

Linköping Studies in Science and Technology
Thesis No. 1107

System Identification of Nonlinear Thermochemical Systems with Dynamical Instabilities

Gunnar Cedersund

Department of Electrical Engineering
Linköpings universitet, S-581 83 Linköping, Sweden
Linköping 2004

**System Identification of Nonlinear Thermochemical Systems with
Dynamical Instabilities**

© 2004 Gunnar Cedersund

gunnar@isy.liu.se
<http://www.fs.isy.liu.se/>
Department of Electrical Engineering,
Linköpings universitet,
S-581 83 Linköping,
Sweden.

ISBN 91-7373-996-0
ISSN 0280-7971
LiU-TEK-LIC-2004:36

Abstract

Thermochemical systems appear in applications as widespread as in combustion engines, in industrial chemical plants and inside biological cells. Science in all these areas is going towards a more model based thinking, and it is therefore important to develop good methods for system identification, especially fit for these kind of systems. The presented systems are described as nonlinear differential equations, and the common feature of the models is the presence of a boundary between an oscillating and a non-oscillating region, i.e. the presence of a bifurcation.

If it is known that a certain input signal brings the system to a bifurcation manifold, and this is the case for many thermochemical systems, this knowledge can be included as an extra constraint in the parameter estimation. Except for special cases, however, this constraint can not be obtained analytically. For the general case a reformulation, adding variables and equally many constraints, have been done. This formulation allows for efficient use of standard techniques from constrained optimization theory. For systems with large state spaces the parameter vector describing the initial state becomes big (sometimes > 1000), and special treatment is required. New theory for such treatment have been shown, and the results are valid for systems operating close to a Hopf bifurcation. Through a combined center manifold and normal form reduction, the initial state is described in minimal degrees of freedom. Experiment designs are presented that force the minimal degrees of freedom two be 2 or 3, independently of the dimension of the state space. The initial state is determined by solving a sub-problem for each step in the ordinary estimation process. For systems starting in stationary oscillations the normal form reduction reveals the special structure of this sub-problem. Therefore it can be solved in a straight-forward manner, that does not have the problem of local minima, and that does not require any integration of the differential equations. It is also shown how the knowledge, coming from the presence of a bifurcation, can be used for model validation. The validation is formulated as a test quantity, and it has the benefit that it can work also with uncalibrated sensors, i.e. with sensors whose exact relation to the state variables is not known.

Two new models are presented. The first is a multi-zonal model for cylinder pressure, temperature and ionization currents. It is a physically based model with the main objectives of understanding the correlation between the ionization curve and the pressure peak location. It is shown that heat transfer has a significant effect on this relation. It is further shown that the combination of a geometrically based heat transfer model and a dynamical NO-model predicts the correct relationship between the pressure and ionization peak location within one crank angle degree. The second developed model is for the mitogenic response to insulin in fat cells. It is the first developed model for this specific pathway and the model has been compared and estimated to experimental data. Finally, a 16-dimensional model for activated neutrophils has been used to generate virtual data, on which the presented methods have been applied, and on which the performance of the methods were demonstrated.

Preface

This work has been carried out at the department of Electrical Engineering, division of Vehicular Systems, Linköpings universitet, Sweden and has been financially supported by the Foundation for Strategic Research and Nationellt FlygForskningsProgram. It is my pleasure to say that the finishing of this report marks my being halfway accomplished with my Ph.D. studies.

Since I intend to stay within the academic world also after my Ph.D. time, I consider the time I am now spending a preparation period for the years to come. There are a few abilities I consider very important to develop during this preparation time, and that I would like to mention here. The first is the ability to find, and follow through, sound, long-term, research questions, and the second thing is to be able to do this, at least partly, through others - by collaboration and through supervision. During this period I have had the opportunity to experience supervision both from the perspective of the supervised, and, through all the master students I have had, from the perspective of the supervisor. I have therefore had plenty of opportunities to see a lot of both the potential of well-functioning supervision, as well as the dangers of supervision being non-functioning. All these experiences will become great wisdoms for the years to come, and to all of these experiences, and towards all the people that have been involved in them, I feel deep gratefulness for all the treasures they have shown me.

Another thing I consider important to learn during this period is scientific writing. For the finishing of this thesis I have had the possibility of getting input from quite many people, and there are a few I would like to especially mention. First I would like to thank my supervisor, Lars Nielsen, for all good general inputs of what distinguishes good science from less good science, and especially during the final structuring of this material I have found the interactions very fruitful. Secondly, I would like to thank my former supervisor Carsten Knudsen, with whom I am still collaborating, for the continued interactions that we have had. Especially during the early phases of the development of new ideas, and for help with technical details regarding nonlinear dynamics, the help has been very valuable. Further there are some people that have contributed with fruitful inputs to specific parts of the thesis, and that has definitely contributed to making it a better work (and not in the least, by being good friends). These people include Lars Eriksson, Peter Strålfors, Erik Frisk, Lars Åslund, Marcus Klein and Mattias Krysander.

I would also like to express my gratitude to all people that are a part of my life, but not a part of my research life. I have always had a strong urge to maintain my balance by doing widely different things, meeting people with diverse background, and in totally different ways. In that way I can experience more aspects of Who I Really Am, and for the love of all this people, expressed to me in this way, I am also deeply grateful.

Finally I would like to come back to my view of why we are working at all, with anything, and my view of how it most beneficially can be done. I will do it by restating a part of the wonderful poem 'the Prophet' by Kahlil Gibran:

You have been told also life is darkness, and in your weariness you echo what was said by the weary.

And I say that life is indeed darkness save when there is urge.

And all urge is blind save when there is knowledge.

And all knowledge is vain save when there is work;

And all work is empty save when there is love.

And when you work with love you bind yourself to yourself, and to one another, and to The All

This thesis is a part of my work.

Linköping, May 2004

Gunnar Cedersund

Contents

1	Introduction	1
1.1	Background	1
1.2	Outline and contributions	6
2	Theoretical Background	9
2.1	Notations	9
2.2	System identification	10
2.3	Bifurcations	13
2.3.1	Some simple one-dimensional bifurcations	15
2.3.2	The Hopf bifurcation	17
2.4	The Brusselator	20
3	Model Simplification around a Hopf Bifurcation	23
3.1	Center manifold and normal form theory	23
3.2	A combined transformation	28
3.2.1	For a general bifurcation	28
3.2.2	Derivation for the Hopf bifurcation	33
3.3	Extensions of existing theory	37
3.3.1	Extension to fifth order	38
3.3.2	Extensions to more than one parameter	38
3.4	Transformation of the Brusselator	39
3.5	Error analysis	43
3.6	Conclusions	46

4	Improved Parameter Estimation	51
4.1	The problem	52
4.2	Starting the estimation	52
4.2.1	Reformulating the problem	53
4.2.2	Finding a feasible point	54
4.3	The first optimization method - by elimination	56
4.3.1	General description	56
4.3.2	The algorithm	58
4.3.3	Properties of the algorithm	59
4.4	The second optimization method - a reduced gradient method	60
4.4.1	General description	60
4.4.2	The algorithm	61
4.4.3	Convergence properties	62
4.5	Identification of the Brusselator	62
4.6	Conclusions	68
5	Experimental Designs for Estimating the Initial Value Parameters	69
5.1	Simplifications by starting in a steady state	70
5.2	Simplifications close to a Hopf bifurcation	72
5.3	The complete algorithm	74
5.3.1	Properties of the algorithm	76
5.4	Quenching data	76
5.4.1	What is quenching data?	77
5.5	Simplifying the initial value sub-problem	78
5.6	Finding the initial values in the Brusselator	82
5.7	Conclusions	84
6	Qualitative Model Validation	87
6.1	Detecting erroneous estimations	88
6.2	The test quantity	89
6.2.1	The idea	89
6.2.2	Constant inputs	90
6.2.3	Time-varying inputs	94
6.3	Validation of the Brusselator	97
6.3.1	The two problems	97
6.3.2	Applying the QTQ algorithm	98
6.4	Conclusions	99
7	A Multi-zonal Model for Pressure, Temperature and Ionization Currents	101
7.1	Introduction	102
7.1.1	Basic engine concepts	102
7.1.2	The importance of predicting the Pressure Peak Location	103
7.1.3	Using the spark plug as an ionization sensor	105
7.2	The model	108

7.2.1	The multi-zonal combustion model	108
7.2.2	Addition of geometry and heat transfer	110
7.2.3	The Mass-flow	112
7.2.4	The Saitzkoff-Reinmann model	114
7.2.5	Dynamical NO concentration	115
7.3	Model evaluation	117
7.3.1	The experimental data	117
7.3.2	Validation of the model without heat transfer	118
7.3.3	Addition of heat transfer	119
7.3.4	Model sensitivity analysis	121
7.3.5	Addition of Dynamical NO-model	128
7.4	Conclusions	128
8	A Model for Insulin Signalling to Map-kinase Control in Fat Cells	131
8.1	Background	131
8.1.1	The mitogenic response to insulin in fat cells	132
8.1.2	Chemical kinetics revisited	133
8.2	The model	134
8.2.1	The first step, including insulin, IR and IRS	134
8.2.2	The second step, including Sos and Grb2	135
8.2.3	The third step, including GDP and Ras	137
8.2.4	The fourth step, the MAPK-cascade	139
8.3	In silico experiments with the model	141
8.3.1	Time-course analysis	141
8.3.2	Dose-response curves	141
8.3.3	Search for experimentally obtainable oscillations	143
8.4	Experimental verification	146
8.5	Parameter estimation	147
8.6	Discussion	149
8.7	Conclusions	150
9	Utilizing the Newly Developed Techniques on a Full-scale Model	153
9.1	The spatio-dynamics of neutrophils	153
9.2	The Olsen model	154
9.2.1	A bifurcation analysis	156
9.3	Compartmentalization	158
9.3.1	The compartmentalized models	160
9.3.2	Simulated waves	162
9.4	Model simplification	164
9.5	Forming a qualitative test quantity	166
9.6	A parameter estimation example with simulated data	168
9.7	Conclusions	170
10	Summary and Conclusions	173

A Sufficient Information for Regeneration of all Plots	177
A.1 From Chapter 3	177
A.2 From Chapter 4	178
A.3 From Chapter 5	178
A.4 From Chapter 6	179
A.5 From Chapter 7	179
A.5.1 Creation of a new zone	180
A.6 From Chapter 8	180
A.7 From Chapter 9	182
References	187

Introduction

1.1 Background

Both the thrill and the difficulty with studying thermochemical systems, is that it requires knowledge from so many diverse disciplines. Thermochemical systems also appear in a wide range of applications, and this makes the study even more interesting. This thesis is devoted to *system identification of thermochemical systems*, with the main focus on *model construction* and *parameter estimation*. Here in the beginning, to introduce the whole subject, a few examples of different thermochemical systems are given. When reading the examples it will become clear that they all share a certain type of nonlinearity. The feature they all share is the possibility for a transition between a stable stationary state and a stable oscillating state, i.e. the presence of a *dynamic instability*. The transition between the states will be caused by variations in choice of operating point, change in outer conditions or by the choice of control algorithm. This type of nonlinearity is commonly occurring in thermochemical systems, and it is the main feature pursued in this thesis.

Engine knock

The chemical energy contained in fuel injected in the cylinder of a combustion engine is transformed into heat and mechanical work during the combustion phase. This process usually takes less than 0.01s, and it is a highly non-equilibrium phenomenon (Heywood, 1988). There is, however, a limit when the process becomes too non-equilibrium, and this is when there is no longer a continuous spreading of a laminar flame, but an almost instantaneous auto-ignition of all unburned fuel-air

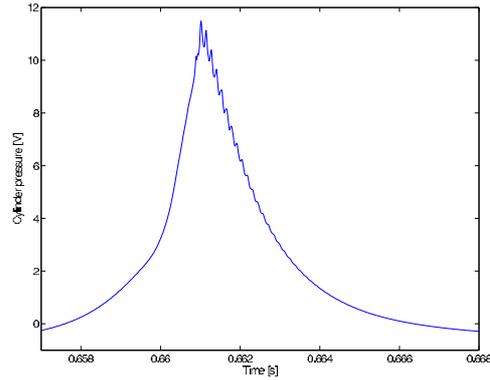


Figure 1.1 *Plot showing the oscillations in the intra-cylinder pressure during engine knock. It is the small oscillations superposed on the slow variation that are due to the knock. If the knock oscillations become too big, they are highly dangerous to the engine. However, an engine usually increases in efficiency as one approaches the knock conditions, and therefore knock causes a trade-off between engine safety and efficiency. In series produced engines today a cylinder pressure sensor is not available, and that is another reason to develop good models for the cylinder pressure that are efficient at these critical points.*

mixture (Heywood, 1988). This phenomenon takes place when the pressure and temperature inside the cylinder become too high, and auto-ignition of this kind might lead to acoustic waves, travelling inside the cylinder. These pressure waves can make a little knocking sound and are therefore called engine knock. They can be seen on a pressure plot (see Figure 1.1), and if they oscillations are big they can be severely damaging to the engine. This danger is weighted against engine efficiency and performance, since and for most operating situations the engine has the highest performance just at the onset of these instabilities. Therefore it is of high importance to understand these instabilities, and to develop an understanding for how they can best be treated. For combustion engines in production today there are no pressure sensors showing the oscillations directly (like in Figure 1.1), and therefore other sensors has to be used. One sensor that has this potential is the ionization sensor, and a model describing the relationship between the ionization current, that the sensor can measure, and the cylinder pressure, is developed in this thesis.

Other technical systems with combustion instabilities

The dynamical instability in the combustion engine explained above, is a special case of a general feature, with the potential of appearing in most combus-

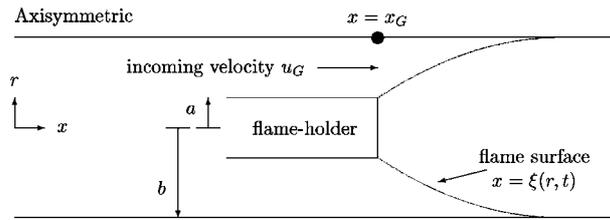


Figure 1.2 One of the simplest set-ups that shows dynamical instabilities in a combustion system.

tion processes, occurring in closed combustion chambers (Fichera, Losenno and Pagano, 2001). There are also many other technical applications where these limit cycles appear, and in many of these cases they are also considered as problematical. One of the simplest examples where these oscillations occur is shown in Figure 1.2. Further, in Figure of a rocket engine, where these oscillations also occur, is shown in Figure 1.3. There are models that describe these oscillations, and also some attempts at dampening them by applying control algorithms (Zaida and Graf, 1998).

Biological instabilities

Life is essentially a non-equilibrium phenomena, and in nature everything oscillates, or at least is affected by something that oscillates. In nature it is therefore more a question of whether there are oscillations at a specific time-scale than whether there are oscillations at all, in a specific system. One major cycle that goes through the entire biological system is the circadian rhythm, the rhythm of the day, but there are many other time-scales. Examples of larger time-scale are the lunar, the year cycles, and the several-year cycles. The latter appear in e.g. predator-prey systems (Sturis and Knudsen, 1996). Examples of shorter time-scales are the oscillations of the blood glucose level, at the time-scale of hours (Sturis, Knudsen, O'Meara, Thomsen, Mosekilde, Vauter and Polonsky, 1995), and the metabolic oscillations governing the glycolysis, at the time-scale of seconds (Hynne, Danø and Sørensen, 2001). There are many examples of more complicated instabilities, like bursting and chaotic attractors (see e.g. Figure 1.4) that appear as natural operating points. Often subsystems with such complex phenomena are also coupled to each other. An example of coupled bursting systems are the coupled β cells (Sørensen, Petersen and Aslanidi, 2004), and coupled nephrons is an example of coupled chaotic attractors (Holstein-Rathlou, Yip, Sosnovtseva and Mosekilde, 2001). For such coupled systems the situation is even more complex. Out of

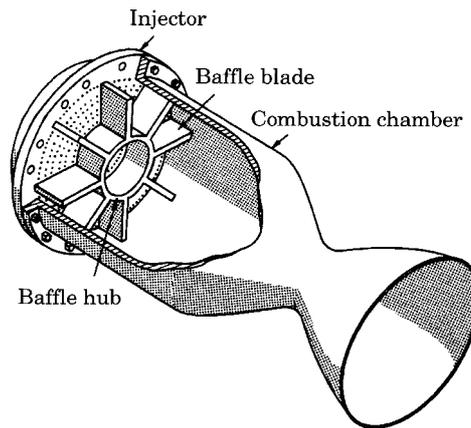


Figure 1.3 *A rocket engine of this kind is one of the many technical applications that show sustained thermochemical oscillations.*

all of this, two questions often comes to ones mind.

- Why is nature so often choosing oscillating, and sometimes even more complex attractors as operating points; why is it beneficial?
- Are there things here that we, as engineers, could learn when constructing our own control strategies, or when choosing our own operating points?

Today, the onset of oscillations, or even worse, the onset of chaotic or bursting oscillations are generally avoided in control strategies. This is because these points, and the states that lie beyond them, are denoted as unstable states, and unstable states are considered as non-attractive when building control strategies. Therefore, if the two questions above could be positively answered, this might affect our whole view of how to construct human-made systems.

Chemical reactors

Apart from the technical thermochemical systems mentioned above, chemical plants consists another big and important class of applications that falls within the class of thermochemical systems. Chemical plants are production systems where the ingoing components, as well as the desired outcome, consists of chemicals or substrates. There are also biochemical reactors, where one uses real biological components, like e.g. cells, to govern most of the reactions, and the only things that in those cases needs to be regulated by the operator are the conditions for the cells. An example of the latter is the production of beer (Nguyen, 2004), and there the yeast

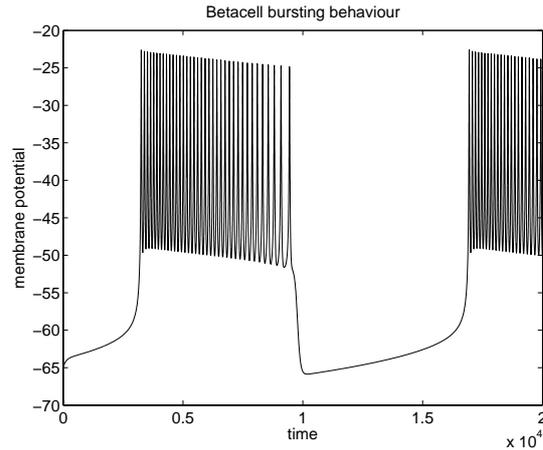


Figure 1.4 *Plot showing the bursting behaviour of β cells. In the plot one and a half burst is shown. When the cells are secreting insulin to the blood, these bursts must be present. Hence it has become clear that this complex behaviour is the stationary state that β cells operate in when they are active. A big question is why nature has found it beneficial to work around such a complex attractor. Another question is whether we also could benefit from developing control strategies for our human-made systems, by letting them operate around these kind of complex attractors. These two questions are two of the main questions of this thesis.*

cells used for the production, display oscillations in many of the chemicals (Hynne et al., 2001). However, most chemical plants that are human controlled altogether are kept at a steady state operating point, but as argued above, this might be a non-optimal way of control, and a strategy that might eventually be replaced by others. Finally, it should be said that much of the early developments of model predictive control was done for chemical reactors. The reason for that was slow time-scales of these systems, that allowed for computer simulations on-line. For the same reason, chemical reactors might be one of the prime targets when trying to apply the methods developed in this thesis to real industrial examples.

Identification of nonlinear systems

This thesis will not study control of thermochemical systems, but will stay at the level of system identification and diagnosis. However, control theory and system identification are subjects that are closely linked to each other, and if one can develop good models for a system, that will be beneficial also for control problems.

System identification contains everything that has to do with systemizing the available knowledge of a system into a useful form. Usually this compact form will be a model and in this thesis, the model structure will always be a system of nonlinear ordinary differential equations (ODEs). The two parts of system identification that will be in main focus in this thesis are model construction and parameter estimation. The first part, model construction, is presented in the form of developed models (Chapters 7, 8 and 9). The second part, parameter estimation, has been attacked on the method level, and there are presented ways of improving the estimation, both during the experiment design phase (Chapter 5), during the estimation process (Chapters 4 and 5), and during the validation phase (Chapter 6). Nonlinear thermochemical systems, of the kind mentioned above, often give rise to too complex models for ordinary estimation methods to be directly applicable. Therefore improved methods, that are especially fit for these kind of systems are greatly desired. The common feature that has been utilized is the existence of a nonlinearity, or more specifically, the existence of a border between oscillating and non-oscillating states. Loosely, these borders are called bifurcations, and there is a rich theory of how the area around these borders can be described. The main results of this thesis are hence obtained by the application of bifurcation theory, to the problem of system identification of nonlinear thermochemical systems.

1.2 Outline and contributions

After this introductory chapter, there is one more background chapter, Chapter 2, that gives a general introduction to the theory used in the thesis. It also introduces most of the notation that will be used during the rest of the thesis. One of the main problems with accomplishing this thesis has been to combine the worlds of bifurcation theory and electro-mechanical engineering, and effort has been put into finding good notations. This in itself, is therefore the first non-trivial contribution of the thesis. Chapter 3 goes on with a review of center manifold, and normal form, theory, and with a review of a nice method that allows for easy calculations of these transformations. The second half of Chapter 3 then extends the theory and methods to include higher order terms in the transformation, and hence allows for improvements of the quantitative agreement between the original and the simplified system. The improvements are shown by application of the theory to a simple test case. The next three chapters, 4, 5, and 6, present new methods for improved parameter estimation. Chapter 4 presents two algorithms showing how one can include the information that a known input signal brings the system to a bifurcation. It is shown that these methods reduces the degrees of freedom describing the search space. An example, where this reduced degree of freedom makes the difference between a successful optimization and an unsuccessful is given. This improvement is one of the main contributions of the thesis. Chapter 5 presents theory, that allows the parameters describing the start state, x_0 , to be described by minimal (usually 2 or 3) degrees of freedom. It is presented experimental designs that make the given theory applicable. To understand the importance of

this simplification, one must remember that for large systems, and especially if they are also compartmentalized, the x_0 parameter vector can have dimensions between 15 and 1000. These improvements, that actually are a continuation of the developments of Chapter 3 and 4, make these results, according to the author, the contribution of the thesis with the highest potential. Chapter 6 presents a way to rule out false estimations, caused by either the convergence to an erroneous local minima, faults in model structure, or faults in one of the sensors that has been used for the estimation. It is shown how this knowledge can be utilized to form a qualitative test quantity, and this test quantity has the benefit that it can work also with uncalibrated, or non-modelled, sensors. This concludes the method development part of the thesis. Chapter 7 describes the first model that has been developed. It is a model for cylinder pressure, temperature and ionization currents. The main contribution of this model compared to previously published models is that it can predict the relationship between pressure and ionization peak location within one crank angle degree (CAD), while having a non-adiabatic combustion model. It is also shown why previously published models has had a characteristic delay of > 2 CAD, compared to measurement data. Chapter 8 contains another model developed, namely one for the intra-cellular response to insulin in human fat cells. This is the first model for this pathway that has been published, and it is shown to agree with published experimental data. A 16-dimensional model for activated neutrophils was chosen as a large-scale test-case for the new methods developed in the earlier chapters. This is the content of Chapter 9. Finally, in Chapter 10, the conclusions of the thesis are summarized, and in the appendix sufficient information to regenerate all the plots appearing in the thesis is given.

Theoretical Background

This chapter reviews some basic theory that has to do with representation and system identification of nonlinear systems in general, i.e. the theory on which Chapters 3, 4, 5, 6 and 9, are built. The chapter starts by introducing the notations that will be used to describe a system of nonlinear ordinary differential equations (ODEs). In Section 2.2, a short introduction to system identification is given, and Section 2.3 gives a short introduction to the theory of bifurcations. It is the combination of these two theories that will lead to some of the main results in the later chapters. Finally, in Section 2.4, the Brusselator is introduced. This system contains a simple bifurcation, and will serve as the first example for testing of new methods.

2.1 Notations

This thesis takes bifurcation theory from nonlinear dynamics and applies it to system identification. These two areas do not use the same notations, and the translation between two is non-trivial. For instance are bifurcations usually not defined for non-autonomous systems whereas control engineering systems deals almost exclusively with this kind of systems. This problem is treated in Section 2.3. Here is introduced the basic notations that will be used, and it has been chosen as a compromise between the notations at use in system identification literature and that in bifurcation literature. First we will introduce the notations we will use for a nonlinear ODE.

A nonlinear ODE

Let the state vector be denoted \mathbf{x} , and let the dimension of \mathbf{x} be denoted \mathbf{n} , i.e. $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$. Further, let the control signals be denoted \mathbf{u} . The vector \mathbf{u} is a function of time and its value is known and controllable. Assume that the other inputs to the system are constant, and included in the parameter vector \mathbf{p}_x . Let the dynamics of the system be governed by a nonlinear, smooth, function f . With these notations the system of differential equations is given by

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{p}_x, \mathbf{u}) \quad (2.1)$$

Sometimes it will be beneficial to consider the \mathbf{u} and \mathbf{p}_x vector together. Let this pair be denoted $\boldsymbol{\mu}$

$$\boldsymbol{\mu} = (\mathbf{p}_x, \mathbf{u})$$

Note that $\boldsymbol{\mu}$ is a time-varying vector, and that all the time-dependence lies in \mathbf{u} . Let the sensor signals be denoted \mathbf{y} . Assume that the sensor values are a function of $\boldsymbol{\mu}$, \mathbf{x} , and perhaps some additional parameters \mathbf{p}_y . Further, let the functional relationship for the vector signals be denoted $\mathbf{y}(\cdot)$

$$\mathbf{y} = \mathbf{y}(\mathbf{x}, \boldsymbol{\mu}, \mathbf{p}_y) \quad (2.2)$$

The time when the simulation, or when the time series intended for the identification starts, can be given or chosen as part of the experiment design. Let this start time be denoted t_{start} and let the state vector at this time be denoted $\mathbf{x}(t_{\text{start}})$. Let \mathbf{x}_0 be the parameter vector giving the start values

$$\mathbf{x}(t_{\text{start}}) = \mathbf{x}_0 \quad (2.3)$$

In most of the thesis we will have chosen the start time so that it is zero, i.e. $t_{\text{start}} = 0$. In those cases $\mathbf{x}(0) = \mathbf{x}_0$.

Equation (2.1), (2.2) and (2.3) fully specifies the system, and we now write these equations together for future reference. To fully specify which vectors are dependent on time, and which are not, the time-dependence is here explicit

$$\dot{\mathbf{x}}(t) = f(\mathbf{x}(t), \boldsymbol{\mu}(t)) = f(\mathbf{x}(t), \mathbf{p}_x, \mathbf{u}(t)) \quad (2.4a)$$

$$\mathbf{y}(t) = \mathbf{y}(\mathbf{x}(t), \mathbf{p}_x, \mathbf{u}(t), \mathbf{p}_y) \quad (2.4b)$$

$$\mathbf{x}(t_{\text{start}}) = \mathbf{x}_0 \quad (2.4c)$$

2.2 System identification

System identification contains everything that has to do with systemizing the available knowledge of a system into a useful form. Usually this compact form will be a model, and in this thesis the model will be a system of the general form (2.4). In this section is given a short introduction to the system identification loop, i.e. to the steps that are usually carried out when doing system identification. There are choices between different approaches and methods at the different steps, and here only a few of them are described. The overall presentation of this section is inspired by the book Ljung (1999).

The system identification loop

The available knowledge going into the identification loop usually consists of two parts: i) a set of experimental data Z^N , and ii) knowledge about the actual processes. The first step in physically based system identification process is to extract the knowledge about the way the processes work into the form of the equations. This form, or set of alternative forms, is called the *model structure* and at this initial step it contains uncertainties. These uncertainties take the form of undetermined parameters, and transitions between different forms are also formulated as variations in parameters. The idea is that these values shall be determined by the data set Z^N . Another component of the system identification procedure is therefore a *selection rule* that determines which specific parameter set that best captures all the information in the data set. Once the best model has been found it is time for the last step in the system identification procedure, the *validation step*. The validation step uses some data not included in the previous identification process, to see how well the chosen model predicts the data in this data set. If the chosen model does a sufficiently good job in this test, it is approved. If the model does not manage the validation step good enough, you have to start over at a previous step in the system identification loop (which might include doing new experiments).

The process for nonlinear ODEs

Now follows a more detailed description of the various steps in the system identification loop. At each step it will be described how the step can be carried out for the specific case of nonlinear ODEs. These various choices will later be referred to as the 'standard method' though it should be emphasized that there exist other methods, also applicable to this class. The specific model structure of equation (2.4) brings many difficulties, and at each step some of these difficulties are highlighted.

Step 1: Choosing model structure

Let the vector of all parameters be denoted \mathbf{p} . So far three possible parts of this parameter vector have been introduced: \mathbf{p}_x , \mathbf{p}_y and \mathbf{x}_0 . In many cases the only parameters that are interesting are the parameters \mathbf{p}_x , and the other parameters are only included to be able to form the objective function (i.e. to simulate the system etc.). For this reason it might be beneficial to introduce all parameters that are not the \mathbf{p}_x parameters. Let this vector be denoted \mathbf{p}_{nx} , i.e.

$$\mathbf{p} = (\mathbf{p}_x, \mathbf{p}_{nx}) \quad (2.5)$$

In this thesis \mathbf{p}_{nx} will have different compositions depending on the situation. Referring to equation (2.4) gives $\mathbf{p}_{nx} = (\mathbf{p}_y, \mathbf{x}_0)$, but in Chapter 4 and 5 \mathbf{p}_{nx} will also include n additional parameters, \mathbf{x}_b , which means that then $\mathbf{p}_{nx} = (\mathbf{p}_y, \mathbf{x}_0, \mathbf{x}_b)$. Let us now turn to the problem of forming a model structure based on these parameters.

For linear time-discrete systems the model structure is usually formulated as a predictor (Ljung, 1999),

$$\hat{\mathbf{y}}(t|\mathbf{p}) = \mathbf{g}^{\text{pred}}(\mathbf{p}, Z^{t-1}) \quad (2.6)$$

i.e. the knowledge of all the data up until the previous time is available, and useful, to predict the present output. For a nonlinear ODE it is not easy to use the outgoing data once the simulation has started. Therefore the Y part in, Z^{t-1} , disappears and g^{pred} is a function of the parameters, p , and the inputs U only. Such a model structure is generally referred to as an output error, or a *simulation model* (Ljung, 1999).

The true system is assumed to behave like

$$\dot{x}_0 = f(x_0, \mu_0) = f(x_0, p_{x,0}, u) \quad (2.7a)$$

$$y = y(x_0, \mu_0, p_{y,0}) + v \quad (2.7b)$$

$$x(t_{\text{start}}) = x_{0,0} \quad (2.7c)$$

where the dimensions and meanings of the ingoing variables are the same as for (2.4), except for the measurement noise, v . The subscript in e.g. p_0 denotes the true value of p . Finally the measured input is assumed to be true, i.e. $u = u_0$. This means that the true system is assumed to behave like the model structure, for a specific choice of parameters, and the only difference regards the measurements. One could consider the idea that the true system also has process noise but the implications of that is outside the scope of this thesis.

Step 2: The experimental data

The experimental data consists, for both linear and nonlinear model structures, of two parts: $Z^N = (U^N, Y^N)$. Here U^N contains the inputs $U^N = \{u(t)\}_{t=1}^N$, and Y^N contains the outputs $Y^N = \{y(t)\}_{t=1}^N$. Z^N is, however, generated during one or several experiments, and the important task of designing these to generate as informative time series as possible is highly dependent on both the chosen model structure and the chosen methods at the different steps in the identification loop. For the specific model structure of (2.4) and the methods suggested here, some important difficulties to consider are the following:

- i) Open simulation is done, i.e. the prediction does not use previous y s. Therefore continuously perturbing a system might not be the same good idea as it is for prediction error based methods.
- ii) Another major problem with system identification of large realistic nonlinear ODE's is that you have to estimate the initial values x_0 along with the desired parameters p_x and p_y . For large systems, and especially for compartmentalized systems, the extra degrees of freedom coming from the x_0 parameters is a serious problem, and often makes the problem unsolvable (Schittkowski, 2002). The increase in degrees of freedom also leads to worse identifiability of the desired parameters. The problems are mainly due to the fact that the objective function $V_N(p, Z^N)$ has many local minima for this particular class of model structures.

The second issue here deals with the problem of estimating the initial values, and argues that this problem is absolutely essential to solve, or at least diminish,

in order to do system identification of large nonlinear ODEs. With clever experimental design, these parameters can be constrained in various ways and in this thesis Chapter 5 is devoted to development of such methods.

Step 3: The selection criteria

The simplest selection criteria is to find the parameters that minimize the sum of the square of differences between the measured and the simulated output. This means that if the estimated parameter value is denoted \hat{p} , it can be identified in the following way:

$$\hat{p} = \arg \min_p V_N(p, Z^N) \quad (2.8)$$

where

$$V_N(p, Z^N) = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t|p))^2 \quad (2.9)$$

One could consider using other norms. The L_1 norm has the benefit that it will force many of the parameters to be zero, which is beneficial when identifying e.g. intra-cellular networks (Gustafsson, Hörnquist and Lombardi, 2003). Sometimes it might also be beneficial to put a harder emphasis on the agreement on some specific parts of the time series (Ljung, 1999). The finding of the minima is usually solved by an ordinary iterative optimization method, like e.g. the Levenberg-Marquardt (Nocedal and Wright, 1999). Another standard book on the subject of parameter estimation is Schittkowski (2002).

Step 4: The validation step

The validation step can be done in many ways. The simplest way is to say that the error function $V_N(\hat{p}, Z_{\text{validate}}^N)$ should be below a given value, but there are also other test (Ljung, 1999). Here Z_{validate}^N is a special part of the given time series that has not been used during the estimation phase. Another way of validating a system could be to see whether the qualitative behaviour of the system is as it should. This is the idea behind Chapter 6. Further, in Chapter 4, a way is presented that ensures that no such contradiction exists, for a specific class of qualitative behaviours, already at the selection phase. That means that with the different methods that are presented in this thesis, one can choose whether to implement new optimization algorithms, or whether to do an additional after-treatment, using the original algorithms. If the validation phase approves the estimated parameters, \hat{p} , the system identification process terminates. If not, one has to reconsider which of the previous steps can be done differently, and there re-enter the identification loop once again.

2.3 Bifurcations

Many of the results in this thesis involves the existence, and usually also the nearby presence, of a bifurcation. Usually the bifurcation will be of a special type: a super-critical Hopf bifurcation. It is therefore beneficial to, here in the beginning, make

clear what a bifurcation is in general and, more specifically, what a supercritical Hopf bifurcation is.

A bifurcation in general

The word bifurcation comes from a Latin word that means branch, and a branch is exactly what a bifurcation is. It is a branching between two topologically non-equivalent flows under the variation of the parameter μ in equation (2.4a). 'Topologically equivalent' is defined as follows (Wiggins, 1990).

Definition 2.1. *Two μ values, μ_1 and μ_2 , with corresponding fix points (x_1, μ_1) and (x_2, μ_2) , have topologically equivalent flows if there exists a homeomorphic mapping $\xi : U \rightarrow V$ that preserves orientation but not necessarily parametrization in time. Here U and V are neighborhoods of the first and second fix point, respectively.*

Two entire systems are topologically equivalent if their attractors (fix points, limit cycles, ...) are equally many and of the same stability. The attractors usually move or change appearance under all variations in μ , but when they change stability or have been created/disappeared, a bifurcation is said to have occurred. Since the attractors govern the qualitative long-term behaviour of the system, you might loosely say that a bifurcation manifold divides two qualitatively different regions in the state-parameter space. The exact formulation of a (local) bifurcation is the following (Wiggins, 1990).

Definition 2.2. *A fix point (x, μ) of a one-parameter family of one-dimensional vector fields is said to undergo a local bifurcation at $\mu = 0$ if the flow for μ near zero and x near zero is not topologically equivalent to the flow for x near zero and at $\mu = 0$.*

There are bifurcations which are not local. They are called global bifurcations and are outside the scope of this thesis. Finally let us specify what we will mean by a bifurcation when we have time-varying inputs.

Definition 2.3. *A fixed point $(x, \mu) = (x, p_x, u)$, with constant u , of a one-parameter family of one-dimensional vector fields is said to undergo a local bifurcation at $\mu = 0$ if the flow for μ near zero and x near zero is not topologically equivalent to the flow for x near zero at $\mu = 0$.*

As can be seen the only thing that has been added is the constraint that u should be constant. This does not mean that only constant u systems are considered but that a system with time-varying u is said to be in the same topological state at time t_i as the corresponding autonomous system would be with $u = u(t_i)$. With this definition we can now speak of topological states, and hence also of bifurcations, for non-autonomous systems as well. Now, let us briefly look at some simple one-dimensional bifurcations, before we go on to the Hopf bifurcation, which is the central bifurcation in this thesis.

2.3.1 Some simple one-dimensional bifurcations

Non-trivial bifurcations require that the function f in (2.4a) is nonlinear, but not that x is multi-dimensional. There are, on the contrary, some quite illustrative bifurcations that can occur already in one-dimensional systems. A nice feature from what is called the center manifold, and normal form, theory is that all generic features of a bifurcation can be extracted from its appearance in the simplest equations - the normal forms. This means that even though the three one-dimensional bifurcations introduced here can just as well occur in arbitrarily large system, for our purpose it is sufficient to study their one-dimensional normal forms. The three one-dimensional bifurcations that will be mentioned are called the *saddle-node* bifurcation, the *transcritical* bifurcation, and the *pitchfork* bifurcation, since they are the only simple¹ one-dimensional bifurcations.

The saddle-node bifurcation

The normal form of the saddle-node bifurcation is (2.10), and the corresponding bifurcation diagram is the top plot of Figure 2.1.

$$\dot{x} = \mu - x^2 \quad (2.10)$$

By inspection of the diagram, or by studying the equations, can be seen that, for positive values of μ , there are two solutions to the equation (2.10) with $\dot{x} = 0$. These are $x = \pm\sqrt{\mu}$ and hence they collide at $\mu = 0$ and this is where the bifurcation occurs. The bifurcation point $(x, \mu) = (0, 0)$ is a non-hyperbolic fix point (it has eigenvalues to the Jacobian with zero real part), as it is for all local bifurcations, and in this particular bifurcation the fix point ceases to exist once μ turns negative. Viewed from the other side, this is of course equivalent to the creation of two fix points out of the 'blue sky', and this fact has caused this bifurcation to sometimes be named *blue-sky* bifurcation. Another common name, which more has to do with the shape of the bifurcation diagram, is *fold* bifurcation (Strogatz, 1994).

The transcritical bifurcation

The normal form for the transcritical bifurcation is equation (2.11) and the corresponding bifurcation diagram is the bottom plot of Figure 2.1.

$$\dot{x} = \mu x - x^2 \quad (2.11)$$

As can be seen, either from looking at the equations, or from Figure 2.1, for this bifurcation there are two fix points on both sides of the bifurcation point $(x, \mu) = (0, 0)$. One fix point is at $x = 0$ and one at $x = \mu$, and what happens at the bifurcation is that these exchange stability. Compared to e.g. the saddle node-bifurcation this is a relatively harmless bifurcation.

¹Here the real term for simple bifurcation is co-dimension 1 bifurcation. When accepting bifurcations of higher co-dimension there can be arbitrarily many types, already for one-dimensional systems. These bifurcations are, however, much more rarely appearing in real systems.

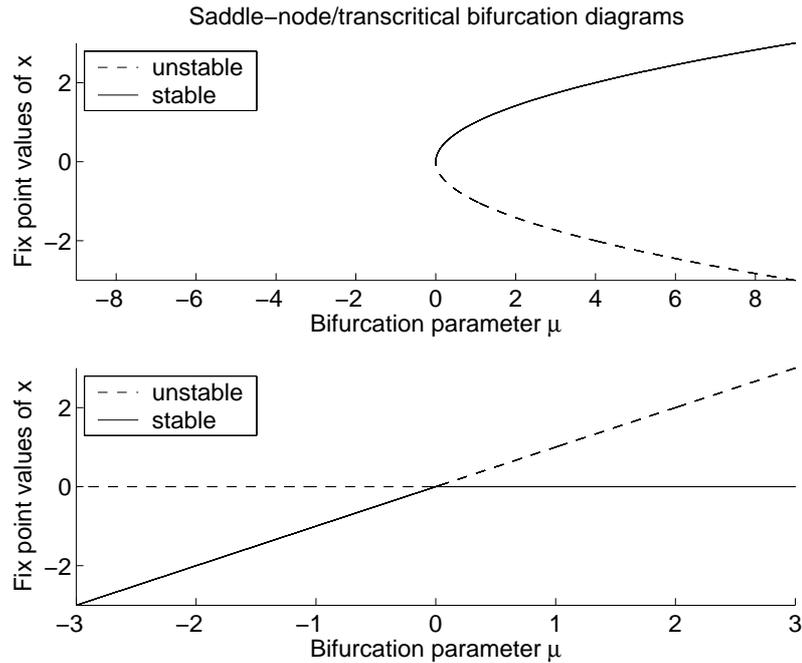


Figure 2.1 Bifurcation diagram over the saddle-node (top), and the transcritical (bottom) bifurcation. What is plotted is the value of the fix point for each value of μ . For the saddle-node bifurcation, a stable and an unstable bifurcation co-exists for positive values of μ . At $\mu = 0$, i.e. at the bifurcation point, the two fix points collide, turn non-hyperbolic, and for negative values of μ there is no fix point at all. For the transcritical bifurcation there are always two fix points, one at $x = 0$ and one at $x = \mu$. At the bifurcation $\mu = 0$ they exchange stability.

The pitchfork bifurcation

The pitchfork bifurcation has received its name from the form of the bifurcation diagram. The bifurcation diagram is plotted in Figure 2.2, and the normal form equation is given in (2.12).

$$\dot{x} = \mu x \pm x^3 \quad (2.12)$$

Just as for the transcritical bifurcation there is a fix point at $x = 0$ that changes stability at the bifurcation point $(x, \mu) = (0, 0)$. What happens in this case, however, depends on whether the bifurcation is supercritical or subcritical, and this depends on whether the sign before the third order term is negative or positive, respectively. In the supercritical case the stable fix point turns unstable just as two stable fix point are created, and this case is therefore just as harmless as the

transcritical bifurcation. The only thing that might be a bit uncertain is whether the system will go to the upper or lower fix point, and if only one of these are desired. But if the bifurcation is detected, and the system can be perturbed, this is usually not a problem. The subcritical case on the other hand is as dramatic as the saddle-node bifurcation. In this case the stable fix point is closed in, and finally annihilated by the two surrounding unstable fix points. After the bifurcation there is no stable fix point at all, and just as after the saddle-node bifurcation, the system will have to go elsewhere.

Remarks

After having introduced these different bifurcations I would like to conclude with some personal remarks. If one studies systems with bifurcations and considers only the linearizations, all these bifurcations will look the same: at the bifurcation a single pole passes the imaginary axis. This is something generally avoided in control situations, and control systems are designed so that this should never happen. This short introduction to the various scenarios that can happen, all caused by a single pole passing the imaginary axis, should make clear that this is not always necessary. When it is necessary depends on the higher order terms, and without a bifurcation analysis one can not determine the effect of these terms. If the bifurcation should turn out to be one of the more dramatic ones, like the saddle-node bifurcation, the fears were motivated, but if it should turn out to be a e.g. transcritical bifurcation, nothing serious will happen to the system and it should be quite safe to operate at, and beyond, this point. In fact, one of the main conclusions of this thesis is that it might be quite *beneficial* to operate close to a bifurcation (see Chapters 3 to 6 and 9). One example of a successful parameter estimation for a large-scale model in systems biology, was also done for yeast cells operating close to a bifurcation (see Hynne et al. (2001) or Figure 2.3). These cells were operating close to a Hopf bifurcation.

2.3.2 The Hopf bifurcation

A general Hopf bifurcation

The Hopf bifurcation was first considered by Henri Poincaré (see e.g. Poincaré (1952)) in the very beginning of the 20th century, further investigated by Andronov a little bit later into the 20th century, and its properties finally exhausted by Hopf (see e.g. Wiggins (1990) for historical review). The bifurcation is therefore sometimes, and more correctly, referred to as the Poincaré-Andronov-Hopf bifurcation. It is a two-dimensional bifurcation, which means that it can only exist in systems with two or more state variables. Just as for the one-dimensional bifurcations all systems with a Hopf bifurcation are (locally) topologically equivalent to the corresponding normal form, and each system can be mapped back from the normal form to the original space (see Chapter 3 for an analysis of how to do this, and Chapter 5 and 9 for two real problem applications utilizing this transformation).

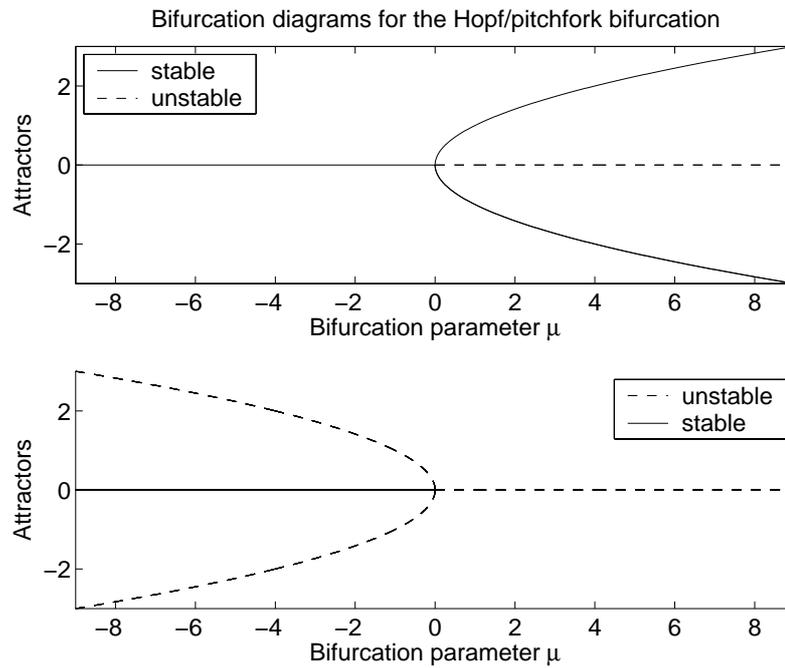


Figure 2.2 Bifurcation diagram for the two Hopf or the two pitchfork bifurcations. For the pitchfork bifurcation the interpretation of the lying parable is as two fix points that at the bifurcation meets with the fix point in the middle, and for the Hopf bifurcation the lying parable is the amplitude of the oscillations. For the Hopf bifurcation it is therefore clear that at the bifurcation the amplitude of the oscillations have grown infinitely small and the bifurcation is merely the transition from fix oscillations to stationary steady state. Notice the extreme difference between the supercritical case (upper plot) and the subcritical case (lower plot). In the supercritical case the stable fix point is replaced by another stable attractor, and from a distance nothing has happened. In the subcritical case, however, the stable fix-point has been eaten by the surrounding unstable attractor, and after the bifurcation there is no stable attractor in this area anymore, and the system will have to go elsewhere.

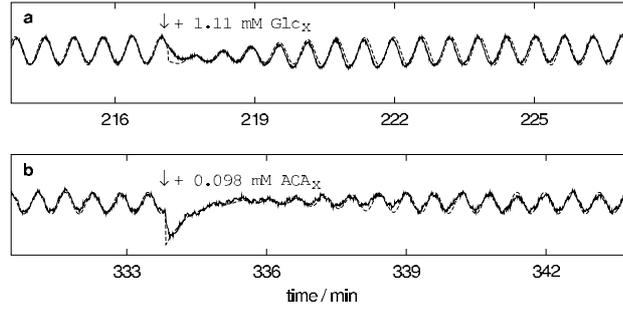


Figure 2.3 Plot showing simulated vs. experimental data from a quenching experiment, i.e. from a system that has been perturbed when operating close to a Hopf bifurcation. For a quenching experiment to substantially change the oscillations, as is done in both these time series, both the timing and the amplitude of the perturbation needs to have very precise values. Considering that this data is taken from a large-scale bio-chemical model, and during the verification step (see Section 2.2) one must conclude that the agreement is quite good. This result is one of the previous indications that it might be beneficial to operate around a Hopf bifurcation. Picture included from with permission from the authors.

The normal form for the Hopf bifurcation is

$$\dot{z} = (i\omega_0 + \mu\sigma_1)z + g_3z|z|^2 \quad (2.13)$$

where $z, g_3, \sigma_1 \in \mathbb{C}$ and $\omega_0, \mu \in \mathbb{R}$. Always in a Hopf bifurcation the eigenvalues of the Jacobian, on the center manifold, passes the imaginary axis at the bifurcation and this means that a limit cycle grows out of a fix point. The bifurcation occurs at $\mu = 0$, and the frequency of the limit cycle at the bifurcation is ω_0 . When regarding stability, however, the sign of $\text{Re}(g_3)$ determines which of the two completely different variations of the bifurcation will occur. For $\text{Re}(g_3) < 0$ we have a supercritical Hopf bifurcation and for $\text{Re}(g_3) > 0$ a subcritical.

A supercritical Hopf bifurcation

The Hopf bifurcation in equation (2.13) is supercritical if $\text{Re}(g_3) < 0$, and then the qualitative nature of the bifurcation diagram is as shown in the top plot of Figure 2.2. The bifurcation occurs at $(z, \mu) = (0, 0)$ and for μ values below this point there is only a stable fix-point. At the bifurcation point the fix point becomes

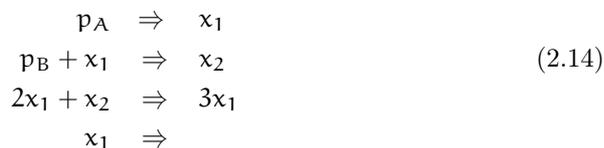
non-hyperbolic, and for positive values the stable fix point has been replaced by a stable limit cycle and the fix-point inside the limit cycle is an unstable spiral (fix point with complex eigenvalues). As μ is increased the limit cycle grows. In the beginning it grows proportional to the square root of the distance to the bifurcation point (here $\mu = 0$). In a time plot this corresponds to an increase in amplitude of the oscillations. This is a local bifurcation, and only affects the local dynamics. The invariant manifold is still attracting, and it is not until the system has moved a little bit away from the bifurcation that the oscillations become visible. This distance can be very small, but for all practical, and even non-analytical, cases it is never infinitely small. This bifurcation is therefore a relatively harmless bifurcation, even though the poles has passed the imaginary axes. It is also occurring in many thermochemical systems, and some examples of this was given already in Chapter 1.

A subcritical Hopf bifurcation

Even though a Hopf bifurcation is detected by a complex pair of eigenvalues of the Jacobian from the linearization around a stable fix point as a control parameter is varied, it is not necessarily the relatively harmless case described above. Just as for the pitchfork bifurcation there is a more 'critical' variation of the bifurcation, and just as for the pitchfork bifurcation it is called subcritical². The bifurcation diagram is shown in the bottom plot of Figure 2.2, and there one sees that at the bifurcation the two invariant manifolds coincide and after the bifurcation the unstable limit cycle has become the new, and unstable, fix point.

2.4 The Brusselator

The Brusselator is a little example that in the simplest case, which is considered here, consists of only two chemicals, x_1 and x_2 . The reactions that take place are the following



i.e. there is an inflow, or a production of the chemical x_1 with the speed p_A . Then x_1 can react with p_B to form x_2 , with the velocity $p_B x_1$ (assuming ordinary mass action kinetics), x_2 can turn into x_1 , by collision with x_1 , with the velocity $x_1^2 x_2$, and finally there is an outflow, or a consumption of x_1 , with the velocity x_1 .

²To unravel the relationship between the pitchfork and the Hopf bifurcation, write equation (2.13) in polar coordinates and find that the amplitude part of the equation is undergoing a pitchfork bifurcation of type (2.12).

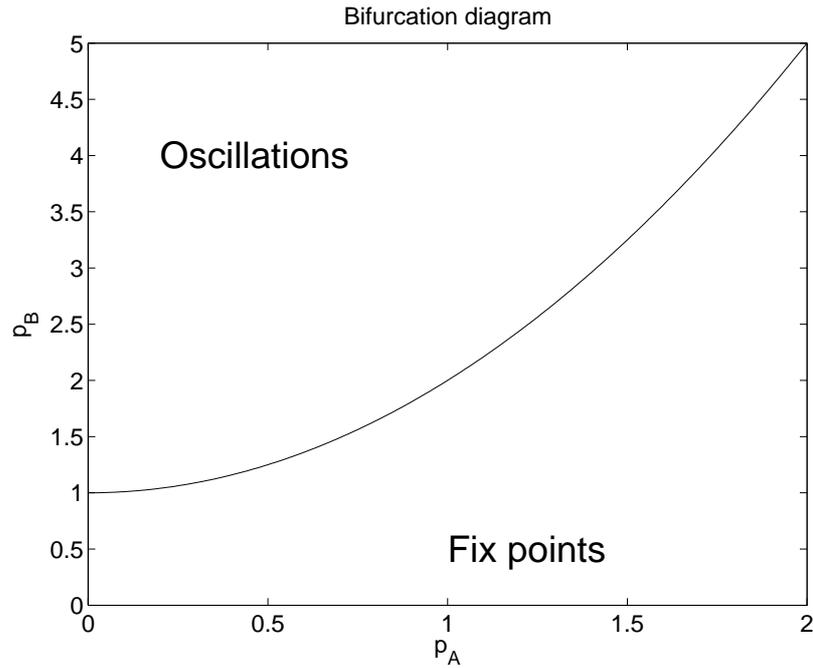


Figure 2.4 A Bifurcation diagram for the Brusselator model. It is known that the bifurcation occurs at $p_B = 1 + p_A^2$ and that above this line the steady state of the system is a limit cycle and below this line it is a fix point.

Writing this as differential equations gives

$$\dot{x}_1 = p_A + x_1^2 x_2 - p_B x_1 - x_1 \quad (2.15a)$$

$$\dot{x}_2 = p_B x_1 - x_1^2 x_2 \quad (2.15b)$$

These two equations correspond to equation (2.4a) and sometimes we will assume that the inflow can be controlled, and sometimes not. That means that p_A and p_B will always belong to the μ vector, but not always to the p_x vector, in (2.4). The dynamics, however, will always be described by equation (2.15). For this system the stationary dynamics has been fully exhausted analytically (see e.g Strogatz (1994)). For each pair (p_A, p_B) there is only one fix point, and it is situated at $(x_1, x_2) = (p_A, p_B/p_A)$. If the fix point is stable it is the only stationary state in the system, and if it unstable it is because it has passed a supercritical Hopf bifurcation. This takes place at the line $p_B = 1 + p_A^2$, and in Figure 2.4 can be seen that above this line the systems stationary state is oscillating, and below it, the stationary state is a fix point.

Model Simplification around a Hopf Bifurcation

This is the first chapter with contributions in it. The first two sections are mainly a review of existing theory, but from Section 3.3.1 and on there are new results. The review in Section 3.1 is mainly taken from the textbook Wiggins (1990), even though it has been rewritten a bit to fit into the rest of the thesis. While Section 3.1 introduces the concepts and basic notations of center manifold and normal form transformations in general, Section 3.2 is taken from the specific article Ipsen, Hynne and Sørensen (1998), and deals with the derivation of the hands-on problem of finding the transformation coefficients. The results starts with an extension of the theory in Ipsen et al. (1998) in Section 3.3 and then, in Section 3.4, the theory is implemented and tested on the Brusselator. The simplifications turns out to work fine and an empirical estimate of how far away one can go from the bifurcation point is the content of the final results section, Section 3.5. Beside the stand alone results of this chapter, it also lays a theoretical basis for some of the coming results. The results in the second half of Chapter 5 depends heavily on these simplifications, and in Chapter 9 the theory is applied to a real, spatio-temporal, nonlinear thermochemical model.

3.1 Center manifold and normal form theory

This chapter is only concerned with equation (2.4a), and with constant \mathbf{u} , i.e. with constant μ . The system of study can hence be formulated

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mu) = \mathbf{J} \cdot \mathbf{x} + \mathbf{f}^{\text{nonl}}(\mathbf{x}, \mu) \quad (3.1)$$

where $\mathbf{x} \in \mathbb{R}^n$ and $\boldsymbol{\mu} \in \mathbb{R}^s$. Here we will also assume that equation (3.1) has a local bifurcation at $(\mathbf{x}, \boldsymbol{\mu}) = (0, 0)$. If this has been accomplished by a translation of the origin in the $\mathbb{R}^{n \times s}$ space, the original state-parameter vector will be denoted $(\mathbf{x}_{\text{phys}}, \boldsymbol{\mu}_{\text{phys}})$. In almost all of the derivation, however, $\boldsymbol{\mu} = 0$ will be the only case considered, but perturbations in $\boldsymbol{\mu}$ space, called the unfolding directions, are important to understand, and the way to extend the theory is sketched, and the corresponding results are stated, in the end of Section 3.2.1.

