Technische Universität München
Zentrum Mathematik

Monomial Dynamical and Control Systems over a Finite Field and Applications to Agent-based Models in Immunology

Edgar Wilfried Delgado-Eckert

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Peter Rentrop
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Rupert Lasser
	2.	Univ.-Prof. Dr. Bodo Pareigis, Ludwig-Maximilians-Universität München (schriftliche Beurteilung)
	3.	Univ.-Dr. Michael D. Shapiro, Tufts University, Boston / USA

Die Dissertation wurde am 25.01.2008 bei der Technischen Universität eingereicht und durch die Fakultät für Mathematik am 16.05.2008 angenommen.

Hay poesía donde hay poetas.
Hay ciencia donde hay científicos.
No podemos institucionalizar las ciencias ni las artes.

Juan Manual Briceño Guerrero.
(Venezuelan philosopher and writer)

Vorwort

Die vorliegende Arbeit entstand während meiner Tätigkeit als Promotionsstipendiat am Zentrum Mathematik der Technischen Universität München und als wissenschaftlicher Assistent am Virginia Bioinformatics Institute at Virginia Polytechnic Institute and State University, USA, sowie am Pathology Department at Tufts University's Medical School, Boston, USA.

Acknowledgements

In the chronological order of this project, I would like to express my gratitude to the following persons:

Prof. Dr. Rupert Lasser for placing his trust in my abilities as a young scientist, thus making possible this project.

Prof. Dr. Johannes Müller for introducing me to the exciting field of biomathematics.

Prof. Dr. Bodo Pareigis for introducing me to the study of monomial dynamical systems and for teaching me abstract nonsense. For all his kind interest and help.

Prof. Dr. Reinhard Laubenbacher for opening the doors of academia in the United States of America.

PhD candidate Miguel Colón-Vélez for interesting discussions.

Prof. Dr. Karen Duca for placing her trust in me and making me part of the PathSim project.

Dr. Michael Shapiro for being the person who had the strongest influence on me in the USA. For being an artist of mathematics and a mathematical artist.

Prof. Dr. David Thorley-Lawson for being a brave biologist who opened the doors of his Lab to mathematicians. For making me part of his team and teaching me lots of biology.

Dr. Omar Colón-Reyes for a very fruitful academic interaction and collaboration at the University of Puerto Rico, Mayagüez.

The graduate students at Dr. David Thorley-Lawson's Lab, in particular, Jill Roughan and Jared Hawkins for all their support.

The immunology program staff of Tufts University's Sackler School of Graduate Biomedical Sciences for their daily support.

In a more personal vein, I would like to gratefully thank my family for their constant inspiration, love and support.

Last but not least, I would like to show appreciation to Kristen Lavallee and her family for all their love and support.

Abstract

In this dissertation the following three main topics are presented:

1. The study of time discrete monomial dynamical and control systems over a finite field.
2. The analysis of a recently developed method ([65]) for reverse engineering the dynamical properties of an observed system. (The method uses the modeling paradigm of time discrete dynamical systems over a finite field.)
3. The description and parameter space exploration of the stochastic agent-based model *Path-Sim (*[36], [104]), which attempts to model and simulate the interaction of the Epstein-Barr virus and parts of the human immune system.

The specifics regarding 1.: A monomial dynamical system $f : K^n \to K^n$ over a finite field K is a nonlinear deterministic time discrete dynamical system with the property that each component function $f_i : K^n \to K$ is a monic nonzero monomial function. In this work we provide an algebraic and graph theoretic framework to study the dynamic properties of monomial dynamical systems over a finite field, in particular, the structure of cyclic trajectories. Within this framework, characterization theorems for fixed point systems (systems in which all trajectories end in steady states) are proved. In particular, we present an algorithm of polynomial complexity to test whether a given monomial dynamical system over a finite field is a fixed point system. Furthermore, theorems that complement previous work are presented and alternative proofs to previous results are supplied.

The formalism introduced in the framework mentioned above also constitutes the basis for the study of *monomial control systems*, i.e. mappings $g : K^n \times K^m \to K^n$, where $m \in \mathbb{N}$ is the number of control inputs, such that every component function $g_i : K^n \times K^m \to K$ is a *monic nonzero monomial function* in the state variable $x \in K^n$ and the control variable $u \in K^m$. Within this study, necessary and sufficient conditions for the controllability of such systems are proved. Additionally, a method for synthesizing a monomial state feedback controller is presented, which, under the assumption of controllability, imposes a desired state transition structure on monomial control systems in the closed-loop.

The specifics regarding 2.: This topic is concerned with the practical use of time discrete dynamical systems over a finite field as a modeling paradigm for biological phenomena. [65] developed a top-down reverse engineering algorithm for this paradigm. This algorithm can be seen as a parameter identification algorithm, where the parameters tuned according to the available data are abstract quantities and don't represent any physical or biological magnitude. Herein, we will refer to it as the LS-algorithm. In this thesis, a mathematical framework is developed that allows the study of the LS-algorithm in depth. This framework is based on a result (also presented in this work) that relates the concept of orthogonality and the canonical representatives for residue classes of a polynomial ideal.

Our aim is to identify minimal requirements on data sets to be used with the LS-algorithm and to characterize optimal data sets. We found minimal requirements on a data set based on how

many terms the functions to be reverse engineered display. Furthermore, we identified optimal data sets, which we characterized using a geometric property called "general position". Moreover, we developed a constructive method to generate optimal data sets, provided a codimensional condition is fulfilled. In addition, we present a generalization of the LS-algorithm that does not depend on the choice of a term order. For this method we derived a formula for the probability of finding the correct model, provided the data set used is optimal. We analyzed the asymptotic behavior of the probability formula for a growing number of variables n (i.e. interacting chemicals). Unfortunately, this formula converges to zero as fast as r^{q^n}, where $q \in \mathbb{N}$ and $0 < r < 1$. Therefore, even if an optimal data set is used and the restrictions in using term orders are overcome, the reverse engineering problem remains unfeasible, unless prodigious amounts of data are available. Such large data sets are experimentally impossible to generate with today's technologies.

The specifics regarding 3.: We provide an exposition of the agent-based model of Epstein–Barr Virus infection developed by Dr. David Thorley-Lawson and his collaborators [36], [104]. Once the model is explained, the joint work (performed during the period of time the author spent at Dr. David Thorley-Lawson's research group) of parameter space exploration is presented. The resulting biological interpretations and potential biologically relevant insights are elucidated.

The connection between these three main topics is explained in the introduction.

Contents

Introduction

Computer simulation and mathematical modeling are receiving increased attention as alternative approaches for providing insight into biological systems [41]. An important potential area of application is the increasingly complex field of immunology, and in particular, the study of viral pathogenesis. This approach is specially attractive in cases of human diseases for which applicable animal models are lacking. To date, most simulations of viral pathogenesis have tended to focus on HIV [94], [88], [87], [78], [82], [46], and employ mathematical models based on differential equations. However, there are reasons to mistrust the spatial homogeneity and well-mixed assumptions that underlie continuous models based on ordinary differential equations [85], [86], [100], despite the success of such models in immunology and virology [95], [19], [88], [80], [81], [39], [13], [27], [82]. Disease processes are spatially distributed. Indeed, it seems likely that this spatial distribution is often critical in determining the course of infection, as has been argued by many, including [7], [37]. As an alternative to ordinary differential equation models, agent-based modeling is increasingly being recognized as a viable way to simulate biological processes [58], [59], [61], [79] (See also [17], [9], [10], [12], [18], [20], [21], [51], [60], [71], [72], [77], [91], [103], [108], [109], [101], [16], [3], [4], [41], [43], [70], [114], [102], [107] for agent-based modeling approaches in the fields of immunology and pathology). The main advantage is that the "agent" paradigm complies by definition with the discrete and finite character of biological structures and entities such as organs, cells, and pathogens. This makes it more accurate, from the point of view of scientific modeling. It is also less abstract since the simulated objects, processes, and interactions usually have a straightforward biological interpretation and the spatial structure of the anatomy can be modeled meticulously. The stochasticity inherent to chemical and biological processes can be incorporated in a natural way. Moreover, agent-based models are typically local in nature, allowing the global picture to emerge from local interactions. Lastly, it is generally much easier to incorporate qualitative or semi-quantiative information into rule sets for discrete models than it is for such data to be converted to accurate rate equations.

The major drawback of using agent-based models is that there is no satisfying mathematical theory that allows for their analysis. As a consequence, currently, scientists must rely on multiple computer simulations of the model and statistical analysis of their output to assess the likely dynamical properties of the model. Developing a mathematical theory that would allow the analysis of the dynamical properties of agent-based models remains an important goal in the field. One of the aims of this thesis is to provide some contributions to that end.

Motivation

Based on a widely accepted biological model of Epstein–Barr Virus (EBV) infection (in humans), Dr. David Thorley-Lawson and his collaborators developed an agent-based model and computer simulation (*PathSim*, Pathogen Simulation) of the interaction between the virus and the host's immune system [36], [104].

I joined the PathSim project at the end of 2004, when Dr. Reinhard Laubenbacher, one of Dr. David Thorley-Lawson's collaborators, engaged me as a graduate student in his research group at the Virginia Bioinformatics Institute at Virginia Tech. Since PathSim is a stochastic agent-

based computer simulation, Dr. Laubenbacher's idea was to use the average output of PathSim as data to construct (or to "reverse engineer") a deterministic, time discrete dynamical system over a suitable finite field. To this end, he and some of his graduate students had developed some specific methods[1][65], [35]. This approach needed to be analyzed and tested, so, I performed the mathematical analysis of the reverse engineering method described in [65]. This analysis is presented in Part II of this thesis; (see also [32]).

One of the fundamental assumptions in Dr. Laubenbacher's reverse engineering program was that the deterministic, time discrete dynamical system over a finite field obtained through the reverse engineering method would reflect key dynamical properties of PathSim. Under this assumption, Dr. Laubenbacher proposed that control variables could be introduced or identified, turning the dynamical system into a *control system*. In this context, the natural issue arises as to whether the dynamical properties of such a system can be predicted. This is how the PathSim project motivated the mathematical study of deterministic, time discrete dynamical and control systems over a finite field. Part I of this thesis is devoted to the study of a particular class of dynamical and control systems over a finite field.

Some types of agent-based models and cellular automata can be interpreted as time discrete dynamical system over a finite field. In this sense, the study of such dynamical systems contributes directly to a deeper understanding of agent-based models and cellular automata. The reverse engineering method proposed by Laubenbacher and his students represents an indirect way by which the knowledge about such dynamical systems could help analyze the dynamics of agent-based models.

My involvement in the PathSim project also included an active participation in the use and interpretation of PathSim, especially after September 2006, when I joined Dr. David Thorley-Lawson's research group at the Pathology Department of Tufts University's Medical School, Boston. In Part III of this thesis, a brief description of PathSim and its capabilities is provided.

Mathematical background and motivation

The deterministic mathematical modeling efforts within the PathSim project involve the core idea of describing the different states of a biological system using the elements of a finite nonempty set X. A common fundamental assumption in a time discrete (and time invariant) deterministic modeling approach is that the future state of the system is a function of its current state. In other words, there is a function $f : X \to X$ such that the future state of the system, described by $x_{n+1} \in X$, and the current state of the system, described by $x_n \in X$, satisfy the following relationship

$$x_{n+1} = f(x_n)$$

Once such a function has been found, given an initial state of the system represented by x_0, the evolution of the system is described by the iteration of the function f. Since the sequence of values generated by this iteration represents the evolution of the system modeled, any mathematical technique that describes and predicts the dynamics of such a function (so called *dynamical system*) is highly desirable, at least from the point of view of the modeler.

To study the dynamics of such a dynamical system mathematically, it is necessary to focus on more specific classes of systems. One way to accomplish this is to add some mathematical structure to the set X, for instance, by endowing X with a topological or algebraic structure. In the vein of an algebraic approach, one could, for example, try to endow the set X with the algebraic structure of a *finite field* (with the binary operations $+$ and $\cdot$). A well-known result states that this is possible if and only if there is a prime number $p \in \mathbb{N}$ and a natural number $m \in \mathbb{N}$ such that the cardinality $|X|$ of the set X satisfies

$$|X| = p^m \tag{1}$$

[1] These methods will be described in Chapter 5.

This result imposes a limitation on our algebraic approach. However, on the other hand, if the cardinality of the set X satisfies the condition (1), it turns out that every function $f : X \to X$ is a *polynomial function* of bounded degree. (We will prove this remarkable result in Chapter 1.) This peculiarity allows for the use of a large tool box of algebraic (and even graph theoretic, as we will see in Chapter 2) techniques to study the dynamical system. As a matter of fact, mathematicians and engineers have studied such systems, in particular the linear systems, i.e. the linear space endomorphism of the vector space X over the finite field X and their higher dimensional analogues; [38], [45], [26]. From an algebraic point of view, it also seems interesting to study the *monoid endomorphisms* of the multiplicative monoid $(X, \cdot)$ and their higher dimensional analogues. These are precisely the monomial dynamical systems (to be defined below), which are studied in depth in this dissertation.

An extension of the idea of a dynamical system is the concept of a *control system*. In a control system, besides the state of the system (described by a variable $x \in X$), also intervention in the system is quantified. This intervention is represented by a *control or input variable* $u \in U$ contained in a so called control set (or space) U. Formally, such a system has the form

$$f : X \times U \to X$$

Given an initial state of the system represented by x_0, and a sequence $u_0, u_1, ..., u_t, ... \in U$ of control inputs, the system evolves according to the law

$$x_{n+1} = f(x_n, u_n)$$

One of the classical problems of control theory is the *controllability problem*, which is concerned with the existence of a suitable sequence of control inputs such that the system evolves towards a desired state (or set of states). This sequence of control inputs could, for instance, be generated as a function $g : X \to U$ of the current state

$$u_i = g(x_i)$$

Such a function is called a *feedback law*. Another important control theoretic issue is the design of suitable feedback laws, such that the so called *closed-loop system*

$$\begin{array}{rcl} h & : & X \to X \\ x & \mapsto & f(x, g(x)) \end{array}$$

satisfies predetermined dynamical properties.

Mathematical control theory is a growing field. Especially for linear functions $f : X \times U \to X$ (when the sets X and U are endowed with a vector space structure), very satisfying answers to the problems mentioned above (and to other, similar problems) have been found. See, for instance, [106], and in the framework of finite fields, [92]. In this thesis, we extend the concept of monomial dynamical systems to *monomial control systems* (to be defined below) and perform a control theoretic study of them.

Contributions of this Dissertation

The first main contribution of this work is to provide an algebraic and graph theoretic framework to study the dynamic properties of monomial dynamical systems over a finite field. These are mappings $f : K^n \to K^n$, where K is a finite field and $n \in \mathbb{N}$, such that every component function $f_i : K^n \to K$ is a *monic nonzero monomial function*. Within this framework, characterization theorems for fixed point systems (systems in which all trajectories end in steady states) are proved. In particular, an algorithm of polynomial complexity is presented, which tests whether a given monomial dynamical system over a finite field is a fixed point system [33]. Furthermore,

theorems that complement previous work by [23], [22] and [24] are presented, and alternative proofs to previous results are supplied. Many of these theorems discuss the structure of cyclic trajectories. The formalism introduced in our framework also constitutes the basis for the study of *monomial control systems*, i.e. mappings $g : K^n \times K^m \to K^n$, where $m \in \mathbb{N}$ is the number of control inputs, such that every component function $g_i : K^n \times K^m \to K$ is a *monic nonzero monomial function* in the state variable $x \in K^n$ and the control variable $u \in K^m$.

A further important novelty presented in this thesis is the control theoretic study of monomial control systems. In particular, necessary and sufficient conditions for the controllability of such systems are proved. Additionally, a method for synthesizing a monomial state feedback controller is presented, which, under the assumption of controllability, imposes a desired state transition structure on monomial control systems in the closed-loop.

The third main result in this thesis is concerned with the practical use of time discrete dynamical systems over a finite field as a modeling paradigm for biological phenomena. [65] developed a top-down reverse engineering algorithm for this paradigm. This algorithm can be seen as a parameter identification algorithm, where the parameters tuned according to the available data are abstract quantities and don't represent any physical or biological magnitude. Herein, we will refer to it as the LS-algorithm. In this thesis, a mathematical framework is developed that allows the study of the LS-algorithm in depth. This framework is based on a result that relates the concept of orthogonality and the canonical representatives for residue classes of a polynomial ideal. This result itself constitutes a pure algebraic contribution of this work [31]. Having expressed the steps of the LS-algorithm in our framework, concrete answers to the following questions are provided:

1. What are the minimal requirements on data sets?

2. Can data sets be characterized in such a way that "optimal" data sets can be identified? (Optimality meaning that the algorithm performs better using such a data set compared to its performance using other data sets.)

The second question is related to the *design of experiments* and optimality is characterized in terms of *quantity and quality* of the data sets. Furthermore, a generalization of the LS-algorithm that does not depend on the choice of a term order is introduced. For this method, a formula for the probability of finding the correct model is derived, provided the data set used satisfies an optimality criterion. In addition, the asymptotic behavior of the probability formula is analyzed for a growing number of variables n (i.e. interacting entities modeled) [32].

In the last part of this thesis the reader will find an exposition of the agent-based model of Epstein–Barr Virus infection developed by Dr. David Thorley-Lawson and his collaborators. Once the model is explained, the joint work (performed during the period of time the author spent at Dr. David Thorley-Lawson's research group) of parameter space exploration is presented. The resulting biological interpretations and potential biologically relevant insights are elucidated. This represents a co-contribution of this thesis to both the biological and biomedical sciences [104].

Outline

This thesis is subdivided in 3 parts. Part I is devoted to the theory of monomial dynamical and control systems over a finite field. It comprises three chapters: Chapter 1 introduces time discrete finite dynamical systems as well as time discrete dynamical systems over a finite field. Chapter 2 is devoted to the study of the dynamics of monomial dynamical systems over a finite field. Chapter 3 studies monomial control systems over a finite field with emphasis on Boolean monomial control systems.

Part II deals with the practical use of time discrete dynamical systems over a finite field as a modeling paradigm for biological phenomena. It starts with Chapter 4, which is an excursus into the relationship between the concept of orthogonality and the canonical representatives for

residue classes of a polynomial ideal. This material provides the basis for Chapter 5, in which the LS-algorithm and its performance are studied in depth.

Part III elucidates the biological research that motivated the mathematical developments presented in this dissertation. Chapter 6 provides an exposition of the agent-based model PathSim. Chapter 7 presents the results of the parameter space exploration as well as some of its biological interpretations and potential consequences.

Part I

The theory of monomial dynamical and control systems over a finite field

Chapter 1

Time discrete finite dynamical systems

1.1 Definition, characteristics and general dynamical properties

Definition 1 *Let X be a nonempty finite set and $n \in \mathbb{N}$ a natural number. An n-dimensional time invariant time discrete finite dynamical system is a mapping*

$$f : X^n \to X^n$$

Remark 2 *Due to the finiteness of X it is obvious that the trajectory*

$$x,\ f(x),\ f^2(x), ...$$

of any point $x \in X^n$ contains at most $|X^n| = |X|^n$ different points and therefore becomes either cyclic or converges to a single point $y \in X$ with the property $f(y) = y$ (i.e. a fixed point of f).

1.1.1 The phase space and the period number

Definition 3 (Notational Definition) *A directed graph $G = (V_G,\ E_G,\ \pi_G : E_G \to V_G \times V_G)$ that allows self loops and parallel directed edges is called* digraph.

Definition 4 *Let $G = (V_G,\ E_G,\ \pi_G)$ be a digraph. Two vertices $a, b \in V_G$ are called* connected *if there is a $t \in \mathbb{N}_0$ and (not necessarily different) vertices $v_1, ..., v_t \in V_G$ such that*

$$a \to v_1 \to v_2 \to ... \to v_t \to b$$

In this situation we write $a \rightsquigarrow_s b$, where s is the number of directed edges involved in the sequence *from a to b (in this case $s = t + 1$). Two sequences $a \rightsquigarrow_s b$ of the same length are considered different if the directed edges involved are different or the order at which they appear is different, even if the visited vertices are the same. As a convention, a single vertex $a \in V_G$ is always connected to itself $a \rightsquigarrow_0 a$ by an empty sequence of length* 0.

Definition 5 *Let $G = (V_G,\ E_G,\ \pi_G)$ be a digraph and $a, b \in V_G$ two vertices. A sequence $a \rightsquigarrow_s b$*

$$a \to v_1 \to v_2 \to ... \to v_t \to b$$

is called a path, *if no vertex v_i is visited more than once. If $a = b$, but no other vertex is visited more than once, $a \rightsquigarrow_s b$ is called a* closed path.

Definition 6 *Let X be a nonempty finite set, $n \in \mathbb{N}$ a natural number and $f : X^n \to X^n$ a time discrete finite dynamical system. The* phase space *of f is the digraph with node set X^n, arrow set E defined as*

$$E := \{(x, y) \in X^n \times X^n \mid f(x) = y\}$$

and vertex mapping

$$\begin{array}{rcl} \pi & : & E \to X^n \times X^n \\ (x, y) & \mapsto & (x, y) \end{array}$$

Remark 7 *The phase space consists of closed paths of different lengths between* 1 *(i.e. loops centered on fixed points) and $|X^n| = |X|^n$ and directed trees that end each one at exactly one closed path. The nodes in the directed trees correspond to* transient states *of the system. In particular, if f is bijective*[1]*, every point $x \in X^n$ is contained in a closed path and the phase space is the union of disjoint closed paths. Conversely, if every point in the phase space is contained in a closed path, then f must be bijective.*

Definition 8 *Let X be a nonempty finite set, $n \in \mathbb{N}$ a natural number and $f : X^n \to X^n$ a time discrete finite dynamical system. We define*

$$\begin{array}{rcl} f^0 & : & = id : X^n \to X^n \\ x & \mapsto & x \end{array}$$

and for $t \in \mathbb{N}$ we recursively define

$$\begin{array}{rcl} f^t & : & X^n \to X^n \\ x & \mapsto & f(f^{t-1}(x)) \end{array}$$

Given a time discrete finite dynamical system $f : X^n \to X^n$, we can find in the phase space the longest path ending in a closed path. Let $m \in \mathbb{N}_0$ be the length of this path. It is easy to see, that for any $s \geq m$ the time discrete finite dynamical system $f^s : X^n \to X^n$ has the following properties

1. $\forall\ x \in X^n$, $f^s(x)$ is a node contained in one closed path of the phase space.
2. If T is the least common multiple of all the lengths of closed paths displayed in the phase space, then it holds
$$f^{s+\lambda T} = f^s\ \forall\ \lambda \in \mathbb{N}$$
and
$$f^{s+i} \neq f^s\ \forall\ i \in \{1, ..., T-1\}$$
We call T the *period number* of f. If $T = 1$, f is called a *fixed point system*.

Definition 9 *Let X be a nonempty finite set, $n \in \mathbb{N}$ a natural number and $f : X^n \to X^n$ a time discrete finite dynamical system. Furthermore, let $s \in \mathbb{N}$ with $s \leq |X^n|$. A closed path of length s in the phase space of f is called a* cycle of length s*. We refer to the total number of cycles and their legths in the phase space of f as the* cycle structure of f.

Definition 10 *Let X be a nonempty finite set, $n \in \mathbb{N}$ a natural number and $f : X^n \to X^n$ a time discrete finite dynamical system. Furthermore, let $s \in \mathbb{N}$ with $s \leq |X^n|$. A point $\xi \in X^n$ is said to show* s-periodicity under f *if $f^s(\xi) = \xi$.*

[1] Note that for any map from a finite set into itself, surjectivity is equivalent to injectivity.

Remark 11 *Let $t \in \mathbb{N}$ be an integer multiple of s. A point $\xi \in X^n$ that shows s-periodicity under f also shows t-periodicity under f ($f^t(\xi) = f^{\lambda s}(\xi) = (f^s)^\lambda(\xi) = (f^s \circ ... \circ f^s)(\xi) = \xi$). As a consequence, a point $\xi \in X^n$ is contained in a cycle of length t in the phase space of f if and only if ξ satisfies the equation $f^t(\xi) = \xi$, but $f^d(\xi) \neq \xi$ for any integer divisor d of t with $d < t$. Once such a point has been found, necessarily further $t-1$ pairwise different points $\xi_1, ..., \xi_{(t-1)} \in X^n$ with $f^t(\xi_i) = \xi_i$ and $f^d(\xi_i) \neq \xi_i$ must exist, namely, the points $f(\xi)$, $f^2(\xi), ..., f^{t-1}(\xi)$.*

1.1.2 The dependency graph

Definition 12 *Let M be a nonempty finite set. Furthermore, let $n := |M|$ be the cardinality of M. A* numeration *of the elements of M is a bijective mapping*

$$f : M \to \{1, ..., n\}$$

Given a numeration f of the set M we write

$$M = \{f_1, ..., f_n\}$$

where the unique element $x \in M$ with the property $f(x) = i \in \{1, ..., n\}$ is denoted as f_i.

Definition 13 *Let X be a nonempty finite set, $n \in \mathbb{N}$ a natural number and $f : X^n \to X^n$ a time discrete finite dynamical system. Furthermore, let $G = (V_G, E_G, \pi_G)$ be a digraph with vertex set V_G of cardinality $|V_G| = n$. The digraph G is called* dependency graph *of f iff a numeration $a : M \to \{1, ..., n\}$ of the elements of V_G exists such that $\forall\ i, j \in \{1, ..., n\}$ the following holds*

$$\exists\ e \in E_G : \pi_G(e) = (a_i, a_j) \Leftrightarrow f_i \text{ depends on } x_j$$

Remark 14 *In the chapter about monomial dynamical systems we will introduce a slightly different and more specific definition of dependency graph.*

1.2 Time discrete dynamical systems over a finite field

Given a time discrete finite dynamical system $f : X^n \to X^n$, how can the period number T be calculated? An obvious brute force procedure would be to actually determine the structure of the phase space by evaluating f on each of the $|X|^n$ different points of the space X^n. However, this quickly becomes computationally intractable, even for small dimension n. If the set X is endowed with an algebraic structure and we consider certain classes of time discrete finite dynamical systems, we might have more mathematical structure and tools to solve this problem. Indeed, if X can be endowed with the algebraic structure of a finite field, all component functions f_i of a system $f : X^n \to X^n$ are polynomial functions in n variables. We will show this remarkable and well-known result in the next subsection. (See, for instance, pages 368-369 in [67] for a different proof.)

1.2.1 The ring of polynomial functions in n variables over $\mathbf{F}_q$ and the vector space of functions $\mathbf{F}_q^n \to \mathbf{F}_q$

Definition 15 (Notational Definition) *Since for every finite field K there is a prime number $p \in \mathbb{N}$ (the characteristic of K) and a natural number $n \in \mathbb{N}$ such that for the number of elements $|K|$ of K it holds*

$$|K| = p^n$$

we will denote a finite field with $\mathbf{F}_q$, where q stands for the number of elements of the field (See, for instance, [67]). Clearly, q is a power of the (prime) characteristic of the field.

Definition 16 *Let $n \in \mathbb{N}$ be a natural number. An n-tuple $\alpha = (\alpha_1, ..., \alpha_n) \in (\mathbb{N}_0)^n$ is called multi index.*

Definition 17 (Notational definition) *We call a commutative Ring $(R, +, \cdot)$ with multiplicative identity $1 \neq 0$ and the binary operations $\cdot$ and $+$ just Ring R.*

Definition 18 *Let R be a ring and $n \in \mathbb{N}$ a natural number. For $n > 1$, the elements of the Cartesian product R^n are marked by arrows, e.g. $\vec{x} \in R^n$. Let $\alpha \in (\mathbb{N}_0)^n$ be a multi index. For a product of powers of the form $x_1^{\alpha_1} \cdot ... \cdot x_n^{\alpha_n} \in R$ we write*

$$\overrightarrow{x}^{\alpha} := x_1^{\alpha_1} \cdot ... \cdot x_1^{\alpha_n}$$

Definition 19 *Let R be a ring, $n \in \mathbb{N}$ a natural number and $m \in \mathbb{N}_0$ a non negative integer. Furthermore let $a_i \in R$, $i = 0, ..., m$ be elements of the ring R and $\alpha_j = (\alpha_{j1}, ..., \alpha_{jn}) \in \mathbb{N}_0^n$, $j = 0, ..., m$ multi indexes of length n. A mapping of the form*

$$\begin{array}{rcl} g & : & R^n \to R \\ \vec{x} & \mapsto & g(\overrightarrow{x}) := \sum_{j=0}^{m} a_j \overrightarrow{x}^{\alpha_j} \end{array}$$

is called polynomial function over R in n R-valued variables. *If $m = 0$ and $a_0 \neq 0$ then g is also called a* monomial function over R in n R-valued variables.

Theorem 20 (and Definition) *Let R be a ring and $n \in \mathbb{N}$ a natural number. The set*

$$PF_n(R) := \{g \mid g : R^n \to R \text{ is polynomial}\}$$

together with the common operations $+$ and $\cdot$ of addition and multiplication of mappings is a ring. This ring is called ring of all polynomial functions over R in n R-valued variables.

Proof. The easy proof is left to the reader. ■

Theorem 21 (and Definition) *Let K be an arbitrary field and $n \in \mathbb{N}$ a natural number. The set of all functions*

$$f : K^n \to K$$

together with the common operations of addition of mappings and scalar multiplication is a vector space over K. We denote this vector space with $F_n(K)$.

Proof. The easy proof is left to the reader. ■

Definition 22 *Let $n, q \in \mathbb{N}$ be natural numbers. Further let $>$ be a total ordering on $(\mathbb{N}_0)^n$. The according to $>$ decreasingly ordered set*

$$M_q^n := \{\alpha \in (\mathbb{N}_0)^n \mid \alpha_j < q \ \forall \ j \in \{1, ..., n\}\}$$

of all n-tuples with entries smaller than q is denoted by $M_q^n \subset (\mathbb{N}_0)^n$.

Remark 23 *In order to avoid a too complicated notation, we skip the appearance of the order relation $>$ in the symbol for this set. It is easy to prove, that M_q^n contains exactly q^n n-tuples. We will index the n-tuples in M_q^n starting with the biggest and ending with the smallest:*

$$\alpha_1 > \alpha_2 > ... > \alpha_{q^n}$$

Definition 24 *For any fixed natural numbers $n, q \in \mathbb{N}$ and for each multi index $\alpha \in M_q^n$ consider the monomial function*

$$\begin{array}{rcl} g_{nq\alpha} & : & K^n \to K \\ \vec{x} & \mapsto & g_{nq\alpha}(\overrightarrow{x}) := \overrightarrow{x}^{\alpha} \end{array}$$

All these monomial functions $g_{nq\alpha}$, $\alpha \in M_q^n$ are called fundamental monomial functions.

Lemma 25 *Let $\mathbf{F}_q$ be a finite field. The vector space $F_n(\mathbf{F}_q)$ has the dimension q^n.*

Proof. Since $\mathbf{F}_q$ contains exactly q elements, there are exactly q^n n-tuples contained in $\mathbf{F}_q^n$. For each n-tuple $\vec{y} \in \mathbf{F}_q^n$ consider the characteristic function

$$\begin{array}{rcl} \mathbf{1}_{\{\vec{y}\}} & : & \mathbf{F}_q^n \to \mathbf{F}_q \\ \vec{x} & \mapsto & \mathbf{1}_{\{\vec{y}\}}(\vec{x}) := \left\{ \begin{array}{c} 1 \text{ if } \vec{x} = \vec{y} \\ 0 \text{ otherwise} \end{array} \right. \end{array}$$

Obviously, every function $f : \mathbf{F}_q^n \to \mathbf{F}_q$ can be written as a linear combination of all characteristic functions. In addition, the q^n characteristic functions $(\mathbf{1}_{\{\vec{y}\}})_{\vec{y} \in \mathbf{F}_q^n}$ are linearly independent, as can be easily concluded from evaluating the expression

$$\sum_{\vec{y} \in \mathbf{F}_q^n} \lambda_{i(\vec{y})} \mathbf{1}_{\{\vec{y}\}} \equiv 0$$

at any value $\vec{x} \in \mathbf{F}_q^n$ ■

Theorem 26 *Let $\mathbf{F}_q$ be a finite field. A basis for the vector space $F_n(\mathbf{F}_q)$ is given by the fundamental monomial functions*

$$(g_{nq\alpha})_{\alpha \in M_q^n}$$

Proof. First, we show that $(g_{nq\alpha})_{\alpha \in M_q^n}$ are linearly independent. For this purpose, we will use induction on the number of variables n : ■

When $n = 1$ we have

$$\{g_{1q\alpha}(x) \mid \alpha \in M_q^1\} = \{1, x^1, x^2, ..., x^{q-1}\}$$

Now consider a linear combination

$$\lambda_0 + \lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_{q-1} x^{q-1},\ \lambda_i \in \mathbf{F}_q,\ i = 0, ..., q-1$$

where not all λ_i are equal to 0. If it holds

$$\lambda_0 + \lambda_1 x^1 + \lambda_2 x^2 + ... + \lambda_{q-1} x^{q-1} = 0 \ \forall\ x \in \mathbb{Z}_q^1$$

then we could construct a nonzero polynomial in $\mathbf{F}_q[\tau]$ of degree less or equal to $q-1$ having $|\mathbf{F}_q| = q$ distinct roots. This is a contradiction to the well known fact, that the number of roots of a nonzero polynomial $h \in R[\tau]$ over an integral domain R is bounded by the degree of the polynomial. Now let $n > 1$ and assume that the linear independence is given for $n-1$. Using the multi indexes in M_q^n to index the coefficients, we can write a linear combination of the $(g_{nq\alpha})_{\alpha \in M_q^n}$ as

$$\sum_{\alpha \in M_q^n} \lambda_\alpha g_{nq\alpha}(\vec{x}) = \sum_{\alpha \in M_q^n} \lambda_\alpha \overrightarrow{x}^{\alpha},\ \lambda_\alpha \in \mathbf{F}_q,\ \alpha \in M_q^n$$

By collecting the various powers of x_n, we can write the above expression in the form

$$\sum_{i=0}^{q-1} h_i(x_1, ..., x_{n-1}) x_n^i$$

where the h_i, $i = 0, ..., q-1$ are the following linear combinations of the monomial functions $(g_{(n-1)q\beta})_{\beta \in M_q^{n-1}}$

$$h_i = \sum_{\beta \in M_q^{n-1}} \lambda_{(\beta_1,...,\beta_{n-1},i)} g_{(n-1)q\beta} \ \forall \ i \in \{0, ..., q-1\}$$

Now if it holds

$$\sum_{\alpha \in M_q^n} \lambda_\alpha \overrightarrow{x}^\alpha = 0 \ \forall \ \vec{x} \in \mathbf{F}_q^n$$

then in particular

$$\sum_{i=0}^{q-1} h_i(x_1, ..., x_{n-1}) x_n^i = 0 \ \forall \ \vec{x} \in \mathbf{F}_q^n$$

The same reasoning as in in the case $n = 1$ forces all the polynomial functions h_i, $i = 0, ..., q-1$ to be the zero function:

$$h_i(\vec{y}) = \sum_{\beta \in M_q^{n-1}} \lambda_{(\beta_1,...,\beta_{n-1},i)} g_{(n-1)q\beta}(\vec{y}) = 0 \ \forall \ \vec{y} \in \mathbb{Z}_q^{n-1}, \ i = 0, ..., q-1$$

From the induction hypothesis now follows for every $i \in \{0, ..., q-1\}$

$$\lambda_{(\beta_1,...,\beta_{n-1},i)} = 0 \ \forall \ \beta \in M_q^{n-1}$$

and therefore

$$\lambda_\alpha = 0 \ \forall \ \alpha \in M_q^n$$

Now, it follows from Remark 23 and from the linear independence of the monomial functions $(g_{nq\alpha})_{\alpha \in M_q^n}$, that $\left|\left\{g_{nq\alpha} \mid \alpha \in M_q^n\right\}\right| = q^n$, which is the dimension of $F_n(\mathbf{F}_q)$ according to the previous lemma.

Remark 27 *The basis elements in the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ are ordered according to the order relation $>$ used to order the n-tuples in the set M_q^n. That means (see Remark 23)*

$$(g_{nq\alpha})_{\alpha \in M_q^n} = (g_{nq\alpha_i})_{i \in \{1,...,q^n\}}$$

Corollary 28 *Let $\mathbf{F}_q$ be a finite field. Then for the* sets *$F_n(\mathbf{F}_q)$ and $PF_n(\mathbf{F}_q)$ it holds*

$$F_n(\mathbf{F}_q) = PF_n(\mathbf{F}_q)$$

Proof. The inclusion

$$PF_n(\mathbf{F}_q) \subseteq F_n(\mathbf{F}_q)$$

is given by the definitions. Now, since every function $f \in F_n(\mathbf{F}_q)$ can be written as a linear combination of the fundamental monomial functions $(g_{nq\alpha})_{\alpha \in M_q^n}$ we have

$$f \in PF_n(\mathbf{F}_q)$$

and therefore

$$F_n(\mathbf{F}_q) \subseteq PF_n(\mathbf{F}_q)$$

■

1.2.2 Currently available techniques for dynamics forecast

We come back to the more specific question: Given a time discrete finite dynamical system $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$, over a finite field $\mathbf{F}_q$, how can the period number T be calculated? If $T > 1$, can the cycle structure be accurately described? For instance, can the number of cycles and their lengths be determined without actually constructing the phase space?

To the author's best knowledge, the systems for which a complete and satisfying theory exists are the linear systems, (i.e. $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ is a linear map). See the seminal work of Bernard Elspas [38] and also [45] for an excellent and more mathematical exposition. [93] and [92] present applications of the Boolean case in control theory. Furthermore, the affine case was studied by [73]. An interesting contribution was made by Paul Cull ([26]), who extended the considerations to nonlinear functions, and showed how to reduce them to the linear case. However, the drawback of this method is that, if the nonlinear system has dimension n and the field has q elements, then the linear system has dimension q^n. It is also very difficult to see directly the effect of the specific nonlinear functions on the state space structure.

