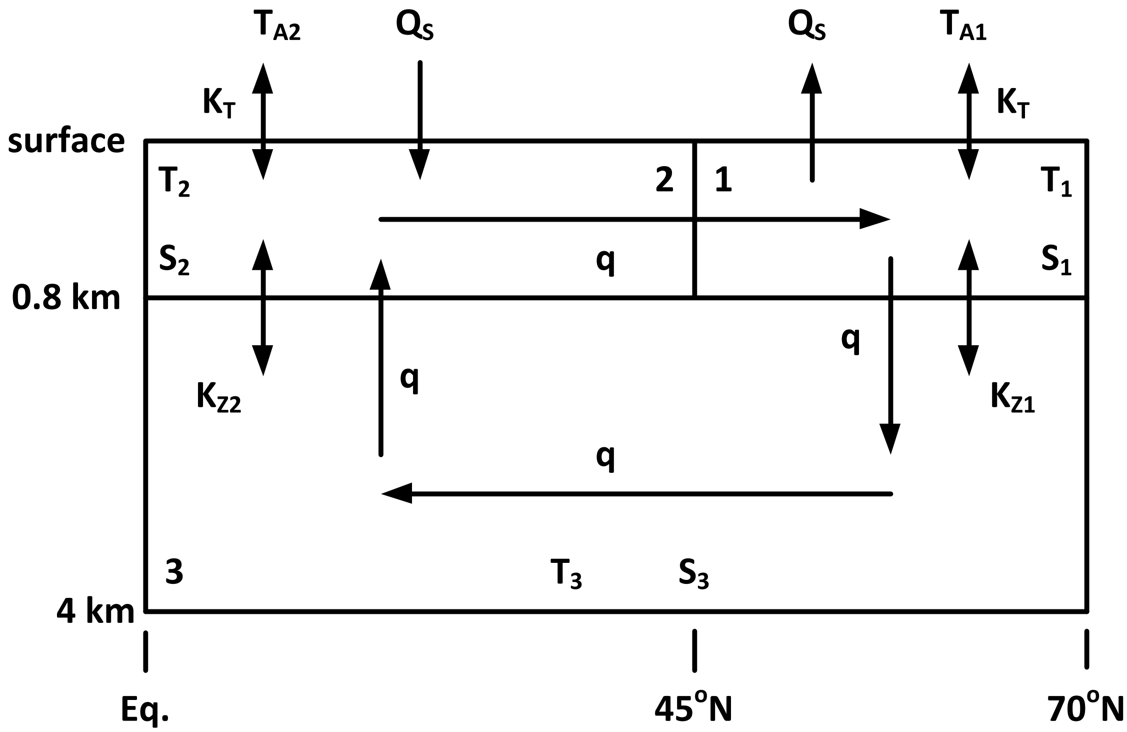


Figure 4.3. The geometry of the North Atlantic three box model

This results in the following, non-dimensional equations for the ocean model:

$$r_1 \frac{dT_1'}{dt} = \frac{1}{2} q'(T_2' - T_3') + K_T'\left(T_{A_1}' - T_1'\right) - K_{Z_1}'(T_1' - T_3') \tag{4.3.9}$$

$$r_2 \frac{dT_2'}{dt} = \frac{1}{2} q'(T_3' - T_1') + K_T'\left(T_{A_2}' - T_2'\right) - K_{Z_2}'(T_2' - T_3') \tag{4.3.10}$$

$$r_3 \frac{dT_3'}{dt} = \frac{1}{2} q'(T_1' - T_2') + K_{Z_1}'(T_1' - T_3') + K_{Z_2}'(T_2' - T_3') \tag{4.3.11}$$

$$r_1 \frac{dS_1'}{dt} = \frac{1}{2} q'(S_2' - S_3') - K_{Z_1}'(S_1' - S_3') - Q_S' \tag{4.3.12}$$

$$r_2 \frac{dS_2'}{dt} = \frac{1}{2} q'(S_3' - S_1') - K_{Z_2}'(S_2' - S_3') + Q_S' \tag{4.3.13}$$

$$r_3 \frac{dS_3'}{dt} = \frac{1}{2} q'(S_1' - S_2') + K_{Z_1}'(S_1' - S_3') + K_{Z_2}'(S_2' - S_3') \tag{4.3.14}$$

Where the prime denotes a non-dimensional variable and

$$r_i = \frac{V_i}{V_1+V_2+V_3},\ i = 1, 2, 3$$

$$K'_T = \frac{K_T}{q^*},\ K'_{Z_j} = \frac{K_{Z_j}}{q^*}$$

$$Q'_s = \frac{\beta Q_s}{\alpha T_0 q^*} \qquad (4.3.15)$$

$$q' = \frac{k\alpha T_0}{q^*}[(T'_2 - T'_1) - (S'_2 - S'_1)]$$

Table 4.1. The values of the parameters used for all integrations of the ocean model.

Dimensional constant	Dimensional value	Non-dimensional constant	Non-dimensional value
V_1	$0.832 \times 10^{16} m^3$	r_1	0.060624
V_2	$2.592 \times 10^{16} m^3$	r_2	0.188866
V_3	$10.3 \times 10^{16} m^3$	r_3	0.75051
K_T	$10.5 \times 10^6 m^3 s^{-1}$	K'_T	1.05
K_Z	$1.0 \times 10^6 m^3 s^{-1}$	K'_Z	0.10
q^*	$10 \times 10^6 m^3 s^{-1}$		
α	$1.0 \times 10^{-4} K^{-1}$		
β	$8.0 \times 10^{-4} psu^{-1}$		
T_0	$278.15\ K$		
μ	$4.0 \times 10^{10} m^3 s^{-1}$		

By setting the left-hand sides of equations (4.3.9) – (4.3.14) to zero and expanding the resultant equations in powers of q, we can derive a 5th order equation describing the steady-state solutions of the ocean model (see Appendix F).

$$q^5 + 4K_Z(K_T + 2K_Z)q^3 + 4\lambda K_Z\left[2Q_s - K_T\left(T_{A_2} - T_{A_1}\right)\right]q^2$$
$$+ 16K_Z^3(K_T + K_Z)q + 32\lambda K_Z^2(K_T + K_Z)Q_s \qquad (4.3.23)$$
$$- 16\lambda K_Z^2 K_T\left(T_{A_2} - T_{A_1}\right) = 0$$

where $K_Z = K_{Z_1} = K_{Z_2}$ and $\lambda = \frac{\mu\alpha T_0}{q^*}$.