The center manifold is an invariant manifold

Let $\mathbf{x}(t, \mathbf{x}(0) = \mathbf{x}_0)$ mean the trajectory determined by the starting value \mathbf{x}_0 and the dynamic equations (3.1). Then the definition of an invariant manifold is as follows.

Definition 3.1. *Let $S \subset \mathbb{R}^n$ be a set, then S is said to be an invariant manifold under the vector field (3.1), if for any $\mathbf{x}(0) \in S$ we have $\mathbf{x}(t, \mathbf{x}(0) = \mathbf{x}_0) \in S$ for all $t > 0$.*

In ordinary terms this means that a system starting on an invariant manifold will, as time evolves, always remain on the manifold. Around a non-hyperbolic¹ fix point one can introduce the center and the stable spaces as follows.

Definition 3.2. *For the dynamical system (3.1) with a non-hyperbolic fix point at origin the following spaces can be defined*

$$\mathcal{E}^c = \text{Span}_{\mathbb{R}}(\mathbf{u}_i)_{i=1}^r \quad (3.2)$$

$$\mathcal{E}^s = \text{Span}_{\mathbb{R}}(\mathbf{v}_i)_{i=1}^{n-r} \quad (3.3)$$

where $\{\mathbf{u}_i\}_{i=1}^r$ and $\{\mathbf{v}_i\}_{i=1}^{n-r}$ are the (right) eigenvectors, of the Jacobian J in (3.1), corresponding to the eigenvalues with zero and negative real part, respectively. \mathcal{E}^c and \mathcal{E}^s are called the center and the stable spaces, respectively. Finally $\text{Span}_{\mathbb{R}}$ is only allowing linear combinations that has complex conjugated coefficients in front of complex conjugated eigenvectors.

Remark: Since the eigenvalues are defined as the solution to a real polynomial equation, they will either be real or complex conjugated. The same will be true for the eigenvectors. By the above definition of $\text{Span}_{\mathbb{R}}$, this means that the spaces \mathcal{E}^c and \mathcal{E}^s will always be real. One could therefore define the spaces \mathcal{E}^c and \mathcal{E}^s as spanned by real coefficients to real eigenvectors (Wiggins, 1990). Then the real eigenvectors would be the real and imaginary part of the complex conjugated \mathbf{u}_i 's and \mathbf{v}_i 's. The reason this form is chosen is that we want to be consistent with the derivation taken from Ipsen et al. (1998).

¹A non-hyperbolic fix point is a fix point with a corresponding Jacobian having eigenvalues with zero real part. Hence, the stability of non-hyperbolic fix points has to be determined by higher order analysis. All local bifurcation points, i.e. all bifurcations studied in this thesis, are also non-hyperbolic fix points.

Denote the vectors in the stable and center space with x_s and z , respectively. Let the elements in z be denoted z_i and the elements in x_s be denoted $x_{s,i}$. Then

$$z = \sum_{i=1}^r z_i u_i \quad (3.4)$$

$$x_s = \sum_{i=1}^{n-r} x_{s,i} v_i \quad (3.5)$$

where u_i and v_i are defined in Definition 3.2. Remember the complex conjugated condition on the coefficient, given in Definition 3.2, making both x_s and z real vectors (but with complex elements).

In the same way as for the stable and unstable spaces we could define an unstable space, but this is non-existing for all local bifurcations brought up in this thesis. Finally, about the eigenvectors, should be said that the corresponding left eigenvectors of u_i and v_i will be denoted u_i^* , and v_i^* , respectively. They have been chosen to be bi-orthogonal, i.e.

$$u_i^* \cdot u_j = \delta_{ij} \quad (3.6)$$

where δ_{ij} is the ordinary Kronecker delta. The subspaces \mathcal{E}^c and \mathcal{E}^s are tangent spaces to corresponding invariant manifolds. An invariant manifold that has \mathcal{E}^c as tangent space is called a center manifold. The definition is as follows (Wiggins, 1990).

Definition 3.3. *The dynamical system (3.1) with a non-hyperbolic fix point at origin, has a center manifold, \mathcal{W}^c , if \mathcal{W}^c is an invariant manifold that can be locally represented as*

$$\mathcal{W}^c = \{x_z \in \mathbb{R}^n | x_z = z + h(z), \forall |z| < \delta\} \quad (3.7)$$

where z is defined in equation (3.4) and $h(z)$ is an $\mathbb{R}^n \rightarrow \mathbb{R}^n$ function.

This means that all points on the center manifold, that are sufficiently close to the bifurcation, can be described by r coordinates instead of n . One important implication of this is that the center manifold will have the same dimension as the center space. Figure 3.1 illustrates the idea for a two-dimensional system with a one-dimensional center manifold. For the local one-dimensional bifurcations described in Section 2.3.1 the center manifold is only one-dimensional and for the Hopf bifurcation described in Section 2.3.2 the center manifold is two-dimensional. Now the interesting question is what the dynamics on the center manifold looks like.

Dynamics on the center manifold

To understand the benefits of the center manifold reduction one should look at the dynamics. The properties of the dynamics are described by two important

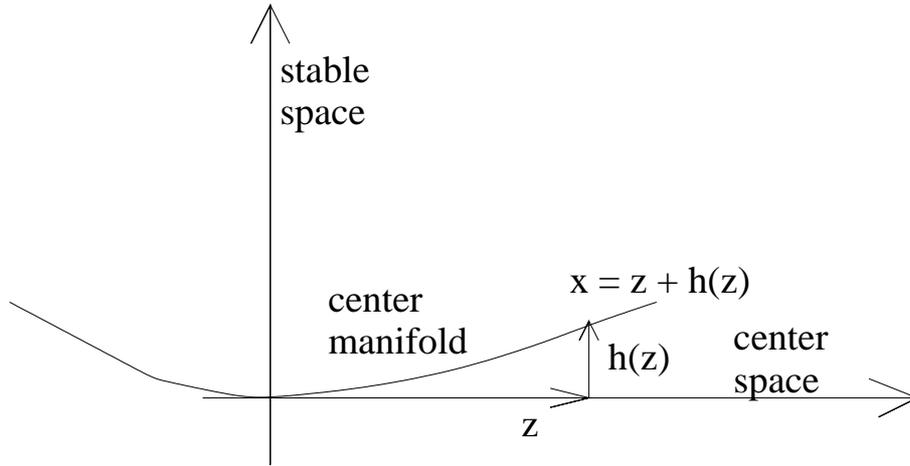


Figure 3.1 The mapping from the center space, \mathcal{E}^c , where $z = \sum_i z_i u_i$ lies to the center manifold \mathcal{W}^c of equation (3.7). As can be seen the center manifold is tangent the center space at the bifurcation, and $h(z)$ is the vector difference between z and x_z .

theorems, and these can be understood once a coordinate transformation from (x_1, \dots, x_n) to $(z_1, \dots, z_r, x_{s,1}, \dots, x_{s,n-r})$ has been done. The original differential equation can in these new coordinates be written

$$\begin{aligned}\dot{z} &= J_z \cdot z + g(z, x_s) \\ \dot{x}_s &= J_x \cdot x_s + g_{x_s}(z, x_s)\end{aligned}\tag{3.8}$$

Theorem 3.1. Existence and dynamics on the manifold

For a system like (3.8) there exists a center manifold. The dynamics of (3.8) on the center manifold, with $|z|$ sufficiently small, can be described by

$$\dot{z} = J_z \cdot z + g(z, h(z))\tag{3.9}$$

Proof. See Wiggins (1990) □

Theorem 3.1 shows the existence of a center manifold and that the dynamics on it is governed by the same functions as before with x_s replaced by $h(z)$. The next theorem deals with the dynamics also when the system starts a little bit away from the manifold.

Theorem 3.2. Dynamics close to the manifold

i) Suppose the fix point $(0,0)$ of (3.9) is stable (asymptotically stable); then the fix point of (3.8) is also stable (asymptotically stable). ii) Suppose the fix point $(0,0)$

of (3.9) is stable. Then if $(z'(t), x_s(t))$ is a solution of (3.8), with $(z'(0), x_s(0))$ sufficiently small, there is a solution $z(t)$ of (3.9) such that as $t \rightarrow \infty$

$$z'(t) = z(t) + O(e^{-\gamma t}) \quad (3.10)$$

$$x_s(t) = h(z(t)) + O(e^{-\gamma t}) \quad (3.11)$$

where γ is a positive constant.

Proof. See Wiggins (1990). □

Theorems 3.1 and 3.2 are important since they say that the dynamics of the original space, around a bifurcation point, can be described by the dynamics on the center manifold alone. The function h takes you back from the center coordinates z_i to the full physical space x . If you start on the center manifold, once you have obtained the functions g and h , you can know *exactly* the dynamics in the full space by simulating the reduced r -dimensional space only, and if you start off, but sufficiently close to the manifold, the r -dimensional simulation gives all information apart from an exponential decay, i.e. the exponential decay to the center manifold will go in parallel with the dynamics on the center manifold.

Normal form transformations

By making transformations of the kind

$$z' = z + \psi_2(z) + \psi_3(z) + \dots \quad (3.12)$$

where ψ_i is an i :th order term in z , you can change the dynamic equation of the center manifold dynamics from (3.9) to

$$\dot{z}' = Az' + g'(z') \quad (3.13)$$

and by choosing the transformation cleverly you can reduce terms from a Taylor expansion of g' . There are, however, some terms that cannot be reduced and these are called resonant terms. When a form has been chosen that contains only resonant terms, i.e. when all terms that can be eliminated by choosing other variables have been eliminated, the dynamics is said to be described by its normal form. When the dimension of the state vector, z , is as low as possible, this form is the simplest form that can describe this dynamics. It was this form that was used to describe the local one- and two-dimensional bifurcations in Section 2.3.

Now all the important components and concepts have been explained, and they sum up to the following.

- Sufficiently close to a bifurcation, everything, except for an ordinary exponential decay, regarding the dynamics is governed by the dynamics on the center manifold (Theorems 3.1 and 3.2).
- The center manifold can be described by the center space coordinates, i.e. those pointing in directions where the corresponding eigenvalues have a zero real part. The relationship is $x = z + h(z)$, where h is a nonlinear function.

- By choosing a clever transformation h in (3.7), the dynamic equations on the center manifold can contain only the resonant terms, and that means that all systems undergoing the same bifurcation can be reduced to the same normal form.

A systematic way to find this h function for all local bifurcations is found in the next section. The transformation table for the Hopf bifurcation, allowing the actual transformation for an arbitrary system to be calculated easily up to third order, is also included.

3.2 A combined transformation

This section presents the results of the paper Ipsen et al. (1998), where a systematic approach of how to make a model reduction (simplification) is given. This reduction includes both the center manifold (reducing the dimension of the dynamics) and the normal form reduction (simplifying the expression by reducing all non-resonant terms), for an arbitrary bifurcation, in the same transformation. Furthermore, the process only involves the solving of linear equations and is therefore easy to implement and to use in practice. In the previous section was seen that in order to complete a transformation from equation (3.1), with a bifurcation point in the origin, to the center manifold dynamics (3.9), and to be able to utilize it (i.e. to be able to transform back again), two things are sought

- function h , in equation (3.7), transforming from the point on the center space to points in the physical space, and
- function g , of equation(3.9) governing the dynamics in the reduced space, \mathcal{E}^c . The function h should be chosen in such a way that only resonant terms in g remain. Therefore a *resonance condition* should also be put up.

Section 3.2.1 gives the relations for a general local bifurcation. The results to both the questions listed above are given in Section 3.2.1 with and without unfolding directions. In Section 3.2.2 these general relationships are applied to the Hopf bifurcation to generate a transformation table that, in most practical cases, will be the only thing one needs to understand. The transformation table says how to use the original differential equation to get the h , g functions, and therefore also, implicitly, what the resonance condition is.

3.2.1 For a general bifurcation

The article Ipsen et al. (1998) starts by deriving the relationships without the parameters, and then just puts up the relationships with the possibility of variations in μ included. The same approach will be done here.

Without parameters

Since the basic theory and notations are already presented above, the first new thing in the article is to expand f^{nonl} , g and h in a Taylor expansion according to

$$f^{\text{nonl}}(z + h(z)) = \sum_{\mathbf{p}} f_{\mathbf{p}} z^{\mathbf{p}}, \quad h(z) = \sum_{\mathbf{p}} h_{\mathbf{p}} z^{\mathbf{p}} \quad g(z, (h(z))) = \sum_{\mathbf{p}} g_{\mathbf{p}} z^{\mathbf{p}} \quad (3.14)$$

Here \mathbf{p} is an index set defined by

$$\mathbf{p} = (p_1, \dots, p_r), \quad z^{\mathbf{p}} = \prod_{i=1}^r z_i^{p_i} \quad (3.15)$$

Here the functions expanded are all strictly nonlinear (no linear terms included), and hence the sums in (3.14) all start at $|\mathbf{p}| = 2$, where $|\mathbf{p}| = \sum p_i$. The interpretation of the sums should be clear after the following example.

Example 3.1

Suppose that $r = 2$, which is the case for e.g. the Hopf bifurcation.

Then $\mathbf{p} \in \{(0, 2), (1, 1), (2, 0)\}$, $z = (z_1, z_2)$ and the expansions in equation (3.14) become:

$$\begin{aligned} f(z + h(z)) &= f_{(2,0)} z_1^2 + f_{(1,1)} z_1 z_2 + f_{(0,2)} z_2^2 + \dots & (3.16) \\ g(z, h(z)) &= g_{(2,0)} z_1^2 + g_{(1,1)} z_1 z_2 + g_{(0,2)} z_2^2 + \dots \\ h(z) &= h_{(2,0)} z_1^2 + h_{(1,1)} z_1 z_2 + h_{(0,2)} z_2^2 + \dots \end{aligned}$$

Exchanging g with its Taylor expansion and multiplying from the left with u_i^* , in equation (3.9) gives

$$u_i^* \cdot \dot{z} = u_i^* \cdot (J \cdot z + \sum_{\mathbf{p}} g_{\mathbf{p}} z^{\mathbf{p}}) \quad (3.17)$$

\Leftrightarrow

$$\dot{z}_i = \lambda_i z_i + \sum_{\mathbf{p}} (u_i^* \cdot g_{\mathbf{p}}) z^{\mathbf{p}} \quad (3.18)$$

Scalar equations for elements of z , like (3.18), are called *amplitude equations*, and (3.18) is hence the i :th amplitude equation. For the Hopf bifurcation the two amplitude equations will be complex conjugations of each other and hence they will both look like (2.13).

Now let us go on with the actual determinations of g and h . A partial result is shown in the following lemma.

Lemma 3.3. *The following relations between the sums describing f , g and h hold*

$$\sum_{\mathbf{p}} g_{\mathbf{p}} z^{\mathbf{p}} = (J - (\mathbf{p} \cdot \lambda)I) \cdot \sum_{\mathbf{p}} h_{\mathbf{p}} z^{\mathbf{p}} + \sum_{\mathbf{p}} \phi_{\mathbf{p}} z^{\mathbf{p}} \quad (3.19)$$

where

$$\sum_{\mathbf{p}} \phi_{\mathbf{p}} z^{\mathbf{p}} = \sum_{\mathbf{p}} f_{\mathbf{p}} z^{\mathbf{p}} - \sum_{\mathbf{p}} \sum_{j=1}^r h_{\mathbf{p}} z^{\mathbf{p}} \frac{p_j}{z_j} u_j^* \sum_{\mathbf{p}'} g_{\mathbf{p}'} z^{\mathbf{p}'} \quad (3.20)$$

Proof. The proof mainly contains a lot of algebraic manipulations. They are as follows. Inserting the relation $x = z + h(z)$ into the original equation (3.1) (without parameters for now) gives

$$\left(I + \frac{\partial h}{\partial z}\right) \cdot \dot{z} = J \cdot z + J \cdot h(z) + f^{\text{nonl}}(z + h(z)) \quad (3.21)$$

Remembering that $\dot{z} = J \cdot z + g(z, h(z))$ (from (3.9)), and the expansions in (3.14), the left hand side of (3.21) becomes

$$\begin{aligned} (3.21).\text{LHS} &= \left(I + \frac{\partial}{\partial z} \left(\sum_{\mathbf{p}} h_{\mathbf{p}} z^{\mathbf{p}}\right)\right) (J \cdot z + \sum_{\mathbf{p}} g_{\mathbf{p}} z^{\mathbf{p}}) = \\ &= \left(I + \sum_{\mathbf{p}} \sum_{j=1}^r h_{\mathbf{p}} z^{\mathbf{p}} \frac{p_j}{z_j} u_j^*\right) \left(\sum_{k=1}^r z_k \lambda_k u_k + \sum_{\mathbf{p}} g_{\mathbf{p}} z^{\mathbf{p}}\right) = \end{aligned} \quad (3.22)$$

$$\begin{aligned} &= \sum_{k=1}^r z_k \lambda_k u_k + \sum_{\mathbf{p}} g_{\mathbf{p}} z^{\mathbf{p}} + \sum_{\mathbf{p}} h_{\mathbf{p}} z^{\mathbf{p}} \cdot \sum_{i=1}^r p_i \lambda_i + \\ &\quad \sum_{\mathbf{p}} \sum_{j=1}^r h_{\mathbf{p}} z^{\mathbf{p}} \frac{p_j}{z_j} u_j^* \cdot \sum_{\mathbf{p}'} g_{\mathbf{p}'} z^{\mathbf{p}'} \end{aligned} \quad (3.23)$$

By putting in the expansions (3.14), the right hand side of (3.21) becomes

$$(3.21).\text{RHS} = \sum_{k=1}^r z_k \lambda_k u_k + J \cdot \sum_{\mathbf{p}} h_{\mathbf{p}} z^{\mathbf{p}} + \sum_{\mathbf{p}} f_{\mathbf{p}} z^{\mathbf{p}} \quad (3.24)$$

Putting the right and left hand sides equal and identifying the introduced entity $\phi_{\mathbf{p}}$ according to (3.20), the given relationship (3.19) is obtained. \square

Corollary 3.4. *The following relationships between the expansion coefficients of equation (3.14) holds*

$$g_{\mathbf{p}} = (J - (\mathbf{p} \cdot \lambda) \mathbf{I}) \cdot h_{\mathbf{p}} + \phi_{\mathbf{p}} \quad (3.25)$$

Proof. Identifying all coefficients of $z^{\mathbf{p}}$ in (3.19) gives the formula. \square

Now this relationship holds sufficient information to find the resonance condition as well as the desired recurrence relations for finding $h_{\mathbf{p}}$, $g_{\mathbf{p}}$. Once having introduced the auxiliary projection and orthogonal complement projection functions R and Q , respectively, according to

$$R \cdot x = \sum_i (u_i^* \cdot x) u_i \quad (3.26)$$

$$Q = I - R \cdot x = I - \sum_i (u_i^* \cdot x) u_i \quad (3.27)$$

By setting $\mathbf{g}_{\mathbf{p}} = 0$ in Corollary 3.4 we can draw the following conclusions.

i) The resonance condition for the normal form and center manifold condition combined, without including the parameters is

$$\mathbf{p} \cdot \lambda = \sum_{j=1}^r p_j \lambda_j = \lambda_i \quad (3.28)$$

i.e. if there is a $\lambda_i \in \lambda = (\lambda_1, \dots, \lambda_r)$ so that the relation (3.28) is fulfilled, for this \mathbf{p} and \mathbf{u}_i direction, $\mathbf{u}_i^* \mathbf{g}_{\mathbf{p}}$ can not be eliminated.

ii) When the resonance condition is not fulfilled choosing the elements of $\mathbf{h}_{\mathbf{p}}$ according to

$$\mathbf{u}_i^* \cdot \mathbf{h}_{\mathbf{p}} = \frac{\mathbf{u}_i^* \cdot \phi_{\mathbf{p}}}{\sum_{j=1}^r p_j \lambda_j - \lambda_i} \quad (3.29)$$

or for the whole vector by solving

$$(\mathbf{J} - (\mathbf{p} \cdot \lambda) \mathbf{I}) \cdot \mathbf{h}_{\mathbf{p}} = -\phi_{\mathbf{p}} \quad (3.30)$$

will cause the corresponding $\mathbf{g}_{\mathbf{p}}$ to be zero.

iii) When the resonance condition is fulfilled, for this \mathbf{p} , $\mathbf{g}_{\mathbf{p}}$ can not be eliminated by any choice of $\mathbf{h}_{\mathbf{p}}$, and $(\mathbf{J} - (\mathbf{p} \cdot \lambda) \mathbf{I})$ is not invertible. For these cases the following solution is suggested in Ipsen et al. (1998)

$$(\mathbf{J} - (\mathbf{p} \cdot \lambda) \mathbf{I}) \cdot \mathbf{h}_{\mathbf{p}} = -\mathbf{Q} \phi_{\mathbf{p}} \quad (3.31)$$

$$\mathbf{R} \cdot \mathbf{h}_{\mathbf{p}} = 0 \quad (3.32)$$

$$\mathbf{g}_{\mathbf{p}} = (\mathbf{J} - (\mathbf{p} \cdot \lambda) \mathbf{I}) \cdot \mathbf{h}_{\mathbf{p}} + \phi_{\mathbf{p}} \quad (3.33)$$

This equation does not always have a solution, so it will in practice be solved by using the pseudo-inverse, obtained by singular value decomposition (SVD).

With parameters

When variations in parameters are also considered, the differential equation is extended to include the parameters as well. The parameters are supposed constant, i.e. $\dot{\boldsymbol{\mu}} = 0$, and therefore these new zero-directions are also a part of the center space. The name will however be kept, since there are some fundamental differences between \mathbf{z} and $\boldsymbol{\mu}$, and sometimes one needs to distinguish between them. The derivations are almost identical for this case and therefore the general results and needed auxiliaries are just stated here, without derivations.

The center manifold is described by a mapping from \mathbf{z} and $\boldsymbol{\mu}$ to \mathbf{x} according to

$$\mathcal{W}^c = \{(\mathbf{x}, \boldsymbol{\mu}) | \mathbf{x} = \mathbf{z} + \mathbf{h}(\mathbf{z}, \boldsymbol{\mu})\} \quad (3.34)$$

and the extension of equation (3.9) is

$$\dot{z} = J \cdot z + g(z, h(z), \mu) \quad (3.35)$$

The Taylor expansions of (3.14) become

$$\begin{aligned} f^{\text{nonl}}(z + h(z, \mu)) &= \sum_{\mathbf{p}\mathbf{q}} h_{\mathbf{p}\mathbf{q}} z^{\mathbf{p}} \mu^{\mathbf{q}}, & h(z, \mu) &= \sum_{\mathbf{p}\mathbf{q}} h_{\mathbf{p}\mathbf{q}} z^{\mathbf{p}} \mu^{\mathbf{q}}, \\ g(z, h(z), \mu) &= \sum_{\mathbf{p}\mathbf{q}} g_{\mathbf{p}\mathbf{q}} z^{\mathbf{p}} \mu^{\mathbf{q}}, \end{aligned} \quad (3.36)$$

where \mathbf{q} is the analog of \mathbf{p} , but for the parameters, i.e.

$$\begin{aligned} \mathbf{p} &= (p_1, \dots, p_r), & z^{\mathbf{p}} &= \prod_{i=1}^r z_i^{p_i}, \\ \mathbf{q} &= (q_1, \dots, q_s), & \mu^{\mathbf{q}} &= \prod_{k=1}^s \mu_k^{q_k}, \end{aligned} \quad (3.37)$$

The corresponding relation of the one in Corollary 3.4 is

$$\mathbf{g}_{\mathbf{p}\mathbf{q}} = (J - \mathbf{p} \cdot \lambda) \cdot \mathbf{h}_{\mathbf{p}\mathbf{q}} + \Phi_{\mathbf{p}\mathbf{q}} \quad (3.38)$$

where

$$\Phi_{\mathbf{p}\mathbf{q}} = f_{\mathbf{p}\mathbf{q}} + \sum_{\mathbf{p}'\mathbf{q}'} h_{\mathbf{p}\mathbf{q}} \sum_{j=1}^r u_j^* \cdot g_{(\mathbf{p}-\mathbf{p}'+\delta_j)(\mathbf{q}-\mathbf{q}')} \quad (3.39)$$

and the conclusions i) to iii) given without parameters become

i) The resonance condition for the normal form and center manifold condition combined is

$$\mathbf{p} \cdot \lambda = \sum_{j=1}^r p_j \lambda_j = \lambda_i \quad (3.40)$$

i.e. if there is a $\lambda_i \in \lambda$ so that the relation (3.40) is fulfilled, for this \mathbf{p} , $\mathbf{g}_{\mathbf{p}\mathbf{q}}$ can not be eliminated (for any \mathbf{q}).

ii) When the resonance condition not is fulfilled, choosing the elements of $\mathbf{h}_{\mathbf{p}\mathbf{q}}$ according to

$$\mathbf{u}_i^* \cdot \mathbf{h}_{\mathbf{p}\mathbf{q}} = \frac{\mathbf{u}_i^* \cdot \Phi_{\mathbf{p}\mathbf{q}}}{\sum_{j=1}^r p_j \lambda_j - \lambda_i} \quad (3.41)$$

or for the whole vector by solving

$$(J - (\mathbf{p} \cdot \lambda)\mathbf{I}) \cdot \mathbf{h}_{\mathbf{p}\mathbf{q}} = -\Phi_{\mathbf{p}\mathbf{q}} \quad (3.42)$$

will cause the corresponding $\mathbf{g}_{\mathbf{p}\mathbf{q}}$ to be zero.

iii) When the resonance condition is fulfilled, for this \mathbf{p} , $\mathbf{g}_{\mathbf{p}\mathbf{q}}$ can not be eliminated by any choice of $\mathbf{h}_{\mathbf{p}\mathbf{q}}$, and $(J - \mathbf{p} \cdot \lambda\mathbf{I})$ is not invertible. For overcoming this

problem the same approach as for the case without parameters is suggested also here (Ipsen et al., 1998)

$$(J - (\mathbf{p} \cdot \lambda)\mathbf{I}) \cdot \mathbf{h}_{\mathbf{p}\mathbf{q}} = -Q\phi_{\mathbf{p}\mathbf{q}} \quad (3.43)$$

$$\mathbf{R} \cdot \mathbf{h}_{\mathbf{p}\mathbf{q}} = 0 \quad (3.44)$$

$$\mathbf{g}_{\mathbf{p}\mathbf{q}} = (J - \mathbf{p} \cdot \lambda) \cdot \mathbf{h}_{\mathbf{p}\mathbf{q}} + \phi_{\mathbf{p}\mathbf{q}} \quad (3.45)$$

where SVD is still used to solve the over-determined equation. Finally, let us consider the problem of transforming back to the original physical space. As mentioned in the introduction to this chapter (just below (3.1)) we have denoted the physical coordinates $(x_{\text{phys}}, \mu_{\text{phys}})$. Let, in these coordinates the bifurcation point be denoted $(x_{\text{phys},b}, \mu_{\text{phys},b})$. Notice that this is the same bifurcation point that lies in the origin in the coordinates used in (3.1). With these notations the transformation from the center space coefficients z_i to the physical coordinates $(x_{\text{phys}}, \mu_{\text{phys}})$ are as follows

$$x_{\text{phys}} = x_{\text{phys},b} + \sum_{i=1}^r z_i \mathbf{u}_i + \sum_{\mathbf{p}\mathbf{q}} h_{\mathbf{p}\mathbf{q}} z^{\mathbf{p}} (\mu_{\text{phys}} - \mu_{\text{phys},b})^{\mathbf{q}}. \quad (3.46)$$

This is just the modification of (3.34) with $h(z, \mu)$ replaced by the expansion in (3.36). Notice that μ has been replaced by $(\mu_{\text{phys}} - \mu_{\text{phys},b})$ and that this replacement should be done at all places where the parameters appear. Let us now turn to the specific problem of determining this transformation for a Hopf bifurcation.

3.2.2 Derivation for the Hopf bifurcation

In this subsection the derivation and interpretation of Table 3.1, that describes all steps to be taken when finding the amplitude equations for a system close to a Hopf bifurcation, is given. Table 3.1 and its derivation is taken from the article Ipsen et al. (1998), and in this article the derivation is done up to third order expansions in state space, and in parameter space, only linear terms are included, and μ is one-dimensional ($s = 1$). These limitations will later be expanded to fifth order in state space (Table 3.2), and to include more than one parameter.

Including third order terms in x , and only up to first order for those terms including the parameter μ , the nonlinear function f^{nonl} of equation (3.1) has the following Taylor expansion

$$f^{\text{nonl}}(x, \mu) = \mathbf{F}_{\mu} \cdot \mu + \mathbf{F}_{x\mu}(x, \mu) + \frac{1}{2!} \mathbf{F}_{xx}(x, x) + \frac{1}{3!} \mathbf{F}_{xxx}(x, x, x) + \dots \quad (3.47)$$

where the terms of the right-hand side are defined according to

$$\begin{aligned}
\mathbf{F}_{xx}(\mathbf{u}, \mathbf{v}) &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{x_{\text{p h y s}}, b} \mathbf{u}_i \mathbf{v}_j \\
\mathbf{F}_{xxx}(\mathbf{u}, \mathbf{v}, \mathbf{w}) &= \sum_{i,j,k=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} \Big|_{x_{\text{p h y s}}, b} \mathbf{u}_i \mathbf{v}_j \mathbf{w}_k \\
\mathbf{F}_{\mu} \cdot \mu &= \sum_{i=1}^s \frac{\partial f}{\partial \mu_i} \Big|_{x_{\text{p h y s}}, b} \mu_i, \\
\mathbf{F}_{x\mu}(\mathbf{u}, \mu) &= \sum_{i=1}^n \sum_{j=1}^s \frac{\partial^2 f}{\partial x_i \partial \mu_j} \Big|_{x_{\text{p h y s}}, b} \mathbf{u}_i \mu_j
\end{aligned} \tag{3.48}$$

where n is the dimension of x , and μ is one-dimensional which means that $s = 1$. The vectors $x_{\text{p h y s}}$ and $\mu_{\text{p h y s}}$ are the same as those used in (3.46).

Let us now go into the derivations specific for the Hopf bifurcation. The Hopf bifurcation has a two-dimensional center space (without counting the parameters), and the corresponding eigenvalues are two purely imaginary, complex conjugated, eigenvalues,

$$\lambda = (\lambda_1, \lambda_2) = (i\omega_0, -i\omega_0) \tag{3.49}$$

where $\omega_0 \in \mathbb{R}$. The corresponding eigenvectors are

$$\mathbf{u} = (\mathbf{u}_1, \mathbf{u}_2) = (\mathbf{u}_1, \bar{\mathbf{u}}_1) = (\mathbf{u}, \bar{\mathbf{u}}) \tag{3.50}$$

where the last equality simply means that we can drop the index. The \mathbf{p} vector is two-dimensional and the μ vector is one-dimensional

$$\mathbf{p} = (p_1, p_2) \quad \mu = (\mu) \tag{3.51}$$

The resonance condition (3.40) says that only those \mathbf{p} that satisfy any of the following two relations will appear with a $\mathbf{g}_{\mathbf{p}\mathbf{q}}$ contribution in the amplitude equations. With (3.49), (3.40) becomes

$$\begin{aligned}
p_1 \lambda_1 + p_2 \lambda_2 &= (p_1 - p_2) \lambda_1 = \lambda_1 \\
p_1 \lambda_1 + p_2 \lambda_2 &= (p_2 - p_1) \lambda_2 = \lambda_2
\end{aligned} \tag{3.52}$$

where the first of the equations in (3.52) is fulfilled only if

$$p_1 = p_2 + 1 \tag{3.53}$$

and the second only if

$$p_2 = p_1 + 1 \tag{3.54}$$

Since the two options correspond to the two eigenvalues, and therefore to the two orthogonal directions, only one of them will appear in each amplitude equation.

Here we arbitrarily pick the first of these equations, and know that the other amplitude equation is simply the complex conjugate of the first. Applying condition (3.53) to the amplitude equation (3.18) with $i = 1$ gives

$$\dot{z}_1 = \lambda_1 z_1 + \sum_{\mathbf{p}\mathbf{q}} g_{\mathbf{p}\mathbf{q}} z^{\mathbf{p}} \mu^{\mathbf{q}} = \quad (3.55)$$

$$= (i\omega_0 + g_{101})z_1 \mu + g_{210}z_1^2 z_2 + \dots \quad (3.56)$$

where higher order terms than this are dropped in Ipsen et al. (1998). Since $z_1 = \bar{z}_2$, the index will be dropped according to

$$(z_1, z_2) = (z, \bar{z}) \quad (3.57)$$

Notice that z will now both denote the scalar, and complex amplitude, and the real and n -dimensional vector. For this reason we will in the subsequent chapters keep the index for the scalar and let z always mean the n -dimensional vector. We can further simplify the notation by letting $g_{\mathbf{p}\mathbf{0}}$ be denoted $g_{|\mathbf{p}|}$ and g_{101} be denoted σ_1 . With these notational changes (3.55) becomes

$$\dot{z} = (i\omega_0 + \sigma_1 \mu)z + g_3 |z|^2 z \quad (3.58)$$

Once the amplitude equation has been established, what remains is the determination of the equation (3.42) to be solved for the non-resonant terms and (3.43) together with (3.45) to be solved for the resonant terms. Then, once these equations have been solved, for the specific system in question, the transformation is complete. The equations to be solved up til third order in $\mathbf{p}\mathbf{q} = \mathbf{p}\mathbf{0}$ expansions, and to exponent one in those terms containing the parameter, has been given in Table 3.1. This table is the main theorem of this section, and it is reached by the following two lemmas.

Lemma 3.5. *The coefficients of f^{nonl} in the expansion (3.36) are*

$$f_{200} = \frac{1}{2} \mathbf{F}_{xx}(\mathbf{u}, \mathbf{u}) \quad (3.59)$$

$$f_{110} = \mathbf{F}_{xx}(\mathbf{u}, \bar{\mathbf{u}}) \quad (3.60)$$

$$f_{300} = \mathbf{F}_{xx}(\mathbf{u}, h_{200}) + \frac{1}{6} \mathbf{F}_{xxx}(\mathbf{u}, \mathbf{u}, \mathbf{u}) \quad (3.61)$$

$$f_{300} = \mathbf{F}_{xx}(\mathbf{u}, h_{110}) + \mathbf{F}_{xx}(\bar{\mathbf{u}}, h_{200}) + \frac{1}{2} \mathbf{F}_{xxx}(\mathbf{u}, \mathbf{u}, \bar{\mathbf{u}}) \quad (3.62)$$

$$f_{001} = \mathbf{F}_{\mu} \quad (3.63)$$

$$f_{101} = \mathbf{F}_{x\mu} \cdot \mathbf{u} + \mathbf{F}_{xx}(\mathbf{u}, h_{001}) \quad (3.64)$$

The proof is too long to be included but instead an example of how the first relation is found is included. Doing the same for all the other terms would constitute the proof.

Example 3.2 Finding f_{200}

Finding f_{200} is the same as replacing \mathbf{x} by

$$\begin{aligned} \mathbf{x} &= \sum_{i=1}^r z_i \mathbf{u}_i + \sum_{\mathbf{p}\mathbf{q}} h_{\mathbf{p}\mathbf{q}} z^{\mathbf{p}} \mu^{\mathbf{q}} = \\ &= \mathbf{u}z + \overline{\mathbf{u}}\bar{z} + h_{200}z^2 + h_{110}|z|^2 + h_{020}\bar{z}^2 + h_{001}\mu + h_{101}z\mu + h_{011}\bar{z}\mu \end{aligned} \quad (3.65)$$

in equation (3.47), simplifying and identifying the coefficient in front of the z^2 term. For higher order terms this is most easily done by a software like Mathematica but for this simple example it can be illustrative to see the contributions from the individual terms in (3.47).

- $\mathbf{F}_{\mu} \cdot \mu$ and $\mathbf{F}_{\mathbf{x}\mu}(\mathbf{x}, \mu)$ cannot have any contributions since they will only have terms, once expanded, including a parameter μ . We are looking for z^2 terms.
- $\frac{1}{2}\mathbf{F}_{\mathbf{x}\mathbf{x}}(\mathbf{x}, \mathbf{x})$ is more promising. We find

$$\begin{aligned} \mathbf{F}_{\mathbf{x}\mathbf{x}}(\mathbf{u}z + \mathcal{O}(\bar{z}, \mu, z^2), \mathbf{u}z + \mathcal{O}(\bar{z}, \mu, z^3)) &= \\ \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} u_i z_i u_j z_j + \mathcal{O}(\bar{z}, \mu, z^3) &= \frac{1}{2}\mathbf{F}_{\mathbf{x}\mathbf{x}}(\mathbf{u}, \mathbf{u})z^2 + \mathcal{O}(\bar{z}, \mu, z^3) \end{aligned} \quad (3.66)$$

where $\mathcal{O}(\bar{z}, \mu, z^3)$ means all terms with at least one of the given arguments to the given order.

- $\frac{1}{2}\mathbf{F}_{\mathbf{x}\mathbf{x}\mathbf{x}}(\mathbf{x}, \mathbf{x}, \mathbf{x}) = \mathcal{O}(\bar{z}, \mu, z^3)$ and hence no terms are possible.

Hence $f_{200} = \frac{1}{2}\mathbf{F}_{\mathbf{x}\mathbf{x}}(\mathbf{u}, \mathbf{u})$ and this is what we wanted to show with our little example.

Lemma 3.6. $\phi_{\mathbf{p}\mathbf{q}} = f_{\mathbf{p}\mathbf{q}}$ for all terms included up til the orders included in this section

Proof. The second sum does not contribute for any order since either the $h_{\mathbf{p}'\mathbf{q}'}$ term will be zero or the $g_{(\mathbf{p}-\mathbf{p}'+\delta_j)(\mathbf{q}-\mathbf{q}'')}$ will be zero (due to the nonlinearity of h or g or to that very few terms of g are indeed nonzero). In general, however this does not hold, and should higher order terms have been included some $\phi_{\mathbf{p}\mathbf{q}}$ might have been different from $f_{\mathbf{p}\mathbf{q}}$ (see e.g. Table 3.2). \square

Theorem 3.7. Table 3.1 contains all necessary equations for transforming an arbitrary system with a Hopf bifurcation to and from its amplitude equation. However, only terms up to third order in $\mathbf{p}\mathbf{q} = \mathbf{p}\mathbf{0}$ terms and only up to single exponents in unfolding terms are included.

Proof. All terms follow directly from equations (3.42) for the non-resonant terms and (3.43) together with (3.45) for the resonant terms. \square

Transformation table for Hopf bifurcation	
A	$x = uz + \bar{u}\bar{z} + h_{200}z^2 + h_{110} z ^2 + h_{020}\bar{z}^2 + h_{210} z ^2z + h_{120} z ^2\bar{z} + h_{300}z^3 + h_{030}\bar{z}^3 + h_{001}\mu + (h_{101}z + h_{011}\bar{z})\mu$
B	$z^2 : (J - \gamma_2 I) \cdot h_{200} = -\frac{1}{2}F_{xx}(u, u)$
	$h_{020} = \bar{h}_{200}$
	$ z ^2 : (J - \gamma_0 I) \cdot h_{110} = -F_{xx}(u, \bar{u})$
	$z^3 : (J - \gamma_3 I) \cdot h_{300} = -F_{xx}(u, h_{200}) - \frac{1}{6}F_{xxx}(u, u, u)$
	$h_{030} = \bar{h}_{300}$
	$z z ^2 : (J - \gamma_1 I) \cdot h_{210} = -Q \cdot (F_{xx}(u, h_{110}) + F_{xx}(\bar{u}, h_{200}) + \frac{1}{2}F_{xxx}(u, u, \bar{u}))$
	$u^* \cdot h_{210} = 0$
	$h_{120} = \bar{h}_{210}$
	$\mu : (J - \gamma_0 I) \cdot h_{001} = -F_\mu$
	$\mu z : (J - \gamma_1 I) \cdot h_{101} = -Q \cdot (F_{x\mu}(u) + F_{xx}(u, h_{001}))$
	$u^* \cdot h_{101} = 0$
	$h_{011} = \bar{h}_{101}$
C	$g_3 = u^* \cdot F_{xx}(u, h_{110}) + u^* \cdot F_{xx}(\bar{u}, h_{200}) + \frac{1}{2}u^* \cdot F_{xxx}(u, u, \bar{u})$
	$\sigma_1 = u^* \cdot F_{x\mu}(u) + u^* \cdot F_{xx}(u, h_{001})$
D	$\dot{z} = (i\omega_0 + \sigma_1 \mu)z + g_3 z z ^2$
	$\gamma_k = k * i\omega_0$
	$Q \cdot x = x - (u^* \cdot x)u$

Table 3.1 In this table, all the information that is needed to make a transformation between an arbitrary system with a Hopf bifurcation, and the normal form, is collected. A is the way (3.46) looks for the Hopf bifurcation. Notice that z here is the scalar, complex, coefficient z_1 , with the index dropped according to (3.57). B contains the linear equations needed to calculate the h s in the transformation in A. C describes how to calculate the resonant coefficients, needed to form the amplitude equation. The amplitude equation is given under D. $F_{xx}(u, v)$ and all similar terms are defined in equations (3.47) and (3.48).

3.3 Extensions of existing theory

The derivation of Table 3.1 in the previous section, did only include terms up to third order in $\mathbf{p}\mathbf{q} = \mathbf{p}\mathbf{0}$ terms and only μz and μ terms in the unfolding directions, where μ was assumed to be one-dimensional. In this section we will extend Table 3.1 in two ways. First we will include up to 5th order in the $\mathbf{p}\mathbf{0}$ terms and second we will consider varying more than one parameter.

3.3.1 Extension to fifth order

The resonance condition derived for the $z_1(=z)$ equation $p_1 = p_2 + 1$ is valid up to an arbitrary high order and the inclusion of the extra term in the amplitude equation (as an extension of (3.58)) is simply

$$\dot{z} = (i\omega_0 + \sigma_1\mu)z + g_3|z|^2z + g_5|z|^4z + \dots \quad (3.67)$$

where $g_5 = g_{320}$. There is hence only one new resonant term g_5 , and an expansion only up to 4'th order would have caused no new terms to appear.

Table 3.2 (included last in the chapter) contains all necessary equations for calculations of the $h_{\mathbf{p}\mathbf{0}}$ and the $g_{\mathbf{p}\mathbf{0}}$ up to $|\mathbf{p}| \leq 5$. The way it is derived is along the exact same lines as those outlined in Section 3.2.2. First the extended expansion of x (compare equation (3.66))

$$\begin{aligned} x = & uz + \bar{u}\bar{z} + h_{200}z^2 + h_{110}|z|^2 + h_{020}\bar{z}^2 + h_{210}|z|^2z + h_{120}|z|^2\bar{z} + \\ & + h_{300}z^3 + h_{030}\bar{z}^3 + h_{400}z^4 + h_{040}\bar{z}^4 + h_{220}|z|^4 + h_{130}|z|^2\bar{z}^2 + \\ & + h_{310}|z|^2z^2 + h_{500}z^5 + h_{050}\bar{z}^5 + h_{410}z^3|z|^2 + h_{140}\bar{z}^3|z|^2 + h_{230}|z|^4\bar{z} + \\ & + h_{320}|z|^4z + h_{001}\mu + (h_{101}z + h_{011}\bar{z})\mu \end{aligned}$$

is included in the extended expansion of $f^{\text{nonl}}(x, \mu)$ (compare equation (3.47))

$$\begin{aligned} f^{\text{nonl}}(x, \mu) = & \mathbf{F}_\mu \cdot \mu + \mathbf{F}_{x\mu}(x, \mu) + \frac{1}{2!}\mathbf{F}_{xx}(x, x) + \frac{1}{3!}\mathbf{F}_{xxx}(x, x, x) + \\ & + \frac{1}{4!}\mathbf{F}_{xxxx}(x, x, x, x) + \frac{1}{5!}\mathbf{F}_{xxxxx}(x, x, x, x, x) + \dots \quad (3.68) \end{aligned}$$

Then this expression is simplified, and the coefficients $f_{\mathbf{p}\mathbf{q}}$ are identified (as in Lemma 3.5). For this extended expansion $\phi_{\mathbf{p}\mathbf{q}} = f_{\mathbf{p}\mathbf{q}}$ does unfortunately not hold and the full expression (3.39) has to be used. Finally the equations (3.42) are formed for the non-resonant terms and the equations (3.43) and (3.45) are formed for the resonant terms. All terms also included in Table 3.1 are identical to those in Table 2. The new terms are however substantially more complex in their appearance. The difference in outcome between 3'rd, and 5'th order for the example of the Brusselator is included in the last section of this chapter.

3.3.2 Extensions to more than one parameter

When one extra parameter is included, we have $s = 2$, $\mathbf{q} = (q_1, q_2)$ and $\mu = (\mu_1, \mu_2)$, in the general derivations of Section 3.2. For the Hopf bifurcation this means that each term $f_{\mathbf{p}\mathbf{q}}$, $g_{\mathbf{p}\mathbf{q}}$ and $h_{\mathbf{p}\mathbf{q}}$ will have four indices, and e.g. f_{2011} is the coefficient of the $z_1^2\mu_1\mu_2$ term in the expansion of $f(z + h(z))$. The derivation of the transformation table, with $s = 2$, is almost identical to that in 3.2.2. Here are mentioned only the differences.

The resonance condition is concerned only with the \mathbf{p} vector and is hence identical compared to before. The amplitude equation for the first z -element $z_1(=z)$

is hence (compare (3.58))

$$\begin{aligned}\dot{z} &= i\omega_0 z + g_{1010}\mu_1 z + g_{1001}\mu_2 z + g_{2100}|z|^2 z = \\ &= i\omega_0 z + \sigma_{10}\mu_1 z + \sigma_{01}\mu_2 z + g_3 z |z|^2\end{aligned}\quad (3.69)$$

The second and third order terms of $\mathbf{h}_{\mathbf{p}\mathbf{q}}$ are the same as in the Table 3.1 (except for the extra index zero). The unfolding terms are, however, different and the new equations are

$$\mathbf{J} \cdot \mathbf{h}_{0010} = -\mathbf{F}_{\mu_1} \quad (3.70)$$

$$\mathbf{J} \cdot \mathbf{h}_{0001} = -\mathbf{F}_{\mu_2} \quad (3.71)$$

$$(\mathbf{J} - \gamma_1 \mathbf{I}) \cdot \mathbf{h}_{1010} = -\mathbf{Q}(\mathbf{F}_{x\mu_1} \cdot \mathbf{u} + \mathbf{F}_{xx}(\mathbf{u}, \mathbf{h}_{0010})) \quad (3.72)$$

$$\mathbf{u}^* \cdot \mathbf{h}_{1010} = 0 \quad (3.73)$$

$$\mathbf{h}_{0110} = \bar{\mathbf{h}}_{1010} \quad (3.74)$$

$$(\mathbf{J} - \gamma_1 \mathbf{I}) \cdot \mathbf{h}_{1001} = -\mathbf{Q}(\mathbf{F}_{x\mu_2} \cdot \mathbf{u} + \mathbf{F}_{xx}(\mathbf{u}, \mathbf{h}_{0001})) \quad (3.75)$$

$$\mathbf{u}^* \cdot \mathbf{h}_{1001} = 0 \quad (3.76)$$

$$\mathbf{h}_{0101} = \bar{\mathbf{h}}_{1001} \quad (3.77)$$

Finally the transformation back to the physical space is the following

$$\begin{aligned}\mathbf{x} &= \mathbf{u}z + \bar{\mathbf{u}}\bar{z} + \mathbf{h}_{2000}z^2 + \mathbf{h}_{1100}|z|^2 + \mathbf{h}_{0200}\bar{z}^2 + \mathbf{h}_{0010}\mu_1 + \\ &+ \mathbf{h}_{0001}\mu_2 + \mathbf{h}_{1010}z\mu_1 + \mathbf{h}_{1001}z\mu_2 + \mathbf{h}_{0110}\bar{z}\mu_1 + \mathbf{h}_{0101}\bar{z}\mu_2\end{aligned}\quad (3.78)$$

3.4 Transformation of the Brusselator

In this section the developed theory will be applied to the simple example called the Brusselator introduced in Chapter 2. The differential equations are

$$\dot{x}_1 = p_A + x_1^2 x_2 - p_B x_1 - x_1 \quad (3.79)$$

$$\dot{x}_2 = p_B x_1 - x_1^2 x_2 \quad (3.80)$$

The condition for a Bifurcation to occur is $p_B = 1 + p_A^2$, and the fix point for a given (p_A, p_B) values lies at $(p_A, p_B/p_A)$. The bifurcation point chosen to do the expansion around in this example is $(x_1, x_2, p_A, p_B) = (1, 2, 1, 2)$. The Brusselator is already two-dimensional, and hence there is no room for contracting directions. The center space, the center manifold and the physical space will thus cover the same points, although not described in the same way. However, since the two transformations, the center manifold reduction and the normal form simplification, are done in one connected step, it is not possible (or at least not beneficial) to utilize this.

The transformation is done up to third and fifth order, and in the next section the difference between the original system, and the transformation back from the

two amplitude equations are compared.

Example 3.3 Using Table 3.1

At the chosen point the Jacobian is

$$J(1, 2, 1, 2) = \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} \quad (3.81)$$

which has as its two eigenvalues $\lambda = \pm i$. The right and left eigenvectors of the eigenvalue $+i$ are

$$\mathbf{u} = (-1 - i, 2)^T \quad (3.82)$$

$$\mathbf{u}^* = \frac{1}{4}(2i, 1 + i) \quad (3.83)$$

In Table 3.1 can be seen that the second order term equations require the terms $\phi_{200} = \frac{1}{2}\mathbf{F}_{xx}(\mathbf{u}, \mathbf{u})$ and $\phi_{110} = \mathbf{F}_{xx}(\mathbf{u}, \bar{\mathbf{u}})$, they become

$$\begin{aligned} \frac{1}{2}\mathbf{F}_{xx}(\mathbf{u}, \mathbf{u}) &= \frac{1}{2} \sum_{i,j=1}^2 \frac{\partial^2 f}{\partial x_i \partial x_j} \Big|_{(1,2,1,2)} \mathbf{u}_i \mathbf{u}_j = \\ &= \frac{1}{2} \begin{pmatrix} 2 \cdot 2 \\ -2 \cdot 2 \end{pmatrix} (-1 - i)(-1 - i) + \frac{1}{2} \begin{pmatrix} 2 \cdot 1 \\ -2 \cdot 1 \end{pmatrix} (-1 - i)2 + \\ &\quad + \frac{1}{2} \begin{pmatrix} 2 \cdot 1 \\ -2 \cdot 1 \end{pmatrix} (-1 - i)2 + \frac{1}{2} \begin{pmatrix} 0 \\ 0 \end{pmatrix} 2 \cdot 2 = (-4, 4)^T \end{aligned} \quad (3.84)$$

$$\mathbf{F}_{xx}(\mathbf{u}, \bar{\mathbf{u}}) = (0, 0)^T \quad (3.85)$$

With these terms the top three linear equations in Table 3.1 for \mathbf{h}_{200} , \mathbf{h}_{020} and \mathbf{h}_{110} become

$$\begin{aligned} \left(\begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} - 2i \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \cdot \begin{pmatrix} \mathbf{h}_{200,1} \\ \mathbf{h}_{200,2} \end{pmatrix} &= - \begin{pmatrix} -4 \\ 4 \end{pmatrix} \\ \begin{pmatrix} \mathbf{h}_{020,1} \\ \mathbf{h}_{020,2} \end{pmatrix} &= \begin{pmatrix} \bar{\mathbf{h}}_{200,1} \\ \bar{\mathbf{h}}_{200,2} \end{pmatrix} \end{aligned} \quad (3.86)$$

$$\begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} \cdot \begin{pmatrix} \mathbf{h}_{110,1} \\ \mathbf{h}_{110,2} \end{pmatrix} = - \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3.87)$$

and once solved they give

$$\mathbf{h}_{200} = \bar{\mathbf{h}}_{020} = \frac{1}{3}(8i, -4 - 8i)^T \quad (3.88)$$

$$\mathbf{h}_{110} = (0, 0)^T \quad (3.89)$$

The third order terms (starting with $z|z|^2$; in Table 3.1) require in the same way

$$\phi_{300} = \mathbf{F}_{xx}(\mathbf{u}, \mathbf{h}_{200}) \frac{1}{6} \mathbf{F}_{xxx}(\mathbf{u}, \mathbf{u}, \mathbf{u}) = (8 + 12i, -8 - 12i)^T$$

$$Q\phi_{210} = Q \cdot (\mathbf{F}_{xx}(\mathbf{u}, \mathbf{h}_{110}) + \mathbf{F}_{xx}(\bar{\mathbf{u}}, \mathbf{h}_{200})) + \frac{1}{2} \mathbf{F}_{xxx}(\mathbf{u}, \mathbf{u}, \bar{\mathbf{u}}) = \frac{1}{3}(8 + 10i, 2 - 18i)^T$$

to put up the $z|z|^2$ equations in Table 3.1. The solutions to these equations are

$$\begin{aligned} h_{300} = \bar{h}_{030} &= \frac{1}{2}(9 - 6i, -7 + 9i)^T \\ h_{210} = \bar{h}_{120} &= \frac{1}{3}(5 - 4i, -10 - i)^T \end{aligned}$$

The unfolding terms (μ and μz equations in Table 3.1) require

$$\phi_{001} = F_\mu = (-1, 1)^T \quad (3.90)$$

$$Q\phi_{101} = Q \cdot (F_{x\mu}(u) + F_{xx}(u, h_{001})) = -\frac{1}{2}(1 + i, -2i)^T \quad (3.91)$$

These unfolding equations, once put up, gives

$$h_{101} = \bar{h}_{011} = \frac{1}{4}(-1 + i, 2)^T \quad (3.92)$$

$$h_{001} = (0, 1)^T \quad (3.93)$$

Finally the two resonant terms of g become

$$g_3 = u^* \cdot \phi_{210} = (-3, -\frac{i}{3}) \quad (3.94)$$

$$\sigma_1 = u^* \cdot \phi_{101} = \frac{1}{2} \quad (3.95)$$

With these terms calculated the transformation under A in Table 3.1, combined with the transformation back to the physical space in equation (3.46), maps the normal form coordinates $(z, \mu) = (0.1 + 0i, 0)$ on the physical coordinates $(x_1, x_2) = (0.812, 2.36033)$.

Let us now see what happens if we use Table 3.2 instead?

Example 3.4 Using Table 3.2

There is very little qualitative differences between how Table 3.2 is used and how Table 3.1 was used in the previous example. The same point $(1, 2, 1, 2)$ is used and hence the terms, and equations, already included in Table 3.1 will be identical. The new terms are

$$h_{400} = \bar{h}_{400} = \left(-\frac{1456}{135} - \frac{16i}{3}, -\frac{1636}{135} - \frac{356i}{135}\right)^T \quad (3.96)$$

$$h_{310} = \bar{h}_{130} = \left(\frac{5}{3} - \frac{4i}{3}, -3 - \frac{i}{3}\right)^T \quad (3.97)$$

$$h_{220} = \left(\frac{16}{9}, -\frac{232}{9}\right)^T \quad (3.98)$$

$$h_{500} = \bar{h}_{050} = \left(\frac{925}{648} - \frac{18205i}{648}, -\frac{761}{108} - \frac{4505i}{162}\right)^T \quad (3.99)$$

$$h_{410} = \bar{h}_{140} = \left(-\frac{2011}{360} - \frac{6989i}{360}, -\frac{1507}{180} - \frac{328i}{15}\right)^T \quad (3.100)$$

$$h_{320} = \bar{h}_{230} = \left(-\frac{190}{27} - \frac{871i}{54}, -\frac{139}{6} - \frac{491i}{54}\right)^T. \quad (3.101)$$

and as seen the term h_{220} is real just as it should.

Now the question that remains to be answered in this chapter is how well the simplified dynamics mimics the real behaviour. A related question is how big the error is from not including the fifth order term when only using Table 3.1. These two questions are answered in the next section.