For monomial systems, i.e. $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ is such that every component f_i is a *monic monomial function* in n variables, groundbreaking results were achieved by [23] and [22]. It is one of the aims of this dissertation to extend and supplement those results. Moreover, we provide a novel mathematical formalism to study monomial systems. Chapter 2 is devoted to the theory of monomial systems.

We finish this chapter with a short review of the linear systems theory:

The linear theory is based on the fact that there are subspaces $U, W \subseteq \mathbf{F}_q^n$ with the property $\mathbf{F}_q^n = U \oplus W$ such that the subgraph of trees of transient states and the subgraph of all cycles in the phase space of a linear system $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ can be linked to a nilpotent linear mapping $f_N : U \to U$ and a bijective linear mapping $f_B : W \to W$, respectively.

The length and structure of the trees of transient states can now be obtained by analyzing the Jordan canonical form of the matrix associated to the mapping f_N by means of Theorems 2 and 3 of [45].

The lengths and number of cycles can be calculated from the factorization in so called *elementary divisor polynomials* of the characteristic polynomial of the mapping f_B by repeated use of Elspas' formula (see [38] and Theorem 5 of [45]).

In [93] the calculation of the Jordan canonical form representing the mapping f_N is avoided by using a slightly different approach that takes advantage of the *Smith form* of a matrix. However, the authors do not elaborate on the computational aspects of calculating the Smith form.

Chapter 2

Monomial dynamical systems over a finite field

As discussed at the end of the previous chapter, there exists a satisfying theory that explains the dynamical properties of linear systems $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$, i.e. the vector space endomorphisms of the vector space $\mathbf{F}_q^n$ over the field $\mathbf{F}_q$. From an algebraic point of view, it also seems interesting to study the monoid endomorphisms of the multiplicative monoid $(\mathbf{F}_q, \cdot)$ as well as the monoid endomorphisms of the direct product[1] $\prod_{i=1}^{n}(\mathbf{F}_q, \cdot)$. Due to the well known fact that the multiplicative group $(\mathbf{F}_q^*, \cdot)$ is cyclic, it turns out, that the monoid endomorphisms of the direct product $\prod_{i=1}^{n}(\mathbf{F}_q, \cdot)$ are precisely the monomial mappings, i.e. mappings $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ such that every component f_i is a monic monomial function in n variables.

Since the component functions f_i of time discrete finite dynamical systems $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ over a finite field $\mathbf{F}_q$ are polynomial functions, (i.e. linear combinations of monic monomial functions), the combined knowledge about linear systems and monomial systems could represent the starting point for a complete theory of time discrete finite dynamical systems over a finite field.

In this chapter we provide key results towards a better understanding of monomial systems.

Throughout this chapter, and in contrast to Chapter 1 and Part II of this thesis, we will denote the elements of the Cartesian product $\mathbf{F}_q^n$ as $x \in \mathbf{F}_q^n$, neglecting the vector arrow.

2.1 What are monomial dynamical systems?

In this section we will introduce n-dimensional monomial dynamical systems over a finite field $\mathbf{F}_q$. Moreover, we will show that the study of monomial dynamical systems is actually the study of the monoid endomorphisms of the direct product $\prod_{i=1}^{n}(\mathbf{F}_q, \cdot)$. To this end we will prove the equality of the set of all monoid endomorphisms of the multiplicative monoid $(\mathbf{F}_q, \cdot)$ and the set of all one dimensional monomial systems over the finite field $\mathbf{F}_q$. This set becomes a monoid if it is endowed with the binary operations of composition and product, respectively.

Definition 29 *Let $\mathbf{F}_q$ be a finite field. The set*

$$E_q := \{0, ...q-1\} \subset \mathbb{N}_0$$

is called the exponents set of the field $\mathbf{F}_q$.

[1] The direct product $\prod_{i=1}^{n}(M, \cdot)$ of a monoid $(M, \cdot)$ is defined as the set $\prod_{i=1}^{n} M$ (cartesian product) endowed with the binary operation $\cdot$ defined as $(x \cdot y)_i := x_i y_i$.

Definition 30 *Let $\mathbf{F}_q$ be a finite field. A map $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ is called a monic monomial dynamical system over $\mathbf{F}_q$ if for every $i \in \{1, ..., n\}$ there exists a tuple $(F_{i1}, ..., F_{in}) \in E_q^n$ such that*

$$f_i(x) = x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n$$

We will call a monic monomial dynamical system just monomial dynamical system.

Remark 31 *We exclude in the definition of monomial dynamical system the possibility that one of the functions f_i is equal to the zero function. This is not a loss of generality because of the following: If we were studying a dynamical system $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ where one of the functions, say f_j, was equal to zero, then, for every initial state $x \in \mathbf{F}_q^n$ after one iteration the system would be in a state $f(x)$ whose jth entry is zero. In all subsequent iterations the value of the jth entry would remain zero. As a consequence, the long term dynamics of the system are reflected in the projection*

$$\begin{array}{rcl} \pi_{\widehat{j}} & : & \mathbf{F}_q^n \to \mathbf{F}_q^{n-1} \\ y & \mapsto & \pi_{\widehat{j}}(y) := (y_1, ..., y_{j-1}, y_{j+1}, ..., y_n)^t \end{array}$$

and it is sufficient to study the system

$$\begin{array}{rcl} \widetilde{f} & : & \mathbf{F}_q^{n-1} \to \mathbf{F}_q^{n-1} \\ y & \mapsto & \begin{pmatrix} f_1(y_1, ..., y_{j-1}, 0, y_{j+1}, ..., y_n) \\ \vdots \\ f_{j-1}(y_1, ..., y_{j-1}, 0, y_{j+1}, ..., y_n) \\ f_{j+1}(y_1, ..., y_{j-1}, 0, y_{j+1}, ..., y_n) \\ \vdots \\ f_n(y_1, ..., y_{j-1}, 0, y_{j+1}, ..., y_n) \end{pmatrix} \end{array}$$

In general, this system $\widetilde{f}$ could contain component functions equal to the zero function, since every component f_i that depends on the variable x_j would become zero. As a consequence, the procedure described above needs to be applied several times until the lower n'-dimensional system obtained does not contain component functions equal to zero. The long term dynamics of f are reflected in the projection to an n'-dimensional subspace, in particular, all the cycles and fixed points of f are located in this lower dimensional space. Moreover, points located outside this lower dimensional subspace are transient states of the system. It is also possible that this repeated procedure yields the one dimensional zero function. In this case, we can conclude that the original system f is a fixed point system with $(0, ..., 0) \in \mathbf{F}_q^n$ as its unique fixed point. Note that this procedure reduces the dimension by s where $0 \leq s \leq n$. As a consequence, the procedure needs to be iterated at most n times.

Definition 32 (Notational Definition) *Let be $n, m \in \mathbb{N}$ natural numbers and S a set. The set of all $m \times n$ matrices (m* rows and n columns*) with entries in S is denoted by $M(m \times n; S)$.*

Definition 33 *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers. The set*

$$MF_m^n(\mathbf{F}_q) := \{f : \mathbf{F}_q^m \to \mathbf{F}_q^n \mid \exists \ F \in M(n \times m; E_q) : f_i(x) := x_1^{F_{i1}} ... x_m^{F_{im}} \ \forall \ x \in \mathbf{F}_q^m\}$$

is called the set of n-dimensional monomial mappings in m variables.

Lemma 34 *Let $\mathbf{F}_q$ be a finite field. Then the multiplicative group*[2] *$(\mathbf{F}_q^*, \cdot)$ is cyclic.*

[2] $\mathbf{F}_q^* := \mathbf{F}_q \backslash \{0\}$

Proof. The proof of this well known fact can be found, for instance, in Theorem 2.8 of [67]. ■

The validity of the following theorem represents an algebraic justification for the conventional definition

$$0^0 = 1 \tag{2.1}$$

(where the zero in the exponent is the entire number $0 \in \mathbb{Z}$, whereas $0, 1 \in \mathbf{F}_q$). The next theorem would be namely false if we set a different value for the expression 0^0. In real analysis, the convention (2.1) has a justification in terms of rendering the function

$$\begin{array}{rcl} h & : & \mathbb{R} \to \mathbb{R} \\ x & \mapsto & x^0 \end{array}$$

continuous at the point $x = 0$.

Theorem 35 *Let $\mathbf{F}_q$ be a finite field. Furthermore, let $End(\mathbf{F}_q, \cdot)$ be the set of monoid endomorphisms of the multiplicative monoid $(\mathbf{F}_q, \cdot)$. Then the following set theoretic equality holds*

$$End(\mathbf{F}_q, \cdot) = MF_1^1(\mathbf{F}_q)$$

Proof. To show the inclusion $End(\mathbf{F}_q, \cdot) \subseteq MF_1^1(\mathbf{F}_q)$ we consider an arbitrary monoid endomorphism $f : (\mathbf{F}_q, \cdot) \to (\mathbf{F}_q, \cdot)$. The goal is to show that there is an $a(f) \in E_q$ such that

$$f(x) = x^{a(f)} \ \forall \ x \in \mathbf{F}_q$$

To prove this, consider a generator $u \in \mathbf{F}_q^*$ of the cyclic group $(\mathbf{F}_q^*, \cdot)$ (see Lemma 34). Since the order of the cyclic finite group $(\mathbf{F}_q^*, \cdot)$ is $q-1$, for every element $x \in \mathbf{F}_q^*$ there is an $i \in \{1, ..., q-1\}$ such that

$$x = u^i$$

Now we consider the two following cases:

$f(u) = 0$. This case is not possible, since $f(1) = f(u^{q-1}) = f(u)^{q-1} = 0 \neq 1$, i.e. f wouldn't be a monoid homomorphism.

$f(u) \neq 0$. In this case $f(u) \in \mathbf{F}_q^*$ and therefore $\exists \ a \in \{1, ..., q-1\}$ such that

$$f(u) = u^a$$

As a consequence, $\forall \ x \in \mathbf{F}_q^*$ we have

$$f(x) = f(u^i) = f(u)^i = (u^a)^i = (u^i)^a = x^a \tag{2.2}$$

For $x = 0$ we have two possibilities: $f(0) = 0$ or $f(0) \neq 0$. In the former situation, we immediately have

$$f(x) = x^a \ \forall \ x \in \mathbf{F}_q$$

On the other hand, if $f(0) \neq 0$, it follows

$$f(0) = f(0 \cdot 0) = f(0) \cdot f(0)$$

thus

$$f(0) = 1$$

and consequently

$$1 = f(0 \cdot u) = f(0) \cdot f(u) = f(u)$$

By equation (2.2) this implies

$$f(x) = 1 \ \forall x \in \mathbf{F}_q^*$$

Summarizing, we have $f \equiv 1$. In this situation, the convention (2.1) saves the proof by allowing

$$f(x) = x^0 \ \forall \ x \in \mathbf{F}_q$$

(any other exponent $b \in E_q \backslash \{0\}$ would yield $f(0) = 0^b = 0$). Given 2.1, the converse inclusion $MF_1^1(\mathbf{F}_q) \subseteq End(\mathbf{F}_q, \cdot)$ follows from the fact that all the functions contained in $MF_1^1(\mathbf{F}_q)$, namely

$$\begin{array}{rcl} f_i & : & \mathbf{F}_q \to \mathbf{F}_q \\ x & \mapsto & x^i \end{array}$$

where $i = 0, ..., q-1$, obviously satisfy the axioms of a monoid homomorphism. ■

Corollary 36 *Let $\mathbf{F}_q$ be a finite field and $n \in \mathbb{N}$ a natural number. Furthermore, let $End(\prod_{i=1}^{n}(\mathbf{F}_q, \cdot))$ be the set of monoid endomorphisms of the direct product*[3] *$\prod_{i=1}^{n}(\mathbf{F}_q, \cdot)$. Then the following set theoretic equality holds*

$$End(\prod_{i=1}^{n}(\mathbf{F}_q, \cdot)) = MF_n^n(\mathbf{F}_q)$$

Proof. To show the inclusion $End(\prod_{i=1}^{n}(\mathbf{F}_q, \cdot)) \subseteq MF_n^n(\mathbf{F}_q)$ we consider an arbitrary monoid endomorphism $f : \prod_{i=1}^{n}(\mathbf{F}_q, \cdot) \to \prod_{i=1}^{n}(\mathbf{F}_q, \cdot)$. The goal is to show that there is an $F \in M(n \times m; E_q)$ such that for each $i \in \{1, ..., n\}$

$$f_i(x) = x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n$$

Consider for each $i, j \in \{1, ..., n\}$ the function defined as

$$\begin{array}{rcl} h_{ij} & : & \mathbf{F}_q \to \mathbf{F}_q \\ x & \mapsto & f_i(\gamma_j(x)) \end{array}$$

where $\gamma_j : \mathbf{F}_q \to \mathbf{F}_q^n$ is a mapping such that $\forall \ x \in \mathbf{F}_q$

$$\gamma_j(x)_k = \begin{cases} 1 \text{ if } k \neq j \\ x \text{ if } k = j \end{cases}$$

Since f is a monoid endomorphism, it follows from the definition of h_{ij}

$$h_{ij}(1) = f_i(1, ..., 1) = 1$$

and

$$h_{ij}(xy) = f_i(\gamma_j(xy)) = f_i(\gamma_j(x) \cdot \gamma_j(y)) = f_i(\gamma_j(x)) f_i(\gamma_j(y)) = h_{ij}(x) h_{ij}(y)$$

Consequently $h_{ij} : \mathbf{F}_q \to \mathbf{F}_q$ is a monoid endomorphism $\forall \ i, j \in \{1, ..., n\}$. By Theorem 35 we know $\exists \ F_{ij} \in \{0, ..., q-1\}$ such that

$$h_{ij}(x) = x^{F_{ij}} \ \forall \ x \in \mathbf{F}_q$$

Again, since f is a monoid endomorphism, it follows from the definition of $h_{ij} \ \forall \ y \in \mathbf{F}_q^n$

$$f_i(y) = f_i(\prod_{j=1}^{n} \gamma_j(y_j)) = \prod_{j=1}^{n} f_i(\gamma_j(y_j)) = \prod_{j=1}^{n} h_{ij}(y_j) = \prod_{j=1}^{n} y_j^{F_{ij}}$$

Given 2.1, the converse inclusion $MF_n^n(\mathbf{F}_q) \subseteq End(\prod_{i=1}^{n}(\mathbf{F}_q, \cdot))$ follows from the fact that all the functions contained in $MF_n^n(\mathbf{F}_q)$ obviously satisfy the axioms of a monoid homomorphism. ■

[3] The direct product $\prod_{i=1}^{n}(M, \cdot)$ of a monoid $(M, \cdot)$ is defined as the set $\prod_{i=1}^{n} M$ (cartesian product) endowed with the binary operation $\cdot$ defined as $(x \cdot y)_i := x_i y_i$.

Remark 37 *In the next section we will show that the set $MF_n^n(\mathbf{F}_q)$ can become a monoid in two different ways: The monoid $(MF_n^n(\mathbf{F}_q), \circ)$, where $\circ$ is the composition of endomorphisms; and the commutative monoid $(MF_n^n(\mathbf{F}_q), *)$, where $*$ is the component-wise multiplication defined as*

$$(f * g)_i(x) := f_i(x)g_i(x)$$

Moreover, it turns out that these two binary operations satisfy distributivity properties, i.e. $(MF_n^n(\mathbf{F}_q), *, \circ)$ is a semiring with identity elements with respect to each binary operation.

2.2 Algebraic and graph theoretic formalism

In this section we will introduce the monoid $(MF_n^n(\mathbf{F}_q), \circ)$ of n-dimensional monomial dynamical systems over a finite field $\mathbf{F}_q$, where $\circ$ is the composition of such systems. Furthermore, we will introduce the commutative monoid $(MF_n^n(\mathbf{F}_q), *)$, where $*$ is the component-wise multiplication defined as

$$(f * g)_i(x) := f_i(x)g_i(x)$$

In addition, we will show that these two binary operations satisfy distributivity properties, i.e. $(MF_n^n(\mathbf{F}_q), *, \circ)$ is a semiring with identity elements with respect to each binary operation. Moreover, we will prove that this semiring is isomorphic to a certain semiring of matrices. This result establishes on the one hand, that the composition $f \circ g$ of two monomial dynamical systems f, g is completely captured by the product $F \cdot G$ of their corresponding matrices. On the other hand, it also shows that the component-wise multiplication $f * g$ is completely captured by the sum $F + G$ of the corresponding matrices.

Finally, we will introduce the concept of dependency graph of a monomial dynamical system f and prove that the adjacency matrix of the dependency graph is precisely the matrix F associated with f via the isomorphism mentioned above. This finding allows us to link topological properties of the dependency graph with the dynamics of f. We start with a short step by looking at the exponents of monomial dynamical systems:

As proved in Chapter 1, every function $h : \mathbf{F}_q^n \to \mathbf{F}_q$ is a polynomial function in n variables where no variable appears to a power higher or equal to q. Calculating the composition of a dynamical system $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ with itself, we face the situation where some of the exponents exceed the value $q - 1$ and need to be reduced according to the well-known rule

$$a^q = a \ \forall \ a \in \mathbf{F}_q \tag{2.3}$$

For instance, if we have $q = 3$ and x^{11} then we would write

$$x^{11} = (x^3)^3 x^2 = x^3 x^2 = x x^2 = x^3 = x \ \forall \ x \in \mathbf{F}_3$$

This process can be accomplished systematically if we look at the polynomial $\tau^p \in \mathbf{F}_q[\tau]$[4] (where $p > q$), as described in the Lemma and Definition below. But first we need an auxiliary result:

Lemma 38 *Let $\mathbf{F}_q$ be a finite field and $a \in \mathbb{N}_0$ a nonnegative integer. Then*

$$x^a = 1 \ \forall \ x \in \mathbf{F}_q \backslash \{0\} \Leftrightarrow \exists \ \lambda \in \mathbb{N}_0 : a = \lambda(q-1)$$

Proof. If $a = \lambda(q-1)$ then $x^a = x^{\lambda(q-1)} = (x^{(q-1)})^\lambda = 1 \ \forall \ x \in \mathbf{F}_q \backslash \{0\}$ by (2.3). Now assume $x^a = 1 \ \forall \ x \in \mathbf{F}_q \backslash \{0\}$ and write $a = \alpha(q-1) + s$ with suitable $\alpha \in \mathbb{N}_0$ and $0 \leq s \leq (q-1)$. Then it follows

$$1 = x^a = x^{\lambda(q-1)+s} = x^{\lambda(q-1)} x^s = x^s \ \forall \ x \in \mathbf{F}_q \backslash \{0\}$$

As a consequence, the polynomial $\tau^s - \tau^0 \in \mathbf{F}_q[\tau]$ has $|\mathbf{F}_q| - 1 = q - 1 \geq s = \deg(\tau^s - \tau)$ roots in $\mathbf{F}_q$ and must be therefore of degree $s = q - 1$. Thus $a = (\alpha + 1)(q-1)$. ■

[4] $\mathbf{F}_q[\tau]$ is the ring of polynomials over $\mathbf{F}_q$.

Lemma 39 (and Definition) *Let $\mathbf{F}_q$ be a finite field and $c \in \mathbb{N}_0$ a nonnegative integer. The degree of the (unique) remainder of the polynomial division $\tau^c \div (\tau^q - \tau)$ is called $red_q(c)$. $red_q(c)$ satisfies the following properties*

1. $red_q(red_q(c)) = red_q(c)$
2. $red_q(c) = 0 \Leftrightarrow c = 0$
3. *For $a, b \in \mathbb{N}_0$, $x^a = x^b \ \forall \ x \in \mathbf{F}_q \Leftrightarrow red_q(a) = red_q(b)$*
4. *For $a, b \in \mathbb{N}$, $red_q(a) = red_q(b) \Leftrightarrow \exists \ \alpha \in \mathbb{Z} : a = b + \alpha(q-1)$*

Proof. By the division algorithm there are unique $g, r \in \mathbf{F}_q[\tau]$ with either $r = 0$ or $\deg(r) < \deg(\tau^q - \tau)$ such that

$$\tau^c = g(\tau^q - \tau) + r$$

If we look at the corresponding polynomial functions[5] defined on $\mathbf{F}_q$, it follows by (2.3)

$$x^c = \widetilde{r}(x) \ \forall \ x \in \mathbf{F}_q \tag{2.4}$$

In particular, $r \neq 0$. From the division process it is also clear that r must be a monomial and we conclude $r = \tau^{red_q(c)}$ with $red_q(c) < q$. The first property follows trivially from the fact $red_q(c) < q$. The second property follows immediately from evaluating the equation $x^c = x^{red_q(c)}$ (i.e. equation (2.4)) at the value $x = 0$. The third property is shown as follows: By the division algorithm $\exists_1 \ g_a \ , \ g_b \ , r_a \ , \ r_b \in \mathbf{F}_q[\tau]$ such that

$$\begin{aligned} \tau^a &= g_a(\tau^q - \tau) + r_a = g_a(\tau^q - \tau) + \tau^{red_q(a)} \\ \tau^b &= g_b(\tau^q - \tau) + r_b = g_b(\tau^q - \tau) + \tau^{red_q(b)} \end{aligned} \tag{2.5}$$

From $x^a = x^b \ \forall \ x \in \mathbf{F}_q$ now we have

$$x^{red_q(a)} = x^{red_q(b)} \ \forall x \in \mathbf{F}_q$$

and since $red_q(a), red_q(b) < q$ we get $red_q(a) = red_q(b)$. On the other hand, from $red_q(a) = red_q(b)$ it would follow from equations (2.5)

$$\tau^a - g_a(\tau^q - \tau) = \tau^b - g_b(\tau^q - \tau)$$

and thus by (2.3)

$$x^a = x^b \ \forall \ x \in \mathbf{F}_q$$

Last we prove the fourth claim: If $red_q(a) = red_q(b)$, then by 3. we have

$$x^a = x^b \ \forall \ x \in \mathbf{F}_q$$

Now assume wlog $a \geq b$ and $d := a - b \in \mathbb{N}_0$. Then the last equation can be written as

$$x^b x^d = x^b \ \forall \ x \in \mathbf{F}_q$$

yielding

$$x^d = 1 \ \forall \ x \in \mathbf{F}_q \backslash \{0\}$$

[5] If $r \in \mathbf{F}_q[\tau]$ is a polynomial of degree n, i.e. $r = \sum_{i=0}^{n} a_i \tau^i$, then $\widetilde{r}$ is defined as the polynomial function

$$\begin{aligned} \widetilde{r} \quad &: \quad \mathbf{F}_q \to \mathbf{F}_q \\ x \quad &\mapsto \quad \sum_{i=0}^{n} a_i x^i \end{aligned}$$

By Lemma 38 we have $\exists\ \alpha \in \mathbb{N}_0 : d = \alpha(q-1)$ and therefore $a = b + \alpha(q-1)$ or $b = a - \alpha(q-1)$. Now assume the converse, namely $\exists\ \alpha \in \mathbb{Z} : a = b + \alpha(q-1)$. Assume wlog $\alpha \geq 0$ (otherwise consider $b = a - \alpha(q-1)$). Then we would have

$$\tau^a = \tau^{\alpha(q-1)}\tau^b$$

and thus by Lemma 38

$$x^a = x^b\ \forall\ x \in \mathbf{F}_q \backslash \{0\}$$

Since $a, b > 0$ we also have

$$x^a = x^b\ \forall\ x \in \mathbf{F}_q$$

■

Remark 40 *From the properties above we have* $x^a = x^{red_q(a)}\ \forall\ x \in \mathbf{F}_q$.

When calculating the composition of dynamical systems $f, g : \mathbf{F}_q^n \to \mathbf{F}_q^n$, one needs to add and multiply exponents. Similarly, when calculating the product $f * g$ of dynamical systems one needs to add exponents. Therefore, it is pertinent to formalize this "exponents arithmetic" based on the reduction algorithm described by the previous lemma. Indeed, we can endow the set

$$E_q = \{0, 1, ..., (q-2), (q-1)\}$$

with the algebraic structure of a commutative semiring with identity. This is shown in the theorem below.

Lemma 41 *Let* $\mathbf{F}_q$ *be a finite field and* $a, b \in \mathbb{N}_0$ *nonnegative integers. Then it holds*

$$red_q(ab) = red_q(red_q(a)red_q(b))$$

Proof. We have $\forall\ x \in \mathbf{F}_q$

$$x^{ab} = (x^a)^b = (x^{red_q(a)})^{red_q(b)} = x^{red_q(a)red_q(b)}$$

and by the previous lemma

$$red_q(ab) = red_q(red_q(a)red_q(b))$$

■

Lemma 42 *Let* $\mathbf{F}_q$ *be a finite field and* $a, b \in \mathbb{N}_0$ *nonnegative integers. Then it holds*

$$red_q(a+b) = red_q(red_q(a) + red_q(b))$$

Proof. By the division algorithm $\exists_1\ g_a\ ,\ g_b\ ,\ g_{a+b}\ ,\ r_a\ ,\ r_b\ ,\ r_{a+b} \in \mathbf{F}_q[\tau]$ such that

$$\begin{aligned}
\tau^a &= g_a(\tau^q - \tau) + r_a = g_a(\tau^q - \tau) + \tau^{red_q(a)} \\
\tau^b &= g_b(\tau^q - \tau) + r_b = g_b(\tau^q - \tau) + \tau^{red_q(b)} \\
\tau^{a+b} &= g_{a+b}(\tau^q - \tau) + r_{a+b} = g_{a+b}(\tau^q - \tau) + \tau^{red_q(a+b)}
\end{aligned}$$

From the first two equations follows

$$\tau^{a+b} = g_a g_b(\tau^q - \tau)^2 + g_a r_b(\tau^q - \tau) + r_a g_b(\tau^q - \tau) + \tau^{red_q(a)+red_q(b)}$$

Applying the division algorithm to $\tau^{red_q(a)+red_q(b)}$ we can say $\exists_1\ g_r\ ,\ r_r \in \mathbf{F}_q[\tau]$ such that

$$\begin{aligned}
\tau^{a+b} &= g_a g_b(\tau^q - \tau)^2 + g_a r_b(\tau^q - \tau) + r_a g_b(\tau^q - \tau) + g_r(\tau^q - \tau) + r_r \\
&= (g_a g_b(\tau^q - \tau) + g_a r_b + r_a g_b + g_r)(\tau^q - \tau) + \tau^{red_q(red_q(a)+red_q(b))}
\end{aligned}$$

From the uniqueness of quotient and remainder it follows

$$\tau^{red_q(a+b)} = \tau^{red_q(red_q(a)+red_q(b))}$$

and consequently

$$red_q(a+b) = red_q(red_q(a) + red_q(b))$$

■

Theorem 43 (and Definition) *Let* $\mathbf{F}_q$ *be a finite field. The set*

$$E_q = \{0, 1, ..., (q-2), (q-1)\} \subset \mathbb{Z}$$

together with the operations of addition $a \oplus b := red_q(a+b)$ *and multiplication* $a \bullet b := red_q(ab)$ *is a commutative semiring with identity* 1*. We call this commutative semiring the* exponents semiring *of the field* $\mathbf{F}_q$*.*

Proof. First we show that E_q is a commutative monoid with respect to the addition $\oplus$. The reduction modulo the ideal $\langle \tau^q - \tau \rangle$ ensures that E_q is closed under this operation. Additive commutativity follows trivially from the definition. The associativity is shown using Lemma 42 and the fact that $c \in E_q \Leftrightarrow c = red_q(c)$

$$\begin{aligned}
(a \oplus b) \oplus c &= red_q(a+b) \oplus c \\
&= red_q(red_q(a+b) + c) \\
&= red_q(red_q(a+b) + red_q(c)) \\
&= red_q((a+b) + c) \\
&= red_q(a + (b+c)) \\
&= red_q(red_q(a) + red_q(b+c)) \\
&= red_q(a + red_q(b+c)) \\
&= a \oplus red_q(b+c) \\
&= a \oplus (b \oplus c)
\end{aligned}$$

It is trivial to see that 0 is the additive identity element. E_q is also a commutative monoid with respect to the multiplication $\bullet$: The reduction modulo the ideal $\langle \tau^q - \tau \rangle$ ensures that E_q is closed under this operation. Multiplicative commutativity as well as the fact that 1 is the multiplicative identity follow trivially from the definition. The associativity is shown using Lemma 41 and the fact that $c \in E_q \Leftrightarrow c = red_q(c)$

$$\begin{aligned}
(a \bullet b) \bullet c &= red_q(red_q(ab)c) \\
&= red_q(red_q(ab) red_q c) \\
&= red_q((ab)c) \\
&= red_q(a(bc)) \\
&= red_q(red_q(a) red_q(bc)) \\
&= red_q(a red_q(bc)) \\
&= a \bullet red_q(bc) \\
&= a \bullet (b \bullet c)
\end{aligned}$$

The distributivity is shown as follows

$$\begin{aligned} a \bullet (b \oplus c) &= red_q(a(b \oplus c)) \\ &= red_q(ared_q(b+c)) \\ &= red_q(red_q(a)red_q(b+c)) \\ &= red_q(a(b+c)) \\ &= red_q(ab+ac) \\ &= red_q(red_q(ab)+red_q(ac)) \\ &= red_q(ab) \oplus red_q(ac) \\ &= (a \bullet b) \oplus (a \bullet c) \end{aligned}$$

■

Remark 44 *From the point of view of the solvability of linear equations it would be convenient to have the commutative semiring E_q as a subsemiring of a commutative ring or integral domain. The straightforward extension of the set E_q is to introduce negative powers that can be defined on a nonzero field element $x \in \mathbf{F}_q \backslash \{0\}$ as $x^{-p} := \overline{x}^p$, where $p \in \mathbb{N}_0$ and $\overline{x} \in \mathbf{F}_q$ denotes the multiplicative inverse of x. The exponent reduction according to (2.3) is naturally defined as $red_q(-p) := -red_q(p)$. Unfortunately, this natural extension of the set E_q does not yield a ring, because the property of Lemma 42 does not hold for negative powers. For instance, we could try to extend the set E_2 by including the negative exponent -1. The table of pairwise addition would look like*

$\oplus$	0	1	-1
0	0	1	-1
1	1	1	0
-1	-1	0	χ

Now, no matter which of the three values $-1, 0, 1$ we choose for χ, (the result of the operation $(-1) \oplus (-1)$), we end up with an algebraic structure that transgresses associativity:

$$(1 \oplus (-1)) \oplus (-1) = 0 \oplus (-1) = -1 \neq 1 \oplus ((-1) \oplus (-1)) = \begin{cases} 1 \oplus -1 = 0 \\ 1 \oplus 0 = 1 \\ 1 \oplus 1 = 1 \end{cases}$$

That such an extension is not possible as a matter of principle is shown with help of the Grothendieck construction (see, for instance, §7 of [64]): Assume there is a semiring isomorphism

$$\iota : (E_q, \oplus, \cdot) \to R$$

of the semiring $(E_q, \oplus, \cdot)$ into a ring $(R, +, *)$. This semiring homomorphism induces a monoid[6] *homomorphism*

$$\iota : (E_q, \oplus) \to (R, +)$$

from the underlying commutative monoid $(E_q, \oplus)$ into the Abelian group $(R, +)$. Now, as shown in the Appendix, there is an Abelian group $G((E_q, \oplus))$ (the so called Grothendieck group), such that every monoid homomorphism

$$\kappa : (E_q, \oplus) \to G((E_q, \oplus))$$

from the additive monoid $(E_q, \oplus)$ into the group $G((E_q, \oplus))$ must be noninjective. *Furthermore, according to the Grothendieck construction, there exists a monoid homomorphism*

$$\gamma : (E_q, \oplus) \to G((E_q, \oplus))$$

[6] $\iota(0) = \iota(0+0) = \iota(0) + \iota(0) \Rightarrow \iota(0) = 0_R \in R$, since $\iota(0) \in R$ has an inverse.

having the following universal property: If $f : (E_q, \oplus) \to U$ is a monoid homomorphism into an Abelian group U, then there is a unique group homomorphism $f_* : G((E_q, \oplus)) \to U$ such that the following diagram commutes

$$\begin{array}{ccc} (E_q,\oplus) & \longrightarrow_{\gamma} & G((E_q,\oplus)) \\ & {\scriptstyle f}\searrow & \downarrow_{f_*} \\ & & U \end{array}$$

(for the proof, see §7 of [64]). Now, if we replace U by the group $(R, +)$ and f by $\iota : (E_q, \oplus) \to (R, +)$, we obtain

$$\begin{array}{ccc} (E_q,\oplus) & \longrightarrow_{\gamma} & G((E_q,\oplus)) \\ & {\scriptstyle \iota}\searrow & \downarrow_{\iota_*} \\ & & (R,+) \end{array}$$

Thus, we can write $\iota : (E_q, \oplus) \to (R, +)$ as $\iota = \iota_* \circ \gamma$. Since γ is not injective (see Appendix), ι cannot be injective either.

Lemma 45 *Let $n \in \mathbb{N}$ be a natural number, $\mathbf{F}_q$ be a finite field and E_q the exponents semiring of $\mathbf{F}_q$. The set $M(n \times n; E_q)$ of $n \times n$ quadratic matrices with entries in the semiring E_q together with the operation $+$ of matrix addition over E_q is a commutative monoid.*

Proof. The matrix addition is defined in terms of the operation $\oplus$ on the matrix entries, i.e. for two matrices $A, B \in M(n \times n; E_q)$ we define $C := A + B$ as

$$C_{ij} := A_{ij} \oplus B_{ij}$$

Now the claim follows directly from the previous theorem and the fact that the zero matrix 0 constitutes the identity element. ■

Lemma 46 *Let $n \in \mathbb{N}$ be a natural number, $\mathbf{F}_q$ be a finite field and E_q the exponents semiring of $\mathbf{F}_q$. The set $M(n \times n; E_q)$ of $n \times n$ quadratic matrices with entries in the semiring E_q together with the operation $\cdot$ of matrix multiplication over E_q is a monoid.*

Proof. The matrix multiplication $\cdot$ is defined in terms of the operations $\oplus$ and $\bullet$ on the matrix entries, therefore $M(n \times n; E_q)$ is closed under multiplication. To show the associativity, consider $A, B, C \in M(n \times n; E_q)$. According to the definition of matrix product we have for $D := A \cdot B$, $E := (A \cdot B) \cdot C$, $F := B \cdot C$ and $G := A \cdot (B \cdot C)$

$$D_{ij} = A_{i1} \bullet B_{1j} \oplus A_{i2} \bullet B_{2j} \oplus ... \oplus A_{in} \bullet B_{nj}$$

and therefore

$$\begin{aligned} E_{kl} &= D_{k1} \bullet C_{1l} \oplus D_{k2} \bullet C_{2l} \oplus ... \oplus D_{kn} \bullet C_{nl} \\ &= (A_{k1} \bullet B_{11} \oplus A_{k2} \bullet B_{21} \oplus ... \oplus A_{kn} \bullet B_{n1}) \bullet C_{1l} \oplus ... \\ &\quad ... \oplus (A_{k1} \bullet B_{1n} \oplus A_{k2} \bullet B_{2n} \oplus ... \oplus A_{kn} \bullet B_{nn}) \bullet C_{nl} \\ &= ((A_{k1} \bullet B_{11}) \bullet C_{1l} \oplus ... \oplus (A_{kn} \bullet B_{n1}) \bullet C_{1l}) \oplus ... \\ &\quad ... \oplus ((A_{k1} \bullet B_{1n}) \bullet C_{nl} \oplus ... \oplus (A_{kn} \bullet B_{nn}) \bullet C_{nl}) \\ &= (A_{k1} \bullet (B_{11} \bullet C_{1l}) \oplus ... \oplus A_{kn} \bullet (B_{n1} \bullet C_{1l})) \oplus ... \\ &\quad ... \oplus (A_{k1} \bullet (B_{1n} \bullet C_{nl}) \oplus ... \oplus A_{kn} \bullet (B_{nn} \bullet C_{nl})) \\ &= A_{k1} \bullet (B_{11} \bullet C_{1l} \oplus ... \oplus B_{1n} \bullet C_{nl}) + A_{k2} \bullet (B_{21} \bullet C_{1l} \oplus ... \oplus B_{2n} \bullet C_{nl}) \oplus ... \\ &\quad ... \oplus A_{kn} \bullet (B_{n1} \bullet C_{1l} \oplus ... \oplus B_{nn} \bullet C_{nl}) \\ &= A_{k1} \bullet F_{1l} + A_{k2} \bullet F_{2l} \oplus ... \oplus A_{kn} \bullet F_{nl} = G_{kl} \end{aligned}$$

The identity element is obviously the unit matrix I. ■

Remark 47 (and Definition) *Since the entries for the matrix product $D = A \cdot B$ are defined as*

$$D_{ij} = A_{i1} \bullet B_{1j} \oplus A_{i2} \bullet B_{2j} \oplus ... \oplus A_{in} \bullet B_{nj}$$

according to the definitions of the operations $\bullet$ and $\oplus$ we can write

$$\begin{aligned} D_{ij} &= red_q(A_{i1}B_{1j}) \oplus red_q(A_{i2}B_{2j}) \oplus ... \oplus red_q(A_{in}B_{nj}) \\ &= red_q(red_q(A_{i1}B_{1j}) + red_q(A_{i2}B_{2j}) + ... + red_q(A_{in}B_{nj})) \end{aligned}$$

Now, by Lemma 42 we have

$$D_{ij} = red_q(A_{i1}B_{1j} + A_{i2}B_{2j} + ... + A_{in}B_{nj})$$

As a consequence, if we define the following reduction operation for matrices with nonnegative integer entries

$$\begin{aligned} mred_q &: M(n \times n; \mathbb{N}_0) \to M(n \times n; E_q) \\ A_{ij} &\mapsto red_q(A_{ij}) \end{aligned}$$

then the following property holds for $U, V \in M(n \times n; \mathbb{N}_0)$ and $W := UV \in M(n \times n; \mathbb{N}_0)$

$$\begin{aligned} mred_q(W)_{ij} &= red_q(W_{ij}) \\ &= red_q(U_{i1}V_{1j} + ... + U_{in}V_{nj}) \\ &= red_q(red_q(U_{i1}V_{1j}) + ... + red_q(U_{in}V_{nj})) \\ &= red_q(U_{i1}V_{1j}) \oplus ... \oplus red_q(U_{in}V_{nj}) \\ &= red_q(red_q(U_{i1})red_q(V_{1j})) \oplus ... \oplus red_q(red_q(U_{in})red_q(V_{nj})) \\ &= red_q(U_{i1}) \bullet red_q(V_{1j}) \oplus ... \oplus red_q(U_{in}) \bullet red_q(V_{nj}) \\ &= (mred_q(U)mred_q(V))_{ij} \end{aligned}$$

In other words

$$mred_q(UV) = mred_q(U) \cdot mred_q(V)$$

It can be easily shown that $M(n \times n; \mathbb{N}_0)$ is a monoid and $mred_q : M(n \times n; \mathbb{N}_0) \to M(n \times n; E_q)$ a monoid homomorphism. In addition, by 2. of Lemma 39 we can conclude

$$mred_q(A) = 0 \Leftrightarrow A = 0 \tag{2.6}$$

Theorem 48 *Let $n \in \mathbb{N}$ be a natural number, $\mathbf{F}_q$ be a finite field and E_q the exponents semiring of $\mathbf{F}_q$. Then $(M(n \times n; E_q), +, \cdot)$ is a semiring with identity elements with respect to each binary operation.*

Proof. Given the two previous lemmas, we only need to show that the two binary operations satisfy distributivity properties: Consider $A, C, D \in M(n \times n;\ E_q)$. According to the definitions we have for $B := C + D$

$$\begin{aligned} (A \cdot (C + D))_{ij} &= A_{i1} \bullet B_{1j} \oplus A_{i2} \bullet B_{2j} \oplus ... \oplus A_{in} \bullet B_{nj} \\ &= A_{i1} \bullet (C_{1j} \oplus D_{1j}) \oplus A_{i2} \bullet (C_{2j} \oplus D_{2j}) \oplus ... \oplus A_{in} \bullet (C_{nj} \oplus D_{nj}) \\ &= A_{i1} \bullet C_{1j} \oplus A_{i1} \bullet D_{1j} \oplus A_{i2} \bullet C_{2j} \oplus A_{i2} \bullet D_{2j} \oplus ... \oplus A_{in} \bullet C_{nj} \oplus A_{in} \bullet D_{nj} \\ &= (A_{i1} \bullet C_{1j} \oplus A_{i2} \bullet C_{2j} \oplus ... \oplus A_{in} \bullet C_{nj}) \oplus (A_{i1} \bullet D_{1j} \oplus A_{i2} \bullet D_{2j} \oplus ... \oplus A_{in} \bullet D_{nj}) \\ &= (A \cdot C)_{ij} + (A \cdot D)_{ij} \end{aligned}$$

where the distributivity properties of E_q where used. The right-distributivity is shown analogously. ■

We recall the following definition from the previous section:

Definition 49 *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers. The set*

$$MF_m^n(\mathbf{F}_q) := \{f : \mathbf{F}_q^m \to \mathbf{F}_q^n \mid \exists\ F \in M(n \times m; E_q) : f_i(x) := x_1^{F_{i1}} ... x_m^{F_{im}}\ \forall\ x \in \mathbf{F}_q^m\}$$

is called the set of n-dimensional monomial mappings in m variables.