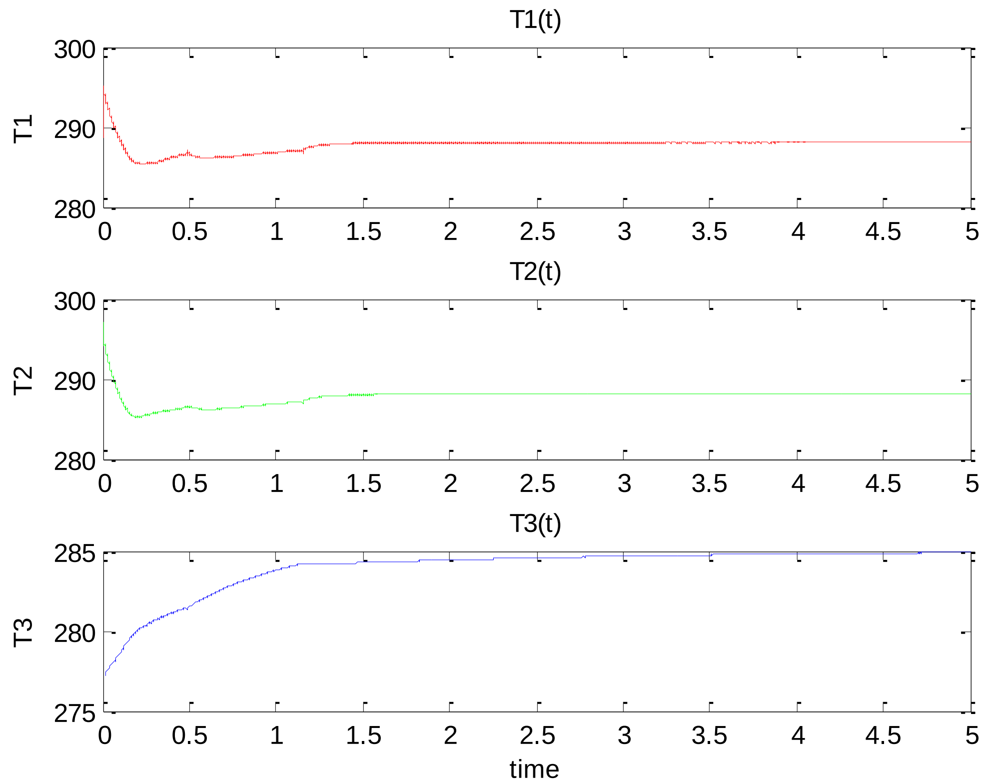


Figure 4.4. Plot of values of temperature (T) of the model

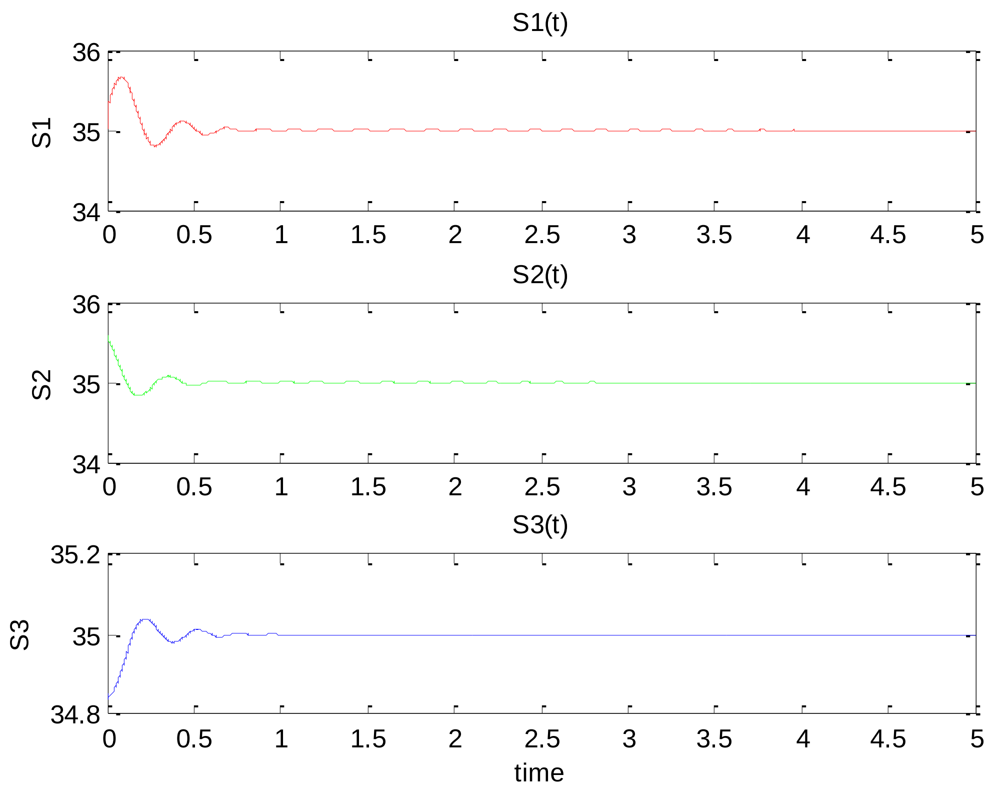


Figure 4.5. Plot of values of temperature (S) of the model

4.4. Conclusions

In this chapter, we focused our attention on the 3-box models of the ocean by adding a third box to Stommel (1961) two box model. Rooth (1982) developed the first three-box model which consists of two polar boxes and an equatorial box. Using Rooth's model, a highly idealized two-hemisphere ocean such as Atlantic or Pacific can be studied. Another type of cross equatorial box model for the global thermohaline circulation was devised by Welander (1986) who combined two Stommel box models back to back, where the equatorial boxes combine into a single cross-equatorial fully mixed box. The last three box model, presented by Roebber (1996), is single-hemisphere ocean (the North Atlantic Ocean). By limiting the study to a single ocean basin, the results obtained will not address the global dynamics of the ocean circulation.

Chapter 5

Four-box models

5.1. General four-box model

Zickfeld et al. (2003) presents a four-box extension of the classic Stommel model (Stommel 1961). The model configuration is shown in Figure 1. It consists of four well-mixed boxes, representing the southern, tropical, northern and deep Atlantic, respectively.

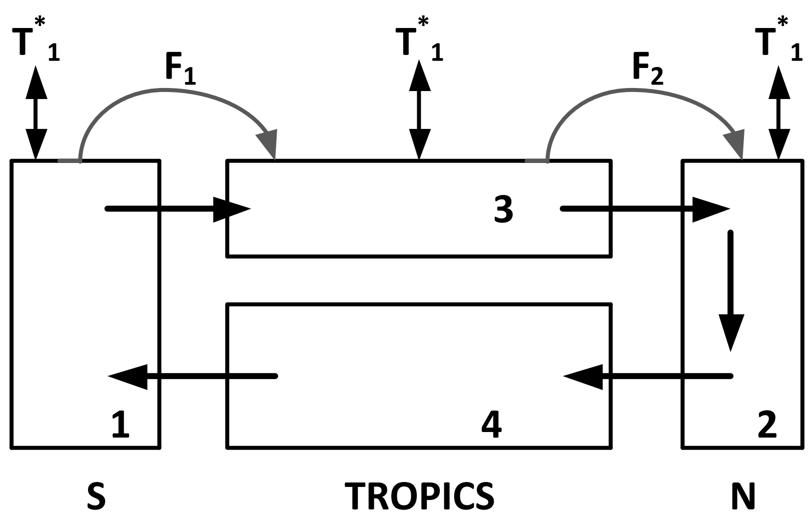


Figure 5.1. Schematic of the four-box model of the Atlantic thermohaline circulation

In this four-box model, the temperatures of boxes 1, 2 and 3 are relaxed toward the values T_1^*, T_2^* and T_3^*, respectively. The salinities are forced by the freshwater fluxes F_1 and F_2. The meridional volume transport m is proportional to the density difference $\rho_2 - \rho_1$ between boxes 1 and 2:

$$m = \frac{k(\rho_2 - \rho_1)}{\rho_0} = k[\beta(S_2 - S_1) - \alpha(T_2 - T_1)] \tag{5.1.1}$$

where $S_2 - S_1$ and $T_2 - T_1$ are the north-south salinity and temperature gradients, respectively, k is a hydraulic constant linking volume transport m to the density difference, ρ_0 is the reference density, α and β are thermal and haline expansion coefficients.

The dynamics of temperatures T_i and salinities S_i of the four boxes are given by:

$$\dot{T}_1 = \frac{m}{V_1}(T_4 - T_1) + \lambda_1(T_1^* - T_1) \tag{5.1.2}$$

$$\dot{T}_2 = \frac{m}{V_2}(T_3 - T_2) + \lambda_2(T_2^* - T_2) \tag{5.1.3}$$

$$\dot{T}_3 = \frac{m}{V_3}(T_1 - T_3) + \lambda_3(T_3^* - T_3) \tag{5.1.4}$$

$$\dot{T}_4 = \frac{m}{V_4}(T_2 - T_4) \tag{5.1.5}$$

$$\dot{S}_1 = \frac{m}{V_1}(S_4 - S_1) + \frac{S_0 F_1}{V_1} \tag{5.1.6}$$

$$\dot{S}_2 = \frac{m}{V_2}(S_3 - S_2) - \frac{S_0 F_2}{V_2} \tag{5.1.7}$$

$$\dot{S}_3 = \frac{m}{V_3}(S_1 - S_3) - \frac{S_0(F_1 - F_2)}{V_3} \tag{5.1.8}$$

$$\dot{S}_4 = \frac{m}{V_4}(S_2 - S_4) \tag{5.1.9}$$

Here V_i are box volumes, λ_i thermal coupling constants, and T_i^* the temperatures towards which the southern, northern and tropical boxes are relaxed. F_1 and F_2 are the freshwater fluxes (multiplied by a reference salinity S_0) into the tropical and northern Atlantic, respectively.

5.2. Equilibrium analysis

We now consider equilibrium conditions by setting the left-hand sides of equations (5.1.2) – (5.1.9) to zero. From (5.1.2) and (5.1.9), the deep box plays no role $(T_2 = T_4, S_2 = S_4)$, and the steady-state equations for temperature and salinity are:

$$\frac{m}{V_1}(T_2 - T_1) = \lambda_1(T_1 - T_1^*) \tag{5.2.1}$$

$$\frac{m}{V_2}(T_3 - T_2) = \lambda_2(T_2 - T_2^*) \tag{5.2.2}$$

$$\frac{m}{V_3}(T_1 - T_3) = \lambda_3(T_3 - T_3^*) \tag{5.2.3}$$

$$m(S_2 - S_1) = -S_0 F_1 \tag{5.2.4}$$

$$m(S_3 - S_2) = S_0 F_2 \tag{5.2.5}$$

$$m(S_1 - S_3) = S_0(F_1 - F_2) \tag{5.2.6}$$

Combining (5.1.1) and (5.2.4), we obtain a quadratic equation in m:

$$m^2 + k\alpha(T_2 - T_1)m + k\beta S_0 F_1 = 0 \tag{5.2.7}$$

Using symbolic toolbox in Matlab, we find that

$$T_2 - T_1 = \frac{f_1(m)}{f_2(m)} \tag{5.2.8}$$

where

$$\begin{aligned} f_1(m) = (&\lambda_1\lambda_2 T_2^* V_1 V_2 - \lambda_1\lambda_3 T_1^* V_1 V_3 - \lambda_1\lambda_2 T_1^* V_1 V_2 + \lambda_1\lambda_3 T_3^* V_1 V_3)m \\ &+ \lambda_1\lambda_2\lambda_3 T_2^* V_1 V_2 V_3 - \lambda_1\lambda_2\lambda_3 T_1^* V_1 V_2 V_3 \end{aligned} \tag{5.2.9}$$

$$\begin{aligned} f_2(m) = (&\lambda_1 V_1 + \lambda_2 V_2 + \lambda_3 V_3)m^2 \\ &+ (\lambda_1 V_1 \lambda_2 V_2 + \lambda_1 V_1 \lambda_3 V_3 + \lambda_2 V_2 \lambda_3 V_3)m + \lambda_1 V_1 \lambda_2 V_2 \lambda_3 V_3 \end{aligned} \tag{5.2.10}$$

Equation (5.2.7) in combination with (5.2.8) results in a 4th-order equation of m. The structure of the physical solutions is identical to that discussed in Rahmstorf (1996) and Titz et al. (2002) for the prescribed box temperature T_i.

Equation (5.2.7) shows that the equilibrium overturning depends only on F_1, not on F_2. This is due to the fact that under steady-state conditions it makes no difference whether the freshwater enters the northern box via atmospheric transport or oceanic advection. F_2 merely determines the salinity gradient between the tropical and the northern boxes.

5.3. Conclusions

In this chapter, a four-box model of the Atlantic thermohaline circulation is presented. The model is an interhermispheric extension of the classic Stommel (1961) model. We then study the steady-state conditions of the thermohaline circulation.

Chapter 6

Conclusion

In this thesis, we have provided a comprehensive survey of the box models. Using highly simplified box models, we focused on the current understanding of the ocean's thermohaline circulation, especially on its stability and variability.

We reviewed the results from a hierarchy of models, ranging in complexity from simple two-box models to four-box model. Starting with the descriptions of models, we derived the governing equations. We then looked for the conditions of existence of multiple equilibria of the thermohaline circulation. We have underlined the importance of thermal and freshwater flux on the thermohaline circulation.

One advantage of low-order models in this thesis is that a detailed analysis of the dynamics can be made. In particular, we studied the mechanisms of two-box models involved in transition between steady states and the model sensitivities with respect to the control parameters and initial conditions. The results of this study might be used to support the formulation of more complex modelling efforts.

References

An Editorial Comment. (2000). "The thermohaline ocean circulation: A system with dangerous thresholds?" *Climatic Change* 46: 247-256.

An Editorial Review Essay. (2005). "Thermohaline circulation changes: A question of risk assessment." *Climatic Change* 68: 241-247.

Bernardo, M., C. Budd, A.R. Champneys, and P. Kowalczyk. (2008). *Piecewise-smooth Dynamical Systems: Theory and Applications.* Springer.

Birchfield, G. E. (1989). "A coupled ocean-atmosphere climate model: temperature versus salinity effects on the thermohaline circulation." *Climate Dynamics* 4: 57-71.