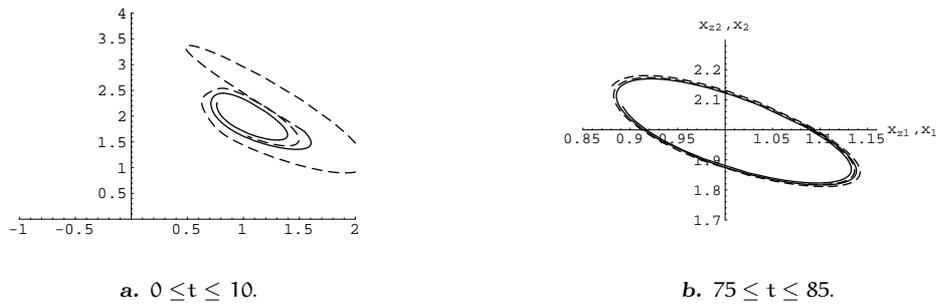


Figure 3.2 Phase plots of the Brusselator model. The dashed line belongs to the transformed system, with up to third order terms included. The solid line corresponds to the directly solved system. The two simulations started at the same initial state, corresponding to $z(0) = 1$. As the figures shows, the transformed system behaves differently at the beginning compared to the directly simulated system. However, after a while the two systems converge to each other. The operating point is the same as the one in Section 3.4, and the settings of the integrator can be found in Appendix A.

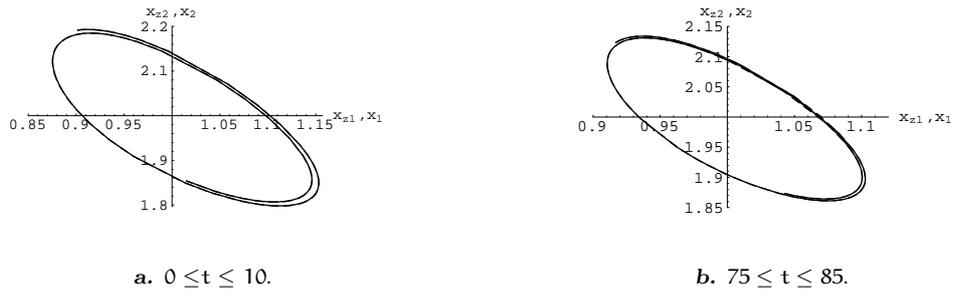


Figure 3.3 Phase-plots of the Brusselator model. The dashed line belongs to the transformed system, with terms up to third order included, and the solid line to the directly solved equations. The two simulations started at the same initial state, corresponding to $z(0) = 0.1$. The phase diagram differs very little from each other, and this is because the initial condition is closer to the fix-point (cf Figure 3.2). The operating point is the same as the one in Section 3.4, and the settings of the integrator can be found in Appendix A.

3.5 Error analysis

When doing a model simplification it is sought to estimate the error that is the price of the simplification. These contributions come from two sources. The first source of errors is the approximation of not including the stable directions, i.e. the projection on to the center manifold. This error is ideally non-existing if the system starts on the center manifold, but since the center manifold is only approximated there will always be a contribution from this truncation effect. The second source of error is the normal form transformation. This is ideally an isomorphic mapping between two descriptions of the dynamics on the center manifold but since also this transformation is truncated, there will be an error contribution also from this part. As with all Taylor expansions these truncations have errors of at least the first non-included order in the corresponding expansion. In practice these two contributions are not separable and it is also difficult to analyse when a particular order comes through. Therefore two pragmatic approaches are used.

Error in time series

The first method is based on the Root Mean Square (RMS) error which is defined as

$$\text{RMS} = \sqrt{\frac{\sum_{i=0}^N (x(t_i) - x_z(t_i))^2}{N}} \quad (3.102)$$

where x_z is the one obtained through the transformation back from the center manifold (see equation (3.7)), and x is obtained through direct simulation of (2.1). Notice that for the Brusselator the bifurcation point does not lie in origin, and it is therefore necessary to do the full transformation (3.46). The RMS error, used like this, is the averaged error between two integrations starting in the same point (which is chosen to lie on the center manifold, $x(t=0) = x_z(z(t=0))$). In (3.102) N is the number of points in the model for which the RMS error is evaluated. RMS error is applied to the transformation of the Brusselator in the previous section and then $N = 100$.

Before doing the analysis of (3.102) it might be instructive to just look at some plots comparing x and x_z for various initial values. In Figure 3.2 the initial value is far away from the bifurcation point, $z(0) = 1$ and the corresponding perturbation in x is almost 0.5 in both x_1 and x_2 . In the beginning x and x_z goes in totally opposite directions. Nevertheless, after some time, the systems have approached the non-hyperbolic fix point and then the errors are hardly visible. If the systems starts closer to the bifurcation point, like in Figure 3.3 (where $z(0) = 0.1$), the difference between x and x_z is not visible to the eye. Generally can be said that for perturbations less than ten per cent ($|x| \leq 0.2$), the difference have not been visible. Figure 3.4 shows simulations where this perturbation has been done in state space, together with a perturbation of 0.2 in the bifurcation parameter p_B . As can be seen, even this large perturbation in parameter space affects surprisingly little, considering that only first order terms are included in the unfolding directions.

The feeling that perturbations less than 10 per cent gives no visible outcome is further studied by the usage of the RMS measure, defined above. RMS plots are made at 441 different initial values of $z(t)$ for the Brusselator, and the axes named Im and Re shows the imaginary and real part of the initial value of $z(t)$. The initial values are taken with a step size of 0.1 and then the RMS values are calculated for the specific initial value and finally a mesh is made in Matlab which describes how the error varies for different initial values of $z(t)$. Such a plot is shown in Figure 3.5. There can be seen that there is no direction where the mismatch is significantly higher than in a randomly chosen direction. Therefore it seems that the observations done in the simulations above were the typical situation one would encounter.

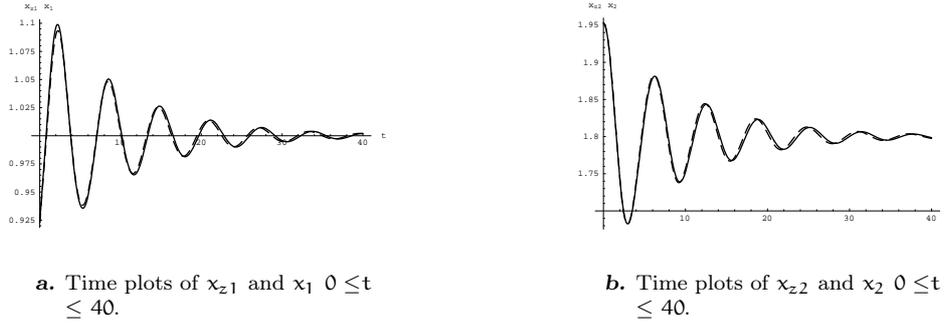


Figure 3.4 Time plots of the two dimensional Brusselator when the bifurcation parameter is changed to $p_B = 1.8$, p_A is still equal to 1. The initial value is $z(0) = 0.1$. The plots shows that the transformation still approximates the Brusselator, even if the value of the bifurcation parameter is changed. The values of the parameters and the settings of the integrator can be found in Appendix A.

Error in derivatives

Another way of estimating the error obtained from the simplification is by comparing the error in \dot{x} directly. Figure 3.6 is a plot showing this error and it has been obtained as follows. For each z value tested, (3.7) gives a symbolic expression of $x_z(z)$. At this point the difference

$$\frac{d}{dt}(x - x_z) = \dot{x} - \frac{\partial x_z}{\partial z} \cdot \dot{z} = \dot{x} - \dot{x}_z \cdot (J\dot{z} + g(z, h(z))) \quad (3.103)$$

is calculated. The first equality follows from the chain rule, and the second from equation (3.9). Here the right hand side was easily calculated by the symbolic differentiation functions in Mathematica. The error for each z_1 obtained in (3.103) is then plotted against the corresponding z value in the same way as for the RMS error.

Differences between 3'rd and 5'th order expansions

The question what the difference in error between including and not including the terms up to fifth order is, was also raised. The difference between third and fifth order is largest at the beginning of the time series. The fifth order follow the transient better than the third order. This was shown by calculating the RMS error for 100 points in the time interval $0 \leq t \leq 1$. In this time interval the RMS error for the fifth order transformation is $RMS_5 = 0.003353$ and for the third order it is $RMS_3 = 0.126121$. In the time interval $9 \leq t \leq 10$ $RMS_5 = 0.007313$ and

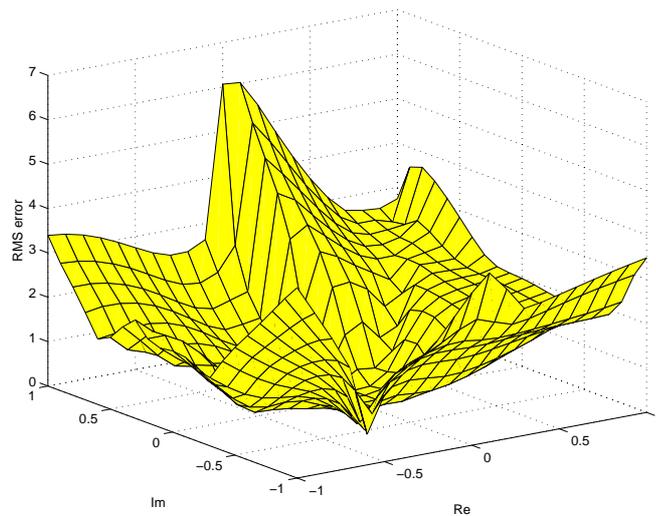


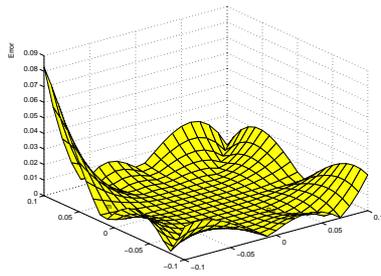
Figure 3.5 *RMS error plot of the Brusselator at different initial values of $z(t)$. As can be seen there is no direction in which the agreement is significantly worse than any other. This is not a proof, but it indicates that one could for this system and operating point, simulate the transformed system instead of the original for perturbation sizes less than the 10% observed in the time-series analysis.*

$\text{RMS}_3 = 0.058932$. Time plots for the Brusselator at different initial values of $z(t)$ are shown in Figure 3.7.

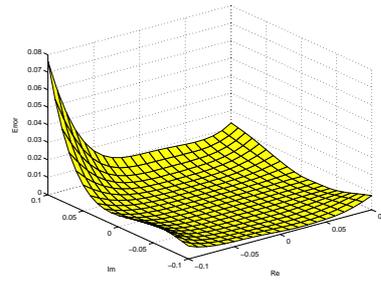
3.6 Conclusions

The conclusions from this chapter are:

- That the center manifold theory and normal form theory combined as described in Section 3.2 can indeed be used to do a local transformation of a dynamical system that is valid in the neighborhood of a Hopf bifurcation with perturbations in both parameter and state space. The simplified equation will always be two-dimensional and, even more, will always have the same form (equation (3.58)).
- That tables describing how to do this transformation for fifth order terms in z and for an arbitrary number of parameter directions have been generated



a. $|\dot{\mathbf{x}} - \dot{\mathbf{x}}_z|$ values of third order.

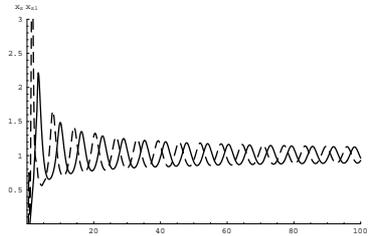


b. $|\dot{\mathbf{x}} - \dot{\mathbf{x}}_z|$ values of fifth order.

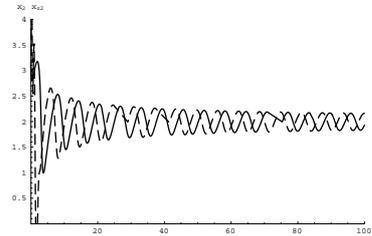
Figure 3.6 *The difference in derivatives of the two dimensional Brusselator for different \mathbf{z} values. The difference is smaller in the fifth order, and this mean that the derivatives of the fifth order transformation will follow the numerically solved system better than the third order transformation and will therefore be more accurate.*

(see Table 3.2). Previously only up to third order terms and for a single parameter has been published.

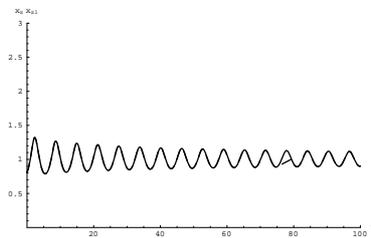
- That with the aid of Mathematica these transformations can easily be done for a simple system like the Brusselator.
- That for the Brusselator the error is not symmetric towards perturbations in different directions, but that the difference between a simulation generated by the simplified system and the original system was not visible to the eye for perturbations of less than 10%. That means that the transformation is valid, not only in a mathematical sense, but for situations that might be used in applications.
- That for larger perturbations than 10% the error eventually increases in its visibility, but that this can be postponed by adding the fifth order terms in the expansion (see Figure 3.6).



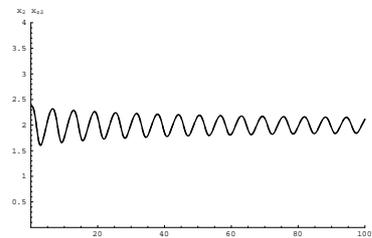
a. Time plots for $x_{z1}(t)$ and $x_1(t)$ at the bifurcation point where $z(0)=0.5$.



b. Time plots for $x_{z2}(t)$ and $x_2(t)$ at the bifurcation point where $z(0)=0.5$.



c. Time plots for $x_{z1}(t)$ and $x_1(t)$ at the bifurcation point where $z(0)=0.1$.



d. Time plots for $x_{z2}(t)$ and $x_2(t)$ at the bifurcation point where $z(0)=0.1$.

Figure 3.7 Time plots of the Brusselator of the fifth order. The dashed line belongs to the transformed system and the solid line to the numerically solved. The figures shows The Brusselator solved at two different initial values of $z(t)$, and the one closer to the fix point will of course give the best result. The values of the parameters and the settings of the integrator can be found in Appendix A.

Transformation table for Hopf bifurcation	
$x = uz + \bar{u}\bar{z} + h_{200}z^2 + h_{110} z ^2 + h_{020}\bar{z}^2 + h_{210} z ^2z + h_{120} z ^2\bar{z} + h_{300}z^3 + h_{030}\bar{z}^3 + h_{400}z^4 + h_{040}\bar{z}^4 + h_{220} z ^4 + h_{130} z ^2\bar{z}^2 + h_{310} z ^2z^2 + h_{500}z^5 + h_{050} z ^5 + h_{410}z^3 z ^2 + h_{140}\bar{z}^3 z ^2 + h_{230} z ^4\bar{z} + h_{320} z ^4z + h_{001}\mu + (h_{101}z + h_{011}\bar{z})\mu$	
z^2	$(J - \gamma_2 I) \cdot h_{200} = -\frac{1}{2}F_{xx}(u, u)$ $h_{020} = \bar{h}_{200}$
$ z ^2$	$(J - \gamma_0 I) \cdot h_{110} = -F_{xx}(u, \bar{u})$
z^3	$(J - \gamma_3 I) \cdot h_{300} = -F_{xx}(u, h_{200}) - \frac{1}{6}F_{xxx}(u, u, u)$ $h_{030} = \bar{h}_{300}$
$z z ^2$	$(J - \gamma_1 I) \cdot h_{210} = -Q \cdot (F_{xx}(u, h_{110}) + F_{xx}(\bar{u}, h_{200}) + \frac{1}{2}F_{xxx}(u, u, \bar{u}))$ $u^* \cdot h_{210} = 0$ $h_{120} = \bar{h}_{210}$
z^4	$(J - \gamma_4 I) \cdot h_{400} = -\frac{1}{24}F_{xxxx}(u, u, u, u) - \frac{1}{2}F_{xxx}(u, u, h_{200}) - \frac{1}{2}F_{xx}(h_{200}, h_{200}) - F_{xx}(u, h_{300})$ $h_{040} = \bar{h}_{400}$
$ z ^4$	$(J - \gamma_0 I) \cdot h_{220} = -\frac{1}{2}F_{xx}(h_{110}, h_{110}) - F_{xx}(u, h_{120}) - F_x(h_{020}, h_{200}) - F_{xx}(u, h_{210}) - \frac{1}{2}F_{xxx}(u, u, h_{020}) - F_{xxx}(u, u^*, h_{110}) - \frac{1}{2}F_{xxx}(u^*, u^*, h_{200}) - \frac{1}{4}F_{xxxx}(u, u, u^*, u^*) + 2(\bar{u}^* \cdot \phi_{210}) \cdot h_{020} + (\bar{u}^* \cdot \phi_{210} + u^* \cdot \phi_{210}) \cdot h_{110} + 2(u^* \cdot \phi_{210})h_{200}$
$z^2 z ^2$	$(J - \gamma_2 I) \cdot h_{310} = -F_{xx}(h_{110}, h_{200}) - F_{xx}(u, h_{210}) - F_{xx}(u^*, h_{300}) - \frac{1}{2}F_{xxx}(u, u, h_{110}) - F_{xxx}(u, u^*, h_{200}) - \frac{1}{6}F_{xxxx}(u, u, u, u^*) + (\bar{u}^* \cdot \phi_{210}) \cdot h_{110} + 2(u^* \cdot \phi_{210}) \cdot h_{200}$ $h_{130} = \bar{h}_{310}$
z^5	$(J - \gamma_5 I) \cdot h_{500} = -F_{xx}(h_{200}, h_{300}) - F_{xx}(u, h_{400}) - \frac{1}{2}F_{xxx}(u, h_{200}, h_{200}) - \frac{1}{2}F_{xxx}(u, u, h_{300}) - \frac{1}{2}F_{xxxx}(u, u, u, h_{200}) - \frac{1}{120}F_{xxxxx}(u, u, u, u, u)$ $h_{050} = \bar{h}_{500}$
$z^3 z ^2$	$(J - \gamma_3 I) \cdot h_{410} = -F_{xx}(h_{200}, h_{210}) - F_{xx}(h_{110}, h_{300}) - F_{xx}(u, h_{310}) - F_{xx}(u^*, h_{400}) - F_{xxx}(u^*, h_{110}, h_{200}) - \frac{1}{2}F_{xxx}(u, h_{200}, h_{200}) - \frac{1}{2}F_{xxx}(u, u, h_{210}) - F_{xxx}(u, u^*, h_{300}) - \frac{1}{6}F_{xxxx}(u, u, u, h_{110}) - \frac{1}{2}F_{xxxx}(u, u, u^*, h_{200}) - \frac{5}{120}F_{xxxxx}(u, u, u, u, u^*) + (\bar{u}^* \cdot \phi_{210}) \cdot h_{210} + 3(u^* \cdot \phi_{210}) \cdot h_{300}$ $h_{140} = \bar{h}_{410}$

Table 3.2 Transformation table for the Hopf bifurcation with up to fifth order terms included. Continued in Table 3.3.

Transformation table for Hopf bifurcation	
$z z ^4$	$ \begin{aligned} (J - \gamma_1 I) \cdot h_{320} = & -Q(F_{xx}(h_{120}, h_{200}) + F_{xx}(h_{110}, h_{210}) + F_{xx}(u, h_{220}) \\ & + F_{xx}(h_{020}, h_{300}) + F_{xx}(u^*, h_{310}) + \frac{1}{2}F_{xxx}(u, h_{110}, h_{110}) \\ & + \frac{1}{2}F_{xxx}(u, u, h_{120}) + F_{xxx}(u, h_{020}, h_{200}) \\ & + F_{xxx}(u^*, h_{110}, h_{200}) + F_{xxx}(u, u^*, h_{210}) \\ & + \frac{1}{2}F_{xxx}(u^*, u^*, h_{300}) + \frac{1}{6}F_{xxxx}(u, u, u, h_{020}) \\ & + \frac{1}{2}F_{xxxx}(u, u, u^*, h_{110}) + \frac{1}{2}F_{xxxx}(u, u^*, u^*, h_{200}) \\ & + \frac{1}{12}F_{xxxxx}(u, u, u, u^*, u^*) - 2(\overline{u^*} \cdot \phi_{210})h_{120} \\ & - (\overline{u^*} \cdot \overline{\phi_{210}} + 2u^* \cdot \phi_{210}) \cdot h_{210} \\ & - 3(u^* \overline{\phi_{210}}) \cdot h_{300} \\ h_{230} = & \overline{h_{320}} \\ \mu : (J - \gamma_0 I) \cdot h_{001} = & -F_\mu \\ (J - \gamma_1 I) \cdot h_{101} = & -Q \cdot (F_{x\mu}(u, h_{200}) \cdot u + F_{xx}(u, h_{001})) \\ \mu z : u^* \cdot h_{101} = & 0 \\ h_{011} = & \overline{h_{101}} \end{aligned} $
g_3	$u^* \cdot F_{xx}(u, h_{110}) + u^* \cdot F_{xx}(\overline{u}, h_{200}) + \frac{1}{2}u^* \cdot F_{xxx}(u, u, \overline{u})$
g_5	$ \begin{aligned} & F_{xx}(h_{120}, h_{200}) + F_{xx}(h_{110}, h_{210}) + F_{xx}(u, h_{220}) + F_{xx}(h_{020}, h_{300}) \\ & + F_{xx}(u^*, h_{310}) + \frac{1}{2}F_{xxx}(u, h_{110}, h_{110}) + \frac{1}{2}F_{xxx}(u, u, h_{120}) \\ & + F_{xxx}(u, h_{020}, h_{200}) + F_{xxx}(u^*, h_{110}, h_{200}) + F_{xxx}(u, u^*, h_{210}) \\ & + \frac{1}{2}F_{xxx}(u^*, u^*, h_{300}) + \frac{1}{6}F_{xxxx}(u, u, u, h_{020}) + \frac{1}{2}F_{xxxx}(u, u, u^*, h_{110}) \\ & + \frac{1}{2}F_{xxxx}(u, u^*, u^*, h_{200}) + \frac{1}{12}F_{xxxxx}(u, u, u, u^*, u^*) - 2(\overline{u^*} \cdot \phi_{210})h_{120} \\ & - (\overline{u^*} \cdot \overline{\phi_{210}} + 2u^* \cdot \phi_{210}) \cdot h_{210} - 3(u^* \overline{\phi_{210}})h_{300} \\ \sigma_1 = & u^* \cdot F_{x\mu} \cdot u + u^* \cdot F_{xx}(u, h_{001}) \end{aligned} $
$\dot{z} = (i\omega_0 + \sigma_1 \mu)z + g_3 z z ^2 + g_5 z z ^4$	
	$ \begin{aligned} \gamma_k &= k * i\omega_0 \\ Q \cdot x &= x - (u^* \cdot x)u \end{aligned} $

Table 3.3 Continuation from Table 3.2.

Improved Parameter Estimation

This is the first of three chapters presenting new identification methods for systems of the general structure (2.4), utilizing some kind of bifurcation theory. For the special model structure of (2.4) the parameter vector p that one seeks to estimate naturally divides in three sub-vectors, p_x , x_0 , and p_y (see Section 2.1). In this chapter we will show results for handling the first of these vectors, the p_x vector. In the next chapter we will extend the methods to apply also for the x_0 vector and in Chapter 6 we will show results for handling some problematic p_y vectors. The method in this chapter is applicable if one knows a specific input, u_b , for which the system undergoes a Hopf bifurcation. It will be shown how this knowledge can be transformed into a constrained optimization problem, allowing a reduction in the degrees of freedom describing the p_x vector. Two different methods for doing this are presented. The first method is an elimination method and it shows the benefit of reducing the search space with one degree of freedom most clearly. However, it has the drawback of forcing the user to choose one specific parameter to eliminate. The second method is a reduced gradient method, and it does not have this drawback. It, however, only assures linear convergence.

The chapter starts by rehearsing the basic notations and giving an initial formulation of the problem. The next section reformulates the question so that it can be fit into the existing constrained optimization formalism, and also introduces a way to find a feasible starting point for the estimation. Then the elimination and the projection methods are introduced, and their basic properties are shown. The two methods are tried out on the Brusselator, and their performance are compared to a straightforward approach, like the one described in Section 2.2. The chapter ends with a summary of the conclusions.

4.1 The problem

We will start with a short rehearsal of the key equations presented in Chapter 2, before we turn to the specific problem that this chapter is interested in.

In this chapter we will assume that the time series Z^N is given, and that it has been generated from a system of the general structure (cf. (2.4))

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}) = \mathbf{f}(\mathbf{x}, \mathbf{p}_x, \mathbf{u}) \quad (4.1a)$$

$$\mathbf{y} = \mathbf{y}(\mathbf{x}, \boldsymbol{\mu}, \mathbf{p}_y) \quad (4.1b)$$

$$\mathbf{x}(t_{\text{start}}) = \mathbf{x}_0 \quad (4.1c)$$

The three unknown parameters are collected in a single parameter vector (cf. (2.5))

$$\mathbf{p} = (\mathbf{p}_x, \mathbf{p}_{n_x}) = (\mathbf{p}_x, \mathbf{p}_y, \mathbf{x}_0) \quad (4.2)$$

These parameters are to be estimated from the time series according to the least square selection criteria (cf (2.8))

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} V_N(\mathbf{p}, Z^N) \quad (4.3)$$

where (cf (2.9))

$$V_N(\mathbf{p}, Z^N) = \frac{1}{N} \sum_{t=1}^N (\mathbf{y}(t) - \hat{\mathbf{y}}(t|\mathbf{p}))^2 \quad (4.4)$$

Now, what is added in this chapter is that we know that the system is at a Hopf bifurcation for a specific input \mathbf{u}_b . This value might have been obtained through an analysis prior to (or posterior to) the collection of the time series, but it could also have been detected by an analysis of the collected time series. If the latter is the case, let t_b denote the time at which this detection was done. Then t_b and \mathbf{u}_b are related by the following relation

$$\mathbf{u}_b = \mathbf{u}(t_b) \quad (4.5)$$

We, however, emphasize that it is not necessary to connect \mathbf{u}_b with a specific time. Now the question that this chapter poses is the following:

How can one include the knowledge that \mathbf{u}_b corresponds to a Hopf bifurcation point, in the identification problem (4.3)?

4.2 Starting the estimation

In this section is described how to reformulate the problem just described into an ordinary constrained optimization problem. For reasons that will be clear later, it will be important to start the optimization processes with an initial estimate that is a bifurcation point, and fortunately there are continuation methods that can do this. Continuation methods are a part of the applied bifurcation theory, and one common method is described in Section 4.2.2.

4.2.1 Reformulating the problem

To understand how to formulate the extra condition that the system is at a Hopf bifurcation for $\mathbf{u} = \mathbf{u}_b$, we first need to recall the basic features of a Hopf bifurcation. The two main features defining a Hopf bifurcation that we will later utilize are the following

- A Hopf bifurcation point $(\mathbf{x}_b, \boldsymbol{\mu}_b)$ for a dynamical system (4.1a) is a non-hyperbolic fix point with exactly two purely imaginary critical eigenvalues (zero real part) to the Jacobian $\mathbf{J} = \nabla_{\mathbf{x}} f(\mathbf{x}_b, \boldsymbol{\mu}_b)$.
- As the system leaves the bifurcation point, through variations in one of the parameters $\mu_{b,i}$, the complex conjugated pair of eigenvalues will leave the imaginary axis. This latter condition is called the crossing-criteria. If one of the critical eigenvalues is denoted λ_1 this criteria can be written as (Wiggins, 1990)

$$\frac{d(\operatorname{Re}(\lambda_1(\mathbf{x}_b, \boldsymbol{\mu}_b)))}{d\mu_{b,i}} \neq 0 \quad (4.6)$$

From the first of these two criteria we can construct the following test for a bifurcation point.

Let $\lambda_i(\mathbf{x}, \boldsymbol{\mu})$ denote the i :th eigenvalue to the Jacobian of (4.1a) evaluated at the point $(\mathbf{x}, \boldsymbol{\mu})$. Further let F^{Hopf} be a scalar function defined by the following product

$$F^{\text{Hopf}}(\mathbf{x}, \boldsymbol{\mu}) := \prod_{i < j} (\lambda_i(\mathbf{x}, \boldsymbol{\mu}) + \lambda_j(\mathbf{x}, \boldsymbol{\mu})) \quad (4.7)$$

Then it follows from the first of the two conditions for a Hopf bifurcation, mentioned above, that if $(\mathbf{x}_b, \boldsymbol{\mu}_b)$ is a Hopf bifurcation it must fulfill the following relations

$$f(\mathbf{x}_b, \boldsymbol{\mu}_b) = 0 \quad (4.8)$$

$$F^{\text{Hopf}}(\mathbf{x}_b, \boldsymbol{\mu}_b) = 0 \quad (4.9)$$

The first of these two relations is fulfilled if and only if $(\mathbf{x}_b, \boldsymbol{\mu}_b)$ is a fix-point, and since a Hopf bifurcation point is also a fix-point, then (4.8) is fulfilled for all Hopf bifurcation points. The equation (4.9), on the other hand, is not fulfilled by all fix-points. However, since a Hopf bifurcation has one pair of purely imaginary eigenvalues, one factor in (4.7) will be zero. Therefore also the entire product, i.e. F^{Hopf} will be zero. We thus conclude that both (4.8) and (4.9) will be fulfilled if $(\mathbf{x}_b, \boldsymbol{\mu}_b)$ is a Hopf bifurcation point. Let us now see how we can use this test to reformulate the problem given in Section 4.1 into a constrained optimization problem.

In the problem given in Section 4.1 we knew that the system was operating at a Hopf bifurcation point for a specific \mathbf{u}_b . That means that for the estimated parameter $\hat{\mathbf{p}}_{\mathbf{x}}$ to fulfill the given constraint there must exist a state \mathbf{x}_b such that $(\mathbf{x}_b, \mathbf{p}_{\mathbf{x}}, \mathbf{u}_b)$ fulfills (4.8) and (4.9). Let us re-state this in a more compact way. Let \mathbf{c}^{Hopf} denote the $n+1$ -dimensional vector containing both the condition in (4.8)

and in (4.9). Further let it always be evaluating the \mathbf{u} part of the $\boldsymbol{\mu}$ vector at \mathbf{u}_b , i.e.

$$\mathbf{c}^{\text{Hopf}}(\mathbf{x}, \mathbf{p}) := \mathbf{c}^{\text{Hopf}}(\mathbf{x}, \mathbf{p}_x) := (\mathbf{f}^\top(\mathbf{x}, \mathbf{p}_x, \mathbf{u}_b), \mathbf{F}^{\text{Hopf}}(\mathbf{x}, \mathbf{p}_x, \mathbf{u}_b))^\top \quad (4.10)$$

Now we can reformulate the original estimation problem (4.3) as follows. The estimated parameter, $\hat{\mathbf{p}}$ is determined as follows

$$\hat{\mathbf{p}} := \arg \min_{\mathbf{p}} \{V_N(\mathbf{p}, Z^N); \exists \mathbf{x}_b \text{ so that } \mathbf{c}^{\text{Hopf}}(\mathbf{x}_b, \mathbf{p}_x) = 0\} \quad (4.11)$$

To utilize the techniques from constrained optimization we need to formulate the constraint into an equality/inequality constraint on the free parameters only. This is done by including the state vectors \mathbf{x}_b as additional parameters in the optimization method. That means that we let the the parameter vector \mathbf{p} include the following components

$$\mathbf{p} = (\mathbf{p}_x, \mathbf{p}_y, \mathbf{x}_0, \mathbf{x}_b) \quad (4.12)$$

With this \mathbf{p} vector a final reformulation of the problem can be made

Reformulation: The estimated parameter, $\hat{\mathbf{p}}$, defined by (4.12) is identified in the following way:

$$\hat{\mathbf{p}} := \arg \min_{\mathbf{p}} \{V_N(\mathbf{p}, Z^N)\} \quad (4.13a)$$

$$\mathbf{c}^{\text{Hopf}}(\mathbf{x}_b, \mathbf{p}_x) = 0 \quad (4.13b)$$

The original problem has now been transformed to the standard nonlinear optimization problem 4.13a with the nonlinear constraint (4.13b).

Equation (4.13) is the formulation of the problem that will be used in this chapter. The most apparent drawback of this formulation is that it has introduced n new parameters, compared to the original estimation problem. As will be shown in the rest of the chapter, this drawback is outweighed by the presence of the extra condition. The extra condition vector \mathbf{c}^{Hopf} has $n + 1$ dimensions, i.e. one dimension higher than that of the new parameter vector \mathbf{x}_b , and therefore the net effect of adding both \mathbf{x}_b and \mathbf{c}^{Hopf} is a reduction of the degrees of freedom describing the parameter vector.

The two methods presented to solve (4.13) will both require that the starting point fulfills the condition. For bifurcations there is fortunately a rich literature with methods for finding such points (Allgower and Georg, 1990). These techniques are called *continuation* techniques, and one of these is described in the next section.

4.2.2 Finding a feasible point

In this section a one-dimensional continuation method is presented. Continuation methods allows for following (continuation) of the attractors (e.g. fix points and limit cycles) as the parameter $\boldsymbol{\mu}$ varies, i.e. it allows for bifurcation diagrams, like e.g. Figure 2.1 and 2.2, to be drawn. In general the varied parameters are a subset

to the μ vector, and can hence be multi-dimensional. For our purposes, however, we only need to vary one parameter at a time. Let the parameter that is varied be denoted μ_k . First the three parts of the method are briefly described and then the actual algorithm is given. The method is described as it is implemented in Janet (Janet, 2002). For a more formal introduction to the subject, and to other similar methods, the reader is referred to Allgower and Georg (1990).

Extending the system

First the parameter, μ_k , on which the continuation shall be done has to be chosen. Let then this parameter be considered as a state variable, i.e. let the vector $x = (x_1, \dots, x_n)^t$ be extended to $x_{\text{ext}} = (x_1, \dots, x_n, p_k)$, and let the parameter vector be diminished from $\mu = (\mu_1, \dots, \mu_s)$ to $\mu_{\text{dim}} = (\mu_1, \dots, \mu_{k-1}, \mu_{k+1}, \dots, \mu_s)$. This is only a redistribution of the same number of parameters and variables and a point (x, p) is therefore equal to a point $(x_{\text{ext}}, \mu_{\text{dim}})$ if the corresponding values are the same. Therefore the condition for an extended point $x_{\text{ext},f}$ to be a fix point is transformed from (4.8) to

$$f(x_{\text{ext},f}, \mu_{\text{dim}}) = 0 \quad (4.14)$$

which is the basis for calculating the predictor step.

The predictor step

The predictor step is calculated by finding a step in x_{ext} , that does not change the value of f , were it a plane. Therefore it is the solution of the following equation

$$\begin{pmatrix} \frac{df}{dx} & \frac{df}{d\mu_k} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta \mu_k \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (4.15)$$

Here the constant 1 is arbitrary and the whole last equation is simply added to obtain a system of equations which has as many unknowns as equations. Since the system also is linear it has an easily obtainable solution. The solution vector is, once obtained, added to the original extended vector to obtain the predicted value, i.e.

$$x_{\text{ext}}^{i+1,\text{pred}} = x_{\text{ext}}^i + (\Delta x, \Delta \mu_k) * h \quad (4.16)$$

where h is the somehow chosen step length.

The corrector step

The predicted value $x_{\text{ext}}^{i+1,\text{pred}}$ is, however, not necessarily on the bifurcation manifold, i.e. the function $f(x_{\text{ext}}^{i+1,\text{pred}}, \mu_{\text{dim}})$ is not necessarily equal to zero. Here Newton-Rhapson iterations finds the x_{ext} that solves the equation (4.14) and the outcome of these iterations is the corrected value $x_{\text{ext}}^{i+1,\text{corr}}$.

The algorithm

To summarize the three steps, making up one continuation step they are here written as an algorithm. The auxiliary algorithm, $\text{Newton}(x, f, J, \text{acc})$, means the Newton-Rhapson root-finding algorithm, for the function f , with Jacobian J , starting at x , and with the solution, x^* , satisfying $|f(x)| < \text{acc}$.

Algorithm 4.1. One Continuation Step

Input: The starting parameter values of the chosen continuation parameter μ_k^i as well as of the full parameter vector μ^i , corresponding fix point (x_f^i, μ^i) , step length h (including direction) and correction accuracy acc .

1. Extending the system: $x_{\text{ext}} = (x_{f,1}, \dots, x_{f,n}, p_k^i)$
 $\mu_{\text{dim}} = (\mu_1, \dots, \mu_{k-1}, \mu_{k+1}, \dots, \mu_r)$

2. Calculating the predictor step: Solve

$$\begin{pmatrix} \frac{df}{dx_{\text{ext},1}} & \cdots & \frac{df}{dx_{\text{ext},n+1}^i} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \Delta x_{\text{ext},1} \\ \vdots \\ \Delta x_{\text{ext},n+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 1 \end{pmatrix} \quad (4.17)$$

3. Taking the predictor step: $x_{\text{ext,pred}} = x_{\text{ext}} + \Delta x_{\text{ext}} * h_1$
4. Correcting: $x_{\text{ext,corr}} = \text{Newton}(x_{\text{ext,pred}}, f(x_{\text{ext}}, p_{\text{dim}}), \nabla_{x_{\text{ext}}} f(x_{\text{ext}}, \mu_{\text{dim}}), \text{acc})$
5. return $x_{\text{ext,corr}}$

Output: Parameter value μ_k^{i+1} after one continuation step and the corresponding fix point (x_f^{i+1}, μ_k^{i+1}) .

In Figure 4.1 a one-dimensional continuation plot is included as a first step in the algorithm depicted there. It is from the Brusselator example of 4.5, and the one-dimensional continuation line is the vertical line going from $(p_A, p_B) = (0.8, 1.8)$ down to $(p_A, p_B) = (0.8, 1.64)$, which is the bifurcating p_B value for this p_A .

4.3 The first optimization method - by elimination

This section describes the first method to solve (4.13). First, the steps are explained, then the steps are combined into the complete algorithm, and third, the properties of the method are shown.

4.3.1 General description

In this subsection the three parts of the algorithm are explained. The three steps are: the initial projection, the reduction of the eliminated parameters, and ordinary optimization from the feasible point in the reduced space.

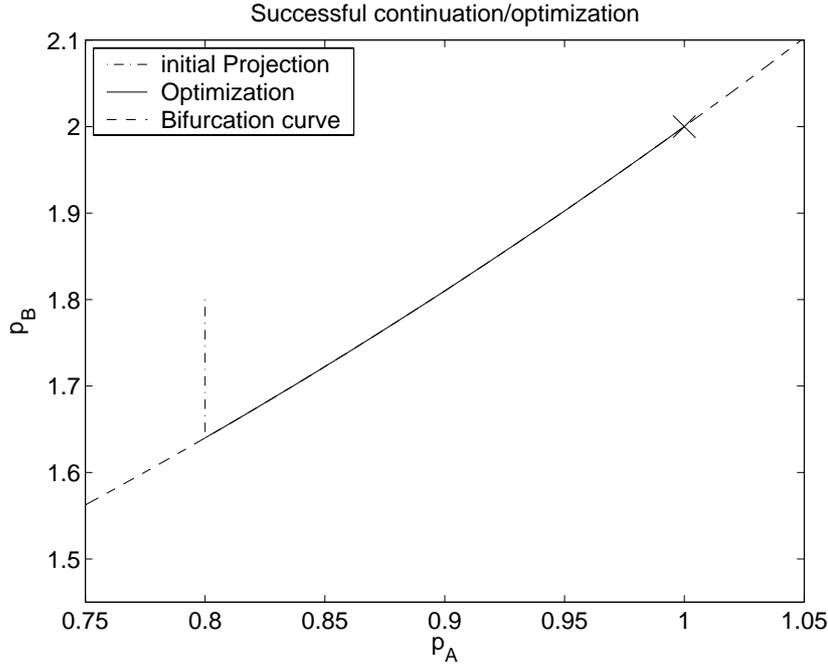


Figure 4.1 Plot of two continuations done when applying the new methods to the second sample in Section 4.5. First a projection down to the bifurcation manifold is done, this is the vertical line and it is done through the one-dimensional continuation techniques described in Section 4.2.2. Then an optimization along the manifold is done, until the optimum just above the point (1,2) is found.

Finding a feasible point

The way the elimination is done here relies on that the initial estimate, $\mathbf{p}^0 = (p_x^0, p_y^0, x_0^0, x_b^0)$, is on, or at least close to, the bifurcation manifold. This means that we want

$$\mathbf{c}^{\text{Hopf}}(\mathbf{p}_x^0, \mathbf{x}_b^0) = 0 \quad (4.18)$$

If this should not be the case for the initial estimate, i.e. if it would be unfeasible, it would have to be projected onto the bifurcation manifold as described in Algorithm 4.1. When \mathbf{c}^{Hopf} is sufficiently close to zero the exact point on the bifurcation manifold can be found by means of the Newton algorithm. Note that one could not do Newton iterations all the way since the function \mathbf{c}^{Hopf} only has gradients pointing in towards the bifurcation manifold in closest vicinity of the manifold.

Reducing the search space

There are two variables to eliminate, x_b and a chosen parameter from p_x , arbitrary called $p_{x,k}$. Then these will be rewritten as functions of the remaining p_x parameters. Let the parameters be denoted

$$p = (p_{x,1}, \dots, p_{x,k}(p_{x,free}), \dots, p_y, x_0, x_b(p_{x,free}))$$

where

$$p_{x,free} := (p_{x,1}, \dots, p_{x,k-1}, p_{x,k+1}, \dots) \quad (4.19)$$

and where the dependent parameter is chosen to fulfill the constraint (4.10) for each $p_{x,free}$. In practice the dependent parameter $p_{x,k}$ will be determined by an iterative process like e.g. the Newton-Rhapson method. Let the eliminated variables be denoted

$$p_{elim}(p_{x,free}) := (p_{x,k}(p_{free}), x_b(p_{free})) \quad (4.20)$$

Ordinary optimization in the reduced space

In this algorithm, the choice of ordinary iteration method in the reduced system is arbitrary. Once the reduction to the reduced space $(p_{x,free}, p_y, x_0)$ has been done the problem is again an ordinary unconstrained optimization problem. In the algorithm, the Levenberg-Marquardt (LM) method has been chosen (Nocedal and Wright, 1999).

4.3.2 The algorithm

By putting all three steps above together the algorithm is as follows

Algorithm 4.2. EliminationSolution

Input: Initial estimate of all parameters, $p^0 = (p_x^0, p_y^0, x_0^0, x_b^0)$, step length (sign gives direction), h , parameter for projection, $p_{x,k}$, tolerances acc_1 , acc_2 , acc_3 , perturbation size, $pert$.

A Initial projection:

- A1 if $c^{Hopf}(x_b^0, p_x^0) < acc_1$, then $(x_b^i, p_{x,k}^i) = (x_b^0, p_{x,k}^0)$, goto A4*
- A2 else $(x_b^i, p_{x,k}^i) = (x_b^0, p_{x,k}^0)$*
- A3 while $c^{Hopf}(x_b^i, p_x^i) > acc_1$ do $((x_b^{i+1}, p_{x,k}^{i+1}) = \text{OneContinuationStep}(x_b^i, p_{x,k}^i, p_{x,k}^i, p_x^i))$
 $(x_b^i, p_x^i) = (x_b^{i+1}, p_{x,k}^{i+1})$*
- A4 $(x_b^i, p_{x,k}^i) = \text{Newton}((x_b^i, p_{x,k}^i), c^{Hopf}, \text{NumPert}_{x_b, p_{x,k}}(c^{Hopf}, pert), acc_2)$*
- A5 $p^i = (p_{x,1}^0, \dots, p_{x,k}^i, \dots, p_y^0, x_0^0, x_b^i)$*

B Eliminating the $n + 1$ parameters:

- $\mathbf{p} = (\mathbf{p}_{x,1}, \dots, \mathbf{p}_{x,k}(\mathbf{p}_{x,\text{free}}), \dots, \mathbf{p}_y, \mathbf{x}_0, \mathbf{x}_b(\mathbf{p}_{x,\text{free}}))$
- $\mathbf{p}_{x,\text{free}} := (\mathbf{p}_{x,1}, \dots, \mathbf{p}_{x,k-1}, \mathbf{p}_{x,k+1}, \dots)$
- $\mathbf{p}_{\text{elim}}(\mathbf{p}_{x,\text{free}}) := (\mathbf{p}_{x,k}(\mathbf{p}_{x,\text{free}}), \mathbf{x}_b(\mathbf{p}_{x,\text{free}}))$
- where $\mathbf{p}_{\text{elim}} = \text{Newton}((\mathbf{x}_b, \mathbf{p}_{x,k}), \mathbf{c}^{\text{Hopf}}, \text{NumPert}_{\mathbf{p}_{\text{elim}}}(\mathbf{c}^{\text{Hopf}}, \text{pert}), \text{acc}_2)$
determines \mathbf{p}_{elim} given $\mathbf{p}_{x,\text{free}}$

C Ordinary optimization:

- return: $\mathbf{p} = \text{LM}((\mathbf{p}_{x,\text{free}}^i, \mathbf{p}_y, \mathbf{x}_0), \text{acc}_3)$

Output: estimated parameter $\hat{\mathbf{p}}$, where $\mathbf{c}^{\text{Hopf}}(\hat{\mathbf{p}}_x, \hat{\mathbf{x}}_b) < \text{acc}$, and $V_N(\hat{\mathbf{p}}|Z^N)$ has a local minimum, when constrained to the bifurcation manifold.

4.3.3 Properties of the algorithm

The two basic features of the method are now stated.

Theorem 4.1. Soundness of method

If a point $(\mathbf{p}_{x,}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*})$ is a minimum of the original estimation problem (4.3), and it is on the bifurcation manifold, it is also a minimum of the constrained optimization problem (4.13), given the same time series Z^N .*

Proof. Assume $\mathbf{p}_* = (\mathbf{p}_{x,*}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*})$ is a minimum of the original estimation problem (4.3), that is also a Hopf bifurcation point. The argumentation in Section 4.2.1 showed that this point will also fulfill $\mathbf{c}^{\text{Hopf}}(\mathbf{p}_*) = 0$ of (4.13). That the original point is a minimum means that

$$V_N((\mathbf{p}_{x,*}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*}) \leq V_N((\mathbf{p}_{x,*}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*}) + \delta\mathbf{p}, Z^N) \text{ for all } \delta\mathbf{p} < \epsilon \quad (4.21)$$

and hence also the following relation will be fulfilled

$$V_N((\mathbf{p}_{x,*}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*}) \leq V_N((\mathbf{p}_{x,*}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*}) + \delta\mathbf{p}, Z^N) \text{ for all } \delta\mathbf{p} < \epsilon \\ \text{where } \delta\mathbf{p} \text{ are restricted by } \mathbf{c}^{\text{Hopf}}(\mathbf{p}_{x,*}, \mathbf{p}_{y,*}, \mathbf{x}_{0,*}) + \delta\mathbf{p} = 0 \quad (4.22)$$

That means that the given point will be a minimum also to (4.13). \square

Remark: Notice that even though the algorithm gives an answer that fulfills (4.13b) this does not mean that the parameter vector describes a Hopf bifurcation. This is because (4.13b) is only a necessary constraint, and the output of the algorithm could also be another non-hyperbolic point than a Hopf bifurcation point. A final check, with e.g. a center manifold reduction, a transversality check, or a check with time series, for the output parameter $\hat{\mathbf{p}}$, must hence be done. This means that one cannot guarantee that all local minima that are outside of the bifurcation manifold, will be eliminated as solutions, even though, for minima far away from the manifold, this will be the case. There might, however, also be added new local minima to the search space, due to the restriction.

Theorem 4.2. Benefits of method

The Algorithm 4.2 converges to a minimum $(\mathbf{p}_{x,}, \mathbf{x}_{b,*})$ on the surface $c^{\text{Hopf}} = 0$ with the same convergence as the method chosen under C in the algorithm. The dimension that the method under C has to work with is one less than the same algorithm would have had to work with when applied to the original non-constrained problem.*

Proof. In the original non-constrained problem $(\mathbf{p}_x, \mathbf{p}_y, \mathbf{x}_0)$ are varied independently of each other. In Algorithm 4.2 only $(\mathbf{p}_{x,\text{free}}, \mathbf{p}_y, \mathbf{x}_0)$ are considered independent of each other, and since $\dim(\mathbf{p}_x) = \dim(\mathbf{p}_{x,\text{free}}) + 1$, the degrees of freedom for search of minima have been reduced by one. \square

4.4 The second optimization method - a reduced gradient method

The second optimization method presented is a reduced gradient method. Just as the first method this method also reduces the degrees of freedom with $n + 1$ but it does not do it with the same parameters all the time. On the contrary, it does it with all parameters, and differently for each step in the integration. This method will be introduced in the same way as the previous one was, i.e. first with a general description in words, then with the actual algorithm, and finally with statements and proofs of its properties. The initial projection is the same as before, but the rest of the method is taken from Evtushenko (1985).

4.4.1 General description

The method consists of two parts, the initial projection and the extended Cauchy method.

Finding a feasible point

Also this method needs to start at the bifurcation manifold and hence an initial projection has to be done. This is done in exactly the same way as for the previous method using a one dimensional continuation method and a Newton step.

The extended Cauchy equation

The extended Cauchy equation is written (see e.g. Evtushenko (1985)):

$$\frac{d\mathbf{p}}{ds} = -[\nabla_{\mathbf{p}} V_N(\mathbf{p}, Z^N) + (\nabla_{\mathbf{p}} c(\mathbf{p}))^T \mathbf{w}(\mathbf{p})] \quad (4.23)$$

Here we demand $\frac{dc(\mathbf{p})}{ds} = 0$ for all $\mathbf{p}(s)$ solved by (4.23). Hence

$$\frac{dc(\mathbf{p})}{ds} = -\nabla_{\mathbf{p}} c(\mathbf{p}) [\nabla_{\mathbf{p}} V_N(\mathbf{p}, Z^N) + (\nabla_{\mathbf{p}} c(\mathbf{p}))^T \mathbf{w}(\mathbf{p})] = 0 \quad (4.24)$$

and if the matrix $(\nabla_p c(p))(\nabla_p c(p))^T$ is invertible, the matrix w can be solved to:

$$w(p) = -[(\nabla_p c(p))(\nabla_p c(p))^T]^{-1}(\nabla_p c(p)\nabla_p V_N(p, Z^N)) \quad (4.25)$$

Putting this back into equation (4.23) gives

$$\begin{aligned} \frac{dp}{ds} &= -\nabla_p V_N(p, Z^N) + (\nabla_p c(p))^T [(\nabla_p c(p))(\nabla_p c(p))^T]^{-1} (\nabla_p c(p)\nabla_p V_N(p, Z^N)) = \\ &= -M(p)\nabla_p V_N(p, Z^N) \end{aligned} \quad (4.26)$$

where the projection operator M is defined according to:

$$M(p) = I - N(p) \quad (4.27)$$

$$N(p) = (\nabla_p c(p))^T [(\nabla_p c(p))(\nabla_p c(p))^T]^{-1} \nabla_p c(p) \quad (4.28)$$

4.4.2 The algorithm

Putting the two steps above together, the algorithm becomes:

Algorithm 4.3. Extended Cauchy

Input: An initial estimate $p^0 = (p_x^0, p_y^0, x_0^0, x_b^0)$, a continuation parameter $p_{x,k}$, a step length (and direction) h , tolerances and perturbations acc_i and $pert_j$.

A Initial projection:

- A1 if $c^{Hopf}(x_b^0, p_x^0) < acc_1$, then $(x_b^i, p_{x,k}^i) = (x_b^0, p_{x,k}^0)$, goto A4
- A2 else $(x_b^i, p_{x,k}^i) = (x_b^0, p_{x,k}^0)$
- A3 while $c^{Hopf}(x_b^i, p_x^i) > acc_1$ do $((x_b^{i+1}, p_{x,k}^{i+1}) = \text{OneContinuationStep}(x_b^i, p_{x,k}^i, h, p_{x,k}))$
 $(x_b^i, p_x^i) = (x_b^{i+1}, p_{x,k}^{i+1})$
- A4 $(x_b^i, p_{x,k}^i) = \text{Newton}((x_b^i, p_{x,k}^i), c^{Hopf}, \text{NumPert}_{x_b, p_{x,k}}(c^{Hopf}, pert), acc_2)$
- A5 $p^i = (p_{x,1}^0, \dots, p_{x,k}^i, \dots, p_y^0, x_0^0, x_b^i)$

B Solving the Cauchy problem:

- B1 Calculate $\nabla_p V_N(p, Z^N) = \text{NumPert}_x(V_N(p, Z^N), pert)$,
and $M(p)$ according to (4.27).
 $\text{RHS}(p) = -M(p)\nabla_p V_N(p, Z^N)$
- B2 $p_{pred}^{i+1} = \text{ODESolver}(p, \text{RHS}(p), acc_3)$
- B3 $p_{corr}^{i+1} = \text{Newton}(p_{pred}^{i+1}, c^{Hopf}, \text{NumPert}_{p_{ext}}(c^{Hopf}, pert), acc_2)$
 $- p^{i+1} = p_{corr}^{i+1}$
- B4 if $|V_N(p^{i+1}, Z^N) - V_N(p^i, Z^N)| > acc_4 * V_N(p^i, Z^N)$
then $p^i = p^{i+1}$ goto B1
 $-$ else return: $\hat{p} = p^{i+1}$

Output: An estimated state \hat{p} where $c^{Hopf}(\hat{p}) < acc_2$

4.4.3 Convergence properties

The following two theorems, taken from the book by Evtushenko (1985), shows the convergence properties and rate of the method.

Introduced the set

$$\Omega := \{\mathbf{p} \in \mathbb{R}^I; V_N(\mathbf{p}, Z^N) \leq V_N(\mathbf{p}^0, Z^N), \quad c(\mathbf{p}) = 0\} \quad (4.29)$$

Now the first convergence theorem can be stated as

Theorem 4.3. Convergence Theorem

Let the functions defining the problem (4.13) be differentiable on an open set containing Ω , where $c(\mathbf{p}) = 0$ satisfies the constraint qualifications; let the local minimum of $V_N(\mathbf{p}, Z^N)$ be attained on Ω at the unique point \mathbf{x}_ . Then the solutions of the system (4.26) converge to the point \mathbf{x}_* as $s \rightarrow \infty$.*

Proof. See Evtushenko (1985). □

The second theorem tells what the convergence rate is:

Theorem 4.4. *The convergence rate of solving a corresponding Cauchy problem, like (4.26), by means of ordinary integration, is linear.*

Remark: Hence it might be advisable to end the iterations with a few steps with an ordinary quasi-Newton method.

4.5 Identification of the Brusselator

The Brusselator is once again the model used to test the new method. The differential equations are the same as always, but the scenario differs a little from example to example. Let the sensor noise be denoted $\mathbf{n} \in N(0, \sigma)$, and let the sensor be measuring the variable x_2 in (2.15). Then (2.15) in the form of (2.4) is

$$\dot{x}_1 = p_A + x_1^2 x_2 - p_B x_1 - x_1 \quad (4.30a)$$

$$\dot{x}_2 = p_B x_1 - x_1^2 x_2 \quad (4.30b)$$

$$\mathbf{y} = x_2 + \mathbf{n} \quad (4.30c)$$

The scenario is as follows. A couple of engineers have the Brusselator in their chemical tank and they have the possibility to slowly control p_A , and p_B , but their absolute values are not yet known. The task is to calibrate the actuator, i.e. to estimate p_A and p_B . They decided to fix p_A and vary p_B in such a way that the system lies in the vicinity of the bifurcation manifold. Two different samples are collected:

Sample 1: Here they stopped just before the bifurcation $((p_A, p_B) = (1, 1.99))$. The time series is shown in the upper plot in Figure 4.2. Initial values $(x_1(0), x_2(0)) = (1, 1.99)$

Sample 2: Now they stopped just after the bifurcation $((p_A, p_B) = (1, 2.01))$. The time series is shown in the lower plot in Figure 4.2. Initial values $(x_1(0), x_2(0)) = ((1.00, 1.88))$.

To generate the time series the Runge-Kutta fourth order method has been used with time-step $h = 0.1$. The variance of the noise was $\sigma = 0.1$. Now we will see how well an optimization method performs with and without adding the constraint c^{Hopf} .