Lemma 50 *Let $\mathbf{F}_q$ be a finite field and $n, m, r \in \mathbb{N}$ natural numbers. Furthermore, let $f \in MF_n^m(\mathbf{F}_q)$ and $g \in MF_m^r(\mathbf{F}_q)$ with*

$$\begin{aligned} f_i(x) &= x_1^{F_{i1}} ... x_n^{F_{in}}\ \forall\ x \in \mathbf{F}_q^n,\ i = 1, ..., m \\ g_j(x) &= x_1^{G_{j1}} ... x_m^{G_{jm}}\ \forall\ x \in \mathbf{F}_q^n,\ j = 1, ..., r \end{aligned}$$

where $F \in M(m \times n; E_q)$ and $G \in M(r \times m;\ E_q)$. Then for their composition $g \circ f : \mathbf{F}_q^n \to \mathbf{F}_q^r$ it holds

$$(g \circ f)_k(x) = \prod_{j=1}^{n} x_j^{\ red_q(\sum_{l=1}^{m} G_{kl} F_{lj})}\ \forall\ x \in \mathbf{F}_q^n,\ k \in \{1, ..., r\}$$

Proof. From the definition it follows for every $k \in \{1, ..., r\}$

$$(g \circ f)_k(x) = \prod_{l=1}^{m} (f_l(x))^{G_{kl}} = \prod_{l=1}^{m} (\prod_{j=1}^{n} x_j^{F_{lj}})^{G_{kl}}$$

For a fixed but arbitrary $m \in \mathbb{N}$ we will prove the claim using induction on the number of variables n of the mapping $g \circ f$. For $n = 1$ we have

$$(g \circ f)_k(x) = \prod_{l=1}^{m} (x_1^{F_{l1}})^{G_{kl}} = \prod_{l=1}^{m} x_1^{G_{kl} F_{l1}} = x_1^{\sum_{l=1}^{m} G_{kl} F_{l1}} = x_1^{\ red_q(\sum_{l=1}^{m} G_{kl} F_{l1})}$$

(see Remark 40), thus the claim holds for 1 variable. Now we consider the case of $n+1$ variables:

$$\begin{aligned} (g \circ f)_k(x) &= \prod_{l=1}^{m} (\prod_{j=1}^{n+1} x_j^{F_{lj}})^{G_{kl}} \\ &= \prod_{l=1}^{m} (x_{(n+1)}^{F_{l(n+1)}} \prod_{j=1}^{n} x_j^{F_{lj}})^{G_{kl}} \\ &= \prod_{l=1}^{m} \left(x_{(n+1)}^{G_{kl} F_{l(n+1)}} (\prod_{j=1}^{n} x_j^{F_{lj}})^{G_{kl}} \right) \\ &= \prod_{l=1}^{m} (x_{(n+1)}^{G_{kl} F_{l(n+1)}}) \prod_{l=1}^{m} (\prod_{j=1}^{n} x_j^{F_{lj}})^{G_{kl}} \end{aligned}$$

and by induction hypothesis

$$\begin{aligned}
&= x_{(n+1)}^{\sum_{l=1}^{m} G_{kl}F_{l(n+1)}} \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} G_{kl}F_{lj})} \\
&= x_{(n+1)}^{\sum_{l=1}^{m} G_{kl}F_{l(n+1)}} \prod_{j=1}^{n} x_j^{\sum_{l=1}^{m} G_{kl}F_{lj}} \\
&= \prod_{j=1}^{n+1} x_j^{\sum_{l=1}^{m} G_{kl}F_{lj}} \\
&= \prod_{j=1}^{n+1} x_j^{red_q(\sum_{l=1}^{m} G_{kl}F_{lj})}
\end{aligned}$$

■

Remark 51 (and Lemma) *If we generalize the matrix multiplication defined on the monoid $M(n \times n;\, E_q)$ for matrices $F \in M(m \times n; E_q)$ and $G \in M(r \times m;\, E_q)$ then we can write*

$$(g \circ f)_k(x) = \prod_{j=1}^{n} x_j^{(G \cdot F)_{kj}} \;\forall\; x \in \mathbf{F}_q^n,\; k \in \{1, ..., r\}$$

To see this, apply the Lemmas 42 and 50 as well as the definitions of $\oplus$ and $\bullet$ to $\prod_{j=1}^{n} x_j^{(G \cdot F)_{kj}}$:

$$\begin{aligned}
\prod_{j=1}^{n} x_j^{(G \cdot F)_{kj}} &= \prod_{j=1}^{n} x_j^{(G_{k1} \bullet F_{1j} \oplus ... \oplus G_{km} \bullet F_{mj})} \\
&= \prod_{j=1}^{n} x_j^{red_q(G_{k1}F_{1j}) \oplus ... \oplus red_q(G_{km}F_{mj})} \\
&= \prod_{j=1}^{n} x_j^{red_q(red_q(G_{k1}F_{1j}) + ... + red_q(G_{km}F_{mj}))} \\
&= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} G_{kl}F_{lj})} \\
&= (g \circ f)_k(x)
\end{aligned}$$

Theorem 52 *Let $\mathbf{F}_q$ be a finite field. The set*

$$MF_n^n(\mathbf{F}_q) := \{f : \mathbf{F}_q^n \to \mathbf{F}_q^n \mid \exists\; F \in M(n \times n; E_q) : f_i(x) := x_1^{F_{i1}} ... x_n^{F_{in}} \;\forall\; x \in \mathbf{F}_q^n\}$$

of all monomial dynamical systems over $\mathbf{F}_q$ together with the composition $\circ$ of mappings is a monoid.

Proof. By Lemma 50 the set $MF_n^n(\mathbf{F}_q)$ is closed under composition. Composition of mappings is trivially associative. The identity function

$$\begin{aligned}
Id \quad &: \quad \mathbf{F}_q^n \to \mathbf{F}_q^n \\
x \quad &\mapsto \quad x
\end{aligned}$$

is a monomial system and is therefore the identity element of the monoid $(MF_n^n(\mathbf{F}_q), \circ)$. ■

Lemma 53 (and Definition) *Let $\mathbf{F}_q$ be a finite field and $n, m, r \in \mathbb{N}$ natural numbers. Furthermore, let $f \in MF_n^n(\mathbf{F}_q)$ and $g \in MF_n^n(\mathbf{F}_q)$ with*

$$\begin{aligned} f_i(x) &= x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n, \ i = 1, ..., n \\ g_j(x) &= x_1^{G_{j1}} ... x_n^{G_{jn}} \ \forall \ x \in \mathbf{F}_q^n, \ j = 1, ..., n \end{aligned}$$

where $F \in M(n \times n; E_q)$ and $G \in M(n \times n;\ E_q)$. Then for their component-wise multiplication $g * f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ defined as

$$(g * f)_i(x) := g_i(x) f_i(x)$$

it holds

$$(g * f)_i(x) = \prod_{j=1}^{n} x_j^{red_q(G_{ij}+F_{ij})} \ \forall \ x \in \mathbf{F}_q^n, \ i \in \{1, ..., n\}$$

Proof. The claim follows directly from the exponents rules on the finite field $\mathbf{F}_q$ and Remark 40. ■

Remark 54 *From the definition of $\oplus$ it follows easily*

$$(g * f)_i(x) = \prod_{j=1}^{n} x_j^{(G+F)_{ij}} \ \forall \ x \in \mathbf{F}_q^n, \ i \in \{1, ..., n\}$$

Theorem 55 *Let $\mathbf{F}_q$ be a finite field. The set*

$$MF_n^n(\mathbf{F}_q) := \{f : \mathbf{F}_q^n \to \mathbf{F}_q^n \mid \exists \ F \in M(n \times n; E_q) : f_i(x) := x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n\}$$

of all monomial dynamical systems over $\mathbf{F}_q$ together with the multiplication $*$ of mappings is a commutative monoid.

Proof. By Lemma 53 the set $MF_n^n(\mathbf{F}_q)$ is closed under multiplication. Now the claim follows from the commutativity and associativity of the multiplication in $\mathbf{F}_q$ and the fact that the *one function*

$$\begin{aligned} \mathbf{1} &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto (1, ..., 1)^t \end{aligned}$$

is a monomial system and obviously constitutes the identity element of the monoid $(MF_n^n(\mathbf{F}_q), *)$. ■

Theorem 56 *Let $n \in \mathbb{N}$ be a natural number and $\mathbf{F}_q$ be a finite field. Then $(MF_n^n(\mathbf{F}_q), *, \circ)$ is a semiring with identity elements with respect to each binary operation.*

Proof. Given Theorems 55 and 52, we only need to show that the two binary operations satisfy distributivity properties: Consider $f, g, h \in MF_n^n(\mathbf{F}_q)$ with

$$\begin{aligned} f_i(x) &= x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n, \ i = 1, ..., n \\ g_j(x) &= x_1^{G_{j1}} ... x_n^{G_{jn}} \ \forall \ x \in \mathbf{F}_q^n, \ j = 1, ..., n \\ h_k(x) &= x_1^{H_{k1}} ... x_n^{H_{kn}} \ \forall \ x \in \mathbf{F}_q^n, \ k = 1, ..., n \end{aligned}$$

where $F, G, H \in M(m \times n; E_q)$. According to the definitions and by Lemmas 41, 42 as well as Remark 40 we have

$$\begin{aligned}
(g \circ (h * f))_k &= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} G_{kl}(H_{lj} \oplus F_{lj}))} \\
&= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} red_q(G_{kl}) red_q(H_{lj}+F_{lj}))} \\
&= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} G_{kl}(H_{lj}+F_{lj}))} \\
&= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} G_{kl}H_{lj}+G_{kl}F_{lj})} \\
&= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} red_q(G_{kl}H_{lj})+red_q(G_{kl}F_{lj}))} \\
&= \prod_{j=1}^{n} x_j^{red_q(\sum_{l=1}^{m} red_q(G_{kl}H_{lj})+\sum_{l=1}^{m} red_q(G_{kl}F_{lj}))} \\
&= \prod_{j=1}^{n} x_j^{\sum_{l=1}^{m} red_q(G_{kl}H_{lj})+\sum_{l=1}^{m} red_q(G_{kl}F_{lj})} \\
&= \prod_{j=1}^{n} (x_j^{\sum_{l=1}^{m} red_q(G_{kl}H_{lj})} x_j^{\sum_{l=1}^{m} red_q(G_{kl}F_{lj})}) \\
&= \prod_{j=1}^{n} (x_j^{\sum_{l=1}^{m} red_q(G_{kl}H_{lj})}) \prod_{j=1}^{n} (x_j^{\sum_{l=1}^{m} red_q(G_{kl}F_{lj})}) \\
&= \prod_{j=1}^{n} (x_j^{red_q(\sum_{l=1}^{m} red_q(G_{kl}H_{lj}))}) \prod_{j=1}^{n} (x_j^{red_q(\sum_{l=1}^{m} red_q(G_{kl}F_{lj}))}) \\
&= \prod_{j=1}^{n} (x_j^{red_q(\sum_{l=1}^{m} G_{kl}H_{lj})}) \prod_{j=1}^{n} (x_j^{red_q(\sum_{l=1}^{m} G_{kl}F_{lj})}) \\
&= (g \circ h)_k(x)(g \circ f)_k(x)
\end{aligned}$$

This shows

$$(g \circ (h * f)) = (g \circ h) * (g \circ f)$$

The right-distributivity is shown analogously. ■

Theorem 57 *The monoids* $(M(n \times n;\, E_q), \cdot)$ *and* $(MF_n^n(\mathbf{F}_q), \circ)$ *are isomorphic.*

Proof. From the definition of $MF_n^n(\mathbf{F}_q)$ it is clear that the mapping

$$\begin{aligned}
\Psi &: M(n \times n; E_q) \to MF_n^n(\mathbf{F}_q) \\
G &\mapsto \Psi(G)
\end{aligned}$$

such that

$$\Psi(G)_i(x) := x_1^{G_{i1}} ... x_n^{G_{in}} \text{ for } i = 1, ..., n$$

is a bijection. Moreover, $\Psi(I) = id$. In addition, by Remark 51 it follows easily that

$$\Psi(F \cdot G) = \Psi(F) \circ \Psi(G) \ \forall \ F, G \in M(n \times n; E_q)$$

■

Corollary 58 *The semirings* $(M(n \times n;\ E_q), +, \cdot)$ *and* $(MF_n^n(\mathbf{F}_q), *, \circ)$ *are isomorphic.*

Proof. Consider the mapping

$$\begin{array}{rcl} \Psi & : & M(n \times n; E_q) \to MF_n^n(\mathbf{F}_q) \\ G & \mapsto & \Psi(G) \end{array}$$

such that

$$\Psi(G)_i(x) := x_1^{G_{i1}} ... x_n^{G_{in}} \text{ for } i = 1, ..., n$$

defined in the previous proof. By Remark 54, Ψ also satisfies

$$\Psi(F + G) = \Psi(F) * \Psi(G) \ \forall \ F, G \in M(n \times n; E_q)$$

■

Remark 59 (and Definition) *For a given monomial dynamical system* $f \in MF_n^n(\mathbf{F}_q)$ *the matrix* $F := \Psi^{-1}(f)$ *is called the* corresponding matrix *of the system* f*. For a matrix power in the monoid* $M(n \times n; E_q)$ *we use the notation* $F^{\cdot m}$*. By induction it can be easily shown*

$$\Psi^{-1}(f^m) = F^{\cdot m}$$

Remark 60 (and Definition) *The image of the* $n \times n$ *zero matrix* $0 \in M(n \times n; E_q)$ *under the isomorphism* Ψ *has the property*

$$\Psi(0)(x)_i = 1 \ \forall \ x \in \mathbf{F}_q^n$$

we call this monomial function the one function $\mathbf{1} := \Psi(0)$.

Remark 61 *The image of the unit matrix* $I \in M(n \times n; E_q)$ *under the isomorphism* Ψ *has the property*

$$\Psi(I)(x)_i = x_i \ \forall \ x \in \mathbf{F}_q^n$$

i.e. $\Psi(I) = id$.

Now we turn our attention to some graph theoretic considerations. We recall the following definition from Chapter 1:

Definition 62 *Let* M *be a nonempty finite set. Furthermore, let* $n := |M|$ *be the cardinality of* M*. A* numeration *of the elements of* M *is a bijective mapping*

$$f : M \to \{1, ..., n\}$$

Given a numeration f *of the set* M *we write*

$$M = \{f_1, ..., f_n\}$$

where the unique element $x \in M$ *with the property* $f(x) = i \in \{1, ..., n\}$ *is denoted as* f_i*.*

Definition 63 *Let $f \in MF_n^n(\mathbf{F}_q)$ be a monomial dynamical system and $G = (V_G,\, E_G,\, \pi_G)$ a digraph (recall Definition 3) with vertex set V_G of cardinality $|V_G| = n$. Furthermore, let $F := \Psi^{-1}(f)$ be the corresponding matrix of f. The digraph G is called* dependency graph *of f iff a numeration $a : M \to \{1, ..., n\}$ of the elements of V_G exists such that $\forall\, i, j \in \{1, ..., n\}$ there are **exactly** F_{ij} directed edges $a_i \to a_j$ in the set E_G, i.e.*

$$\left|\pi_G^{-1}((a_i, a_j))\right| = F_{ij}$$

Remark 64 *It is easy to show that if G and H are dependency graphs of f then G and H are isomorphic. In this sense we speak from* the *dependency graph of f and denote it by $G_f = (V_f,\, E_f,\, \pi_f)$.*

We recall the following two definitions from Chapter 1:

Definition 65 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a digraph. Two vertices $a, b \in V_G$ are called* connected *if there is a $t \in \mathbb{N}_0$ and (not necessarily different) vertices $v_1, ..., v_t \in V_G$ such that*

$$a \to v_1 \to v_2 \to ... \to v_t \to b$$

In this situation we write $a \rightsquigarrow_s b$, where s is the number of directed edges involved in the sequence *from a to b (in this case $s = t + 1$). Two sequences $a \rightsquigarrow_s b$ of the same length are considered different if the directed edges involved are different or the order at which they appear is different, even if the visited vertices are the same. As a convention, a single vertex $a \in V_G$ is always connected to itself $a \rightsquigarrow_0 a$ by an empty sequence of length* 0.

Definition 66 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a digraph and $a, b \in V_G$ two vertices. A sequence $a \rightsquigarrow_s b$*

$$a \to v_1 \to v_2 \to ... \to v_t \to b$$

is called a path*, if no vertex v_i is visited more than once. If $a = b$, but no other vertex is visited more than once, $a \rightsquigarrow_s b$ is called a* closed path*.*

Definition 67 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a digraph. Two vertices $a, b \in V_G$ are called* strongly connected *if there are natural numbers $s, t \in \mathbb{N}$ such that*

$$a \rightsquigarrow_s b \text{ and } b \rightsquigarrow_t a$$

In this situation we write $a \rightleftharpoons b$.

Theorem 68 (and Definition) *Let $G = (V_G,\, E_G,\, \pi_G)$ be a digraph. $\rightleftharpoons$ is an equivalence relation on V_G called* strong equivalence*. The equivalence class of any vertex $a \in V_G$ is called a* strongly connected component *and denoted by $\overleftrightarrow{a} \subseteq V_G$.*

Proof. Due to the convention $a \rightsquigarrow_0 a$ the relation $\rightleftharpoons$ is reflexive. Symmetry follows trivially from the definition of $\rightleftharpoons$. Transitivity follows from

$$\begin{aligned} a &\rightleftharpoons b \text{ and } b \rightleftharpoons c \\ &\Leftrightarrow\ a \rightsquigarrow_s b \text{ and } b \rightsquigarrow_t a \text{ and } b \rightsquigarrow_u c \text{ and } c \rightsquigarrow_v b \\ &\Rightarrow\ a \rightsquigarrow_{s+u} c \text{ and } c \rightsquigarrow_{v+t} a \\ &\Leftrightarrow\ a \rightleftharpoons c \end{aligned}$$

■

Definition 69 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a digraph and $a \in V_G$ one of its vertices. The* strongly connected component *$\overleftrightarrow{a} \subseteq V_G$ is called* trivial *iff $\overleftrightarrow{a} = \{a\}$ and there is no edge $a \to a$ in E_G.*

Definition 70 *Let $G = (V_G, E_G, \pi_G)$ be a digraph with vertex set V_G of cardinality $|V_G| = n$ and $V_G = \{a_1, ..., a_n\}$ a numeration of the elements of V_G. The matrix $A \in M(n \times n; \mathbb{N}_0)$ whose entries are defined as*

$$A_{ij} := \text{number of edges } a_i \to a_j \text{ contained in } E_G$$

for $i, j = 1, ..., n$ is called adjacency matrix *of G with the numeration a.*

Theorem 71 *Let $G = (V_G, E_G, \pi_G)$ be a digraph with vertex set V_G of cardinality $|V_G| = n$ and $V_G = \{a_1, ..., a_n\}$ a numeration of the elements of V_G. Furthermore, let $A \in M(n \times n; \mathbb{N}_0)$ be its adjacency matrix (with the numeration a), $m \in \mathbb{N}$ a natural number and $B := A^m \in M(n \times n; \mathbb{N}_0)$ the mth power of A. Then $\forall\ i, j \in \{1, ..., n\}$ the entry B_{ij} of B is equal to the number of different sequences $a_i \rightsquigarrow_m a_j$ of length m.*

Proof. We prove the claim using induction on m. For $m = 1$ the claim holds due to the definition of adjacency matrix. Now assume the claim holds for mth power of A and consider the $(m+1)$th power of A :

$$C := A^{m+1} = AA^m$$

Since the entry C_{ij}, $i, j \in \{1, ..., n\}$ is computed as

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \tag{2.7}$$

and every sequence $a_i \rightsquigarrow_{m+1} a_j$ necessarily uses as the first edge an edge connecting a_i with one of its neighbors a_k, the expression (2.7) indeed counts all possible different sequences $a_i \rightsquigarrow_{m+1} a_j$ of length $m + 1$. ■

Remark 72 *Let $f \in MF_n^n(\mathbf{F}_q)$ be a monomial dynamical system. Furthermore, let $G_f = (V_f, E_f, \pi_f)$ the dependency graph of f and $V_f = \{a_1, ..., a_n\}$ the associated numeration of the elements of V_f. Then, according to the definition of dependency graph, $F := \Psi^{-1}(f)$ (the corresponding matrix of f) is precisely the adjacency matrix of G_f with the numeration a. Now, by Remarks 59 and 47 we can conclude*

$$\Psi^{-1}(f^m) = mred_q(F^m) \tag{2.8}$$

2.3 Characterization of fixed point systems

The results proved in the previous section allow us to link topological properties of the dependency graph with the dynamics of f. We will exploit this feature in this subsection to prove some characterizations of fixed point systems stated in terms of connectedness properties of the dependency graph. In the course of these investigations we will identify a class of monomial dynamical systems, namely the $(q-1)$*-fold redundant monomial systems* (to be defined below), that allows for a very satisfying characterization of fixed point systems inside the class. A trivial example of $(q-1)$-fold redundant systems are the Boolean systems, (i.e. monomial systems $f \in MF_n^n(\mathbf{F}_2)$).

Theorem 73 *Let $\mathbf{F}_q$ be a finite field and $f \in MF_n^n(\mathbf{F}_q)$ a monomial dynamical system. Then f is a fixed point system with $(1, ..., 1)^t \in \mathbf{F}_q^n$ as its only fixed point if and only if its dependency graph only contains trivial strongly connected components .*

Proof. By Remark 72, $F := \Psi^{-1}(f)$ is the adjacency matrix of the dependency graph of f. If the dependency graph does not contain any nontrivial strongly connected components, every sequence $a \rightsquigarrow_s b$ between two arbitrary vertices can be at most of length $n - 1$. (A sequence that revisits a vertex would contain a closed sequence, which is strongly connected.) Therefore, by theorem 71

$\exists\ m \in \mathbb{N}$ with $m \leq n$ such that $F^m = 0$ (the zero matrix in $M(n \times n;\ \mathbb{N}_0)$). Now, according to equation (2.8) we have

$$\Psi^{-1}(f^m) = mred_q(F^m) = mred_q(0) = 0$$

and consequently

$$\Psi^{-1}(f^r) = 0\ \forall\ r \geq m$$

Thus

$$f^r = \mathbf{1}\ \forall\ r \geq m$$

If, on the other hand, there is an $m \in \mathbb{N}$ such that

$$f^{m+\lambda} = f^m = \mathbf{1}\ \forall\ \lambda \in \mathbb{N}$$

applying the isomorphism Ψ^{-1} (see Remark 59) we obtain

$$F^{\cdot(m+\lambda)} = F^{\cdot m} = 0\ \forall\ \lambda \in \mathbb{N}$$

and (see equation (2.8))

$$mred_q(F^{m+\lambda}) = mred_q(F^m) = 0\ \forall\ \lambda \in \mathbb{N}$$

It follows from equation (2.6) of Remark 47

$$F^{m+\alpha} = 0\ \forall\ \alpha \in \mathbb{N}_0$$

Now by theorem 71 there are no sequences $a \rightsquigarrow_s b$ between any two arbitrary vertices a, b of length larger than $m-1$. As a consequence, there cannot be any nontrivial strongly connected components in the dependency graph of f. ■

Definition 74 *A monomial dynamical system $f \in MF_n^n(\mathbf{F}_q)$ whose dependency graph contains nontrivial strongly connected components is called* coupled monomial dynamical system.

Definition 75 *Let $G = (V_G,\ E_G,\ \pi_G)$ be a digraph, $m \in \mathbb{N}$ a natural number and $a, b \in V_G$ two vertices. The number of different sequences of length m from a to b is denoted by $s_m(a,b) \in \mathbb{N}_0$.*

Remark 76 *Let $G = (V_G,\ E_G,\ \pi_G)$ be a digraph with vertex set V_G of cardinality $n := |V_G|$ and $V_G = \{a_1, ..., a_n\}$ a numeration of the elements of V_G. Furthermore, let $m \in \mathbb{N}$ be a natural number and $A \in M(n \times n;\ \mathbb{N}_0)$ the adjacency matrix of G with the numeration a. Then by Theorem 71 we have*

$$s_m(a_i, a_j) = (A^m)_{ij}$$

Theorem 77 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. Then f is a fixed point system if and only if there is an $m \in \mathbb{N}$ such that the following two conditions hold*

1. *For every pair of nodes $a, b \in V_f$ with $a \rightsquigarrow_m b$ there exists for every $\lambda \in \mathbb{N}$ an $a_\lambda \in \mathbb{Z}$ such that $s_{m+\lambda}(a,b) = s_m(a,b) + a_\lambda(q-1) \neq 0$.*
2. *For every pair of nodes $a, b \in V_f$ with $s_m(a,b) = 0$ it holds $s_{m+\lambda}(a,b) = 0\ \forall\ \lambda \in \mathbb{N}$.*

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices. If f is a fixed point system, $\exists\ m \in \mathbb{N}$ such that

$$f^{m+\lambda} = f^m\ \forall\ \lambda \in \mathbb{N}$$

By applying the homomorphism Ψ^{-1} we get (see Remark 59)

$$F^{\cdot(m+\lambda)} = F^{\cdot m}\ \forall\ \lambda \in \mathbb{N} \tag{2.9}$$

By Remark 72 it follows

$$mred_q(F^{m+\lambda}) = mred_q(F^m) \ \forall \ \lambda \in \mathbb{N}$$

Let $i, j \in \{1, ..., n\}$. If, on the one hand, $(F^{\cdot m})_{ij} = 0$ then by (2.9) we would have $(F^{\cdot(m+\lambda)})_{ij} = 0$ $\forall \ \lambda \in \mathbb{N}$. Consequently, by 2. of Lemma 39 we have

$$(F^{m+\alpha})_{ij} = 0 \ \forall \ \alpha \in \mathbb{N}_0$$

Now, by theorem 71 there are no sequences $a_i \rightsquigarrow_s a_j$ of length larger than $m-1$. In other words, 2. follows. If, on the other hand, $(F^{\cdot m})_{ij} \neq 0$ then by (2.9) we would have $(F^{\cdot(m+\lambda)})_{ij} = (F^{\cdot m})_{ij} \neq 0$ $\forall \ \lambda \in \mathbb{N}$. Consequently, by 2. and 4. of Lemma 39 $\exists \ a_\lambda \in \mathbb{Z}$ such that

$$(F^{m+\lambda})_{ij} = (F^m)_{ij} + a_\lambda(q-1) \ \forall \ \lambda \in \mathbb{N}$$

In other words, 1. follows. To show the converse we start from the following fact: Given 1. and 2. and according to Theorem 71 and Remark 72

$$\text{If } (F^m)_{ij} = 0, \text{ then } (F^{m+\lambda})_{ij} = (F^m)_{ij} \ \forall \ \lambda \in \mathbb{N}$$

and

$$\text{if } (F^m)_{ij} \neq 0, \text{ then } \exists \ a_\lambda \in \mathbb{Z} : (F^{m+\lambda})_{ij} = (F^m)_{ij} + a_\lambda(q-1) \neq 0 \ \forall \ \lambda \in \mathbb{N}$$

Now by 2. and 4. of Lemma 39 we have

$$mred_q(F^{m+\lambda}) = mred_q(F^m) \ \forall \ \lambda \in \mathbb{N}$$

and by 72

$$F^{\cdot(m+\lambda)} = F^{\cdot m} \ \forall \ \lambda \in \mathbb{N}$$

Thus, after applying the isomorphism Ψ

$$f^{m+\lambda} = f^m \ \forall \ \lambda \in \mathbb{N}$$

■

The following parameter for digraphs was introduced by [23]:

Definition 78 *Let $G = (V_G, E_G, \pi_G)$ be a digraph and $a \in V_G$ one of its vertices. The number*

$$\mathcal{L}_G(a) := \min_{\substack{a \rightsquigarrow_u a \\ a \rightsquigarrow_v a \\ u \neq v}} |u - v|$$

is called the loop number of a. If there is no sequence of positive length from a to a, then $\mathcal{L}_G(a)$ is set to zero.

Remark 79 *Note that the loop number $\mathcal{L}_{G'}(a)$ of the vertex a in a graph $G' = (V_G, E'_G, \pi'_G)$ may have a different value.*

Lemma 80 (and Definition) *Let $G = (V_G, E_G, \pi_G)$ be a digraph and $a \in V_G$ one of its vertices. If $\overleftrightarrow{a}$ is nontrivial then for every $b \in \overleftrightarrow{a}$ it holds*

$$\mathcal{L}_G(b) = \mathcal{L}_G(a)$$

Therefore, we introduce the loop number of strongly connected components as

$$\mathcal{L}_G(\overleftrightarrow{a}) := \mathcal{L}_G(a)$$

Proof. Let $\mathcal{L}_G(a) = t$. Therefore there are sequences $a \rightsquigarrow_r a$ and $a \rightsquigarrow_s a$ such that $|r - s| = t$. Since $\overleftrightarrow{a}$ is strongly connected, there are sequences $a \rightsquigarrow_u b$ and $b \rightsquigarrow_v a$ and we can construct the following sequences

$$\begin{aligned} b &\rightsquigarrow {}_{v+r+u}b = b \rightsquigarrow_v a \rightsquigarrow_r a \rightsquigarrow_u b \\ b &\rightsquigarrow {}_{v+s+u}b = b \rightsquigarrow_v a \rightsquigarrow_s a \rightsquigarrow_u b \end{aligned}$$

Now from

$$|v + r + u - (v + s + u)| = |r - s|$$

we have due to the minimality of the loop number

$$\mathcal{L}_G(b) \leq \mathcal{L}_G(a)$$

By symmetry the claim follows. ■

Remark 81 *The loop number of any trivial strongly connected component is equal to zero, due to the convention made in the definition of loop number.*

Corollary 82 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. If f is a fixed point system then the loop number of each of its nontrivial strongly connected components is equal to 1.*

Proof. Let $m \in \mathbb{N}$ be as in the statement of the previous theorem. Let $\overleftrightarrow{a} \subseteq V_f$ be a nontrivial strongly connected component. For every $b \in \overleftrightarrow{a}$ we have that b is strongly connected with itself. Therefore, for every $s \in \mathbb{N}$ there is a $t \geq s$ such that $b \rightsquigarrow_t b$. In particular, there must be a $u \in \mathbb{N}$ with $u > m$ such that $b \rightsquigarrow_u b$, i.e. $s_u(b, b) \geq 1$. By 2. of the previous theorem we know that $s_m(b, b) \neq 0$, otherwise $s_u(b, b) = 0$. Now from 1. of the previous theorem we know

$$\exists\, a_\lambda \in \mathbb{Z} : s_{m+\lambda}(b, b) = s_m(b, b) + a_\lambda(q - 1) \neq 0 \;\forall\; \lambda \in \mathbb{N}$$

and in particular

$$s_{m+\lambda}(b, b) \neq 0 \;\forall\; \lambda \in \mathbb{N}$$

Therefore, $\forall\; \lambda \in \mathbb{N}$ there are sequences $b \rightsquigarrow_{m+\lambda} b$. Thus $\mathcal{L}_{G_f}(\overleftrightarrow{a}) = \mathcal{L}_{G_f}(b) = 1$. ■

Definition 83 *Let $G = (V_G, E_G, \pi_G)$ be a digraph and $a, b \in V_G$ two vertices. The vertex a is called* recurrently connected *to b, if for every $s \in \mathbb{N}$ there is a $u \geq s$ such that $a \rightsquigarrow_u b$.*

Lemma 84 *Let $G = (V_G, E_G, \pi_G)$ be a digraph with vertex set V_G of cardinality $n := |V_G|$. Two vertices $a, b \in V_G$ are connected through a sequence $a \rightsquigarrow_t b$ of length $t > n - 1$ if and only if a is recurrently connected to b.*

Proof. If there is a sequence $a \rightsquigarrow_t b$ of length $t > n - 1$, then it necessarily revisits one of its vertices, in other words, there is a $c \in V_G$ such that

$$a \rightsquigarrow_t b = a \to ... \to c \to ... \to c \to ... \to b$$

Now a sequence $a \rightsquigarrow_{t'} b$ can be constructed that repeats the loop around c as many times as desired. The converse follows immediately from the definition of recurrent connectedness. ■

Remark 85 *Let $G = (V_G, E_G, \pi_G)$ be a digraph with vertex set V_G of cardinality $n := |V_G|$. Then for any two vertices $a, b \in V_G$ it holds: Either a is recurrently connected to b or there is an $m \in \mathbb{N}$ with $m \leq n$ such that no sequence $a \rightsquigarrow_t b$ of length $t \geq m$ exists.*

Lemma 86 *Let $G = (V_G,\ E_G,\ \pi_G)$ be a digraph and $U \subseteq V_G$ a nontrivial strongly connected component. Furthermore, let $t := \mathcal{L}_G(U)$ be the loop number of U. Then for each $a, b \in U$ there is an $m \in \mathbb{N}$ such that the graph G contains sequences $a \rightsquigarrow_{m+\lambda t} b$ of length $m + \lambda t\ \forall\ \lambda \in \mathbb{N}$.*

Proof. See the proof of Proposition 4.5 in [23]. This is an interesting and not straightforward proof! ■

Theorem 87 *Let $G = (V_G,\ E_G,\ \pi_G)$ be a digraph containing nontrivial strongly connected components. If the loop number of every nontrivial strongly connected component is equal to 1 then there is an $m \in \mathbb{N}$ such that **any** pair of vertices $a_i, a_j \in V_G$ with a_i recurrently connected to a_j satisfies*

$$s_{m+\lambda}(a_i, a_j) > 0\ \forall\ \lambda \in \mathbb{N}_0$$

Proof. Let $V_G = \{a_1, ..., a_n\}$ be the numeration of the vertices and $a_i, a_j \in V_G$. If a_i is recurrently connected to a_j, then necessarily there is a sequence $a_i \rightsquigarrow_s a_j$ that visits a vertex contained in a nontrivial strongly connected component. In other words, $\exists\ a_k \in V_f$ and a sequence $a_i \rightsquigarrow_s a_j$ such that $\overleftrightarrow{a_k}$ is nontrivial and

$$a_i \rightsquigarrow_s a_j = a_i \to ... \to a_k \to ... \to a_j$$

By Lemma 86 there is a $m_k \in \mathbb{N}$ such that there are sequences $a_k \rightsquigarrow_{m_k+\lambda} a_k\ \forall\ \lambda \in \mathbb{N}_0$. Now $\forall\ \lambda \in \mathbb{N}_0$ we can construct a sequence

$$a_i \rightsquigarrow_{s_\lambda} a_j = a_i \to ... \to a_k \rightsquigarrow_{m_k+\lambda} a_k \to ... \to a_j$$

Now, if we consider among all pairs $i, j \in \{1, ..., n\}$ such that $a_i \in V_G$ is recurrently connected to $a_j \in V_G$ the maximum m of all values m_k, we can state: $\exists\ m \in \mathbb{N}$ such that any pair of recurrently connected vertices $a_i, a_j \in V_G$ satisfies

$$s_{m+\lambda}(a_i, a_j) > 0\ \forall\ \lambda \in \mathbb{N}_0$$

■

Theorem 88 *Let $\mathbf{F}_2$ be the finite field with two elements, $f \in MF_n^n(\mathbf{F}_2)$ a* Boolean *coupled monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. f is a fixed point system if and only if the loop number of each nontrivial strongly connected component of G_f is equal to 1.*

Proof. The necessity follows from Corollary 82. Now assume that each nontrivial strongly connected component of G_f has loop number 1 and let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices. Furthermore let $F := \Psi^{-1}(f)$ be the corresponding matrix and consider vertices $a_i, a_j \in V_f$. By Remark 85, either a_i is recurrently connected to a_j or there is an $u_0 \in \mathbb{N}$ with $u_0 \leq n$ such that no sequence $a_i \rightsquigarrow_t a_j$ of length $t \geq u_0$ exists. If the latter is the case, then

$$(F^{u_0+\lambda})_{ij} = 0\ \forall\ \lambda \in \mathbb{N}_0$$

On the other hand, if a_i is recurrently connected to a_j, then by Theorem 87 there is an $m_0 \in \mathbb{N}$ such that

$$(F^{m_0+\lambda})_{ij} \neq 0\ \forall\ \lambda \in \mathbb{N}_0$$

Therefore, we have for $m := \max(m_0, u_0)$ that

$$(F^{m+\lambda})_{ij} \neq 0\ \forall\ \lambda \in \mathbb{N}_0 \text{ or } (F^{m+\lambda})_{ij} = 0\ \forall\ \lambda \in \mathbb{N}_0$$

Summarizing we have by 2. of Lemma 39

$$mred_q(F^{m+\lambda}) = mred_q(F^m)\ \forall\ \lambda \in \mathbb{N}$$

and by 72

$$F^{\cdot(m+\lambda)} = F^{\cdot m} \ \forall \ \lambda \in \mathbb{N}$$

Thus, after applying the isomorphism Ψ

$$f^{m+\lambda} = f^m \ \forall \ \lambda \in \mathbb{N}$$

■

Remark 89 *The statements of the previous theorems together with the Remark 31 about zero functions as components constitute the statement of Theorem 6.1 in [23].*

In the following two corollaries we provide alternative proofs to the claims made in Corollary 6.3 and Theorem 6.5 of [23]:

Corollary 90 (and Definition) *Let* $\mathbf{F}_2$ *the finite field with two elements and* $f \in MF_n^n(\mathbf{F}_2)$ *the coupled monomial dynamical system defined by*

$$\begin{aligned} f_1(x) &= x_1^{a_{11}} \\ f_i(x) &= (\prod_{j=1}^{i-1} x_j^{a_{ij}}) x_i^{a_{ii}}, \ i = 2, ..., n \end{aligned}$$

where $a_{ij} \in E_q$, $i = 1, ..., n$, $j = 1, ..., i-1$. *Such a system is called a* Boolean triangular system. *Boolean triangular systems are always fixed point systems.*

Proof. From the structure of f it is easy to see that every strongly connected component of the dependency graph of f is either trivial or has loop number 1. ■

Corollary 91 *Let* $\mathbf{F}_2$ *the finite field with two elements,* $f \in MF_n^n(\mathbf{F}_2)$ *a fixed point system and* $j, i \in \{1, ..., n\}$. *Consider the system* $g \in MF_n^n(\mathbf{F}_2)$ *defined as* $g_k(x) = f_k(x) \ \forall \ k \in \{1, ..., n\} \backslash j$ *and* $g_j(x) = x_i f_j(x) \ \forall \ x \in \mathbf{F}_2^n$. *Then* g *is a fixed point system if there is no sequence* $a_i \rightsquigarrow_s a_j$ *from* a_i *to* a_j *or if* $\overleftrightarrow{a_i}$ *or* $\overleftrightarrow{a_j}$ *are nontrivial.*

Proof. Let $G_g = (V_g, E_g, \pi_g)$ be the dependency graph of g. If $i = j$ then E_g contains the self loop $a_i \to a_i$ and $\overleftrightarrow{a_i}$ becomes nontrivial (if it wasn't already) with loop number 1. If $i \neq j$ then we have two cases: If there is no sequence $a_i \rightsquigarrow_s a_j$, then adding the edge $a_j \to a_i$ (which might be already there) doesn't affect $\overleftrightarrow{a_i} \neq \overleftrightarrow{a_j}$. If there is a sequence $a_i \rightsquigarrow_s a_j$ then adding the edge $a_j \to a_i$ (which might be already there) forces $\overleftrightarrow{a_i} = \overleftrightarrow{a_j}$. Now since by hypothesis $\overleftrightarrow{a_i}$ or $\overleftrightarrow{a_j}$ are nontrivial and f is a fixed point system, then

$$\mathcal{L}_{G_g}(\overleftrightarrow{a_i}) = \mathcal{L}_{G_g}(\overleftrightarrow{a_j}) = 1$$

■

Definition 92 *Let* $\mathbf{F}_q$ *be a finite field,* $f \in MF_n^n(\mathbf{F}_q)$ *a monomial dynamical system and* $G_f = (V_f, E_f, \pi_f)$ *its dependency graph.* f *is called a* $(q-1)$-fold redundant monomial system *if there is an* $N \in \mathbb{N}$ *such that for* ***any*** *pair* $a, b \in V_f$ *with* a *recurrently connected to* b, *the following holds:*

$$\forall \ m \geq N \ \exists \ \alpha_{abm} \in \mathbb{N}_0 : s_m(a, b) = \alpha_{abm}(q-1)$$

Remark 93 *Note that any* Boolean *monomial dynamical system* $f \in MF_n^n(\mathbf{F}_2)$ *is* $(2-1)$*-fold redundant.*

Lemma 94 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled $(q-1)$-fold redundant monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Then f is a fixed point system if the loop number of each nontrivial strongly connected component of G_f is equal to* 1.