Birchfield, G. E., M. Wyant, and H. Wang. (1990). "A coupled ocean-atmosphere box model of the Atlantic Ocean: a bimodal climate response." *Journal of Marine Systems* 1: 197-208.

Bryan, F. (1986). "High-latitude salinity effects and interhemispheric thermohaline circulation." *Nature* 323: 301-304.

Cessi, P. (1994). "A simple box model of stochastically forced thermohaline flow." *Journal of Physical Oceanography* 24: 1911-1920.

Cessi, P., and W. R. Young. (1992). "Multiple equilibria in two-dimensional thermohaline circulation." *Journal of Fuild Mechanics* 241: 291-309.

Dieci, L., and L. Lopez. (2012). "A survey of numerical methods for IVPs of ODEs with discontinuous right-hand side." *Journal of Computational and Applied Mathematics* 236: 3967–3991.

Filippov, A.F. (1988). *Differential Equations with Discontinuous Righthand Sides.* Kluwer Academic Publications.

Gill, A. E. (1982). *Atmosphere-Ocean Dynamics.* Academic Press.

Hansen, J., and et al. (1983). "Efficient three-dimensional global models for climate studies: Models I and II." *Monthly Weather Review* 111 (4): 609-662.

Hofmann, M., and S. Rahmstorf. (2009). "On the stability of the Atlantic meridional overturning circulation." *Proceedings of the National Academy of Sciences of the United States of America* 106 (49): 20584-20589.

Huang, R. X., J. R. Luyten, and H. M. Stommel. (1992). "Multiple equilibrium states in combined thermal and saline circulation." *Journal of Physical Oceanography* 22: 231-246.

Kuhlbrodt, T., and et al. (2009). "An integrated assessment of changes in the thermohaline circulation." *Climatic Change* 96: 489-537.

Lorenz, E. N. (2006). "An atractor embedded in the atmosphere." *Tellus* 58A: 425-429.

Lorenz, E. N. (1984). "Irregularity: a fundamental property of the atmosphere." *Tellus* 36A: 98-110.

Lucarini, V., and P. H. Stone. (2005). "Thermohaline circulation stability: A box model study. Part I: Uncoupled model." *Journal of Climate* 18: 501-513.

Lucarini, V., and P. H. Stone. (2005). "Thermohaline circulation stability: A box model study. Part II: Coupled atmosphere-ocean model." *Journal of Climate* 18: 514-529.

Marotzke, J. (1991). "Influence of convective adjustment on the stability of the thermohaline circulation." *Journal of Physical Oceanography* 21: 903-907.

Marotzke, J. (1989). "Instabilities and multiple steady states of the thermohaline circulation." In *Oceanic Circulation Models: Combining Data and Dynamics*, by D. L. T Anderson and J. Willebrand, 501-511. Kluwer Academic Publishers.

Marotzke, J., and J. Willebrand. (1991). "Multiple equilibria of the global thermohaline circulation." *Journal of Physical Oceanography* 21: 1372-1385.

Marotzke, J., P. Welander, and J. Willebrand. (1988). "Instability and multiple steady states in a meridional-plane model of the thermohaline circulation." *Tellus* 40A: 162-172.

Rahmstorf, S. (1995). "Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle." *Nature* 378: 145-149.

Rahmstorf, S. (2002). "Ocean circulation and climate during the past 120,000 years." *Nature* 419: 207-214.

Rahmstorf, S. (1996). "On the freshwater forcing and transport of the Atlantic thermohaline circulation." *Climate Dynamics* 12: 799-811.

Rahmstorf, S. (2003). "The current climate." *Nature* 421: 699.

Rahmstorf, S. (2007). "Thermohaline ocean circulation." In *Encyclopedia of Quaternary Science*, by S. A. Elias, 739-750. Elsevier.

Rahmstorf, S., and et al. (2005). "Thermohaline circulation hysteresis: A model intercomparison." *Geophysical Research Letters* 32.

Rahmstorf, S., and J. Willebrand. (1995). "The role of temperature feedback in stabilizing the thermohaline circulation." *Journal of Physical Oceanography* 25: 787-805.

Roebber, P. J. (1995). "Climate variability in a low-order coupled atmosphere-ocean model." *Tellus* 47A: 473-494.

Rooth, C. (1982). "Hydrology and ocean circulation." *Progress in Oceanography* 11 (2): 131-149.

Ruddick, B., and L. Zhang. (1989). "The mythical thermohaline oscillator?" *Journal of Marine Research* 47: 717-746.

Srokosz, M., M. Baringer, H. Bryden, S. Cunningham, T. Delworth, S. Lozier, J. Marotzke, and R. Sutton. (2012). "Past, present, and future changes in the atlantic meridional overturning circulation." *Bulletin of the American Meteorological Society*, November: 1663-1676.

Stommel, H. M. (1993). "A conjectural regulating mechanism for determining the thermohaline structure of the oceanic mixed layer." *Journal of Physical Oceanography* 23: 142-148.

Stommel, H. M., and W. R. Young. (1993). "The average T-S relation of a stochastically forced box model." *Journal of Physical Oceanography* 23: 151-158.

Stommel, H. (1961). "Thermohaline convection with two stable regimes of flow." *Tellus* 13 (2): 224-230.

Tardif, R., and G. J. Hakim. (2014). "Coupled atmosphere-ocean data assimilation experiments with a low-order climate order." *Climatic Dynamics* 43 (5-6): 1631-1643.

Tardif, R., G. J. Hakim, and C. Snyder. (2014). "Coupled atmosphere-ocean data assimilation experiments with a low-order climate model." *Climate Dynamics* 43: 1631-1643.

Thual, O., and J. C. McWilliams. (1992). "The catastrophe structure of thermohaline convection in a two-dimensional fluid model and a comparison with low-order box models." *Geophysical & Astrophysical Fluid Dynamics* 64: 67-95.

Titz, S., T. Kuhlbrodt, S. Rahmstorf, and U. Feudel. (2002). "On freshwater-dependent bifurcations in box models of the interhemispheric thermohaline circulation." *Tellus* 54A: 89-98.

Utkin, V.I. (1992). *Sliding Modes in Control and Optimization.* New York: Springer Verlag.

Veen, L. V. (2002). *Time scale interaction in low-order climate models.* University of Utrecht, PhD Dissertation.

Veen, L. V., T. Opsteegh, and F. Verhulst. (2001). "Active and passive ocean regimes in a low-order climate model." *Tellus* 53A: 616-628.

Walin, G. (1985). "The thermohaline circulation and the control of ice ages." *Palaeogeography, Palaeoclimatology, Palaeoecology* 50: 323–332.

Weaver, A. J., and T. M. C. Hughes. (1994). "Rapid interglacial climate fluctuations driven by North Atlantic ocean circulation." *Nature* 367: 447-450.

Weaver, A. J., and T. M. C. Hughes. (1992). "Stability and variability of the thermohaline circulation and its link to climate." *Trends in Physical Oceanography* 1: 15-70.

Weaver, A. J., J. Marotzke, P. F. Cummins, and E. S. Sarachik. (1993). "Stability and variability of the thermohaline circulation." *Journal of Physical Oceanography* 23: 39-60.

Webster, P. J., and et al. (1998). "Monsoons: Process, predictability, and the prospects for prediction." *Journal of Geophysical Research* 103 (C7): 14,451-14,510.

Welander, P. (1982). "A simple heat-salt oscillator." *Dynamics of Atmospheres and Oceans* 6: 233-242.

Welander, P. (1981). "Oscillations in simple air-water system driven by a stabilizing heat flux." *Dynamics of Atmospheres and Oceans* 5: 269-280.

Welander, P. (1986). "Thermohaline effects in the ocean circulation and related simple models." In *Large-Scale Transport Processes in Oceans and Atmosphere*, by J. Willebrand and D. L. T. Anderson, 163-200. Springer Netherlands.

Whitehead, J. A. (1995). "Thermohaline ocean processes and models." *Annual Review of Fluid Mechanics* 27: 89-113.

Wunsch, C. (2005). "Thermohaline loops, Stommel box models, and the Sandström theorem." *Tellus* 57A: 84-99.

Zhang, J. (2007). *Box and low-order models of the thermohaline circulation.* University of Bremen, MSC. Thesis.

Zickfeld, K., T. Slawig, and S. Rahmstorf. (2004). "A low-order model for the response of the Atlantic thermohaline circulation to climate change." *Ocean Dynamics* 54: 8-26.

Appendix A

A second-order equation of state for sea-water

It is well known (Appendix A in Gill (1982)) that the density of sea water, denoted by $\rho(S,T,p)$, critically depends on the salinity S, temperature T, and pressure p. This pressure is the sum of the atmosphere pressure and that due to the column of sea water at a given depth. Since the atmosphere pressure is the same at all depth, by convention, pressure p at the sea surface is taken to be zero. Thus, $\rho(S,T,0)$ denotes the density of the surface layer of the sea as a function of S and T. In the analysis of thermohaline circulation, one is often interested in the density distribution in a shallow layer at the sea surface whose depth is a small fraction of the overall average depth of an ocean basin. For simplicity in the analysis, we ignore the effect of pressure on the density in this shallow layer and postulate that the density variation in this shallow layer is indeed given by $\rho(S,T) = \rho(S,T,0)$.

Following Appendix A in Gill (1982) the variation of ρ with respect to S and T is given by the sum of two polynomial functions:

$$\rho(S,T) = \rho_w(T) + \rho_S(T) \tag{A.1}$$

$$\rho_w(T) = \sum_{i=0}^{5} a_i T^i \tag{A.2}$$

$$\rho_S(T) = A(T)S + B(T)S^{\frac{3}{2}} + CS^2 \tag{A.3}$$

where

$$A(T) = \sum_{i=0}^{4} A_i T^i \tag{A.4}$$

$$B(T) = \sum_{i=0}^{2} B_i T^i \tag{A.5}$$

and C is a constant.

The values of the coefficients a_i, A_i, B_i and C are given in Table A.1. The first term in (A.1) $\rho_w(T)$ captures the density variation in pure water as a function of temperature and the second term $\rho_S(T)$ denotes the variation of density resulting from temperature variation when salinity is present.

A three dimensional plot of the variation of $\rho(S,T)$ for $-5 \leq T \leq 30$ and $28 \leq S \leq 38$ is given in figure A.1. It can be verified from figures A.2 and A.3, that for a given T, $\rho(S,T)$ as a function of S is nearly linear, but for a given S, $\rho(S,T)$ as a function of T is clearly nonlinear concave with a maximum near $T = 4$.