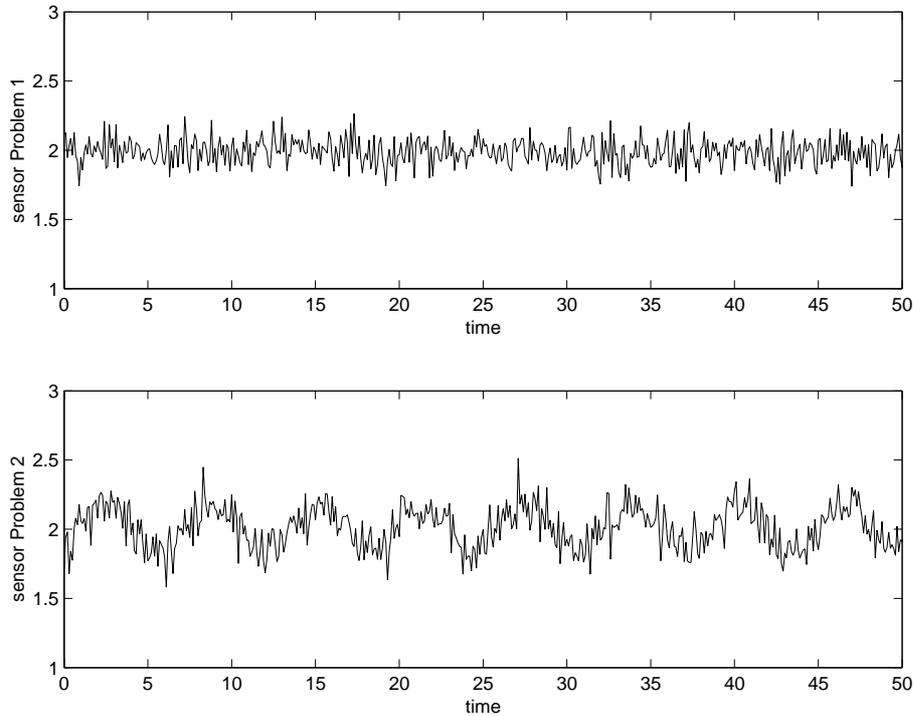


Figure 4.2 The first 50 seconds of the time series used for sample 1 and 2. For the first sample the control parameters to be determined are put just before the oscillations starts and in problem 2 just after. The series have been integrated and the parameters necessary for their regeneration are $(p_A, p_B) = (1, 1.99)$ (Sample 1), $= (1, 1.88)$ (Sample 2), $(x_1(0), x_2(0)) = (1, 1.99)$ (Sample 1), $= (1.00, 1.88)$ (Problem 2). The integrator is a fix step RK(4), step length 0.1 and $\sigma = 0.1$, $t \in [0, 300]$.

Results before adding c^{Hopf}

We will study the effect of adding the new constraint on the convergence regions, but not on other performance measures like the convergence rates etc. Finding the convergence regions is equivalent to finding the basins of attraction for the correct minimum, when studying the Cauchy formulation of (4.3). With sufficiently short iteration steps this determination will not depend on the optimization method. The equivalent Cauchy formulation is studied, and it is solved with a fourth order Runge-Kutta method. In terms of the original formulation (4.3) this is equivalent to a steepest descent method. The integrator is imbedded in a trust-region method to ensure that it does not take too long time-steps.

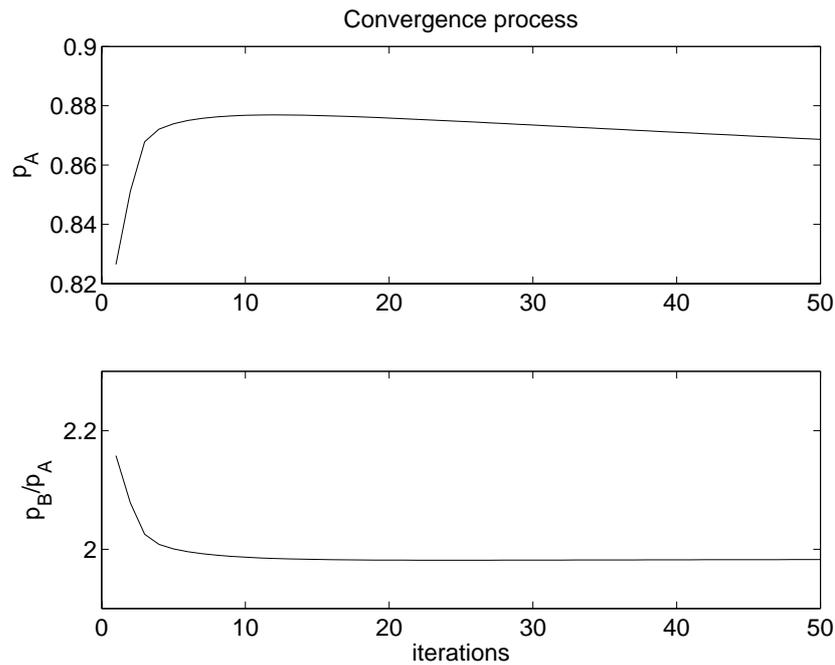


Figure 4.3 The typical behaviour of an estimation process for the first sample. The fraction $\frac{p_A}{p_B}$ converges to the correct value, while the exact values of p_A and p_B are missed. The convergence to p_A and p_B is also slower than that to the fraction. Since this convergence process is for Sample 1, there is not sufficient information in the time-series to find the correct value of p_A and p_B . However, the same phenomenon appears also for the other sample.

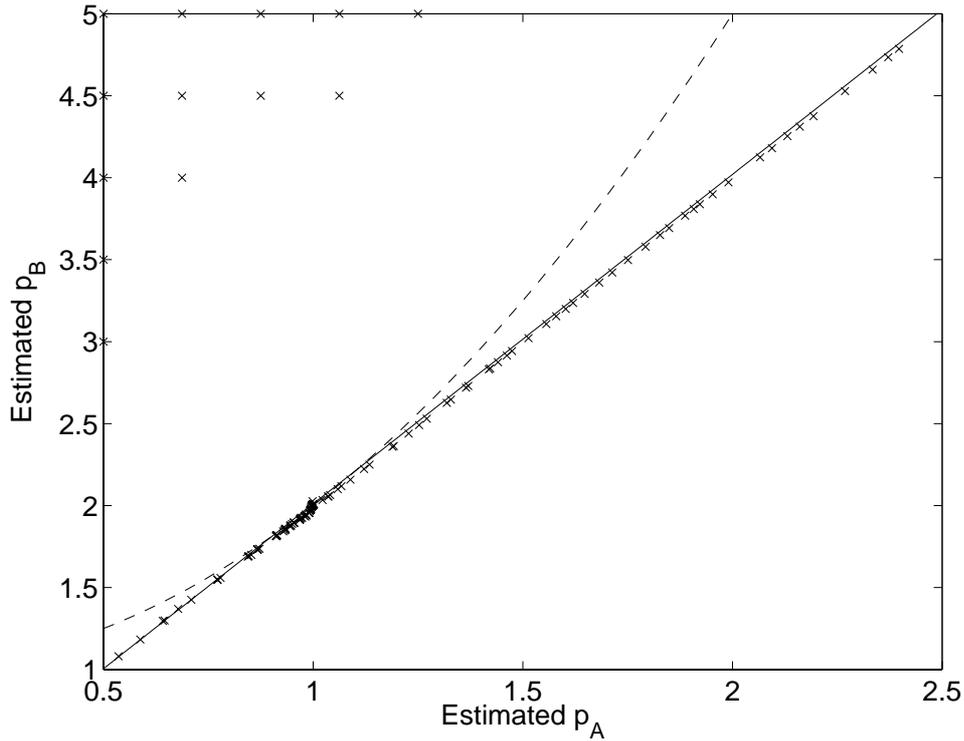


Figure 4.4 *The points to which the optimization process has converged. As can be seen there are a number of points that have converged to the correct value, $(p_A, p_B) = (1, 2.01)$. It can, however, also clearly be seen that the estimates \hat{p}_A and p_B are correlated. When adding the constraint c^{Hopf} all the estimates lie clearly on a line.*

Sample 1: Figure 4.3 shows a typical convergence process. Since there is no dynamics included in this sample, only the fix point value of x_2 is measured. This value is known to be equal to the fraction p_B/p_A , and since both these values are assumed unknown, only the correct fraction can be estimated. For Sample 1 this is equal to 1.99, and as can be seen in Figure 4.3 the algorithm converges quickly to a fraction value that is very close to the correct one. It then keeps varying p_A and p_B but keeps the fraction constant. Since the true value of p_A is 1, it can also be seen that the algorithm brings p_A in the wrong direction.

Sample 2: Here there is dynamics included in the time-series, although not any transients. Therefore more information than the correct fraction p_B/p_A can be

estimated. However, as can be seen in Figure 4.4, it is only the estimations that have started in the vicinity of the correct value $(p_A, p_B) = (1, 2.01)$, that will converge to the global minimum. Initial estimates further away will estimate the fraction p_A/p_B correctly, had the time-series been a steady state. This is understandable when looking at the objective function in Figure 4.5; it has a clear valley corresponding to the correct fraction, and only less distinct global minimum in this valley.

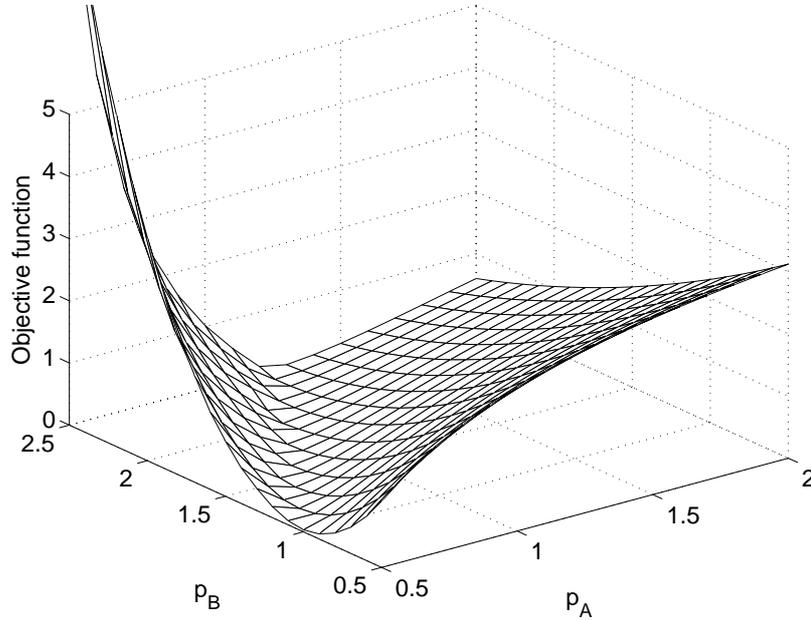


Figure 4.5 The objective function for the second sample. Here the correct value of $x_1(0)$ has been used when calculating the objective function. As can be seen there is a valley along the line $p_B/p_A = 2$, i.e. the same relation as in Sample 1 is still dominant. Now, however, there is also a small global minimum appearing in the valley, and therefore the correct value is obtained for the starting estimates close to the correct one (see Figure 4.4).

Results after adding c^{Hopf}

When adding the new parameters x_b together with the new constraint c^{Hopf} and solving the problem with one of the two suggested methods, a typical convergence process looks like the one in Figure 4.1. As can be seen the initial estimate is first

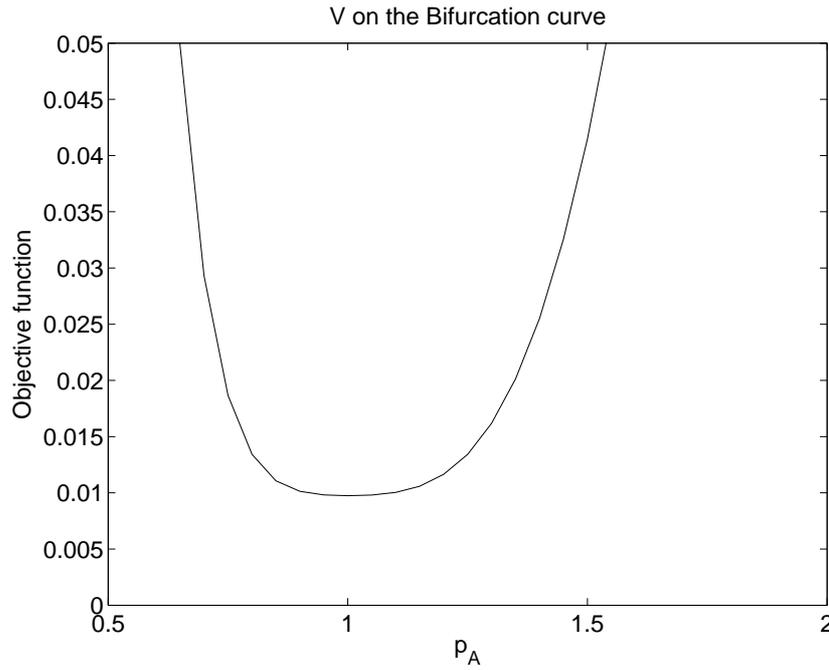


Figure 4.6 *The objective function on the bifurcation curve, plotted against the parameter p_A . Since there is only one degree of freedom left p_A uniquely determines the objective function. When comparing this objective function with the one obtained without using c^{Hopf} (see Figure 4.5) one sees that the c^{Hopf} constraint does indeed simplify the problem.*

projected down to the bifurcation manifold ($p_B = 1 + p_A^2$) and then the optimization process carries on, staying on the manifold. In Figure 4.6 the objective function on the bifurcation curve is plotted. As can be seen it is easy to find the only minimum, and therefore all points converged to the same minimum, the correct one. When comparing this result with the uncertain estimation in Figure 4.4, we conclude that the optimization processes have been improved. When comparing Figure 4.6 with Figure 4.5 we understand that the improvement is due to the reduced degree of freedom in the search space.

4.6 Conclusions

System identification of nonlinear ODEs is a difficult problem. One of the difficulties comes from the local minima in the objective function. This, and other, problems are simplified if one can reduce the search space. Here have been shown two methods that allow for a reduction in the degrees of freedom describing the p_x parameters, i.e. the parameters appearing the differential equation (4.1a). The methods are applicable for systems known to operate close to a supercritical Hopf bifurcation. The main idea is to introduce n new parameters x_b describing the fix-point corresponding to the current estimate of the p_x parameters. Then one can introduce $n + 1$ new constraints, c^{Hopf} , and the net result is that one degree of freedom has been reduced. The price for the reduction is that one has to use more advanced optimization methods. The two presented methods are two standard constrained optimization methods: one elimination method and one reduced gradient method. The former has the advantage of translating the problem back to an unconstrained optimization problem, and therefore allowing the user to choose an appropriate method for the transformed problem. The second method has the advantage of not forcing the user to choose a specific parameter to eliminate. The advantage of the new methods has been shown on the Brusselator.

Experimental Designs for Estimating the Initial Value Parameters

This is the second of three chapters that presents new identification techniques for systems of the general structure (2.4), utilizing bifurcation theory. We started the previous chapter by considering the common subdivision of \mathbf{p} into the three sub-vectors, \mathbf{p}_x , \mathbf{x}_0 and \mathbf{p}_y

$$\mathbf{p} = (\mathbf{p}_x, \mathbf{x}_0, \mathbf{p}_y) \quad (5.1)$$

In the previous chapter was shown how one could utilize the knowledge that a certain input value, \mathbf{u}_b , corresponds to a bifurcation point. The knowledge was utilized by introducing the n new parameters \mathbf{x}_b , which allowed for the given knowledge to be formulated by $n + 1$ new constraints \mathbf{c}^{Hopf} . Since there was one more constraint than new parameters, the net effect was a reduction in degrees of freedom describing the \mathbf{p}_x parameters, and two methods was presented that could handle the constraints.

In this chapter we will introduce new ways of handling the \mathbf{x}_0 vector. This parameter vector, in contrast to the \mathbf{p}_x vector, is not particularly interesting in itself, but is estimated more by necessity when forming the objective function. However, for systems with large state-spaces, due to e.g. a detailed description of the dynamics, or to the inclusion of spatial aspects, this vector can become big (>1000). In such cases, this vector alone can make a system identification problem unsolvable, even when the desired number of parameters is not so large. Therefore, methods that avoids estimating the \mathbf{x}_0 parameters as ordinary parameters are sought.

One well-known such method is to let the system start in a steady state. In this chapter is presented a way of generalizing this simple idea to systems operating on the center manifold, close to a supercritical Hopf bifurcation. In Chapter 3 we saw that the center manifold theory allows for a two-dimensional description of

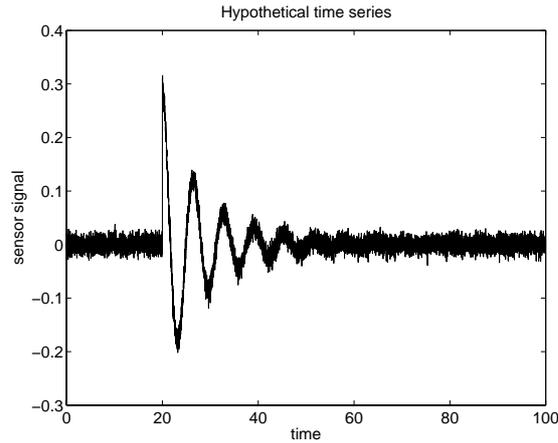


Figure 5.1 *An time series that starts in a steady state. Here the system is only singularly perturbed, but such a simple following is not necessary. If the time series starts with a stationarity, the methods of section 5.1 can always be applied, irrespectively of what happens during the rest of the time series.*

any dynamics on the center manifold, irrespectively of how large the original state-space is. Since we already developed methods that keep track of the (x_b, μ_b) vector during each step in the optimization process, it is straightforward to do the center manifold reduction at each step, and then with these parameters, describe the initial state vector. This is the basic idea behind the new methods presented in this chapter, and it is thus a combination of the methods from Chapter 3 and Chapter 4.

The material is structured as follows. In the next section the idea with starting the system in a steady state is recalled. Section 5.2 gives the basic idea, and the basic relations behind this chapter. Then, in Section 5.3 and 5.4 the method is specified for systems with arbitrary initial values (on the center manifold), and for systems with initial stationary oscillations, respectively. The chapter ends with a summary of the conclusions.

5.1 Simplifications by starting in a steady state

As a background to the contribution in the next section, this section recalls the well known advantages of starting at steady state. As in the previous chapter the

model structure is of the general form (cf. equation (2.4))

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}) = \mathbf{f}(\mathbf{x}, \mathbf{p}_x, \mathbf{u}) \quad (5.2a)$$

$$\mathbf{y} = \mathbf{y}(\mathbf{x}, \boldsymbol{\mu}, \mathbf{p}_y) \quad (5.2b)$$

$$\mathbf{x}(t_{\text{start}}) = \mathbf{x}_0 \quad (5.2c)$$

where we remind of the convention that the time is chosen so that $t_{\text{start}} = 0$ unless otherwise specified. The estimated parameters are determined by (cf (2.8))

$$\hat{\mathbf{p}} = \arg \min_{\mathbf{p}} V_N(\mathbf{p}, Z^N) \quad (5.3)$$

where (cf (2.9))

$$V_N(\mathbf{p}, Z^N) = \frac{1}{N} \sum_{t=1}^N (\mathbf{y}(t) - \hat{\mathbf{y}}(t|\mathbf{p}))^2 \quad (5.4)$$

What we are seeking now are ways of handling the parameters \mathbf{x}_0 .

In Figure 5.1 is shown a time series from a system that starts in a steady state and that is perturbed at time t_{pert} . After the perturbation the system is allowed to operate freely according to the dynamical equations (5.2a). Here the perturbation is assumed to be known. Hence if one knows the state of the system before the perturbation, one knows the state after the perturbation. Let \mathbf{u}_p be the known perturbation, and assume that the perturbation was done momentary. Then it can be formulated as an extra term $+\mathbf{u}_p\delta(t-t_{\text{pert}})$ in the differential equations. When a situation like this is possible to create, the problem of estimating initial values is as easy as it can get. No initial values has to be estimated at all, even though there are dynamics in the time series. This comes from the fact that the presence of the initial stationarity gives the n extra constraints

$$\mathbf{f}(\mathbf{x}_0, \boldsymbol{\mu}(t_{\text{steady}})) = \mathbf{0} \quad (5.5)$$

where t_{steady} is any time during which the stationarity is present. If there are only steady states (no transients) included in the time-series Z^N , the problem can also be reformulated to a nonlinear equation problem. If this latter approach is not chosen, the \mathbf{x}_0 parameters can be eliminated by e.g. the Newton method in the same way as the parameters \mathbf{x}_b were eliminated in Section 4.3. The situation of Figure 5.1 is, however, often not obtainable, and might sometimes, for other reasons, not be optimal (even from an identification point of view). Common reasons are the following

- Sometimes it is not experimentally possible to perturb the system and reveal its dynamics, and even a single perturbation might sometimes not reveal enough of the dynamics of the system. This leads to problems with identifiability, and to high uncertainties (covariances), of at least some of the parameters.

- The perturbation is not possible to do momentary. If this problem is significant the state after the perturbation will have to be assumed unknown, and hence a full state vector $\mathbf{x}(t_{\text{pert}})$ has to be estimated anyway. However, the perturbation will always be non-momentary, and hence this effect will only be a matter of degree.
- Sometimes it is not possible to have the system at a stationary state. Either because it is too expensive, or because it is technically impossible, experiments of this type cannot be done. Examples of the latter problem can be historical or economical models.

When it is not possible (or sufficient) to obtain a situation like the one described in Figure 5.1, it might still be possible (and sufficient) to bring the system close to a bifurcation, and then make observations there. To make advantage of this possibility is the topic of the rest of the chapter.

5.2 Simplifications close to a Hopf bifurcation

The initial value parameters \mathbf{x}_0 will be parametrized using minimal degrees of freedom. For each iterative step in the $(\mathbf{p}_x, \mathbf{p}_y)$ estimation \mathbf{x}_0 will be determined in a separate initial value optimization step, using the minimal parameters. The objective of this section is to explain the reasons for and feasibility of this approach.

Parametrizing \mathbf{x}_0 through \mathbf{z}_1 and \mathbf{p}_μ

To understand the common idea of the methods presented, first recall the transformation derived in Chapter 3. There it was shown that a point, \mathbf{x}_z , in the real space, but on the center manifold, can be described as (cf. equation (3.46))

$$\mathbf{x}_z = \mathbf{x}_b + \mathbf{z} + \mathbf{h}(\mathbf{z}_1, \mathbf{p}_\mu) \quad (5.6)$$

where $\mathbf{z} = \mathbf{z}_1 \mathbf{u} + \bar{\mathbf{z}}_1 \bar{\mathbf{u}}$ is a real vector in the center space and $\mathbf{p}_\mu = \boldsymbol{\mu} - \boldsymbol{\mu}_b$ where $(\mathbf{x}_b, \boldsymbol{\mu}_b)$ is the bifurcation point in the state-parameter space. Finally \mathbf{h} is the vector between the center space and the center manifold (see Figure 3.1). That means that if the system is known (or chosen, through the experiment design) to be operating on the center manifold at the time $t = t_{\text{start}}$, the initial state parameter vector \mathbf{x}_0 can be parametrized as follows

$$\mathbf{x}_0(\mathbf{p}_x, \mathbf{x}_b, \mathbf{z}_1, \mathbf{p}_\mu) = \mathbf{x}_b + \mathbf{z} + \mathbf{h}(\mathbf{z}_1, \mathbf{p}_\mu) \quad (5.7)$$

where $(\mathbf{x}_b, \mathbf{p}_x, \mathbf{u}(t_{\text{start}}))$ would be at the bifurcation point, and $\boldsymbol{\mu} = (\mathbf{p}_x, \mathbf{u}(t_{\text{start}})) + \mathbf{p}_\mu$ would be the real (time varying) value of $\boldsymbol{\mu}$.

Dimensionality of the parametrization

In equation (5.7) \mathbf{x}_0 have been parameterized by four different parameter vectors. Two of these, \mathbf{x}_b and \mathbf{p}_x , are already estimated in each step by the methods described in the previous section, and hence they do not add new complexity, if the

new methods are based on the methods of Chapter 4. The other two, z_1 and p_μ , are not a part of the estimation processes described in the previous chapter, and hence they come in as new parameters. Let us now look at the dimension of these two new parameter vectors. For a Hopf bifurcation, z_1 is always a complex scalar, and for the other one-dimensional bifurcations described in Section 2.3.1, z_1 is a real scalar. The vector p_μ on the other hand has, in general, the full dimension of μ , s , but as we shall see below, experiments can often be designed so that p_μ can be chosen as one-dimensional. In any case, if $s + 2 \leq n$, the parameter x_0 have more dimensions than the new parameters that are needed to parameterize it. Hence it is a reduction in degrees of freedom to exchange x_0 for z_1 and p_μ , according to (5.7).

Optimization problems caused by straight-forward use of z_1 and p_μ

The major problem in utilizing this parametrization right away is that there is not a continuous mapping from x_b and μ_b to x_0 , i.e. for an infinitesimal change in x_b or μ_b , but with z_1 and p_μ kept still, the change in $x_0(p_x, x_b, z_1, p_\mu)$ will not always be infinitesimal. This problem comes from the fact that the center manifold transformation (5.7) is only a local transformation and that a new transformation is done for each bifurcation point (x_b, μ_b) . This is a problem since the transformation is not unique. For the same bifurcation point, e.g. two different choices of eigenvectors u and \bar{u} , in the center manifold transformation, give rise to two different parameterizations of the center manifold, and hence to two different values of x_0 given the same z_1 and p_μ . Now follows a description of how to come around this problem, and still utilize (5.7) to parameterize x_0 in the optimization processes.

The initial value optimization sub-problem

Even though $x_0(p_x, x_b, z_1, p_\mu)$ is not uniquely determined vector, but is determined only after the choice of parametrization, the center manifold, and the dynamics on it does not depend on the choice of parametrization. The center manifold lies in the original physical space \mathbb{R}^n , and the dynamics is determined by the original equation (5.2a). The center manifold reduction (5.6) is just a way of describing this dynamics, and even though the descriptions may vary, the dynamics will be unique. This can be compared to describing a circle in different coordinate systems. The vector (r, θ, ϕ) is not a well-defined vector until one also has determined the coordinate axis, but independently of the choice of coordinate axis, the circle in the real space will always be the same. This means that a minimization of the objective function over all points on the center manifold

$$\min_{z_1, p_\mu} V(x_0(p_x, x_b, z_1, p_\mu)) \quad (5.8)$$

also leads to a unique value, i.e. a result independent of the parametrization. The idea is therefore to solve the sub-problem (5.8), over the two parameters $z_1 \in \mathbb{C}$ and $p_\mu \in \mathbb{R}$, to determine $x_0(p_x, x_b, z_1, p_\mu)$ in (5.7), for each value of the parameters, p_x , x_b and p_y . We now summarize this idea in two steps as follows.

- 1 Replace x_0 from being a freely varied parameter to an auxiliary determined by the two parameters z_1 and p_μ according to

$$x_0(p_x, x_b, z_1(p_x, p_y, x_b), p_\mu(p_x, p_y, x_b)) = x_b + z(p_x, p_y, x_b) + h(z_1(p_x, p_y, x_b), p_\mu(p_x, p_y, x_b)) \quad (5.9)$$

- 2 The two parameters z_1 and p_μ are determined from the parameters p_x, p_y and x_b according to:

$$z_1(p_x, p_y, x_b) = \arg \min_{z_1, p_\mu} V_N(p_x, p_y, x_0(p_x, x_b, z_1, p_\mu), Z^N) \quad (5.10a)$$

$$p_\mu(p_x, p_y, x_b) = \arg \min_{p_\mu, z_1, p_\mu} V_N(p_x, p_y, x_0(p_x, x_b, z_1, p_\mu), Z^N) \quad (5.10b)$$

5.3 The complete algorithm

In this section the idea presented in Section 5.2 will be written out algorithm, and the basic properties of the method will be shown. In the previous chapter x_b was determined in two different ways, first through elimination and then through a reduced gradient method. Here we will also have use of the x_b parameter, and we will determine it by the first of these methods, the elimination method. The algorithm of this section will work with the most general case studied above, and in the next sections we will go to a special case. For the general case the estimated parameters, \hat{p} , are found as those solving

$$\hat{p} = \arg \min_p V_N(p, Z^N) \quad (5.11)$$

where

$$p = (p_{x,free}, p_{x,elim}(p_{x,free}), p_y, x_0(p_x(p_{x,free}), p_y, x_b(p_{x,free}))) \quad (5.12)$$

and where (cf Section 4.3):

$$p_{x,free} := (p_{x,1}, \dots, p_{x,k-1}, p_{x,k+1}, \dots) \quad (5.13)$$

$$p_{x,elim}(p_{x,free}) := (p_{x,k}(p_{x,free}), x_b(p_{x,free})) \quad (5.14)$$

and, by recalling (5.10)

$$x_0(p_x, p_y, x_b) := x_b + z(p_x, p_y, x_b) + h(z_1(p_x, p_y, x_b), p_\mu(p_x, p_y, x_b)) \quad (5.15a)$$

$$z_1(p_x, p_y, x_b) := \arg \min_{z_1, p_\mu} V(p_x, p_y, x_0(p_x, x_b, z_1, p_\mu), Z^N) \quad (5.15b)$$

$$p_\mu(p_x, p_y, x_b) := \arg \min_{p_\mu, z_1, p_\mu} V_N(p_x, p_y, x_0(p_x, x_b, z_1, p_\mu), Z^N) \quad (5.15c)$$

The elimination method has the benefit of calling an ordinary optimization method, for the eliminated system, and hence it does not limit the user to using only one type of optimization methods, or to one rate of convergence etc.

Since elimination methods calls an ordinary optimization method for the optimization in the reduced space, the most important thing to specify is how to calculate the full parameter vector, given only the parameters that are considered free to vary. Here the only parameters that are considered independent of each other are $p_{x,free}$ and p_y , and given previous estimates of the dependent parameters, $z_{1,0}$, p_μ and p_{elim} , the dependent parameters will be estimated in the following way:

Algorithm 5.1. Fullp($p_{x,free}, p_y, p_{elim}^i, z_{1,0}^i, p_\mu^i, acc$)

Input: The freely varied parameters $p_{x,free}$ and p_y , initial estimates of the dependent parameters $p_{elim}^i, z_{1,0}^i$ and p_μ^i , acceptances acc_i

- 1 $p_{elim} = \text{Newton}(p_{elim}^i, c^{\text{Hopf}}, \text{NumPert}_{p_{elim}}(c^{\text{Hopf}}, \text{pert}), acc)$
- 2 $\hat{z}_1, \hat{p}_\mu = \text{LM}((z_{1,0}^i, p_\mu^i), V_N((p_x, p_y, x_0(p_x, x_b, z_1, p_\mu)), Z^N), \nabla_{z_1, p_\mu} V_N, acc)$
where $x_0(p_x, x_b, z_1, p_\mu)$ is decided according to (5.15).
- 3 *Return:* $x_0(p_x, x_b, \hat{z}_1, \hat{p}_\mu)$ according to (5.15).

Output: The parameters $p_{elim}(p_{x,free})$, and $x_0(p_x, p_y, x_b)$ calculated according to (5.14) to (5.15).

Since the full parameter set p can hence be determined by the freely varied parameters $p_{x,free}$ and p_y only, p will for the rest of this subsection be denoted $p(p_{x,free}, p_y)$. Hence the objective function can be written $V_N(p(p_{x,free}, p_y), Z^N)$ and the following main algorithm will look as follows.

Algorithm 5.2. Fullbifurcation – optimization algorithm

Input: A time series Z^N collected when the system has, at $t = 0$ been operating close to a Hopf bifurcation, a model structure of the form (5.2a), initial estimates of the needed parameters $p_{x,free}^0, p_y^0, p_{elim}^0, z_1$ and p_μ^0 , acceptances and perturbation sizes, acc_i and pert_j .

A Initial projection:

- A1 if $c^{\text{Hopf}}(x_b^0, p_x^0) < acc_1$, then $(x_b^i, p_{x,k}^i) = (x_b^0, p_{x,k}^0)$, goto A4
- A2 else $(x_b^i, p_{x,k}^i) = (x_b^0, p_{x,k}^0)$
- A3 while $c^{\text{Hopf}}(x_b^i, p_x^i) > acc_1$ do $((x_b^{i+1}, p_{x,k}^{i+1}) = \text{OneContinuationStep}(x_b^i, p_{x,k}^i, p_{x,k}^i, p_x^i))$
 $(x_b^i, p_x^i) = (x_b^{i+1}, p_{x,k}^{i+1})$
- A4 $(x_b^i, p_{x,k}^i) = \text{Newton}((x_b^i, p_{x,k}^i), c^{\text{Hopf}}, \text{NumPert}_{x_b, p_{x,k}}(c^{\text{Hopf}}, \text{pert}_1), acc_2)$
- A5 $p^i = (p_{x,1}^0, \dots, p_{x,k}^i, \dots, p_y^0, x_0^0, x_b^i)$

B Ordinary optimization:

- B1 $(\hat{p}_{x,free}, \hat{p}_y) = \text{LM}((p_{x,free}, p_y), V_N(p(p_{x,free}, p_y), \text{NumPert}_{p_{x,free}, p_y}(V_N(p(p_{x,free}, p_y), \text{pert}_2), acc))$
- B2 $\hat{p} = \text{Fullp}(\hat{p}_{x,free}, \hat{p}_y, p_{elim}^j, z_{1,0}^j, p_\mu^j, acc_3)$, where $p_{elim}^j, z_{1,0}^j, p_\mu^j$ are obtained from the last step of the LM iterations.

B3 Return \hat{p}

Output: An estimated parameter \hat{p} , fulfilling $c^{\text{Hopf}}(\hat{x}_b, \hat{p}_x) < \text{acc}_2$.

Remark: The μ value at each point in time will be $\mu(t) = (\hat{p}_x, u(t)) + \hat{p}_\mu$.

5.3.1 Properties of the algorithm

The soundness of the method is given by the following theorem.

Theorem 5.1. Soundness

For time series collected sufficiently close to the bifurcation point, and on the center manifold, the minima of (5.11) coincides with the minima of (2.8).

Proof. That the system operates sufficiently close to the manifold means that the error from approximating the center manifold with an expansion will not be significant. The minima will coincide if

$$\min_{u,v} f(u,v) = \min_u \min_v f(u(v), v)$$

and this is one of the basic results of optimization theory (Nocedal and Wright, 1999). \square

The benefit of the method is most apparent if the time series have been collected by an experiment of the following type.

Definition 5.1. *By a bifurcation experiment we will mean an experiment where the system at $t = 0$ is at a point from where the system can come to the Hopf bifurcation manifold by varying one parameter μ_b only. Further the distance from the bifurcation manifold is so small that the center manifold approximation is valid.*

The benefit of the method can then be described by the following theorem.

Theorem 5.2. Benefit

Assume an experiment of the type described in definition 5.1 has given rise to the time series Z^N . Then the algorithm 5.2 will reduce the degrees of freedom describing the x_0 parameters from n , which is the dimension of the state equations, to 3.

Proof. Equation (5.15a) gives the relationship. When a bifurcation experiment, as described in Definition 5.1 has been done, p_μ , in equation (5.15a) can be chosen as $p_\mu = (0, \dots, p_{\mu,b}, 0, \dots)$, i.e. it can be described by one degree of freedom only. Since z_1 will always be two-dimensional, the total degrees of freedom describing x_0 in (5.15a) will be 3. \square

5.4 Quenching data

Quenching data can be obtained by a specific experiment performed close to a Hopf bifurcation. For this type of data a further simplification of the initial value sub-problem can be obtained. In this section quenching experiments and quenching data are introduced, whereas the theory will follow in the next section.

5.4.1 What is quenching data?

Let us now go back to quenching data of the type successfully used in (Hynne et al., 2001). In Figure 5.2 is shown the result of a successful validation with the model developed in that paper. The same plot was shown in Chapter 2 (in Figure 2.3), and then it was included as a motivation for developing methods for this kind of data. The methods used in the original article did not allow time series to be included in the estimation data, and this was one of the main improvements that could be done. The method used there was also not formulated as a standard optimization problem, and therefore they did not have access to the standard machinery available for these problems. This was the other main improvement needed. Since then, in Chapters 3 to 5, we have developed general methods for systems operating around Hopf bifurcations, and now we are finally ready to turn back, and see what we have achieved. As can be seen in Figure 5.2, these time series starts with stationary oscillations, i.e. with the system operating on the center manifold. Then, at a known point and in a known way, the system is perturbed away from the stationary state and the response is studied. The experimental set-up used to collect this data has been considered the standard set-up supposed available throughout the whole thesis. This set-up has as its main features that

- The system can be brought to a supercritical Hopf bifurcation by variations in a single control variable (in (Hynne et al., 2001): through the external glycolysis concentration)
- Variables can be perturbed with a single pulse, and some of them can be perturbed so much that the effect on the oscillations is clearly visible (by hitting the stable manifold). Since many directions can be perturbed, the given data is, in this aspect, rich of information.
- One has only possibility to measure time-series of one of the chemicals, and this is the major limitation of the data. These time-series are obtained e.g. through fluorescence techniques, and therefore, compared to the time-scales appearing in the dynamics of the system, one can get arbitrarily high sample frequencies, and thus also arbitrarily many data points.

Finally should be added a comment on why the agreement in Figure 5.2 is so impressive, and why quenching data is so rich of information. As can be seen in the schematic drawing in Figure 5.3 a successful quenching of oscillations requires a precise perturbation. The oscillation quenching is due to the system hitting the stable manifold (see Section 3.1) and when this is hit the system will follow this manifold (since it is an invariant manifold) toward to unstable fix point, and then (since there is noise and since it will never hit the manifold perfectly) slowly spiral out towards the stable limit cycle. If one changes from the successful perturbation only slightly, in either phase or amplitude, the stable manifold will be missed and the projection of the oscillations to the measured variable will hardly change at all. This high demand on the precision of the perturbation is the reason why the

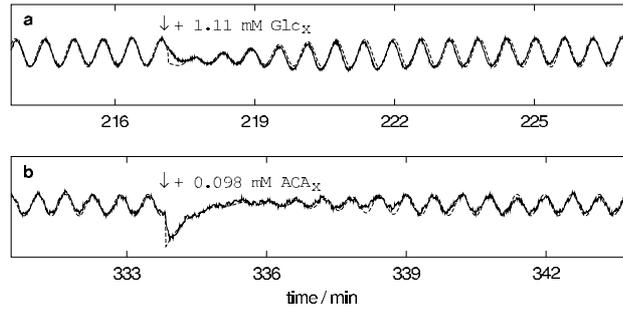


Figure 5.2 Plot showing simulated vs. experimental data from a quenching experiment, i.e. from a system that has been perturbed when operating close to a Hopf bifurcation. For a quenching experiment to substantially change the oscillations, as is done in both these time series, both the timing and the amplitude of the perturbation needs to have precise values. Considering that this data is taken from a large-scale bio-chemical model, and the comparison is from the verification step (see Section 2.2) one concludes that the agreement is quite good. This result is one of the previous indications that it might be beneficial to operate around a Hopf bifurcation. Picture included with permission from the authors.

agreement in Figure 5.2 is so impressive, and why there is such rich information in quenching data. If one knows Z^N is quenching data, some of the steps in the general algorithm, described in Section 5.3, can be simplified. Now follows a description of how this can be done.

5.5 Simplifying the initial value sub-problem

One thing that can be simplified in the special case of quenching data is the optimization sub-problem in step 2 in Algorithm 5.1

$$\min_{z_1, p_\mu} V(x_0(p_x, x_b, z_1, p_\mu))$$

This is done in two steps, first by parametrizing z_1 in p_μ and an additional parameter α , and then by utilizing the interpretation of these two parameters when performing the optimization.

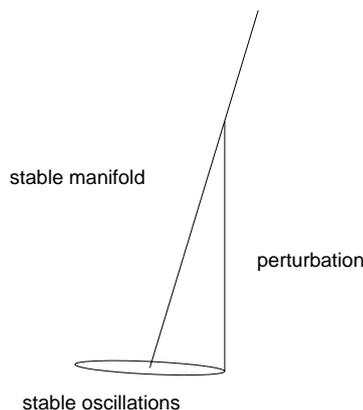


Figure 5.3 This figure shows how the oscillations are quenched in a quenching experiments. The system is perturbed from the stationary oscillations in one of the coordinate directions (by e.g. adding a chemical). If the system is perturbed in the exact right amplitude and the exact correct phase the stable manifold, originating from the unstable fix point in the middle of the oscillations, will be hit. Since this is an invariant manifold the system will follow it to the fix point and only after this spiral out to the oscillations again. This particular effect is highly sensitive to variations in both the phase and the amplitude of the perturbation and this is the reason why this kind of data is highly informative.

Formulation in minimal degrees of freedom

For quenching data the initial state vector $\chi(0)$ lies not only on the center manifold, but in the stationary limit cycle. Let us now restate the dynamic equation, in its normal form, on this manifold and also its explicit solution. Let $z_1 \in \mathbb{C}$ denote the first center space coordinate, $\omega_0 \in \mathbb{R}$ the imaginary part of the first critical eigenvalue, and let $\sigma_1 \in \mathbb{C}$ and $g_3 \in \mathbb{C}$ be the two complex (scalar) resonant terms. Finally let $p_\mu \in \mathbb{R}$ be the distance from the bifurcation point in the single control parameter that has been varied in the quenching experiment. Then the dynamics on the center manifold can (with only up to third order terms included) be described by (cf. equation (3.58))

$$\dot{z}_1 = (\omega_0 i + p_\mu \sigma_1) z_1 + g_3 |z_1|^2 z_1 \quad (5.16)$$

Now let $R \in \mathbb{C}$ and $\alpha \in \mathbb{R}$. An analytical solution to (5.16), describing the stationary behaviour, is given by

$$z_1(t) = \begin{cases} 0 & \text{if } p_\mu \sigma_1 \leq 0 \\ R \exp(i(\omega_1 t + \alpha)) & \text{if } p_\mu \sigma_1 > 0 \end{cases}$$

where R , ω_1 and α are determined by

$$0 = p_\mu \sigma_1 + \operatorname{Re}(g_3)|R|^2 \quad (5.17a)$$

$$\omega_1 = \omega_0 + \operatorname{Im}(g_3)|R|^2 \quad (5.17b)$$

$$z_1(0) = R \exp(i\alpha) \quad (5.17c)$$

By deriving the expression for $z_1(t)$ with respect to time, and substituting the result in equation (5.16) one sees that these expression fulfills the dynamics.

Let Z^{stat} denote the initial stationary part of the oscillations, i.e. if t_{stat} denotes the time the stationarity ends, then

$$Z^{\text{stat}} = \{u(t), y(t)\}_{t=0}^{t_{\text{stat}}} \quad (5.18)$$

From equation (5.17c), (5.2b) and equation (5.6) the simulated correspondence to Z^{stat} can be generated as follows

$$y = y(x, \mu, p_y) \quad (5.19a)$$

$$x = x_b + z + h(z_1, p_\mu) \quad t = 0, \dots, t_{\text{stat}} \quad (5.19b)$$

$$z_1 = R \exp(i(\omega_1 t + \alpha)) \quad (5.19c)$$

$$0 = p_\mu \sigma_1 + \operatorname{Re}(g_3)|R|^2 \quad (5.19d)$$

$$\omega_1 = \omega_0 + \operatorname{Im}(g_3)|R|^2 \quad (5.19e)$$

and where p_μ and α are still to be determined. Notice that this means that we can *calculate* $\hat{y}(t|p)$ (in (5.4)) for the whole interval $t = 0, \dots, t_{\text{stat}}$. Notice also that we only have to compute those rows in equation (5.6) that are needed to form the x_i s that can be measured. Loosely we can say that the dynamics is calculated, in the two-dimensional space where it occurs, and then projected to only those directions that are measured. Specifically, we can do this projection for the time $t = t_{\text{stat}} = 0$. This means that we can simplify (5.15) into

$$x_0(p_x, p_y, x_b) := x_b + z(p_x, p_y, x_b) + h(z_1(p_x, p_y, x_b), p_\mu(p_x, p_y, x_b)) \quad (5.20a)$$

$$z_1(p_x, p_y, x_b) = R(p_x, p_y, x_b) \exp(i\alpha(p_x, p_y, x_b)) \quad (5.20b)$$

$$R(p_x, p_y, x_b) := \arg \min_{R, p_\mu, \alpha} V_N(p_x, p_y, x_0(p_x, x_b, \alpha, p_\mu), Z^{\text{stat}}) \quad (5.20c)$$

$$\alpha(p_x, p_y, x_b) := \arg \min_{\alpha, p_\mu, \alpha} V_N(p_x, p_y, x_0(p_x, x_b, \alpha, p_\mu), Z^{\text{stat}}) \quad (5.20d)$$

$$p_\mu(p_x, p_y, x_b) := \arg \min_{p_\mu, p_\mu, \alpha} V_N(p_x, p_y, x_0(p_x, x_b, \alpha, p_\mu), Z^{\text{stat}}) \quad (5.20e)$$

Notice that the difference between equation (5.15) and (5.20) is that we now only have to minimize over two degrees of freedom, p_μ and α . Previously it was three degrees of freedom, p_μ and z_1 . The two degrees of freedom, p_μ and α , corresponds to two clearly interpretable quantities, the amplitude and the phase, and this can be utilized when solving the optimization problem. Let us now see how this can be done.

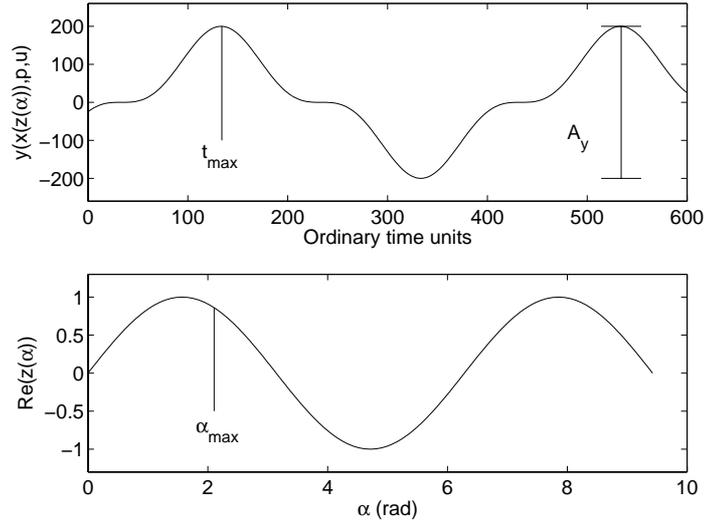


Figure 5.4 This figure shows how t_{\max} and A_y in Algorithm 5.3 is defined. As can be seen the physical oscillations can be much more complex than the oscillations in the complex plane.

The optimization

When solving the optimization sub-problem (5.20c) to (5.20e) we can solve it through an ordinary optimization algorithm, like e.g. the Levenberg-Marquardt method used earlier. Here, however, we know the analytical solution (equation (5.19c)), and the interpretation of the two degrees of freedom (R is the amplitude and α is the phase) in the normal form equation, and this can be utilized to perform the optimization easier. Other reasons are the desire to have more control of what is going on, and to avoid the problem of getting stuck in local minima.

We will assume that we have determined the amplitude of the oscillations, A_y , and the a specific maximum time, t_{\max} , from the time-series of measured of one of the measured variables (see Figure 5.4). As can be seen in Figure 5.4, the steady state oscillations in the complex plane are nice, but the oscillations in the physical space are more complex. It is, however, assumed that there is a unique maximum in the measured oscillations, considering one period only. It should also be added that as one approaches the bifurcation, from the unstable side, both the amplitude and the complexity of the oscillations in the measured variable decreases. This comes from the fact that for small z , i.e. for small distances from the bifurcation point, (p_x, x_b) , only the linear terms in the transformation h will be dominant. This is also the reason why the amplitude will grow monotonically as p_μ is increased. These two properties are the reason why α and R (in equation (5.17c)) can be determined from the amplitude A_y and the time-point t_{\max} by the following simple algorithm.

Algorithm 5.3. *Input: The amplitude A_y , the time at which the measured variable is at its maximum, t_{\max} , the necessary row in the h function, increments Δp_μ and $\Delta\alpha$*

```

 $p_\mu = 0, A(p_\mu) = 0$ 
while  $A(p_\mu) < A_y$ 
     $\alpha = 0, y(\alpha) = 0, y_{\min} = \infty, y_{\max} = 0$ 
    while  $\alpha \in (0, 2\pi)$ 
        if  $y(\alpha, p_\mu, t_{\max}) < y_{\min}$   $y_{\min} = y(\alpha), \alpha_{\min} = \alpha$ 
        if  $y(\alpha, p_\mu, t_{\max}) > y_{\max}$   $y_{\max} = y(\alpha), \alpha_{\max} = \alpha$ 
        *  $\alpha = \alpha + \Delta\alpha$ 
    end while
     $A(p_\mu) = y_{\max} - y_{\min}$ 
     $p_\mu = p_\mu + \Delta p_\mu$ 
end while
• return:  $R(p_x, p_y, x_b) = R(p_\mu), \alpha = \alpha_{\max}, p_\mu(p_x, p_y, x_b) = p_\mu$ 
Output: Values of  $R(p_x, p_y, x_b), \alpha(p_x, p_y, x_b)$  and  $p_\mu(p_x, p_y, x_b)$  according to (5.20c) to (5.20e).

```

Remark: Note that there are only two degrees of freedom for the three parameters R, p_μ and α . In this algorithm this was solved by only varying p_μ and α , and determine $R(p_\mu)$ according to (5.17a).

The idea behind the algorithm is to utilize that the amplitude grows monotonically as the distance p_μ increases. The amplitude for each p_μ is simply determined by going through all angles α and detecting the min and max values, measured in y , and finally by taking the difference between the two. When this is done for increasing distances p_μ the amplitude will increase until it eventually hits the measured amplitude. This algorithm is simple and has a clearly defined stopping criterion.

5.6 Finding the initial values in the Brusselator

Let us now go back to the example in Chapter 4.5. In that example were collected two different time-series: one on the stable side of the bifurcation (Sample 1), and one on the unstable (Sample 2). Both of the time series were close to the bifurcation, and this information was shown to be sufficient to determine the correct parameters $(p_A, p_B) \sim (1, 2)$, in both cases. However, in both cases the time series were treated as stationary time series, i.e. the method described in Section 5.1 was used. That means that there were not estimated any corresponding initial values for Sample 2, and this problem we turn to now. The problem is solved

by Algorithm 5.3. By inspection of Sample 2, the measured amplitude, A_y , in Algorithm 5.3 was estimated as $A_y = 0.3270$. The value of t_{\max} , in the same algorithm, was estimated as $t_{\max} = 2.6s$. The estimation of the parameters p_x and x_b done earlier gave $(\hat{p}_x, \hat{x}_b) = (1, 2, 1, 2)$, and the normal form transformation around this bifurcation point was done in Section 3.4. The relationship between the distance, p_μ from the bifurcation point, and the amplitude of the oscillations, is given in Figure 5.5. As can be seen the amplitude grows monotonically with the distance beyond the interesting region, and the assumption behind Algorithm 5.3 is thus valid. In Figure 5.5 can further be seen that the normal form approximation for distances around the correct distance ($p_\mu = 0.01$) is a good approximation, even though the error eventually increases. The correct distance value was estimated with an accuracy of $1E - 4$ (which was the step length Δp_μ), and the phase α was estimated to 3.79 rad. The final agreement is shown in Figure 5.6. As can be seen the agreement between the estimated and measured output is good already from the beginning of the time series. The correct initial values have thus been estimated.

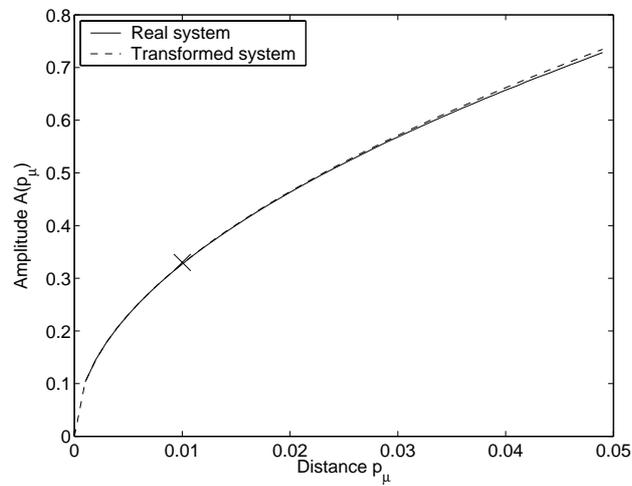


Figure 5.5 The amplitude of the stationary oscillations as a function of the distance from the bifurcation point. The plot is taken from the problem on the Brusselator described in Section 5.6. That the amplitude is growing monotonically with the distance is, however, a generic feature of the Hopf bifurcation. This is the feature utilized in Algorithm 5.3.

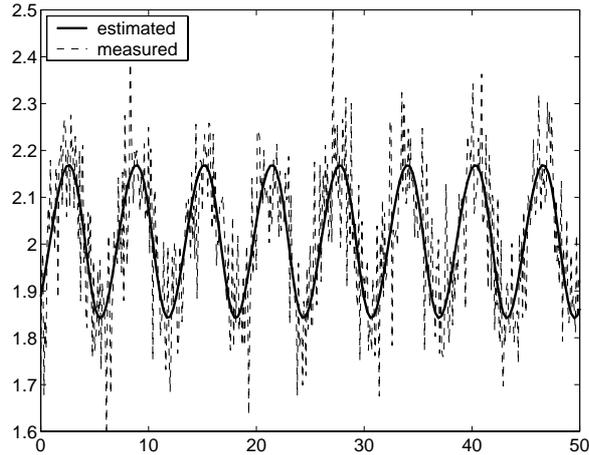


Figure 5.6 Plot showing the agreement between the estimated system and the original time series for the second sample introduced in Section 4.5. The initial parameters were estimated by Algorithm 5.3.

5.7 Conclusions

This is the second of three chapters that deals with estimation of the parameter vector $\mathbf{p} = (\mathbf{p}_x, \mathbf{x}_0, \mathbf{p}_y)$, and in this chapter ways of handling the initial state parameters \mathbf{x}_0 have been presented. A well known method is to let the system start at a steady state, when this is possible. This simple method has been generalized to allow time series starting on an arbitrarily place on the center manifold surrounding a Hopf bifurcation. This is done by parametrizing \mathbf{x}_0 in minimal degrees of freedom, \mathbf{z}_1 and \mathbf{p}_μ , using equation (5.7). There are problems associated with straightforward usage of the new parameters \mathbf{z}_1 and \mathbf{p}_μ . These have been overcome by solving the initial value optimization sub-problem (5.10), for each step in the $(\mathbf{p}_x, \mathbf{p}_y)$ optimization.

There exists published data, e.g. quenching data (see Figure 5.2), that starts not only on the center manifold, but in stationary oscillations. For such data the initial value optimization sub-problem can be further simplified. For the case of stationary oscillations, on the center manifold, one can parametrize the parameter vector \mathbf{z}_1 , through the two parameters α and \mathbf{p}_μ . Then \mathbf{x}_0 is described by two degrees of freedom only, and these correspond directly to the phase and the amplitude of the measured oscillations. It has been shown how the reformulation $\mathbf{z}(\mathbf{p}_\mu, \alpha)$ can be used to perform the initial value optimization sub-problem in a straight-forward way, with the advantage of avoiding any numerical integration. The possibility of using α is due to the combination of the center manifold and normal form reduction, which reveals the special structure of the Hopf bifurcation.

The parametrization of x_0 in p_μ and z_1 , or in p_μ and α , can be performed independently of the original dimension of x_0 . It has thus been shown how the handling of the initial state parameters can be simplified when working close to Hopf bifurcation.

Qualitative Model Validation

The last two chapters have presented new methods for the identification step in the system identification loop, that incorporate knowledge of where there is a bifurcation in the system. We will now look for another way of using this information. An alternative to include extra knowledge in the identification step, is to use knowledge to reject estimations, during the validation step in the system identification loop (see Section 2.2). This chapter presents a test quantity that, by determining the qualitative behaviour of a system, can be used for model validation. Here qualitative determination of a system means determination of if and when a system has oscillations, and if and when a system has steady state. Determination of the qualitative behaviour of the real system will be done through spectral analysis (or by mere inspection), and determination of the qualitative behaviour of the estimated system will be done with continuation methods and bifurcation theory. The method has the advantage that it can use uncalibrated sensors for determination of the qualitative behaviour of the real system. An uncalibrated sensor corresponds to unknown p_y parameters, or sometimes even to unknown $y(\cdot)$ functions in equation (2.2). We can therefore now, by again considering the parameter vector $p = (p_x, p_y, x_0)$, see the relationship between the last three chapters. Chapter 4 considered reductions of the p_x parameters, which often are the only parameters one really seeks. This reduces the search space which in turn might e.g. eliminate some erroneous local minima. Chapter 5 (with usage of theory from Chapter 3) introduced new ways of handling the x_0 parameter. For models with large state-space descriptions and with few sensors these parameters can otherwise be a severe problem. Now, this chapter introduces ways of getting around problems associated with the remaining parameters, namely those in the p_y vector, and also the

case when the functional form of $\mathbf{y}(\mathbf{x}, \mathbf{p}, \mathbf{u})$ is unknown. This means that after this chapter we have shown identification improvements for all parts of the parameter vector \mathbf{p} for systems operating in the neighborhood of a Hopf bifurcation.