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. Consider two arbitrary vertices $a_i, a_j \in V_f$. By Remark 85, either a_i is recurrently connected to a_j or there is an $m_0 \in \mathbb{N}$ with $m_0 \leq n$ such that no sequence $a \rightsquigarrow_t b$ of length $t \geq m_0$ exists. If the latter is the case, then

$$(F^{m_0+\lambda})_{ij} = 0 \ \forall \ \lambda \in \mathbb{N}_0$$

On the other hand, if a_i is recurrently connected to a_j, then by Theorem 87 there is an $m_1 \in \mathbb{N}$ such that

$$s_{m_1+\gamma}(a_i, a_j) > 0 \ \forall \ \gamma \in \mathbb{N}_0 \tag{2.10}$$

Consider now $m_2 := \max(n, m_1)$. Due to the universality of m_1 in the expression (2.10), for any pair of vertices $a_i, a_j \in V_G$ with a_i recurrently connected to a_j there is a sequence $a_i \rightsquigarrow_{m_2+\gamma} a_j$ of length $m_2+\gamma$, in particular $s_{(m_2+\gamma)}(a_i, a_j) > 0 \ \forall \ \gamma \in \mathbb{N}_0$. Now, let N be the constant in Definition 92 and $m_3 := \max(N, m_2)$. Now, by hypothesis, $\exists \ \alpha_{ij\gamma} \in \mathbb{N}$ such that

$$s_{(m_3+\gamma)}(a_i, a_j) = \alpha_{ij\gamma}(q-1) \ \forall \ \gamma \in \mathbb{N}_0$$

Thus

$$\begin{aligned} s_{(m_3+\gamma)}(a_i, a_j) &= \alpha_{ij\gamma}(q-1) = \alpha_{ij0}(q-1) + (\alpha_{ij\gamma} - \alpha_{ij0})(q-1) \\ &= s_{m_3}(a_i, a_j) + (\alpha_{ij\gamma} - \alpha_{ij0})(q-1) \ \forall \ \gamma \in \mathbb{N}_0 \end{aligned}$$

Summarizing, since $m_0 \leq n \leq m_2 \leq m_3$, we can say $\forall \ i, j \in \{1, ..., n\}$, depending on whether a_i and a_j are recurrently connected or not,

$$(F^{m_3+\lambda})_{ij} = 0 \ \forall \ \lambda \in \mathbb{N}_0$$

or

$$\exists \ a_\lambda \in \mathbb{Z} : (F^{m_3+\lambda})_{ij} = (F^{m_3})_{ij} + a_\lambda(q-1) \neq 0 \ \forall \ \lambda \in \mathbb{N}_0$$

Now, by 2. and 4. of Lemma 39 it follows

$$mred_q(F^{m_3+\lambda}) = mred_q(F^{m_3}) \ \forall \ \lambda \in \mathbb{N}$$

and by 72

$$F^{\cdot(m_3+\lambda)} = F^{\cdot m_3} \ \forall \ \lambda \in \mathbb{N}$$

Thus, after applying the isomorphism Ψ

$$f^{m_3+\lambda} = f^{m_3} \ \forall \ \lambda \in \mathbb{N}$$

■

Theorem 95 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled $(q-1)$-fold redundant monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Then f is a fixed point system if and only if the loop number of each nontrivial strongly connected component of G_f is equal to* 1.

Proof. The claim follows immediately from Lemma 94 and Corollary 82. ■

Theorem 96 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Then f is a fixed point system if the following properties hold*

1. *The loop number of each nontrivial strongly connected component of* G_f *is equal to* 1.
2. *For each nontrivial strongly connected component* $\overleftrightarrow{a} \subseteq V_f$ *and arbitrary* $b, c \in \overleftrightarrow{a}$,

$$s_1(b,c) \neq 0 \Rightarrow s_1(b,c) = q-1$$

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices. Consider two vertices $a_i, a_j \in V_f$ such that a_i is recurrently connected to a_j. Then by Theorem 87 there is an $m_1 \in \mathbb{N}$ such that

$$s_{m_1+\gamma}(a_i, a_j) > 0 \ \forall \ \gamma \in \mathbb{N}_0 \tag{2.11}$$

Consider now $m_2 := \max(n, m_1)$. Due to the universality of m_1 in the expression (2.11), for any pair of vertices $a_i, a_j \in V_G$ with a_i recurrently connected to a_j there is a sequence $a_i \rightsquigarrow_{m_2+\gamma} a_j$ of length $m_2+\gamma$. Since $m_2+\gamma > n-1$, necessarily $\exists\ a_{k_\gamma}, a_{l_\gamma} \in \overleftrightarrow{a_{k_\gamma}}$ such that $\overleftrightarrow{a_{k_\gamma}}$ is nontrivial and

$$a_i \rightsquigarrow_{(m_2+\gamma)} a_j = a_i \to ... \to a_{k_\gamma} \rightsquigarrow_t a_{l_\gamma} \to ... \to a_j \tag{2.12}$$

(t depends on i, j and γ). Now, by hypothesis, every two directly connected vertices $a, b \in \overleftrightarrow{a_{k_\gamma}}$ are directly connected by exactly $q-1$ directed edges. Therefore, for any sequence $a_{k_\gamma} \rightsquigarrow_t a_{l_\gamma}$ of length $t \in \mathbb{N}$ there are $(q-1)^t$ different copies of it and we can conclude $\exists\ \alpha \in \mathbb{N}$ such that $s_t(a_{k_\gamma}, a_{l_\gamma}) = \alpha(q-1)$. As a consequence, there are $\alpha(q-1)$ different copies of the sequence (2.12). Since we are dealing with an arbitrary sequence $a_i \rightsquigarrow_{(m_2+\gamma)} a_j$ of fixed length $m_2+\gamma$, $\gamma \in \mathbb{N}_0$ we can conclude that $\exists\ \alpha_{ij\gamma} \in \mathbb{N}$ such that

$$s_{(m_2+\gamma)}(a_i, a_j) = \alpha_{ij\gamma}(q-1) \ \forall \ \gamma \in \mathbb{N}_0$$

Thus f is a coupled $(q-1)$-fold redundant monomial dynamical system and the claim follows from Lemma 94. ■

Corollary 97 *Let* $\mathbf{F}_2$ *be the finite field with two elements,* $f \in MF_n^n(\mathbf{F}_2)$ *a* Boolean *monomial dynamical system and* $F := \Psi^{-1}(f) \in M(n \times n;\ E_2)$ *its corresponding matrix. Furthermore, let* $\mathbf{F}_q$ *be a finite field and* $g \in MF_n^n(\mathbf{F}_q)$ *the monomial dynamical system whose corresponding matrix* $G := \Psi^{-1}(g) \in M(n \times n;\ E_q)$ *satisfies* $\forall\ i, j \in \{1, ..., n\}$

$$G_{ij} = \begin{cases} q-1 \text{ if } F_{ij} = 1 \\ 0 \text{ if } F_{ij} = 0 \end{cases}$$

If f *is a fixed point system then* g *is a fixed point system too.*

Proof. Let $G_f = (V_f,\ E_f,\ \pi_f)$ be the dependency graph of f. By the definition of g, one can easily see that the dependency graph $G_g = (V_g,\ E_g,\ \pi_g)$ of g can be generated from G_f by adding $q-2$ identical parallel edges for every existing edge. Obviously G_f and G_g have the same strongly connected components. If G_f doesn't contain any nontrivial strongly connected components, then G_g wouldn't contain any either and by Theorem 73 g would be a fixed point system. If, on the other hand, G_f does contain nontrivial strongly connected components, then by Theorem 88 each of those components would have loop number 1. From the definition of g it also follows for any pair of vertices $a, b \in E_g$

$$s_1(a,b) \neq 0 \Rightarrow s_1(a,b) = q-1$$

By the previous theorem g would be a fixed point system. ■

Example 98 (and Corollary) *Let* $\mathbf{F}_q$ *be a finite field and* $f \in MF_n^n(\mathbf{F}_q)$ *the coupled monomial dynamical system defined by*

$$\begin{aligned} f_1(x) &= x_1^{q-1} \\ f_i(x) &= (\prod_{j=1}^{i-1} x_j^{a_{ij}}) x_i^{q-1},\ i = 2, ..., n \end{aligned}$$

where $a_{ij} \in E_q$, $i = 1, ..., n$, $j = 1, ..., i-1$ are not further specified exponents. Such a system is called triangular*. It is easy to see that the dependency graph of f contains n one vertex nontrivial strongly connected components. Each of them has a $(q-1)$-fold self loop. Therefore, by the previous Theorem, f must be a fixed point system.*

Theorem 99 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Then f is a fixed point system if for every vertex $a \in V_f$ that is recurrently connected to some other vertex $b \in V_f$ the edge $a \to a$ appears exactly $q-1$ times in E_f, i.e.*

$$\left|\pi_f^{-1}((a,a))\right| = q-1$$

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. Consider two vertices $a_i, a_j \in V_f$ such that a_i is recurrently connected to a_j. Then by Theorem 87 there is an $m_1 \in \mathbb{N}$ such that

$$s_{m_1+\gamma}(a_i, a_j) > 0 \ \forall \ \gamma \in \mathbb{N}_0 \tag{2.13}$$

Consider now $m_2 := \max(n, m_1)$. Due to the universality of m_1 in the expression (2.13), for any pair of vertices $a_i, a_j \in V_G$ with a_i recurrently connected to a_j there is a sequence $a_i \rightsquigarrow_{m_2+\gamma} a_j$ of length $m_2 + \gamma$. Consider one particular sequence $a_i \rightsquigarrow_{m_2+\gamma} a_j$ of length $m_2 + \gamma$ and call it $w_\gamma := a_i \rightsquigarrow_{m_2+\gamma} a_j$. By hypothesis there are exactly $q-1$ directed edges $a_i \to a_i$. Therefore, there are $q-1$ copies of the sequence w_γ. Since we are dealing with an arbitrary sequence $a_i \rightsquigarrow_{(m_2+\gamma)} a_j$ of fixed length $m_2 + \gamma$, $\gamma \in \mathbb{N}_0$ we can conclude that $\exists\ \alpha_{ij\gamma} \in \mathbb{N}$ such that

$$s_{(m_2+\gamma)}(a_i, a_j) = \alpha_{ij\gamma}(q-1) \ \forall \ \gamma \in \mathbb{N}_0$$

Thus f is a coupled $(q-1)$-fold redundant monomial dynamical system and the claim follows from Lemma 94. ■

Example 100 (and Corollary) *Let $\mathbf{F}_q$ be a finite field and $f \in MF_n^n(\mathbf{F}_q)$ a monomial dynamical system such that the diagonal entries of its corresponding matrix $F := \Psi^{-1}(f)$ satisfy*

$$F_{ii} = q-1 \ \forall \ i \in \{1, ..., n\}$$

Since every vertex satisfies the requirement of the previous theorem, f must be a fixed point system. This result generalizes our previous result about triangular monomial dynamical systems *$g \in MF_n^n(\mathbf{F}_q)$ defined as*

$$\begin{aligned} g_1(x) &= x_1^{q-1} \\ g_i(x) &= (\prod_{j=1}^{i-1} x_j^{a_{ij}}) x_i^{q-1}, \ i = 2, ..., n \end{aligned}$$

Lemma 101 *Let $n \in \mathbb{N}$ be a natural number and $A \in M(n \times n; \mathbb{R})$ a real matrix. In addition, let A be diagonalizable over $\mathbb{C}$. Then $A^m = A \ \forall \ m \in \mathbb{N}$ if and only if the characteristic polynomial $charpoly(A)$ of A can be written as*

$$charpoly(A) = a(\lambda - 1)^s \lambda^t$$

where $a \in \mathbb{R} \backslash \{0\}$.

Proof. Since A is diagonalizable, there is an invertible matrix $S \in M(n \times n; \mathbb{C})$ such that

$$A = SDS^{-1} \tag{2.14}$$

where $D \in M(n \times n; \mathbb{C})$ is a diagonal matrix. As a consequence,

$$\begin{aligned} charpoly(A) &= \det(A - \lambda I) = \det(SDS^{-1} - \lambda SS^{-1}) \\ &= \det(S(D - \lambda I)S^{-1}) \\ &= \det(S)\det(D - \lambda I)\det(S^{-1}) \\ &= \det(D - \lambda I) = \prod_{i=1}^{n}(D_{ii} - \lambda) \end{aligned} \tag{2.15}$$

Now assume $A^m = A \ \forall \ m \in \mathbb{N}$.Then it follows $\forall \ m \in \mathbb{N}$

$$SD^mS^{-1} = A^m = A = SDS^{-1}$$

and therefore

$$D^m = D \ \forall \ m \in \mathbb{N}$$

from which it follows

$$D_{ii}^m = D_{ii} \ \forall \ m \in \mathbb{N}, \ i \in \{1, ..., n\}$$

and thus $D_{ii} = 0$ or $D_{ii} = 1 \ \forall \ i \in \{1, ..., n\}$. From equation (2.15) we can conclude, $charpoly(A) = a(\lambda - 1)^s\lambda^t$ with $a \in \{-1, 1\}$. On the other hand, if $charpoly(A) = a(\lambda - 1)^s\lambda^t$, the eigenvalues of A are 0 or 1 and therefore the zeros of equation (2.15) must be 0 or 1. In other words, $D_{ii} = 0$ or $D_{ii} = 1 \ \forall \ i \in \{1, ..., n\}$. Therefore, by equation (2.14) we have $\forall \ m \in \mathbb{N}$

$$F^m = SD^mS^{-1} = SDS^{-1} = F$$

■

Theorem 102 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $F := \Psi^{-1}(f) \in M(n \times n; E_q)$ its corresponding matrix. If the matrix F (viewed as a real matrix $F \in M(n \times n; \mathbb{N}) \subset M(n \times n; \mathbb{R})$) has the characteristic polynomial*

$$charpoly(F) = a(\lambda - 1)^s\lambda^t \tag{2.16}$$

where $a \in \mathbb{Z}\backslash\{0\}$, and the geometric multiplicity of the eigenvalues 0 and 1 is equal to the corresponding algebraic multiplicity, then f is a fixed point system.

Proof. It is a well-known linear algebraic result that if there is a basis of eigenvectors of a matrix, the matrix is diagonalizable. By the hypothesis this is the case for F. Therefore, by the previous Lemma

$$F^m = F \ \forall \ m \in \mathbb{N}$$

Now, by Remarks 59 and 47 we consequently have $\forall \ m \in \mathbb{N}$

$$\Psi^{-1}(f^m) = F^{\cdot m} = mred_q(F^m) = mred_q(F) = F$$

After applying the isomorphism Ψ we get

$$f^m = f \ \ \forall \ m \in \mathbb{N}$$

■

Remark 103 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $F := \Psi^{-1}(f) \in M(n \times n; E_q)$ its corresponding matrix. The matrix F viewed as the adjacency matrix of the dependency graph $G_f = (V_f, E_f, \pi_f)$ of f satisfies*

$$F^m = F \ \forall \ m \in \mathbb{N}$$

if and only if for each pair of vertices $a, b \in V_f$ the value $s_m(a, b)$ is constant for all $m \in \mathbb{N}$. In other words, a and b are either disconnected or for every length $m \in \mathbb{N}$ they are connected with the same degree of redundancy.

Example 104 *Consider the monomial system $g \in MF_5^5(\mathbf{F}_3)$ defined by the matrix*

$$G := \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

It is easy to show that

$$charpoly(G) = (\lambda - 1)^3 \lambda^2$$

However, g is not a fixed point system. This shows that the condition (2.16) alone is not sufficient.

2.4 An algorithm of polynomial complexity to identify fixed point systems

According to our definition of monomial dynamical system $f \in MF_n^n(\mathbf{F}_q)$, the possibility that one of the functions f_i is equal to the zero function is excluded (see Definition 30 and Remark 31). Therefore, the following algorithm is designed for such systems. However, in this algorithmic framework it would be convenient to include the more general case (as defined in [22] and [23]), i.e. the case when some of the functions f_i can indeed be equal to the zero function. In the vein of Remark 31 this actually only requires some type of preprocessing. The preprocessing algorithm will be described and analyzed in the Appendix.

Our algorithm is based on the following observation made by Dr. Michael Shapiro about general time discrete finite dynamical systems: By Remark 2, a chain of transient states in the phase space of a time discrete finite dynamical system $f : X^n \to X^n$ can contain at most $s := |X^n| - 1 = |X|^n - 1$ transient elements. Therefore, to determine whether a system is a fixed point system, it is sufficient to establish whether the mappings f^r and f^{r+1} are identical for any $r \geq s$. In the case of a monomial system $f \in MF_n^n(\mathbf{F}_q)$, due to Theorem 57, we only need to look at the corresponding matrices $F^{\cdot r}$, $F^{\cdot r+1} \in M(n \times n; E_q)$. Computationally it is more convenient to generate the following sequence of powers

$$F^{\cdot 2},\ (F^{\cdot 2})^{\cdot 2} = F^{\cdot 4},\ (F^{\cdot 4})^{\cdot 2} = F^{\cdot 8},\ (F^{\cdot 8})^{\cdot 2} = F^{\cdot 16}, ..., F^{\cdot (2^t)}$$

To achieve the "safe" number of iterations $\left|\mathbf{F}_q^n\right| - 1 = q^n - 1$ we need to make sure

$$2^t \geq q^n - 1$$

This is equivalent to

$$t \geq \log_2(q^n - 1)$$

To obtain a natural number we use the ceil function

$$t := ceil(\log_2(q^n - 1)) \tag{2.17}$$

Thus we have, due to the monotonicity of the log function,

$$t < \log_2(q^n - 1) + 1 \leq \log_2(q^n) + 1 = n \log_2(q) + 1$$

The algorithm is fairly simple: Given a monomial system $f \in MF_n^n(\mathbf{F}_q)$ and its corresponding matrix $F := \Psi^{-1}(f) \in M(n \times n; E_q)$

1. With t as defined above (2.17), calculate the matrices $A := F^{\cdot 2^{(2^t)}}$ and $B := FA$. This step requires $t + 1$ matrix multiplications.

2. Compare the n^2 entries A_{ij} and B_{ij}. This step requires at most n^2 comparisons (this maximal value is needed in the case that f is a fixed point system).

3. f is a fixed point system if and only if the matrices A and B are equal.

It is well known that matrix multiplication requires $2n^3 - n^2$ addition or multiplication operations. Since $t + 1 < n \log_2(q) + 2$, the number of operations required in step 1 is bounded above by

$$(2n^3 - n^2)(n \log_2(q) + 2)$$

Summarizing, we have the following upper bound $N(n, q)$ for the number of operations in steps 1 and 2

$$N(n, q) := (2n^3 - n^2)(n \log_2(q) + 2) + n^2$$

For a fixed size q of the finite field $\mathbf{F}_q$ used it holds

$$\lim_{n\to\infty} \frac{N(n, q)}{n^4} = 2 \log_2(q)$$

and we can conclude $N(n, q) \in O(\ n^4)$ for a fixed q. The asymptotic behavior for a growing number of variables and growing number of field elements is described by

$$\lim_{\substack{n\to\infty \\ q\to\infty}} \frac{N(n, q)}{n^4 \log_2(q)} = 2$$

Thus, $N(n, q) \in O(\ n^4 \log_2(q))$ for $n, q \to \infty$.
It is pertinent to comment on the arithmetic operations performed during the matrix multiplications. Since the matrices are elements of the matrix monoid $M(n \times n; E_q)$, the arithmetic operations are operations in the semiring E_q. By the Lemmas 42 and 41 the addition resp. the multiplication operation on E_q requires an integer number addition[7] resp. multiplication and a reduction as defined in Lemma 39. The reduction $red_q(a)$ of an integer number $a \in \mathbb{N}_0$, $a \geq q$ is obtained as the degree of the remainder of the polynomial division $\tau^a \div (\ \tau^q - \tau)$. According to 4.6.5 of [53] this division requires

$$O(2(\deg(\tau^a) - \deg(\tau^q - \tau))) = O(2(a - q))$$

integer number operations. However, we know that the reductions $red_q(.)$ are applied to the result of (regular integer) addition or multiplication of elements of E_q and therefore

$$a - q \leq \begin{cases} 2(q-1) - q = q - 2 \\ (q-1)^2 - q = q^2 - q + 1 \end{cases}$$

As a consequence, in the worst case scenario, one addition resp. multiplication in the monoid E_q requires $O(q)$ resp. $O(q^2)$ regular integer number operations.
Since E_q is a finite set and only the results of n^2 pairwise additions and n^2 pairwise multiplications are needed, while the algorithm is running, these numbers are of course stored in a table after the first time they are calculated.

2.5 The cycle structure of monomial systems with strongly connected dependency graph

We start this section with a review and a detailed exposition of the definitions and results obtained by [23] on strongly connected graphs:

[7] See Chapter 4 of [53] for a detailed description of integer number representation and arithmetic in typical computer algebra systems.

2.5.1 Strongly connected graphs and the loop number

The most general setting for the following definitions and statements would include the case of trivial strongly connected graphs, i.e. a graph containing only one vertex and no edges. Such a graph, seen as dependency graph, corresponds to a univariate monomial dynamical system $f : \mathbf{F}_q \to \mathbf{F}_q$ such that $f \equiv 1$. In other words, not an interesting system. Therefore, we will consider only nontrivial strongly connected graphs.

Lemma 105 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph. Furthermore, let $t := \mathcal{L}_G(V_G) \geq 0$ be its loop number and $a, b \in V_G$ arbitrary vertices. Then for any pair of sequences $a \rightsquigarrow_m b$ and $a \rightsquigarrow_{m'} b$ contained in G there is an $\alpha \in \mathbb{N}_0$ such that*

$$\left|m - m'\right| = \alpha t$$

Proof. See the proof of Lemma 4.3 in [23]. ■

Corollary 106 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph. Furthermore, let $t := \mathcal{L}_G(V_G) \geq 0$ be its loop number and $a \in V_G$ an arbitrary vertex. Then for any closed sequence $a \rightsquigarrow_m a$ there is a $\alpha \in \mathbb{N}_0$ such that*

$$m = \alpha t$$

Proof. See the proof of Corollary 4.4 in [23]. ■

Lemma 107 (and Definition) *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number. For any $a, b \in V_G$ the relation $\approx$ defined by*

$$a \approx b :\Leftrightarrow \exists \text{ a sequence } a \rightsquigarrow_{\alpha t} b \text{ with } \alpha \in \mathbb{N}_0$$

is an equivalence relation called loop equivalence*. The loop equivalence class of an arbitrary vertex $a \in V_G$ is denoted by $\widetilde{a}$.*

Proof. See the proof of Lemma 4.6 in [23]. ■

Lemma 108 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number. Then the partition of V_G defined by the loop equivalence $\approx$ contains exactly t loop equivalence classes.*

Proof. See the proof of Lemma 4.7 in [23]. ■

Definition 109 *Let $G = (V_G, E_G, \pi_G)$ be a digraph, $a \in V_G$ an arbitrary vertex and $m \in \mathbb{N}$ a natural number. Then the set*

$$N_m(a) := \{b \in V_G : \exists\ a \rightsquigarrow_m b\}$$

is called the set of neighbors of order m.

Remark 110 *From the definitions it is clear that*

$$\widetilde{a} = \bigcup_{\alpha \in \mathbb{N}_0} N_{\alpha t}(a)$$

Theorem 111 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number and $\widetilde{a} \subseteq V_G$ an arbitrary loop equivalence class of V_G. Then for any $b, b' \in \widetilde{a}$ the following holds*

1. $N_m(b) \cap N_{m'}(b') = \emptyset$ *for* $m, m' \in \mathbb{N}$ *such that* $1 \leq m, m' < t$ *and* $m \neq m'$.
2. $N_m(b) \cap \widetilde{a} = \emptyset$ *for* $m \in \mathbb{N}$ *such that* $1 \leq m < t$.
3. *For every fixed* $m \in \mathbb{N}$ *such that* $1 \leq m \leq t$ $\exists\, c \in V_G : \bigcup_{b \in \widetilde{a}} N_m(b) = \widetilde{c}$.

Proof. Assume there was a vertex $c \in V_G$ and paths $b \rightsquigarrow_m c$ and $b' \rightsquigarrow_{m'} c$, where $m, m' \in \mathbb{N}$ such that $1 \leq m, m' < t$ and $m \neq m'$. Since $b, b' \in \widetilde{a}$, there is a sequence $b \rightsquigarrow_{\lambda t} b'$ with $\lambda \in \mathbb{N}_0$. Now consider the sequence $b \rightsquigarrow_{\lambda t} b' \rightsquigarrow_{m'} c$. By Lemma 105, there would be an $\alpha \in \mathbb{N}_0$ such that

$$\left|\lambda t + m - m'\right| = \alpha t$$

Wlog assume $m > m'$. As a consequence we would have

$$m = m' + (\alpha - \lambda)t$$

and thus $m = m'$ or $m \geq t$, a contradiction. This shows 1.

Let $c \in \widetilde{a}$. Then, by the definition of the class $\widetilde{a}$, there is a sequence $b \rightsquigarrow_{\alpha t} c$ with $\alpha \in \mathbb{N}_0$. Now, if there was a sequence $b \rightsquigarrow_m c$ with $m \in \mathbb{N}$ such that $1 \leq m < t$ then by Lemma 105 there would be a $\beta \in \mathbb{N}_0$ such that

$$|m - \alpha t| = \beta t$$

and thus $m = 0$ or $m \geq t$, a contradiction. This shows 2.
To show 3. consider an arbitrary pair of vertices $b, b' \in \widetilde{a}$ and vertices $c, c' \in V_G$ such that there are sequences $b \rightsquigarrow_m c$ and $b' \rightsquigarrow_m c'$ of length $m \in \mathbb{N}$ with $1 \leq m \leq t$. Since G is strongly connected, there is a sequence $c \rightsquigarrow_p b$ of length $p \in \mathbb{N}_0$. Now, by Corollary 106, the length $m + p$ of the sequence

$$b \rightsquigarrow_m c \rightsquigarrow_p b$$

satisfies

$$m + p = \alpha t \text{ with } \alpha \in \mathbb{N} \tag{2.18}$$

We also know that since $b, b' \in \widetilde{a}$, there is a sequence $b \rightsquigarrow_{\lambda t} b'$ with $\lambda \in \mathbb{N}_0$. Again, since G is strongly connected, there is a sequence $c' \rightsquigarrow_q c$ of length $q \in \mathbb{N}_0$. Now we consider the closed sequence

$$c \rightsquigarrow_p b \rightsquigarrow_{\lambda t} b' \rightsquigarrow_m c' \rightsquigarrow_q c$$

of length $p + \lambda t + m + q$. Again, by Corollary 106, this length satisfies

$$p + \lambda t + m + q = \gamma t \text{ with } \gamma \in \mathbb{N}_0$$

and with equation (2.18) we have

$$q = (\gamma - \alpha - \lambda)t$$

Therefore $c \approx c'$. This shows $N_m(b) \subseteq \widetilde{c}$ and thus $\bigcup_{b \in \widetilde{a}} N_m(b) \subseteq \widetilde{c}$. Now consider an arbitrary vertex $d \in \widetilde{c}$. Since G is strongly connected, there is a sequence $b \rightsquigarrow_s d$ of length $s \in \mathbb{N}_0$. In addition, from $c \in \widetilde{c}$ we know that there is a sequence $d \rightsquigarrow_{\delta t} c$ with $\delta \in \mathbb{N}_0$. Now we consider the closed sequence

$$b \rightsquigarrow_s d \rightsquigarrow_{\delta t} c \rightsquigarrow_p b$$

of length $s + \delta t + p$. As before, by Corollary 106, this length satisfies

$$s + \delta t + p = \omega t \text{ with } \omega \in \mathbb{N}_0$$

and with equation (2.18) we have

$$s = (\omega - \delta - \alpha)t + m$$

From this equation we can follow that the sequence $b \rightsquigarrow_s d$ visits a vertex $e \in V_G$ after a distance of $(\omega - \delta - \alpha)t$ edges, i.e. $e \in \widetilde{b} = \widetilde{a}$. Then the sequence continues in form of a sequence $e \rightsquigarrow_m d$. In other words, $\exists\ e \in \widetilde{a} : d \in N_m(e)$. Therefore, $\widetilde{c} \subseteq \bigcup\limits_{b \in \widetilde{a}} N_m(b)$. ■

Remark 112 *It is worth mentioning that since V_G is strongly connected and nontrivial, $N_m(b) \neq \emptyset\ \forall\ m \in \mathbb{N}$, $b \in V_G$. Moreover, from 1. it follows easily*

$$\left(\bigcup_{b \in \widetilde{a}} N_m(b)\right) \cap \left(\bigcup_{b \in \widetilde{a}} N_{m'}(b)\right) = \emptyset \text{ for } m, m' \in \mathbb{N} \text{ such that } 1 \leq m, m' < t \text{ and } m \neq m'$$

and because of 2. of course

$$\widetilde{a} = \bigcup_{b \in \widetilde{a}} N_t(b)$$

Given one loop equivalence class $\widetilde{a} \subseteq V_G$, the set of all the t loop equivalence classes can be ordered in the following manner

$$\widetilde{a}_i := \widetilde{a},\ \widetilde{a}_{i+1} = \bigcup_{b \in \widetilde{a}_i} N_1(b), ...\widetilde{a}_{i+j} = \bigcup_{b \in \widetilde{a}_i} N_j(b), ...\widetilde{a}_{i+t-1} = \bigcup_{b \in \widetilde{a}_i} N_{t-1}(b) \tag{2.19}$$

For any $c \in \bigcup\limits_{b \in \widetilde{a}_i} N_{t-1}(b)$ it must hold $N_1(c) \subseteq \widetilde{a}_i$ (if $N_1(c) \cap \widetilde{a}_j \neq \emptyset$ with $j \neq i$, then $\widetilde{a}_i = \widetilde{a}_j$). Thus, the graph G can be visualized as

$$\widetilde{a}_i \rightrightarrows \widetilde{a}_{i+1} \rightrightarrows \cdots \rightrightarrows \widetilde{a}_{i+j} \rightrightarrows \widetilde{a}_{(i+j+1) \bmod t} \rightrightarrows ... \rightrightarrows \widetilde{a}_{i+t-1} \rightrightarrows \widetilde{a}_{(i+t) \bmod t}$$

Due to the fact $\widetilde{a} = \bigcup\limits_{b \in \widetilde{a}} N_t(b)\ \forall\ a \in V_G$, we can conclude that the claims of the previous lemma still hold if the sequence lengths m and m' are replaced by the more general lengths $\lambda t + m$ and $\lambda' t + m'$. In other words, it holds for any $b, b' \in \widetilde{a}$ and $\lambda, \lambda' \in \mathbb{N}_0$

1. *$N_{\lambda t+m}(b) \cap N_{\lambda' t+m'}(b') = \emptyset$ for $m, m' \in \mathbb{N}$ such that $1 \leq m, m' < t$ and $m \neq m'$.*
2. *$N_{\lambda t+m}(b) \cap \widetilde{a} = \emptyset$ for $m \in \mathbb{N}$ such that $1 \leq m < t$.*
3. *$\exists\ c \in V_G : \bigcup\limits_{b \in \widetilde{a}} N_{\lambda t+m}(b) = \widetilde{c}$ for $m \in \mathbb{N}$ such that $1 \leq m \leq t$.*

and consequently

$$\left(\bigcup_{b \in \widetilde{a}} N_{\lambda t+m}(b)\right) \cap \left(\bigcup_{b \in \widetilde{a}} N_{\lambda' t+m'}(b)\right) = \emptyset \text{ for } m, m' \in \mathbb{N} \text{ such that } 1 \leq m, m' < t \text{ and } m \neq m'$$

and

$$\widetilde{a} = \bigcup_{b \in \widetilde{a}} N_{\lambda t}(b)$$

Corollary 113 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number and $V_G = \{a_1, ..., a_n\}$ a numeration of the vertices. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_G$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_G$ of V_G, i.e. $\forall\ k \in \{0, ..., t-1\}\ |C_k| = |\widetilde{a}_k|$ and $a_j \in \widetilde{a}_k\ \forall\ j \in C_k$. Then for any natural number $s \in \mathbb{N}$ such that $s \leq t$ and each $i \in \{0, ..., t-1\}$ the following holds*

$$\bigcup_{j \in C_i} N_{\lambda t+s}(a_j) = \widetilde{a}_{(i+s) \bmod t}\ \forall\ \lambda \in \mathbb{N}$$

Proof. This follows immediately from $\widetilde{a}_k = \bigcup_{b\in\widetilde{a}_k} N_{\lambda t}(b)$ and the definition of the order (2.19). ■