Table A.1. The values of the coefficients $\boldsymbol{a_i}$, $\boldsymbol{A_i}$, $\boldsymbol{B_i}$ and C.

a_0	999.842594	A_0	0.824493	B_0	-5.72466×10^{-3}
a_1	6.793952×10^{-2}	A_1	-4.0899×10^{-3}	B_1	1.0227×10^{-4}
a_2	-9.095290×10^{-3}	A_2	7.6438×10^{-5}	B_2	-1.6546×10^{-6}
a_3	1.001685×10^{-4}	A_3	-8.2467×10^{-7}		
a_4	-1.120083×10^{-6}	A_4	5.3875×10^{-9}	C	4.8314×10^{-4}
a_5	6.536330×10^{-9}				

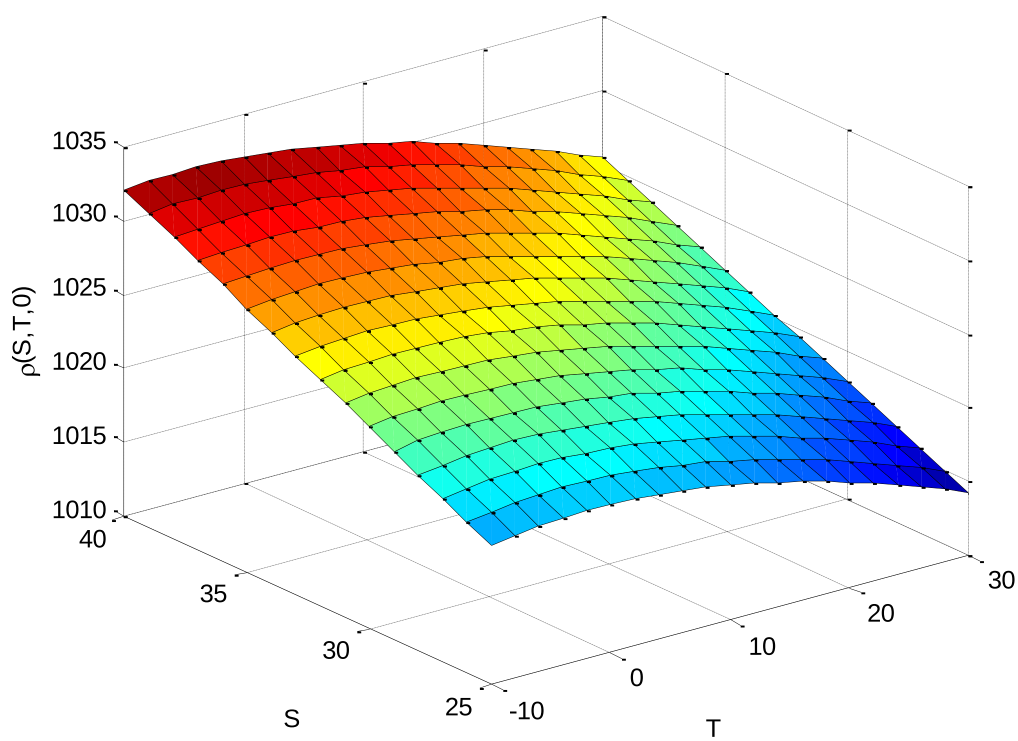


Figure A.1. Three dimensional plot of the variation of $\boldsymbol{\rho(S,T)}$ for $\boldsymbol{-10 \leq T \leq 30}$ and $\boldsymbol{25 \leq S \leq 40}$

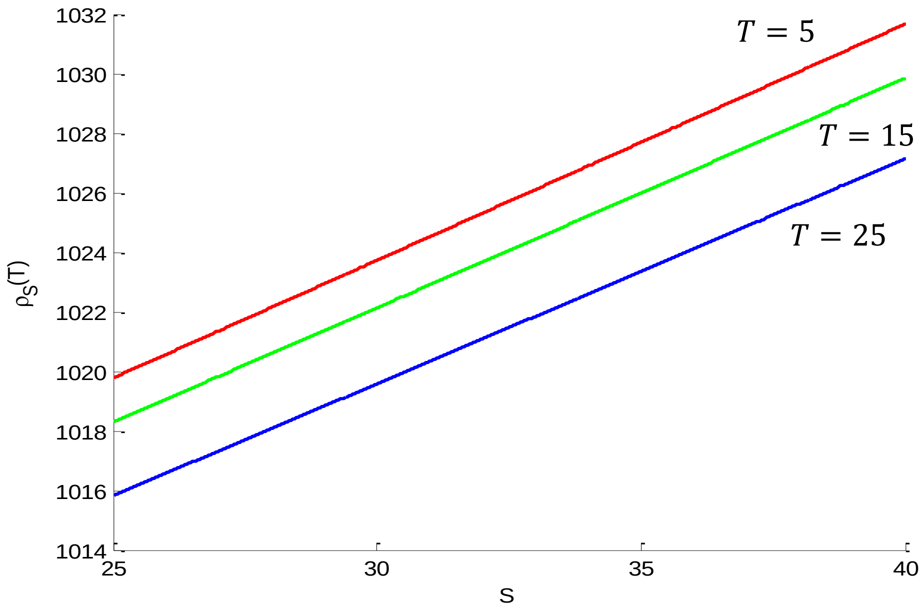


Figure A.2. Given T, plot $\boldsymbol{\rho}(\boldsymbol{S}, \boldsymbol{T})$ as a function of $\boldsymbol{S}$

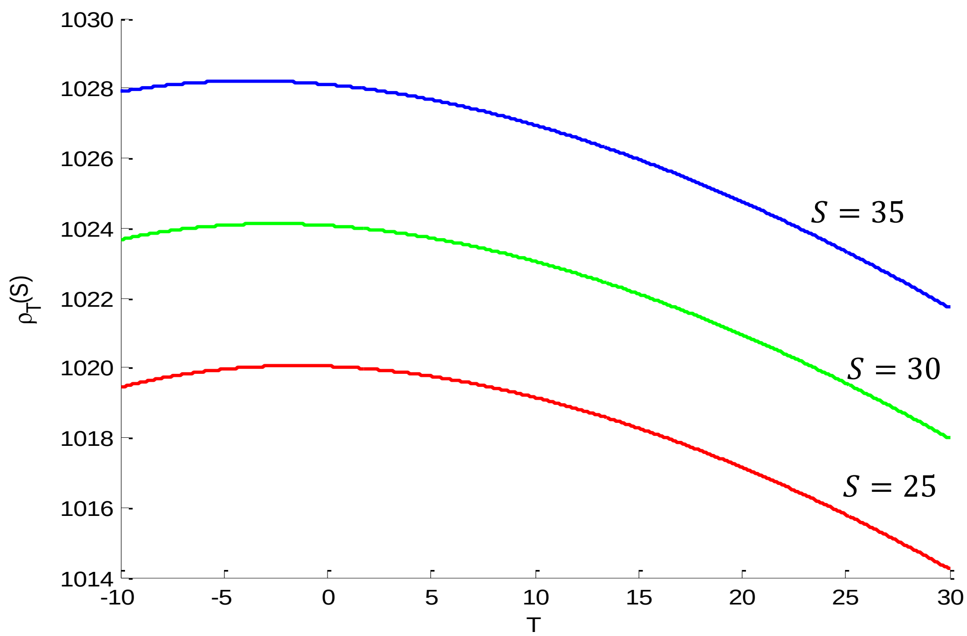


Figure A.3. Given S, plot $\boldsymbol{\rho}(\boldsymbol{S}, \boldsymbol{T})$ as a function of $\boldsymbol{T}$

Table A.2. Expressions and the values of derivatives evaluated at $(\boldsymbol{S_0}, \boldsymbol{T_0})$.

$\frac{\partial \rho_w}{\partial T_0} = \sum_{i=1}^{5} a_i i T^{i-1} = -0.0160$	$\frac{\partial^2 \rho_w}{\partial T_0^2} = \sum_{i=2}^{5} a_i i(i-1) T^{i-2} = -0.0155$
$\frac{\partial \rho_S}{\partial S_0} = A(T) + \frac{3}{2} B(T) S^{\frac{1}{2}} + 2CS = 0.7930$	$\frac{\partial^2 \rho_S}{\partial S_0^2} = \frac{3}{4} B(T) S^{-\frac{1}{2}} + 2C = 0.0003$
$\frac{\partial \rho_S}{\partial T_0} = S \sum_{i=1}^{4} A_i i T^{i-1} + S^{\frac{3}{2}} \sum_{i=1}^{2} B_i i T^{i-1} = -0.1007$	$\frac{\partial^2 \rho_S}{\partial T_0^2} = S \sum_{i=2}^{4} A_i i(i-1) T^{i-2} + 2B_2 S^{\frac{3}{2}} = 0.0039$
$\frac{\partial^2 \rho_S}{\partial S_0 \partial T_0} = \sum_{i=1}^{4} A_i i T^{i-1} + \frac{3}{2} S^{\frac{1}{2}} \sum_{i=1}^{2} B_i i T^{i-1} = -0.0026$	