In Section 6.1 a specification of the problem and conditions of this chapter is given. In Section 6.2.1 the basic idea behind the test quantity is explained, and in Sections 6.2.2 to 6.2.3, its exact formulation for constant and time-varying inputs is given. In Section 6.3, the test quantity is applied to our standard example, the Brusselator, and in Section 6.4 comes the conclusions.

6.1 Detecting erroneous estimations

Now follows a description of the situation in which this chapter will bring a contribution. First is assumed that the dynamics of the system is modelled by a nonlinear ODE.

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{p}_x, \mathbf{u}) + \mathbf{f}_x = \mathbf{f}(\mathbf{x}, \boldsymbol{\mu}) + \mathbf{f}_x \quad (6.1)$$

Here the interpretation of all the ingoing symbols are the same as for equation (2.4a), except for \mathbf{f}_x which is a fault signal, describing the difference between the fault-free and the faulty dynamics. It is also assumed that there are only two qualitatively different stationary states in the system, oscillations and fix points, and that there is only one bifurcation in the system, a supercritical Hopf bifurcation. It is, however, not necessary that neither the bifurcation manifolds nor the basins of attractions have been determined. It is further assumed that there exists at least two groups of sensors in the system, \mathbf{y}_D and \mathbf{y}_P

$$\mathbf{y}_D = \mathbf{y}_D(\mathbf{x}, \boldsymbol{\mu}, \mathbf{p}_y) + \mathbf{f}_D \quad (6.2)$$

$$\mathbf{y}_P = \mathbf{y}_P(\mathbf{x}, \boldsymbol{\mu}, \mathbf{p}_y) + \mathbf{f}_P \quad (6.3)$$

Here the sensors \mathbf{y}_D are used to observe the dynamical behaviour of the system, and the sensors \mathbf{y}_P are used for the parameter estimation. The two groups are not necessarily disjoint and they both have sensor faults, \mathbf{f}_D and \mathbf{f}_P , modelled as additive signals.

In the problem one has collected a time-series

$$Z^N = \{\mathbf{U}, \mathbf{Y}_D, \mathbf{Y}_P\} = \{\mathbf{u}(t), \mathbf{y}_D(t), \mathbf{y}_P(t)\}_{t=0}^N$$

Here it is the \mathbf{U} and \mathbf{Y}_P part that has been used for the parameter estimation. Denoting the estimation procedure ϕ we thus have

$$\phi : \{\mathbf{Y}_P, \mathbf{U}\} \rightarrow \hat{\mathbf{p}}_x = \mathbf{p}_{x,0} + \mathbf{f}_{est} \quad (6.4)$$

It is assumed that the estimation has already been done. Here the fault \mathbf{f}_{est} means that instead of having estimated the true value $\mathbf{p}_{x,0}$ the value $\mathbf{p}_{x,0} + \mathbf{f}_{est}$ has been estimated. The fault might be due to a failed estimation, due to e.g. the convergence to an erroneous local minima. Such a fault is interesting to detect in

itself if the estimated parameter is used for something else, like e.g. in a control algorithm. The detection of such a fault is also interesting as a part of the validation step in the system identification loop. Apart from being the result of an erroneous convergence the fault can also be due to one of the faults f_x and f_p , and a detection of f_{est} is hence useful also for detection of these two faults, i.e. for usage in a diagnosis system. For a sensor to be one of the sensors denoted y_D is only required that it is dependent on some of the dynamic variables, i.e.

$$\frac{\partial y_D}{\partial x_i} \neq 0 \quad (6.5)$$

for some x_i that is oscillating on the unstable side of the bifurcation. Here should be emphasized that the exact relation $y_D(\cdot)$ is not necessarily known, i.e. the sensor does not have to be calibrated. It is, however, assumed that y_D is continuous.

Now, the problem that this chapter is attacking is to find ways of detecting faults f_{est} (defined by equation (6.4)), by utilizing one of the uncalibrated sensors y_D combined with the knowledge of which bifurcations that exists in the system.

6.2 The test quantity

In this section the test quantity is defined. First the basic idea is introduced and the high-level steps of the defining algorithm are given. Then follows a more detailed description of the three steps in the algorithm, first for constant, and then for time-varying inputs.

6.2.1 The idea

The basic idea behind the test quantity is very simple and can be summarized as follows. Even though the exact functional form of $y_D(\cdot)$ is not known, the continuity properties allows for a qualitative determination of the system, like e.g. when the system is oscillating. Further, since all the attractors and bifurcations appearing in the system are assumed to be known, continuation methods combined with basic bifurcation theory, allows for determination of all qualitative behaviours consistent with the estimated \hat{p}_x . If the observed qualitative behaviour of the system is not among the possible ones, there is a contradiction. The idea is that such contradictions shall be detected with this test quantity. If, for instance, your study of the sensor y_D tells you that the system is displaying stationary oscillations at time $t = t_i$, but your study of the topology, corresponding to the estimated $(\hat{p}_x, u(t_i))$ says that there are only stable fix points, you have such a contradiction. A high-level description of the test quantity is given in Algorithm 6.1.

Algorithm 6.1. QTQ

1. Determine the observed qualitative behaviour, Q_o , of the system by usage of the given time-series Y_D and U

2. Determine the possible qualitative behaviours, Q_p , determined by the estimated parameters \hat{p}_x
3. if the observed qualitative behaviour is among the possible ones return false, else return true

The difficulty with carrying out the different steps is highly problem specific. For some systems simple checks will be sufficient, while for others it will be much more difficult. Here is only given some initial methods, that are applicable to a certain class of systems, and that can say something during periods when certain conditions on the inputs are fulfilled. What makes it difficult to use arbitrary inputs is that it is necessary that the *stationary* state of the system is determined, not just a transient phenomenon. The easiest case will be when the external inputs, u and d , to the system are constant, and when there is only one supercritical Hopf bifurcation in the system. This case is considered first. Then the case of simple time-variations on the input is considered and in Chapter 9 the a simple variation of the test quantity is suggested for a large-scale model.

6.2.2 Constant inputs

Observing the qualitative behaviour of the system

Let $q_o(t)$ be the observed qualitative behaviour at time t . Let the possible observations be: no stationarity, stationary oscillations and steady state. Further let these three states correspond to the three sets $\{\}, \{\text{osc}\}$ and $\{\text{fix}\}$, respectively. We thus seek to determine the following time series

$$\{q_o(t)\}_{t=1}^N \quad q_o(t) \in \{\{\}, \{\text{osc}\}, \{\text{fix}\}\}$$

where

$$q_o(t) = \begin{cases} \{\} & \text{if transient detected} \\ \{\text{osc}\} & \text{if stable oscillations detected} \\ \{\text{fix}\} & \text{if steady state detected} \end{cases} \quad (6.6)$$

by using the given time series Y_D .

If this test quantity is used in online processes the determination will have to be fully automatized, and then this problem can be quite challenging. In this presentation, however, we will have in mind cases similar to those of using quenching data off-line (see Section 5.4) and then we will be able to do the this step in Algorithm 6.1 by simply looking at the time-series. We will therefore, for this step, only give some of the basic ideas to consider when implementing this step.

The major difficulty when determining Q_o is to ensure that it is really the steady state behaviour of the system that has been observed, and not just a transient. This is important since bifurcation analysis, which is used to determine Q_p , only gives steady state information of the system, and since the transient and the steady state behaviour of a system are not necessarily the same. If e.g. the stationary state to a system is a spiral (i.e. a fix point with complex eigenvalues) the transient to this fix

point will be oscillating even though the steady state is non-oscillating. Further, all fix points undergoing a Hopf bifurcation are spirals close to the bifurcation. In the case of constant inputs, however, the problem is simplified. Then it will suffice to determine where the transient ends, and then do spectral analysis on these parts only.

Determining the possible qualitative behaviours of the system

Let $q_p(t)$ denote the set of all possible qualitative behaviours at time t . When there is only a supercritical Hopf bifurcation in the system q_p will be made out of two parts: one constant $q_{p,c}$ and one time-varying $q_{p,t}(t)$

$$q_p(t) = \{q_{p,c}, q_{p,t}(t)\} \quad (6.7)$$

The first part, $q_{p,c}$ makes up all constant attractors in the system, i.e. those present for all μ . The second part is determined by the relationship between $\mu(t)$ and the bifurcation manifold. If $\mu(t)$ is sufficiently far away from the manifold on the unstable side, the oscillations are assumed visible and $q_{p,t} = \{\text{osc}\}$. If $\mu(t)$ is less than a certain distance from the manifold, but still on the unstable side, both oscillations and steady-state would be consistent observations, i.e. $q_{p,t}(t) = \{\text{osc}, \text{fix}\}$. Let this critical distance be denoted δ . The need for this distance follows from the fact that a supercritical Hopf bifurcation starts with zero amplitude oscillations at the bifurcation (see Section 2.3.2). If $\mu(t)$ is on the stable side of the bifurcation manifold we assume that $q_{p,t}(t) = \{\text{fix}\}$. Here should be noted that if process noise had been included in equation (6.1), a similar uncertainty distance would have to have been included also on the stable side. This is due to the fact that the system then has precursors to the bifurcation, i.e. peaks in the power spectra might be observable already before the bifurcation has been reached Jeffries and Wiesenfeld (1985). An algorithm that does all this is given as Algorithm 6.2.

Algorithm 6.2. $Q_p 1$

Input: The estimated parameters \hat{p}_x , the input $u(t) = u$, a bifurcating fix point (x^0, μ^0) , a minimum distance from the bifurcation curve δ for the oscillations to be definitely visible, step lengths and tolerances h and acc , and the background attractors $q_{p,c}$

1. $x^* = \text{InterContinue}((x^0, \mu^0), (\hat{p}_x, u))$
2. $\lambda = \text{Eigenvalues } \nabla f(x^*, \hat{p}_x, u)$
3. *if* ($\exists i, j$ so $\lambda_i = \bar{\lambda}_j$ and $\text{Re}(\lambda_i) > 0$)
 then if $\text{Distance}(x^*, \hat{p}_x, u) < \delta$
 then $q_{p,t} = \{\text{osc}, \text{fix}\}$
 else $q_{p,t} = \{\text{osc}\}$
 else $q_{p,t} = \{\text{fix}\}$

4. return $\mathbf{q}_p = \{\mathbf{q}_{p,c}, \mathbf{q}_{p,t}\}$

5. return $(\mathbf{x}^*, \widehat{\mathbf{p}}_{\mathbf{x}}, \mathbf{u})$

Output: The qualitative behaviours, \mathbf{q}_p , consistent with the given $(\widehat{\mathbf{p}}_{\mathbf{x}}, \mathbf{u})$; the bifurcating fix point corresponding to the given $\widehat{\mathbf{p}}_{\mathbf{x}}, \mathbf{u}$

The first step in the algorithm finds the state vector that corresponds to the observed \mathbf{u} . Since there might be parallel attractors in the state space, this is done by starting at a point one knows belongs to the right attractor (\mathbf{x}^0, μ^0) and following this attractor to the correct \mathbf{x}^* . This is the function of the `InterContinue` method. The next step in the algorithm checks whether the fix point is stable or unstable. If it is stable, it returns `{fix}`, and if it is unstable it checks whether it is sufficiently far away to guarantee that there should be oscillations. Now follows a few remarks on how the algorithm can sometimes be simplified.

Remark 1: Since the system only contains a supercritical Hopf bifurcation, it can never leave the joint stable fix point/stable limit cycle attractor if it has once reached it. Therefore, if the system is known to have reached this attractor at some point during the experiment, the background behaviours, $\mathbf{q}_{p,c}$, can be set to the null set `{}`.

Remark 2: If there are no other fix points in the system other than the one undergoing the bifurcation, one can find the fix point $(\mathbf{x}^*, \widehat{\mathbf{p}}_{\mathbf{x}}, \mathbf{u})$ without knowing a starting point (\mathbf{x}^0, μ) that is a bifurcating fix point. Ordinary integration, and/or a root-finding algorithm like the Newton-Raphson method, will then be sufficient.

Remark 3: Sometimes the amplitude of the oscillations will be very low also far away from the bifurcation manifold. In those cases the distance δ will have to be set to infinity, or the actual amplitude of the oscillations will have to be determined with some other method.

If the condition in Remark 2 is not fulfilled the method `InterContinue` needs yet to be specified. It allows for a transition from the fix point that is known to bifurcate to the bifurcating fix point at the given $(\widehat{\mathbf{p}}_{\mathbf{x}}, \mathbf{u})$. An easy way to solve this method is by continuation in one direction at a time. This is shown in Algorithm 6.3 suggests.

Algorithm 6.3. `InterContinue`

Input: A fix point (\mathbf{x}^*, μ^*) that is known to be bifurcating. A parameter μ that one seeks a bifurcating fix point for, step lengths h and tolerances acc

1. $\mu_{Diff} = \mu^* - \mu$, $\mathbf{x}^i = \mathbf{x}^*$, $\mu^i = \mu^*$

2. for $i = 1, \dots, s$

 while $|\mu_{Diff,i}| > acc$

$$\begin{aligned} (x^i, \mu^i) &= \text{OneContinuationStep}(x^i, \mu^i, h) \\ \mu_{\text{Diff},i} &= \mu_{\text{Diff},i} - h \end{aligned}$$

3. return x^i

Output: The bifurcating fix point x^i , corresponding to the given μ , and that lies on the same attractor that (x^*, μ^*) does.

Similarly the Distance sub-algorithm, that estimates the distance from the bifurcation curve to the current fix point can be implemented in many ways. Algorithm 6.4 takes the minimum distance when checking the distance along the coordinate axes only. It has the benefit of being fast, especially if it is known to be sufficient to check in one or a few directions (see Section 6.3). The drawback is of course that the algorithm might give unrealistic values for some cases. An extreme case that illustrates this is given in Figure 6.1. A more sophisticated, but computationally more demanding, method would be to continue simultaneously in all directions and stop when the first direction hits the bifurcation manifold. This latter choice would be equal to a numerical approximation of the L_2 norm.

Algorithm 6.4. Distance

Input: A fix point (x^*, μ) that might undergo a Hopf bifurcation when μ is varied in some certain directions, a maximum length tMax, a step length h and desired accuracies acc.

1. $t = 0$, shouldIStop = false;
 2. for $i=1, \dots, s$
 - $x_+^j = x^*$, $x_-^j = x^*$, $\mu_{i+}^j = (\mu, \mu_i)$ $\mu_{i-}^j = (\mu, \mu_i)$
 3. while($t < \text{tMax}$)
 - for $i=1, \dots, s$
 - $(x_+^{j+1}, \mu_{i+}^{j+1}) = \text{OneContinuationStep}(x_+^j, \mu_{i+}^j, h)$
 - if $c^{\text{Hopf}}(x_+^{j+1}, \mu_{i+}^{j+1}) < \text{acc}$, then shouldIStop = true;
 - $(x_-^{j+1}, \mu_{i-}^{j+1}) = \text{OneContinuationStep}(x_-^j, \mu_{i-}^j, -h)$
 - if $c^{\text{Hopf}}(x_-^{j+1}, \mu_{i-}^{j+1}) < \text{acc}$, then shouldIStop = true;
 - $t = t + 1$;
- end for
4. return $d = t * h$;

Output: An estimated distance d, between the given fix point and the bifurcation manifold

Finally the check whether the observed qualitative behaviours, q_o is among the possible ones, q_p , for each time is a simple inclusion check. Now we will see what happens if one allows time-varying inputs to the system.

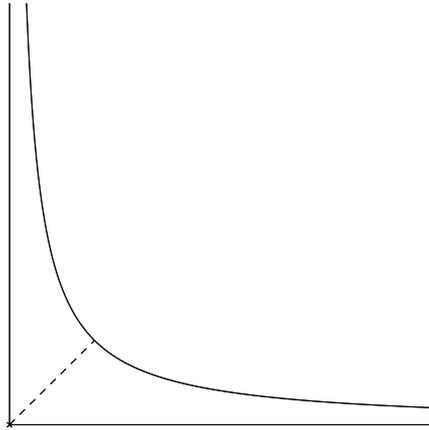


Figure 6.1 This figure illustrates the potential problem with estimating the distance from the bifurcation manifold with the simple Algorithm 6.4. In the figure the estimated point is marked with an x and the algorithm continues in all axis directions. The problem is that even though the true distance to the manifold is quite small (dashed line), the algorithm will estimate it to infinitely large. In the case of quenching experiments, however, one knows that the bifurcation can always be reached by a fairly small perturbation in a single control variable, and then this problem will not appear.

6.2.3 Time-varying inputs

When the input signals u are allowed to vary over time the algorithms above has to be altered in different ways. Time-varying inputs also opens up for the possibility of changes between different qualitative behaviours at different time-points, i.e. it opens up for the passing of bifurcations, which also complicates the algorithms. As in the previous section we will first study the problem of determining the observed behaviours and then go to the problem of determining the possible behaviours. Here, however, we will have a more complex comparison problem, and therefore we will treat this separately in the last part of this section.

Observing the qualitative behaviours of the system

Also now we will mainly assume that this step can be done in some simple way, like e.g. when utilizing quenching data off-line. We still seek to determine $q_p(t)$ which can have one of the three values $\{\}$, $\{\text{osc}\}$ and $\{\text{fix}\}$, and they will still denote the three observations: transient behaviour, stationary oscillations, and steady state. The only thing that will change is that we demand that $u(t)$ is constant for any of the values $\{\text{osc}\}$ and $\{\text{fix}\}$ to be possible. However, it might happen that for

two different constant \mathbf{u} windows, two different stationary states will be observed. If this is the case we will have detected an observed bifurcation. Let the vector of this detected passings be denoted `detectedPassings`. It should be noted that the adding of the presence of a bifurcation could also have been included in the constant \mathbf{u} problem. Then, however, the qualitative change would have to have been observed before the collection of the time-series. How this adding is included into the process is described in Algorithm 6.6.

Determining the possible qualitative behaviours of the system

We still let $\mathbf{q}_p(\mathbf{t})$ denote the set of all possible qualitative behaviours at time \mathbf{t} . Now, however, all predicted passings of bifurcation manifold are detected as well. They are collected in the array `predictedPassings`. The new algorithm is as follows.

Algorithm 6.5. Q_p2

Input: The estimated parameters $\hat{\mathbf{p}}_x$, the time-series \mathbf{U} , a bifurcation fix point (\mathbf{x}^0, μ^0) , a minimum distance δ from the bifurcation manifold for the oscillations to be definitely visible, step lengths and tolerances \mathbf{h} and \mathbf{acc} , the background attractors $\mathbf{q}_{p,c}$.

for $i = 1, \dots, N$

1 First the old check as if \mathbf{u} had been constant

$$(\mathbf{q}_p(i), \mathbf{x}^*(i)) = Q_p1(\hat{\mathbf{p}}_x, \mathbf{u}(i), (\mathbf{x}^*(i-1), \mu(i-1)), \delta)$$

2 Then check whether the bifurcation manifold has been passed

$$\text{if } (\text{osc} \in \mathbf{q}_{p,t}(i) \wedge \text{osc} \notin \mathbf{q}_{p,t}(i-1)) \vee \\ (\text{osc} \in \mathbf{q}_{p,t}(i-1) \wedge \text{osc} \notin \mathbf{q}_{p,t}(i))$$

then `predictedPassings.append(i)`

end if

end for

• return $\{\mathbf{q}_p(\mathbf{t})\}_{\mathbf{t}=1}^N, \text{predictedPassings}$

Output: The predicted qualitative behaviours, the predicted passings of the bifurcation

Comparing the predicted and observed behaviours

Since now both the predicted and the observed behaviours contain an array describing the passings of the bifurcation the last step in the overall algorithm, the comparison step, will become more complicated than a simple inclusion check. Since one might also know that there should be bifurcations for certain \mathbf{u} s (like

in the examples in Section 4.5) one might sometimes also want to add such a knowledge to the test quantity (or to the verification step). Let the indices in the time series where the observed and the predicted passings of bifurcation points be denoted `observedPassings` and `predictedPassings`, respectively. Further, let the uncertainty, measured in u , of how accurate one can determine the bifurcation point be denoted u_{unc} . Finally let the known point where the system is undergoing a Hopf bifurcation be denoted u_b , and the corresponding state vectors be denoted x_b (cf Chapter 4). With these notations we present Algorithm 6.6 for doing the last step in the high-level Algorithm 6.1.

Algorithm 6.6. Comparison

Input: A predicted \widehat{p}_x , observed qualitative behaviours $\{q_o(t)\}_{t=1}^N$, predicted qualitative behaviours $\{q_p(t)\}_{t=1}^N$, observed passings of the bifurcation manifold `observedPassings`, predicted passings of the manifold `predictedPassings`, known u_{bif} for which the system should undergo a bifurcation, estimated uncertainties in the bifurcation determinations u_{unc} .

1. First the simple inclusion check

```

for  $i = 1, \dots, N$ 
  if  $q_o(t) \in q_p(t)$ 
    then  $q(t) = \text{true}$ 
    else  $q(t) = \text{false}$ 
  end if
end for

```

2. Then the passings of the bifurcation points

```

for  $i = 1, \dots, \text{observedPassings.length}$ 
  if  $\|u(\text{observedPassings}(i)) - u(\text{predictedPassings}(i))\| > \|u_{\text{unc}}\|$ 
    then  $q(\text{observedPassings}(i)) = \text{true}$ 
  end if

```

3. Then the a priori known bifurcation points

```

 $x_{\text{bif}} = \text{InterContinuate}((x^0, \mu^0), \widehat{p}_x, u_{\text{bif}})$ 
 $\lambda = \text{Eigenvalues} \nabla f(x_{\text{bif}}, \widehat{p}_x, u_{\text{bif}})$ 
if  $(\exists i, j \text{ so } \lambda_i = \bar{\lambda}_j \wedge \text{Re}(\lambda_i) > 0)$ 
  then  $\text{isUnstable} = \text{true}$ 
  else  $\text{isUnstable} = \text{false}$ 
end if
for  $i = 1, \dots, u_{\text{unc.length}}$ 
   $x_{\text{bif}, i+} = \text{InterContinuate}((x^0, \mu^0), \widehat{p}_x, u_{\text{bif}} + u_{\text{unc}, i})$ 

```

```

     $x_{\text{bif},i-} = \text{InterContinuate}((x^0, \mu^0), \hat{p}_x, u_{\text{bif}} - u_{\text{unc},i})$ 
     $\lambda_{i+} = \text{Eigenvalues} \nabla f(x_{\text{bif},i+}, \hat{p}_x, u_{\text{bif}} + u_{\text{unc},i})$ 
     $\lambda_{i-} = \text{Eigenvalues} \nabla f(x_{\text{bif},i-}, \hat{p}_x, u_{\text{bif}} - u_{\text{unc},i})$ 
    if  $(\exists j, k \text{ so } \lambda_{i+,j} = \bar{\lambda}_{i+,k} \wedge \text{Re}(\lambda_{i+,j}) > 0)$ 
        then  $\text{isUnstable+} = \text{true}$ 
        else  $\text{isUnstable+} = \text{false}$ 
    end if
    if  $(\exists j, k \text{ so } \lambda_{i-,j} = \bar{\lambda}_{i-,k} \wedge \text{Re}(\lambda_{i-,j}) > 0)$ 
        then  $\text{isUnstable-} = \text{true}$ 
        else  $\text{isUnstable-} = \text{false}$ 
    end if
    if  $\text{isUnstable+} \neq \text{isUnstable} \vee \text{isUnstable-} \neq \text{isUnstable}$ 
        then  $\text{isDetected} = \text{true}$ 
    end if
    end for
    if  $\text{isDetected} \neq \text{true}$ 
        then  $q(t) = \text{true} \forall t$ 
    end if
4. return QTQ :  $t \rightarrow q(t)$ 

```

Output: the value of the qualitative test quantity for all values of t . Here $q(t_i) = \text{true}$ means that the test quantity has reacted at time t_i .

The inclusion check in Algorithm 6.6 is straightforward. The reason the predicted passings and observed passings are compared twice is that they might have different lengths, and that they therefore will give different agreements in depending on whose index should be gone though. The last part of the algorithm is simply a continuation in all different directions that there is an uncertainty u_{unc} in, to see whether there is a bifurcation within the region of uncertainty. Similarly to the Distance algorithm there are better alternatives for a general problem, and the norm problem described in Figure 6.1 will be present also here. However, in the case of quenching data, a single direction will be sufficient to check, and then this feature is not a problem.

6.3 Validation of the Brusselator

6.3.1 The two problems

Let us now again come back to the analysis of the two time-series presented in Section 4.5. The first of these samples was collected just before the bifurcation and the second just after. For both the samples a parameter estimation was done.

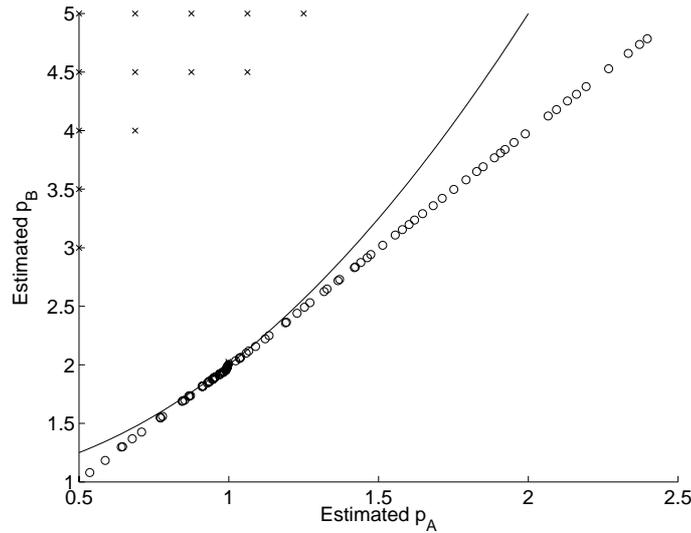


Figure 6.2 *The estimated values for Sample 2 in Section 6.3. Those point denoted with an x corresponds to an oscillating stationary state, and those denoted with an o to a steady state. Here the real system was oscillating and therefore all the o's are erroneous estimates that will be captured by the new test quantity. In this case it captured over 60% of unsuccessful estimations by this qualitative check.*

The two problems was solved with two different methods, one that utilized the knowledge of the presence of the bifurcation and one that did not. The former method converged to the correct value for all initial estimates, while the latter method only converged to the correct value for those initial estimates starting sufficiently close to the correct value. Let us now see how many of the erroneous parameter estimates that will be detected with the methods of Section 6.2.3.

6.3.2 Applying the QTQ algorithm

Figure 4.4 shows the points to which the unconstrained method converged, when analysing Sample 2, the one collected on the unstable side of the bifurcation manifold. Let us now re-evaluate these estimates from the perspective of a qualitative analysis.

Step 1: Determining the observed qualitative behaviour

The corresponding y_D time-series for Sample 2 was plotted in Figure 4.2. Since this time-series has already been collected we can solve the first step in the QTQ

algorithm by mere inspection. As we see in the plot, Sample 2 contains stationary oscillations during the whole series, i.e. no transients and always clearly visible oscillations. We therefore set $\mathbf{q}_o(t) = \{\text{osc}\}$ for all t .

Step 2: Determining the possible qualitative behaviours

Since there is a constant input to this problem we can drop the InterContinue algorithm and set $\mathbf{q}_{p,c} = \{\}$. By some initial simulations we estimate the distance from the bifurcation manifold for the oscillations to be visible to be $\delta = 0.005$. With these settings the estimated values can be divided in three groups, corresponding to the three possible values of $\mathbf{q}_{p,t}$. Such a division is shown in Figure 6.2. Approximately 65% of the estimated parameters correspond to a stationary state. These erroneous estimates are captured by the QTQ algorithm. It should be added that many of the estimates that lie in the steady state regime (below the bifurcation line in Figure 2.4), were initiated in the unsteady regime (above the bifurcation line). We conclude that an estimation algorithm does not automatically guarantee that the qualitative behaviour of a system is consistent with the observed qualitative behaviour, even when the observation and estimation is based on the same time-series, and even when the initial estimate is in a qualitatively consistent regime. This qualitative check must therefore be added explicitly, either during the validation phase, as is suggested in this chapter, or directly during the estimation phase.

Step 3: Comparing the results of the two steps

The qualitative comparison has already been discussed above. Let us therefore turn to the other part of the comparison, the one that checks whether the bifurcations are where they should. For this sample, there has been no detected bifurcations in the time-series, but the system has been tuned to lie just on the unstable side of the bifurcation (see Section 4.5). A bifurcation has therefore been detected prior to the collection of the time-series, and the estimated value is $(\hat{p}_A, \hat{p}_B - 0.01)$. The estimated uncertainty is $u_{unc} = 0.005$ (only p_B is considered as controlled). With these values put into Algorithm 6.6 89% of the estimates will be rejected. Note that we have to use Algorithm 6.6 even though there are constant inputs. This is due to the fact that we have detected bifurcations prior to the collected time series. As can be seen in Figure 6.3 the remaining estimates are very centralized, and all lie close to correct parameter value, $(p_A, p_B) = (1, 2.01)$. It is concluded that the presented QTQ algorithm has the potential to reject a substantial part of the erroneous estimates. It is therefore an alternative to changing the estimation algorithm itself, as suggested in Chapter 4 and 5. The drawback is that with this method one has to do more optimizations.

6.4 Conclusions

In this chapter has been presented a way of validating the result of an estimation process. The method studies the dynamics of one of the sensors to see if and when the system is in an oscillatory or in a steady stationary state. This can be done

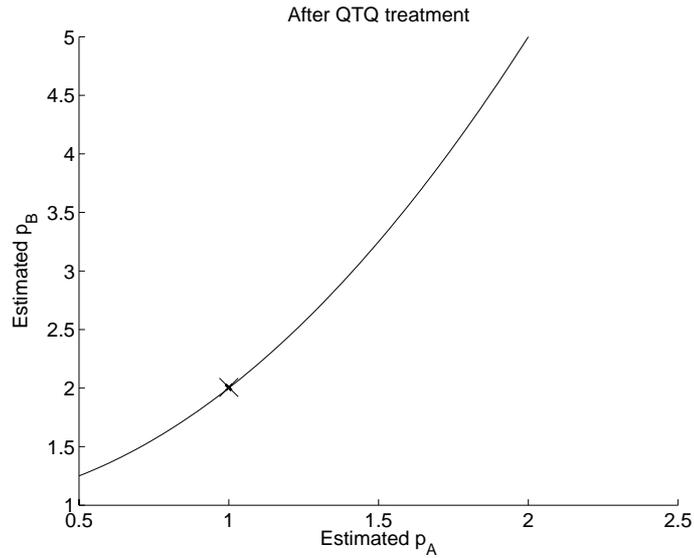


Figure 6.3 *The remaining estimates after the ones captured by the QTQ algorithm has been removed. The estimates are marked with an \times . As can be seen the remaining estimates are centralized around the correct estimate $(p_A, p_B) = (1, 2.01)$*

also in the case of uncalibrated sensors, i.e. with p_y or $y(\cdot)$ is unknown. Then continuation methods combined with bifurcation theory and supposed knowledge of the topology of the system is used to determine whether it is possible that the system has this stationary behaviour at these instances. It is also checked whether the detected passings, or positions, of a Hopf bifurcation are consistent with the estimated parameters. If this should not be the case the found estimation is rejected. The test quantity will be sensitive to faults due to the convergence to a local minima (see equation (6.4)), faults appearing in the underlying ODE, and faults in the used sensors.

A Multi-zonal Model for Pressure, Temperature and Ionization Currents

This chapter is the first that deals with system identification in practice, i.e. with the development of specific models for specific applications. The model presented is a multi-zonal model for cylinder pressure, temperature, and ionization currents. The model is designed for a spark-ignited combustion engine and is covering the time between the Inlet Valve Closure (IVC) and the Exhaust Valve Open (EVO). The analysis of the model has focused on the relationship between the pressure peak location (PPL) and the ionization peak location (IPL). Previous models have systematically predicted the $IPL > 2$ CAD late, and the reason for this is analysed. The final model has an experimental agreement for the prediction of IPL within one crank angle degree (CAD), which is comparable to measurement uncertainty. The final model, and the analysis during the development are the main contributions of the chapter.

The material is structured as follows. First the basic concepts of an Otto-cycle are reviewed in Section 7.1. Then the research trace that this chapter is following is zoomed in by first explaining the importance of knowing the position of the pressure curve relative to top dead center (TDC), and then by giving a short review of previous attempts to predict this by means of measuring the ionization current. The new model is presented in Section 7.2, and in Section 7.3, the model is analyzed in a number of ways. Finally, in Section 7.4, the results of the chapter are summarized.

7.1 Introduction

This section starts by giving the basic concepts of an Otto engine. Then is explained why it is important to know the position of the pressure curve relative to the TDC. A cheap way of finding this position is by using the spark plug as an ionization sensor. This is cheap since the sensor is already present in production engines for knock detection and misfire, and it therefore only requires further signal interpretation. The section ends with a short review of previous modelling work in the area, and with a high-lighting of the observed problem that the rest of the chapter is attacking.

7.1.1 Basic engine concepts

The Otto-cycle

The four-stroke Otto engine is a machine that converts chemical energy in the fuel into mechanical energy and heat. As the name implies it operates in four strokes, see Figure 7.1. In the first stroke (a) the piston is moving downwards as the intake valve is open. The air-fuel mixture flows into the cylinder due to the pressure difference in the intake manifold and the cylinder, and this difference is kept by the movement of the piston. When the piston reaches a point around the Bottom Dead Center (BDC) the intake valve closes (this point is denoted IVC), and as the piston moves up (b) it compresses the air-fuel mixture. When the piston reaches a certain point, before the Top Dead Center (TDC), i.e. its upper turning point, an electric circuit creates a spark in the gap between the two electrodes in the spark plug (see Figure 7.7). The point is called the ignition angle, and starts the combustion phase. During the combustion phase the piston passes TDC and goes into the expansion phase when the piston moves downward (c). It is during this phase that the internal energy of the uncombusted gas is converted to mechanical energy. Approximately when the piston reaches its lower turning point the exhaust valve opens (the exact point is denoted EVO) and as the piston moves up (d) the combusted gas is pushed out in the exhaust pipe. After this stroke the Otto-cycle goes back to (a), which completes the four-stroke cycle. The model that will be presented in this chapter will only cover the compression, combustion and expansion phases, i.e. the part between the IVC and the EVO.

When modelling an Otto-cycle, usually the equations are written as a function of the position of the crank. The most common notation is the angle between the crank and the axis of the cylinder, see Figure 7.2. This angle will be referred to as the Crank Angle Degree (CAD), and our convention will be to have it negative before TDC, zero at TDC, and positive after TDC. In Figure 7.3 we see one of the most common ways of displaying the cylinder processes, a plot of the cylinder pressure against the time measured in CAD. We will now go on by explaining why this is such a relevant plot and why it is important to know the position of this curve relative to TDC. The reader who wants more background material regarding general engine concepts is referred to e.g. one of the textbooks Heywood (1988) or Nielsen and Eriksson (2004).

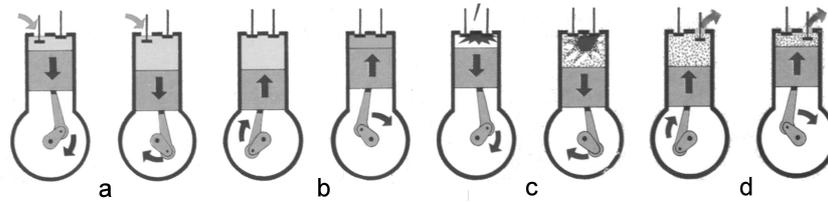


Figure 7.1 *The four-stroke Otto cycle. In the first stroke (a) the air-fuel mixture is pressed into the cylinder due to the relatively high pressure in the intake manifold. After the intake stroke the compression stroke (b) begins. At some point before Top Dead Center the ignition angle is passed, and during the expansion phase (c) the combustion processes are converting the chemical energy in the fuel into mechanical energy and heat. Finally during the exhaust phase (d) the combusted mixture is exhausted.*

7.1.2 The importance of predicting the Pressure Peak Location

The ignition timing problem

In Spark Ignited (SI) engines ignition timing is an important control signal for the combustion efficiency. The ignition timing alone affects almost every engine output. In nearly all of today's production engines there is no feedback from the combustion to the ignition timing; the spark advance is based on a pre-calibrated system. For such a pre-calibrated system to work well it has to take into account the many parameters that effect the best timing. Such parameters include e.g. engine speed, load, air/fuel ratio, fuel characteristics, EGR, coolant temperature, air temperature and humidity. Taking this into account implies two problems: i) not all these parameters are measured in today's production engines ii) even if all these parameters were measured a calibration scheme that determines the optimal ignition timing with respect to all these parameters would be too expensive to obtain through engine tests. Hence, compromises leading to less than optimal control, have to be done. Another reason for compromises is that the calibration scheme has to ensure that engine knock does not occur, and that reasonable performance over the whole operating range is obtained. A fundamentally different approach to using a pre-calibrated scheme is to measure the cylinder pressure and use it for feedback in the control algorithm.

Using the pressure curve

The optimal ignition timing will position the pressure curve in a certain way with respect to TDC. If the peak of this curve (the pressure peak location (PPL)) comes

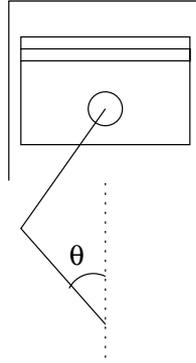


Figure 7.2 Figure showing how the angle θ is defined in this thesis. The indicated θ is negative and located approximately where the ignition takes place (10-30 Crank Angle Degrees (CAD) Before Top Dead Center (BTDC)).

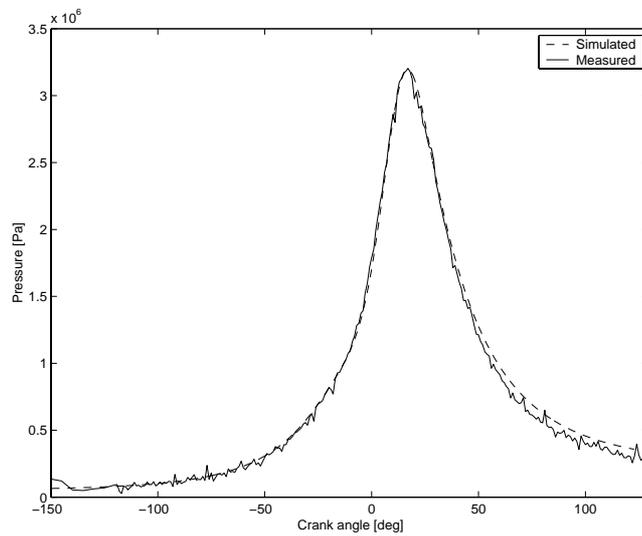


Figure 7.3 A simulated pressure plotted with against measured pressure, both as a function of Crank Angle Degree (CAD). Ignition angle: 27 CAD BTDC. See Appendix A for other parameters.

too early, too much work will be lost during the compression phase. Also if the peak comes too late the extracted work from the engine will be less than optimal. Usually the optimal position is measured as the one giving the highest torque

(Maximum Break Torque (MBT)), and depending on e.g. the engine design the PPL corresponding to MBT varies between 12 and 20 CAD after TDC (Hubbard, Dobson and Powell, 1976). This is supposed to be due to the varying heat flow to the walls for the different engine designs. However, with respect to the time-varying parameters that affect one and the same engine it has been shown that PPL is almost constant at MBT (Hubbard et al., 1976). A spark control algorithm that maintains a constant position for the pressure peak (called the peak pressure algorithm), results in an ignition timing that is within 2 CAD of optimum. The algorithm also gives optimal timing for large changes in parameters that influence the flame speed, such as the air/fuel ratio and the air humidity ((Hubbard et al., 1976);(Glaser and Powell, 1981)). One manufacturer has implemented the peak pressure concept in combination with knock control, and reported a 10% improvement in power and a 5% improvement in efficiency (Sawamoto, Kawamura, Kita and Matshushita, 1987).

One major problem stopping more manufacturers from using such an algorithm today is that a pressure sensor is currently not installed, and installing such a sensor would be both expensive and cumbersome, due to the short life-time of these sensors. However, for estimating PPL also other sensors can be used. Such a possibility lies in using the spark plug as an ionization sensor and an overview of the current status with this approach is now given.

7.1.3 Using the spark plug as an ionization sensor

Due to the high temperature and pressure in the cylinder during the combustion phase, free ions are formed. These ions are measured by applying a voltage across the two electrodes of the spark plug, and by measuring the resulting current. In Figure 7.4 the result of such a measurement is shown. As can be seen the obtained signal is highly complex, and it has been shown to contain much useful information of various features of the combustion processes. The signal is usually divided into several phases. The first phase occurs during the ignition, the second during the early development of the flame, and the last includes the late development of the flame. These phases are denoted ignition phase, flame-front phase, and the post-flame phase, respectively. Two of the most common usages of this signal are knock and misfire detection (Auzins, Johansson and Nytomt, 1995). However, the air-fuel ratio and many other physical properties of the combustion components are also visible in the signal (Reinmann, Saitzkoff and Mauss, 1997). Since there is much (and still much unused) information in this signal, and since the signal is complex, it is important to understand the physical phenomena that govern the generation of this current. Hence physical models are sought. Over the years a number of such models have been developed and usually they have focused on different properties of the signal. Some examples of modelling attempts are the Yoshiyama-Tomita model (Yoshiyama, Tomita and Hamamoto, 2000), that bases the model on bomb experiments; the Calcote model (Calcote, 1957), that models the spark plug as a Langmuir probe; and different attempts utilizing the Saha equation (Saitzkoff, Reinmann, Berglind and Glavmo, 1996), (Franke, 2002), (Andersson, 2002). In

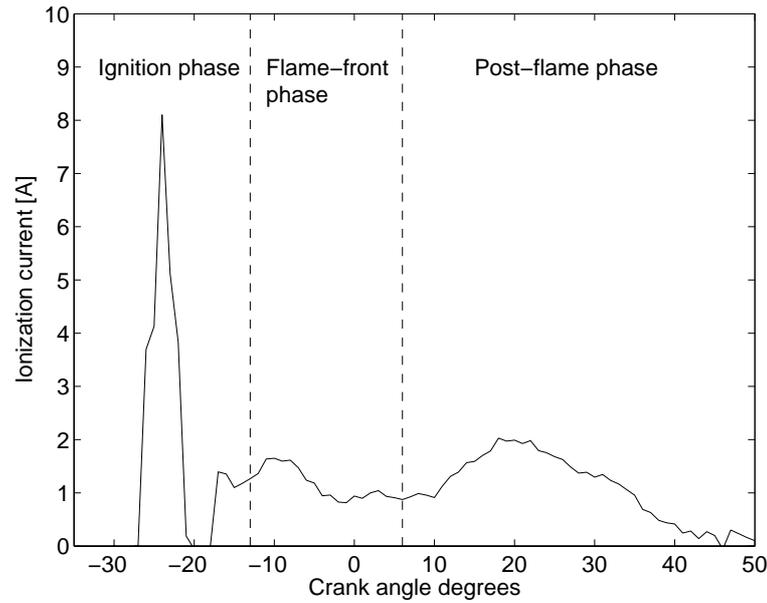


Figure 7.4 Example of a measured ionization current with its three characteristic phases.

this chapter it is the latter research trace that has been followed. One reason for this is that this family of models are all mainly concerned with understanding the last part of the ionization curve, the post-flame part. It is namely this part that is believed to have the highest correlation with the pressure development. Now follows a review of previous works using this part of the curve to predict the PPL.

Previous modelling of the relation between PPL and IPL

Analogously with the acronym PPL there is also an acronym for Ionization Peak Location, IPL, and when using it we will always mean the peak in the post-flame part of the curve. It is not always that this peak is a local maximum value of the signal. This happens if the strength of this peak is lower than the down-slope from the flame-front phase, and this happens for some parts of the operating range. This in combination with the poor signal-to-noise ratio in the ionization signal leads to an inherent uncertainty already when determining IPL. An early attempt at solving this problem was done in Eriksson, Nielsen and Nytomt (1996) and Eriksson, Nielsen and Glavenius (1997), where also the usage of the prediction of PPL in the peak pressure algorithm was exploited. The idea was based on the resemblance between the simulated ionization curve from the physically based model presented in Saitzkoff et al. (1996), and a Gauss function. When analyzing

the ionization signal it was expressed as a signal in two parts

$$I(\theta) = f(\theta) + \alpha_1 \exp\left(-\frac{1}{\alpha_2}(\theta - \alpha_3)^2\right) = \beta_1 \exp\left(-\frac{1}{\beta_2}(\theta - \beta_3)^2\right) + \alpha_1 \exp\left(-\frac{1}{\alpha_2}(\theta - \alpha_3)^2\right) \quad (7.1)$$

where the clock modelled with α 's is the part describing the post-flame part, and $f(\theta)$, which is also modelled as one or two clock functions, models the flame-front contribution. The symbols α_i and β_i are model parameters. As can be seen in Figure 10 in Eriksson et al. (1997) and in Figure 13 in Eriksson et al. (1996) the agreement between the second peak, when fitted to data, and the measured pressure peak, for the same data is good, i.e. centered around the $x = y$ line, except for the very late ignition angles. This good agreement shows that there is high correlation between the post-flame peak and the PPL. This method is based on a number of assumptions that implies that the predicted IPL and the predicted PPL coincides.

Further analysis of the assumptions underlying the original Saitzkoff model has been done in Franke (2002). This analysis includes an analysis of alternative ion sources to NO in the post-flame phase. One of the conclusions has been that alkali metals in the atmosphere make a major contribution to thermal ionization, also at temperatures characteristic of the combustion of diluted mixtures. In Franke (2002) a further analysis of the relationship between the pressure curve and the post-flame part of the ionization signal has also been given. This has been done e.g. by imaging nitric oxide and hydroxyl radicals. The effect of gas flow on this relationship has been studied, and it has been shown that gas flow does effect the relationship between the pressure and ionization curves in some situations. There was shown experimental recordings where the PPL and IPL do not coincide, but where a model equating the IPL by the PPL would predict the IPL a few degrees late (see Figure 4.28 in Franke (2002)). Finally the following linear relationship between the PPL and IPL was suggested

$$\text{IPL} = a + b\text{PPL} + c\bar{u} \quad (7.2)$$

where, a , b and c , are constant parameters. The last term models the effect of gas flow, where \bar{u} is the mean velocity of the gas flow. It was, however, noticed that in a number of situations the last term, $c\bar{u}$, made no significant contribution.

Another recent study of the relationship between PPL and IPL was done in Andersson (2002). There the original temperature and pressure part of the Saitzkoff model was replaced by a number of other one- and two-zone models, however, all of them without including heat transfer. Also here a late prediction of the IPL was observed (see e.g. Figure 6.5 in Andersson (2002)). It was also shown that the choice of underlying heat model can change the amplitude of the simulated ionization current by several orders of magnitude, apart from moving the predicted IPL several CAD. Another recent work is Andersson (2004), and although it differs from Andersson (2002) in some aspects, it also emphasizes that the underlying temperature model has a significant effect on the simulated ionization signal.

In this chapter we will continue this study of the relationship between PPL and IPL. Based on the previous observation that the choice of underlying temperature model is important, a multi-zonal model has been chosen as the underlying

temperature model for a model, that is otherwise like the original Saitzkoff model (Saitzkoff et al., 1996). The presented multi-zonal model is more detailed than the models tested in the previous works by Andersson, and it will also have the possibility of including heat transfer, something that is known to have a significant effect on the temperature development. A zonal model was chosen above a partial differential equation since it was believed that such a model structure would lead too far away from control oriented modelling. The model is presented in Section 7.2 and the analysis and experimental comparison is done in Section 7.3.

7.2 The model

7.2.1 The multi-zonal combustion model

Since the ionization model builds on a combustion model, the latter is presented first and the former in the next section. The theory is from Nilsson and Eriksson (2001). This paper presents a general model describing the thermodynamic basis of a thermo-dynamical model. It, however, includes no suggestions of how to determine the heat transfer to the walls, the mass transfer between the zones, and also no details of how to interpret the different zones geometrically. These details will be added in the subsequent subsections.

The thermodynamic part of the model is valid under the following assumptions

- The combusted zone has such a high temperature that all reactions can be considered to be in equilibrium, with respect to the time-scale of the temperature and pressure developments.
- The uncombusted zone is so cold that all reactions can be considered frozen, with respect to the time-scale of the temperature and pressure developments.
- Each zone is a well-stirred tank reactor
- Let V be the volume of the i :th zone, p the pressure, m the mass, R the gas constant, and T the temperature. Let further the index i , like in m_i , indicate that the mass belongs to the i :th zone. Then the ideal gas law in its differentiated form, (7.3), is assumed valid for each zone

$$V_i dp + p dV_i = R_i T_i dm_i + m_i T_i dR_i + m_i R_i dT_i \quad (7.3)$$

- Let dU be the change in internal energy, dQ the heat change, and dW the work. Then energy conservation (7.4) is valid for each zone

$$dU = dW + dQ \quad (7.4)$$

- The pressure can be treated as a function of time only, i.e. modelled as the same in the in all zones.

These assumptions leads to the following system of nonlinear DAEs

$$\mathbf{A} \mathbf{d}\mathbf{x} = \mathbf{B} \quad (7.5)$$

where

$$\mathbf{d}\mathbf{x} = [\mathbf{d}p \ \mathbf{d}V_1 \ \mathbf{d}T_1 \ \dots \ \mathbf{d}V_N \ \mathbf{d}T_N]^T \quad (7.6)$$

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & \dots & 1 & 0 \\ \mathbf{a}_1 & p & \mathbf{b}_1 & \dots & 0 & 0 \\ \mathbf{c}_1 & p & \mathbf{d}_1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{a}_N & 0 & 0 & \dots & p & \mathbf{b}_N \\ \mathbf{c}_N & 0 & 0 & \dots & p & \mathbf{d}_N \end{pmatrix}$$

$$\mathbf{B} = \begin{pmatrix} \mathbf{d}V \\ \delta Q_1 + \sum_{i \neq 1} (R_i T_i - h_{e,1} + R_1 T_1) \mathbf{d}m_{1i} \\ \vdots \\ \delta Q_N + \sum_{i \neq N} (R_i T_i - h_{e,N} + R_N T_N) \mathbf{d}m_{Ni} \end{pmatrix}$$

and where the auxiliary variables \mathbf{a}_i , \mathbf{b}_i , \mathbf{c}_i and \mathbf{d}_i are defined as follows

$$\mathbf{a}_i = V_i \left(1 - \frac{p}{R_i} \left(\frac{\partial R_i}{\partial p} \right)_{T_i} \right) \quad (7.7)$$

$$\mathbf{b}_i = -m_i \left(R_i + T_i \left(\frac{\partial R_i}{\partial T_i} \right)_p \right)$$

$$\mathbf{c}_i = -m_i T_i \left(\frac{T_i}{p} \left(\frac{\partial R_i}{\partial T_i} \right)_p + \left(\frac{\partial R_i}{\partial p} \right)_{T_i} \right) \quad (7.8)$$

$$\mathbf{d}_i = m_i \left(c_p - R_i - T_i \left(\frac{\partial R_i}{\partial T_i} \right)_p \right)$$

See Nilsson and Eriksson (2001) for more details concerning the derivations. In the matrices appear a few variables whose dependence have not been written out. The $\mathbf{d}V$ variable appearing in the \mathbf{B} matrix is a mere function of the time, and the geometry of the engine. The expression for \mathbf{V} can be found in e.g. Nielsen and Eriksson (2004). The variables R , h_e (enthalpy), c_p (specific heat), and the partial derivatives $(\frac{\mathbf{d}R}{\mathbf{d}T})_p$ and $(\frac{\mathbf{d}R}{\mathbf{d}p})_T$ are all determined by the ordinary state-variables p , T and V . In practice they are calculated by look-up tables generated by the program CHEPP (Eriksson, 2004). The last two unknowns, $\mathbf{d}Q$ and $\mathbf{d}m$, are treated in Section 7.2.2 and 7.2.3, respectively.

In practice when solving the system \mathbf{A} is inverted, after a suitable scaling given in Nilsson and Eriksson (2001), and then the system is solved with an ordinary ODE solver. The system is, however, non-standard also in such a way that it does not have the same number of zones (i.e states) during the whole simulation.

First there is only one zone, an uncombusted zone. At the ignition a new zone, a combusted zone, is created. This zone then grows as more and more uncombusted air-fuel mixture is burned, and this is formulated as a mass-flow from the unburned zone to the burned zone (see Section 7.2.3). After the mass of the burned zone has reached a certain limit, the zone is divided into two. This is done by creating a very thin zone outside the first zone, which now has all the mass-flow from the unburned zone. The mass-flow will cause this zone to grow, and divide, in exactly the same way as the first zone did. In Figure 7.5 the ordering of the zones is displayed. See Appendix A.5 for more details on the simulations, and also for more details of how and when new zone is created. Now we will turn to the problem of specifying dQ and then to specifying dm .

7.2.2 Addition of geometry and heat transfer

First we will give the general equation that is used to calculate the heat transfer, the Newtons law of cooling. Then we will specify how the three parts, h_w , A and ΔT , determining the cooling equation are calculated. This will lead us to the introduction of a geometrical interpretation of the otherwise only topologically coupled zones.

Newtons law of cooling

The temperature difference, ΔT , between the in-cylinder gas and the cylinder walls, is the source of the energy flow. The base for the heat transfer modelling is Newton's law of cooling, which says that the heat transmitted through the gas-cylinder contact area A per unit time is given by the following expression

$$\frac{dQ}{dt} = h_w A \Delta T \quad (7.9)$$

First we will present the determination of h and then we will give the three methods used to determine A . The last of this will include the introduction of a geometrical interpretation. Finally the ΔT determination is given.

Heat transfer coefficient

There are different ways of determining the coefficient h_w in (7.9). We have chosen the expression presented in Woschni (1967), but with the coefficient C_1 and C_2 adapted to SI engines. The model is also converted into SI units.

$$h_w = \frac{253B^{-0.2}C_1p^{0.8} \left(\frac{0.0034(p-p_m)T_{IVC}V_{disp}C_2}{p_{IVC}V_{IVC}} + U_p \right)^{0.8}}{T^{0.53}} \quad (7.10)$$

Here p_f is the pressure in a firing cycle and p_m is the pressure in a motored cycle, a cycle without combustion. C_1 and C_2 are motor-type dependent constants. U_p is the mean velocity of the piston. The values and units used are given in Table 7.1.

B	9.0e-2 [m]
C ₁	2 []
p	[Pa]
p _m	[Pa]
T _{IVC}	363 [K]
V _{disp}	4.94e-4 [m ³]
C ₂	0.44 []
p _{IVC}	6.5e4 [Pa]
V _{IVC}	[m ³]
U _p	4.97 [m/s]
T	[K]

Table 7.1 Parameter list for the heat transfer model in equation (7.10).

Various area models

Let us now turn to the problem of determining A in equation (7.9). Three different methods for this are presented. The first two methods has the advantage of the overall cylinder pressure not being dependent on the number of zones, and the third has the advantage of being the most realistic. The first and most simplest method tested is to consider A_i to be directly dependent of the zonal mass

$$A_i = \frac{m_i}{m_{\text{tot}}} A_{\text{tot}} \quad (7.11)$$

Here m_i is the zones own mass, and m_{tot} is the total mass of the gas in the cylinder. A_i is the contact area between the i :th zone and the cylinder walls, and A_{tot} is the total cylinder wall area exposed to the fuel-gas mixture. The second and more advanced method tested scales the area with the zones volume. The method is based on the observation that the area of a sphere is proportional to its volume according to $A \propto V^{\frac{2}{3}}$. Since the most simple interpretation of the zones are that they are spherically shaped, a method to calculate the contact area from this relation is implemented. The equation for the i :th zone is:

$$A_i = \frac{V_i^{\frac{2}{3}}}{\sum_i V_i^{\frac{2}{3}}} A_{\text{tot}} \quad (7.12)$$

Where V_i is the i :th zones volume. The most advanced method for calculating A_i , used to give the zones their real (in the most simple case, spherical) form, i.e. by start introducing a real geometry in the model. This has been done below.

Translating topology to geometry

In the geometry that has been included in this model the burned zones are assumed to be spherical shells with center on the cylinders symmetry axis. The first zone, that always exists, is the unburned zone. When the combustion starts, the first

burned zone is created. The center is between the two electrodes in the spark plug, which is assumed to be positioned on the symmetry axis of the cylinder. As the piston moves the center is also assumed to move, to approximate the compression and the expansion of the gas. The movement of the center is modelled in such a way that the ratio dh/dp is held constant. Here dh is the distance between the cylinder-head and the center, and dp the distance between the piston and the center (see Figure 7.5). The inner and the outer radius of each zone (see Figure 7.5), are calculated recursively for each time-point in the following manner.