Remark 114 *If* $\mathrm{lcm}(s,t) < st$, *i.e.* $\exists\ r \in \mathbb{N}$ *with* $r < t$ *such that* $\mathrm{lcm}(s,t) = rs$, *then the sets* $\widetilde{a}_0, ..., \widetilde{a}_{t-1}$ *can be arranged in* $q := t/r$ *families with no repetitions except for the first and last class*

$$\begin{gathered}\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{rs \bmod t}\\ \widetilde{a}_1, \widetilde{a}_{1+s}, \widetilde{a}_{1+2s}, ..., \widetilde{a}_{(1+rs) \bmod t}\\ \vdots\\ \widetilde{a}_{q-1}, \widetilde{a}_{q-1+s}, \widetilde{a}_{q-1+2s}, ..., \widetilde{a}_{(q-1+rs) \bmod t}\end{gathered}$$

where the vertices in $\widetilde{a}_{j+ks}$ *and* $\widetilde{a}_{j+(k+1)s}$ *are connected by sequences of length* $\lambda t + s$. *Moreover, no shorter family of this type can be constructed. To see this, assume that there is a shortest family containing* $0 < r < t$ *loop-classes where the vertices are connected by sequences of length* $\lambda t + s$

$$\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{rs \bmod t} = \widetilde{a}_0$$

Then it follows

$$rs \bmod t = 0 \Leftrightarrow \exists\ \lambda \in \mathbb{N} : rs = \lambda t$$

and since r *is minimal we have* $\mathrm{lcm}(s,t) = rs < ts$. *Of course, every of the classes* $\widetilde{a}_0, \widetilde{a}_1, ..., \widetilde{a}_{s-1}$ *yields such a family, though, not necessarily a different one. The number of different families is given by the quotient*

$$q := \frac{t}{r} = \frac{ts}{rs} = \frac{ts}{\mathrm{lcm}(s,t)} \in \mathbb{N}$$

(For any class $\widetilde{a}_j \notin \{\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{(r-1)s}\}$, *the family*

$$\widetilde{a}_j, \widetilde{a}_{(j+s) \bmod t}, \widetilde{a}_{(j+2s) \bmod t}, ..., \widetilde{a}_{j+rs \bmod t} = \widetilde{a}_j$$

cannot contain any of the classes $\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{(r-1)s}$ *since this would yield the contradiction* $\widetilde{a}_{j+rs \bmod t} = \widetilde{a}_j \in \{\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{(r-1)s}\}$. *The same argument can be now applied to a class* $\widetilde{a}_k \notin \{\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{(r-1)s}, \widetilde{a}_j, \widetilde{a}_{(j+s) \bmod t}, \widetilde{a}_{(j+2s) \bmod t}, ..., \widetilde{a}_{j+(r-1)s \bmod t}\}$. *This process can be continued till no more classes outside a family are left. As a consequence, the number* α *of different families satisfies* $t = \alpha r$, *thus* $\alpha = t/r$*) For the converse, if there is an* $r \in \mathbb{N}$ *with* $r < t$ *such that* $\mathrm{lcm}(s,t) = rs$, *then for every* $i \in \{0, ...s-1\}$ *it holds*

$$i + rs = i + \lambda t = i \bmod t$$

Therefore, each of the s *classes* $\widetilde{a}_0, \widetilde{a}_1, ..., \widetilde{a}_{s-1}$ *yields a family*

$$\widetilde{a}_i, \widetilde{a}_{(i+s) \bmod t}, \widetilde{a}_{(i+2s) \bmod t}, ..., \widetilde{a}_{(i+rs) \bmod t} = \widetilde{a}_i$$

where the vertices in $\widetilde{a}_{j+s}$ *and* $\widetilde{a}_{j+2s}$ *are connected by sequences of length* $\lambda t + s$. *Were there a shorter family*

$$\widetilde{a}_i, \widetilde{a}_{(i+s) \bmod t}, \widetilde{a}_{(i+2s) \bmod t}, ..., \widetilde{a}_{(i+r's) \bmod t} = \widetilde{a}_i$$

with $r' < r$ *then it would follow*

$$(i + r's) \bmod t = i$$

which is equivalent to $r's = \lambda t$ *and thus* $r's < \mathrm{lcm}(s,t)$, *a contradiction. Again, the number of different families is given by the quotient* $q := t/r$.
If, on the other hand, $\mathrm{lcm}(s,t) = ts$, *then any family where the vertices in* $\widetilde{a}_{j+ks}$ *and* $\widetilde{a}_{j+(k+1)s}$ *are connected by sequences of length* $\lambda t + s$ *must contain all the* t *classes. Were there a shorter family*

$$\widetilde{a}_i, \widetilde{a}_{(i+s) \bmod t}, \widetilde{a}_{(i+2s) \bmod t}, ..., \widetilde{a}_{(i+r's) \bmod t} = \widetilde{a}_i$$

with $r' < t$ then it would follow

$$(i + r's) \bmod t = i$$

which is equivalent to $r's = \lambda t$ and thus $r's < \operatorname{lcm}(s,t)$, a contradiction. Consequently, the only family that can be constructed is

$$\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{ts \bmod t}$$

From the results presented in Remark 112 we may ask whether the properties listed there already characterize a strongly connected digraph $G = (V_G,\, E_G,\, \pi_G)$ with a certain loop number $\mathcal{L}_G(V_G) > 0$. In other words, the question arises whether a strongly connected digraph whose vertex set V_G can be partitioned in t (nonempty) classes such that the properties listed in Remark 112 are satisfied, automatically satisfies $\mathcal{L}_G(V_G) = t$. It turns out, that this is not sufficient as the following example shows

Example 115 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a hexagon, i.e. $V_G = \{a_0, ..., a_5\}$, $E_G = \{e_0, ..., e_5\}$ and $\pi_f(e_i) = (a_i, a_{(i+1) \bmod})\ \forall\ i \in \{0, ..., 5\}$. Then $\mathcal{L}_G(V_G) = 6$. Now define the following classes*

$$\widehat{a}_0 := \{a_0, a_3\},\ \widehat{a}_1 := \{a_1, a_4\},\ \widehat{a}_2 := \{a_2, a_5\}$$

It is easy to verify that each class $\widehat{a}_i$, $i \in \{0, .., 2\}$ satisfies the following properties for any $b, b' \in \widehat{a}_i$ and $\lambda, \lambda' \in \mathbb{N}_0$

1. *$N_{\lambda 3+m}(b) \cap N_{\lambda' 3+m'}(b') = \emptyset$ for $m, m' \in \mathbb{N}$ such that $1 \leq m, m' < 3$ and $m \neq m'$.*
2. *$N_{\lambda 3+m}(b) \cap \widehat{a}_i = \emptyset$ for $m \in \mathbb{N}$ such that $1 \leq m < 3$.*
3. *$\exists\ j \in \{0, ..2\} : \bigcup\limits_{b \in \widehat{a}_i} N_{\lambda 3+m}(b) = \widehat{a}_j$ for $m \in \mathbb{N}$ such that $1 \leq m \leq 3$.*

Moreover, we have $\forall\ \lambda \in \mathbb{N}_0$

$$\widehat{a}_i = \bigcup_{b \in \widehat{a}_i} N_{\lambda 3}(b)$$

This could suggest that $\mathcal{L}_G(V_G) = 3$ which is, as we know, not the case. The missing property that would force $\mathcal{L}_G(V_G) = 3$ is provided by the following theorem concerning closed paths on the graph G :

Theorem 116 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial and $V_G = \{a_1, ..., a_n\}$ a numeration of the vertices. Furthermore, let $U := \{a_{i_1}, ..., a_{i_k}\} \subseteq V_G$ be the subset of vertices such that there is a closed path $a_{i_j} \rightsquigarrow_{l(j)} a_{i_j}$ contained in the graph G. Then the loop number $\mathcal{L}_G(V_G)$ satisfies*

$$\mathcal{L}_G(V_G) = \gcd(l(1), ..., l(k))$$

Proof. See the proof of Theorem 4.13 in [23]. ■

Remark 117 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Assume that the vertex set V_G can be partitioned into t (nonempty) classes*

$$\widehat{a}_0, \widehat{a}_1, ..., \widehat{a}_{t-1} \subseteq V_G$$

such that each class $\widehat{a}_i$, $i \in \{0, .., t-1\}$ satisfies the following properties for any $b, b' \in \widehat{a}_i$ and $\lambda, \lambda' \in \mathbb{N}_0$

1. *$N_{\lambda t+m}(b) \cap N_{\lambda' t+m'}(b') = \emptyset$ for $m, m' \in \mathbb{N}$ such that $1 \leq m, m' < t$ and $m \neq m'$.*
2. *$N_{\lambda t+m}(b) \cap \widehat{a}_i = \emptyset$ for $m \in \mathbb{N}$ such that $1 \leq m < t$.*

3. $\exists\ j \in \{0, ..t-1\} : \bigcup_{b \in \widehat{a}_i} N_{\lambda t+m}(b) = \widehat{a}_j$ *for* $m \in \mathbb{N}$ *such that* $1 \leq m \leq t$.

Moreover, we assume that $\forall\ \lambda \in \mathbb{N}_0$

$$\widehat{a}_i = \bigcup_{b \in \widehat{a}_i} N_{\lambda t}(b)$$

Then it follows that the length l *of any closed path must satisfy*

$$\exists\ \alpha_l \in \mathbb{N} : l = \alpha_l t$$

Now let $U := \{a_{i_1}, ..., a_{i_k}\} \subseteq V_G$ *be the subset of vertices such that there is a closed path* $a_{i_j} \rightsquigarrow_{l(j)} a_{i_j}$ *contained in the graph* G. *Then by the previous Theorem we have*

$$\mathcal{L}_G(V_G) = \gcd(l(1), ..., l(k)) = \gcd(\alpha_{l(1)}t, ..., \alpha_{l(k)}t) = t \gcd(\alpha_{l(1)}, ..., \alpha_{l(k)})$$

As a consequence, for $\mathcal{L}_G(V_G) = t$ *to hold, the condition*

$$\gcd(\alpha_{l(1)}, ..., \alpha_{l(k)}) = 1$$

must be fulfilled. This is precisely the additional property needed in the previous Example, which failed to be satisfied since in the Example the length of any closed path was $6 = 2 \cdot 3$. *The condition* $\gcd(\alpha_{l(1)}, ..., \alpha_{l(k)}) = 1$, *i.e. the numbers* $\alpha_{l(1)}, ..., \alpha_{l(k)} \in \mathbb{N}$ *are relatively prime, is in particular always fulfilled if one of the* $\alpha_{l(j)}$ *is equal to* 1.

We finish this Subsection reviewing one Theorem proved by [23]:

Theorem 118 *Let* $G = (V_G,\ E_G,\ \pi_G)$ *be a strongly connected digraph such that* V_G *is nontrivial and* $V_G = \{a_1, ..., a_n\}$ *a numeration of its vertices. Furthermore, let* $t := \mathcal{L}_G(V_G) > 0$ *be its loop number and* $\widetilde{a} = \{a_{i_1}, ..., a_{i_r}\} \subseteq V_G$ *an arbitrary loop equivalence class of* V_G *with cardinality* $r := |\widetilde{a}|$. *Then for any vertex* $a_{i_k} \in \widetilde{a}$ *there is an* $m \in \mathbb{N}$ *such that there is a sequence* $a_{i_k} \rightsquigarrow_{(m+\lambda)t} a_{i_j}$ *of length* $(m + \lambda)t\ \forall\ j \in \{1, ..., r\}$ *and* $\lambda \in \mathbb{N}$.

Proof. See the proof of Corollary 4.8 in [23]. ■

2.5.2 The cycle structure of Boolean monomial systems with strongly connected dependency graph

We start this subsection with two statements about general (not only Boolean) monomial systems over $\mathbf{F}_q$. These two simple results have interesting consequences for Boolean and $(q-1)$-*fold redundant monomial systems*. For pedagogic reasons we devote the rest of this subsection to the analysis of the cycle structure of Boolean monomial systems with strongly connected dependency graph. In particular, we show that for Boolean monomial systems with strongly connected dependency graph, the loop number and the period number coincide. Moreover, we provide alternative proofs of results presented in [23] and complement those results with a theorem on the number of cyclic trajectories of a given length (Theorem 132). In the next subsection we perform the more general analysis of $(q-1)$-*fold redundant monomial systems* with strongly connected dependency graph, obtaining analogical results. Since Boolean systems are trivial examples of $(q-1)$-fold redundant systems, the results of this subsection are actually a consequence of the more general theorems proved in the next subsection.

Lemma 119 *Let* $\mathbf{F}_q$ *be a finite field,* $f \in MF_n^n(\mathbf{F}_q)$ *a coupled monomial dynamical system and* $G_f = (V_f,\ E_f,\ \pi_f)$ *its dependency graph. Furthermore, let* G_f *be strongly connected with loop number* $t := \mathcal{L}_{G_f}(V_f) \geq 1$. *Then there is an* $\alpha \in \mathbb{N}$ *such that the period number* T *of* f *satisfies*

$$T = \alpha \mathcal{L}_{G_f}(V_f)$$

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. According to the definition of period number, there is an $m \in \mathbb{N}$ such that for every $s \geq m$ it holds

$$f^{s+\lambda T} = f^s \ \forall \ \lambda \in \mathbb{N}$$

Now, applying the isomorphism Ψ^{-1} we have by Remark 59

$$F^{\cdot(s+\lambda T)} = F^{\cdot(s)} \ \forall \ \lambda \in \mathbb{N}, s \geq m$$

In particular we have

$$F^{\cdot(s+T)} = F^{\cdot(s)} \ \forall \ s \geq m$$

which is equivalent to (see Remark 47)

$$mred_q(F^{s+T}) = mred_q(F^s) \ \forall \ s \geq m$$

By Remark 72 and 2. of Lemma 39 we can conclude that for every sequence of length $s \geq m$

$$a_i \rightsquigarrow_s a_j$$

contained in the graph G_f, (and there is certainly one such sequence for some $s \geq m$, since F^s cannot be the zero matrix), there must be a sequence

$$a_i \rightsquigarrow_{s+T} a_j$$

of length $s + T$ as well. Now, by Lemma 105, we have

$$T = \alpha \mathcal{L}_{G_f}(V_f) \text{ with } \alpha \in \mathbb{N}$$

■

Example 120 *Let $f \in MF_3(\mathbf{F}_5)$ be the monomial system defined by*

$$\begin{array}{rcl} f & : & \mathbf{F}_5^3 \to \mathbf{F}_5^3 \\ \vec{x} & \mapsto & f(\vec{x}) := (x_2, x_1x_3, x_1) \end{array}$$

It is easy to verify that the dependency graph of f is strongly connected with loop number equal to 1. However, the phase space of f displays closed paths of length 7 and 14, therefore, the period number T is equal to 14.

Definition 121 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a monomial dynamical system and $s \in \mathbb{N}$ a natural number. We denote the set of solutions in $\mathbf{F}_q^n$ of the equation $f^s(x) = x$ by $V_{\mathbf{F}_q}(f^s(x) - x)$.*

Definition 122 *Let $m \in \mathbb{N}$ be a natural number. We denote with*

$$D(m) := \{d \in \mathbb{N} : d \text{ divides } m\}$$

the set of all positive divisors of m.

Corollary 123 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$. In addition, let T be the period number of f, $s \in D(T)$ and α as in the previous Lemma. Then it holds*

$$V_{\mathbf{F}_q}(f^s(x) - x) \subseteq V_{\mathbf{F}_q}(f^{\alpha t}(x) - x)$$

Proof. Since $s \in D(T)$, $\exists\ \beta \in \mathbb{N} : T = \beta s$. Thus,

$$f^{\beta s} = f^T = f^{\alpha t}$$

As a consequence, we have

$$V_{\mathbf{F}_q}(f^s(x) - x) \subseteq V_{\mathbf{F}_q}(f^{\beta s}(x) - x) = V_{\mathbf{F}_q}(f^{\alpha t}(x) - x)$$

■

Lemma 124 *Let* $\mathbf{F}_2$ *be the finite field with two elements,* $f \in MF_n^n(\mathbf{F}_2)$ *a* Boolean *coupled monomial dynamical system and* $G_f = (V_f,\ E_f,\ \pi_f)$ *its dependency graph. Furthermore, let* G_f *be strongly connected with loop number* $t := \mathcal{L}_{G_f}(V_f) \geq 1$ *and* $V_f = \{a_1, ..., a_n\}$ *the numeration of the vertices. In addition, let* $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ *be the* t *loop equivalence classes ordered according to (2.19) and* $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ *the partition of the set* $\{1, ..., n\}$ *induced by the partition* $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ *of* V_f*, i.e.* $\forall\ k \in \{0, ..., t-1\}\ |C_k| = |\widetilde{a}_k|$ *and* $a_j \in \widetilde{a}_k\ \forall\ j \in C_k$*. Then there is an* $m \in \mathbb{N}$ *such that* $\forall\ \lambda \in \mathbb{N}$

$$f_i^{(m+\lambda)t}(x) = \prod_{j \in C_k} x_j\ \forall\ i \in C_k,\ k = 0, ..., t-1$$

Proof. Let $F := \Psi^{-1}(f)$ be the corresponding matrix of f. Let $k \in \{0, ..., t-1\}$. By Remark 112 it holds $\forall\ \lambda \in \mathbb{N}$

$$\widetilde{a}_k = \bigcup_{b \in \widetilde{a}_k} N_{\lambda t}(b)$$

In addition, by Theorem 118 $\exists\ m_k \in \mathbb{N}$ such that for any pair of vertices $a_i, a_j \in \widetilde{a}_k$ and $\forall\ \lambda \in \mathbb{N}$ there is a sequence $a_i \rightsquigarrow_{(m_k+\lambda)t} a_j$ of length $(m_k + \lambda)t$. Let $m := \max(m_0, ..., m_{t-1})$. From these facts we can conclude, that the matrix $F^{(m+\lambda)t}$ has the following properties $\forall\ i \in C_k$, $k = 1, ..., t-1$ and $\forall\ \lambda \in \mathbb{N}$

$$(F^{(m+\lambda)t})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_k$$

and

$$(F^{(m+\lambda)t})_{il} \neq 0\ \forall\ l \in C_k$$

By Remark 72 and 2. of Lemma 39 it follows $\forall\ i \in C_k$, $k = 1, ..., t-1$ and $\forall\ \lambda \in \mathbb{N}$

$$(F^{\cdot(m+\lambda)t})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_k$$

and

$$(F^{\cdot(m+\lambda)t})_{il} = 1\ \forall\ l \in C_k$$

Now, applying the isomorphism Ψ we have (see Remark 59)

$$(f^{(m+\lambda)t})_i(x) = \prod_{j \in C_k} x_j\ \forall\ i \in C_k$$

■

Theorem 125 *Let* $\mathbf{F}_2$ *be the finite field with two elements,* $f \in MF_n^n(\mathbf{F}_2)$ *a* Boolean *coupled monomial dynamical system and* $G_f = (V_f,\ E_f,\ \pi_f)$ *its dependency graph. Furthermore, let* G_f *be strongly connected with loop number* $t := \mathcal{L}_{G_f}(V_f) \geq 1$ *and* $s \in \mathbb{N}$ *a natural number. In addition, let* $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ *be the* t *loop equivalence classes ordered according to (2.19) and* $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ *the partition of the set* $\{1, ..., n\}$ *induced by the partition* $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ *of* V_f*. Then any point* $\xi \in \mathbf{F}_2^n$ *showing* s*-periodicity under* f*, i.e.* $f^s(\xi) = \xi$*, satisfies the following property*

$$\xi_i = \xi_j\ \forall\ i, j \in C_k,\ k = 0, ..., t-1$$

Proof. Let $m \in \mathbb{N}$ be as in Lemma 124 and $u, v \in \mathbb{N}$ such that $ut = vs$. (This is always possible due to the existence of the $\mathrm{lcm}(s,t)$.) Now choose $\alpha \in \mathbb{N}$ such that $\alpha u > m$. Then we have $f^{\alpha ut} = f^{\alpha vs}$ and by Lemma 124

$$f_i^{\alpha vs}(x) = \prod_{j \in C_k} x_j \ \forall\ i \in C_k,\ k = 0, ..., t-1$$

In particular, for the s-periodic point ξ it holds for each $k \in \{0, ..., t-1\}$

$$f_i^{\alpha vs}(\xi) = (f^s)_i^{\alpha v}(\xi) = \xi_i = \prod_{j \in C_k} \xi_j \ \forall\ i \in C_k$$

As a consequence

$$\xi_i = \xi_j \ \forall\ i, j \in C_k,\ k = 0, ..., t-1$$

■

Theorem 126 *Let* $\mathbf{F}_2$ *be the finite field with two elements,* $f \in MF_n^n(\mathbf{F}_2)$ *a* Boolean *coupled monomial dynamical system and* $G_f = (V_f, E_f, \pi_f)$ *its dependency graph. Furthermore, let* G_f *be strongly connected with loop number* $t := \mathcal{L}_{G_f}(V_f) \geq 1$. *Then there is an* $m \in \mathbb{N}$ *such that* $\forall\ \lambda \in \mathbb{N}$

$$V_{\mathbf{F}_2}(f^t(x) - x) = V_{\mathbf{F}_2}(f^{(m+\lambda)t}(x) - x)$$

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ of V_f, i.e. $\forall\ k \in \{0, ..., t-1\}\ |C_k| = |\widetilde{a}_k|$ and $a_j \in \widetilde{a}_k\ \forall\ j \in C_k$. By Remark 112 it holds

$$\widetilde{a}_k = \bigcup_{b \in \widetilde{a}_k} N_t(b)$$

From this fact we can conclude, that the matrix F^t has the following properties $\forall\ i \in C_k$, $k = 1, ..., t-1$

$$(F^t)_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_k$$

and

$$\exists\ l \in C_k : (F^t)_{il} \neq 0$$

By Remark 72 and 2. of Lemma 39 it follows $\forall\ i \in C_k$, $k = 1, ..., t-1$

$$(F^{\cdot t})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_k$$

and

$$\exists\ l \in C_k : (F^{\cdot t})_{il} = 1$$

As before, we can conclude

$$(f^t)_i(x) = \prod_{j \in C_k} x_j^{\epsilon_i(j)}\ \forall\ i \in C_k$$

where ϵ_i are *nonzero* functions

$$\epsilon_i : C_k \to \{0, 1\} \subset \mathbb{N}$$

Now let $m \in \mathbb{N}$ be as in Lemma 124. Then by Lemma 124 we have $\forall\ \lambda \in \mathbb{N}$

$$f_i^{(m+\lambda)t}(x) = \prod_{j \in C_k} x_j\ \forall\ i \in C_k,\ k = 0, ..., t-1$$

From the structure of the functions $f^{(m+\lambda)t}$ and f^t and Theorem 125 it is clear that any solution $\xi \in \mathbf{F}_2^n$ of the equation $f^{(m+\lambda)t}(x) = x$ also solves the equation $f^t(x) = x$. In other words

$$V_{\mathbf{F}_2}(f^{(m+\lambda)t}(x) - x) \subseteq V_{\mathbf{F}_2}(f^t(x) - x)$$

The inclusion

$$V_{\mathbf{F}_2}(f^t(x) - x) \subseteq V_{\mathbf{F}_2}(f^{(m+\lambda)t}(x) - x)$$

follows from the fact $f^{(m+\lambda)t} = (f^t)^{(m+\lambda)}$. The claim follows. ■

Corollary 127 *Let $\mathbf{F}_2$ be the finite field with two elements, $f \in MF_n^n(\mathbf{F}_2)$ a* Boolean *coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$. In addition, let T be the period number of f and $s \in D(T)$. If the phase space of f contains a cycle of length s, then s must divide $\mathcal{L}_{G_f}(V_f)$.*

Proof. By Corollary 123 there is an $\alpha \in \mathbb{N}$ such that

$$V_{\mathbf{F}_2}(f^s(x) - x) \subseteq V_{\mathbf{F}_2}(f^{\alpha t}(x) - x) \tag{2.20}$$

Now let $m \in \mathbb{N}$ be as in the previous Theorem. If $\alpha > m$ set $\beta := 1$ otherwise choose $\beta \in \mathbb{N}$ such that $\alpha\beta > m$. Then we have

$$V_{\mathbf{F}_2}(f^{\alpha t}(x) - x) \subseteq V_{\mathbf{F}_2}(f^{\alpha\beta t}(x) - x) \tag{2.21}$$

and from (2.20) and (2.21) and by the previous Theorem it follows

$$V_{\mathbf{F}_2}(f^s(x) - x) \subseteq V_{\mathbf{F}_2}(f^t(x) - x)$$

If the phase space of f contains a cycle of length s, i.e. if there are s *different* points $\xi_0, ..., \xi_{s-1} \in \mathbf{F}_2^n$ with

$$f(\xi_i) = \xi_{(i+1) \bmod s}$$

then from $\xi_0, ..., \xi_{s-1} \in V_{\mathbf{F}_2}(f^s(x) - x) \subseteq V_{\mathbf{F}_2}(f^t(x) - x)$ it follows that $s \leq t$ and s divides t. ■

Lemma 128 *Let $\mathbf{F}_2$ be the finite field with two elements, $f \in MF_n^n(\mathbf{F}_2)$ a* Boolean *coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$ and $s \in \mathbb{N}$ a natural number such that $s \leq t$. In addition, let $r \in \mathbb{N}$ be such that $\operatorname{lcm}(s, t) = rs$. Then the equation $f^s(x) = x$ has exactly $2^{\frac{t}{r}}$ solutions in $\mathbf{F}_2^n$.*

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ of V_f. We consider two cases. First, assume $\operatorname{lcm}(s, t) < st$, i.e. $\exists\, r \in \mathbb{N}$ with $r < t$ such that $\operatorname{lcm}(s, t) = rs$. Then, by Remark 114 the sets $\widetilde{a}_0, ..., \widetilde{a}_{t-1}$ can be arranged in $q := t/r$ families

$$\begin{gathered}
\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{rs \bmod t} \\
\widetilde{a}_1, \widetilde{a}_{1+s}, \widetilde{a}_{1+2s}, ..., \widetilde{a}_{(1+rs) \bmod t} \\
\vdots \\
\widetilde{a}_{q-1}, \widetilde{a}_{q-1+s}, \widetilde{a}_{q-1+2s}, ..., \widetilde{a}_{(q-1+rs) \bmod t}
\end{gathered}$$

where the vertices in $\widetilde{a}_{j+s}$ and $\widetilde{a}_{j+2s}$ are connected by sequences of length $\lambda t + s$. Moreover, no shorter family of this type can be constructed. From these facts we can conclude, that the matrix F^s has the following properties $\forall\, i \in C_k$, $k = u, u+s, u+2s, ..., u+(r-1)s$, $u = 0, ..., q-1$

$$(F^s)_{ij} = 0 \;\forall\; j \in \{1, ..., n\} \backslash C_{(k+s) \bmod t}$$

and

$$\exists\ l \in C_{(k+s) \bmod t} : (F^s)_{il} \neq 0$$

By Remark 72 and 2. of Lemma 39 it follows $\forall\ i \in C_k,\ k = 1, ..., t-1$

$$(F^{\cdot s})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_{(k+s) \bmod t}$$

and

$$\exists\ l \in C_{(k+s) \bmod t} : (F^{\cdot s})_{il} = 1$$

Now, applying the isomorphism Ψ we have (see Remark 59)

$$(f^s)_i(x) = \prod_{j \in C_{(k+s) \bmod t}} x_j^{\epsilon_i(j)}\ \forall\ i \in C_k$$

where ϵ_i are *nonzero* functions

$$\epsilon_i : C_{(k+s) \bmod t} \to \{0, 1\} \subset \mathbb{N}$$

As a consequence, for every fixed $u \in \{0, ..., q-1\}$ and $k = u, u+s, u+2s, ..., u+(r-1)s$ any solution $\xi \in \mathbf{F}_2^n$ of the equation $f^s(x) = x$ satisfies

$$\xi_i = \prod_{j \in C_{(k+s) \bmod t}} \xi_j^{\epsilon_i(j)}\ \forall\ i \in C_k \tag{2.22}$$

By Theorem 125 we also know that

$$\xi_l = \xi_i\ \forall\ i, l \in C_k,\ k = u, u+s, u+2s, ..., u+(r-1)s \tag{2.23}$$

Now, if $\xi_i = 1\ \forall\ i \in C_u$, by (2.22) and (2.23), it must follow, that $\xi_l = 1\ \forall\ l \in C_{(u+s) \bmod t}$. The same argument applied $r-1$ times lets us conclude $\xi_i = 1\ \forall\ i \in C_k,\ k = u, u+s, u+2s, ..., u+(r-1)s$. If, on the other hand, $\xi_i = 0\ \forall\ i \in C_u$, by (2.22) we have that $\exists\ v \in C_{(u+s) \bmod t} : \xi_v = 0$ and by (2.23) $\xi_l = 0\ \forall\ l \in C_{(u+s) \bmod t}$. The same argument applied $r-1$ times lets us conclude $\xi_i = 0$ $\forall\ i \in C_k,\ k = u, u+s, u+2s, ..., u+(r-1)s$. Summarizing, since every $u \in \{0, ..., q-1\}$ represents one of the above q families, there are exactly $2^q = 2^{\frac{t}{r}}$ solutions of $f^s(x) = x$ in $\mathbf{F}_2^n$.
The second case is when $\operatorname{lcm}(s, t) = ts$. Here, by Remark 114 the sets $\widetilde{a}_0, ..., \widetilde{a}_{t-1}$ can be arranged in one single family

$$\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{ts \bmod t}$$

The same argument as used above for a fixed value of u yields that, in this case, the only solutions of $f^s(x) = x$ in $\mathbf{F}_2^n$ are $(1, ..., 1), (0, ..., 0) \in \mathbf{F}_2^n$. Therefore, the number of solutions is equal to $2 = 2^{\frac{t}{t}}$. ■

Corollary 129 *Let $\mathbf{F}_2$ be the finite field with two elements, $f \in MF_n^n(\mathbf{F}_2)$ a* Boolean *coupled monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$ and $s \in \mathbb{N}$ a natural number such that $s \in D(t)$. Then the equation $f^s(x) = x$ has exactly 2^s solutions in $\mathbf{F}_2^n$.*

Proof. Since s divides t, $\exists\ r \in \mathbb{N} : t = rs$. Thus $\operatorname{lcm}(t, s) = t = rs$ and $t/r = s$. The claim follows from the previous Lemma. ■

Remark 130 *In particular, if $\{1, d_1, ..., d_u, t\} \subset \mathbb{N}$ is the set of divisors of t in ascending order, then the number of solutions of $f^s(x) = x$ in $\mathbf{F}_2^n$ grows monotonically from 2^1 to 2^t for $s = 1, d_1, ..., d_u, t$. More generally, if $s < t$, then $\operatorname{lcm}(s, t) > s$ and thus $\operatorname{lcm}(s, t) = rs$ with $r \geq 2$. As a consequence of the previous Lemma, for $s \in \{1, ..., t\}$ the number of solutions in $\mathbf{F}_2^n$ of the equation $f^s(x) = x$ takes its maximal value for $s = t$.*

The next theorem shows that in the Boolean case, period number and loop number coincide, provided the dependency graph is strongly connected:

Theorem 131 *Let $\mathbf{F}_2$ be the finite field with two elements, $f \in MF_n^n(\mathbf{F}_2)$ a* Boolean *coupled monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) > 1$. Then the period number T of f satisfies*

$$T = \mathcal{L}_{G_f}(V_f)$$

Moreover, the phase space of f contains cycles of all lengths $s \in D(T)$.

Proof. By Corollary 127 the length s of any cycle displayed in the phase space of f divides t, in particular, it holds $s \leq t$. Now let $\{d_0 := 1, d_1, ..., d_u, d_{u+1} := t\} \subset \mathbb{N}$ be the set of divisors of t in ascending order. By Remark 130 we know that

$$\left|V_{\mathbf{F}_2}(f^{d_i}(x) - x)\right| > \left|V_{\mathbf{F}_2}(f^{d_j}(x) - x)\right| \ \forall\ i, j \in \{1, ..., u+1\} : i > j$$

Therefore, the phase space of f indeed contains cycles of length $d_i\ \forall\ i \in \{1, ..., u+1\}$. Summarizing we can say that the the phase space of f only contains cycles of length $d_i\ \forall\ i \in \{1, ..., u+1\}$. From the definition we know $T = \mathrm{lcm}(1, d_1, ..., d_u, t)$ and thus

$$T = \mathcal{L}_{G_f}(V_f)$$

■

Theorem 132 (and Definition) *Let $\mathbf{F}_2$ be the finite field with two elements, $f \in MF_n^n(\mathbf{F}_2)$ a* Boolean *coupled monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) > 1$. In addition, let $s \in \mathbb{N}$ be a natural number and denote by Z_s the number of cycles of length s displayed by the phase space of f. Then it holds for any $d \in \mathbb{N}$*

$$Z_d = \begin{cases} \frac{2^d - \sum\limits_{j \in D(d) \setminus d} Z_j}{d} \text{ if } d \in D(t) \\ 0 \text{ if } d \notin D(t) \end{cases}$$

Proof. The claim follows immediately from Theorem 131 and Corollary 129. ■

Remark 133 *In particular, if the loop number $t = \mathcal{L}_{G_f}(V_f)$ is a prime number, then the phase space of f only displays cycles of length t and 1 (fixed points). More precisely*

$$Z_t = \frac{2^t - 2}{t}$$

and

$$Z_1 = 2$$

2.5.3 The cycle structure of $(q-1)$-fold redundant monomial systems

In this subsection we study the cycle structure of $(q-1)$*-fold redundant monomial systems with strongly connected dependency graph.* For this purpose, let's recall the definition:

Definition 134 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. f is called a* $(q-1)$-fold redundant monomial system *if there is an $N \in \mathbb{N}$ such that for **any** pair $a, b \in V_f$ with a recurrently connected to b, the following holds:*

$$\forall\ m \geq N\ \exists\ \alpha_{abm} \in \mathbb{N}_0 : s_m(a, b) = \alpha_{abm}(q-1)$$

Remark 135 *Note that any* Boolean *monomial dynamical system $f \in MF_n^n(\mathbf{F}_2)$ is $(2-1)$-fold redundant.*

Lemma 136 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a* $(q-1)$-fold redundant *coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$ and $V_f = \{a_1, ..., a_n\}$ the numeration of the vertices. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ of V_f. Then there is an $m \in \mathbb{N}$ such that $\forall\ \lambda \in \mathbb{N}$*

$$f_i^{(m+\lambda)t}(x) = \prod_{j \in C_k} x_j^{q-1}\ \forall\ i \in C_k,\ k = 0, ..., t-1$$

Proof. Let $F := \Psi^{-1}(f)$ be the corresponding matrix of f. Let $k \in \{0, ..., t-1\}$. By Remark 112 it holds $\forall\ \lambda \in \mathbb{N}$

$$\widetilde{a}_k = \bigcup_{b \in \widetilde{a}_k} N_{\lambda t}(b)$$

In addition, by Theorem 118 $\exists\ m_k \in \mathbb{N}$ such that for any pair of vertices $a_i, a_j \in \widetilde{a}_k$ and $\forall\ \lambda \in \mathbb{N}$ there is a sequence $a_i \rightsquigarrow_{(m_k+\lambda)t} a_j$ of length $(m_k + \lambda)t$. Let $m' := \max(m_0, ..., m_{t-1})$ and $N \in \mathbb{N}$ as in the previous Definition. Now choose $\gamma \in \mathbb{N}$ such that $(m' + \gamma)t \geq N$ and set $m := m' + \gamma$. From this information we can conclude, that the matrix $F^{(m+\lambda)t}$ has the following properties $\forall\ i \in C_k$, $k = 1, ..., t-1$ and $\forall\ \lambda \in \mathbb{N}$

$$(F^{(m+\lambda)t})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_k$$

and

$$\exists\ \alpha_{il((m+\lambda)t)} \in \mathbb{N} : (F^{(m+\lambda)t})_{il} = \alpha_{il((m+\lambda)t)}(q-1) \forall\ l \in C_k$$

By Remark 72 and 2. and 4. of Lemma 39 it follows $\forall\ i \in C_k$, $k = 1, ..., t-1$ and $\forall\ \lambda \in \mathbb{N}$

$$(F^{\cdot(m+\lambda)t})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_k$$

and

$$(F^{\cdot(m+\lambda)t})_{il} = q-1\ \forall\ l \in C_k$$

Now, applying the isomorphism Ψ we have (see Remark 59)

$$(f^{(m+\lambda)t})_i(x) = \prod_{j \in C_k} x_j^{q-1}\ \forall\ i \in C_k$$

■

Theorem 137 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$ and $s \in \mathbb{N}$ a natural number. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ of V_f. Then any point $\xi \in \mathbf{F}_q^n$ showing s-periodicity under f, i.e. $f^s(\xi) = \xi$, satisfies the following property*

$$\xi_i = 1\ \forall\ i \in C_k \text{ or } \xi_i = 0\ \forall\ i \in C_k$$

Proof. Let $m \in \mathbb{N}$ be as in Lemma 136 and $u, v \in \mathbb{N}$ such that $ut = vs$. (This is always possible due to the existence of the $\operatorname{lcm}(s, t)$.) Now choose $\alpha \in \mathbb{N}$ such that $\alpha u > m$. Then we have $f^{\alpha ut} = f^{\alpha vs}$ and by Lemma 136

$$f_i^{\alpha vs}(x) = \prod_{j \in C_k} x_j^{q-1}\ \forall\ i \in C_k,\ k = 0, ..., t-1$$

In particular, for the s-periodic point ξ it holds for each $k \in \{0, ..., t-1\}$

$$f_i^{\alpha vs}(\xi) = (f^s)_i^{\alpha v}(\xi) = \xi_i = \prod_{j \in C_k} \xi_j^{q-1} \ \forall \ i \in C_k$$