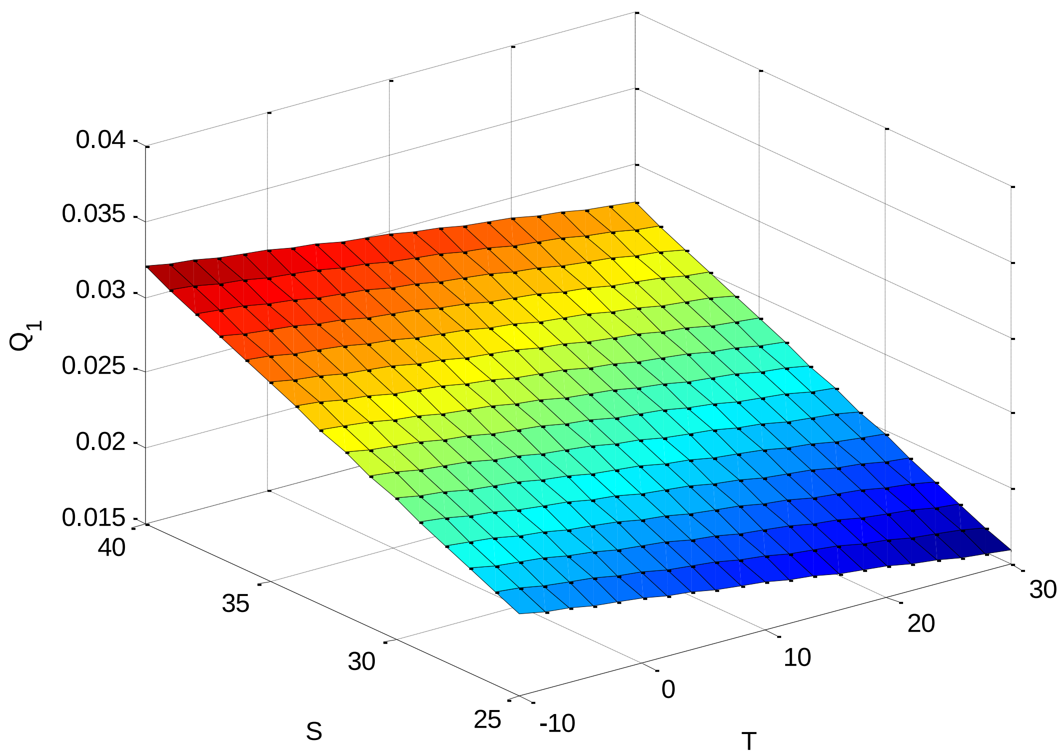


Figure A.4. A plot $\boldsymbol{Q_1}$ $\boldsymbol{vs.}$ $(\boldsymbol{S}, \boldsymbol{T})$

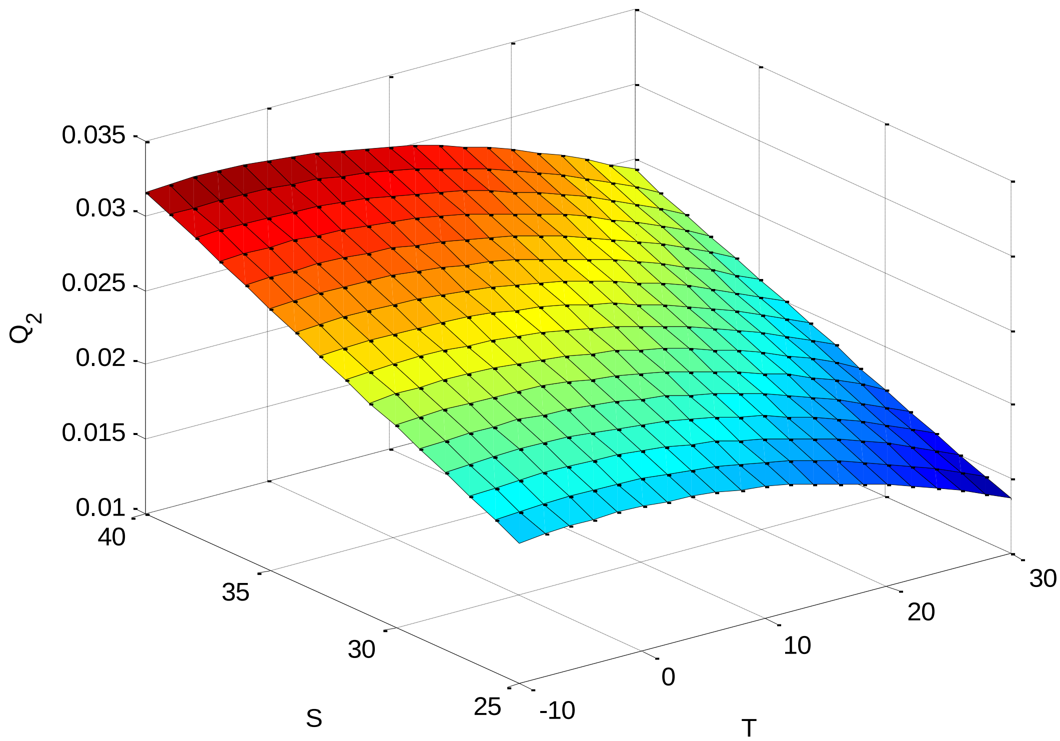


Figure A.5. A plot $\boldsymbol{Q_2}$ $\boldsymbol{vs.}$ $(\boldsymbol{S}, \boldsymbol{T})$

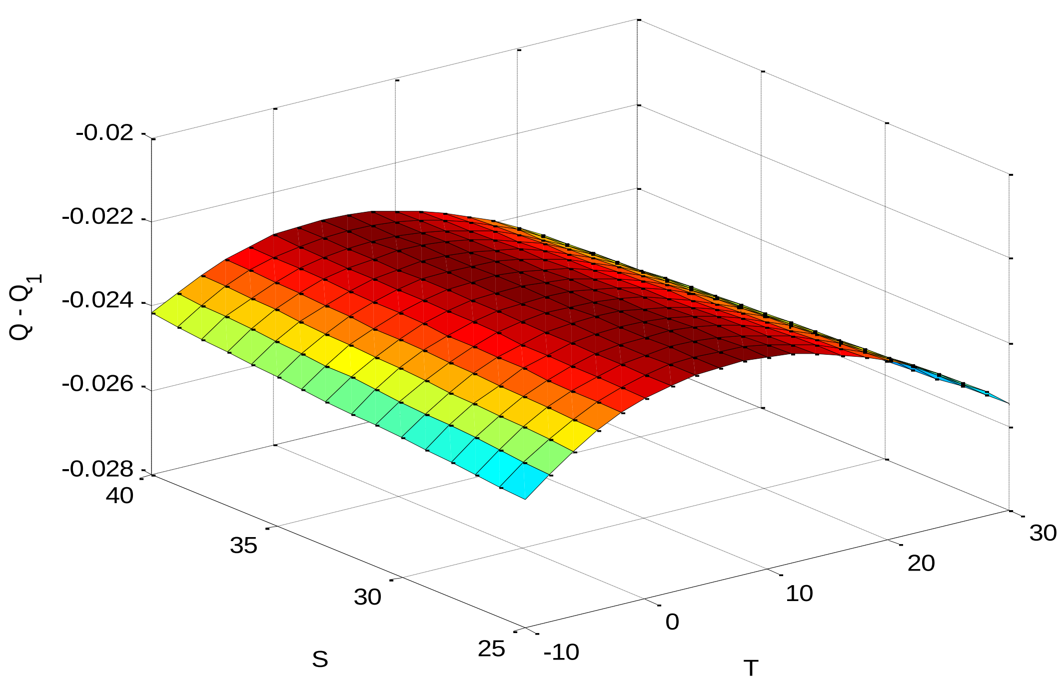


Figure A.6. A plot $\boldsymbol{Q - Q_1}$ $\boldsymbol{vs.}$ $(\boldsymbol{S}, \boldsymbol{T})$

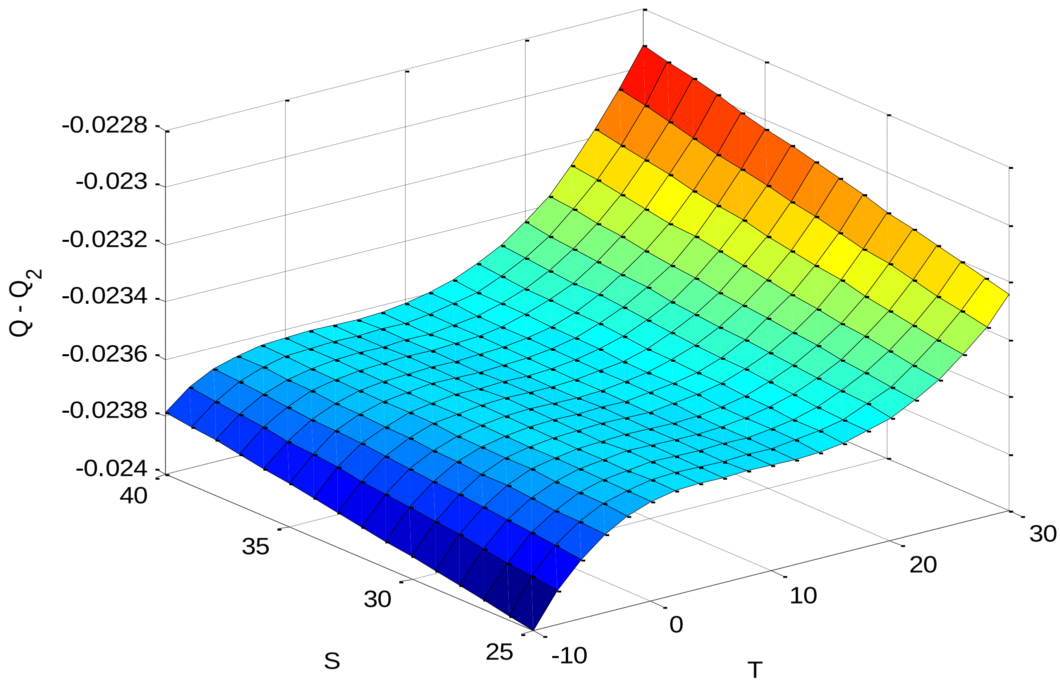


Figure A.7. A plot $\boldsymbol{Q_2 - Q\ vs.\ (S,T)}$

Let $S_0 = 35$ and $T_0 = 5$ be the reference salinity and temperature. The value of $\rho(T_0, S_0, 0)$ computed using (A.1) – (A.5) is 1027.7.

In the following, we seek to approximate $\rho(S,T)$ in (A.1) – (A.5) by a second-order approximation around (S_0, T_0). Expanding $\rho(S,T)$ in a second-order Taylor series around (S_0, T_0), we readily obtain

$$\rho(S,T) \approx \rho(S_0,T_0) + \frac{\partial\rho}{\partial S_0}(S - S_0) + \frac{\partial\rho}{\partial T_0}(T - T_0) + \frac{1}{2}\frac{\partial^2\rho}{\partial S_0^2}(S - S_0)^2$$
$$+ \frac{\partial^2\rho}{\partial S_0 \partial T_0}(S - S_0)(T - T_0) + \frac{1}{2}\frac{\partial^2\rho}{\partial T_0^2}(T - T_0)^2 \tag{A.6}$$

where $\frac{\partial\rho}{\partial S_0}$ is the value of $\frac{\partial\rho}{\partial S}$ at $S = S_0$, and similarly for other derivatives. Clearly,

$$\frac{\partial\rho}{\partial S_0} = \frac{\partial\rho_S}{\partial S_0}, \frac{\partial\rho}{\partial T_0} = \frac{\partial\rho_w}{\partial T_0} + \frac{\partial\rho_S}{\partial T_0}$$

$$\frac{\partial^2\rho}{\partial S_0^2} = \frac{\partial^2\rho_S}{\partial S_0^2}, \frac{\partial^2\rho}{\partial T_0^2} = \frac{\partial^2\rho_w}{\partial T_0^2} + \frac{\partial^2\rho_S}{\partial T_0^2}$$

$$\text{and } \frac{\partial^2 \rho}{\partial S_0 \partial T_0} = \frac{\partial^2 \rho_S}{\partial S_0 \partial T_0} \tag{A.7}$$

In the analysis of thermohaline circulation, we are particularly interested in the magnitude and sign of ratio

$$Q = \frac{\rho(S_1, T_1) - \rho(S_2, T_2)}{\rho(S_0, T_0)} \tag{A.8}$$

for $25 \leq S_1, S_2 \leq 40$ and $-10 \leq T_1, T_2 \leq 30$, which decides the flow.