1. Integrate the differential equations (7.5) to obtain the volume of each zone.
2. The first burned zone (labelled with 2 in Figure 7.5) is simply a sphere and hence there is only one radius, r_1 , to determine, and this is given uniquely by the volume and the equation for the sphere.
3. The calculated r_1 is also the inner radius for the second combusted zone. The outer radius of the second combusted zone, r_2 , is therefore uniquely determined by the volume of the zone, the radius r_1 and the equation for a truncated spherical shells.
4. The calculated r_2 is also the inner radius for the third combusted zone. The outer radius of the third combusted zone, r_3 , is therefore uniquely determined by the volume of the zone, the radius r_2 and the equation for a truncated spherical shell.
5. The calculated r_3 is also the inner radius for the fourth zone, etc.

When the zones are truncated, and in what way, is determined by the specific geometry of the cylinder, and the time-development of the combustion. The formulas giving the relationship between the radii and the volumes of each zone is just the difference between the corresponding expressions for simple geometrical shapes, spheres and cylinders, and these can be found in e.g. Råde and Westergren (1990).

Determining ΔT

With A determined by any of three techniques mentioned above, it is easy to determine ΔT , which is the only remaining part in equation (7.9). It is simply given by the difference between the current temperature of the zone and the temperature of the walls that the zone is in contact with. Here all walls are assumed to have the same temperature $T_{\text{wall}} = 470\text{K}$.

Now we will turn to the problem of determining the mass-flow, that was also not included in Nilsson and Eriksson (2001), but which is needed to simulate equation (7.5).

7.2.3 The Mass-flow

There are two types of mass-flows included in the model. The first mass-flow simulates the combustion, i.e. a mass-flow from the uncombusted zone to the combusted

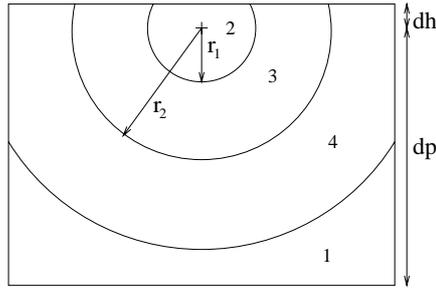


Figure 7.5 This shows the shells in a typical situation. Zone 3 has inner radius r_1 and outer radius r_2

zones. The second mass-flow simulates the high turbulence in the cylinder, and this mass-flow is a mass-transport between the different burned zones. The progress of the combustion is modelled by the Vibe function

$$x_b(\theta) = \frac{m_b}{m_{tot}} = 1 - e^{-\alpha \left(\frac{\theta - |\theta_0|}{\Delta\theta} \right)^{m+1}} \quad (7.13)$$

where x_b is the mass-fraction burned, i.e. the ratio between the burned mass, m_b , and the total mass, m_{tot} , as a function of the crank angle θ . Here θ_0 is the angle at which the combustion starts and $\Delta\theta$ is the combustion duration. Finally α and m are adjustable parameters. This is an experimentally developed function, and it is taken from Heywood (1988)). As can be seen in Figure 7.5 the zones are related in such a way that it is assumed that all mass-flow simulating the combustion goes only to the newest burned zone. That means that in Figure 7.5 there is only a mass-flow of the first type between the zone labelled 1 and the zone labelled 4.

There is also a mass-flow included between the burned zones and it is modelled as follows. From Heywood (1988) is given the rule of thumb that the mass-flow in the cylinder is well approximated by half the mean piston speed, i.e.

$$v = S \frac{\omega}{2\pi} \quad (7.14)$$

where S is the stroke and ω is the engine speed in radians per second. This expression gives values in agreement with other figures in Heywood (1988) (e.g. Figure 8-1), and the value of v was therefore chosen to 2.5 m/s. The mass-flow is then given by the contact area between the zones, and it is assumed that the flow is equal in both directions (see Figure 7.6). That means that this flow mostly serves as a sort of heat-transfer between the zones. Now all elements of the A and B matrices in equation (7.5) are specified and equation (7.5) can be formed. Let us therefore turn to the problem of calculating the ionization current from the state-vector \mathbf{x} in (7.6).

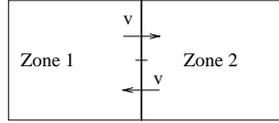


Figure 7.6 Due to in-cylinder gas movements, the zones are exchanging mass. Mass, with the temperature of zone 1, is flowing from zone 1 to zone 2. In return the zone gets mass from zone 2 with its temperature.

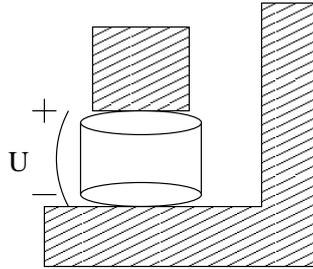


Figure 7.7 The two electrodes of the spark plug with the voltage U applied. The cylinder shaped volume contains the ions and the free electrons are conducting the current.

7.2.4 The Saitzkoff-Reinmann model

As argued in the introduction to this chapter we are following the research trace modelling the post-flame part of the ionization signal by thermal ionization. The model is based on the Saha equation, and this usage was originally suggested in Saitzkoff et al. (1996). It assumes that all the current flows in a cylinder shaped control volume put between the spark plug electrodes as shown in Figure 7.7. The derived relation is (see Saitzkoff et al. (1996) for other assumptions)

$$I = U \frac{\pi r^2}{d} \frac{e^2}{\sigma m_e \sqrt{\frac{8kT}{\pi m_e}}} \sqrt{\phi_s} \sqrt{\frac{2 \left(\frac{2\pi m_e kT}{h^2}\right)^{\frac{3}{2}} \frac{B_1}{B_0} e^{-\frac{E_1}{kT}}}{n_{tot}}} \quad (7.15)$$

$$\phi_s = \frac{[NO] \cdot 10^6}{n_{tot}/N_A} \quad (7.16)$$

$$n_{tot} = \frac{p}{RT} N_A \quad (7.17)$$

The way to calculate ϕ_s , and the parameter values given in 7.3, are taken from Andersson (2002). The interpretations of the parameters are given in Table 7.2

ϕ_s	Ratio of NO in gas mixture []	m_e	Electron mass [kg]
U	Measurement voltage [V]	B_i	Internal partition function
r	Radius of measurement cylinder [m]	E_1	Ionization energy for 1st order ionization [J]
d	Length of measurement cylinder [m]	n_{tot}	Total particle density density [1/m ³]
σ	Collision cross section [m ²]	k	Boltzmann's constant [J/K]
[NO]	NO Concentration [mol/cm ³]	h_p	Planck's constant [Js]
\tilde{R}	Universal gas constant []	e	Unit charge constant [As]
T	Temperature of gas [K]	N_A	Avogadro constant [molecule/mol]
p	Cylinder pressure [Pa]		

Table 7.2 Parameter list for the Saitzkoff-Reinmann equation.

U	80 [V]
r	1 [mm]
d	1 [mm]
σ	0.1 [Å ²]
m_e	9.31×10^{-31} [kg]
$\frac{B_i}{B_0}$	1 []
E_1	9.25 [eV]
k	1.38×10^{-23} [J/K]
h_p	6.63×10^{-34} [Js]
e	1.6×10^{-19} [As]
N_A	6.022×10^{23} [mol ⁻¹]

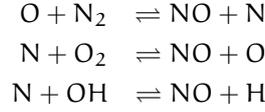
Table 7.3 Parameter values in Saitzkoff-Reinmann model.

In the expression the concentration of NO is needed. Here two ways of determining it have been compared. One assumes the NO concentration to be constant ($\Phi_s = 0.01$) and the other determines it by a simple chemical network. This latter approach is explained in the following section.

7.2.5 Dynamical NO concentration

Usually one divides the reactions occurring in a cylinder in three rate-classes: frozen, equilibrium and rate-limiting (Heywood, 1988). When simulating the thermodynamical development in Section 7.2.1 it was assumed that all reactions in the cumbusted zones were equilibrium reactions. Now we will have to relax this assumption for a few reactions. It should here be noted that since there is no feedback from the calculated NO-concentration to the pressure, volume or temperature, it does not lead to any logical inconsistencies when doing this. Just as in

Saitzkoff et al. (1996) and Andersson (2002), the extended Zeldovich mechanism is used to describe the NO-dynamics. It was originally presented in Zeldovich, Sadovnikov and Frank-Kamenetskii (1974), but this presentation is taken from Heywood (1988). The assumed reactions are



With the following two assumptions:

1. The content of N is small and changes slowly compared to the content of NO.
2. Concentrations of O, O₂, OH, H and N₂ can be approximated by their equilibrium concentrations.

the expression for NO formation is

$$\frac{d[\text{NO}]}{dt} = \frac{2R_1(1 - ([\text{NO}]/[\text{NO}]_e)^2)}{1 + ([\text{NO}]/[\text{NO}]_e)R_1/(R_2 + R_3)} \quad (7.18)$$

where

$$\begin{aligned} R_1 &= k_1^+ [\text{O}]_e [\text{N}_2]_e = k_1^- [\text{NO}]_e [\text{N}]_e \\ R_2 &= k_2^+ [\text{N}]_e [\text{O}_2]_e = k_2^- [\text{NO}]_e [\text{O}]_e \\ R_3 &= k_3^+ [\text{N}]_e [\text{OH}]_e = k_3^- [\text{NO}]_e [\text{H}]_e \end{aligned}$$

The concentration [] is in the unit [mol/cm³] and the reaction rate constants are listed in Table 7.4. The concentration [NO] is defined as

$$[\text{NO}] = \frac{N_{\text{NO}}}{V} \quad (7.19)$$

where N_{NO} is the quantity of NO in [mol] distributed in the volume V . The equilibrium concentrations are calculated by the CHEPP package (Eriksson, 2004). If V_b is the volume of the burned zone at question, equation (7.18) can be written as

$$\frac{d[\text{NO}]}{dt} = \frac{2R_1(1 - ([\text{NO}]/[\text{NO}]_e)^2)}{1 + ([\text{NO}]/[\text{NO}]_e)R_1/(R_2 + R_3)} - [\text{NO}] \frac{1}{V_b} \frac{dV_b}{dt} \quad (7.20)$$

We have now described all parts and all variations of the models that have been used to do the analysis and we are therefore ready to go to the results part.

	Rate constant	$\frac{\text{cm}^3}{\text{mol} \times \text{s}}$
k_1^+	$7.6 \times 10^{13} e^{-38000/T}$	
k_1^-	1.6×10^{13}	
k_2^+	$6.4 \times 10^9 \times T e^{-3150/T}$	
k_2^-	$1.5 \times 10^9 e^{-19500/T}$	
k_3^+	4.1×10^{13}	
k_3^-	$2.0 \times 10^{14} e^{-23650/T}$	

Table 7.4 Reaction rate constants for NO formation.

7.3 Model evaluation

The Saitzkoff model has been used in a number of previous works. One common feature, in the previous works, is that the simulated IPL is 2 CAD or more after the measured IPL (as discussed in the introduction or see e.g. Table 4.3, 4.4 or Figure 6.5 in Andersson (2002)). These models have all lacked heat-transfer, and therefore it is first seen (in Section 7.3.2) whether this feature will be present also for this model, setting the heat transfer to zero. Then the effect of adding first the heat transfer (in Section 7.3.3) and later the NO-dynamics (in Section 7.3.5) is investigated. First, however, a description of the data and of the comparison technique.

7.3.1 The experimental data

The data was collected on a 2.0 liter turbo charged gasoline engine. The engine was kept at 2000 revolutions per minute, and a scan was made in ignition angle and air/fuel ratio. The scan in ignition angle was between 27.0 and 18.1 CAD BTDC, which corresponds to MBT to MBT + 9. During this scan λ was kept at 1. During the scan in λ , which was between 0.88 and 1.06, the ignition angle was kept at 27.0. The parameters defining the geometry are given in Table 7.9 and the values of the parameters are given in Table 7.7.

Throughout the chapter the comparison between the model output and measurements has been done as follows. First the median cycles, with respect to PPL, were chosen for each operating point. This restriction was chosen since the combustion progress is described by the Vibe function, and since this function has problems fitting the extreme cycles. The next step was to tune the cycle-to-cycle dependent parameters to the specific cycle of study. Since some of the parameters depend on each other there are in effect not so many different degrees of freedom. Here these were described by variations of the residual gas fraction, x_{res} , and the combustion duration $\Delta\theta$. From these two parameters the initial temperature was calculated according to the following formula.

$$T_{\text{ivc}} = 1350 * x_{\text{res}} + 300 * (1 - x_{\text{res}})$$

Other cycle-to-cycle dependent parameters are the fraction of the fuel that was

combusted, η , and the α and m parameters in the Vibe function (7.13). The latter were, however, for all cycles except those mentioned in the Appendix A, set to the values given in Table 7.7. This would, however, not be possible if other operating ranges had been included in the data. After the model has been adjusted to fit the measured pressure curve for a specific cycle the model's IPL is compared to the measured IPL, for the same cycle. For all the operating points there was a clear peak in the post-flame phase, and the exact value of the IPL was obtained through visual interpolation.

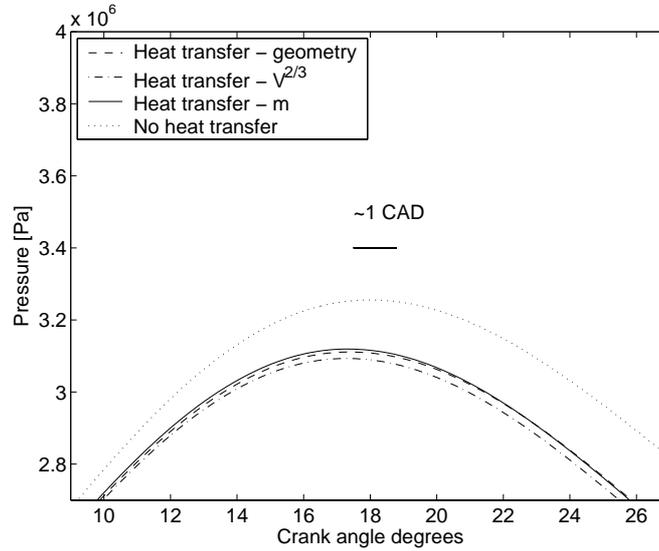


Figure 7.8 A comparison between the different models for calculation of the heat transfer. As a reference we see a pressure plot with no heat transfer. The choice of model only has a small influence on the cylinder pressure. See Appendix A for parameters.

7.3.2 Validation of the model without heat transfer

In this section the same discrepancy as in the earlier implementations of the Saitzkoff model will be reproduced, but this time with the multi-zonal model described in Section 7.2. Figure 7.9 and 7.10 shows the result of a simulation with the heat transfer set to zero, and with a static NO-model ($\phi_s = 0.01$). In this simulation 8-10 zones were used, and as can be seen also with this model is the predicted IPL approximately 2 CAD later than the measured one. The reason the number of zones is not a constant is that the zones divide when they reach a certain mass limit, m_b , and how many reaches this limit depends on all the other simulation parameters. The mass-limit used is given in Table 7.7. The pressure curves can

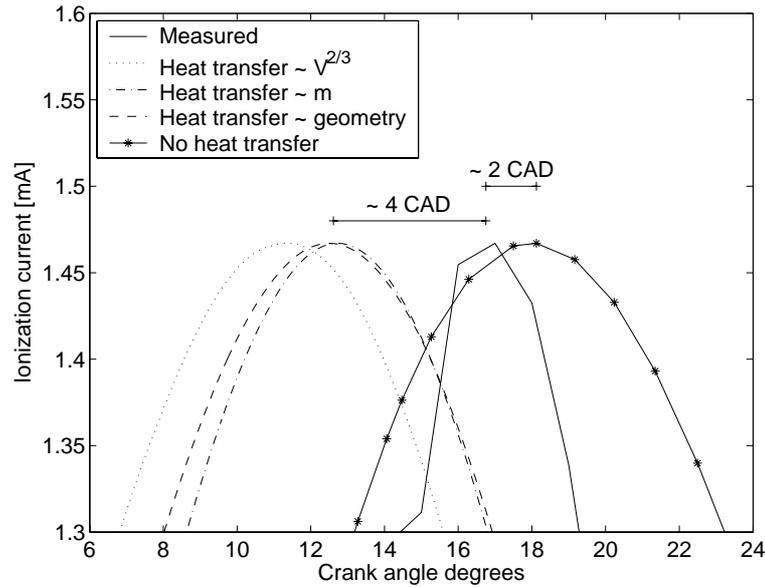


Figure 7.9 Simulations done with different heat transfer models and without heat transfer. The data is compared to the measured data, and all curves has been scaled to have the same maximum value. It can be seen that simulations with heat transfer result in a ionization current peak location about 4 CAD earlier than the measured peak location. But simulation without heat transfer results in a ionization current peak location about 2 CAD later than the measured peak location. The result of the latter simulation is consistent with previously published results, and also serves as a validation for the given implementation

be made to agree well for the compression and the early combustion phase. For the end of the combustion phase the simulated data is higher than the measured, but this is reasonable since there is no heat transfer included in the model. This pressure validation (see Figure 7.3 for a pressure validation with heat transfer included) together with the reproduction of the late IPL prediction observed earlier, serves as a validation of the implementation.

7.3.3 Addition of heat transfer

A plot comparing the pressure plots for the different heat transfer models, introduced in Section 7.2.2 is shown in Figure 7.8, and a plot comparing the ionization curves for the same models is shown in Figure 7.9. As can be seen the curves divide in two groups, the one without heat transfer in one group, and the three different

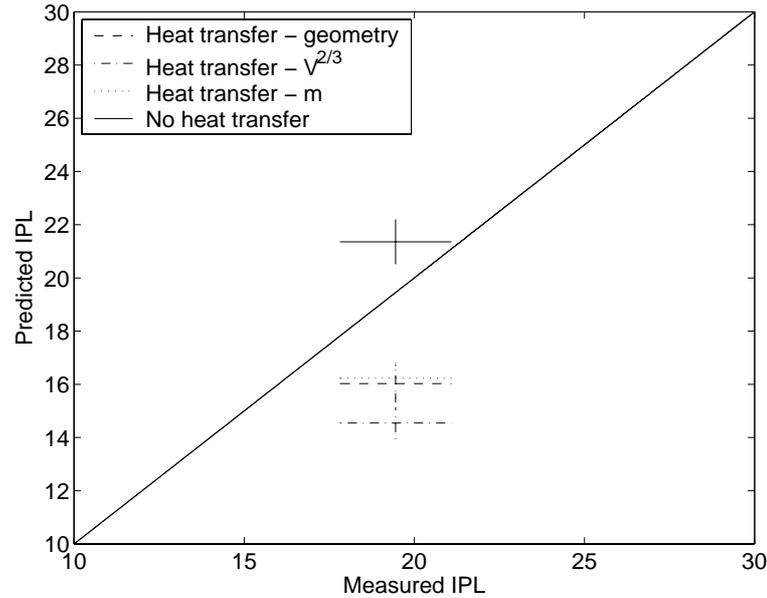


Figure 7.10 A comparison between measured and predicted IPL for the different heat transfer models. As can be seen all models with heat release systematically predicts the IPL too early, while, the one model without heat release predicts it to late. We can also see that volume based model (as in (7.12)), lies 1-2 CAD even further away from the x-y line. The chosen operating points is $(\theta_0, \lambda) = (24.1, 1)$.

models with heat transfer in the other. In the pressure graph (Figure 7.8), the difference between the PPLs for all the models including heat transfer is insignificant, and the difference between the PPL for any of these models, and the PPL without heat transfer is about 1 CAD. From Figure 7.9 and 7.10 a bigger difference between the different heat transfer models can be observed. The IPL prediction with mass and geometric basis is within one cad, but the volume based calculations differs approximately 1.5 CAD from the two others. What can also be seen, and what is more important, is that the introduction of heat transfer models has moved the IPL 5-8 CAD earlier with respect to the model without heat transfer. The main conclusion drawn from this is that the choice of heat transfer model is not as important as the choice of whether one should include it or not. It is also clear that a model with heat transfer included does no longer have the problem of predicting the IPL too late, but on the contrary, it predicts it too early. This new problem is attacked in the next section. After this section it is only the geometrically based heat transfer model that has been used.

7.3.4 Model sensitivity analysis

It was shown in the previous section that the inclusion of a heat transfer model does remove the old problem of predicting the IPL too late, but instead makes the model predict a too early IPL. This new problem is approached in the following way. First a sensitivity analysis is done for the parameters in the model, with geometrically based heat transfer, that effects the difference IPL-PPL. Also other small variations, like choice of fitting method, and data uncertainty is analysed. Then the dynamical NO-model, described in Section 7.2.5, is added to the model, to see whether it can improve the agreement. These analyses serve two purposes. Firstly, the sensitivity analyses serve as a quantification of how much the various parts of the model affect the relative position of the simulated pressure and ionization curves. Secondly, it serves as an ensurance that the difference observed in the previous section is really significant, i.e. that it can not be solved by small variations in e.g. one of the model's parameters.

Sensitivity to the fitting

The first thing that was tested was how dependent the IPL prediction is on the fitting of the model to the measured pressure curve. Since no automatic process has been constructed, this was easiest checked by letting two persons do a fitting on the same data. The result of this process is shown in Table 7.5, and in Figure 7.11. The figure also includes data from Table 7.6, which is from the other half of the data. For this latter half, however, only one fitting was done. It can be seen in the plot that the difference between the measured and predicted IPL is 3-5 CAD for all operating points and the difference between the two independent fittings is less than one CAD. From this we conclude that the prediction of IPL is not highly sensitive to the fitting process.

Sensitivity to parameter variations

Another sensitivity measure is that of the model with respect to small variations in the parameters. The result of such a sensitivity analysis is presented in Table 7.7. It has been generated in the following way. For each parameter appearing in the model a perturbation with 10% followed by a simulation has been done. For some parameters, however, a perturbation size of 1% was chosen. This was due to an either unfortunate formulation of the parameter value, making a 10% perturbation unrealistically big, or to interdependence between parameters, leading to simulation problems when perturbing them too much separately from each other. Each such perturbation results in a simulated PPL and a simulated IPL. Let $PPL(x)$ mean the PPL value after a perturbation x , and analogously for $IPL(x)$. Further let the difference between these two values be denoted $\Delta_{pi}(x)$, i.e.

$$\Delta_{pi}(x) = PPL(x) - IPL(x) \quad (7.21)$$

θ_0 [CAD]	Comparison I		Comparison II		Measured	
	Mean	Std	Mean	Std	Mean	Std
27.0	12.36	0.27	12.44	0.39	16.00	2.14
21.1	16.19	0.52	16.03	0.71	19.45	1.64
24.1	18.26	0.40	18.26	0.48	21.65	4.15
18.1	20.74	1.28	20.05	0.85	25.25	2.75

Table 7.5 Data from two independent fittings on the same data. The first column specifies the ignition angle for the working point, and for all points λ was equal to one. The four next columns describe the mean value and standard deviation for the IPL, obtained from the two fittings, for each of the operating points. The last two columns gives the measured mean and standard deviation of the IPL for the same data. The two independent fittings have resulted in different χ_{res} , and $\Delta\theta$, values. The first fitting was done by Karl-Johan Nogenmyr (Nogenmyr, 2003), and the second was done by Daniel Claesson (Claesson, 2004). As can be seen in the data, the standard deviations in the measured values are quite big. This is a consequence of the large cycle-to-cycle variations of the ionization currents. On the other hand the two different estimations have given similar results, and this indicates the robustness to the fitting process.

λ	Comparison III		Measured	
	Mean	Std	Mean	Std
0.8824	11.17	0.44	13.25	1.67
0.9200	11.46	0.62	14.00	2.88
0.9536	12.32	0.50	15.00	1.41
1.0682	14.50	0.75	18.40	0.55

Table 7.6 Mean values and standard deviations of the predicted ionization peaks, compared with the measured IPLs for the same cycles. The data spans four different lambda values and for all of them θ_0 has been 27.2. The simulated peaks are also for this part of the dataset 2–4 CAD earlier then the measured peaks. This comparison was done by Claesson (Claesson, 2004).

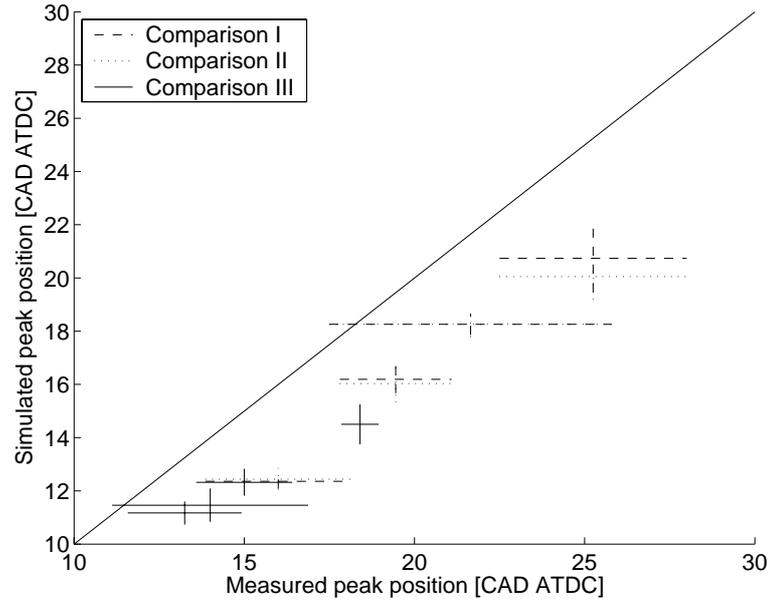


Figure 7.11 The simulated position of the ionization peaks against the measured peaks for three independent fittings, with geometrically based heat transfer but without inclusion of the NO-dynamics. The data is the same as in Tables 7.6 and 7.5. The lengths of the lines are two standard deviations, and the mean value is the middle of the line. The large standard deviations in the measured data are a consequence of the large cycle-to-cycle variations in the measured ionization current data. As can be seen the data from the two fittings are almost identical.

Finally let the difference between $\Delta_{pi}(x)$ before and after the perturbation be denoted $\Delta_{\Delta}(x)$

$$\Delta_{\Delta}(x) = \Delta_{pi}(x) - \Delta_{pi}(0) \quad (7.22)$$

It is this $\Delta_{\Delta}(x)$ that has been observed for each parameter perturbation. The reason why this quantity has been observed, and not e.g. the difference $IPL(x) - IPL(0)$ is that both the PPL and the IPL will change when a parameter is perturbed. Therefore, when doing a parameter perturbation before a fit, with the fit done as described in Section 7.3.1, the cycle-to-cycle dependent parameters will adapt to keep the PPL fix (equal to the measured value). It is therefore only the difference $\Delta_{\Delta}(x)$ that has the potential of changing the agreement between model and experiments, and this is the reason why this is the quantity observed in the sensitivity Table 7.7. The generality of the result from the analysis is based on the assumption that neither the adaption of the cycle-to-cycle dependent parameters nor the choice

Parameter	Value	$\Delta_{\Delta}(-x)$	$\Delta_{\Delta}(+x)$	$x[\%]$
C_1	2	-0.44	0.47	10
C_2	0.44	-0.18	0.33	10
t_{Stop}	0.0736	~ 0	~ 0	10
m	4	0.46	-0.32	10
a	20	0.34	~ 0	10
p_{ivc}	6.5e4	0.18	~ 0	10
T_{ivc}	363	~ 0	0.23	10
ϕ	1	~ 0	~ 0	10
ϕ_{res}	1	~ 0	~ 0	10
x_{res}	0.065	0.11	~ 0	10
$(A/F)_s$	14.7	~ 0	~ 0	10
$\frac{V_{b,i+1}}{V_u}$	1e-5	0.11	~ 0	10
m_{max}	4.7e-5	~ 0	~ 0	10
a	3.9e-2	0.34	~ 0	10
l	15.9e-2	~ 0	~ 0	10
V_d	4.96e-4	-0.18	0.33	10
dt	2e-3	~ 0	~ 0	10
T_{wall}	470	~ 0	0.11	1
ω	200	~ 0	~ 0	1
t_{Start}		~ 0	~ 0	1
θ_0	-27	~ 0	0.11	1
$\Delta\theta$	44	0.17	~ 0	1
h	2e-5	~ 0	~ 0	1
η	0.97	~ 0	~ 0	1
B	9.0e-2	~ 0	0.11	1
r_c	9.25	~ 0	~ 0	1
V_c	6.01e-5	~ 0	0.10	1
S	0.078	~ 0	~ 0	1

Table 7.7 Parameters appearing in the model, as well as their affect on the quantity $\Delta_{\Delta}(x)$ defined in (7.22). The final column gives the perturbation size used for the each parameter.

of operating point will significantly affect the model agreement. The assumption is supported by the relative robustness of both the fitting and the choice of operating point observed in the previous sensitivity analysis (in Figure 7.11).

As can be seen in the sensitivity tables none of the parameters change the difference with more than 0.5 CAD, and most of them only changes the difference with less than 0.1 CAD (which has been set to ~ 0). From this we draw the conclusion that no single perturbation of a parameter can lead to significant improvements of the approximately 4 CADs early prediction of the IPL obtained by adding the heat transfer in Section 7.3.3.

T_{wall}	Temperature at the cylinder wall [K]
C_1	Woschni heat transfer constant []
C_2	Woschni heat transfer constant []
ω	The angular velocity [rad/s]
t_{Start}	Simulation start-time [s]
t_{Stop}	Simulation stop-time [s]
θ_0	Ignition angle [CAD]
$\Delta\theta$	Combustion duration [CAD]
m	The m parameter for the Vibe function []
a	The a parameter for the Vibe function []
p_{ivc}	Pressure at intake valve close [Pa]
T_{ivc}	Temperature at intake valve close [K]
ϕ	Fuel / air equivalence ratio []
ϕ_{res}	Fuel / air equivalence ratio in residual gas []
x_{res}	Residual gas fraction []
$(A/F)_s$	Stoichiometric air/fuel ratio []
h	The time-step of the simulation [s]
$\frac{V_{b,i+1}}{V_u}$	Fraction of the unburned zone used to create a new burned zone []
m_{max}	The maximum mass that a boundary zone might have [kg]
η	Fraction describing how much fuel is left unburnt []

Table 7.8 *Simulation parameters that were increased and decreased in the sensitivity analysis. A short explanation to each parameter is also given.*

a	Crank radius [m]
l	Connecting rod length [m]
B	Cylinder bore [m]
r_c	Compression ratio []
V_d	Displaced volume [m ³]
V_c	Clearance volume [m ³]
S	Piston stroke [m]
dt	Distance spark-gap to cylinder head [m]

Table 7.9 *Engine parameters that were increased and decreased in the sensitivity analysis. A short explanation to each parameter is also given.*

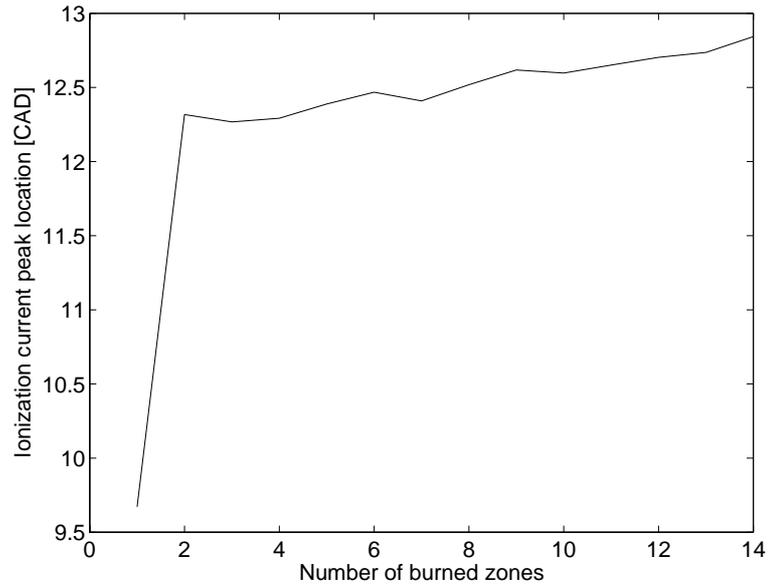


Figure 7.12 The ionization current peak location for simulations done with different number of zones.

Sensitivity to number of zones

Another uncertainty is the effect of the number of zones chosen. Since it is clear that the effect of heat transfer is significant, it is interesting to see whether dividing the heat transfer between few or many zones makes a difference for the zone that contains the spark plug. In Figure 7.12 the IPL for different number of zones is plotted. As can be seen, apart from the change between one and two combusted zones, there is no significant change when changing the number of zones. In all other plots in the thesis 8-10 zones has been used, and for this region the model is concluded to not be sensitive to the number of zones.

Sensitivity to mass-flow

Another uncertainty in the model is the mass-flow. The mass-flow was set with the rule of thumb given in equation (7.14). The value was also in agreement with some experimental recordings in Heywood (1988). The sensitivity to the estimation of the value of v in (7.14) was checked by varying this parameter with $\pm 100\%$. The result is plotted in Figure 7.13, and as can be seen the change in IPL is not significant. It should, however, be added that a greater uncertainty probably lies in the way that the mass-flow has been modelled.

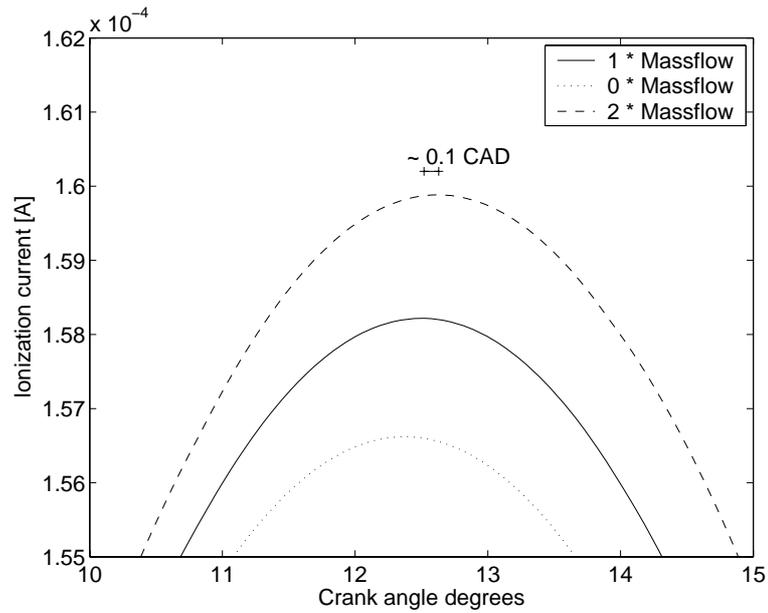


Figure 7.13 Simulated ionization curves for different values of the mass-flow velocity v , approximated by equation (7.14). The value has been perturbed with 100% in both directions, and as can be seen the change of IPL is less than 0.2 CAD. We conclude that this parameter is not important for the prediction.

Sensitivity to data uncertainty

We also need to estimate the uncertainty contribution from the data. Our main objective is to see if the observed difference between the predicted and measured IPL (in Figure 7.11) already lies within the measurement uncertainty or if the agreement can be further improved. That means that we need to find an upper limit to the uncertainty. If we find such a limit that is smaller than the observed difference (from adding the heat transfer in Section 7.3.3) we will conclude that we need to improve the model further. We approximate the uncertainty in the estimation of the timing in the sampling to be less than one CAD. The basis for this is that the data are sampled each CAD. If this approximation is valid, this uncertainty is of an order of magnitude less than the observed discrepancy by adding the heat transfer, and it does not violate the conclusion. Another data uncertainty comes from the calibration of the pressure sensor. An uncertainty in the absolute value of the measured pressure will affect the fitting process deciding the cycle-to-cycle dependent parameters. When doing this fitting there are basically two degrees of freedom, here chosen as χ_{res} and $\Delta\theta$ (see Section 7.3.2). The fitting has been done

by the heuristically based principle of first fitting the position of the pressure peak with $\Delta\theta$ and then fitting the amplitude with the χ_{res} parameter. From Table 7.7 we see that the sensitivity of the observed difference to the χ_{res} parameter is almost zero. From this we draw the conclusion that the observed difference cannot be explained by uncertainties from the pressure sensor calibration. Finally we have a big data uncertainty coming from the noise in the ionization sensor. The effect of the noise is illustrated by the variance plot in Figure 7.11, and as can be seen this effect is less than the observed discrepancy. However, the ionization sensor signal also contains systematic noise components. These come from the multitude of processes that are determining the ionization current. One such is the remaining effect of the flame-front phase. That means that the observed signal is really a superposition of a signal that (most likely) has a negative slope and a signal that has a peak in the post-flame phase. However, independently of the exact shape of the remaining transient from the flame-front part, as long as it has a negative slope the real post-flame peak will lie later than the observed peak. That means that the inclusion of this data uncertainty will (if the assumption of negative slope holds) increase the difference obtained by adding the heat transfer to the model. After this discussion of the various uncertainties we draw the conclusion that the agreement can be further improved, and from the previous sensitivity analysis we conclude that we need to exploit new model structures to obtain it.

7.3.5 Addition of Dynamical NO-model

In Figure 7.14, 7.15 and Table 7.10 can be seen the result of a fitting of the model with respect to the same data as earlier, but after including the dynamical NO-model described in Section 7.2.5. As can be seen, with the difference between the predicted and the measured IPL is less than one CAD for all working points. We therefore conclude that the combination of the heat transfer model, introduced in Section 7.2.2, and the dynamical NO-model, introduced in Section 7.2.5, together make the model predict the IPL with a high accuracy. It should be emphasized that one of the strengths of this result is that the agreement has been best for the most realistic model. All previous models are containing more or less crude approximations of the features included in this model, and if they would have modelled a better agreement with the measurements, we would have been forced to go back to the modelling phase and understand why. Now, however, we have included all the details that we know, and the results is therefore the desired: *A physically based model that describes the relationship between the position of the pressure curve and the measured IPL, with an agreement within one CAD.*

7.4 Conclusions

This chapter has presented a multi-zonal model for cylinder pressure, temperature and ionization currents. It has been shown that the reason previously published models predict a too late IPL position, is that they have not included heat trans-

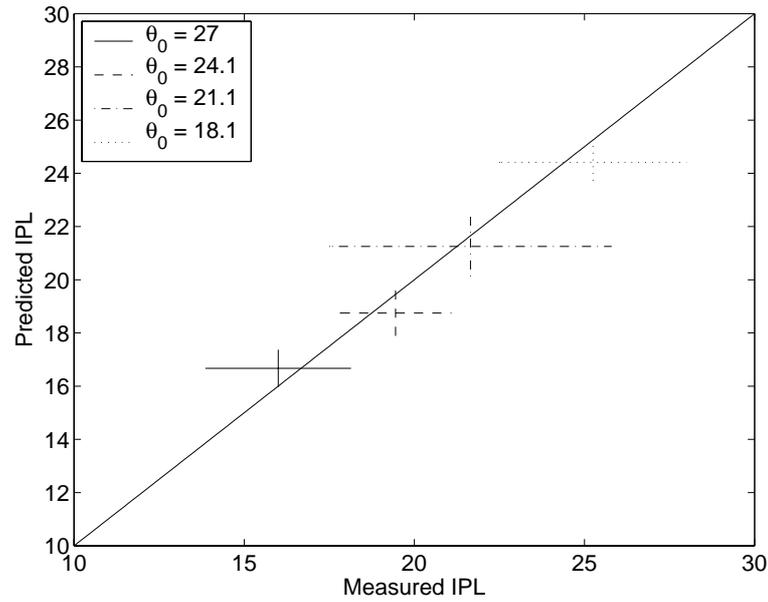


Figure 7.14 The agreement between predicted and measured IPL for four different operating points. The data is taken from Table 7.10 and as can be seen the agreement is within one CAD.

θ_0 [CAD BTDC]	Simulation		Measured	
	Mean	Std	Mean	Std
27.0	16.67	0.70	16.00	2.14
24.1	18.75	0.86	19.45	1.64
21.1	21.25	1.12	21.65	4.15
18.1	24.41	0.69	25.25	2.75

Table 7.10 Data from simulations done with the dynamical NO-model. Now the simulated and measured peak positions of the ionization currents are within 1 CAD. The standard deviations from the simulations are approximately 1 CAD.

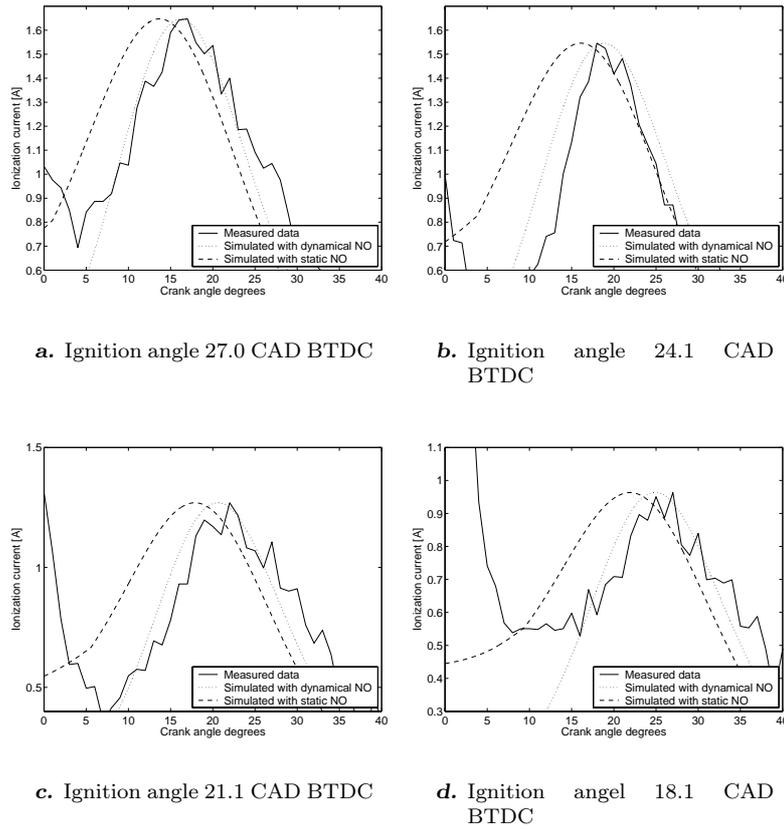


Figure 7.15 The measured- and simulated ionization current for four different ignition angles are plotted. Plotted is the average curves, that is the mean value of all the cycles. As can be seen the simulated plot with dynamical NO shows good agreement with the measured plot.

fer. The presented model includes a realistic heat-transfer, that together with a dynamical NO-model is capable of predicting the measured IPL within one CAD. The final heat transfer model is assuming symmetrical spheres that grows out as the combustion proceeds. However, two other models have also been tested, and it was observed that the difference, when predicting the IPL, was bigger between the models with and without heat transfer, than between the different models including heat transfer.

A Model for Insulin Signalling to Map-kinase Control in Fat Cells

This chapter is the second of the two that deals with system identification in practice, i.e. with development and analysis of specific models in specific applications. While the previous chapter dealt with the development of model for a combustion engine, this chapter is concerned with cell biology. The differences between these two applications show how diverse thermochemical systems might appear, and how general the problem of system identification is. The model presented in this chapter describes the intra-cellular response to insulin in human fat cells. A motivation of why such a model is interesting, together with a short review of the necessary theory is given in the first section. Then, in Section 8.2, the actual model is presented. Section 8.3 presents some basic analysis of the model. The *in silico* experiments include: i) time-series analysis, ii) dose-response curves and iii) a search for experimentally obtainable oscillations. Section 8.4 compares the model simulations with previously published experimental data. The overall agreement is quite good, but the fit is not perfect, and some general things can also be improved. This improvement, together with a sensitivity analysis, is done in Section 8.5. Section 8.6 contains a discussion of the present accuracy of the model and discusses what the model can and cannot be used for now. Finally, as in the other chapters, comes a summary of the results obtained in the chapter.

8.1 Background

This sections starts by giving a general introduction to the specific pathway that will be modelled, and ends by reviewing the necessary theory for understanding how this has been done.

8.1.1 The mitogenic response to insulin in fat cells

The main role of insulin in the human body is to regulate the level of glucose (sugar) in the blood. Insulin is produced in the β -cells in the islands of Langerhans, and the production increases when the blood glucose level rises. When the production of insulin increases, the concentration of insulin in the blood also increases, and this is the signal to the cells that there is glucose in the blood, and that the cells should change their metabolism to use glucose instead. The conventional view of how this information is given to the cells is the following. The free insulin molecules in the blood can dock to certain insulin receptors that are situated all over the cell membrane (boundary separating the inside from the outside of the cell). When an insulin molecule has coupled to such a receptor, the receptor and insulin will form a complex, and this complex will change shape, once the coupling is complete. The complex now has the potential to bind another molecule, a phosphate group. This will cause the insulin-receptor complex to change mode, into an active mode. When the insulin-receptor complex is activated, it has the possibility to bind to another substrate, called insulin receptor substrate (IRS). Once this substrate has bound to the complex, and thus forming an even greater complex, the IRS molecule can also bind a phosphate group. Once this has been done, the IRS molecule will detach from the Insulin-Receptor-IRS complex and allow for this complex to couple to another IRS molecule. The decoupled IRS molecule has still got the phosphate group and it is therefore denoted IRS-P. IRS-P is the active form of IRS, and it can be used to couple to other substrates, and activate them in a similar fashion. These activated enzymes will then be capable of activating yet other enzymes and in this way the original insulin molecule starts a chain reaction, where each step is activating or deactivating some signalling intermediate in a network of substrates (see Figures 8.1 to 8.5 for the network included in this model). In a network there might also be upstream feedbacks back to previous steps in the network (see e.g. Figure 8.5), and it is the presence of negative feedback loops that makes a system capable of having oscillations. The activation signal that proceeds into the cell is usually divided into two major sub-branches. One of these governs the metabolic response and the other governs the mitogenic. The metabolic response has to do with things like energy production and consumption, and the mitogenic response has to do with things like cell growth and division, cell death and protein transcription. It is only the mitogenic response that is spread by phosphorylation cascades as described above. In the model presented in this chapter it is only the mitogenic response of insulin that is included. The model starts with extracellular insulin concentration and ends with the doubly phosphorylated mitogen activated protein kinase (MAPK-PP) (see Figure 8.4 and 8.5). MAPK-PP is a protein kinase (activator) of transcription factors, and once these have been phosphorylated they go into the cell nucleus, and there start transcription of various proteins. This last part is not included in the model. The time-scale of the entire response is in the order of a few minutes, even though some individual reactions equilibrate in the time scale of micro-seconds. One major reason why it is interesting to model the response to insulin is the following. The common disease type II Diabetes Mellitus

(sometimes called age-diabetes), is often caused by insulin non-sensitivity. When this is the case, the disease can not be cured by adding more insulin to the blood, and therefore usually other methods has to be used. Since this kind of diabetes is becoming increasingly more common, and since it is not fully understood what goes wrong in the cell response, research within this area has a high priority. Considering the complexity of the system, it is a great help to complement the experiments with mathematical models. It should, be emphasized that mathematical modelling can in no way replace real experiments. It can only be a means of structuring the given information, generating new hypothesis, be a cheap and fast alternative when testing things not yet tested etc. Now follows a short summary of two standard ways of translating chemical reactions into ODEs.

8.1.2 Chemical kinetics revisited

In this Subsection follows a short review of the basic theory regarding cellular modelling, a more extensive treatment is found in e.g. Heinrich and Schuster (1996). When modelling intra-cellular reactions a common way is by assuming continuous concentrations, and collision-based reactions. That means that one considers reactions of the form



where v is the velocity and the question mark stands for an analytical expression that consists of chemicals (not necessarily A, B or C), and reaction parameters. The simplest case is the law off mass-action and another common case, that is often used when modelling enzymatic reactions, is the Michaelis-Menten expression.

Mass-action kinetics

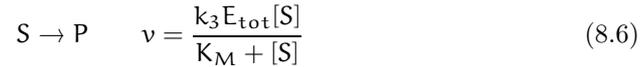
In the case of mass-action kinetics the reaction velocity in (8.1) is considered proportional to the concentration of A and B. It can be written $v = k[A][B]$ where k is a reaction constant that in general depends on things like temperature, pressure, and pH. In cellular models, however, these things are considered constant in time and, hence, so will the reaction parameters.

Michaelis-Menten kinetics

When reactions include enzymes like in the phosphorylation cascade described in Section 8.1.1, one often does a combination of several steps in a way that usually is denoted the Michaelis-Menten approximation. In the simplest case one combines the following four reactions



into the single reaction



This simplification is possible if one can assume that $k_4 \sim 0$, and $\frac{d[SE]}{dt} \sim 0$. The parameter E_{tot} is the total enzyme concentration. For a derivation of the relationship, see e.g. (Heinrich and Schuster, 1996).

8.2 The model

Prior to this work there were no existing (mathematical) model in the literature that described a connected signal transduction pathway from extracellular insulin concentration to double phosphorylated MAPK. There are, however, many models that describe or include parts of the network lying between these two enzymes, and some of these models have been used as a basis for the construction of this model. The model is presented in four parts.

i) The first part starts with extracellular insulin concentration and ends with the activation of the insulin receptor substrate (IRS). This part is assumed common to both the metabolic and the mitogenic response to insulin and this part of the model is mainly taken from Sedeghat, Sherman and Quon (2002) which describes the former of these responses.

ii,iii) The second and third parts of this pathway consists of SOS, Grb2, GDP and Ras, and these steps are taken from the DOQCS database (Doqcs, 2001) and the models presented by Bhalla and Iyenger, (Bhalla and Iyenger, 1999).

iv) The final part of the model is known as the MAPK-cascade, and it is one of the most modelled parts of the pathway. Also here a model by Bhalla and Iyenger has been used as the base model during the construction.

8.2.1 The first step, including insulin, IR and IRS

The first step of the pathway (Figure 8.1) including insulin, the insulin receptor (IR) and the insulin receptor substrate (IRS), is modelled according (Sedeghat et al., 2002). Although their model is designed for the metabolic signalling pathway, these initial reactions are assumed to be the same for the mitogenic pathway (Leng, Karlsson and Zierath, 2004). The concentration of insulin is given as a parameter and acts as input to the system. Extracellular insulin binding to IR is modelled as a reversible reaction (reaction number 12 in Figure 8.1). IR is then autophosphorylated (reaction 13) and can only be dephosphorylated after insulin dissociates from it (reaction 14). The dephosphorylation (reaction 17) is modelled with a multiplicative factor, PTP, which represents the relative activity of all the protein tyrosine phosphatases in the cell that dephosphorylate IR (and IRS, see below). The phosphorylated receptor can also bind a second insulin molecule, in a reversible reaction (reaction 15). Both IR-P forms phosphorylate IRS into IRS-P (reaction 16). This reaction's dependence on the concentration of IR is modelled as a linear dependence on the fraction of phosphorylated receptors $([IR-insP] + [IR-2ins]) / IR_p$,

where IRp is a parameter corresponding to the concentration of phosphorylated receptors during maximal insulin stimulation. The dephosphorylation of IRS-P is modelled with the factor PTP as explained above (reaction 17). Figure 8.1 gives a schematic view of these reactions and their respective rate equations are stated in equations (8.7)-(8.14). Compared to the model in ((Sedeghat et al., 2002)) the endocytosis of insulin receptors into the cell and the further steps relevant only to the metabolic pathway, have been excluded from the present model. A detailed description of the model including all parameters, initial values etc, both for this and all the other steps is given in Appendix A.6.

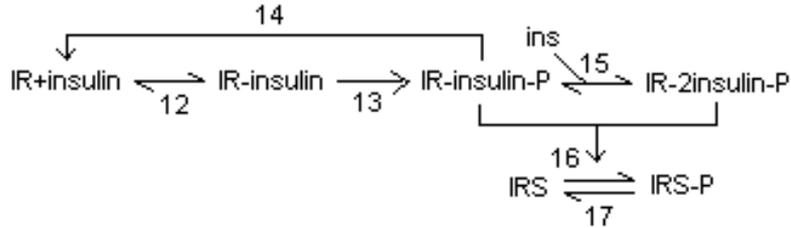


Figure 8.1 Schematic view of the first step in the insulin model. Insulin binds reversibly to the insulin receptor, IR, which then autophosphorylates. A second insulin molecule can also bind to the receptor when it is phosphorylated. The phosphorylated receptors then activates the insulin receptor substrate, IRS. All activation/deactivation is done by phosphorylation.

$$vf_{12} = Kf_{12}[IR][ins] \quad (8.7)$$

$$vb_{12} = Kb_{12}[IR - ins] \quad (8.8)$$

$$v_{13} = K_{13}[IR - ins] \quad (8.9)$$

$$v_{14} = K_{14}PTP[IR - ins - P] \quad (8.10)$$

$$vf_{15} = Kf_{15}[IR - ins - P] \quad (8.11)$$

$$vb_{15} = Kb_{15}[IR - 2ins - P] \quad (8.12)$$

$$v_{16} = K_{16}[IRS]([IR - ins - P] + [IR - 2ins - P])/IRp \quad (8.13)$$

$$v_{17} = K_{17}PTP[IRS - P] \quad (8.14)$$

8.2.2 The second step, including Sos and Grb2

The next part of the pathway includes Sos, Grb2 and the negative feedback induced by MAPK-PP, the last component of the pathway.

Sos and Grb2 reversibly form a complex (reaction number 21 in Figure 8.2), which then binds to IRS-P (reaction 22), from the previous step, and forms the

super complex IRS-P-Sos-Grb2. This binding is modelled with the same parameters as for the binding of the super complex SHCP-Sos-Grb2 in the model of Bhalla and Iyenger ((Bhalla and Iyenger, 1999) and (Bhalla, Ram and Iyenger, 2002)). This approximation is supported by the fact that the two complexes IRS-P-Sos-Grb2 and SCHP-Sos-Grb2 have the same downstream function in the pathway. Further, MAPK-PP can phosphorylate Sos (reaction 20), and Sos-P can bind to Grb2 reversibly (reaction 18), but Sos-P-Grb2 can not bind to IRS-P. This has the effect that the concentration of the complex Sos-Grb2 is reduced, which in turn reduces the concentration of IRS-P-Sos-Grb2. The extension of this is that the concentration of MAPK-PP is reduced, which in turns causes a reduced level of Sos phosphorylation. This feature is a negative feedback and it is one of two included in the model. The dephosphorylation of Sos-P is modelled as an intrinsic reaction, i.e. it is only proportional to the concentration of Sos-P itself (reaction 19). The schematic view of these reactions are shown in Figure 8.2 and their respective rate equations are stated in equations (8.15)-(8.22). Note that MAPK-PP is shortened to KPP in the equations.

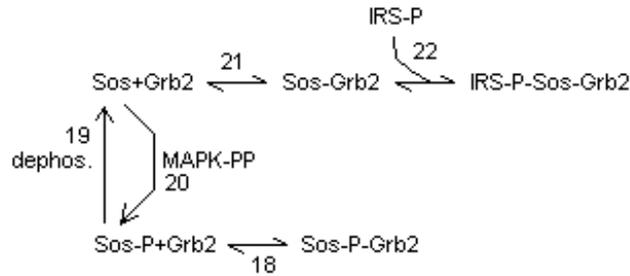


Figure 8.2 Schematic view of the second step in the insulin model. Sos binds to Grb2 and forms a complex which then binds to IRS-P (from the previous step). MAPK-PP (from the last step) phosphorylates Sos which then also can form a complex with Grb2. However this complex can not bind to IRS-P. Hence this reaction is a negative feedback for the system.

$$vf_{18} = Kf_{18}[\text{Sos} - \text{P}][\text{Grb2}] \quad (8.15)$$

$$vb_{18} = Kb_{18}[\text{Sos} - \text{P} - \text{Grb2}] \quad (8.16)$$

$$v_{19} = K_{19}[\text{Sos} - \text{P}] \quad (8.17)$$

$$v_{20} = \frac{V_{20}[\text{KPP}][\text{Sos}]}{[\text{Sos}] + k_{20}} \quad (8.18)$$

$$vf_{21} = Kf_{21}[\text{Sos}][\text{Grb2}] \quad (8.19)$$

$$vb_{21} = Kb_{21}[\text{Sos} - \text{Grb2}] \quad (8.20)$$

$$vf_{22} = Kf_{22}[\text{Sos} - \text{Grb2}][\text{IRSP}] \quad (8.21)$$

$$vb_{22} = Kb_{22}[\text{IRS} - \text{P} - \text{Sos} - \text{Grb2}] \quad (8.22)$$

8.2.3 The third step, including GDP and Ras

The next part of the pathway is the one preceding the MAPK cascade. Its main function is the activation of Ras by exchange of its bound GDP group with a GTP-group. In the model this activation of Ras is catalyzed by the complex Sos-Grb2-IRS-P (reaction 23 in Figure 8.3). This is an important reaction in this step because it is the only reaction coupling the previous steps with the MAPK cascade. Another activator of Ras included in the model is the phosphorylated Guanine nucleotide exchange factor (GEF) (reaction 24). Both GEF and IRS-P-Sos-Grb2 activate Ras by enzymatic reactions. Ras is deactivated by GTPase activation protein (GAP), which also is modelled as an enzymatic reaction (reaction 25). Ras also has an intrinsic deactivation (reaction 26). GEF and GAP can both be phosphorylated by Protein kinase C (PKC) (reactions 28 and 30) and have intrinsic dephosphorylation (reactions 27 and 29). Note that although PKC activates GEF and inactivates GAP the effect of both these phosphorylations on GTP-Ras phosphorylation is positive. Also note that there is no feedback from the rest of the system to neither GAP nor GEF, and hence these two branches (reaction 27-30) gives a certain zero-level phosphorylation degree. Figure 8.3 shows the reactions. The reaction expressions for these steps were taken from models on the DOQCS database (Doqcs, 2001). This database provides several models for this step, including varying number of reactions. However, the approximation of equating IRS-P-Sos-Grb2 with SHCP-Sos-Grb2 had to be done also here. Among the existing models the one with the least number of reactions was chosen for simplicity, and equations (8.23)-(8.30) state the rate equations.