In addition, according to eq. (2.3), $z^{q-1} = 1 \ \forall \ z \in \mathbf{F}_q \backslash \{0\}$. Therefore, it holds for each fixed $k \in \{0, ..., t-1\}$

$$\xi_i = 1 \ \forall \ i \in C_k \text{ or } \xi_i = 0 \ \forall \ i \in C_k$$

■

Theorem 138 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$. Then there is an $m \in \mathbb{N}$ such that $\forall \ \lambda \in \mathbb{N}$*

$$V_{\mathbf{F}_q}(f^t(x) - x) = V_{\mathbf{F}_q}(f^{(m+\lambda)t}(x) - x)$$

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ of V_f. By Remark 112 it holds

$$\widetilde{a}_k = \bigcup_{b \in \widetilde{a}_k} N_t(b)$$

From this fact we can conclude, that the matrix F^t has the following properties $\forall \ i \in C_k$, $k = 1, ..., t-1$

$$(F^t)_{ij} = 0 \ \forall \ j \in \{1, ..., n\} \backslash C_k$$

and

$$\exists \ l \in C_k : (F^t)_{il} \neq 0$$

By Remark 72 and 2. of Lemma 39 it follows $\forall \ i \in C_k$, $k = 1, ..., t-1$

$$(F^{\cdot t})_{ij} = 0 \ \forall \ j \in \{1, ..., n\} \backslash C_k$$

and

$$\exists \ l \in C_k : (F^{\cdot t})_{il} \neq 0$$

As before, we can conclude

$$(f^t)_i(x) = \prod_{j \in C_k} x_j^{\epsilon_i(j)} \ \forall \ i \in C_k$$

where ϵ_i are *nonzero* functions

$$\epsilon_i : C_k \to \{0, 1, ...q-1\} \subset \mathbb{N}$$

Now let $m \in \mathbb{N}$ be as in Lemma 136. Then by Lemma 136 we have $\forall \ \lambda \in \mathbb{N}$

$$f_i^{(m+\lambda)t}(x) = \prod_{j \in C_k} x_j^{q-1} \ \forall \ i \in C_k, \ k = 0, ..., t-1$$

By Theorem 137 any solution $\xi \in \mathbf{F}_q^n$ of the equation $f^{(m+\lambda)t}(x) = x$ satisfies

$$\xi_i = 1 \ \forall \ i \in C_k \text{ or } \xi_i = 0 \ \forall \ i \in C_k$$

Now, from the structure of the function f^t it is clear that any solution $\xi \in \mathbf{F}_q^n$ of the equation $f^{(m+\lambda)t}(x) = x$ also solves the equation $f^t(x) = x$. In other words

$$V_{\mathbf{F}_q}(f^{(m+\lambda)t}(x) - x) \subseteq V_{\mathbf{F}_q}(f^t(x) - x)$$

The inclusion

$$V_{\mathbf{F}_q}(f^t(x) - x) \subseteq V_{\mathbf{F}_q}(f^{(m+\lambda)t}(x) - x)$$

follows from the fact $f^{(m+\lambda)t} = (f^t)^{(m+\lambda)}$. The claim follows. ■

Corollary 139 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$. In addition, let T be the period number of f and $s \in D(T)$. If the phase space of f contains a cycle of length s, then s must divide $\mathcal{L}_{G_f}(V_f)$.*

Proof. By Corollary 123 there is an $\alpha \in \mathbb{N}$ such that

$$V_{\mathbf{F}_q}(f^s(x) - x) \subseteq V_{\mathbf{F}_q}(f^{\alpha t}(x) - x) \tag{2.24}$$

Now let $m \in \mathbb{N}$ be as in the previous Theorem. If $\alpha > m$ set $\beta := 1$ otherwise choose $\beta \in \mathbb{N}$ such that $\alpha\beta > m$. Then we have

$$V_{\mathbf{F}_q}(f^{\alpha t}(x) - x) \subseteq V_{\mathbf{F}_q}(f^{\alpha\beta t}(x) - x) \tag{2.25}$$

and from (2.24) and (2.25) and by the previous Theorem it follows

$$V_{\mathbf{F}_q}(f^s(x) - x) \subseteq V_{\mathbf{F}_q}(f^t(x) - x)$$

If the phase space of f contains a cycle of length s, i.e. if there are s *different* points $\xi_0, ..., \xi_{s-1} \in \mathbf{F}_q^n$ with

$$f(\xi_i) = \xi_{(i+1) \bmod s}$$

then from $\xi_0, ..., \xi_{s-1} \in V_{\mathbf{F}_q}(f^s(x) - x) \subseteq V_{\mathbf{F}_q}(f^t(x) - x)$ it follows that $s \leq t$ and s divides t. ■

Lemma 140 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$ and $s \in \mathbb{N}$ a natural number such that $s \leq t$. In addition, let $r \in \mathbb{N}$ be such that $\operatorname{lcm}(s,t) = rs$. Then the equation $f^s(x) = x$ has exactly $2^{\frac{t}{r}}$ solutions in $\mathbf{F}_2^n$.*

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices and $F := \Psi^{-1}(f)$ be the corresponding matrix of f. In addition, let $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ be the t loop equivalence classes ordered according to (2.19) and $C_0, ..., C_{t-1} \subset \{1, ..., n\}$ the partition of the set $\{1, ..., n\}$ induced by the partition $\widetilde{a}_0, ..., \widetilde{a}_{t-1} \subset V_f$ of V_f. We consider two cases. First, assume $\operatorname{lcm}(s,t) < st$, i.e. $\exists\, r \in \mathbb{N}$ with $r < t$ such that $\operatorname{lcm}(s,t) = rs$. Then, by Remark 114 the sets $\widetilde{a}_0, ..., \widetilde{a}_{t-1}$ can be arranged in $v := t/r$ families

$$\begin{gathered}
\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{rs \bmod t} \\
\widetilde{a}_1, \widetilde{a}_{1+s}, \widetilde{a}_{1+2s}, ..., \widetilde{a}_{(1+rs) \bmod t} \\
\vdots \\
\widetilde{a}_{v-1}, \widetilde{a}_{v-1+s}, \widetilde{a}_{v-1+2s}, ..., \widetilde{a}_{(v-1+rs) \bmod t}
\end{gathered}$$

where the vertices in $\widetilde{a}_{j+s}$ and $\widetilde{a}_{j+2s}$ are connected by sequences of length $\lambda t + s$. Moreover, no shorter family of this type can be constructed. From these facts we can conclude, that the matrix F^s has the following properties $\forall\, i \in C_k$, $k = u, u+s, u+2s, ..., u+(r-1)s$, $u = 0, ..., v-1$

$$(F^s)_{ij} = 0 \;\forall\; j \in \{1, ..., n\} \backslash C_{(k+s) \bmod t}$$

and

$$\exists\ l \in C_{(k+s) \bmod t} : (F^s)_{il} \neq 0$$

By Remark 72 and 2. of Lemma 39 it follows $\forall\ i \in C_k,\ k = 1, ..., t-1$

$$(F^{\cdot s})_{ij} = 0\ \forall\ j \in \{1, ..., n\} \backslash C_{(k+s) \bmod t}$$

and

$$\exists\ l \in C_{(k+s) \bmod t} : (F^{\cdot s})_{il} \neq 0$$

Now, applying the isomorphism Ψ we have (see Remark 59)

$$(f^s)_i(x) = \prod_{j \in C_{(k+s) \bmod t}} x_j^{\epsilon_i(j)}\ \forall\ i \in C_k$$

where ϵ_i are *nonzero* functions

$$\epsilon_i : C_{(k+s) \bmod t} \to \{0, 1, ...q-1\} \subset \mathbb{N}$$

As a consequence, for every fixed $u \in \{0, ..., v-1\}$ and $k = u, u+s, u+2s, ..., u+(r-1)s$ any solution $\xi \in \mathbf{F}_2^n$ of the equation $f^s(x) = x$ satisfies

$$\xi_i = \prod_{j \in C_{(k+s) \bmod t}} \xi_j^{\epsilon_i(j)}\ \forall\ i \in C_k \tag{2.26}$$

By Theorem 137 we also know that

$$\xi_i = 1\ \forall\ i \in C_k \text{ or } \xi_i = 0\ \forall\ i \in C_k,\ \ k = u, u+s, u+2s, ..., u+(r-1)s \tag{2.27}$$

Now, if $\xi_i = 1\ \forall\ i \in C_u$, by (2.26) and (2.27), it must follow, that $\xi_l = 1\ \forall\ l \in C_{(u+s) \bmod t}$. The same argument applied $r-1$ times lets us conclude $\xi_i = 1\ \forall\ i \in C_k,\ k = u, u+s, u+2s, ..., u+(r-1)s$. If, on the other hand, $\xi_i = 0\ \forall\ i \in C_u$, by (2.26) we have that $\exists\ v \in C_{(u+s) \bmod t} : \xi_v = 0$ and by (2.27) $\xi_l = 0\ \forall\ l \in C_{(u+s) \bmod t}$. The same argument applied $r-1$ times lets us conclude $\xi_i = 0\ \forall\ i \in C_k,\ k = u, u+s, u+2s, ..., u+(r-1)s$. Summarizing, since every $u \in \{0, ..., v-1\}$ represents one of the above v families, there are exactly $2^v = 2^{\frac{t}{r}}$ solutions of $f^s(x) = x$ in $\mathbf{F}_q^n$.
The second case is when $\operatorname{lcm}(s,t) = ts$. Here, by Remark 114 the sets $\widetilde{a}_0, ..., \widetilde{a}_{t-1}$ can be arranged in one single family

$$\widetilde{a}_0, \widetilde{a}_s, \widetilde{a}_{2s}, ..., \widetilde{a}_{ts \bmod t}$$

The same argument as used above for a fixed value of u yields that, in this case, the only solutions of $f^s(x) = x$ in $\mathbf{F}_q^n$ are $(1, ..., 1), (0, ..., 0) \in \mathbf{F}_q^n$. Therefore, the number of solutions is equal to $2 = 2^{\frac{t}{t}}$. ■

Corollary 141 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f,\ E_f,\ \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) \geq 1$ and $s \in \mathbb{N}$ a natural number such that $s \in D(t)$. Then the equation $f^s(x) = x$ has exactly 2^s solutions in $\mathbf{F}_q^n$.*

Proof. Since s divides t, $\exists\ r \in \mathbb{N} : t = rs$. Thus $\operatorname{lcm}(t,s) = t = rs$ and $t/r = s$. The claim follows from the previous Lemma. ■

Remark 142 *In particular, if $\{1, d_1, ..., d_u, t\} \subset \mathbb{N}$ is the set of divisors of t in ascending order, then the number of solutions of $f^s(x) = x$ in $\mathbf{F}_q^n$ grows monotonically from 2^1 to 2^t for $s = 1, d_1, ..., d_u, t$. More generally, if $s < t$, then $\operatorname{lcm}(s,t) > s$ and thus $\operatorname{lcm}(s,t) = rs$ with $r \geq 2$. As a consequence of the previous Lemma, for $s \in \{1, ..., t\}$ the number of solutions in $\mathbf{F}_q^n$ of the equation $f^s(x) = x$ takes its maximal value for $s = t$.*

The next theorem shows that in the case of a $(q-1)$-fold redundant coupled monomial dynamical system, period number and loop number coincide, provided the dependency graph is strongly connected:

Theorem 143 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) > 1$. Then the period number T of f satisfies*

$$T = \mathcal{L}_{G_f}(V_f)$$

Moreover, the phase space of f contains cycles of all lengths $s \in D(T)$.

Proof. By Corollary 139 the length s of any cycle displayed in the phase space of f divides t, in particular, it holds $s \leq t$. Now let $\{d_0 := 1, d_1, ..., d_u, d_{u+1} := t\} \subset \mathbb{N}$ be the set of divisors of t in ascending order. By Remark 142 we know that

$$\left|V_{\mathbf{F}_2}(f^{d_i}(x) - x)\right| > \left|V_{\mathbf{F}_2}(f^{d_j}(x) - x)\right| \ \forall\ i, j \in \{1, ..., u+1\} : i > j$$

Therefore, the phase space of f indeed contains cycles of length $d_i\ \forall\ i \in \{1, ..., u+1\}$. Summarizing we can say that the the phase space of f only contains cycles of length $d_i\ \forall\ i \in \{1, ..., u+1\}$. From the definition we know $T = \mathrm{lcm}(1, d_1, ..., d_u, t)$ and thus

$$T = \mathcal{L}_{G_f}(V_f)$$

■

Theorem 144 (and Definition) *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a $(q-1)$-fold redundant coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) > 1$. In addition, let $s \in \mathbb{N}$ be a natural number and denote by Z_s the number of cycles of length s displayed by the phase space of f. Then it holds for any $d \in \mathbb{N}$*

$$Z_d = \begin{cases} \frac{2^d - \sum\limits_{j \in D(d) \setminus d} Z_j}{d} \text{ if } d \in D(t) \\ 0 \text{ if } d \notin D(t) \end{cases}$$

Proof. The claim follows immediately from Theorem 143 and Corollary 141. ■

Remark 145 *In particular, if the loop number $t = \mathcal{L}_{G_f}(V_f)$ is a prime number, then the phase space of f only displays cycles of length t and 1 (fixed points). More precisely*

$$Z_t = \frac{2^t - 2}{t}$$

and

$$Z_1 = 2$$

Example 146 *Consider the system $f \in MF_6^6(\mathbf{F}_5)$ defined as*

$$\begin{aligned} f_1(x) &: = x_2 \\ f_2(x) &: = x_3 \\ f_3(x) &: = x_4 \\ f_4(x) &: = x_5^2 \\ f_5(x) &: = x_6^2 \\ f_6(x) &: = x_1 \end{aligned}$$

It can be easily shown, that the dependency graph $G_f = (V_f, E_f, \pi_f)$ of f is strongly connected with loop number $\mathcal{L}_{G_f}(V_f) = 6$. It is also clear that f is 4-fold redundant. As a consequence and by Theorem 144, the period number of f is equal to 6 and the phase space of f displays 9 cycles of length 6, 2 cycles of length 3, 1 cycle of length 2 and 2 fixed points.

Theorem 147 *Let $\mathbf{F}_q$ be a finite field, $f \in MF_n^n(\mathbf{F}_q)$ a coupled monomial dynamical system and $G_f = (V_f, E_f, \pi_f)$ its dependency graph. Furthermore, let G_f be strongly connected with loop number $t := \mathcal{L}_{G_f}(V_f) > 1$. If for arbitrary $b, c \in V_f$ it holds*

$$s_1(b, c) \neq 0 \Rightarrow s_1(b, c) = q - 1$$

Then the period number T of f satisfies

$$T = \mathcal{L}_{G_f}(V_f)$$

and the phase space of f contains cycles of all lengths $s \in D(T)$.

Proof. Let $V_f = \{a_1, ..., a_n\}$ be the numeration of the vertices. Consider two vertices $a_i, a_j \in V_f$ (a_i is recurrently connected to a_j). Then, for any $r \in \mathbb{N}$ such that there is a sequence $a_i \rightsquigarrow_r a_j$ there are $(q-1)^r$ different copies of it and we can conclude $\exists\ \alpha_{ijr} \in \mathbb{N}$ such that

$$s_r(a_i, a_j) = \alpha_{ijr}(q - 1)$$

Consequently, f is a $(q-1)$-fold redundant coupled monomial dynamical system and the claim follows from Theorem 143. ■

Example 148 *Let $\mathbf{F}_q$ be a finite field and consider the system $f \in MF_7^7(\mathbf{F}_q)$ defined as*

$$\begin{aligned}
f_1(x) &: = x_4^{q-1} \\
f_2(x) &: = x_5^{q-1} \\
f_3(x) &: = x_4^{q-1} \\
f_4(x) &: = x_7^{q-1} \\
f_5(x) &: = x_6^{q-1} \\
f_6(x) &: = x_3^{q-1} \\
f_7(x) &: = x_1^{q-1} x_2^{q-1}
\end{aligned}$$

It can be easily shown, that the dependency graph $G_f = (V_f, E_f, \pi_f)$ of f is strongly connected with loop number $\mathcal{L}_{G_f}(V_f) = 3$. Moreover, f satisfies the condition of the previous theorem. As a consequence and by Theorem 144, the period number of f is equal to 3 and the phase space of f displays two cycles of length 3 and two fixed points.

Remark 149 *The result of the previous theorem is not surprising, since such a monomial system immediately maps any point $\xi \in \mathbf{F}_q^n$ into a point $f(\xi) \in \mathbf{F}_2^n$.*

The study of the cycle structure of more general classes of monomial systems (with or without strongly connected dependency graph) remains the subject of mathematical research. Now, we leave this exciting area to turn our attention to the study of monomial control systems in the next chapter.

Chapter 3

Monomial control systems over a finite field

In the previous chapter, for monomial dynamical systems over a finite field, criteria were presented by means of which the period number as well as the cycle structure of the phase space can be determined. As commonly in mathematics, concessions have to be made in order to obtain strong propositions. Our concessions affect the cardinality of the finite field used and/or the topology of the dependency graph. Therefore, the class of monomial control systems (to be defined below) that we will study in depth will be constrained regarding the underlying finite field and/or the topology of the dependency graph.

When dealing with non-autonomous control systems, control inputs are at one's disposal which can be used for controlling the state evolution. Furthermore, equipped with knowledge about the current state — provided by measurement, for example — the input can be related to an appropriate function of the state, a so-called (static) control law, in order to synthesize desired system properties in a feedback control loop. By virtue of a control law the closed-loop system is rendered autonomous.

Due to the fact that a monomial control system remains monomial under monomial state feedback, the resulting autonomous closed-loop system can be analyzed with the methods presented in the previous Chapter. If the purpose of control is to guarantee certain closed-loop properties then a natural question is to ask for criteria about the existence of a suitable state feedback, and subsequently, how this suitable control law can be chosen.

Throughout this chapter, and in contrast to Chapter 1 and Part II of this thesis, we will denote the elements of the Cartesian product $\mathbf{F}_q^n$ as $x \in \mathbf{F}_q^n$, neglecting the vector arrow.

3.1 General definitions and control theoretic questions studied

Definition 150 *Let $\mathbf{F}_q$ be a finite field, $n \in \mathbb{N}$ a natural number and $m \in \mathbb{N}_0$ a nonnegative integer. A mapping*

$$g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$$

is called time invariant monomial control system over F_q *if for every $i \in \{1, ..., n\}$ there exist two tuples $(A_{i1}, ..., A_{in}) \in E_q^n$ and $(B_{i1}, ..., B_{im}) \in E_q^m$ such that*

$$g_i(x, u) = x_1^{A_{i1}} ... x_n^{A_{in}} u_1^{B_{i1}} ... u_m^{B_{im}} \ \forall \ (x, u) \in \mathbf{F}_q^n \times \mathbf{F}_q^m$$

Remark 151 *In the case $m = 0$, we have $\mathbf{F}_q^m = \mathbf{F}_q^0 = \{()\}$ (the set containing the empty tuple) and thus $\mathbf{F}_q^n \times \mathbf{F}_q^m = \mathbf{F}_q^n \times \mathbf{F}_q^0 = \mathbf{F}_q^n \times \{()\} = \mathbf{F}_q^n$. In other words, g is a monomial dynamical system over $\mathbf{F}_q$. From now on we will refer to a time invariant monomial control system over F_q as monomial control system over F_q.*

Definition 152 *Let X be a nonempty finite set and $n, l \in \mathbb{N}$ natural numbers. The set of all functions*

$$f : X^l \to X^n$$

is denoted with $F_l^n(X)$.

Definition 153 *Let $\mathbf{F}_q$ be a finite field and $l, m, n \in \mathbb{N}$ natural numbers. Furthermore, let E_q be the exponents semiring of $\mathbf{F}_q$ and $M(n \times l;\, E_q)$ the set of $n \times l$ matrices with entries in E_q. Consider the map*

$$\begin{array}{rcl} \Gamma & : & F_m^l(\mathbf{F}_q) \times M(n \times l; E_q) \to F_m^n(\mathbf{F}_q) \\ (f, A) & \mapsto & \Gamma_A(f) \end{array}$$

where $\Gamma_A(f)$ is defined for every $x \in \mathbf{F}_q^m$ and $i \in \{1, ..., n\}$ by

$$\Gamma_A(f)(x)_i := f_1(x)^{A_{i1}} ... f_l(x)^{A_{il}}$$

We denote the mapping $\Gamma_A(f) \in F_m^n(\mathbf{F}_q)$ simply Af.

Remark 154 *Let $l = m$, $id \in F_m^m(\mathbf{F}_q)$ be the identity map (i.e. $id_i(x) = x_i \ \forall \ i \in \{1, ..., m\}$) and $A \in M(n \times m; E_q)$ Then the following relationship between the mapping $Aid \in F_m^n(\mathbf{F}_q)$ and any $f \in F_m^m(\mathbf{F}_q)$ holds*

$$Aid(f(x)) = Af(x) \ \forall \ x \in \mathbf{F}_q^m$$

Remark 155 *Consider the case $l = m = n$. For every monomial dynamical system $f \in MF_n^n(\mathbf{F}_q) \subset F_n^n(\mathbf{F}_q)$ with corresponding matrix $F := \Psi^{-1}(f) \in M(n \times n; E_q)$ it holds*

$$Fid = f$$

On the other hand, given a matrix $F \in M(n \times n; E_q)$ we have

$$\Psi^{-1}(Fid) = F$$

Moreover, the map $\Gamma : F_n^n(\mathbf{F}_q) \times M(n \times n; E_q) \to F_n^n(\mathbf{F}_q)$ is an action of the multiplicative monoid $M(n \times n;\, E_q)$ on the set $F_n^n(\mathbf{F}_q)$. It holds namely, that $If = f \ \forall \ f \in F_n^n(\mathbf{F}_q)$ (which is trivial) and $(A \cdot B)f = A(Bf) \ \forall \ f \in F_n^n(\mathbf{F}_q)$, $A, B \in M(n \times n; E_q)$. To see this consider

$$\begin{array}{rcl} ((A \cdot B)f)_i(x) & = & f_1(x)^{(A \cdot B)_{i1}} ... f_n(x)^{(A \cdot B)_{in}} \\ & = & \prod_{j=1}^{n} f_j(x)^{(A_{i1} \bullet B_{1j} \oplus ... \oplus A_{in} \bullet B_{nj})} \\ & = & (Aid \circ Bid)_i(f(x)) \\ & = & (Aid)_i(Bid(f(x))) \\ & = & (Aid)_i(fB(x)) \\ & = & (A(Bf))_i(x) \end{array}$$

where $id \in F_n^n(\mathbf{F}_q)$ is the identity map (i.e. $id_i(x) = x_i \ \forall \ i \in \{1, ..., n\}$). (cf. with the proof of Theorem 57). As a consequence, $MF_n^n(\mathbf{F}_q)$ is the orbit in $F_n^n(\mathbf{F}_q)$ of id under the monoid $M(n \times n; E_q)$. In particular (see Theorem 57), we have

$$(F \cdot G)id = F(Gid) = f \circ g$$

where $g \in MF_n^n(\mathbf{F}_q)$ is another monomial dynamical system with corresponding matrix $G := \Psi^{-1}(g) \in M(n \times n; E_q)$. This means that the set $MF_n^n(\mathbf{F}_q)$ is also an $M(n \times n;\, E_q)-$set.

Lemma 156 (and Definition) *Let $\mathbf{F}_q$ be a finite field, $n \in \mathbb{N}$ a natural number and $m \in \mathbb{N}_0$ a nonnegative integer. Furthermore, let $id \in F_{(n+m)}^{(n+m)}(\mathbf{F}_q)$ be the identity map (i.e. $id_i(x) = x_i$ $\forall\ i \in \{1, ..., n+m\}$) and $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ a monomial control system over F_q. Then there are matrices $A \in M(n \times n;\ E_q)$ and $B \in M(n \times m;\ E_q)$ such that*

$$((A|B)id)(x,u) = g(x,u)\ \forall\ (x,u) \in \mathbf{F}_q^n \times \mathbf{F}_q^m$$

Where $(A|B) \in M(n \times (n+m);\ E_q)$ is the matrix that results by writing A and B side by side. In this sense we denote g as the monomial control system (A, B) with n state variables and m control inputs.

Proof. This follows immediately from the previous Definitions. ■

Remark 157 (and Definition) *If the matrix $B \in M(n \times m;\ E_q)$ is equal to the zero matrix, then g is called a* control system with no controls. *In contrast to linear control systems (see, for instance, [106], and in the framework of finite fields, [92]), when the input vector $u \in \mathbf{F}_q^m$ satisfies*

$$u = \vec{1} := (1, ..., 1)^t \in\ \mathbf{F}_q^m$$

then no control input is being applied on the system, i.e. the monomial dynamical system over $\mathbf{F}_q$

$$\begin{aligned} \sigma &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto g(x, \vec{1}) \end{aligned}$$

satisfies

$$\sigma(x) = ((A|0)id)(x,u)\ \forall\ (x,u) \in \mathbf{F}_q^n \times \mathbf{F}_q^m$$

where $0 \in M(n \times m;\ E_q)$ stands for the zero matrix.

Definition 158 *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers. A* monomial feedback controller *is a mapping*

$$f : \mathbf{F}_q^n \to \mathbf{F}_q^m$$

such that for every $i \in \{1, ..., m\}$ there exists a tuple $(F_{i1}, ..., F_{in}) \in E_q^n$ such that

$$f_i(x) = x_1^{F_{i1}} ... x_n^{F_{in}}\ \forall\ x \in \mathbf{F}_q^n$$

Remark 159 *We exclude in the definition of monomial feedback controller the possibility that one of the functions f_i is equal to the zero function. The reason for this is that we want to be able to use the same formalism developed for monomial dynamical systems in the previous Chapter (see Remark 31). This convention does not represent an impediment for our goals.*

Now we are able to formulate the first control theoretic problem to be addressed in this thesis:

Problem 160 *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers. Given a monomial control system $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ with measurable state, design a monomial state feedback controller $f : \mathbf{F}_q^n \to \mathbf{F}_q^m$ such that the closed-loop system*

$$\begin{aligned} h &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto g(x, f(x)) \end{aligned}$$

has a desired period number and cycle structure of its phase space. What properties has g to fulfill for this task to be accomplished?

Remark 161 *Note that every component*

$$\begin{array}{rcl} h_i & : & \mathbf{F}_q^n \to \mathbf{F}_q,\ i=1,...,n \\ x & \mapsto & g_i(x, f(x)) \end{array}$$

is a nonzero monic monomial function, i.e. the mapping $h : \mathbf{F}_q^n \to \mathbf{F}_q^n$ is a monomial dynamical system over $\mathbf{F}_q$. As a consequence, the results achieved in the previous Chapter can be used to analyze the dynamical properties of h. Moreover, the following identity holds

$$h = (A + B \cdot F)id$$

where $F \in M(m \times n; E_q)$ is the corresponding matrix of f (see Remark 51), (A, B) are the matrices in Lemma 156 and $id \in F_n^n(\mathbf{F}_q)$ (see Corollary 58). To see this, consider the mapping

$$\begin{array}{rcl} \mu & : & \mathbf{F}_q^m \to \mathbf{F}_q^n \\ u & \mapsto & g(\vec{1}, u) \end{array}$$

where $\vec{1} \in \mathbf{F}_q^n$. From the definition of g it follows that $\mu \in MF_m^n(\mathbf{F}_q)$. Now, since $f \in MF_n^m(\mathbf{F}_q)$, by Remark 51 we have for the composition $\mu \circ f : \mathbf{F}_q^n \to \mathbf{F}_q^n$

$$\mu \circ f = (B \cdot F)id$$

Now its easy to see

$$h = (A + B \cdot F)id$$

Besides the results about $(q-1)$-fold redundant systems, the most significant results proved in the previous Chapter concern Boolean monomial dynamical systems with a strongly connected dependency graph. Therefore, in the next Section we will focus on the solution of Problem 160 for Boolean monomial control systems $g : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2^n$ with the property that the mapping

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

has a strongly connected dependency graph. Although the above representation

$$h = (A + B \cdot F)id$$

of the closed loop system displays a striking structural similarity with linear control systems and linear feedback laws, our approach will completely differ from the well known "Pole-Assignment" method, (see, for instance, [106]).

3.2 Controller design for Boolean monomial control systems

3.2.1 The principle of loop number assignment and first general results

One of the most important results of the previous Chapter is Theorem 131, which states that the loop number of the (strongly connected) dependency graph of a Boolean monomial dynamical system completely determines the period number of the system and its cycle structure. In light of the representation

$$h = (A + B \cdot F)id$$

of the closed loop system

$$\begin{array}{rcl} h_i & : & \mathbf{F}_q^n \to \mathbf{F}_q,\ i=1,...,n \\ x & \mapsto & g_i(x, f(x)) \end{array}$$

(and assuming that the dependency graph of h is strongly connected) the question arises as to how the loop number of the dependency graph of h can be modified by different choices of the matrix F corresponding to the monomial feedback controller f used. In other words, Theorem 131 tells us that the dynamical properties of h could be engineered through a proper *loop number assignment* by means of choosing a suitable matrix F. Of course, the ability to find such an F will be restricted by the shape of the matrices A and B, as illustrated by the example $B = 0$, for which obviously no such F exists. In order to establish which properties A and B must fulfill to make the loop number assignment possible via a suitable choice of F, we will start investigating the variation of the loop number of a strongly connected graph when new directed edges are added to it:

Lemma 162 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial and $t := \mathcal{L}_G(V_G) > 0$ its loop number. Furthermore, let $G' = (V_{G'}, E_{G'}, \pi_{G'})$ be a strongly connected digraph such that $V_G = V_{G'}$, $E_{G'} \supseteq E_G$ and $\pi_{G'}(e) = \pi_G(e)\ \forall\ e \in E_G$. Then the loop number $t' := \mathcal{L}_{G'}(V_{G'}) = \mathcal{L}_{G'}(V_G)$ of G' must divide the loop number $\mathcal{L}_G(V_G)$ of G.*

Proof. From the definition of loop number it follows immediately

$$t' \leq t$$

Let $V_G = \{a_1, ..., a_n\}$ be a numeration of the vertices of G. Furthermore, let $U := \{a_{i_1}, ..., a_{i_k}\} \subseteq V_G$ be the subset of vertices of V_G such that there is a closed path $a_{i_j} \rightsquigarrow_{l(j)} a_{i_j}$ contained in the graph G. By Theorem 116 it holds

$$\mathcal{L}_G(V_G) = \gcd(l(1), ..., l(k))$$

Since the graph G is a subgraph of G', by Corollary 106 for each $j \in \{1, ..., k\}\ \exists\ \alpha_j \in \mathbb{N}$ such that

$$l(j) = \alpha_j t'$$

and thus

$$t = \gcd(\alpha_1 t', ..., \alpha_k t') = t' \gcd(\alpha_1, ..., \alpha_k)$$

■

From the representation $h = (A + B \cdot F)id$ it is easy to see that the dependency graph $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ of the system (see Remark 157)

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

is isomorphic to a subgraph G'_h of the dependency graph $G_h = (V_h, E_h, \pi_h)$ of the system h containing all vertices V_h. Moreover, the bijective correspondence between the sets V_σ and V_h defines pairs of vertices that correspond to the same variable. *In what follows, we won't make any distinction between the vertices V_σ and V_h and we will see G_σ as a subgraph of G_h with $V_\sigma = V_h$.* Having stated that, we are able to clarify what is feasible when it comes to solutions of Problem 160 in view of the available mathematical results and tools:

Theorem 163 *Let $\mathbf{F}_2$ be the finite field with two elements, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2^n$ a* Boolean *monomial control system such that the dependency graph of the system*

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

is strongly connected. Furthermore, let $f : \mathbf{F}_2^n \to \mathbf{F}_2^m$ *be an arbitrary (Boolean) monomial feedback controller,* $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ *the dependency graph of the system* σ *and* $G_h = (V_h, E_h, \pi_h)$ *the dependency graph of the closed loop system*

$$\begin{aligned} h &: \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x &\mapsto g(x, f(x)) \end{aligned}$$

Then the loop number $\mathcal{L}_{G_h}(V_h)$ *divides the loop number* $\mathcal{L}_{G_\sigma}(V_\sigma)$.

Proof. By Lemma 156 there are matrices $A \in M(n \times n;\, E_2)$ and $B \in M(n \times m;\, E_2)$ such that

$$((A|B)id)(x, u) = g(x, u) \;\forall\; (x, u) \in \mathbf{F}_2^n \times \mathbf{F}_2^m$$

Let $F \in M(m \times n; E_2)$ be the corresponding matrix of f (see Remark 51). By Remark 161 we know

$$h = (A + B \cdot F)id$$

Therefore, the dependency graph $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ of the system σ is isomorphic to a subgraph G'_h of the dependency graph $G_h = (V_h, E_h, \pi_h)$ of the system h containing all vertices V_h. Since isomorphic strongly connected graphs must have the same loop number, the claim follows from the previous Lemma 162. ■

Remark 164 *The question as to how the loop number of the dependency graph of* h *can be modified by different choices of the matrix* F *(corresponding to the monomial feedback controller* f *used) can now be partially answered: It can be only modified to values contained in the set* $D(\mathcal{L}_{G_\sigma}(V_\sigma))$ *(see Definition 122). In particular, if* $\mathcal{L}_{G_\sigma}(V_\sigma)$ *is a prime number, it is only possible to stabilize the system* σ *by making* h *a fixed point system via a suitable choice of a monomial feedback law. If* $\mathcal{L}_{G_\sigma}(V_\sigma) = 1$ *then* $\mathcal{L}_{G_h}(V_h) = 1$ *no matter what feedback law is chosen. In the next section we will formulate one necessary and sufficient conditions on the matrix* B *for the loop number* $\mathcal{L}_{G_h}(V_h)$ *of* h *to be modified among the possible set of values* $D(\mathcal{L}_{G_\sigma}(V_\sigma))$.

3.2.2 Controllability of Boolean strongly dependent monomial control systems

Definition 165 *Let* $\mathbf{F}_q$ *be a finite field and* $n, m \in \mathbb{N}$ *natural numbers. A monomial control system* $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ *such that the dependency graph of the system*

$$\begin{aligned} \sigma &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto g(x, \vec{1}) \end{aligned}$$

is strongly connected is called a strongly dependent monomial control system.

Definition 166 *Let* $\mathbf{F}_q$ *be a finite field,* $n, m \in \mathbb{N}$ *natural numbers and* $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ *a strongly dependent monomial control system such that the dependency graph* $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ *of the system*

$$\begin{aligned} \sigma &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto g(x, \vec{1}) \end{aligned}$$

has loop number $t := \mathcal{L}_{G_\sigma}(V_\sigma) > 1$. *Furthermore, let* $f : \mathbf{F}_q^n \to \mathbf{F}_q^m$ *be an arbitrary monomial feedback controller and* $G_{h_f} = (V_{h_f}, E_{h_f}, \pi_{h_f})$ *the dependency graph of the closed loop system*

$$\begin{aligned} h_f &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto g(x, f(x)) \end{aligned}$$

and $t' \in D(\mathcal{L}_{G_\sigma}(V_\sigma))\backslash\{\mathcal{L}_{G_\sigma}(V_\sigma)\}$ *be a divisor of* $\mathcal{L}_{G_\sigma}(V_\sigma)$ *with* $t' < \mathcal{L}_{G_\sigma}(V_\sigma)$. *The system* g *is called* controllable to loop number t' *if the loop number* $\mathcal{L}_{G_{h_f}}(V_{h_f})$ *of* h_f *can be forced to take the value* t' *(see Theorem 163) by means of choosing a suitable monomial feedback controller* f.

Definition 167 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ a strongly dependent monomial control system such that the dependency graph $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ of the system*

$$\begin{aligned} \sigma &: \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x &\mapsto g(x, \vec{1}) \end{aligned}$$

has loop number $t := \mathcal{L}_{G_\sigma}(V_\sigma) > 1$. Furthermore, let $f : \mathbf{F}_q^n \to \mathbf{F}_q^m$ be an arbitrary monomial feedback controller. The system g is called completely loop number controllable *if it is controllable to loop number t' for any $t' \in D(\mathcal{L}_{G_\sigma}(V_\sigma))\backslash\{\mathcal{L}_{G_\sigma}(V_\sigma)\}$.*

Now we introduce some technical definitions that will help us formulate our theorems about loop number controllability:

Definition 168 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $a \in V_G$ be a fixed but arbitrary vertex. The* directed distance

$$d(a \rightsquigarrow b) \in \mathbb{N}_0$$

from a to any other arbitrary vertex $b \in V_G$ is defined as the length of the shortest path connecting a to b. We set $d(a \rightsquigarrow a) := 0 \ \forall \ a \in V_G$. Note that the directed distance is not symmetric.

Definition 169 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number. Consider a fixed but arbitrary loop equivalence class $\widetilde{c} \subset V_G$. The* upstream distance *from $\widetilde{a}$ to any other loop equivalence class $\widetilde{d} \subset V_G$ is defined as*

$$d(\widetilde{c} \rightsquigarrow \widetilde{d}) := \min_{\substack{a \in \widetilde{c} \\ b \in \widetilde{d}}} (d(a \rightsquigarrow b))$$

Note that the upstream distance is not symmetric.