A first-order approximation of Q is given by

$$Q_1 = \frac{1}{\rho(S_0, T_0)} \left[\frac{\partial \rho}{\partial S_0} S + \left(\frac{\partial \rho_w}{\partial T_0} + \frac{\partial \rho_S}{\partial T_0} \right) T \right] = \frac{1}{\rho(S_0, T_0)} (\alpha_S S + \alpha_T T) \tag{A.9}$$

with $S = S_1 - S_2$, $T = T_1 - T_2$, $\alpha_S = \frac{\partial \rho}{\partial S_0}$ and $\alpha_T = \frac{\partial \rho_w}{\partial T_0} + \frac{\partial \rho_S}{\partial T_0}$

From the Table A.2, the values of $\alpha_S = \frac{\partial \rho}{\partial S_0} = 0.7930$ and $\alpha_T = \frac{\partial \rho_w}{\partial T_0} + \frac{\partial \rho_S}{\partial T_0} = -0.0160 - 0.1007 = -0.1167$.

The second-order approximation of Q is given by

$$\begin{aligned} Q_2 = Q_1 + \frac{1}{\rho(S_0, T_0)} \Bigg[& \frac{1}{2} \frac{\partial^2 \rho}{\partial S_0^2} [(S_1 - S_0)^2 - (S_2 - S_0)^2] \\ & + \frac{\partial^2 \rho}{\partial S_0 \partial T_0} [(S_1 - S_0)(T_1 - T_0) - (S_2 - S_0)(T_2 - T_0)] \\ & + \frac{1}{2} \frac{\partial^2 \rho}{\partial T_0^2} [(T_1 - T_0)^2 - (T_2 - T_0)^2] \Bigg] \end{aligned} \tag{A.10}$$

The values of the derivatives evaluated at (S_0, T_0) are given in Table A.2. The plot of $Q_1\ vs.(S, T)$ and $Q_2\ vs.(S, T)$ for $25 \leq S_1, S_2 \leq 40$ and $-10 \leq T_1, T_2 \leq 30$ are given in Figure A.4 and A.5 respectively. Figure A.6 and A.7 contain the errors in these approximations.

In all the analyses of thermohaline circulation, the linear approximation Q_1 in (A.9) is invariably used. It should be interesting to evaluate the benefits of using more accurate second-order approximation.

Appendix B

Models for flow between boxes

By invoking the classical Poiseuille's equation for flow in a pipe, Stommel (1961) modeled the flow in the pipe connecting the E and P boxes (refer to Figure 2.1) as

$$kq = P_E - P_P \tag{B.1}$$

where q is the mass flow in $kg.sec^{-1}$, $P_E - P_P$ is the hydrostatic pressure difference between the E and P-boxes, and k is called the fluid resistance.

In the context of the box models, the pressure difference is induced by the density difference and hence

$$P_E - P_P = Hg(\rho_E - \rho_P) \tag{B.2}$$

where ρ_E and ρ_P are the densities of well mixed seas water in E and P boxes respectively, H is the common height of the boxes and g is the acceleration due to gravity. Substituting (B.2) in (B.1) and dividing both sides by ρ_o, we obtain the volumetric flow

$$\psi = \frac{q}{\rho_o} = \gamma\Delta\rho \tag{B.3}$$

where

$$\Delta\rho = \frac{\rho_E - \rho_P}{\rho_o} \qquad \text{and} \qquad \gamma = \frac{Hg}{k} \tag{B.4}$$

ψ and γ are in units of m^3sec^{-1}.

Referring to equations (A.8) – (A.9), we readily obtain (with $T = T_e - T_p$ and $S = S_e - S_p$) a first-order approximation

$$\Delta\rho = -\alpha T + \beta S \tag{B.5}$$

where $\alpha = 0.1167$ and $\beta = 0.7930$.

The list of different choices for the flow used in the dynamics

	Flow
Stommel (1961)	$q = \frac{\psi}{V}(sec^{-1})$
Cessi (1996)	$Q(\Delta\rho) = R_d + \frac{q}{V}(\Delta\rho)^2(sec^{-1})$
Van Veen (2001)	$f = R_d + \frac{q}{V}\|\Delta\rho\|(sec^{-1})$
Marotzke (1990)	$q = \frac{\gamma}{V}\|\Delta\rho\|(sec^{-1})$

Appendix C

Variation of density of pure water

Following the Appendix A in Gill (1982), the dependence of density $\rho(T)$ of pure water on temperature T is given by the 5th order polynomial

$$\rho(T) = \sum_{i=0}^{5} a_i T^i \tag{C.1}$$

where the values of the coefficients a_i are given in Table C.1. It can be verified that $\rho(T)$ attains its maximum value of 999.9750 very near $T_o = 4$℃.

Define the normalized density difference $\rho(T_s)$ as

$$\Delta\rho(T_s) = \frac{\rho(T_s) - \rho(T_d)}{\rho(T_d)} \tag{C.2}$$

where T_s and T_d are the temperature of the top and bottom layers of the two box model described in Figure 1. It is assumed that T_d is a constant and is set at $T_d = 2$℃.

Let x be the non-dimensional temperature defined by

$$x = \frac{T_s - T_d}{T_a - T_d} \quad \text{or} \qquad T_s = T_d + x(T_a - T_d) \tag{C.3}$$

Substituting (C.3) in (C.1) and (C.2), we get the density in terms of x as

$$\rho(T_s) = \bar{\rho}(x) = \sum_{i=0}^{5} a_i [T_d + x(T_a - T_d)]^i \tag{C.4}$$

and $\Delta\bar{\rho}(x)$ as

$$\Delta\bar{\rho}(x) = \frac{\bar{\rho}(x) - \rho(T_d)}{\rho(T_d)} \tag{C.5}$$

A plot of $\Delta\bar{\rho}(x)$ vs. x for $T_a = 11.5$ and $T_d = 2.0$ is given Figure C.1.

Table C.1. The coefficients of the polynomial in (C.1)

Coefficients	Value
a_0	999.842594
a_1	6.793952×10^{-2}
a_2	-9.095290×10^{-3}
a_3	1.001685×10^{-4}
a_4	-1.120083×10^{-6}
a_5	6.536330×10^{-9}

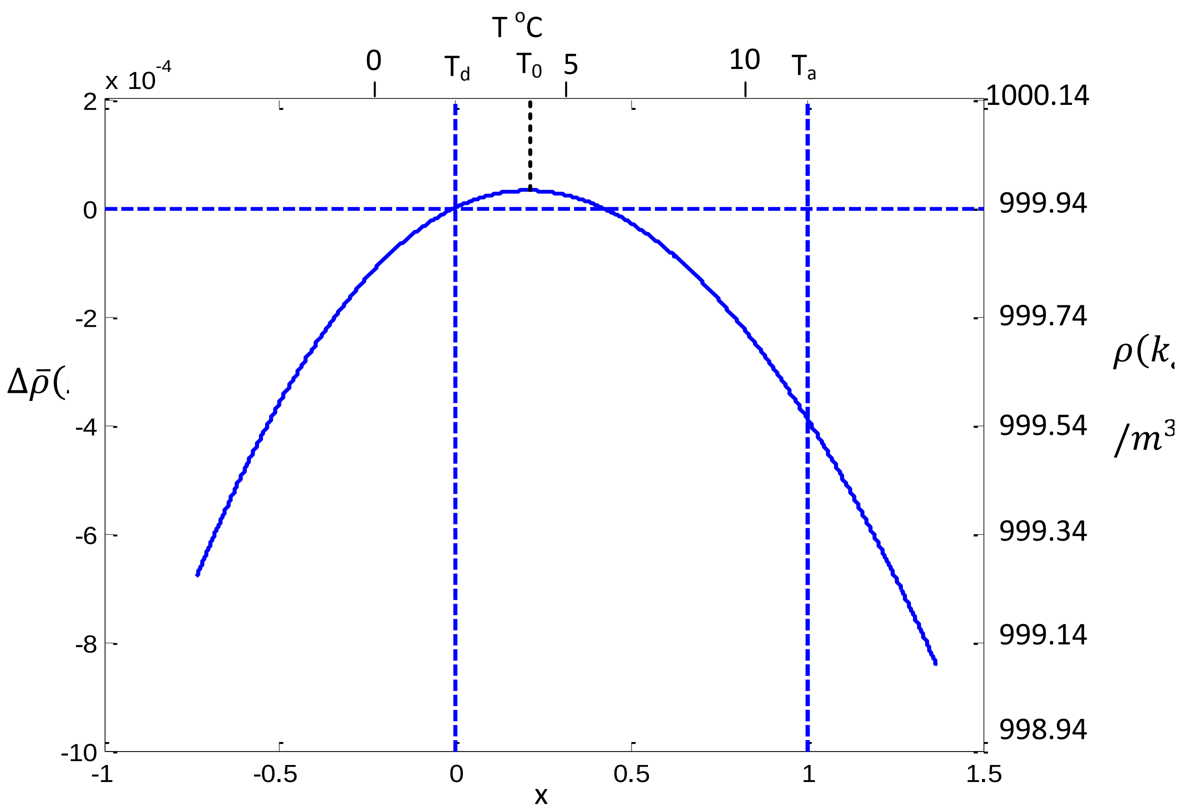


Figure C.1. The plot of $\Delta\bar{\rho}(x)$ vs. x

Appendix D

Piece-wise smooth dynamical system

This appendix provides a short summary of the last literature on the theory of dynamical systems with piece-wise smooth right hand side that is relevant to the analysis of this paper. We follow very closely the developments in the recent survey paper (Section 7) by Dieci and Lopez (2012) and a book (Chapter 2) by di Bernardo et. al. (2008).