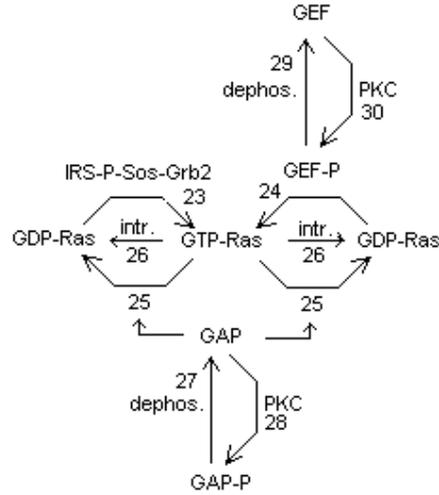


Figure 8.3 Schematic view of the third step in the insulin model. The complex *IRS-P-Sos-Grb2* (from the previous step) activates *GDP-Ras* by phosphorylation into *GTP-Ras*. *GAP* and *GEF* and *PKC* act as background signal from other interacting pathways.

$$v_{23} = \frac{V_{23}[\text{IRS-P-Sos-Grb2}][\text{GDP-Ras}]}{[\text{GDP-Ras}] + k_{23}} \quad (8.23)$$

$$v_{24} = \frac{V_{24}[\text{GEF-P}][\text{GDP-Ras}]}{[\text{GDP-Ras}] + k_{24}} \quad (8.24)$$

$$v_{25} = \frac{V_{25}[\text{GAP}][\text{GTP-Ras}]}{[\text{GTP-Ras}] + k_{25}} \quad (8.25)$$

$$v_{26} = K_{26}[\text{GTP-Ras}] \quad (8.26)$$

$$v_{27} = K_{27}[\text{GAP-P}] \quad (8.27)$$

$$v_{28} = \frac{V_{28}\text{PKC}[\text{GAP}]}{[\text{GAP}] + k_{28}} \quad (8.28)$$

$$v_{29} = K_{29}[\text{GEF-P}] \quad (8.29)$$

$$v_{30} = \frac{V_{30}\text{PKC}[\text{GEF}]}{[\text{GEF}] + k_{30}} \quad (8.30)$$

In Figure 8.4 the first three steps are included to get a full view of the model.

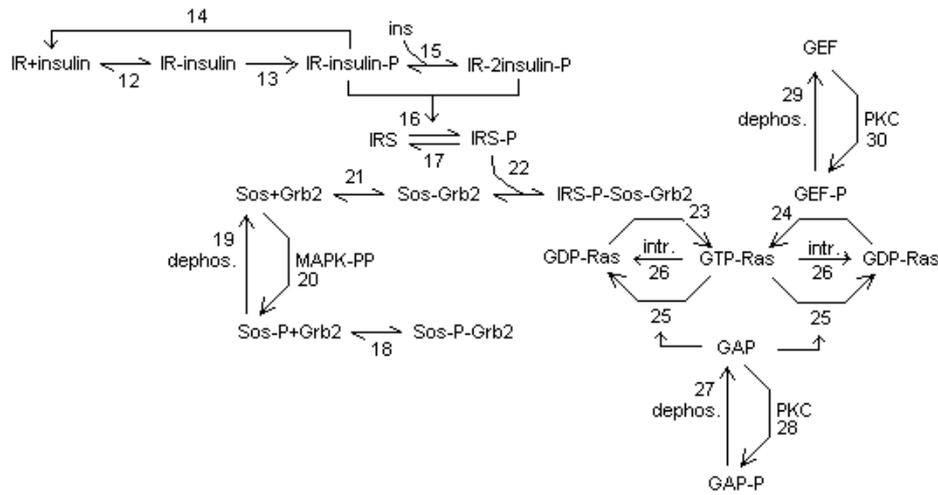


Figure 8.4 Model of the mitogenic insulin pathway previous to the MAPK cascade with all the previously explained steps included.

8.2.4 The fourth step, the MAPK-cascade

Finally, as the last step in the pathway, is the MAPK-cascade. This is the most modelled part of the pathway and there is much interesting dynamics taking place, considering only this part of the pathway. The insulin signals are brought to the MAPK-cascade through GTP-Ras that binds to Raf-P and forms the complex GTP-Ras-Raf-P (reaction 31). This complex then phosphorylates MAPKK twice (reactions 3 and 4). Finally MAPKK-PP phosphorylates MAPK twice and the doubly phosphorylated MAPK-PP is the end result of the insulin signal included in this model. In the cell this protein kinase phosphorylates various transcription factors that eventually enter the cell nucleus and controls transcription of specific genes. There are two negative feedbacks from MAPK-PP, the first phosphorylates Sos to Sos-P, as was explained above, and the second is due to an extra phosphorylation of the active enzyme RAF-P into the inactive form RAF-PP (reaction loop). There is also known to be a positive feedback from MAPK to Ras, but this is not included in the model. The Bhalla and Iyenger model was again chosen in preference to others, because it was believed to be most compatible with the previous steps. This last step in the model is shown in Figure 8.5 and the corresponding rate equations are given in equations (8.31)-(8.44). Note that MAPK and MAPKK

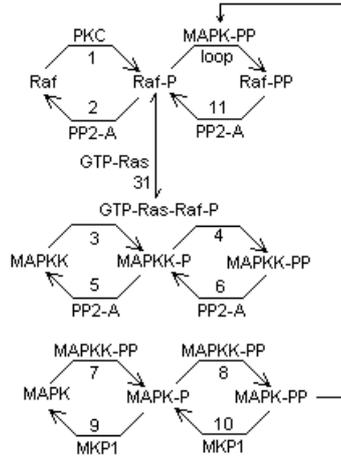


Figure 8.5 Schematic view of the MAPK cascade used in the insulin mitogenic pathway. *GTPRas* (from the previous step) binds to *Raf-P* which then phosphorylates *MAPKK* to *MAPKK-P* and *MAPKK-PP*. *MAPKK-PP* then phosphorylates *MAPK* and *MAPK-P*. *Raf-P* can be phosphorylated by *MAPK-PP* which makes a negative feedback.

are shortened to *K* and *KK*, respectively, in the rate equations.

$$vf_{31} = Kf_{31}[GTP - Ras][Raf - P] \quad (8.31)$$

$$vb_{31} = Kb_{31}[GTP - Ras - Raf - P] \quad (8.32)$$

$$v_1 = \frac{V_1[PKC][Raf]}{[Raf] + k_1} \quad (8.33)$$

$$v_2 = \frac{V_2[PP2 - A][Raf - P]}{[Raf - P] + k_2} \quad (8.34)$$

$$v_3 = \frac{V_3[GTP - Ras - Raf - P][KK]}{[KK] + k_3} \quad (8.35)$$

$$v_4 = \frac{V_4[GTP - Ras - Raf - P][KKP]}{[KKP] + k_4} \quad (8.36)$$

$$v_5 = \frac{V_5[PP2 - A][KKP]}{[KKP] + k_5} \quad (8.37)$$

$$v_6 = \frac{V_6[PP2 - A][KKPP]}{[KKPP] + k_6} \quad (8.38)$$

$$v_7 = \frac{V_7[KKPP][K]}{[K] + k_7} \quad (8.39)$$

$$v_8 = \frac{V_8[\text{KKPP}][\text{KP}]}{[\text{KP}] + k_8} \quad (8.40)$$

$$v_9 = \frac{V_9[\text{MKP1}][\text{KP}]}{[\text{KP}] + k_9} \quad (8.41)$$

$$v_{10} = \frac{V_{10}[\text{MKP1}][\text{KPP}]}{[\text{KPP}] + k_{10}} \quad (8.42)$$

$$v_{\text{loop}} = \frac{V_{\text{loop}}[\text{KPP}][\text{Raf} - \text{P}]}{[\text{Raf} - \text{P}] + k_{\text{loop}}} \quad (8.43)$$

$$v_{11} = \frac{V_{11}[\text{PP2} - \text{A}][\text{Raf} - \text{PP}]}{[\text{Raf} - \text{PP}] + k_{11}} \quad (8.44)$$

8.3 In silico experiments with the model

The analysis of the model has been done in four ways: i) Time-course analysis, ii) Dose-response analysis iii) Search for oscillations, and iv) Sensitivity analysis. i), ii) and iii) are presented in this section and the sensitivity analysis is presented together with the parameter estimation in Section 8.5.

8.3.1 Time-course analysis

Time-courses for the response of three different signalling molecules to a step increase in the extra-cellular insulin concentration from 0.1 to 0.01 μM are shown in Figure 8.6. The response for a signalling molecule in the beginning (IR-ins-P), in the middle (GTPRas-RafP), and for the last step included in the pathway (MAPK-PP) was analysed. As can be seen the major part of the response (10 to 90 %), takes 5-10 minutes. This can be considered a realistic response time compared with what has been found in rat fat cells ((Strålfors and Honnor, 1989), Figure 6). Moreover, the time-courses for the initial steps of the pathway are completed before the ones at the end. There are, however, as yet no experimental data available for fat cells and a detailed verification of the model is therefore not possible. A human fat cell is approximately 100 μm in diameter, and less than 2% of this is accounted for by cytosol (Thom, Stenkub, Karlsson, Önnegren, Nyström, Gustavsson and Strålfors, 2003), and this corresponds to a cytosolic volume of less than 30 000(μm)³ per fat cell. Sub-nanomolar concentrations therefore correspond to less than 20 000 particles per cell, which means that a concentration of 1E-7 μM (for MAPK-PP) or 1E-9 μM (for RAF-PP) represents about 10 or < 1 particle, respectively, per cell. This is not realistic in a situation with maximal insulin stimulation and to fix this is one of two objectives with the parameter estimation done below. The other objective concerns dose-response curves.

8.3.2 Dose-response curves

Dose-response curves showing how the stationary phosphorylation degree varies with the extracellular insulin concentration are shown in Figures 8.7, 8.8 and 8.9.

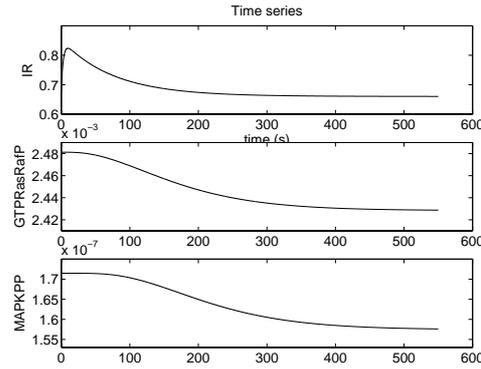


Figure 8.6 Time-series showing the step-response of various signalling molecules to insulin. The insulin step is from 0.1 to $0.01 \mu\text{M}$ and all the other concentrations are given in μM . The response time (10 - 90 %) is approximately 5-10 minutes and this is in accordance with experimental experience. Some of the concentrations, e.g. the MAPK-PP concentration, is lower than realistic. To fix this is one of the objectives of the parameter estimation done in Section 8.5.

The phosphorylation degree is presented in the same way as the corresponding experimental data is. That means that the phosphorylation signal has been normalised to be 100% at the highest stimuli ($\text{ins} \approx 0.1 \mu\text{M}$) and 0% at the lowest stimuli ($\text{ins} \approx 1\text{E-}5 \mu\text{M}$). The actual phosphorylation signal was measured in slightly different ways for the different enzymes.

- The MAPK phosphorylation degree signal is proportional to the concentration of MAPK-PP. The concentration of MAPK-P is not measured.
- The IR phosphorylation degree signal is proportional to the number of phosphorylation groups sitting on IR-molecules. This means that e.g. double phosphorylation of an individual IR-molecule contributes the same amount as two single phosphorylated IR-molecules.
- The IRS phosphorylation degree signal is measured in the same way as the IR signal.

Since the model has the MAPK-PP concentration explicitly included the steady-state concentration of this variable is directly comparable to the experimental signal. The model does not describe double phosphorylation and there is therefore no variable that directly corresponds to the experimental signals for IR and IRS. Here it has been assumed that double phosphorylation happens with the same velocity as the single phosphorylation. It has also been assumed that there is no difference

Variable	Steady state conc. original parameters	Steady state conc estimated parameters
KKP	0.00160429	0.043088
KP	8.39989E-5	0.1118420
KKPP	3.029892E-5	0.02185080
KPP	1.7148E-7	0.175226
SOSP	9.383E-6	0.041455
SOSP-Grb2	1.2764E-5	0.0580822
SOS-Grb2	0.01919	8.499412E-5
GTP-Ras	0.01903	0.063738
GAP	0.001906	0.00190643
GEF-P	0.00118	0.0011840
Raf-P	0.002715	0.013679811
Raf-PP	2.1155E-9	0.0218001
GTP-Ras-Raf-P-c	0.0024812	0.0418525
IR-ins	5.3996E-5	5.399612E-5
IR-ins-P	0.6749	0.6749583
IR-2ins-P	0.2024	0.2024875
IRS-P	0.694927	0.7443392
IRS-PSosGrb2	0.066678	3.163096E-4

Table 8.1 *The steady state values for $ins = 0.1 \mu\text{M}$ for the model with literature parameter values (first column), and with values after the estimation (second column). All concentrations are given in μM , which means that before the estimation some of the steady state values are unrealistically low. As can be seen, after the parameter estimation no steady state value has this problem.*

between finding phosphorylation groups on IR-ins-P and IR-2ins-P. With these assumptions the model equivalence to the experimental signals is given by (IRS-P) and (IR-ins-P + IR-2ins-P). The actual dose response curves has been obtained through the continuation methods in the Janet software (Janet, 2002), and the parameter settings used are described in the appendix.

8.3.3 Search for experimentally obtainable oscillations

Background

In some of the sub-models oscillations has been observed. This has, however, only been done for unrealistically low parameter values of the kloop parameter, appearing in equation (8.43). There might, however, be possible to have oscillations for realistic kloop values by changing other model parameters. If that is possible that would mean that the Hopf bifurcation must also have moved past the limit of realistic kloop values. That means that bifurcation analysis might be an efficient

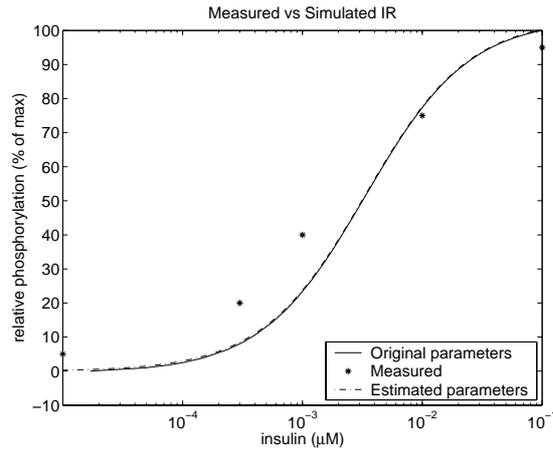


Figure 8.7 Comparison between the experimental and the simulated dose-response curves for IR. The experimental uncertainty is approximately 30%, and the agreement is within the measurement uncertainty both with the original and with the estimated parameters.

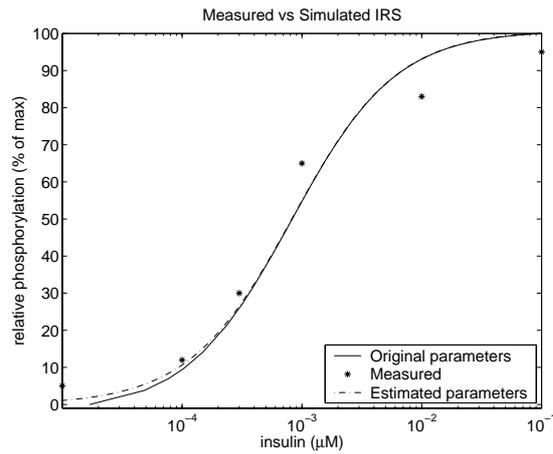


Figure 8.8 Comparison between the experimental and the simulated dose-response curves for IRS. The experimental uncertainty is approximately 30%, and the agreement is within the measurement uncertainty both with the original and with the estimated parameters.

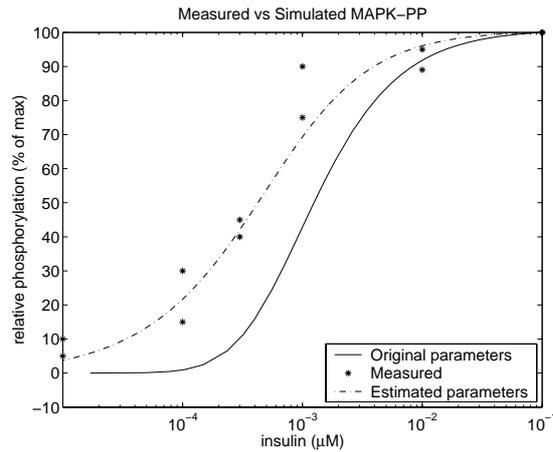


Figure 8.9 Comparison between the experimental and the simulated dose-response curves for MAPK-PP. The experimental uncertainty is about 30%. As can be seen the inflexion point with the original parameters is $5\text{-}10E\text{-}4 \mu\text{M}$ to the right of the measured inflexion point. After the parameter estimation the model inflexion point has been moved $8E\text{-}4 \mu\text{M}$ to the left.

tool to investigate the possibility of this. Should the model investigations show that there are, at least in the model, experimentally obtainable oscillations, this statement can later be checked experimentally. That possibility of experimental comparisons around the Hopf bifurcation would also make the model an excellent test model for the techniques developed in this thesis. The search included two parts: search for oscillating areas and bifurcation analysis.

Search for oscillating regions

The previously observed oscillations have only been seen in some of the sub-models that make up parts of the model presented here. This is natural since there are no previous models describing the full cascade from insulin to MAPK-PP. An immediate question was therefore whether there were oscillations for corresponding parameter values of kloop for the full model. The answer is that there are not. A search was made, also in neighboring parameter areas, but initially no oscillating regions at all were found for the full model. This search was however not exhaustive. The main effort, however, was done on searching in the sub-model for which the oscillations had been previously observed. There the previously observed oscillations were reproduced and a search for other regions was done. It gave, however, not rise to any discovery of new regions not connected with the first region for

which the system oscillated. The bifurcation analysis was therefore only conducted for one area, the one found by lowering the `kloop` parameter.

Bifurcation analysis

The idea was that by varying some of the parameters that can be varied experimentally the value of the parameter `kloop`, which cannot be varied experimentally, for which the bifurcation occurs could be brought back to its realistic value. That means that a bifurcation analysis showing how the Hopf bifurcation manifold lies in the parameter space is desired. Here this analysis was restricted to only varying one parameter at a time, i.e. restricted to a number of two-dimensional bifurcation plots. Such a plot is shown in Figure 8.10, where the bifurcation curve in the `kloop` – `Ktot` plane is plotted. `Ktot` is the total concentration of all the variables in the final level of the cascade, i.e. the sum of `K`, `K – P`, and `K – PP`, and is maybe the most reasonable candidate for allowing higher `kloop` values. As can be seen in the plot the bifurcation curve moves to higher values of `kloop` as `Ktot` increases. The intuition was therefore right in this case, and the remaining question was therefore if the effect would be big enough to bring the oscillations to experimentally obtainable regions. What was needed was for the `Ktot` parameter to move the Hopf bifurcation two orders of magnitude in the `kloop` parameter, but as can be seen in Figure 8.10, `Ktot` can only move the bifurcation point with a few per cent in `kloop`. Therefore the intuition was right, but the quantitative affect was too low. The same feature was found for all other parameters.

Discussion

The conclusions of the previous section was that no experimentally obtainable oscillating regimes, within the time-scale searched for here, were found. They have also never been observed experimentally so the surprise would have been biggest if there would have been found such regions and especially if these regions would have been experimentally verified. Since the parameter space is so huge it has not been possible to make a complete search. There might therefore still be such regions, they just have to be among those regions that have not yet been searched. From a biological point of view it is also fairly reasonable that this particular pathway does not oscillate. The reason for this is that the pathway is regulating e.g. the cell cycle (cell division), and it would be strange if such a decision to divide would oscillate between two states. This result, however, means that another model will have to be used for the final tests of the new methods presented in the previous chapters of this report. This model will be a neutrophil model and it is presented in the beginning of the next chapter.

8.4 Experimental verification

As was described in the previous section the simulated dose-response curves has experimental correspondences for the three proteins IR, IRS and MAPK. This

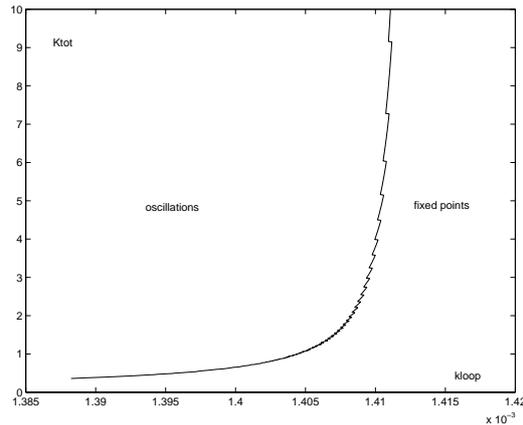


Figure 8.10 Showing how the bifurcation line lies in the $kloop - kTot$ plane for the sub-model of the fat cell model that had shown oscillations. The idea was that an increase in $kTot$ would allow $kloop$ to take higher and more realistic values. The qualitative effect is correct but it is too small. Realistic $kloop$ values lie at least one order of magnitude above the ones with observed oscillations and as can be seen in the plot $kTot$ can only alter the bifurcation point with a few percent. Therefore variations in $kTot$ alone is not sufficient to find oscillations for realistic $kloop$ values.

data is taken from fat cells and was originally published in ((Danielsson, Karlsson, Lystedt, Kjolhede, Gustavsson, Nystrom and Strålfors, 2004)). The experimental data is plotted in the same graphs as the simulated ones (Figure 8.7 to 8.9) and as can be seen the agreement is, considering the uncertainties in the model, quite good. The standard deviation of each individual sample is approximately 30% and therefore the dose-responses of IR and IRS are within the experimental uncertainty. This is almost true for the MAPK curve as well, but as can be seen in Figure 8.9 the inflexion point of the experimental samples lies approximately $5-10E-4\mu M$ to the left of the simulated inflexion point. This disagreement, together with the observation of the unrealistically low concentrations in the previous sections, is the basis for the parameter estimation presented in the next section.

8.5 Parameter estimation

Since there is too little information in the available experimental data to extract information about all parameters a full parameter estimation approach is not necessary. If one can find, in one way or another, a parameter combination that matches

the given demands this is as good as any formalized approach ever can do, given this amount of data. The chosen objectives are the following:

- To move the inflexion point of the MAPK-PP dose response curve 5-10 E-4 μM to the left.
- To keep the inflexion points of the other curves intact.
- To not have any steady-state concentrations less than 1E-5 μM , of activated enzymes (like MAPK-PP) for a high insulin stimulus ($\approx 0.1\mu\text{M}$).
- To not change any individual parameter more than 50% from its literature value.
- To not change more than 10 parameters.

The last two objectives means that we seek to keep as many parameters as close to the literature values as possible. A change with 50% of any parameter is, with todays uncertainty of the kinetic parameters, well within the experimental uncertainty range. There is also a cell-type and operating point dependency for these kinetic parameters, and therefore an adjustment of this kind to match a specific cell-type and a specific operating point is reasonable. The estimation was done as follows. First a sensitivity analysis with perturbation in all the parameters was made. All parameters were perturbed with 50% in both directions and for each perturbed state two things were measured: the steady-state concentration of MAPK-PP and the inflexion point of the MAPK dose response curve. The full sensitivity table is given in the appendix (Table A.2 and A.3), and the important parameters are extracted into Table 8.2. The parameters that have been extracted have been chosen by the following two criteria:

- A 50% perturbation gives rise to at least a -0.3E-4 μM displacement of the inflexion point in the MAPK-PP dose-response curve.
- The parameter is not involved in reactions that has a downstream effect on the IRS and IR curves.

The last criteria is chosen since it was desired not to change the dose-response curves for IR and IRS. From Table 8.2 can be seen that the direction that increases the MAPK-PP concentration is the same as the one that moves its inflexion point to the left. This gives even more confidence to our belief that these two objectives were good to include as a basis for the parameter estimation. The extracted parameters were the ones that were chosen for fitting, and in Table 8.2 their final values can be seen. As can be seen in Table 8.1 and in Figure 8.7, 8.8 and 8.9, the final model satisfies the given objectives. In Figure 8.11 is shown the same step-response time-series for the model with these parameters, as was shown in Figure 8.6 with the original parameters. As can be seen (see also Table 8.1), with these parameters there are no longer any phosphorylated concentrations less than 1E-3 μM . This has the effect that the negative feedbacks from MAPK-PP become more apparent, and

parameter	literature				estimated	
	value	$\frac{K_{PP}(+50\%)}{K_{PP}(0\%)}$	$\frac{K_{PP}(-50\%)}{K_{PP}(0\%)}$	$\Delta I_{KPP}(+50\%)$	$\Delta I_{KPP}(-50\%)$	value
V2	6.0	0.2091	13.4365	$0.1E-4$	$-0.4E-4$	3
PPP2 – A	0.224	0.041589	179.392	$0.1E-4$	$-1.2E-4$	0.112
k1	66.66667	0.2099	13.2984	$0.1E-4$	$-0.4E-4$	33.33
V25	10.0	0.2226	11.1461	$0.3E-4$	$-1.2E-4$	5
k25	1.0104	4.1003	0.0776	$-0.7E-4$	$0.6E-4$	1.5156
V23	0.02	4.1814	0.0791	$-0.5E-4$	$0.3E-4$	0.03
k23	0.50505	0.3216	5.2371	$0.1E-4$	$-0.3E-4$	0.2525

Table 8.2 *The parameters that were chosen for the parameter estimation. The full sensitivity table is given in Table A.2 and A.3. Both there and here the columns are ordered as follows (from left to right): parameter name and literature value, relative change in MAPK-PP steady-state concentration after a $\pm 50\%$ perturbation in this parameter, how many μM the inflexion point of the MAPK-PP dose-response curve has moved after the same perturbations, and the last column contains the final value after the parameter estimation.*

this can be seen e.g. in the damped oscillations in Figure 8.11. The final model is the main contribution of this chapter and now follows a discussion of its current status.

8.6 Discussion

Modelling intra-cellular pathways is an active research field that has just started. At present, lack of well-determined kinetic parameters, uncertainty in how and which intra-cellular signalling intermediates are interacting, and the lack of experimental high-throughput data, all contribute to the difficulty of constructing realistic mathematical models. Realistic models today therefore are highly over-parametrized, with respect to the available data. However, as can be seen in Figure 8.7 to 8.9 a model such as the one presented here can anyway, both before and after the parameter estimation, be agreeing with *in vivo* data. It must, however, be emphasized that this does not mean that any of the parameters, or even that any of the individual reactions, is individually correct, but is only a validation of the model behaviour considered as a whole. This agreement could have been obtained with many other combinations of parameters, and assumptions of the ingoing reactions. This agreement could also have been obtained with a much simpler model, containing e.g. only the measured variables. One could also have modelled the phosphorylation degree for each enzyme instead of considering the

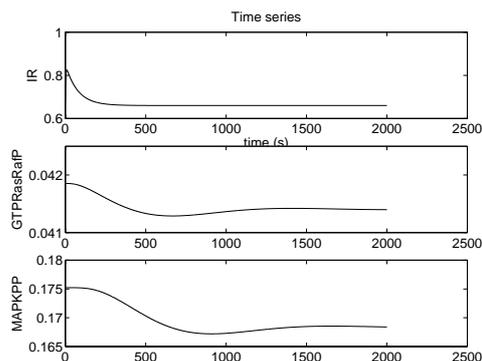


Figure 8.11 Simulations showing the step-response of various chemicals to insulin for the final model. The insulin step is from 0.1 to 0.01 μM and all the other concentrations are given in μM . The response time (10 - 90%) is still approximately 5-10 minutes but as can be seen there are now damped oscillations. These were not present with the original parameters, and they are probably due to that MAPK-PP, which governs the negative feedbacks, now no longer has an unrealistically small value.

different states of the same enzyme separately. Such input-output models have, however, the drawback that they do not include the existing mechanistic understanding of how the reactions take place, and they can therefore also not be used to test different assumptions regarding these mechanisms. These minimal models also has the drawback that they cannot include mechanistic information as it becomes experimentally available. A mechanistic model, like this one, can on the other hand be mapped to a minimal model, and that would be a natural continuation of this work. Regarding uncertainty of the various parts of the model, the steps after IRS and ending up in the MAPK-cascade is probably the one to be considered most uncertain. This is since this is the biggest part of the model for which there has been done no experimental comparison. It is also only here that an approximation of equating two different enzymes with each other has been necessary.

8.7 Conclusions

The main contribution of this chapter is a mathematical model that describes the mitogenic response to insulin. The input to the model is extra-cellular insulin and the model contains a connected pathway all the way to MAPK-PP, which is considered to be the last step before the transcription factors. The model agrees with the existing experimentally obtained dose-response curves, and has a time-scale of the response that is realistic. The final model has all parameters from

literature except those in Table 8.2, which have been changed by 50%. This has been done to further improve the dose-response agreement in Figure 8.9 and to not have any unrealistically small concentrations. It is assumed that the change in parameters has been done within their natural range of variations, and that the adjustments therefore does not make the model more unrealistic. As a help in choosing parameter adjustments a sensitivity analysis has been done (Tables A.2 and A.3), and this also serves as an important model analysis. The model is highly over-parametrized, given the current amount of data, and many parameters are uncertain. Nevertheless, here has been shown that it is possible to get a reasonable agreement with experimental observations. Some of the sub-models has shown oscillations, but a search, both with brute force and with bifurcation analysis, has failed to find any such regions for experimentally obtainable operating points. Therefore another model will be chosen as the large-scale example in the next chapter.

Utilizing the Newly Developed Techniques on a Full-scale Model

In this chapter the new methods developed in Chapters 3 to 6 will be applied to a full-scale model. The chosen model describes the dynamical behaviour of neutrophils and it was first presented in Olsen, Kummer, Kindzelskii and Petty (2003). The original model only contain two compartments and does therefore not capture the spatiality of the dynamics (see Figure 9.1). The chapter starts with a short introduction to neutrophils, and the special behaviour that the modelling work tries to capture. The original two-compartment model, along with some attempts to capture also the spatiality is presented in the next two sections. The methods of Chapter 3, describing how one can simplify a model like this to a two-dimensional description, are applied to the model in Section 9.4. An example where the qualitative test quantity of Chapter 6 is applied to the model is given in Section 9.5, and the new methods enabling improved parameter estimation techniques are tried out in Section 9.6. Finally, in Section 9.7, the conclusions of the chapter are given.

9.1 The spatio-dynamics of neutrophils

Neutrophils are a special kind of white blood cells, and they are part of the human immune system. That means that when an unwanted entity appears in the body, like e.g. a bacteria or a cancer cell, neutrophils go there to terminate it. When neutrophils change to termination mode, they change shape and become oval shaped. Another thing that happens during their activation is the onset of travelling waves, moving along the direction of the cell. The oscillatory change at each point in the cell probably is in form of concentration variations, and one of the chemicals that is oscillating is NADPH. NADPH has fluorescent features that

allows for time-continuous measurements of the concentration at various points in the cell. Experimental recordings of the wave phenomena are shown in Figure 9.1. There can be seen that there are many types of waves, some are starting at one end and travelling over to the other end, and some are starting in the middle, and going out in both directions. There have been observed a wide spectra of different observed wave-like phenomena for the neutrophil, and how and why the cell changes between them is not known. In some articles by Petty (see e.g. Kindzelskii and Petty (2002)) is a description of at least 10 different types of oscillations and waves, varying both in frequency and amplitude, and there is also described how the changes between these types of waves are correlated to changes in the functionality of the cell. Here a model based approach is helpful to help understand what is going on, and also to come with suggestions for how the transitions between the various states are made.

9.2 The Olsen model

The first presentation of a model that captures the dynamics of the behaviour was presented in Olsen et al. (2003), and it will henceforth be referred to as the Olsen model. This model shows oscillations, i.e. periodically time-varying concentrations, and its oscillations occur through a supercritical Hopf bifurcation. The model undergoes this bifurcation through parameter variations, and the four most important of these (the inflow of NADPH, the NADPH oxidase reaction, the presence of melatonin and the peroxidase oxidase reaction) have been confirmed experimentally. It is therefore possible to take the system to an operating point which is at a supercritical Hopf bifurcation, and also to measure a continuous signal (through the fluorescence of NADPH), and the settings described in Chapters 3 to 6 are therefore applicable.

The model consists of two compartments: the phagosome, where all superoxides used to kill the bacterias reside, and the cytosol, where e.g. NADPH resides. The reactions taking place in and between the various compartments are given in Table 9.1. These reactions, together with the rate expression, are sufficient to form the differential equations. Let the subscript *c* in $[\cdot]_c$ mean that the concentration is for a chemical in the cytosol, and the subscript *p* in $[\cdot]_p$ that the chemical is in the phagosome. With the R_i s defined in Table 9.1, the equations for the chemicals

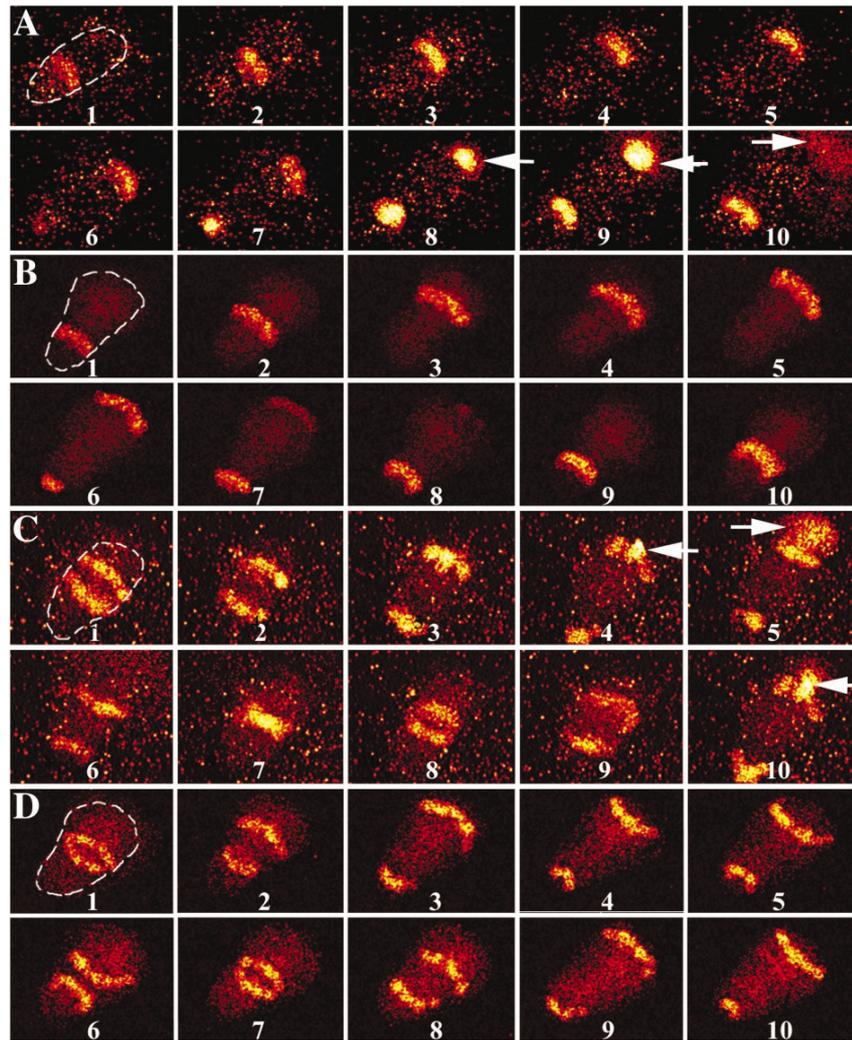


Figure 9.1 Figure showing the experimentally observed waves in the neutrophil. As can be seen the waves can be of very different character. Picture taken from Kindzelskii and Petty (2002).

in the phagosome are

$$\frac{d[\text{per}^{3+}]_p}{dt} = -R_1 + R_3 - R_4 \quad (9.1)$$

$$\frac{d[\text{CoI}]_p}{dt} = R_1 - R_2 + R_6 \quad (9.2)$$

$$\frac{d[\text{CoII}]_p}{dt} = R_2 - R_3 \quad (9.3)$$

$$\frac{d[\text{CoIII}]_p}{dt} = R_4 - R_6 \quad (9.4)$$

$$\frac{d[\text{H}_2\text{O}_2]_p}{dt} = -R_1 + R_5 - R_{15} \quad (9.5)$$

$$\frac{d[\text{O}_2^-]}{dt} = -R_4 - 2R_5 - R_6 - R_{18} + R_{19} \quad (9.6)$$

$$\frac{d[\text{O}_2]_p}{dt} = R_5 + R_6 - R_{14} - 2R_{19} \quad (9.7)$$

$$\frac{d[\text{MLTH}]_p}{dt} = -R_2 - R_3 - R_{16} \quad (9.8)$$

$$\frac{d[\text{MLT}]_p}{dt} = R_2 + R_3 - R_{17} \quad (9.9)$$

and the equations for the chemicals in the cytosol are

$$\frac{d[\text{NADPH}]_c}{dt} = R_{10} - R_{11} - R_{13} - R_{19} \quad (9.10)$$

$$\frac{d[\text{NADP}]_c}{dt} = -2R_9 - R_{12} + R_{13} \quad (9.11)$$

$$\frac{d[\text{H}_2\text{O}_2]_c}{dt} = R_7 + R_{11} + fR_{15} \quad (9.12)$$

$$\frac{d[\text{O}_2^-]_c}{dt} = -2R_7 + R_{12} + fR_{18} \quad (9.13)$$

$$\frac{d[\text{O}_2]_c}{dt} = R_7 + R_8 - R_{11} + R_{12} + fR_{14} \quad (9.14)$$

$$\frac{d[\text{MLTH}]_c}{dt} = R_{13} + fR_{16} \quad (9.15)$$

$$\frac{d[\text{MLT}]_c}{dt} = -R_{13} + fR_{17} \quad (9.16)$$

9.2.1 A bifurcation analysis

The original Olsen model shows oscillations of the kind shown in Figure 9.2. These oscillations disappear as the parameter k_{10} is lowered, and they disappear in a supercritical Hopf bifurcation. Sometimes it can be interesting to see cross-sections of the bifurcation manifold. Such figures are called bifurcation diagrams, and an

Table 9.1 *The reactions taking place in the original model Olsen et al. (2003), together with the corresponding reaction velocity expressions R_i*

Reaction	Rate expression (R_i)	Rate constant
Phagosomal reactions:		
1. $\text{H}_2\text{O}_2 + \text{per}^{3+} \rightleftharpoons \text{coI}$	$k_1[\text{H}_2\text{O}_2]_p - k_{-1}[\text{coI}]$	$k_1 = 5.0\text{E}7\text{M}^{-1}\text{s}^{-1}$ $k_{-1} = 58\text{s}^{-1}$
2. $\text{CoI} + \text{MLTH} \rightarrow \text{coII} + \text{MLT}$	$k_2[\text{coI}][\text{MLTH}]_p$	$k_2 = 5.0\text{E}7\text{M}^{-1}\text{s}^{-1}$
3. $\text{CoII} + \text{MLTH} \rightarrow \text{per}^{3+} + \text{MLT}$	$k_3[\text{coII}][\text{MLTH}]_p$	$k_3 = 4\text{E}3\text{M}^{-1}\text{s}^{-1}$
4. $\text{Per}^{3+} + \text{O}_2^- \rightarrow \text{coIII}$	$k_4[\text{per}^{3+}][\text{O}_2^-]$	$k_4 = 2\text{E}7\text{M}^{-1}\text{s}^{-1}$
5. $2\text{H}^+ + 2\text{O}_2^- \rightarrow \text{H}_2\text{O}_2 + \text{O}_2$	$k_5[\text{O}_2^-]_p^2$	$k_5 = 1\text{E}7\text{M}^{-1}\text{s}^{-1}$
6. $\text{CoIII} + \text{O}_2^- \rightarrow \text{coI} + \text{O}_2$	$k_6[\text{coIII}][\text{O}_2^-]_p$	$k_6 = 1\text{E}7\text{M}^{-1}\text{s}^{-1}$
Cytosolic reactions:		
7. $2\text{H}^+ + 2\text{O}_2^- \rightarrow \text{H}_2\text{O}_2 + \text{O}_2$	$k_7[\text{O}_2^-]_c^2$	$k_7 = 5\text{E}8\text{M}^{-1}\text{s}^{-1}$
8. $\rightleftharpoons \text{O}_2$	$k_8 - k_{-8}[\text{O}_2]_c$	$k_8 = 12.5\mu\text{M}\text{s}^{-1}$ $k_{-8} = 4.5\text{E}-2\text{s}^{-1}$
9. $2\text{NADP} \rightarrow$	$k_9[\text{NADP}]_c^2$	$k_9 = 6\text{E}7\text{M}^{-1}\text{s}^{-1}$
10. $\rightarrow \text{NADPH}$	k_{10}	$k_{10} = 22 - 35\mu\text{M}\text{s}^{-1}$
11. $\text{NADPH} + \text{O}_2 \rightarrow \text{H}_2\text{O}_2$	$k_{11}[\text{NADPH}]_c[\text{O}_2]_c$	$k_{11} = 1\text{M}^{-1}\text{s}^{-1}$
12. $\text{NADP} + \text{O}_2 \rightarrow \text{O}_2^-$	$k_{12}[\text{NADP}]_c[\text{O}_2]_c$	$k_{12} = 5\text{E}7\text{M}^{-1}\text{s}^{-1}$
13. $\text{MLT} + \text{NADPH} \rightarrow \text{NADP} + \text{MLTH}$	$k_{13}[\text{MLT}]_c[\text{NADPH}]_c$	$k_{13} = 1\text{E}7\text{M}^{-1}\text{s}^{-1}$
Diffusion terms:		
14. $\text{O}_{2,p} \rightleftharpoons \text{O}_{2,c}$	$k_{14}([\text{O}_2]_p - [\text{O}_2]_c)$	$k_{14} < 0.1$
15. $\text{H}_2\text{O}_{2,p} \rightleftharpoons \text{H}_2\text{O}_{2,c}$	$k_{15}([\text{H}_2\text{O}_2]_p - [\text{H}_2\text{O}_2]_c)$	$k_{15} = 30\text{s}^{-1}$
16. $\text{MLTH}_p \rightleftharpoons \text{MLTH}_c$	$k_{16}([\text{MLTH}]_p - [\text{MLTH}]_c)$	$k_{16} < 10\text{s}^{-1}$
17. $\text{MLT}_p \rightleftharpoons \text{MLT}_c$	$k_{17}([\text{MLT}]_p - [\text{MLT}]_c)$	$k_{17} < 10\text{s}^{-1}$
18. $\text{O}_{2,p}^- \rightleftharpoons \text{O}_{2,c}^-$	$k_{18}([\text{O}_2^-]_p - [\text{O}_2^-]_c)$	$k_{18} < 0.01\text{s}^{-1}$
NADPH oxidase:		
19. $\text{NADPH}_c + 2\text{O}_{2,p} \rightarrow \text{NADP}_c^+ + 2\text{O}_{2,p}^-$	$\frac{V_\alpha(1+\alpha)}{L+(1+\alpha)^2} \frac{[\text{O}_2]_p}{K_O + [\text{O}_2]_p}$ $\alpha = \frac{[\text{NADPH}]_p}{K_{\text{NADPH}}}$	$V = 288\mu\text{M}\text{s}^{-1}$ $L = 550$ $K_O = 1.5\mu\text{M}$ $K_{\text{NADPH}} = 60\mu\text{M}$

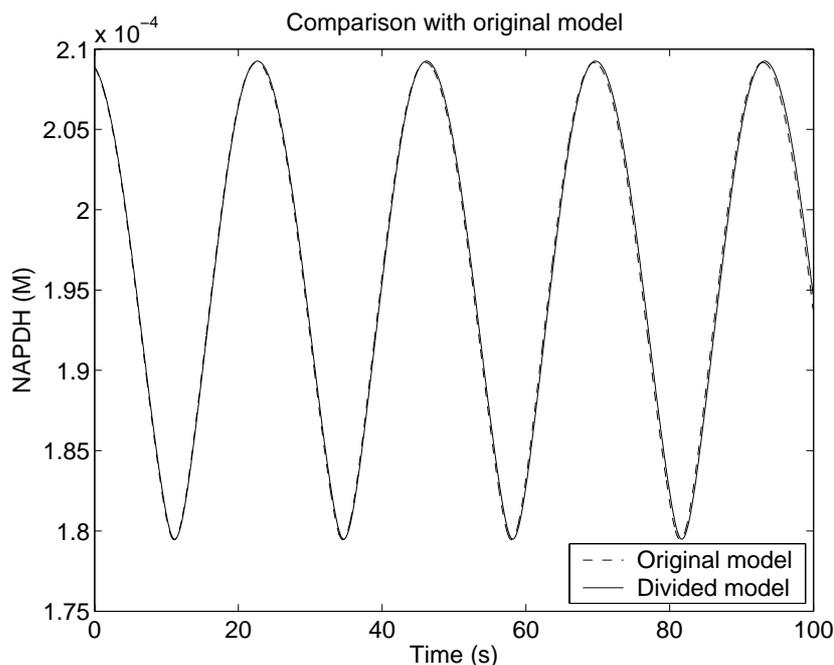


Figure 9.2 Time series showing the oscillations appearing in the Olsen model. There is also seen a comparison between one of the divided models and the original model, where the diffusion values in the divided model have been set to high values.

example is shown in Figure 9.3. Note that this is a figure of the same kind as Figure 2.4, where a bifurcation diagram for the Brusselator was plotted. For that system it was possible to obtain an analytical expression for it ($p_b = 1 + p_\lambda^2$), but here the curve has to be obtained numerically. For large-scale systems it is rarely the case that analytical solutions exist for these curves and it is therefore none of the new methods in this thesis requires an analytical solution of these curves. Figure 9.3 is just one of many bifurcation curves obtained for the system and for more plots from the analysis, see e.g. Karlsson (2003) or Claussen and Philipsen (2004). In Claussen and Philipsen (2004) evidence of multi-stability in the model, is also presented. This means that there are parallel attractors, existing for the same parameter values.

9.3 Compartmentalization

When capturing a phenomenon like the one in Figure 9.1 one needs to have not only the right chemicals involved in the right reactions in the model - they also have to

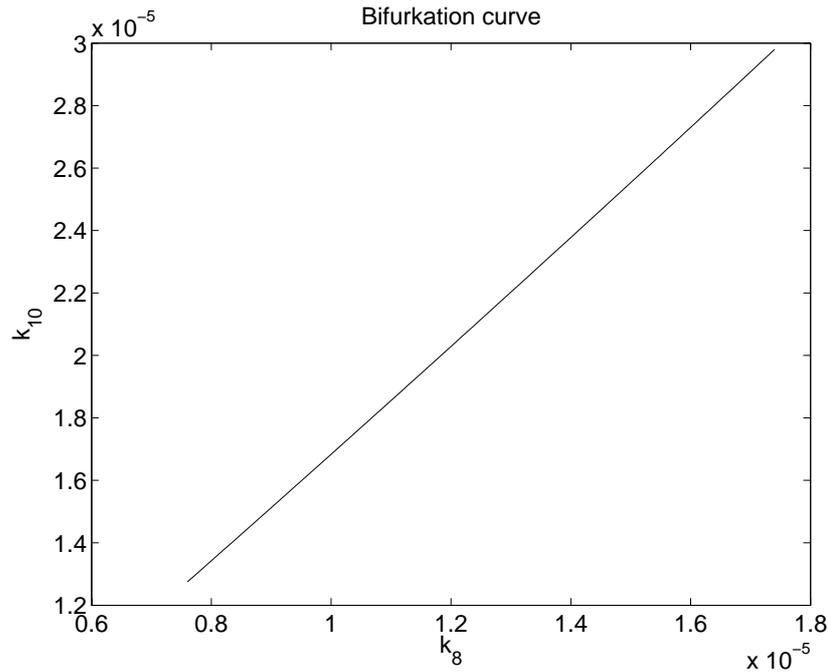


Figure 9.3 The bifurcation diagram for the neutrophil model and the two parameters k_{10} and k_8 . As can be seen the relation is a very smooth function.

be at the right *place*. The models have to supply this information, and if one does not wish to change the model into partial differential equations, an option is to divide the existing sub-cellular compartments into virtual sub-compartments. The term virtual, will henceforth be dropped though it should be clear that the new sub-compartments are not real cellular compartments, like mitochondria or cell nuclei. The divided model will, for infinitely fast diffusion, be identical to the original model. It might therefore turn out that the new model will keep all its good virtues, and only add the spatial knowledge. The division of an existing compartment into sub-compartments has two major consequences concerning the re-implementation of the original flows into, or out from, the non-divided compartments:

- The amount coming into or out from an original compartment is now divided between several sub-compartments. This causes all terms describing such in- and outflows to be changed by a multiplicative factor $1/p$, where p is the number of sub-compartments that are sharing the original flow. The same factor should of course be applied also if the inflow comes from several sub-compartments of the neighboring compartment that the flow comes from.

- The volume of the new sub-compartment is smaller, and hence the effect of an equal inflow, measured in number of particles, gives rise to a larger change, measured in the concentration of the chemical, since concentration is defined as $\text{amount}/\text{volume}$. It means that each term describing a flow in or out of the original compartment should be altered by a multiplicative factor of V_o/V_v , where V_o is the volume of the original compartment and V_v is the volume of the sub-compartment.

The differential equations describing the reactions taking place inside the original compartment is not affected by the division into sub-compartments. This follows since such reactions are multiplied by an equal volume fraction factor on both sides, and that they therefore cancel out. Division into sub-compartments might, however, give rise to so small concentrations that the approximation of continuous concentrations is a non-neglectable source of error. This is especially true if the concentrations of the chemicals in the original compartments are low. Between the sub-compartments there should be added some kind of transport. This is most easily modelled as diffusion, where the flow in a chemical between two compartments is proportional to the difference in the concentrations of this chemical for the two compartments. When the size and contact area of the sub-compartments are equal for all compartments and between all pairs of compartments, respectively, the proportionality constant will be equal between all sub-compartments. When this is not the case the proportionality constants have to be modified specifically for each compartment, but this is not the case in this thesis.

9.3.1 The compartmentalized models

One approach when trying to include the spatiality of the experimentally observed waves in the two-compartment model, is to subdivide the cytosol into sub-compartments according to the guidelines. The phagosome is not divided since the waves do not take place in this compartment, and since it is smaller. The division has been done under two different assumptions. In the first all chemicals have been present with a unique variable for all compartments. Then the new equations are the same as equations (9.1) to (9.16), with a one-dimensional diffusion operator on all the cytosolic chemicals. The system was solved through a one-dimensional fixed-grid division. For the second assumption, however, the new equations are more non-trivially obtained from the last. The second assumption is that all chemicals appearing in the cytosol have an infinitely fast diffusion, except NADP and NADPH. This means that all other chemicals will have the same value in all sub-compartments. Note the similarity between the assumption of infinitely fast pressure waves in the previous chapter, which also lead to a variable being shared between all compartments. To help writing up the modified equations we have introduced the following notation.

- n is number of sub-compartments of the original cytoplasm
- p is number of sub-compartments of the original cytoplasm that are neighbors to the phagosome

- The subscripts ${}_p$ and ${}_c$ means that the chemical belongs to the phagosome or the cytosol, respectively
- R_{19}^β is the NADPH oxidase between the phagosome and the compartment c_β
- R_{20} and R_{21} are the diffusion expressions between the new sub-compartments for NADPH and NADP, respectively
- $C_I = \{c_0, \dots, c_{n-1}\}$ is the set of all sub-compartments
- $C = \{i : c_i \in C_I\} = \{0, 1, 2, \dots, n-1\}$ are all the indices that corresponds to a sub-compartment
- P contain all the indices corresponding to a sub-compartment neighboring the phagosome
- G_i contain all the indices corresponding to a neighboring cytosolic sub-compartment to c_i
- the not mentioned rate reactions R_i are still defined in Table 9.1

With this notation, the new equations are as follows

$$\frac{d[\text{per}^{3+}]_p}{dt} = -R_1 + R_3 - R_4 \quad (9.17)$$

$$\frac{d[\text{CoI}]_p}{dt} = R_1 - R_2 + R_6 \quad (9.18)$$

$$\frac{d[\text{CoII}]_p}{dt} = R_2 - R_3 \quad (9.19)$$

$$\frac{d[\text{CoIII}]_p}{dt} = R_4 - R_6 \quad (9.20)$$

$$\frac{d[\text{H}_2\text{O}_2]_p}{dt} = -R_1 + R_5 - R_{15} \quad (9.21)$$

$$\begin{aligned} \frac{d[\text{O}_2^-]_p}{dt} &= -R_4 - 2R_5 - R_6 - R_{18} + \\ &+ \frac{2}{p} \sum_{\beta \in P} R_{19}^\beta \end{aligned} \quad (9.22)$$

$$\frac{d[\text{O}_2]_p}{dt} = R_5 + R_6 - R_{14} - \frac{2}{p} \sum_{\beta \in P} R_{19}^\beta \quad (9.23)$$

$$\frac{d[\text{MLTH}]_p}{dt} = -R_2 - R_3 - R_{16} \quad (9.24)$$

$$\frac{d[\text{MLT}]_p}{dt} = R_2 + R_3 - R_{17} \quad (9.25)$$

For the chemicals in the i :th compartment:

$$\begin{aligned} \frac{d[\text{NADPH}]_{c_i}}{dt} &= R_{10} - R_{11} - R_{13} - \frac{n}{p}R_{19}^i + \\ &+ \sum_{\beta \in G_i} R_{20}^\beta \end{aligned} \quad (9.26)$$

$$\begin{aligned} \frac{d[\text{NADP}]_{c_i}}{dt} &= -2R_9 - R_{12} + R_{13} + \\ &+ \sum_{\beta \in G_i} R_{21}^\beta \end{aligned} \quad (9.27)$$

and for the chemicals common to all cytosolic compartments:

$$\frac{d[\text{H}_2\text{O}_2]_c}{dt} = R_7 + \frac{1}{n} \sum_{\beta \in C} R_{11}^\beta + fR_{15} \quad (9.28)$$

$$\begin{aligned} \frac{d[\text{O}_2^-]_c}{dt} &= -2R_7 + \frac{1}{n} \sum_{\beta \in C} R_{12}^\beta + \\ &+ fR_{18} \end{aligned} \quad (9.29)$$

$$\begin{aligned} \frac{d[\text{O}_2]_c}{dt} &= R_7 + R_8 - \frac{1}{n} \sum_{\beta \in C} (R_{11}^\beta + R_{12}^\beta) + \\ &+ fR_{14} \end{aligned} \quad (9.30)$$

$$\frac{d[\text{MLTH}]_c}{dt} = \frac{1}{n} \sum_{\beta \in C} R_{13}^\beta + fR_{16} \quad (9.31)$$

$$\frac{d[\text{MLT}]_c}{dt} = -\frac{1}{n} \sum_{\beta \in C} R_{13}^\beta + fR_{17} \quad (9.32)$$

9.3.2 Simulated waves

There has been done an analysis of the new models with the two assumptions for varying values of the diffusion constants. For a detailed analysis of the results, see Karlsson (2003) or Claussen and Philipsen (2004). The general features observed were the following.