Remark 170 *By Remark 112 it holds for every loop equivalence class $\widetilde{c} \subset V_G$*

$$d(\widetilde{c} \rightsquigarrow \widetilde{c}) = 0$$

and for two different *loop classes $\widetilde{a}, \widetilde{b} \subset V_G$*

$$d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{a}) = \mathcal{L}_G(V_G)$$

Lemma 171 *Let $G = (V_G, E_G, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number. Consider three arbitrary loop equivalence classes $\widetilde{a}, \widetilde{b}, \widetilde{c} \subset V_G$. Then there is a $\lambda \in \mathbb{N}_0$ such that*

$$d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) = \lambda t$$

Proof. If only two of the three classes are equal we have $d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) = t$, if all three are equal we have $d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) = 0$. Now assume all three are different. We distinguish between two cases: $d(\widetilde{b} \rightsquigarrow \widetilde{c}) < d(\widetilde{b} \rightsquigarrow \widetilde{a})$ and $d(\widetilde{b} \rightsquigarrow \widetilde{c}) > d(\widetilde{b} \rightsquigarrow \widetilde{a})$ (if $d(\widetilde{b} \rightsquigarrow \widetilde{c}) = d(\widetilde{b} \rightsquigarrow \widetilde{a})$, then by Remark 112 it would follow $\widetilde{a} = \widetilde{c}$) In the first case we have

$$d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) = t$$

In the second case it holds $d(\widetilde{b} \rightsquigarrow \widetilde{c}) = d(\widetilde{b} \rightsquigarrow \widetilde{a}) + d(\widetilde{a} \rightsquigarrow \widetilde{c})$ and we have

$$\begin{aligned} d(\widetilde{a} &\rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) \\ &= d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{a}) + d(\widetilde{a} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) \\ &= 2t \end{aligned}$$

■

Lemma 172 *Let $G = (V_G,\, E_G,\, \pi_G)$ be a strongly connected digraph such that V_G is nontrivial. Furthermore, let $t := \mathcal{L}_G(V_G) > 0$ be its loop number. Consider three arbitrary loop equivalence classes $\widetilde{a}, \widetilde{b}, \widetilde{c} \subset V_G$. Then there is a $\alpha \in \mathbb{N}_0$ such that*

$$d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) = d(\widetilde{a} \rightsquigarrow \widetilde{c}) + \alpha t$$

Proof. By Remark 170 we have

$$-d(\widetilde{a} \rightsquigarrow \widetilde{c}) = d(\widetilde{c} \rightsquigarrow \widetilde{a}) - t$$

and thus

$$d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) - d(\widetilde{a} \rightsquigarrow \widetilde{c}) = d(\widetilde{a} \rightsquigarrow \widetilde{b}) + d(\widetilde{b} \rightsquigarrow \widetilde{c}) + d(\widetilde{c} \rightsquigarrow \widetilde{a}) - t$$

Now the claim follows by the previous lemma. ■

Definition 173 *Let $\mathbf{F}_2$ be the finite field with two elements, $n \in \mathbb{N}$ a natural number and $v \in \mathbf{F}_2^n$ a tuple. The* one set *of v is defined as*

$$E(v) := \{i \in \{1, ..., n\} : v_i = 1\} \subseteq \{1, ..., n\}$$

We will refer to the problem of controllability to loop number 1 as the *stabilization problem* and treat it separately (see below).

Theorem 174 *Let $\mathbf{F}_2$ be the finite field with two elements, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2^n$ a* Boolean *strongly dependent monomial control system such that the dependency graph $G_\sigma = (V_\sigma,\, E_\sigma,\, \pi_\sigma)$ of the system*

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

has loop number $t := \mathcal{L}_{G_\sigma}(V_\sigma) > 1$. Furthermore, let $A \in M(n \times n;\, E_2)$ and $B \in M(n \times m;\, E_2)$ be the matrices such that

$$((A|B)id)(x, u) = g(x, u) \;\forall\; (x, u) \in \mathbf{F}_2^n \times \mathbf{F}_2^m$$

(see Lemma 156) and $t' \in D(\mathcal{L}_{G_\sigma}(V_\sigma))\backslash\{1, \mathcal{L}_{G_\sigma}(V_\sigma)\}$ a divisor of $\mathcal{L}_{G_\sigma}(V_\sigma)$ with $1 < t' < \mathcal{L}_{G_\sigma}(V_\sigma)$. In addition, let $V_\sigma = \{a_1, ..., a_n\}$ be the numeration of the vertices and $\widetilde{a_1}, ..., \widetilde{a_n} \subseteq V_\sigma$ their corresponding loop equivalence classes. If g is controllable to loop number t' then B contains a column B_j different from zero with the property that there is a loop equivalence class $\widetilde{a_k} \subseteq \mathcal{L}_{G_\sigma}(V_\sigma)$ such that $\forall\; s \in E(B_j)\; \exists\; \alpha_s \in \mathbb{N}$ with

$$1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s}) = \alpha_s t'$$

Proof. If g is controllable to loop number t', there is a monomial feedback controller $f^* : \mathbf{F}_2^n \to \mathbf{F}_2^m$ such that the dependency graph $G_{h_{f^*}} = (V_{h_{f^*}},\, E_{h_{f^*}},\, \pi_{h_{f^*}})$ of the corresponding closed loop system

$$\begin{array}{rcl} h_{f^*} & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, f^*(x)) \end{array}$$

has loop number $\mathcal{L}_{G_{h_{f^*}}}(V_{h_{f^*}}) = t'$. Let $F^* \in M(m \times n; E_2)$ be the corresponding matrix of f^* (see Remark 51). From the representation

$$h_{f^*} = (A + B \cdot F^*)id$$

and the fact $\mathcal{L}_{G_{h_{f^*}}}(V_{h_{f^*}}) = t' < \mathcal{L}_{G_\sigma}(V_\sigma)$ it follows that at least one entry $(B \cdot F^*)_{il}$ of the matrix $B \cdot F^*$ must be nonzero. If $(B \cdot F^*)_{il} \neq 0$, from the definition of matrix product we can conclude

that $\exists\ j \in \{1,...,m\}$ such that $B_{ij} = 1$ and $F^*_{jl} = 1$. As a consequence, $f^*_j(x) = \prod_{k\in\Omega} x_k\ \forall\ x \in \mathbf{F}_2^n$ with $\emptyset \neq \Omega \subseteq \{1,...,n\}$. Now, consider the column B_j of B. For every $s \in E(B_j)$ the expression $(h_{f^*})_s(x)$ contains the factor $f^*_j(x) = \prod_{k\in\Omega} x_k\ \forall\ x \in \mathbf{F}_2^n$. This means that the graph $G_{h_{f^*}}$ contains the edges $a_s \to a_k$, $k \in \Omega$. (However, since $t' > 1$, no self loop $a_s \to a_s$ can exist in the graph $G_{h_{f^*}}$). Furthermore, consider in the graph G_σ the loop equivalence classes $\widetilde{a_s}, \widetilde{a_k} \subseteq V_\sigma$, $s \in E(B_j)$, $k \in \Omega$. From the cyclic loop equivalence classes structure of this graph (see Remark 112) we know that for every $s \in E(B_j)$ we can construct a closed path $a_s \rightsquigarrow_{u_s} a_s$ in the graph $G_{h_{f^*}}$ of length

$$u_s = 1 + \lambda_s t + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s})$$

where $\lambda_s \in \mathbb{N}_0$. Now, since t' divides t, $\exists\ c \in \mathbb{N} : t = ct'$. Thus

$$u_s = \lambda_s ct' + 1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s})$$

Moreover, since u_s is the length of a closed path in the graph $G_{h_{f^*}}$, by Theorem 116 t' must divide u_s and therefore t' must divide $1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s})\ \forall\ s \in E(B_j)$ and $\forall\ k \in \Omega$. ■

Remark 175 *The previous theorem also holds in the case $q > 2$, if the definition of $E(v)$ is modified accordingly. However, the control theoretic analysis of non-Boolean control systems would go beyond the scope of this dissertation.*

Theorem 176 *Let $\mathbf{F}_2$ be the finite field with two elements, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2^n$ a* Boolean *strongly dependent monomial control system such that the dependency graph $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ of the system*

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

has loop number $t := \mathcal{L}_{G_\sigma}(V_\sigma) > 1$. Furthermore, let $A \in M(n \times n;\ E_2)$ and $B \in M(n \times m;\ E_2)$ be the matrices such that

$$((A|B)id)(x,u) = g(x,u)\ \forall\ (x,u) \in \mathbf{F}_2^n \times \mathbf{F}_2^m$$

(see Lemma 156) and $t' \in D(\mathcal{L}_{G_\sigma}(V_\sigma))\backslash\{1, \mathcal{L}_{G_\sigma}(V_\sigma)\}$ a divisor of $\mathcal{L}_{G_\sigma}(V_\sigma)$ with $1 < t' < \mathcal{L}_{G_\sigma}(V_\sigma)$. In addition, let $V_\sigma = \{a_1, ..., a_n\}$ be the numeration of the vertices and $\widetilde{a_1}, ..., \widetilde{a_n} \subseteq V_\sigma$ their corresponding loop equivalence classes. Assume that B contains a column B_j different from zero with the property that there is a loop equivalence class $\widetilde{a_k} \subseteq \mathcal{L}_{G_\sigma}(V_\sigma)$ such that $\forall\ s \in E(B_j)$ $\exists\ \alpha_s \in \mathbb{N}$ with

$$1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s}) = \alpha_s t'$$

and, additionally, that among all elements of $D(\mathcal{L}_{G_\sigma}(V_\sigma))$ the biggest one that divides the numbers $1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s})$, $s \in E(B_j)$ is t'. Then g is controllable to loop number t'.

Proof. Consider the column B_j and the set $E(B_j)$. We define the following monomial feedback controller $\widehat{f} : \mathbf{F}_2^n \to \mathbf{F}_2^m$

$$\widehat{f}_i(x) := \begin{cases} x_k \text{ if } i = j \\ 1 \text{ if } i \neq j \end{cases} \quad \forall\ x \in \mathbf{F}_2^n,\ i = 1, ..., m$$

Due to the representation $g = (A|B)id$ we can conclude $\forall\ s \in E(B_j)$ that the function $g_s : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2$ depends on the control variable u_j. As a consequence and due to the structure of $\widehat{f}$, the closed loop system

$$\begin{array}{rcl} h_{\widehat{f}} & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \widehat{f}(x)) \end{array}$$

has the following properties $\forall\ x \in \mathbf{F}_2^n$

$$(h_{\widehat{f}})_i(x) = \begin{cases} \sigma(x)x_k \text{ if } i \in E(B_j) \\ \sigma(x) \text{ if } i \notin E(B_j) \end{cases},\ i = 1, ..., n$$

Therefore, the dependency graph $G_{h_{\widehat{f}}} = (V_{h_{\widehat{f}}},\ E_{h_{\widehat{f}}},\ \pi_{h_{\widehat{f}}})$ is identical to $G_\sigma = (V_\sigma,\ E_\sigma,\ \pi_\sigma)$ except for the additional edges $a_s \to a_k$, $s \in E(B_j)$. It is easy to see that the only closed paths in the graph $G_{h_{\widehat{f}}}$ are the closed paths in G_σ and the paths $a_s \rightsquigarrow_{v_s} a_s$, $s \in E(B_j)$ that actually make use of the edges $a_s \to a_k$, $s \in E(B_j)$. A path that makes use of one of the edges $a_s \to a_k$, $s \in E(B_j)$ can only contain one such edge, otherwise the vertex a_k would be visited more than once. As a consequence, the length of such a path satisfies

$$v_s = d(\widetilde{a_s} \rightsquigarrow \widetilde{a_{s'}}) + \beta_{ss'}t + 1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s}) + \gamma_s t$$

where $s' \in E(B_j)$ and $\beta_{ss'}, \gamma_s \in \mathbb{N}_0$. Now, by Lemma 172 $\exists\ \alpha_{ss'} \in \mathbb{N}_0$ such that

$$v_s = d(\widetilde{a_k} \rightsquigarrow \widetilde{a_{s'}}) + \alpha_{ss'}t + \beta_{ss'}t + 1 + \gamma_s t$$

Summarizing, the length of each of those paths satisfies

$$v_s = 1 + \lambda_s t + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_{s'}})$$

whereas the lengths of the paths in G_σ are multiples of t. Now, since t' divides t, $\exists\ c \in \mathbb{N} : t = ct'$. Thus, we have

$$v_s = \lambda_s c t' + 1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_{s'}})$$

By hypothesis, t' divides all $1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s})$, $s \in E(B_j)$. Therefore, by Theorem 116 $\exists\ \alpha \in \mathbb{N}$ such that

$$\mathcal{L}_{G_{h_{\widehat{f}}}}(V_{h_{\widehat{f}}}) = \alpha t'$$

Were $\alpha > 1$, then by Theorem 163, $\mathcal{L}_{G_{h_{\widehat{f}}}}(V_{h_{\widehat{f}}})$ would be a divisor $\widetilde{t} \in D(\mathcal{L}_{G_\sigma}(V_\sigma))$ with $\widetilde{t} > t'$. Moreover, by Corollary 106 $\mathcal{L}_{G_{h_{\widehat{f}}}}(V_{h_{\widehat{f}}})$ would divide $v_s = \lambda_s \widetilde{ct} + 1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_{s'}})$ and therefore $1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_{s'}})$ as well, a contradiction. ■

Example 177 (and Theorem) *Let $\mathbf{F}_2$ be the finite field with two elements, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2^n$ a* Boolean *strongly dependent monomial control system such that the dependency graph $G_\sigma = (V_\sigma,\ E_\sigma,\ \pi_\sigma)$ of the system*

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

has loop number $t := \mathcal{L}_{G_\sigma}(V_\sigma) > 1$. In addition, let $V_\sigma = \{a_1, ..., a_n\}$ be the numeration of the vertices and $\widetilde{a_1}, ..., \widetilde{a_n} \subseteq V_\sigma$ their corresponding loop equivalence classes. Furthermore, let $A \in M(n \times n;\ E_2)$ and $B \in M(n \times m;\ E_2)$ be the matrices such that

$$((A|B)id)(x, u) = g(x, u)\ \forall\ (x, u) \in \mathbf{F}_2^n \times \mathbf{F}_2^m$$

and assume $\exists\ s \in \{1, ..., n\}$, $r \in \{1, ..., m\}$ such that

$$B_{ij} = \begin{array}{c} 1 \text{ if } i = s \text{ and } j = r \\ 0 \text{ if } i \neq s \text{ or } j \neq r \end{array}$$

Then g is completely loop number controllable. To see this, let $t' \in D(\mathcal{L}_{G_\sigma}(V_\sigma))\backslash\{\mathcal{L}_{G_\sigma}(V_\sigma)\}$ be a divisor of $\mathcal{L}_{G_\sigma}(V_\sigma)$ with $1 < t' < \mathcal{L}_{G_\sigma}(V_\sigma)$. Due to the cyclic loop equivalence classes structure of the graph G_σ (see Remark 112) it is always possible to find a loop equivalence class $\widetilde{a_k}$ such that

$1 + d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s}) = t'$. *As a consequence, the* j*th column* B_j *of* B *satisfies the requirements of the previous theorem. Thus,* g *is controllable to loop number* t'*. If we pick a vertex* $a_l \in \widetilde{a_k}$ *such that* $d(a_l \rightsquigarrow a_k) = d(\widetilde{a_k} \rightsquigarrow \widetilde{a_s})$*, a suitable monomial feedback controller* $f : \mathbf{F}_2^n \to \mathbf{F}_2^m$ *can be*

$$f_i(x) := \begin{cases} x_l \text{ if } i = r \\ 1 \text{ if } i \neq r \end{cases} \quad \forall\ x \in \mathbf{F}_2^n,\ i = 1, ..., m$$

To force the loop number $\mathcal{L}_{G_{h_f}}(V_{h_f})$ *to be equal to one, we could use the monomial feedback controller* $\overline{f} : \mathbf{F}_2^n \to \mathbf{F}_2^m$ *defined as*

$$\overline{f}_i(x) := \begin{cases} x_s \text{ if } i = r \\ 1 \text{ if } i \neq r \end{cases} \quad \forall\ x \in \mathbf{F}_2^n,\ i = 1, ..., m$$

Remark 178 *From the previous example we can see that a single control variable* (u_r) *can be used to completely control the system* g*. It also becomes apparent from the previous theorems that the use of too many control variables, (i.e. to many entries of the matrix* B *are equal to* 1*) actually reduces the controllability of the system. This represents a surprising counterintuitive result. As a consequence, we will develop a loop number assignment algorithm for control systems with one single control variable appearing in only one equation, i.e. the matrix satisfies* $B \in M(n \times 1;\ E_2)$ *and it contains exactly one nonzero entry.*

3.2.3 Control synthesis algorithm for Boolean systems with one single control variable

Let $\mathbf{F}_2$ be the finite field with two elements, $n \in \mathbb{N}$ a natural number and $g : \mathbf{F}_2^n \times \mathbf{F}_2 \to \mathbf{F}_2^n$ a *Boolean* monomial control system. Furthermore, let $A \in M(n \times n;\ E_2)$ and $B \in M(n \times 1;\ E_2)$ be the matrices such that

$$((A|B)id)(x, u) = g(x, u)\ \forall\ (x, u) \in \mathbf{F}_2^n \times \mathbf{F}_2$$

and assume that the matrix B contains exactly one nonzero entry, say B_{k1}. We will assume that the system is given and stored using the matrices A and B. The steps of the algorithm are as follows:

1. Calculate the matrices
$$A^{\cdot 2}, ..., A^{\cdot n}$$

2. Establish whether the dependency graph $G_\sigma = (V_\sigma,\ E_\sigma,\ \pi_\sigma)$ of the system
$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$
is strongly connected. This can be accomplished by calculating the reachability matrix (see, for instance, [44])
$$R := (I + A + A^{\cdot 2} + ... + A^{\cdot n-1})$$
and the value
$$(R^2)_{11} = \sum_{k=1}^{n} R_{1k} R_{k1}$$
The graph is strongly connected if and only if $(R^2)_{11} = n$ (see corollary 5.7a and theorem 5.9 in chapter 5 of [44]). If the graph is strongly connected, proceed to step 3, otherwise this algorithm is not applicable.

3. Calculate the loop number $\mathcal{L}_{G_\sigma}(V_\sigma)$ of the graph $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$. According to the algorithm described in [23], $\mathcal{L}_{G_\sigma}(V_\sigma)$ is the greatest common divisor of the numbers i with $1 \leq i \leq n$, such that $A^{\cdot i}$ has at least one non- zero diagonal entry. If $\mathcal{L}_{G_\sigma}(V_\sigma) > 1$ proceed to step 4, otherwise the system g is not controllable.

4. Calculate the set $D(\mathcal{L}_{G_\sigma}(V_\sigma))$. For every element $t' \in D(\mathcal{L}_{G_\sigma}(V_\sigma))$, by Theorem 177, the system g is controllable to loop number t'.

5. Once a desirable $t' > 1$ has been picked (we will treat the case $t' = 1$ separately), in the matrix $A^{\cdot(t'-1)}$ (which was calculated in step 1) we look at the kth column. Any nonzero entry of the kth column provides a candidate variable for the monomial feedback controller $f : \mathbf{F}_2^n \to \mathbf{F}_2$. If, for instance, $(A^{t'-1})_{ik} \neq 0$, we can set $u = f(x) := x_i$.

6. The closed loop system

$$\begin{array}{rcl} h_f & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, f(x)) \end{array}$$

has the desired phase space structure dictated by the loop number t'.

Step 1 requires $(n-1)(2n^3 - n^2)$ addition or multiplication operations[1]. In Step 2, the calculation of R takes $(n-1)n^2$ addition operations. Moreover, the calculation of $(R^2)_{11}$ requires n multiplications and $n-1$ additions. In step 3, to determine the numbers involved, at most n^2 comparisons are needed. The complexity of the greatest common divisor calculation is polynomial in the $\beta-$length or size of the computer representation of the numbers involved (see section 4.1.5 of [53] for the details). These sizes are bounded above by the fact that the numbers involved are smaller than $n+1$. In step 4, the divisors of $\mathcal{L}_{G_\sigma}(V_\sigma)$ are determined. This step requires integer number factorization algorithms that are not of polynomial complexity. However, if $\mathcal{L}_{G_\sigma}(V_\sigma)$ is small, the heuristics used by most computer algebra systems can keep the calculation time in a reasonable frame. It would go beyond the scope of this dissertation to discuss here integer number factorization methods. We refer to chapter 5 of [53]. In step 5, at most n comparisons are needed.

Summarizing, if we put aside the factorization step required in Step 4, the complexity of the algorithm is dominated by the multiple matrix multiplications of Step 1, which is $O(\ n^4)$.

3.2.4 Stabilization of Boolean monomial control systems

In this subsection we provide a characterization of Boolean control systems (not necessarily strongly dependent) that are stabilizable.

Definition 179 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ a monomial control system. Furthermore, let $f : \mathbf{F}_q^n \to \mathbf{F}_q^m$ be an arbitrary monomial feedback controller. The system g is called* stabilizable *if the closed loop system*

$$\begin{array}{rcl} h_f & : & \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x & \mapsto & g(x, f(x)) \end{array}$$

can be forced to be a fixed point system by means of choosing a suitable monomial feedback controller f.

Definition 180 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers and $g : \mathbf{F}_q^n \times \mathbf{F}_q^m \to \mathbf{F}_q^n$ a monomial control system. Furthermore, let $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ be the dependency graph of the system*

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_q^n \to \mathbf{F}_q^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

[1] See also the analysis of the arithmetic operations in the semiring E_q in Section 2.4.

and $V_\sigma = \{a_1, ..., a_n\}$ *be the numeration of its vertices. Now we label every vertex* $a_i \in V_\sigma$ *(corresponding to the variable* x_i *and its update function* σ_i*) as* critical vertex*, if* g_i *depends on some control variable* u_j*. The resulting graph* $\widehat{G}_\sigma$ *is called the* labeled state dependency graph of g*.*

Theorem 181 *Let* $\mathbf{F}_2$ *be the finite field with two elements,* $n, m \in \mathbb{N}$ *natural numbers and* $g : \mathbf{F}_2^n \times \mathbf{F}_2^m \to \mathbf{F}_2^n$ *a* Boolean *monomial control system such that*

$$\begin{array}{rcl} \sigma & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, \vec{1}) \end{array}$$

is not *a fixed point system. Then* g *is stabilizable if and only if for every strongly connected component* C *of the labeled state dependency graph* $\widehat{G}_\sigma$ *of* g *with loop number bigger than one either of the following conditions holds:*

- C *has a critical vertex.*
- C *is connected by a sequence to a strongly connected component* D*, which contains a critical vertex.*

Proof. Let $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ be the dependency graph of σ and $V_\sigma = \{a_1, ..., a_n\}$ be the numeration of its vertices. Since by hypothesis, σ is not a fixed point system, by Theorem 73 the dependency graph $G_\sigma = (V_\sigma, E_\sigma, \pi_\sigma)$ of σ as well as the labeled state dependency graph of g must contain nontrivial strongly connected components.
In order to prove that g is stabilizable, we have to prove that we can find a feedback controller $f : \mathbf{F}_2^n \to \mathbf{F}_2^m$, such that the closed loop system

$$\begin{array}{rcl} h_f & : & \mathbf{F}_2^n \to \mathbf{F}_2^n \\ x & \mapsto & g(x, f(x)) \end{array}$$

is a fixed point system. The construction of a suitable f is straightforward if we look at the labeled state dependency graph of g : If C is a strongly connected component with loop number bigger than 1 of $\widehat{G}_\sigma$, according to the hypothesis, it either contains a critical vertex or it is connected by a path to a strongly connected component, which contains a critical vertex. In the former case, let a_i be the critical vertex and x_i its corresponding variable. The definition of critical vertex tells us that the function g_i actually depends on $\vec{u}$, i.e. there is a $j \in \{1, ..., m\}$ such that g_i depends on u_j. Now we can add the edge $a_i \to a_i$ to the labeled state dependency graph of g. This modification corresponds to setting the function $f_j(\vec{x}) := x_i$.
Now we will consider the case where C is connected by a path to a strongly connected component, say D, which contains a critical vertex. Let a_k be the critical vertex contained in D and x_k its corresponding variable. Following the same argument as above, we know there is an $l \in \{1, ..., m\}$ such that g_k depends on u_l. Now we can add two edges to the labeled state dependency graph of g, namely, the edge $a_k \to a_k$ and an edge that starts at a_k and points to a vertex contained in the component C, say the vertex a_s (which corresponds to the variable x_s). The edge $a_k \to a_k$ is not strictly necessary if the component D already has loop number equal to 1. Again, this modifications correspond to setting the function $f_l(\vec{x}) := x_k x_s$.
We continue this procedure with every strongly connected component of the labeled state dependency graph of g, that has loop number bigger than one. At the end of this process, for some non empty subset $J \subseteq \{1, ..., m\}$ the functions f_t, $t \in J$ will be defined. For the remaining indices in the set $I := \{1, ..., m\} \backslash J$ we simply set $f_t \equiv 1 \ \forall \ t \in I$.
With the obtained feedback controller f we construct the function h_f. The dependency graph of h_f obviously only differs from the labeled state dependency graph of F by the edges that were

added to critical vertices. It is clear that the edges added pursuit the following purposes: Either to merge two or more strongly connected components into a bigger one or to force the loop number of a strongly connected component to be equal to 1. Therefore, every nontrivial strongly connected component of h_f has loop number one and thus, by Theorem 88, h_f is a fixed point system. This shows that the conditions stated in the theorem are *sufficient* for the control system g to be stabilized.
To show that the conditions are also *necessary*, assume that they do not hold. As a consequence, there is a strongly connected component U of the labeled state dependency graph of g, such that

1. U has loop number bigger than 1.
2. U does not contain a critical vertex.
3. U is not connected by a path to a strongly connected component that contains a critical vertex.

Let $a_{i_1}, ...a_{i_t}$ be all the vertices contained in U and consider the corresponding variables $x_{i_1}, ...x_{i_t}$ and their update functions $\sigma_{i_1}, ...\sigma_{i_t}$. Since U does not contain any critical vertices, the functions $g_{i_1}, ..., g_{i_t}$ cannot depend on any of the control variables $u_1, ..., u_m$. Therefore, for *any* feedback controller $f : \mathbf{F}_2^n \to \mathbf{F}_2^m$ the function h_f has the property

$$(h_f)_{i_q} = g_{i_q} \ \forall \ q \in \{1, ..., t\}$$

Now, assume that by the choice of f the arrows added to the labeled state dependency graph of g create a strong connection between U and additional vertices outside of U. The sequences that make that strong connection possible must contain at least one of the new arrows. As a consequence, a critical vertex c is part of the sequences. Thus, there must be a path from a vertex in $u \in U$ to c. Since, by hypothesis 3, no such path exists in the labeled state dependency graph of g, there must be a new arrow involved in the path from u to c. Consequently, there is a critical vertex c' contained in the path such that there is a shorter path from u to c'. If we repeat this argument, the path becomes shorter and shorter and eventually we are forced to claim that u is a critical vertex. This contradicts 2. Summarizing, the dependency graph of the system h_f contains the strongly connected component U which has loop number bigger than 1. By Theorem 88, h_f cannot be a fixed point system. ■

Remark 182 *The proof of this theorem suggests an algorithm to design a feedback controller that would stabilize a given Boolean monomial control systems, provided the conditions of the theorem are satisfied. However, we won't elaborate on the algorithmic aspects of this verification and the controller design.*

Part II

Reverse engineering time discrete dynamical systems over a finite field

Introductory remarks

Since the development of multiple and simultaneous measurement techniques such as microarray technologies, reverse engineering of biochemical and, in particular, gene regulatory networks has become a more important problem in systems biology . One well-known reverse engineering approach are the top-down methods, which try to infer network properties based on the observed global input-output-response. The observed input-output-response is usually only partially described by available experimental data.

Depending on the type of mathematical model used to describe a biochemical process, a variety of top-down reverse engineering algorithms have been proposed [28], [34], [42]. Each modeling paradigm presents different requirements relative to quality and amount of the experimental data needed. Moreover, for each type of model, a suitable mathematical framework has to be developed in order to study the performance and limitations of reverse engineering methods. For any given modeling paradigm and reverse engineering method it is important to answer the following questions:

1. What are the minimal requirements on data sets?

2. Can data sets be characterized in such a way that "optimal" data sets can be identified? (Optimality meaning that the algorithm performs better using such a data set compared to its performance using other data sets.)

The second question is related to the *design of experiments* and optimality is characterized in terms of *quantity and quality* of the data sets.

[65] developed a top-down reverse engineering algorithm for the modeling paradigm of time discrete finite dynamical systems. Herein, we will refer to it as the LS-algorithm. They apply their method to biochemical networks by modeling the network as a time discrete finite dynamical system, obtained by discretizing the concentration levels of the interacting chemicals to elements of a finite field. One of the key steps of the LS-algorithm includes the choice of a term order. The modeling paradigm of time discrete finite dynamical systems generalizes the Boolean approach [55] (where the field only contains the elements 0 and 1). Moreover, it is a special case of the paradigm described in [110].

Some aspects of the performance of the LS-algorithm were studied by [52] in a probabilistic framework.

In this part of the work we investigate the two questions stated above in the particular case of the LS-algorithm. For this purpose, we developed a mathematical framework[2] that allows us to study the LS-algorithm in depth. Having expressed the steps of the LS-algorithm in our framework, we were able to provide concrete answers to both questions: First, we found minimal requirements on a data set based on how many terms the functions to be reverse engineered display. Second, we identified optimal data sets, which we characterize using a geometric property called "general position". Moreover, we developed a constructive method to generate optimal data sets, provided a codimensional condition is fulfilled.

[2] This framework is described in the next chapter. Since it contains a pure algebraic subject, the reader more interested in mathematical modelling may skip it without any loss.

In addition, we present a generalization of the LS-algorithm that does not depend on the choice of a term order. We call this generalization the *term-order-free reverse engineering method.* For this method we derive a formula for the probability of finding the correct model[3], provided the data set used satisfies an optimality criterion. Furthermore, we analyze the asymptotic behavior of the probability formula for a growing number of variables n (i.e. interacting chemicals). Unfortunately, this formula converges to zero as fast as r^{q^n}, where $q \in \mathbb{N}$ and $0 < r < 1$. Consequently, we conclude that even if an optimal data set is used and the restrictions imposed by the use of term orders are overcome, the reverse engineering problem remains unfeasible, unless experimentally impracticable amounts of data are available. This result discouraged us from including in this thesis any computational and algorithmic aspects of the term-order-free reverse engineering method.

In contrast to [52], we focus here on providing possible criteria for the design of specific experiments instead of assuming that the data sets are generated randomly. Moreover, we do not necessarily assume that information about the actual number of interactions in the biochemical network is available.

[3] We will give a precise definition of "correct model".

Chapter 4

Excursus: Canonical representatives for residue classes of a polynomial ideal and orthogonality

4.1 Some introductory statements

A well known result of B. Buchberger is the existence of the normal form of a polynomial with respect to a polynomial ideal I in the ring of multivariate polynomials over a field K. This result follows from the existence of so called Gröbner bases for polynomial ideals. For a given fixed term ordering, this normal form is unique [66], [15], [14]. In this chapter we present a new way to calculate this normal form, provided the field K is finite and the ideal I is a vanishing ideal, i.e. I is equal to the set of polynomials which vanish in a given set of points X. Our method doesn't pursue establishing a new, especially efficient, algorithm for the computation of such a normal form. Rather, the aim of this chapter is to unveil an interesting way to look at this issue based on the concept of orthogonality.

For orthogonality to apply, we introduce a symmetric bilinear form on a vector space (see, for instance, [98]). A symmetric bilinear form can be seen as a generalized inner product. Some authors have explored vector spaces endowed with generalized forms of inner products. For example, we refer to the following papers: [69],[6],[29], [30], [74], [54], [113].

Having defined a symmetric bilinear form, we are able to introduce the notion of orthogonality and orthonormality. Then we consider the orthogonal solution of a solvable inhomogeneous under-determined linear operator equation. If one thinks of an inhomogeneous under-determined system of linear equations in an Euclidean space, the orthogonal solution is simply the solution that is perpendicular to the affine subspace associated with the system. After going through existence and uniqueness considerations, we come to the main statement of this chapter, namely, that the above mentioned normal form can be obtained as the orthogonal solution of a system of linear equations. That system of equations arises as a linear formulation of the multivariate polynomial interpolation problem.

Based on our literature research, we believe that the study of polynomial algebras in the framework of symmetric bilinear spaces (vector spaces endowed with a symmetric bilinear form) represents a novel approach. Suitable extensions of our method to more general fields (i.e. infinite fields) could open new possibilities for studying problems in the areas of polynomial algebra, computational algebra and algebraic geometry using functional analytic or linear algebraic techniques.

The concept of orthogonal solution is not limited by monomial orders, as it is the case for Gröbner bases calculations. In this sense, our method reveals a wider class of normal forms (with respect to vanishing ideals) in which the normal forms à la Buchberger appear as special cases.

Another application that we will describe in detail in Chapter 5 is the problem of choosing a particular interpolant among all possible solutions of a highly under-determined multivariate

interpolation problem. This is related to the study of the performance of the "reverse engineering" algorithm presented in [65].

The organization of this chapter is the following:

Section 4.2 is devoted to the general definition of *symmetric bilinear spaces* and *orthogonal solutions* of an inhomogeneous linear operator equation. Subsection 4.2.1 covers basic definitions and properties of symmetric bilinear spaces, in particular, the concepts of *orthogonality* and *orthonormality* are introduced. Subsection 4.2.2 introduces the notion of orthogonal solution of a solvable under-determined linear operator equation. Existence and uniqueness of orthogonal solutions are proved and some issues regarding the existence of orthonormal bases are discussed.

Section 4.3 introduces a linear operator called *evaluation epimorphism* and formulates the multivariate polynomial interpolation problem in a linear algebraic fashion.

Section 4.4 covers the more technical aspect of constructing special symmetric bilinear forms. Using that type of symmetric bilinear form will allow us to prove the main result of this chapter in section 4.5.

Section 4.5 is devoted to the statement and proof of our main result. Namely, that the canonical normal form of an arbitrary polynomial f with respect to a vanishing ideal $I(X)$ in the ring of multivariate polynomials over a finite field K can be calculated as the orthogonal solution of a linear operator equation involving the evaluation epimorphism.

For standard terminology, notation and well known results in computational algebraic geometry and commutative algebra we refer to [25] and [8].

4.2 Symmetric bilinear vector spaces and orthogonal solutions of inhomogeneous systems of linear equations

4.2.1 Basic definitions

In this subsection we will introduce the concept of a symmetric bilinear form in a vector space. With this concept it will be possible to define symmetric bilinear vector spaces and orthonormality. Furthermore, some basic properties are briefly reviewed (cf. [98])

Definition 183 *Let V be a vector space over a field K. A* symmetric bilinear form *on V is a mapping*

$$\langle \cdot, \cdot \rangle : V \times V \to K$$

that fulfills the following properties

1. *Bilinearity: For all $w \in V$ the mappings*

$$\begin{array}{rcl} \langle \cdot, w \rangle & : & V \to K \\ v & \mapsto & \langle v, w \rangle \end{array}$$

$$\begin{array}{rcl} \langle w, \cdot \rangle & : & V \to K \\ v & \mapsto & \langle w, v \rangle \end{array}$$

   are linear.

2. *Symmetry: For all $v, w \in V$ holds*

$$\langle v, w \rangle = \langle w, v \rangle$$

Remark 184 (and Definition) *Let V be a finite dimensional vector space over a field K endowed with a bilinear form*

$$s : V \times V \to K$$

Further let $d := \dim(V)$. Due to the bilinearity, s is uniquely defined by its values on all possible pairs (u_i, u_j) of a given basis $(u_1, ..., u_d)$ of V. Indeed, after introducing coordinates with respect to the basis $(u_1, ..., u_d)$, the value $s(v, w)$ of s on two arbitrary vectors

$$v = \sum_{i=1}^{d} x_i u_i$$

and

$$w = \sum_{i=1}^{d} y_i u_i$$

with coordinate vectors $\vec{x}, \vec{y} \in K^d$, can be simply calculated as

$$s(v, w) = \vec{x}^t S \vec{y}$$

where S is the $d \times d$ matrix

$$S_{ij} := s(u_i, u_j), \ i, j \in \{1, ..., d\}$$

with entries in K. If the bilinear form s is symmetric, so the matrix S. The matrix S is called the representing matrix of s with respect to the basis $(u_1, ..., u_d)$. *After fixing a basis $(u_1, ..., u_d)$ of V, it is easy to show, that there is a one-to-one correspondence between the set of all symmetric bilinear forms on V and the set of all $d \times d$ symmetric matrices with entries in K seen as representing matrices with respect to the basis $(u_1, ..., u_d)$.*

Definition 185 *A vector space V over a field K endowed with a symmetric bilinear form*

$$\langle \cdot, \cdot \rangle : V \times V \to K$$

is called a symmetric bilinear space.

Example 186 *Every (real) Euclidean space is due to the positive definiteness of its inner product a symmetric bilinear space.*

Definition 187 *Let V be a symmetric bilinear space. Two vectors $v, w \in V$ are called* orthogonal *if*

$$\langle w, v \rangle = 0$$

In this situation we write $w \perp v$.

Theorem 188 (and Definition) *Let V be a symmetric bilinear space and $W \subseteq V$ a subspace. The set*

$$W^{\perp} := \{v \in V \mid v \perp w \ \forall \ w \in W\}$$

is a subspace of V and is called the orthogonal complement *of W.*

Proof. The easy proof is left to the reader. ■

Remark 189 (and Example) *Contrary to the case of Euclidean or unitary vector spaces, $W^{\perp} \cap W$ is not always equal to the zero vector space $\{0\}$. For instance, consider any finite field K of characteristic 2 and the finite dimensional vector space K^2. Let U be the one dimensional subspace*

$$U := span(\begin{pmatrix} 1 \\ 1 \end{pmatrix})$$

Now let the matrix

$$A := \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

define the following generalized inner product

$$\langle \vec{x}, \vec{y} \rangle := \vec{x}^t A \vec{y}$$

Then for any vector $\vec{u} \in U$ we have

$$\begin{aligned} \left\langle \vec{u}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle &= \left\langle \begin{pmatrix} \lambda \\ \lambda \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\rangle = (\lambda, \lambda) \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \\ &= (\lambda, \lambda) \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0 \end{aligned}$$

This means

$$U^{\perp} = U$$

The well known result

$$V = W^{\perp} \oplus W$$

for an Euclidean or unitary vector space V ($\oplus$ stands for the direct sum) depends on the existence of orthonormal bases for V. As the above example shows, such bases don't always exist in the case of symmetric bilinear spaces.

Definition 190 *Let $d \in \mathbb{N}$ be a natural number and V a d-dimensional symmetric bilinear space over a field K. A basis $(w_1, ...w_d)$ of V is called* orthonormal *if it holds for $i, j \in \{1, ..., d\}$*

$$\langle w_i, w_j \rangle = \delta_{ij} := \begin{cases} 1 \text{ if } i = j \\ 0 \text{ otherwise} \end{cases}$$

Remark 191 *This definition can be extended to symmetric bilinear spaces with countable dimension, but such spaces are not the object of study in this treatise.*

Example 192 *Let $d \in \mathbb{N}$ be a natural number and V a d-dimensional vector space over a field K. Furthermore let $(u_1, ...u_d)$ be a basis of V. Then one can construct a symmetric bilinear form on V by setting*

$$\langle u_i, u_j \rangle := \delta_{ij} \ \forall \ i, j \in \{1, ..., d\}$$

(see also Remark 184.) Here the basis $(u_1, ...u_d)$ is obviously orthonormal.