Let $\mathbb{R}^n$ denote the state space of a dynamical system. Let $h: \mathbb{R}^n \to \mathbb{R}$ be a smooth function that defines a surface Σ in $\mathbb{R}^n$ given by

$$\Sigma = \{x \in \mathbb{R}^n | h(x) = 0\} \tag{D.1}$$

This surface divides $\mathbb{R}^n$ into two regions

$$R_1 = \{x \in \mathbb{R}^n | h(x) < 0\}$$

and

$$R_2 = \{x \in \mathbb{R}^n | h(x) > 0\} \tag{D.2}$$

where

$$\mathbb{R}^n = R_1 \cup R_2 \cup \Sigma \tag{D.3}$$

Let the state $x(t) \in \mathbb{R}^n$ evolve according to the piece-wise smooth dynamics given by

$$\dot{x} = \begin{cases} f_1(x) & if \quad x \in R_1 \\ f_2(x) & if \quad x \in R_2 \end{cases} \tag{D.4}$$

where $f_i: \mathbb{R}^n \to \mathbb{R}^n$ are the smooth vector (flow) fields defined in R_i, $i = 1, 2$.

Let $n(x) \in \mathbb{R}^n$ be the unit normal to Σ at $x \in \Sigma$ defined by

$$n(x) = \frac{h_x(x)}{\|h_x(x)\|} \tag{D.5}$$

where $h_x(x) \in \mathbb{R}^n$ is the gradient of $h(x)$ with respect to x. Let

$$P_i(x) = n^T(x) f_i(x) \tag{D.6}$$

be the inner product that defines the orthogonal projection of the flow field $f_i(x)$ onto $n(x)$. It is tacitly assumed that $f_i(x)$ is not tangential to Σ at the point x, that is, $P_i(x) \neq 0$ at $x \in \Sigma$. Then there are exactly two possibilities: each of the $P_i(x)$ is either positive or negative.

Assuming that the solution $x(t)$ reaches $\xi \in \Sigma$ in finite time, then it will enter the region R_2 (R_1) if it starts in R_1 (R_2) and $P_1(\xi) > 0$ (< 0) since $h\big(x(t)\big)$ would be increasing (decreasing) along the solution $x(t)$. Consequently, we get the following four cases as described in Figure D.1 that leads to the following classification.

1) Transversal intersection: The trajectory $x(t)$ is said to intersect Σ at ξ transversely if

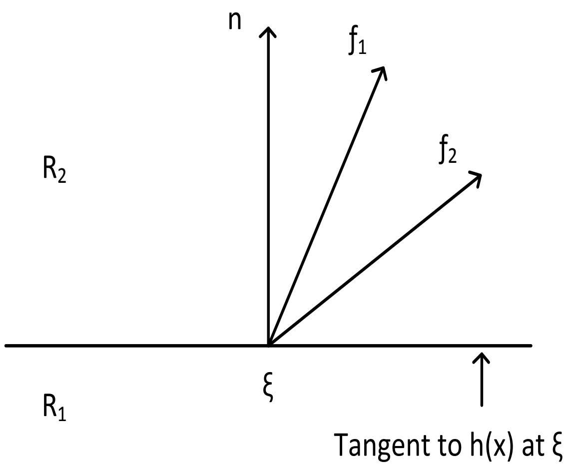
$$P_1(\xi) P_2(\xi) > 0 \tag{D.7}$$

If $P_i(\xi) > 0$ for $i = 1, 2$ as in Figure D.1 (a) the solution $x(t)$ starting from R_1 will cross into R_2. Similarly, If $P_i(\xi) < 0$ for $i = 1, 2$ as in Figure D.1 (b) the solution $x(t)$ starting from R_2 will cross into R_1.

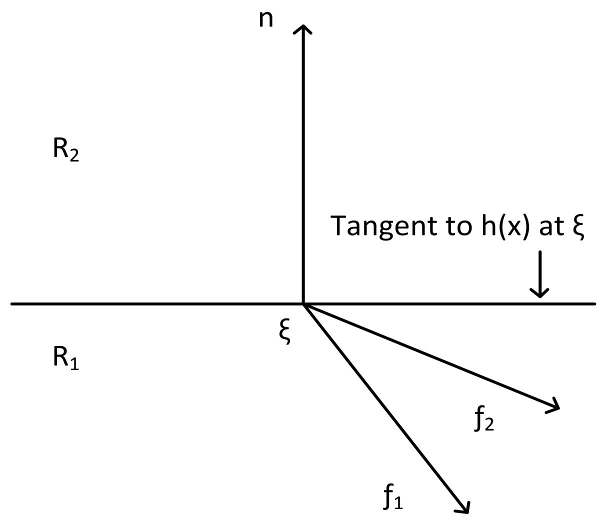
2) Sliding modes: This case is characterized by the condition

$$P_1(\xi) P_2(\xi) < 0 \quad \text{for } \xi \in \Sigma \tag{D.8}$$

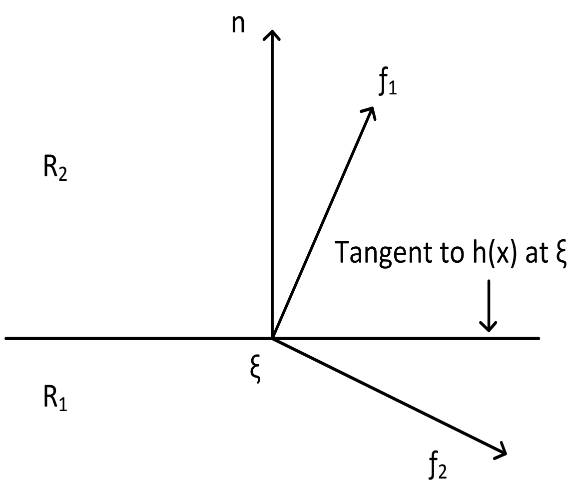
When $P_1(\xi) > 0$ and $P_2(\xi) < 0$ as in Figure D.1 (c), it is called attracting sliding mode, and it is called repulsive sliding if $P_1(\xi) < 0$ and $P_2(\xi) > 0$ as in Figure D.1 (d).



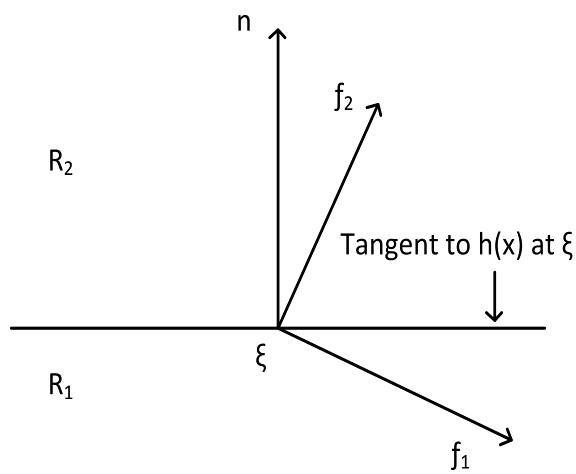
(a) $P_1(\xi) > 0, P_2(\xi) > 0$



(b) $P_1(\xi) < 0, P_2(\xi) < 0$



(c) $P_1(\xi) > 0, P_2(\xi) < 0$



(d) $P_1(\xi) < 0, P_2(\xi) > 0$

Figure D.1. Four possible cases. (a) and (b) define the transversal intersection, (c) and (d) define the sliding cases. Case (c) is called attractive and case (d) is called repulsive.

There are basically two approaches to defining the field at the point of discontinuity for the sliding modes.

Approach 1: In Utkin's equivalent control method (Utkin (1992)), we define the sliding vector field

$$f_U(x) = \frac{f_1(x) + f_2(x)}{2} - \frac{f_1(x) - f_2(x)}{2}\beta(x) \tag{D.9}$$

where

$$\beta(x) = \frac{h_x^T\big(f_1(x) + f_2(x)\big)}{h_x^T\big(f_1(x) - f_2(x)\big)} \tag{D.10}$$

is called the equivalent control.

Approach 2: In Fillipov's method (Fillipov (1988)) define the sliding vector field as the convex combination

$$f_F(x) = \big(1 - \alpha(x)\big)f_1(x) + \alpha(x)f_2(x) \quad \text{for } \xi \in \Sigma \tag{D.11}$$

where

$$\alpha(x) = \frac{h_x^T f_1(x)}{h_x^T\big(f_1(x) - f_2(x)\big)} \tag{D.12}$$

It can be verified that

$$\beta(x) = 2\alpha(x) - 1 \tag{D.13}$$

These two approaches guarantee the existence and the uniqueness of the solution of (D.1).

It is easy to verify that both f_U and f_F are orthogonal to $n(x)$ and hence along the tangent to Σ.

Finally consider a small δ neighborhood of the surface Σ on either side as shown in the Figure D.2. If $y_0 \in R_1$ lies in this neighborhood, then y_1 is decided by $f_1(y_0)$ and the time interval Δt used in integration. Then y_2 is decided by $f_2(y_1)$, and so on, refer to Figure D.2 for details. Then, as $\delta \to 0$, in the limit, $\alpha(x)$ defines the fraction of the time the trajectory spends in the region S_1 (Bernardo et. al. (2008)).