With the assumption of infinitely fast diffusion for all cytosolic chemicals, except NADP and NADPH, the model exhibited waves of the kind exemplified in Figure 9.4, and only for a limited range of diffusion values. For these diffusion values, there was a sufficient phase between the oscillations in the various compartments to imply waves, but there was also present a constant concentration gradient that was higher than the amplitude of the oscillations. Therefore, the only kind of wave that could be observed was the one kind of spreading-in-coming-out type of wave depicted in Figure 9.4. For higher diffusion values, the difference between the compartments gradually disappeared, and they eventually behaved

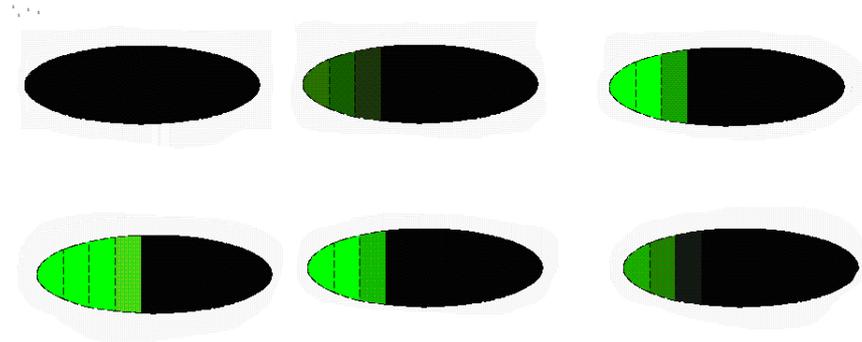
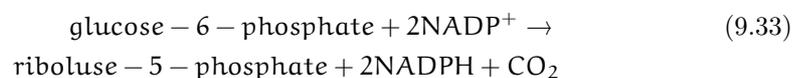


Figure 9.4 *The dynamics of the divided model. Here the cytosolic compartment of the original model has been divided into 9 sub-compartments. The pictures are oriented as the experimental ones of Figure 9.1. Chronologically the first is the top left, the second the top middle and the last is at the bottom to the right. As can be seen there is a wave spreading in from the left. That the most right compartments are all black is because they are closest to the phagosome, and the NADPH oxidase consumes there NADPH. A lower concentration of NADPH corresponds to a darker colour.*

like the two-compartment model (see Figure 9.2). For lower diffusion values, the above mentioned concentration gradients increased even more, and eventually the oscillations disappeared altogether.

With the above assumption dropped the situation was slightly different, and for a specific set of diffusion values there was observed travelling waves. These are depicted in Figure 9.5, and as can be seen the concentration gradient is observed also here. A plausible explanation for the gradient is the following. For the compartments that are neighboring the phagosome there is an additional outflow of NADPH through the NADPH oxidase. There is, however, no corresponding increased inflow for these compartments. The inflow, is believed to be due mainly to the hexose monophosphate shunt (Henderson and Chappel, 1996)



In the original Olsen model this has been modelled as a constant inflow (k_{10}) of NADPH, i.e. independently of the concentration of NADP^+ . A more realistic approximation would be to have it proportional to the concentration of NADP^+ , i.e. to replace k_{10} for

$$k_{10} \rightarrow k'_{10} * [\text{NADP}^+] \quad (9.34)$$

This inclusion would counteract the gradient, and is a natural candidate for future work.

9.4 Model simplification

In this section the theory of Chapter 3 is applied to the undivided neutrophil model. What is examined is how well the dynamics of the transformed two-dimensional system agree with that of the original system. The comparison is first done by comparing simulations and then with RMS analysis.

Analysis by comparing time series

The original Olsen model was brought to a Hopf bifurcation by changing the k_{10} and k_8 to $2.1\text{E-}5$ and $1.25\text{E-}5$, while having all other parameter values as in Table 9.1. Around this point the model was transformed into a one-dimensional complex equation (cf equation (3.58))

$$\dot{z}_1 = (i\omega_0 + \mu\sigma_1)z_1 + g_3|z_1|^2z_1 \quad (9.35)$$

where ω_0 , σ_1 and g_3 are determined by the transformation, and μ is $k_{10} - 2.1\text{E-}5$. The transformation was done with the formulas in Table 3.1. Equation (9.35)

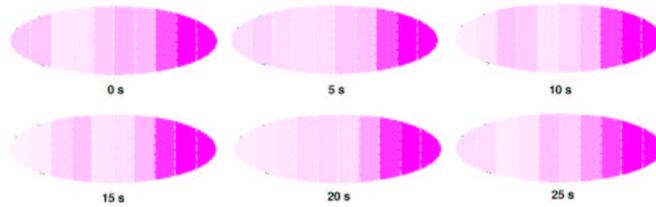


Figure 9.5 Simulated wave behaviour. As can be seen a concentration peak starts at the left and moves to the right, just as in some of the experimental recordings. However, the compartments neighboring the phagosome (at the right end), have a constantly higher lower concentration. A lower concentration of NADPH corresponds to a darker colour.

was then simulated and the result was transformed back through (cf equation (3.46))

$$x_z = x_b + z + h(z_1, \mu) \quad (9.36)$$

where the function h was also determined through Table 3.1. Then the outcome of this, x_z , was compared by simulating with x , which was obtained by simulating the Olsen model directly. Figures 9.6 and 9.7 show examples of such comparisons. In the first of these figures the initial values were chosen as $z(0) = 10^{-6}$, which corresponds to a perturbation in the original space of approximately 50% for most of the variables. As can be seen a perturbation of this size does, for most variables, not cause any difference between the simplified and the original model, that can be seen with mere inspection. This is quite impressive considering that x is 16-dimensional and z is 2-dimensional. However, in Figure 9.7 the initial values have been chosen to $z(0) = 4 * 10^{-6}$ which is four times as far away from the fix point as in the previous figure. Here can be seen that there is a clearly observable difference between the dynamical evolution of x and of x_z . That means that there is a fairly sharp limit where the difference goes from non-observable to clearly observable. Therefore, when using this simplification, e.g. in the optimization subproblems suggested in Chapter 5 (see e.g. equation (5.15)), one should be clear that the system has not been operating beyond this limit.

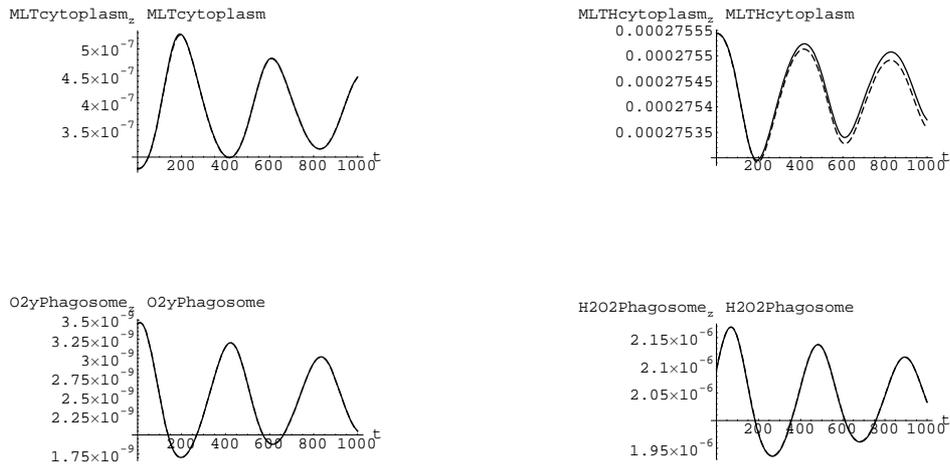


Figure 9.6 Time plots of the undivided Neutrophil model at the initial value $z(0) = 1 \cdot 10^{-6}$. The dashed line corresponds to x_z in (9.36) and the solid line to the directly solved system. As can be seen the agreement is quite good.



Figure 9.7 Time plots of the undivided Neutrophil model at the initial value $z(0) = 4 \cdot 10^{-6}$. The dashed line corresponds to x_z in (9.36) and the solid line to the directly solved system. As can be seen the agreement is not as good as that for $z(0) = 1 \cdot 10^{-6}$, shown in Figure 9.6.

RMS plots

Just as in the small examples of Chapter 3 the time course analysis will here be complemented with a RMS analysis. The definition of RMS is given in equation (3.102) and the result of the analysis is given in Figure 9.8. When studying such a plot it is interesting to see how the error from the transformation depends on the perturbation. What is especially interesting is to find ways to estimate how far from the bifurcation point the approximation can be used, given a certain maximum error. As can be seen the error increases as one leaves the bifurcation point (at $z = 0$), independently of in which direction the perturbation is made. However, the rate at which the error increases depends on the direction, and for the $(\text{Re}(z), \text{Im}(z)) = (1, -1)$ direction the error increases most rapidly. The relationship is smooth and that means that perturbations in 8 maximumly spread directions would have given a good approximation of the relation.

9.5 Forming a qualitative test quantity

Since the model has been shown to include both multi-stability and multiple attractors (Claussen and Philipsen, 2004), the conditions for the qualitative test quantity introduced in Chapter 6 are not fulfilled. We will, however, give an example of how one can utilize the basic ideas behind the test quantity. The exact relationship between the real NADPH concentration and the measured fluorescence is typically not known, and therefore the situation is really the one with an uncalibrated sensor y_D in equation (6.2). Since it is only possible to measure one chemical on-line, there are no other sensors in the system and

$$y = y_D = y_D(\text{NADPH}) \quad (9.37)$$

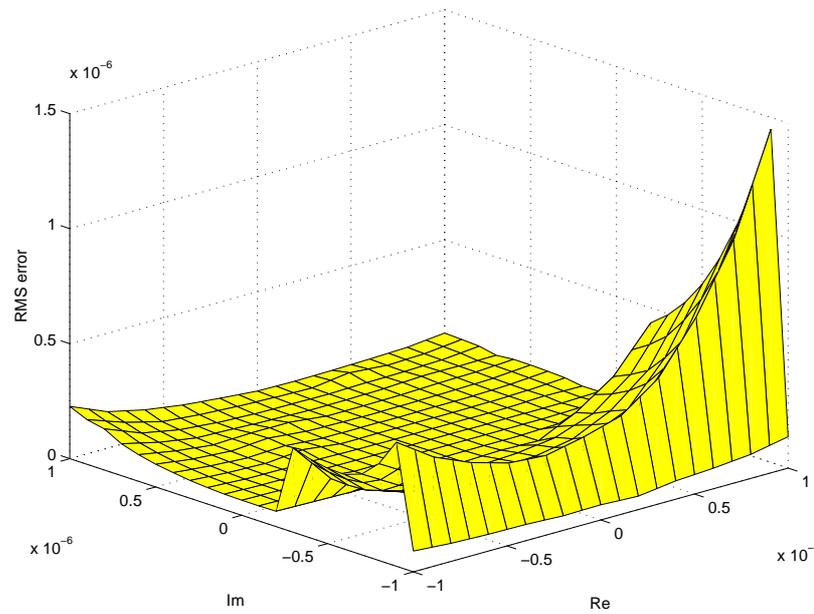


Figure 9.8 *RMS error plots of the Neutrophil model, viewing the agreement between the dynamics in the simplified and the original system, for different initial values. As can be seen the error is highly dependent on the direction in which the perturbation is made. However, the curve is fairly smooth, and perturbations in 8 maximally spread directions would have given a good approximation of the manifold.*

where the functional relationship y_D is not known, but continuous and well-behaved. In this example there is further a possibility to control the inflow of NADPH, i.e. the parameter k_{10} in Table 9.1 is known and controllable. That means that we have the typical situation of a bifurcation experiment. Let δ be the distance from the bifurcation point where the oscillations should be observable. Let u_b be an experimentally determined value, where the bifurcation point is. Since there are no other time-variations in the system we can simply form the test quantity, T_{qual} as

$$T_{\text{qual}} = \begin{cases} \text{true} & \text{if } k_{10} > u_b + \delta \text{ and } y_D \text{ shows a steady state} \\ \text{true} & \text{if } k_{10} < u_b \text{ and } y_D \text{ shows stationary oscillations} \\ \text{false} & \text{otherwise} \end{cases}$$

This test quantity will work if the system is kept so close to the bifurcation point that it will not encounter other bifurcations. The test quantity will detect changes

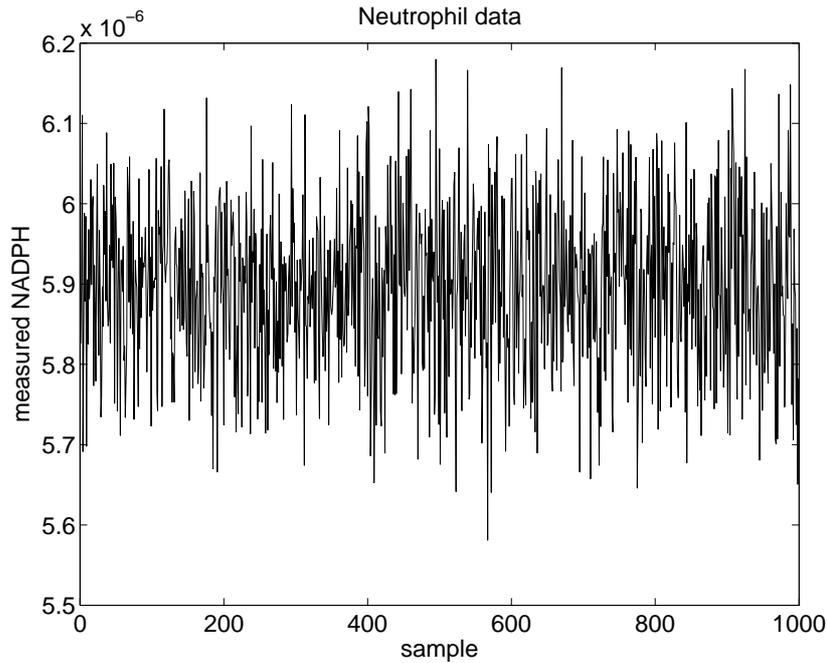


Figure 9.9 The data generated for the estimation example. It shows the [NADPH] concentration when the system is operating close to a Hopf bifurcation.

in the parameters, actuator faults, and oscillating faults in the y_D sensor. It will therefore have all the desired properties of the test quantity described in Section 6.1, and we have thus shown how simple the basic idea of Section 6.2 can sometimes be realized.

9.6 A parameter estimation example with simulated data

In Chapter 4 two methods for adding the extra constraint of a Hopf bifurcation were given. Here one of them, the elimination method, will be tried on the neutrophil model. The application is formulated as an example.

Example 9.1 Consider the following situation. Time series of NADPH can be collected for the neutrophil model, and the parameter k_{10} can be varied continuously, but its real value is not yet known. This is experimentally realistic, since one can not control the inflow directly, but only increase or decrease the activity of the hexose monophosphate shunt. Furthermore, all other parameters are assumed

known, except for the k_8 parameter, governing the inflow of oxygen. To estimate these two parameters, the k_{10} parameter was varied so that the system was operating precisely at the Hopf bifurcation point, and then a time series of NADPH was collected. In this example it is assumed that the real value of NADPH can be measured directly.

The given time series that was generated is plotted in Figure 9.9. The x -axis only specifies which sample value it is. In this example $p = (p_x, x_b, x_0) = (k_{10}, k_8, x_b, x_0)$ and in the elimination Algorithm 4.2, p_k , the eliminated parameter, is chosen as k_{10} . For the specific case of a stationary time series $x_b = x_0$ and therefore both x_0 and x_b are eliminated in the same way. The initial estimate of p_x is $(2.1E-5, 1.25E-5)$ and the initial estimate of $x_b (= x_0)$ is the state space values of the fix point that corresponds to this specific $\mu (= p_x)$ vector. The given start estimate is not at the bifurcation manifold, but it is close enough for the initial projection of the algorithm (step A in Algorithm 4.2) to consist only of the Newton iterations. Since p has been reduced to a 1-dimensional vector the error can simply be plotted as a graph, and this is done in Figure 9.10. There can be seen that the minimum is a very clearly defined point, and that there are no other local minima, and therefore the correct solution was easily obtained, for all chosen initial values. Without the elimination of k_{10} the search-space is two-dimensional, and the search becomes harder. Further, here it can converge to any point on a line including the correct estimate. In Section 4.5 the same situation appeared, even though then the original state space model was much smaller.

Finally some implementation details. The original neutrophil model contains two algebraic relationships, and hence the Jacobian has two constant zero eigenvalues. These gave difficulties, both when finding the inverse of the Jacobian, and when finding the critical eigenvalues, that are zero precisely at the bifurcation point. Therefore the system was rewritten to a 14 dimensional model, where per^{3+} and MLTH_c were obtained as auxiliary functions of the other variables. The relationships are

$$\begin{aligned} [\text{per}^{3+}]_p &= \text{CoTot} - ([\text{coI}]_p + [\text{coII}]_p + [\text{coIII}]_p) \\ [\text{MLTH}]_c &= \text{MLTHTot} - [\text{MLT}]_c - f * ([\text{MLTH}]_p + [\text{MLT}]_p) \end{aligned}$$

where CoTot and MLTHTot are two new parameters in the model describing the total Co-enzyme and the total MLTH concentration, respectively. The second modification made was that the constraint function c^{Hopf} was not obtained by multiplying all eigenvalues, like in expression (4.7), but rather by adding only the two smallest eigenvalues, that were also complex conjugated. That means that (4.7) was reformulated to

$$F^{\text{Hopf}}(x_b, \mu_b) := \lambda_i + \lambda_j$$

where

$$\lambda_i := \arg \min\{\lambda_j : \exists \lambda_k \text{ so } \bar{\lambda}_j = \lambda_k \text{ and } \text{Re}(\lambda_j) > 0\}$$

These adaptations were sufficient to make the algorithm work for this model.

Through this example we have seen that the elimination method, presented in

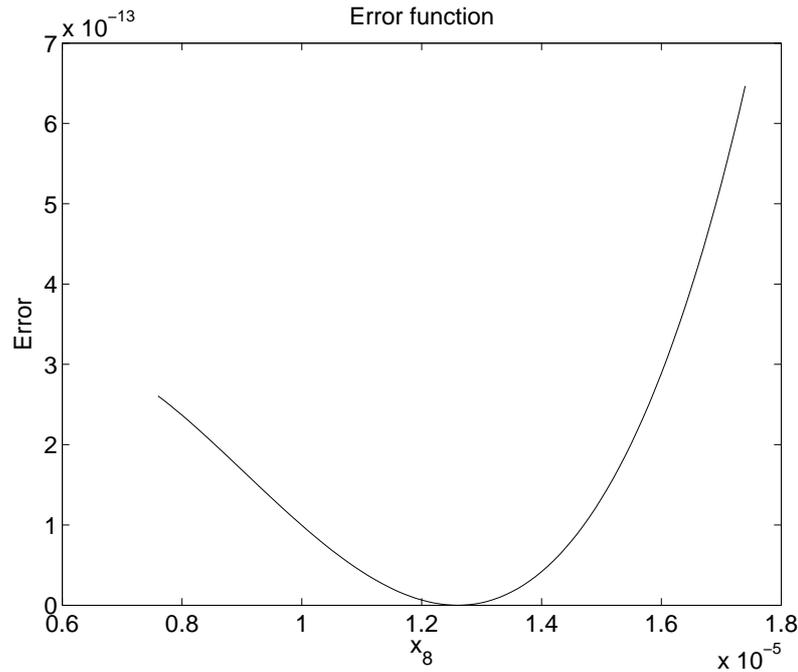


Figure 9.10 *The error as a function of the only freely varied parameter k_8 . As can be seen the error has a very nice and unique minima, and therefore when applying the methods of Chapter 4 and 5 to the simple example given here, the problem is very easy to solve.*

Chapter 4 can work, with small adaptations, also for a 16-dimensional system.

9.7 Conclusions

The conclusions of this chapter are as follows:

- A neutrophil model, with an experimentally obtainable Hopf bifurcation has been implemented and the new methods of this thesis have been applied to it.
- When doing the center manifold reductions derived in Chapter 3 the agreement is quite acceptable for initial values within $z(0) = 1E-6$, corresponding to 50% changes in the physical variables. Outside $z(0) = 4E-6$, which corresponds to 4 times the previous perturbation, the disagreement is large enough so that it can be seen by pure inspection (see Figures 9.6 and 9.7).

- Usually the only available sensor is the fluorescence of NADPH, which gives an unknown relationship to the real NADPH concentration. It has been shown how a qualitative test quantity of Chapter 6 easily can be formed with this sensor.
- A simple parameter estimation scenario for the 16-dimensional model, that clearly shows the benefit of the method, utilizing the bifurcation constraint c^{Hopf} introduced in Chapter 4, was given.

Summary and Conclusions

This thesis has been devoted to system identification of nonlinear thermochemical systems that contain dynamical instabilities. The results naturally divide in development of methodology and development of specific models.

Development of new identification methods

The new methods are utilizing that the system is known to contain a transition between steady states and stationary oscillations. It is also known that the transition between these states occurs through a supercritical Hopf bifurcation. There is a rich theory regarding these transitions, and the presented methods have utilized this theory for the problem of system identification. The model structure has been assumed to be a known nonlinear differential equation. It has further been assumed that some, but not necessarily all, of the dynamics is observable through the sensors, and that all the time-varying signals going into the system are known and controllable. With these assumption the unknown parameters, p , naturally divide in three parts

$$p = (p_x, x_0, p_y)$$

The first of these parameters, p_x , appear in the state equations. Here has been presented a method that reduces the degrees of freedom describing this parameter. The method can be used if one knows an input signal, that corresponds to a point on the bifurcation manifold. It has been shown how this knowledge can be used to form an extra constraint, c^{Hopf} , that is zero for all points on the manifold. Two different methods that can solve the new constrained estimation problem have been presented, and the advantage of adding the constraint has been shown on both a 2-dimensional and on a 16-dimensional model.

The second of the unknown parameters, x_0 , corresponds to the unknown initial state of the system. These parameters are often not interesting in themselves, but for systems with large state-spaces, this vector can make a problem unsolvable, even though the parameters one really seeks are few. Through a combined center manifold and normal form reduction this parameter vector has been described in minimal degrees freedom. We have shown experiment design that, independently of the dimension of the original system, force the minimal degrees of freedom to be 2 or 3. The x_0 parameters are calculated in a separate sub-problem for each step in the optimization algorithm. For systems starting in stationary oscillations the normal form reduction reveals the special structure of this sub-problem. Therefore it can be solved in a straight-forward manner, that does not have the problem of local minima, and that does not require any integration of the differential equations.

The last part of the parameter vector, p_y , corresponds to parameters appearing only in the sensor equations. It has been shown how the same bifurcation knowledge as used to simplify the p_x description can be utilized to validate estimated models. Under continuity assumptions it was demonstrated how this validation check can be done also with unknown p_y values. The validation check was formulated as test quantity which can be used in a diagnosis system, and it was demonstrated how it can function as an alternative to implementing the other developed estimation algorithms.

Development of new models

The first presented model is a zonal model for cylinder pressure, temperature and ionization currents. It is valid during the compression, combustion and expansion phase of a spark-ignited four-stroke Otto-engine. It has had as its main focus to model the relative position between the ionization curve and the pressure curve. A prediction of the position of the pressure curve relative to the top dead center is important information for the engine control algorithm. Today, there are rarely pressure sensors installed in production engines, and therefore using already available ionization sensing (used for e.g. knock and misfire detection) is an attractive alternative. It, however, requires further signal interpretation, and to do this in an optimal way physical models are required. Previous models describing the relationship between the pressure and the ionization positions have predicted the ionization peak location (IPL) > 2 crank angle degrees too late, in relation to the pressure peak location (PPL). Here has been shown that this problem disappears with the inclusion of a realistic heat-transfer in the model. Different heat transfer models has been compared, and it was demonstrated that the different heat transfer models have the same qualitative effect. It has also been shown that the difference is bigger between having and not having a heat transfer models, than between having different heat transfer models. It has also been shown that the combination of a geometrically based heat transfer model with a dynamical NO-model predicts the measured IPL within one crank angle degree.

The second presented model is one for the mitogenic response to insulin in fat cells. This is the first published model for this specific pathway, and the main result

is therefore the model structure itself. The presented model agrees with published data, and after a parameter adjustment the agreement is within measurement uncertainty, and with all steady-state concentrations in a realistic order of magnitude. A similar model to this, the Olsen model for activated neutrophils, was chosen for tests of the newly developed identification methods. The original model captured the oscillatory part of the dynamics only, and a formulation that compartmentalizes the model to include also the spatiality of the observed dynamics was made. Further, a bifurcation analysis revealed that the Olsen model had an experimentally obtainable Hopf bifurcation, and this was chosen as the operating point for the tests of the developed methods, where their performance were demonstrated.

A

Sufficient Information for Regeneration of all Plots

This appendix is here to enable the interested reader to regenerate all important plots of the thesis. If there should be any uncertainties in how the plots has been done, also after reading this appendix, the reader is very welcome to contact the author.

A.1 From Chapter 3

For all the plots in this chapter the simulations were done using a Runge Kutta pair integrator with variable step length, fourth and fifth order. The simulations were done in Mathematica, and the given integrator settings were as follows.

$$\text{PrecisionGoal} = 10 \quad (\text{A.1})$$

$$\text{AccuracyGoal} = 10 \quad (\text{A.2})$$

$$\text{Workingprecision} = 15 \quad (\text{A.3})$$

The initial values are given in terms of $z(0)$ under all figures. The bifurcation point, around which the system has been transformed is the same as in Example 3.4.

For Figure 3.2 and 3.3, the parameters has been chosen to lie in the bifurcation point.

Figure 3.4, on the other hand, has perturbed p_B to 1.8, while keeping p_A still at 1.

Figure 3.5 has also had the parameters at the bifurcation point, and the for this plot the initial values (measured in $z(0)$), are given as two of the axis in the plot. On the last axis (3.102) is plotted.

Also for Figure 3.6 the initial values are given on the axes. However, now the error in derivative (3.103), is on the last axis. In the last figure in the chapter.

Figure 3.7, the initial values are given in the figure text, the integrator is the same as for all other plots, the parameters has been kept at the bifurcation point. However, for the last plot, the expansion up to fifth order has been included. The coefficients for this transformation is given in Example 3.4.

A.2 From Chapter 4

Figure 4.1 has been obtained by applying Algorithm 4.3 on Sample 2 in Section 4.5. The initial estimate was $(p_A, p_B, x_{1,0}, x_{2,0}) = (0.8, 1.8, 1, 1.88)$, and the step length for the projection was 0.001. The accuracies were here fulfilled by doing 15 Newton iterations and checking for convergence problems.

For Figure 4.2 all the details have been included in the caption.

Figures 4.3, 4.4 and 4.5 were all obtained by solving the Sample 2 in Section 4.5, without adding the extra condition c^{Hopf} . Here a fix-step Runge-Kutta method, fourth order was imbedded in a trust-region method. The latter was included to ensure that the new iterations always implied sufficiently lowering in the objective function. Starting estimates were chosen in a small box $p_A \in (0.8, 2.2)$, $p_B \in (1.8, 2.2)$, and in a big box $p_A \in (0.5, 2)$, $p_B \in (1, 5)$ and the grid resolution was chosen as 8x8 for both the boxes. The initial time-step for the Runge-Kutta method was chosen to 0.01.

Figure 4.6 is just the objective function from Sample 2 plotted as a function of the variable p_A when continuing along the bifurcation curve.

A.3 From Chapter 5

The only interesting plots in this Chapter appears in the example. The example is a continuation of the study started in Chapter 4.

Figure 5.5 is showing the amplitude for different distances to the bifurcation point in the Brusselator model. The bifurcation point is the $(x_1, x_2, p_A, p_B) = (1, 2, 1, 2)$. The amplitude generated by the transformed system and the original system are compared. The former is generated through equation (5.17c), with α varied between 0 and 2π . Here no integration was necessary. The amplitude for the real system was generated by solving the Brusselator equations, and waiting out the transient. The transient was approximated to end after 10 000s. The distance,

p_μ , from the bifurcation manifold is in the increasing p_B direction. The numerical integration was done with a standard fix-step Runge-Kutta method. Time-step was set to 0.01.

Figure 5.6 shows the given time-series (denoted Sample 2) and the estimated, and transformed, system. The initial value for the latter was determined by $p_\mu = 0.01$, and $\alpha = 3.79$ rad.

A.4 From Chapter 6

Figures 6.2 and 6.3 are just after-treatments of the result from Section 4.5 and the necessary information for the original generation is given above and for the after-treatment in the text in Section 6.3.

A.5 From Chapter 7

Here a diagonally implicit Runge-Kutta method was used. The parameters used in this algorithm are the following.

$$\text{pert} = 1\text{E} - 12 \quad (\text{A.4})$$

$$\text{initialstep} = 0.00001 \quad (\text{A.5})$$

$$\text{maxstep} = 1 \quad (\text{A.6})$$

$$\text{tolerance} = 1\text{E} - 12 \quad (\text{A.7})$$

$$\text{Jacobianrecyclings} = 10 \quad (\text{A.8})$$

For Figure 7.3 the model with geometrically based heat transfer, but with no NO-dynamics was used. Ignition angle 27 CAD, and $\lambda = 1$. All other parameters are as given in Table 7.7.

Figures 7.8 and 7.9 is obtained with the exact same fitting as in the 7.3. The only thing that has been exchanged is the heat transfer model. All the given models were tested (also the one with no heat transfer), but the fit was only done for the model with geometrically interpretable parameters. No NO-dynamics was included.

Is the same as Figures 7.8 and 7.9, but with more data. Again the fit was done for the heat transfer model with geometrically interpretable geometry. The plotted operating point is the same as the second in Table 7.5. No NO-dynamics was included.

Figure 7.11 is given by Table 7.5 and 7.6. The model with geometrically based heat transfer, and no NO-dynamics was used all the time.

Figure 7.12 was obtained by varying the mass-limit deciding when a zone should be divided. Other settings like in Table 7.7. A geometrically based heat transfer,

with static NO-dynamics was used.

Figure 7.13 was produced by varying the velocity from 2.5m/s to 5 and 0 m/s. Other parameters like in Table 7.7. A geometrically based heat transfer, with static NO-dynamics was used.

Figure 7.14 was taken from the four operating points with varying ignition angle (see legend in figure). The model with dynamical NO was used. The data is taken from 7.10.

Figure 7.15 is the same as the previous figure, but with one randomly chosen cycle for each operating point.

A.5.1 Creation of a new zone

A zone is created when the mass passes the given parameter m_{max} . Then a certain fraction is taken from the unburned zone (parameter $\frac{V_{b,i+1}}{V_u}$), to create the new burned zone. The pressure of the new burned zone is the same, due to the assumption of pressure. The volume is decided as just described. The final thing, the temperature, is decided according to the following formula

$$h_{e,u}(T_u) = h_{e,b}(T_b) \quad (A.9)$$

where $h_{e,u}$ is the enthalpy of the unburned gas, and $h_{e,b}$ is the enthalpy of the burned gas. It should finally be noted that one does not set the number of zones explicitly, but only the division criteria. How many zones will be divided in practice depends on the development of the simulation. This is the reason for the ambiguity when saying that there has been used 8-10 zones.

A.6 From Chapter 8

This section contains a description of the fat cell model as well as sufficient information to regenerate all the plots and tables. Throughout Chapter 8, both for the ordinary simulations and for the dose-response curves obtained through continuation methods, the double order (4,5) Runge-Kutta method has been used as implemented in the Janet software (Janet, 2002). The parameters used in the integrator are: initial time step = 0.00001, absolute error = 1E-11 and relative error = 1E-7. For the numerical calculation of the ordinary and parameter Jacobian, which are needed in the continuation method, the perturbation size has been chosen to $1E-11 + 1E-7x$, for the ordinary Jacobian, and 1E-7 for the parameter Jacobian. Here x is the variable that has been perturbed. The initial values that has been used in all figures are given in Table 8.1, and this corresponds to a steady-state value for $ins = 0.1$. In the used implementation the dependent variables in Table A.1 has been given implicitly by the total concentrations, given in Table A.2 and A.3, and the initial values of the freely varied variables, given in Table 8.1.

The models rate equations

$$\begin{aligned}
\frac{d[\text{IR-ins}]}{dt} &= Kf_{12}[\text{IR}][\text{ins}] - Kb_{12}[\text{IR-ins}] \\
&\quad - K_{13}[\text{IR-ins}] \\
\frac{d[\text{IR-insP}]}{dt} &= K_{13}[\text{IR-ins}] + Kb_{15}[\text{IR-2insP}] \\
&\quad - Kf_{15}[\text{IR-insP}][\text{ins}] \\
&\quad - K_{14}[\text{IR-insP}]\text{PTP} \\
\frac{d[\text{IR-2insP}]}{dt} &= Kf_{15}[\text{IR-insP}][\text{ins}] \\
&\quad - Kb_{15}[\text{IR-2insP}] \\
\frac{d[\text{IRSP}]}{dt} &= K_{16}[\text{IRS}]([\text{IR-insP}] + [\text{IR-2insP}]) / \text{IRp} \\
&\quad - K_{17}\text{PTP}[\text{IRSP}] - Kf_{22}[\text{Sos-Grb2}][\text{IRSP}] \\
&\quad + Kb_{22}[\text{IRSP-Sos-Grb2}] \\
\frac{d[\text{SosP}]}{dt} &= Kb_{18}[\text{SosP-Grb2}] - Kf_{18}[\text{SosP}][\text{Grb2}] \\
&\quad - K_{19}[\text{SosP}] + \frac{V_{20}[\text{KPP}][\text{Sos}]}{[\text{Sos}] + k_{20}} \\
\frac{d[\text{Sos-Grb2}]}{dt} &= Kf_{21}[\text{Sos}][\text{Grb2}] - Kb_{21}[\text{Sos-Grb2}] \\
&\quad - Kf_{22}[\text{Sos-Grb2}][\text{IRSP}] \\
&\quad + Kb_{22}[\text{IRSP-Sos-Grb2}] \\
\frac{d[\text{SosP-Grb2}]}{dt} &= Kf_{18}[\text{SosP}][\text{Grb2}] - Kb_{18}[\text{SosP-Grb2}] \\
\frac{d[\text{IRSP-Sos-Grb2}]}{dt} &= Kf_{22}[\text{Sos-Grb2}][\text{IRSP}] \\
&\quad - Kb_{22}[\text{IRSP-Sos-Grb2}] \\
\frac{d[\text{GTPRas}]}{dt} &= \frac{V_{23}[\text{IRSP-Sos-Grb2}][\text{GDPRas}]}{[\text{GDPRas}] + k_{23}} \\
&\quad - \frac{V_{24}[\text{GEFP}][\text{GDPRas}]}{[\text{GDPRas}] + k_{24}} - \frac{V_{25}[\text{GAP}][\text{GTPRas}]}{[\text{GTPRas}] + k_{25}} \\
&\quad - K_{26}[\text{GTPRas}] - Kf_{31}[\text{RafP}][\text{GTPRas}] \\
&\quad + Kb_{31}[\text{GTPRas-RafP}] \\
\frac{d[\text{GAP}]}{dt} &= K_{27}[\text{GAPP}] - \frac{V_{28}\text{PKC}[\text{GAP}]}{[\text{GAP}] + k_{28}} \\
\frac{d[\text{GEFP}]}{dt} &= \frac{V_{30}\text{PKC}[\text{GEF}]}{[\text{GEF}] + k_{30}} - K_{29}[\text{GEFP}] \\
\frac{d[\text{GTPRas-RafP}]}{dt} &= Kf_{31}[\text{GTPRas}][\text{RafP}] \\
&\quad - Kb_{31}[\text{GTPRas-RafP}] \\
\frac{d[\text{RafP}]}{dt} &= \frac{V_1\text{PKC}[\text{Raf}]}{[\text{Raf}] + k_1} + \frac{V_{11}[\text{PP2-A}][\text{RafPP}]}{[\text{RafPP}] + k_{11}} \\
&\quad - \frac{V_2[\text{PP2-A}][\text{RafP}]}{[\text{RafP}] + k_2} - \frac{V_{10op}[\text{KPP}][\text{RafP}]}{[\text{RafP}] + k_{10op}} \\
&\quad - Kf_{31}[\text{RafP}][\text{GTPRas}] \\
&\quad + Kb_{31}[\text{GTPRas-RafP}] \\
\frac{d[\text{RafPP}]}{dt} &= \frac{V_{10op}[\text{KPP}][\text{RafP}]}{[\text{RafP}] + k_{10op}} \\
&\quad - \frac{V_{11}[\text{PP2-A}][\text{RafPP}]}{[\text{RafPP}] + k_{11}} \\
\frac{d[\text{KKP}]}{dt} &= \frac{V_3[\text{GTPRas-RafP}][\text{KK}]}{[\text{KK}] + k_3} + \frac{V_6[\text{PP2-A}][\text{KKPP}]}{[\text{KKPP}] + k_6} \\
&\quad - \frac{V_5[\text{PP2-A}][\text{KKP}]}{[\text{KKP}] + k_5} \\
&\quad - \frac{V_4[\text{GTPRas-RafP}][\text{KKP}]}{[\text{KKP}] + k_4} \\
\frac{d[\text{KKPP}]}{dt} &= \frac{V_4[\text{GTPRas-RafP}][\text{KKP}]}{[\text{KKP}] + k_4} \\
&\quad - \frac{V_6[\text{PP2-A}][\text{KKPP}]}{[\text{KKPP}] + k_6} \\
\frac{d[\text{KP}]}{dt} &= \frac{V_7[\text{KKPP}][\text{K}]}{[\text{K}] + k_7} + \frac{V_{10}[\text{MKP1}][\text{KPP}]}{[\text{KPP}] + k_{10}} \\
&\quad - \frac{V_9[\text{MKP1}][\text{KP}]}{[\text{KP}] + k_9} - \frac{V_8[\text{KKPP}][\text{KP}]}{[\text{KP}] + k_8} \\
\frac{d[\text{KPP}]}{dt} &= \frac{V_8[\text{KKPP}][\text{KP}]}{[\text{KP}] + k_8} - \frac{V_{10}[\text{MKP1}][\text{KPP}]}{[\text{KPP}] + k_{10}}
\end{aligned}$$

Auxiliary variables used in the model	
[IR]	= [IRtot] - [IR - ins] - [IR - insP] - [IR - 2insP]
[IRS]	= [IRStot] - [IRSP] - [IRSP - Sos - Grb2]
[Sos]	= [Sostot] - [Sos - Grb2] - [IRSP - Sos - Grb2] - [SosP] - [SosP - Grb2]
[Grb2]	= [Grb2tot] - [Sos - Grb2] - [IRSP - Sos - Grb2] - [SosP - Grb2]
[GDPRas]	= [GDPRastot] - [GTPRas] - [GTPRas - RafP]
[GAPP]	= [GAPtot] - [GAP]
[GEF]	= [GEFtot] - [GEFP]
[Raf]	= [Raftot] - [RafP] - [RafPP] [GTPRas - RafP]
[KK]	= [KKtot] - [KKP] - [KKPP]
[K]	= [Ktot] - [KP] - [KPP]

Table A.1 *These are the auxiliary variables used in the rate equations on the previous page. These auxiliaries are due to the presence of algebraic bonds on the dynamics. Another option is to introduce separate differential equations for each of auxiliary variables, and then have these above equations as algebraic constraints. This would lead to a zero index differential algebraic equation system, and since it is numerically more straightforward to solve an ordinary differential equation, this formulation was chosen.*

A.7 From Chapter 9

For the model simplification plots the following integrator settings has been used.

$$k_{12} = 2.105324 \cdot E^{-5} \quad (\text{A.10})$$

$$\text{PrecisionGoal} = 13 \quad (\text{A.11})$$

$$\text{AccuracyGoal} = 13 \quad (\text{A.12})$$

$$\text{Workingprecision} = 15 \quad (\text{A.13})$$

this together with the initial values given in $z(0)$ under each plot, as well as the parameters given in Table 9.1, should be sufficient to regenerate these plots.

Regarding the estimation example in Section 9.6 the used integrator is a diagonally implicit Runge-Kutta method, called SDIRK. It exists in the Janet software

parameter	literature						estimated
	value	$\frac{\text{KPP}(+50\%)}{\text{KPP}(0\%)}$	$\frac{\text{KPP}(-50\%)}{\text{KPP}(0\%)}$	$\Delta I_{\text{KPP}}(+50\%)$	$\Delta I_{\text{KPP}}(-50\%)$	value	
V_2	6.0	0.2091	13.4365	$0.1\text{E}-4$	$-0.4\text{E}-4$	3.0	
PPP2 – A	0.224	0.041589	179.392	$0.1\text{E}-4$	$-1.2\text{E}-4$	0.112	
k_2	15.6565	4.6429	0.0681	$-0.3\text{E}-4$	$0.1\text{E}-4$	–	
V_1	4.0	4.6444	0.0681	$-0.2\text{E}-4$	$0.1\text{E}-4$	–	
PKC	0.02	5.2013	0.0604	$-0.4\text{E}-4$	$0.3\text{E}-4$	–	
Rafttot	0.2	4.8552	0.065	$-0.2\text{E}-4$	$0.1\text{E}-4$	–	
k_1	66.66667	0.2099	13.2984	$0.1\text{E}-4$	$-0.4\text{E}-4$	33.33	
V_5	6.0	0.4489	3.8783	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
k_5	15.6565	2.2156	0.2538	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
V_3	0.105	2.2158	0.2538	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
KKtot	0.18	1.4039	0.4631	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
k_3	0.159091607	0.6589	1.6975	$0.0\text{E}-4$	$-0.0\text{E}-4$	–	
V_6	6.0	0.4447	3.9879	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
k_6	15.6565	2.247	0.2502	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
V_4	0.105	2.247	0.2502	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
k_4	0.159091607	0.4477	3.9097	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
V_9	1.0	0.667	1.9973	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
MKP1	0.0032	0.4447	3.9885	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
k_9	0.066667	1.498	0.5007	$-0.0\text{E}-4$	$0.0\text{E}-4$	–	
V_7	0.15	1.499	0.5003	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
Ktot	0.36	1.0394	0.8978	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
k_7	0.046296667	0.9462	1.0603	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
V_{10}	1.0	0.6668	1.9985	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
k_{10}	0.066667	1.4994	0.5002	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
V_8	0.15	1.4994	0.5002	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
k_8	0.046296667	0.6672	1.9949	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
V_{11}	6.0	1.0	1.0	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
k_{11}	15.6565	1.0	1.0	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
V_{loop}	10.0	1.0	1.0	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
k_{loop}	25.64	1.0	1.0	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
K_{b21}	0.0168	0.7901	1.2912	$1.5\text{E}-4$	$-1.9\text{E}-4$	–	
K_{f21}	0.025	1.1828	0.6346	$-1.2\text{E}-4$	$2.9\text{E}-4$	–	
Grb2tot	1.0	1.1958	0.5903	$-1.3\text{E}-4$	$3.2\text{E}-4$	–	
SOSTot	0.1	3.9235	0.0844	$-0.3\text{E}-4$	$0.1\text{E}-4$	–	
K_{b18}	0.0168	1.0001	0.9996	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
K_{f18}	0.025	0.9998	1.0002	$-0.1\text{E}-4$	$-0.0\text{E}-4$	–	
K_{19}	0.0010	1.0003	0.9992	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
V_{20}	10.0	0.9996	1.0004	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
k_{20}	2.5641	1.0003	0.9992	$-0.0\text{E}-4$	$-0.0\text{E}-4$	–	
K_{27}	0.1	0.9443	1.1806	$-0.0\text{E}-4$	$-0.1\text{E}-4$	–	
GAPtot	0.0020	0.2226	11.1461	$0.4\text{E}-4$	$-1.2\text{E}-4$	–	

Table A.2 The sensitivity analysis used in Section 8.5. The changed parameters are extracted in Table 8.2. Note that the last parameters are given in Table A.3

and the following settings has been used.

$$\text{pert} = 1\text{E} - 12 \quad (\text{A.14})$$

$$\text{initialstep} = 0.00001 \quad (\text{A.15})$$

$$\text{maxstep} = 1 \quad (\text{A.16})$$

$$\text{tolerance} = 1\text{E} - 12 \quad (\text{A.17})$$

$$\text{Jacobianrecyclings} = 10 \quad (\text{A.18})$$

the initial estimate, as well as initial values for the integration, is given in the example.

parameter	literature		estimated			
	value	$\frac{\text{KPP}(+50\%)}{\text{KPP}(0\%)}$	$\frac{\text{KPP}(-50\%)}{\text{KPP}(0\%)}$	$\Delta I_{\text{KPP}}(+50\%)$	$\Delta I_{\text{KPP}}(-50\%)$	value
V ₂₈	25.0	1.0877	0.9173	-0.1E-4	-0.0E-4	-
k ₂₈	101.87499	0.9443	1.1806	-0.0E-4	-0.1E-4	-
K ₂₉	0.1	0.979	1.0637	0.1E-4	-0.3E-4	-
V ₃₀	4.0	1.0317	0.9686	-0.2E-4	0.1E-4	-
GEF _{tot}	0.1	1.0322	0.9684	-0.2E-4	0.1E-4	-
k ₃₀	66.667	0.9791	1.0636	0.1E-4	-0.3E-4	-
K ₂₆	1.0E-4	0.9902	1.0099	-0.0E-4	-0.1E-4	-
V ₂₅	10.0	0.2226	11.1461	0.3E-4	-1.2E-4	5
k ₂₅	1.0104	4.1003	0.0776	-0.7E-4	0.6E-4	1.5156
V ₂₄	0.02	1.0323	0.9684	-0.2E-4	0.1E-4	-
GDPR _{astot}	0.2	3.1889	0.1041	0.1E-4	-0.3E-4	-
k ₂₄	0.50505	0.9829	1.0379	0.0E-4	-0.2E-4	-
V ₂₃	0.02	4.1814	0.0791	-0.5E-4	0.3E-4	0.03
k ₂₃	0.50505	0.3216	5.2371	0.1E-4	-0.3E-4	0.2525
Kb ₃₁	0.5	0.2055	14.1294	0.1E-4	-0.5E-4	-
Kf ₃₁	24.0	4.7655	0.0663	-0.2E-4	0.2E-4	-
Kb ₂₂	0.1	0.5833	1.8943	3.5E-4	-5.0E-4	-
Kf ₂₂	0.49998	1.5102	0.3654	-3.1E-4	6.2E-4	-
Kb ₁₂	0.0033333	1.0	1.0	-0.0E-4	-0.0E-4	-
Kf ₁₂	1.0	1.0025	0.9926	-4.1E-4	12.0E-4	-
IR _{tot}	0.9	1.1059	0.7515	-3.9E-4	10.2E-4	-
K ₁₃	41.666666	1.0	1.0	-0.0E-4	-0.0E-4	-
K ₁₄	0.0033333	0.9963	1.0037	6.0E-4	-6.2E-4	-
PTP	1.0	0.8598	1.1666	13.9E-4	-9.2E-4	-
Kb ₁₅	0.3333333	0.9994	1.0014	-0.1E-4	-0.0E-4	-
Kf ₁₅	1.0	1.0008	0.999	-0.0E-4	-0.1E-4	-
K ₁₇	0.0232662	0.8644	1.1643	5.3E-4	-5.9E-4	-
K ₁₆	0.06933333	1.1059	0.7515	-3.9E-4	10.2E-4	-
IR _{Stot}	1.0	1.5457	0.3324	-3.2E-4	6.0E-4	-
IR _p	0.897	0.8644	1.1643	5.3E-4	-5.9E-4	-

Table A.3 Continuation from Table A.2.

References

- Allgower, E. and Georg, K. (1990), *Numerical Continuation Methods*, number 3-540-12760-7 in 'Springer Series in Computational Mathematics', Springer-Verlag.
- Andersson, I. (2002), Cylinder Pressure and Current Modeling for Spark Ignited Engines, Thesis no. 962, Department of Electrical Engineering, Linköpings Universitet, Linköping, Sweden.
- Andersson, I. (2004), A comparison of combustion temperature models for ionization current modeling in an si engine. SAE paper no. 2004-01-1465.
- Auzins, J., Johansson, H. and Nytomt, J. (1995), Ion-gap sense in miss-fire detection, knock and engine control. SAE paper no. 950004.
- Bhalla, U. S. and Iyenger, R. (1999), 'Emergent properties of networks of biological signaling pathways', *Science* **283**(15), 381–387.
- Bhalla, U. S., Ram, P. T. and Iyenger, R. (2002), 'Map kinase phosphatase as a locus of flexibility in mitogen-activated protein kinase signaling network.', *Science* **297**(5583), 1018–23.
- Calcote, H. F. (1957), *Mechanisms for the Formation of Ions in Flames*, Vol. 1 of *Combustion and Flame*.
- Claesson, D. (2004), Experimental verification of a multi-zonal model for cylinder pressure, temperature and ionization currents, Master's thesis, Linköping University.

- Claussen, A. and Philipsen, K. R. (2004), Modelling and simulation of the neutrophil cell, Polytechnical midterm report, Technical university of Denmark, Lyngby, Denmark.
- Danielsson, A., Karlsson, M., Lystedt, E., Kjolhede, M., Gustavsson, J., Nystrom, F. H. and Strålfors, P. (2004), 'Insulin resistance in human adipocytes after surgery differs from that of diabetes', *Manuscript in preparation*.
- Doqcs (2001), 'Database of quantitative signalling pathways', <http://doqcs.ncbs.res.in>.
- Eriksson, L. (2004), Chemical engineering programming package.
- Eriksson, L., Nielsen, L. and Glavenius, M. (1997), Closed loop ignition control by ionization current interpretation.
- Eriksson, L., Nielsen, L. and Nytomt, J. (1996), Ignition control by ionization current interpretation.
- Evtushenko, Y. G. (1985), *Numerical Optimization Techniques*, number 0-387-90949-4 in 'Translation series in mathematics and engineering', Springer-Verlag, New York.
- Fichera, A., Losenno, C. and Pagano, A. (2001), 'Experimental analysis of thermo-acoustic combustion instability', *Applied Energy* **70**, 179–191.
- Franke, A. (2002), Thesis no. 962, Department of Electrical Engineering, Lunds Universitet, Lund, Sweden.
- Glaser, I. and Powell, J. D. (1981), Optimal closed-loop spark control of an automotive engine. SAE paper No. 810058.
- Gustafsson, M., Hörnquist, M. and Lombardi, A. (2003), Large-scale reverse engineering by the lasso, in 'Workshop on Cellular Networks'.
- Heinrich, R. and Schuster, S. (1996), *The Regulation of Cellular Systems*, number 0412032619, 1st edn, Kluwer Academic Publishers.
- Henderson, L. M. and Chappel, J. B. (1996), 'Nadph oxidase of neutrophils', *Biochem Biophys Acta* **1273**, 87–107.
- Heywood, J. B. (1988), *Internal Combustion Engine Fundamentals*, McGraw-Hill series in mechanical engineering, McGraw-Hill International Editions.
- Holstein-Rathlou, N.-H., Yip, K.-P., Sosnovtseva, O. V. and Mosekilde, E. (2001), 'Synchronization phenomena in nephron-nephron interaction', *Chaos* **11**(2), 417.
- Hubbard, M., Dobson, P. D. and Powell, J. D. (1976), 'Closed loop control of spark advance using cylinder pressure sensor', *Journal of Dynamic System, Measurement, and Control* pp. 414–420.

- Hynne, F., Danø, S. and Sørensen, P. G. (2001), 'Full-scale model of glycolysis in *saccharomyces cerevisiae*', *Biophys Chem* **94**(1-2), 121–63.
- Ipsen, M., Hynne, F. and Sørensen, P. G. (1998), 'Systematic derivation of amplitude equations and normal forms for dynamical systems', *Chaos* **8**(4), 834–851.
- Janet (2002), 'Java nonlinear exploration tools', <http://fys.dtu.dk/~janet>.
- Jeffries, C. and Wiesenfeld, K. (1985), 'Observation of noisy precursors of dynamical instabilities', *Physical Review A* **31**(2), 1077–1984.
- Karlsson, H. (2003), Analysis and spatial extension of a neutrophil model, Master's thesis LiTH-IFM-EX-1205-2003, Department of Physics and Measurement Technology, Biology and Chemistry Linköpings Universitet, Linköping, Sweden.
- Kindzelskii, A. and Petty, H. (2002), 'Apparent role of traveling metabolic waves in oxidant release by living neutrophils', *PNAS* **99**(14), 9207–9212.
- Leng, Y., Karlsson, H. K. and Zierath, J. (2004), 'Insulin signaling defects in type 2 diabetes', *Rev Endocr Metab Disord* **5**(2), 111–7.
- Ljung, L. (1999), *System Identification Theory for the User*, 2nd edn, Prentice-Hall, inc., Upper Saddle River, New Jersey 07458.
- Nguyen, H. T. T. (2004), Engineering of *Saccharomyces cerivisiae* for the production fo L-glycerol 3-phosphate, PhD thesis, Technical University of Berlin.
- Nielsen, L. and Eriksson, L. (2004), *Course Material Vehicular Systems*, Linköping Institute of Technology, Vehicular Systems, ISY.
- Nilsson, Y. and Eriksson, L. (2001), A new formulation of multi-zone combustion engine models, Karlsruhe, Germany, IFAC Workshop: Advances in Automotive Control.
- Nocedal, J. and Wright, S. J. (1999), *Numerical Optimization*, number 0-387-98793-2, Springer-Verlag, New York.
- Nogenmyr, K.-J. (2003), Development and analysis of zonal models for a combustion engine, Master's thesis, Linköping University.
- Olsen, L. F., Kummer, U., Kindzelskii, A. L. and Petty, H. R. (2003), 'A model of the oscillatory metabolism of activated neutrophils', *Biophys J* **84**(4), 69–81.
- Poincaré, J. H. (1952), *Science and Method*, Dover, New York.
- Reinmann, R., Saitzkoff, A. and Mauss, F. (1997), Local air-fuel raio measurements using the spark plug as an ionization sensor. SAE paper no. 970856.
- Råde, L. and Westergren, B. (1990), *BETA Mathematics Handbook*, number 91-44-25052-5, 2nd edn, Studentlitteratur.

- Saitzkoff, A., Reinmann, R., Berglind, T. and Glavmo, M. (1996), An ionization equilibrium analysis of the spark plug as an ionization sensor. SAE paper no. 960337.
- Sawamoto, K., Kawamura, Y., Kita, T. and Matshushita, K. (1987), Individual cylinder knock control by detecting cylinder pressure. SAE paper No. 871911.
- Schittkowski, K. (2002), *Numerical Data Fitting in Dynamical Systems - A Practical Introduction with Applications and Software*, number 1402010796, Kluwer Academic Publishers.
- Sedeghat, A., Sherman, A. and Quon, M. (2002), 'A mathematical model of metabolic insulin signaling pathways', *Am. J. Physiol. Endocrinol. Metab.* **283**, E1084–E1101.
- Sørensen, M. P., Petersen, M. V. and Aslanidi, O. (2004), Coherent dynamics of excitable coupled beta-cells, in A. D. F. Andris Buikis, Raimondas Ciegis, ed., 'Progress in Industrial Mathematics at ECMI 2002', Mathematics in Industry, Springer-Verlag, Berlin Heidelberg, pp. 375–379.
- Strogatz, S. H. (1994), *Nonlinear Dynamics and Chaos*, Addison-Wesley.
- Strålfors, P. and Honnor, R. C. (1989), 'Insulin-induced dephosphorylation of hormone-sensitive lipase', *Eur J Biochem* **182**, 379–385.
- Sturis, J. and Knudsen, C. (1996), 'Modelling seasonal predator-prey interactions', *J Theor Biol* **178**, 99–103.
- Sturis, J., Knudsen, C., O'Meara, N. M., Thomsen, J. S., Mosekilde, E., Vauter, E. V. and Polonsky, K. S. (1995), 'Phase-locking regions in a forced model of slow insulin and glucose concentrations', *CHAOS* **5**, 193–199.
- Thom, H., Stenkub, K. G., Karlsson, M., Önnegren, U., Nyström, F. H., Gustavsson, J. and Strålfors, P. (2003), 'Cell surface orifices of caveolae and localization of caveolin to the necks of caveolae in adipocytes', *Mol Biol Cell* **14**(10), 3967–3976.
- Wiggins, S. (1990), *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, number 2 in 'Texts in Applied Mathematics', 2rd edn, Springer-Verlag New York, Inc.
- Woschni, G. (1967), Universally applicable equation for the instantaneous heat transfer coefficient in the internal combustion engine. SAE paper 670931.
- Yoshiyama, S., Tomita, E. and Hamamoto, Y. (2000), Fundamental study on combustion diagnostics using a spark plug as ion probe. SAE paper no. 2000-01-2828.
- Zaida, S. and Graf, H. (1998), 'Feedback control of combustion oscillations', *Journal of Fluids and Structures* **12**, 491–507.

Zeldovich, Y. A., Sadvnikov, P. Y. and Frank-Kamenetskii, D. A. (1974), *Oxidation in Nitrogen in Combustion*, Academy of Sciences of USSR. Translated by M. Shelef 1970, Ford Motor Company.