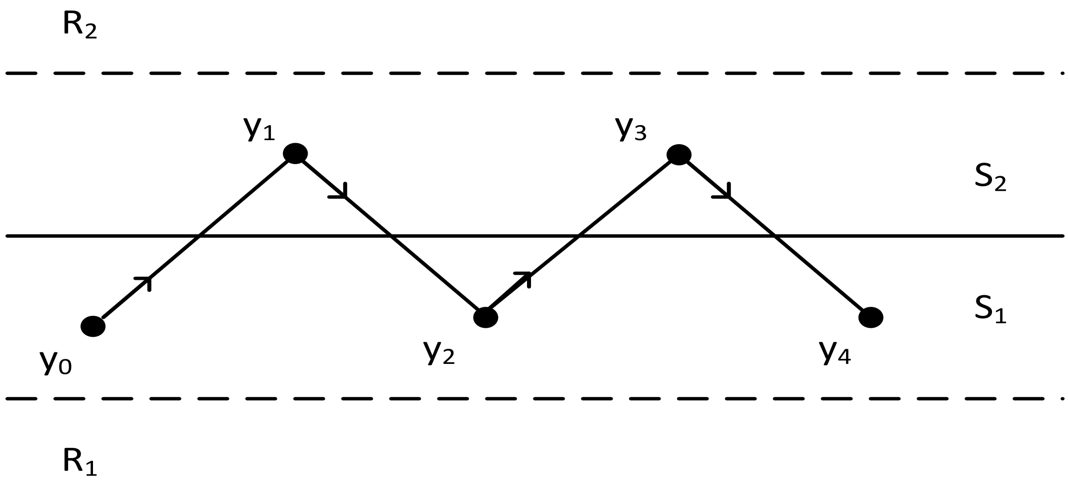


Figure D.2. Oscillation of trajectories in a small neighborhood S_1 and S_2 of the boundary Σ

Appendix E

Finite-time arrival of the solution at the discontinuity x_1

Consider the ODE with DRHS given by

$$\dot{x} = \begin{cases} f_1(x) = 1 - (1 + k_0)x & for \quad x < \overline{x_1}, or\ x > \overline{x_2} \\ f_2(x) = 1 - (1 + k_1)x & for \quad \overline{x_1} < x < \overline{x_2} \end{cases} \tag{E.1}$$

where recall that $\overline{x_1} \approx 0.0352$ and $\overline{x_2} \approx 0.3850$, with an initial condition $x(0) < \overline{x_1}$. In here, we have relabeled x_1 and x_2 in Figure 2 as $\overline{x_1}$ and $\overline{x_2}$ to avoid confusion with x_k, the discretized state of (E.1).

Discretizing (E.1) using a uniform time interval $\Delta > 0$, we get

$$x_{k+1} = \begin{cases} m_0 x_k + \Delta & if \quad x < \overline{x_1}, or\ x > \overline{x_2} \\ m_1 x_k + \Delta & if \quad \overline{x_1} < x < \overline{x_2} \end{cases} \tag{E.2}$$

where

$$m_0 = 1 - (1 + k_0)\Delta \quad \text{and} \quad m_1 = 1 - (1 + k_1)\Delta \tag{E.3}$$

Since $k_0 < k_1$, setting $\Delta < \frac{1}{1+k_1}$, it follows that $0 < m_1 < m_0 < 1$.

Iterating the first equation in (E.2), we get

$$x_k = m_0^k x_0 + \Delta \sum_{i=0}^{k-1} m_0^i \tag{E.4}$$

Hence

$$x^* = \lim_{k \to \infty} x_k = \Delta \sum_{i=0}^{\infty} m_0^i = \frac{\Delta}{1 - m_0} = \frac{1}{1 + k_0} \tag{E.5}$$

Now referring to Case 1 in Table 3, we have $k_0 = 0$ and $k_1 = 35$. Hence, $x^* = 1$ is the equilibrium E_3 in this case. Consequently, if we continue integrating the first equation in (E.2), assuming that there are discontinuities at $\overline{x_1}$ and $\overline{x_2}$, we find that x_k increases monotonically from $x(0) < \overline{x_1}$ to the value $x^* = 1 > \overline{x_1}$. Thus, on its journey from $x(0)$ to x^*, the solution crosses the value $\overline{x_1}$ at some finite time, k^*.

We can indeed quantity k^*. To this end recall that $k_0 = 0$ and $k_1 = 35$ for Case 1 in Table 3. Further, in this case $m_0 = 1 - \Delta$. Hence, from (E.4), it follows that

$$\begin{aligned} x_{k^*} = \overline{x_1} &= m_0^{k^*} x_0 + \Delta \sum_{i=0}^{k^*-1} m_0^i = m_0^{k^*} x_0 + \Delta \frac{1 - m_0^{k^*}}{1 - m_0} \\ &= m_0^{k^*}(x_0 - 1) + 1 \end{aligned} \tag{E.6}$$

Hence,

$$k^* = \frac{\log(1 - \overline{x_1}) - \log(1 - x_0)}{\log m_0} \tag{E.7}$$

For the illustration in Figure 11, $x(0) = -1$, $\overline{x_1} \approx 0.0352$, and $\Delta = 1 - m_0 = 0.0001$, the value of $k^* \approx 7290$, which corresponds to $t \approx 0.7290$.

When the solution $x_k < \overline{x_1}$, the one step increment is given by

$$x_{k+1} - x_k = (m_0 x_k + \Delta) - x_k = \Delta - x_k(1 - m_0) \tag{E.8}$$

Similarly, when $x_k > \overline{x_1}$, the one step increment is given by

$$x_{k+1} - x_k = (m_1 x_k + \Delta) - x_k = \Delta - x_k(1 - m_1) \tag{E.9}$$

Setting $x_k = \overline{x_1} - \delta$ in (E.6) and $x_k = \overline{x_1} + \delta$ in (E.7), it can be verified that

$$\Delta - (\overline{x_1} - \delta)(1 - m_0) > \Delta - (\overline{x_1} + \delta)(1 - m_1) \tag{E.10}$$

or equivalently

$$(\overline{x_1} - \delta)(1 - m_0) < (\overline{x_1} + \delta)(1 - m_1) \tag{E.11}$$

Since

$$\frac{\overline{x_1} - \delta}{\overline{x_1} + \delta} < \frac{1 - m_1}{1 - m_0} = \frac{1}{1 + k_0} \leq 1 \tag{E.12}$$

it follows that the one step upward increment for iterates bellow $\overline{x_1}$ are larger than the one step downward increment for iterates above $\overline{x_1}$.

Appendix F

Steady-state condition of Roebber's model

$$\frac{1}{2}q(T_2 - T_3) + K_T\left(T_{A_1} - T_1\right) - K_Z(T_1 - T_3) = 0 \tag{F.1}$$

$$\frac{1}{2}q(T_3 - T_1) + K_T\left(T_{A_2} - T_2\right) - K_Z(T_2 - T_3) = 0 \tag{F.2}$$

$$\frac{1}{2}q(T_1 - T_2) + K_Z(T_1 - T_3) + K_Z(T_2 - T_3) = 0 \tag{F.3}$$

$$\frac{1}{2}q(S_2 - S_3) - K_Z(S_1 - S_3) - Q_s = 0 \tag{F.4}$$

$$\frac{1}{2}q(S_3 - S_1) - K_Z(S_2 - S_3) + Q_s = 0 \tag{F.5}$$

$$\frac{1}{2}q(S_1 - S_2) + K_Z(S_1 - S_3) + K_Z(S_2 - S_3) = 0 \tag{F.6}$$

$$q = \frac{\mu(\rho_1 - \rho_2)}{\rho_0} = \lambda[(T_2 - T_1) - (S_2 - S_1)] \tag{F.7}$$

$$\lambda = \frac{\mu\alpha T_0}{q^*} \tag{F.8}$$

Subtracting (F.1) from (F.2), we get

$$\frac{1}{2}q(T_1 + T_2) - qT_3 + K_T\left(T_{A_1} - T_{A_2}\right) + (K_T + K_Z)(T_2 - T_1) = 0 \tag{F.9}$$

From (F.3), we obtain

$$\frac{1}{2}q(T_1 - T_2) + K_Z(T_1 + T_2) - 2K_Z T_3 = 0 \tag{F.10}$$

$$T_3 = \frac{1}{4K_Z}q(T_1 - T_2) + \frac{1}{2}(T_1 + T_2) \tag{F.11}$$

Substituting (F.11) into (F.9):

$$\frac{1}{2}q(T_1+T_2)-q\left[\frac{1}{4K_Z}q(T_1-T_2)+\frac{1}{2}(T_1+T_2)\right]+K_T\left(T_{A_1}-T_{A_2}\right) \\ +(K_T+K_Z)(T_2-T_1)=0 \tag{F.12}$$

$$(T_2-T_1)\left(\frac{1}{4K_Z}q^2+K_T+K_Z\right)+K_T\left(T_{A_1}-T_{A_2}\right)=0 \tag{F.13}$$

$$T_2-T_1=\frac{K_T\left(T_{A_2}-T_{A_1}\right)}{\frac{1}{4K_Z}q^2+K_T+K_Z} \tag{F.14}$$

Subtracting (F.4) from (F.5), we get

$$\frac{1}{2}q(S_1+S_2)-qS_3+K_Z(S_2-S_1)-2Q_s=0 \tag{F.15}$$

From (F.3), we obtain

$$\frac{1}{2}q(S_1-S_2)+K_Z(S_1+S_2)-2K_ZS_3=0 \tag{F.16}$$

$$S_3=\frac{1}{4K_Z}q(S_1-S_2)+\frac{1}{2}(S_1+S_2) \tag{F.17}$$

Substituting (F.17) into (F.15):

$$\frac{1}{2}q(S_1+S_2)-q\left[\frac{1}{4K_Z}q(S_1-S_2)+\frac{1}{2}(S_1+S_2)\right]+K_Z(S_2-S_1)-2Q_s=0 \tag{F.18}$$

$$(S_2-S_1)\left(\frac{1}{4K_Z}q^2+K_Z\right)-2Q_s=0 \tag{F.19}$$

$$S_2-S_1=\frac{2Q_s}{\frac{1}{4K_Z}q^2+K_Z} \tag{F.20}$$

Substituting (F.14) and (F.20) into (F.7):

$$q=\lambda\left[\frac{K_T\left(T_{A_2}-T_{A_1}\right)}{\frac{1}{4K_Z}q^2+K_T+K_Z}-\frac{2Q_s}{\frac{1}{4K_Z}q^2+K_Z}\right] \tag{F.21}$$

$$q\left(\frac{1}{4K_Z}q^2 + K_T + K_Z\right)\left(\frac{1}{4K_Z}q^2 + K_Z\right) - \lambda K_T\left(T_{A_2} - T_{A_1}\right)\left(\frac{1}{4K_Z}q^2 + K_Z\right) + 2\lambda Q_s\left(\frac{1}{4K_Z}q^2 + K_T + K_Z\right) = 0 \tag{F.22}$$

$$q^5 + 4K_Z(K_T + 2K_Z)q^3 + 4\lambda K_Z\left[2Q_s - K_T\left(T_{A_2} - T_{A_1}\right)\right]q^2 + 16K_Z^3(K_T + K_Z)q + 32\lambda K_Z^2(K_T + K_Z)Q_s - 16\lambda K_Z^2 K_T\left(T_{A_2} - T_{A_1}\right) = 0 \tag{F.23}$$