Lemma 193 (and Definition) *Let $d \in \mathbb{N}$ be a natural number and V a d-dimensional symmetric bilinear space over a field K. Furthermore let $(w_1, ...w_d)$ be an orthonormal basis of V. Then for every vector $v \in V$ holds*

$$v = \sum_{k=1}^{d} \langle v, w_k \rangle w_k$$

The field elements $\langle v, w_i \rangle \in K$, $i = 1, ..., d$ are the so called Fourier coefficients.

Proof. Since $(w_1, ...w_d)$ is a basis for V, every vector $v \in V$ can be written uniquely in the form

$$v = \sum_{k=1}^{d} \lambda_k w_k$$

with unique coefficients $\lambda_i \in K$, $i = 1, ..., d$. Now for every $j \in \{1, ..., d\}$ we have

$$\langle v, w_j \rangle = \left\langle \sum_{k=1}^{d} \lambda_k w_k, w_j \right\rangle = \sum_{k=1}^{d} \lambda_k \langle w_k, w_j \rangle = \sum_{k=1}^{d} \lambda_k \delta_{kj} = \lambda_j$$

■

4.2.2 Orthogonal solutions of inhomogeneous linear operator equations

Definition 194 *Let $d \in \mathbb{N}$ be a natural number and V a d-dimensional symmetric bilinear space over a field K. Furthermore, let W be an arbitrary vector space over the field K, $T : V \to W$ a not injective linear operator and $w \in W$ a vector with the property*

$$w \in T(V)$$

Now let $m :=$nullity$(T) \in \mathbb{N}$ be the dimension of the kernel of T. A solution $v^* \in V$ of the equation

$$Tv = w$$

is called orthogonal solution*, if for an arbitrary basis $(u_1, ..., u_m)$ of* $\ker(T)$ *the following orthogonality conditions hold*

$$\langle u_i, v^* \rangle = 0 \ \forall \ i \in \{1, ..., m\}$$

Remark 195 *Let $(u_1, ..., u_m)$ be a basis of* $\ker(T)$*. Then each arbitrary vector $u \in \ker(T)$ can be written in the form*

$$u = \sum_{i=1}^{m} \lambda_i u_i$$

with suitable field elements $\lambda_i \in K$. If the orthogonality conditions

$$\langle u_i, v^* \rangle = 0 \ \forall \ i \in \{1, ..., m\}$$

hold for the basis $(u_1, ..., u_m)$, then we have

$$\langle u, v^* \rangle = \left\langle \sum_{i=1}^{m} \lambda_i u_i, v^* \right\rangle = \sum_{i=1}^{m} \lambda_i \langle u_i, v^* \rangle = 0$$

and that means

$$v^* \in \ker(T)^{\perp}$$

In particular, for any other different basis $(w_1, ..., w_m)$ of $\ker(T)$ *it holds*

$$\langle w_j, v^* \rangle = 0 \ \forall \ j \in \{1, ..., m\}$$

Theorem 196 *Let $d \in \mathbb{N}$ be a natural number and V a d-dimensional symmetric bilinear space over a field K. Furthermore, let W be an arbitrary vector space over the field K, $T : V \to W$ a not injective linear operator and $w \in W$ a vector with the property*

$$w \in T(V)$$

If $\ker(T)$ *has an* orthonormal basis*, then the equation*

$$Tv = w$$

has always a unique orthogonal solution $v^* \in V$.

Proof. Let $m :=$nullity$(T) = \dim(\ker(T)) \in \mathbb{N}$ be the dimension of the null space of T and $(u_1, ..., u_m)$ an orthonormal basis of $\ker(T)$. Since $w \in T(V)$, there must exist a solution $\widehat{\xi} \in V$ of $Tv = w$. For any other solution $\xi \in V$ we have

$$T(\xi - \widehat{\xi}) = T(\xi) - T(\widehat{\xi}) = 0$$

and therefore

$$\xi - \widehat{\xi} \in \ker(T)$$

That means that all solutions $\xi \in V$ of $Tv = w$ can be written in the form

$$\xi = \widehat{\xi} + \sum_{i=1}^{m} \lambda_i u_i$$

with the $\lambda_i \in K$, $i = 1, ..., m$ running over all K. In particular, we can construct a very specific solution by choosing the parameters $\lambda_i \in K$, $i = 1, ..., m$ in the following manner

$$\lambda_i := -\left\langle u_i, \widehat{\xi} \right\rangle, \ i = 1, ..., m$$

For this solution

$$v^* := \widehat{\xi} + \sum_{i=1}^{m} -\left\langle u_i, \widehat{\xi} \right\rangle u_i$$

and for every $j \in \{1, ..., m\}$ it holds

$$\begin{aligned} \langle u_j, v^* \rangle &= \left\langle u_j, \widehat{\xi} + \sum_{i=1}^{m} -\left\langle u_i, \widehat{\xi} \right\rangle u_i \right\rangle = \left\langle u_j, \widehat{\xi} \right\rangle + \sum_{i=1}^{m} -\left\langle u_i, \widehat{\xi} \right\rangle \langle u_j, u_i \rangle \\ &= \left\langle u_j, \widehat{\xi} \right\rangle + \sum_{i=1}^{m} -\left\langle u_i, \widehat{\xi} \right\rangle \delta_{ji} = \left\langle u_j, \widehat{\xi} \right\rangle - \left\langle u_j, \widehat{\xi} \right\rangle = 0 \end{aligned}$$

This shows the existence of an orthogonal solution of $Tv = w$. Now let $\widetilde{v} \in V$ be another orthogonal solution of $Tv = w$. Again, since

$$T(v^* - \widetilde{v}) = T(v^*) - T(\widetilde{v}) = 0$$

we can write

$$v^* = \widetilde{v} + \sum_{i=1}^{m} \alpha_i u_i$$

with suitable $\alpha_i \in K$. From the orthogonality conditions for v^* and $\widetilde{v}$ we have $\forall\ j \in \{1, ..., m\}$

$$\begin{aligned} 0 &= \langle u_i, v^* \rangle = \left\langle u_i, \widetilde{v} + \sum_{i=1}^{m} \alpha_i u_i \right\rangle = \langle u_j, \widetilde{v} \rangle + \left\langle u_j, \sum_{i=1}^{m} \alpha_i u_i \right\rangle \\ &= \sum_{i=1}^{m} \alpha_i \langle u_j, u_i \rangle = \sum_{i=1}^{m} \alpha_i \delta_{ji} = \alpha_j \end{aligned}$$

and that means $v^* = \widetilde{v}$. ■

Remark 197 *The existence of an orthonormal basis of* $\ker(T)$ *is crucial for the proof of this theorem. It is important to notice that in a symmetric bilinear space over a general field* K*, the Gram-Schmidt orthonormalization only works if the norm*

$$\|v\| := \sqrt{\langle v, v \rangle}$$

of the vectors used in the Gram-Schmidt process exists in the field K *and is not equal to the zero element. In general terms, the existence of square roots would be assured in a field* K *which satisfies*

$$\forall\ x \in K\ \exists\ y \in K \text{ such that } y^2 = x \tag{4.1}$$

Now, if K *is finite, then (4.1) holds if and only if* $Char(K) = 2$*. To see this, consider*

$$\begin{aligned} \phi &: K \to K \\ x &\mapsto \phi(x) := x^2 \end{aligned}$$

Since K *is finite,* ϕ *is surjective if and only if* ϕ *is injective. If* ϕ *is injective, then* $1 = -1$ *since* $\phi(1) = \phi(-1)$*; that is* $Char(K) = 2$*. The converse follows from the fact that if* $Char(K) = 2$*,*

then ϕ is the Frobenius homomorphism, which is - as is generally known - an automorphism. After fixing a basis $(u_1, ..., u_d)$ for the vector space V, the question whether $\langle v, v\rangle = 0$ for $v \neq 0$ is equivalent to the nontrivial solvability in K^d of the following quadratic form

$$\vec{x}^t A \vec{x} = 0 \tag{4.2}$$

where A is the representing matrix of $\langle ., .\rangle$ with respect to the basis $(u_1, ..., u_d)$ (see Remark 184). In chapter 3, §2 of [67] explicit formulas for the exact number of solutions in K^n of equations of the type (4.2), where A is a $n \times n$ symmetric matrix with entries in a finite field K, can be found. In accordance with Remark 189, the facts stated above show that in the case of a general finite field K, orthonormal bases might not exist.

Corollary 198 *Let $d \in \mathbb{N}$ be a natural number and V a d-dimensional symmetric bilinear space over a field K. Furthermore, let W be an arbitrary vector space over the field K and $T : V \to W$ a not injective linear operator. If* $\ker(T)$ *has an* orthonormal basis*, then the equation*

$$Tv = 0$$

has always the unique orthogonal solution $0 \in V$.

Proof. The zero vector $0 \in V$ satisfies trivially the equation

$$Tv = 0$$

and for any basis $(u_1, ..., u_m)$ of $\ker(T)$ it follows from the bilinearity of the inner product

$$\langle u_i, 0\rangle = 0 \ \forall \ i \in \{1, ..., m\}$$

Now the claim follows from the uniqueness of the orthogonal solution. ■

4.3 Solving the polynomial interpolation problem in $PF_n(\mathbf{F}_q)$

In this section we define the *evaluation epimorphism* of a tuple $(\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$ of points in the space $\mathbf{F}_q^n$. The evaluation epimorphism allows for a linear algebraic formulation of the multivariate polynomial interpolation problem. For this section, recall Definitions 21 and 20 and Corollary 28.

Theorem 199 (and Definition) *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers with $m \leq q^n$. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

*be a tuple of m **different** n-tuples with entries in the field $\mathbf{F}_q$. Then the mapping*

$$\begin{aligned} \Phi_{\vec{X}} \quad &: \quad F_n(\mathbf{F}_q) \to \mathbf{F}_q^m \\ f \quad &\mapsto \quad \Phi_{\vec{X}}(f) := (f(\vec{x}_1), ..., f(\vec{x}_m))^t \end{aligned}$$

is a surjective linear operator. $\Phi_{\vec{X}}$ is called the evaluation epimorphism of the tuple $\vec{X}$.

Proof. The proof of the linearity is left to the reader. Now let $\vec{b} \in \mathbf{F}_q^m$ be an arbitrary vector. Since $m \leq q^n$ we can construct a function

$$g \in F_n(\mathbf{F}_q)$$

with the property

$$g(\vec{x}_i) = b_i \ \forall \ i \in \{1, ..., m\}$$

and that means exactly

$$\Phi_{\vec{X}}(g) = \vec{b}$$

■

Remark 200 (and Corollary) *Since a basis of $F_n(\mathbf{F}_q)$ is given by the fundamental monomial functions $(g_{nq\alpha})_{\alpha\in M_q^n}$, the matrix*

$$A := (\Phi_{\vec{X}}(g_{nq\alpha}))_{\alpha\in M_q^n} \in M(m \times q^n; \mathbf{F}_q)$$

representing the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$ with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$ has always the full rank $m = \min(m, q^n)$. That also means, that the dimension of the $\ker(\Phi_{\vec{X}})$ is

$$\dim(\ker(\Phi_{\vec{X}})) = \dim(F_n(\mathbf{F}_q)) - m = q^n - m$$

Corollary 201 *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers with $m \leq q^n$. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$ and $\vec{b} \in \mathbf{F}_q^m$ a vector. Then the interpolation problem of finding a polynomial function $f \in PF_n(\mathbf{F}_q)$ with the property

$$f(\vec{x}_i) = b_i \ \forall\ i \in \{1, ..., m\}$$

can be solved by solving the system of linear equations

$$A\vec{y} = \vec{b} \tag{4.3}$$

where

$$A := (\Phi_{\vec{X}}(g_{nq\alpha}))_{\alpha\in M_q^n}$$

is the matrix representing the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$ with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$. The entries of a solution vector of the equations (4.3) are the coefficients of the solution with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$.

Proof. Since $F_n(\mathbf{F}_q) = PF_n(\mathbf{F}_q)$, a solution of the interpolation problem can be found by solving the equation

$$\Phi_{\vec{X}}(g) = \vec{b} \tag{4.4}$$

for g, where $\Phi_{\vec{X}}$ is the surjective linear operator

$$\begin{aligned} \Phi_{\vec{X}} \quad &: \quad F_n(\mathbf{F}_q) \to \mathbf{F}_q^m \\ f \quad &\mapsto \quad \Phi_{\vec{X}}(f) := (f(\vec{x}_1), ..., f(\vec{x}_m))^t \end{aligned}$$

of the above theorem. After fixing the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$, equation (4.4) implies the following system of linear equations for the coefficients of the solutions with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$

$$A\vec{y} = \vec{b}$$

where

$$A := (\Phi_{\vec{X}}(g_{nq\alpha}))_{\alpha\in M_q^n}$$

is the matrix representing the map $\Phi_{\vec{X}}$ with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$. According to Remark 200, the matrix A has full rank and therefore a solution of $A\vec{y} = \vec{b}$ always exists. ■

Remark 202 *In the case $m < q^n$ where m is strictly smaller than q^n we have*

$$\dim(\ker(\Phi_{\vec{X}})) = \dim(F_n(\mathbf{F}_q)) - m = q^n - m > 0$$

and the solution of the interpolation problem is not unique. Only in the case $m = q^n$, that means, when for all elements of $\mathbf{F}_q^n$ the corresponding interpolation values are given, the solution is unique. In the most common case $m \ll q^n$, one meaningful way to choose one particular solution among the affine subspace of all solutions is to look for an orthogonal solution, that is a solution that doesn't contain any linear combinations of vectors lying in $\ker(\Phi_{\vec{X}})$. For this purpose we need to define a useful generalized inner product on the vector space $F_n(\mathbf{F}_q)$. In the next section we will explore this issue.

4.4 Construction of special purpose symmetric bilinear forms

Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers with $m < q^n$. Further let

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in \left(\mathbf{F}_q^n\right)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$ and $d := \dim(F_n(\mathbf{F}_q))$. Now consider the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$. By Remark 200 and due to the fact $m < q^n$, the nullity of $\Phi_{\vec{X}}$ is given by

$$s := \dim(\ker(\Phi_{\vec{X}})) = \dim(F_n(\mathbf{F}_q)) - m = q^n - m > 0$$

Now let $(u_1, ..., u_s)$ be a basis of $\ker(\Phi_{\vec{X}}) \subseteq F_n(\mathbf{F}_q)$. By the basis extension theorem, we can extend the basis $(u_1, ..., u_s)$ to a basis

$$(u_1, ..., u_s, u_{s+1}, ..., u_d)$$

of the whole space $F_n(\mathbf{F}_q)$. As in example 192, we can construct a symmetric bilinear form on $F_n(\mathbf{F}_q)$ by setting

$$\langle u_i, u_j \rangle := \delta_{ij} \ \forall \ i, j \in \{1, ..., d\}$$

Here the basis $(u_1, ...u_d)$ is orthonormal and the vectors $(u_{s+1}, ..., u_d)$ are a basis of the orthogonal complement $\ker(\Phi_{\vec{X}})^\perp$ of $\ker(\Phi_{\vec{X}})$. Indeed, according to Lemma 193, every vector $v \in F_n(\mathbf{F}_q)$ can be written as

$$v = \sum_{k=1}^{d} \langle v, u_k \rangle \, u_k$$

If $v \in \ker(\Phi_{\vec{X}})^\perp$, then in particular

$$v \perp u_i \ \forall \ i \in \{1, ..., s\} \Leftrightarrow \langle v, u_i \rangle = 0 \ \forall \ i \in \{1, ..., s\}$$

and that means

$$v = \sum_{k=s+1}^{d} \langle v, u_k \rangle \, u_k$$

In other words, the set $(u_{s+1}, ..., u_d)$ generates $\ker(\Phi_{\vec{X}})^\perp$. The vectors $(u_{s+1}, ..., u_d)$ are as subset of the basis $(u_1, ..., u_s, u_{s+1}, ..., u_d)$ of course linearly independent. In particular, this shows that for the above constructed generalized inner product we have

$$\ker(\Phi_{\vec{X}}) \cap \ker(\Phi_{\vec{X}})^\perp = \{0\} \tag{4.5}$$

In general, the way we extend the basis $(u_1, ..., u_s)$ of $\ker(\Phi_{\vec{X}})$ to a basis

$$(u_1, ..., u_s, u_{s+1}, ..., u_d)$$

of the whole space $F_n(\mathbf{F}_q)$ determines crucially the symmetric bilinear form we get by setting $\langle u_i, u_j \rangle := \delta_{ij} \ \forall \ i, j \in \{1, ..., d\}$. Consequently, the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$ may vary according to the chosen extension $u_{s+1}, ..., u_d \in F_n(\mathbf{F}_q)$. One systematic way to get a basis of the whole space $F_n(\mathbf{F}_q)$ starting with a basis $(u_1, ..., u_s)$ of $\ker(\Phi_{\vec{X}})$ is the following: let

$$(\vec{y}_1, ..., \vec{y}_s)^t \tag{4.6}$$

be the matrix whose rows are the coordinate vectors $\vec{y}_1, ..., \vec{y}_s \in K^d$ of $(u_1, ..., u_s)$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of $F_n(\mathbf{F}_q)$. Now we perform Gauss-Jordan elimination on the matrix (4.6), obtaining the matrix R. Now consider the set $B := \{\vec{e}_1, ..., \vec{e}_d\}$ of canonical unit vectors of the space $\mathbf{F}_q^d$. For every pivot element r_{ij} used during the Gauss-Jordan elimination performed on (4.6), eliminate the canonical unit vector $\vec{e}_j$ from the set B. This yields the set $\tilde{B}$. The coordinate vectors for a basis for the whole space $F_n(\mathbf{F}_q)$ are now given by the the rows of R and the vectors in the set $\tilde{B}$. We call this way of construction of the orthonormal basis for the space $F_n(\mathbf{F}_q)$ the *standard orthonormalization.* We illustrate the algorithm using an example:

Example 203 *Suppose $q = 3$, $\mathbf{F}_3 = \mathbb{Z}_3$, $m = 4$, $d = 3^2 = 9$, $s = 5$ and that after performing Gauss-Jordan elimination on (4.6) we get the following matrix*

$$R := \begin{pmatrix} 1 & 0 & z_{1,3} & 0 & 0 & z_{1,6} & 0 & z_{1,8} & z_{1,9} \\ 0 & 1 & z_{2,3} & 0 & 0 & z_{2,6} & 0 & z_{2,8} & z_{2,9} \\ 0 & 0 & 0 & 1 & 0 & z_{3,6} & 0 & z_{3,8} & z_{3,9} \\ 0 & 0 & 0 & 0 & 1 & z_{4,6} & 0 & z_{4,8} & z_{4,9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & z_{5,8} & z_{5,9} \end{pmatrix} \tag{4.7}$$

(The $z_{i,j} \in \mathbf{F}_q$ stand for unspecified field elements). Then for the extension of the basis we choose the following canonical basis vectors

$$\vec{e}_3, \vec{e}_6, \vec{e}_8, \vec{e}_9 \in \mathbb{Z}_3^9$$

Now we substitute coordinate vectors $(\vec{y}_1, ..., \vec{y}_5)$ of the basis $(u_1, ..., u_5)$ by the rows in the reduced matrix 4.7 (this step is not strictly necessary, but it will be needed to prove the theorems below) and get the following coordinate vectors for a basis for the whole space $F_2(\mathbb{Z}_3)$

$$(\widetilde{\vec{y}_1}, ..., \widetilde{\vec{y}_s}, \vec{y}_{s+1}, ..., \vec{y}_d) := \left(R^t, \vec{e}_3, \vec{e}_6, \vec{e}_8, \vec{e}_9\right)$$

In this specific example we use the standard lexicographic ordering on $(\mathbb{N}_0)^2$ and so we have

$$M_3^2 = \{(2,2), (2,1), (2,0), (1,2), (1,1), (1,0), (0,2), (0,1), (0,0)\}$$

and

$$(g_{23\alpha}(\vec{x}))_{\alpha \in M_3^2} = \left(x_2^2x_1^2, x_2^2x_1, x_2^2, x_2x_1^2, x_2x_1, x_2, x_1^2, x_1, 1\right)$$

Thus the orthonormal basis $(\widetilde{u_1}, ..., \widetilde{u_s}, u_{s+1}, ..., u_d)$ of $F_2(\mathbb{Z}_3)$ evaluated at the point $\vec{x} \in \mathbb{Z}_3^2$ would be

$$\begin{pmatrix} x_2^2x_1^2 + z_{1,3}x_2^2 + z_{1,6}x_2 + z_{1,8}x_1 + z_{1,9} \\ x_2x_1^2 + z_{2,3}x_2^2 + z_{2,6}x_2 + z_{2,8}x_1 + z_{2,9} \\ x_2x_1^2 + z_{3,6}x_2 + z_{3,8}x_1 + z_{3,9} \\ x_2x_1 + z_{4,6}x_2 + z_{4,8}x_1 + z_{4,9} \\ x_1^2 + z_{5,8}x_1 + z_{5,9} \\ x_2^2 \\ x_2 \\ x_1 \\ 1 \end{pmatrix}^t$$

and the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$ is a vector in Span$(x_2^2$, x_2 , x_1 , $1)$.

In the next section, we will establish the exact relationship between the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$ (using the symmetric bilinear form defined above) and the normal form with respect to the vanishing ideal $I(X)$. This relationship can be established if the order relation $>$ used to order the n-tuples in the set M_q^n is a *monomial ordering*. If, more generally, total orderings on $(\mathbb{N}_0)^n$ are used to order the set M_q^n, the set of possible orthogonal solutions of $\Phi_{\vec{X}}(g) = \vec{b}$ can be seen as a wider class of normal forms (with respect to vanishing ideals) in which the "classical" normal forms (attached to monomial orderings) appear as special cases.

4.5 Orthogonal solutions of $\Phi_{\vec{X}}(g) = \vec{b}$ and the normal form with respect to $I(X)$

In this section we will show the main result of this chapter: Given a set of points $X \subset K^n$, an arbitrary polynomial $f \in K[\tau_1, ..., \tau_n]$ and a monomial order $>$, the normal form of f with respect to the vanishing ideal $I(X) \subseteq K[\tau_1, ..., \tau_n]$ can be calculated as the orthogonal solution of

$$\Phi_{\vec{X}}(g) = \vec{b}$$

where $\vec{b}$ is given by

$$b_i := \widetilde{f}(\vec{x}_i),\ i = 1, ..., m$$

The yet undefined notation $\widetilde{f}$ suggests that a mapping between the ring $K[\tau_1, ..., \tau_n]$ of polynomials and the vector space of functions $F_n(\mathbf{F}_q)$ is needed. That mapping will be defined and characterized in the first lemma and theorem of this section. After introducing some notation we arrive at an important preliminary result in Theorem 208, which states how a (particular) basis of $\ker(\Phi_{\vec{X}})$ can be extended to a Gröbner basis of $I(X)$. With that result our goal can be easily reached. Please note that through this section a more technical result stated and proved in the appendix is used.

Lemma 204 (and Definition) *Let K be a field, $n, q \in \mathbb{N}$ natural numbers and $K[\tau_1, ..., \tau_n]$ the polynomial ring in n indeterminates over K. Then the set of all polynomials of the form*

$$\sum_{\alpha \in M_q^n} a_\alpha \tau_1^{\alpha_1} ... \tau_n^{\alpha_n} \in K[\tau_1, ..., \tau_n]$$

with coefficients $a_\alpha \in K$ is a vector space over K. We denote this set with $P_q^n(K) \subset K[\tau_1, ..., \tau_n]$.

Proof. The easy proof is left to the reader. ■

Theorem 205 *Let $\mathbf{F}_q$ be a finite field and $n \in \mathbb{N}$ a natural number. Then the vector spaces $P_q^n(\mathbf{F}_q)$ and $F_n(\mathbf{F}_q)$ are isomorphic.*

Proof. After defining the linear mapping

$$\begin{aligned} \varphi &: P_q^n(\mathbf{F}_q) \to F_n(\mathbf{F}_q) \\ g &= \sum_{\alpha \in M_q^n} a_\alpha \tau_1^{\alpha_1} ... \tau_n^{\alpha_n} \mapsto \varphi(g)(\vec{x}) := \sum_{\alpha \in M_q^n} a_\alpha \overrightarrow{x}^\alpha \end{aligned}$$

the claim follows easily. ■

Remark 206 (and Definition) *The mapping φ is defined on the set $P_q^n(K) \subset K[\tau_1, ..., \tau_n]$, but of course it can naturally be extended to $K[\tau_1, ..., \tau_n]$ as*

$$\begin{aligned} \varphi &: K[\tau_1, ..., \tau_n] \to F_n(\mathbf{F}_q) \\ g &= \sum_{\alpha \in \Gamma} a_\alpha \tau_1^{\alpha_1} ... \tau_n^{\alpha_n} \mapsto \varphi(g)(\vec{x}) := \sum_{\alpha \in \Gamma} a_\alpha \overrightarrow{x}^\alpha \end{aligned}$$

where Γ is a finite set of multi indexes. We denote the image under $\varphi : K[\tau_1, ..., \tau_n] \to F_n(\mathbf{F}_q)$ of a polynomial $g \in K[\tau_1, ..., \tau_n]$ with

$$\widetilde{g} := \varphi(g) \in F_n(\mathbf{F}_q)$$

Definition 207 *Let $d \in \mathbb{N}$ be a natural number, V a d-dimensional vector space over a field K and F a basis of V. Furthermore, let $U \subset V$ be an arbitrary* proper *subspace of V. Now let $s := \dim(U) \in \mathbb{N}$. A basis $(u_1, ..., u_s)$ of U is called a* cleaned kernel basis with respect to the basis *F if the matrix $(\vec{y}_1, ..., \vec{y}_s)^t$ whose rows are the coordinate vectors $\vec{y}_1, ..., \vec{y}_s \in K^d$ of $(u_1, ..., u_s)$ with respect to the basis F is in reduced row echelon form.*

For a tuple $\vec{x} = (x_1, ..., x_n)$ we write $x := \{x_1, ..., x_n\}$ for the set containing all the entries in the tuple $\vec{x}$.

Theorem 208 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers with $m < q^n$ and $>$ a fixed monomial order. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in \left(\mathbf{F}_q^n\right)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$ and $s := \dim(\ker(\Phi_{\vec{X}}))$. In addition, let $(u_1, ..., u_s)$ be a cleaned kernel basis of $\ker(\Phi_{\vec{X}}) \subseteq F_n(\mathbf{F}_q)$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$. Then the family of polynomials

$$\left(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$$

is a Gröbner basis of the vanishing ideal $I(X) \subseteq \mathbf{F}_q[\tau_1, ..., \tau_n]$ with respect to the monomial order $>$.

Proof. The idea of the proof is to show that

$$U := \left(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$$

generates the ideal $I(X)$ and that for any polynomial $g \in I(X)$ the remainder on division of g by U is zero. According to a well known fact about Gröbner bases (see proposition 5.38 of [8]) this is equivalent to U being a Gröbner basis for $I(X)$. For this proof, remember that the fundamental monomial functions $(g_{nq\alpha})_{\alpha \in M_q^n}$ are ordered decreasingly with respect to the order $>$.
Now let $g \in I(X) \subseteq \mathbf{F}_q[\tau_1, ..., \tau_n]$ be an arbitrary polynomial in the vanishing ideal of X. Since

$$(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n)$$

is a universal Gröbner basis for $I(\mathbf{F}_q^n)$ (see Theorem 241 in the appendix), there is a unique $r \in \mathbf{F}_q[\tau_1, ..., \tau_n]$ with the properties

1. No term of r is divisible by any of $LT(\tau_1^q - \tau_1) = \tau_1^q, LT(\tau_2^q - \tau_2) = \tau_2^q, ..., LT(\tau_n^q - \tau_n) = \tau_n^q$. That means in particular $r \in P_q^n(\mathbf{F}_q)$.
2. There is a $q \in I(\mathbf{F}_q^n)$ such that $g = q + r$

This means that when we start to divide g by the (ordered) family U we get the intermediate result

$$g = q + r$$

where the remainder $r \in P_q^n(\mathbf{F}_q)$ and $q \in \langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n \rangle = I(\mathbf{F}_q^n)$. If $r = 0$, then we are done and the remainder $\overline{g}^U$ on division of g by U is zero. If $r \neq 0$, then we know from

$$r = g - q$$

that $r \in I(X)$ ($q \in I(\mathbf{F}_q^n) \subseteq I(X)$) and this is equivalent to

$$\widetilde{r}(\vec{x}) = \varphi(r)(\vec{x}) = 0 \ \forall \ \vec{x} \in \mathbf{F}_q^n \Leftrightarrow \widetilde{r} \in \ker(\Phi_{\vec{X}})$$

Since $(u_1, ..., u_s)$ is a basis for $\ker(\Phi_{\vec{X}})$, there are unique $\lambda_i \in \mathbf{F}_q$, $i = 1, ..., s$ with

$$\widetilde{r} = \sum_{i=1}^{s} \lambda_i u_i$$

Applying the vector space isomorphism $\varphi^{-1} : F_n(\mathbf{F}_q) \to P_q^n(\mathbf{F}_q)$ to this equation yields

$$r = \sum_{i=1}^{s} \lambda_i \varphi^{-1}(u_i)$$

From the requirement on $(u_1, ..., u_s)$ to be a cleaned kernel basis of $\ker(\Phi_{\vec{X}})$ now follows for each $j \in \{1, ..., s\}$, that the leading term

$$LT(\varphi^{-1}(u_j))$$

doesn't appear in the polynomials $\varphi^{-1}(u_i)$, $i \in \{1, ..., s\} \backslash \{j\}$. Consequently, in the expression

$$\sum_{i=1}^{s} \lambda_i \varphi^{-1}(u_i)$$

no cancellation of the leading terms $LT(\varphi^{-1}(u_i))$, $i = 1, ..., s$ can occur. Therefore, the division of $r = \sum_{i=1}^{s} \lambda_i \varphi^{-1}(u_i)$ by $\left(\varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$ must yield

$$r = \sum_{i=1}^{s} \lambda_i \varphi^{-1}(u_i) + 0$$

and the remainder $\bar{g}^U$ on division of g by U is zero. As a consequence,

$$g \in \left\langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s) \right\rangle$$

and since $g \in I(X)$ was arbitrary

$$I(X) \subseteq \left\langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s) \right\rangle$$

The inclusion

$$\left\langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s) \right\rangle \subseteq I(X)$$

is given by the fact $u_1, ..., u_s \in \ker(\Phi_{\vec{X}})$ and Theorem 241. Summarizing we can say

$$\left\langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s) \right\rangle = I(X)$$

and for every $g \in I(X)$ the remainder $\bar{g}^U$ on division of g by U is zero. Now proposition 5.38 of [8] (see also the remarks after corollary 2, chapter 2, § 6 of [25]) proves the claim. ■

Theorem 209 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers with $m < q^n$ and $>$ a fixed monomial order. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$, $\vec{b} \in \mathbf{F}_q^m$ a vector, $d := \dim(F_n(\mathbf{F}_q))$ and $s := \dim(\ker(\Phi_{\vec{X}}))$. In addition, let $(u_1, ..., u_s)$ be a cleaned *kernel basis of* $\ker(\Phi_{\vec{X}}) \subseteq F_n(\mathbf{F}_q)$ *with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$, $(u_1, ..., u_s, u_{s+1}, ..., u_d)$ an orthonormal basis of $F_n(\mathbf{F}_q)$ constructed* using the standard orthonormalization *and $f \in \mathbf{F}_q[\tau_1, ..., \tau_n]$ a polynomial satisfying the interpolation conditions*

$$\widetilde{f}(\vec{x}_j) = b_j \ \forall \ j \in \{1, ..., m\}$$

Furthermore, let $U \subseteq I(X)$ be an arbitrary Gröbner basis of the vanishing ideal $I(X)$ with respect to the monomial order $>$ and v^* the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$. Then

$$\varphi^{-1}(v^*) = \overline{f}^U$$

Proof. If $\varphi^{-1}(v^*) = 0$ then $v^* = 0$ and

$$\vec{b} = \Phi_{\vec{X}}(v^*) = \Phi_{\vec{X}}(0) = \vec{0}$$

In this case we also have

$$\overline{f}^U = 0$$

and therefore

$$\varphi^{-1}(v^*) = \overline{f}^U$$

Assume $\varphi^{-1}(v^*) \neq 0$. Since the remainder on division by a Gröbner basis is independent of which Gröbner basis we use (for a fixed monomial order), the idea of the proof is to show that $\varphi^{-1}(v^*)$ is the unique remainder on division by the Gröbner basis

$$\left(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$$

(see Theorem 208). Now, since $\varphi^{-1}(v^*) \in P_q^n(\mathbf{F}_q)$, no term of $\varphi^{-1}(v^*)$ is divisible by any of the

$$LT(\tau_1^q - \tau_1) = \tau_1^q, LT(\tau_2^q - \tau_2) = \tau_2^q, ..., LT(\tau_n^q - \tau_n) = \tau_n^q$$

If terms of $\varphi^{-1}(v^*)$ would be divisible by

$$LT(\varphi^{-1}(u_1)), ..., LT(\varphi^{-1}(u_s))$$

then after division by the family

$$\left(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$$

we would have

$$\varphi^{-1}(v^*) = \sum_{i=1}^{s} h_i \varphi^{-1}(u_i) + r \tag{4.8}$$

where $h_i, r \in \mathbf{F}_q[\tau_1, ..., \tau_n]$, $i = 1, ..., s$ and either $r = 0$ or no term of r is divisible by the

$$LT(\tau_1^q - \tau_1), ..., LT(\tau_n^q - \tau_n), LT(\varphi^{-1}(u_1)), ..., LT(\varphi^{-1}(u_s))$$

If $r = 0$, then

$$\varphi^{-1}(v^*) = \sum_{i=1}^{s} h_i \varphi^{-1}(u_i)$$

and the polynomial $\varphi^{-1}(v^*)$ vanishes on the set X, that is

$$\varphi(\varphi^{-1}(v^*))(\vec{x}) = v^*(\vec{x}) = 0 \ \forall \ \vec{x} \in X$$

Consequently

$$\vec{b} = \Phi_{\vec{X}}(v^*) = \vec{0}$$

and due to the uniqueness of the orthogonal solution

$$v^* = 0$$

But this is a contradiction to our assumption $\varphi^{-1}(v^*) \neq 0$.
Now if $r \neq 0$, since no term of r is divisible by $LT(\tau_1^q - \tau_1), ..., LT(\tau_n^q - \tau_n)$, then in particular $r \in P_q^n(\mathbf{F}_q)$. Due to the fact, that $(u_1, ..., u_s, u_{s+1}, ..., u_d)$ is a basis for $F_n(\mathbf{F}_q)$, we can write

$$\widetilde{r} = \varphi(r) = \sum_{j=1}^{d} \lambda_j u_j$$

with unique $\lambda_j \in \mathbf{F}_q$, $j = 1, ..., d$. Applying the vector space isomorphism $\varphi^{-1} : F_n(\mathbf{F}_q) \to P_q^n(\mathbf{F}_q)$ to this equation yields

$$r = \sum_{j=1}^{d} \lambda_j \varphi^{-1}(u_j)$$

From the requirement on $(u_1, ..., u_s)$ to be a cleaned kernel basis of $\ker(\Phi_{\vec{X}})$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ and since the basis extension $(u_1, ..., u_s, u_{s+1}, ..., u_d)$ has been constructed using the standard orthonormalization, in the expression

$$\sum_{j=1}^{d} \lambda_j \varphi^{-1}(u_j)$$

no cancellation of the leading terms $LT(\varphi^{-1}(u_k))$, $k = 1, ..., s$ can occur. But r is not divisible by $LT(\varphi^{-1}(u_1)), ..., LT(\varphi^{-1}(u_s))$ and that forces

$$\lambda_k = 0, \ \forall \ k \in \{1, ..., s\}$$

In other words

$$r = \sum_{j=s+1}^{d} \lambda_j \varphi^{-1}(u_j) \Leftrightarrow \widetilde{r} = \varphi(r) = \sum_{j=s+1}^{d} \lambda_j u_j$$

which is equivalent to

$$\widetilde{r} \in \ker(\Phi_{\vec{X}})^{\perp} \tag{4.9}$$

From the equation (4.8) we know that

$$r = \varphi^{-1}(v^*) - \sum_{i=1}^{s} h_i \varphi^{-1}(u_i)$$

and that means

$$\widetilde{r}(\vec{x}) = v^*(\vec{x}) \ \forall \ \vec{x} \in X$$

In other words

$$\Phi_{\vec{X}}(\widetilde{r}) = \vec{b}$$

This together with (4.9) says that $\widetilde{r}$ is an orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$. From the uniqueness now follows

$$v^* = \widetilde{r} \Leftrightarrow \varphi^{-1}(v^*) = r$$

Consequently, no term of the polynomial $\varphi^{-1}(v^*)$ is divisible by any of the leading terms of the elements of the Gröbner basis (see Theorem 208)

$$G := \left(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$$

for the vanishing ideal $I(X)$. Now we define the polynomial

$$h := f - \varphi^{-1}(v^*)$$

Since v^* is a solution of $\Phi_{\vec{X}}(g) = \vec{b}$ and f satisfies the interpolation conditions

$$\widetilde{f}(\vec{x}_j) = b_j \ \forall \ j \in \{1, ..., m\}$$

we have

$$\widetilde{h}(\vec{x}) = \widetilde{f}(\vec{x}) - v^*(\vec{x}) = 0 \ \forall \ \vec{x} \in X \Leftrightarrow h \in I(X)$$

So we have a polynomial $h \in I(X)$ such that

$$f = h + \varphi^{-1}(v^*)$$

By proposition 1, chapter 2, §6 in [25], $\varphi^{-1}(v^*)$ is the unique remainder on division by the Gröbner basis G. It is a well known fact, that the remainder on division by a Gröbner basis is independent of which Gröbner basis we use, as long as we use one fixed particular monomial order. Therefore

$$\overline{f}^U = \overline{f}^G = \varphi^{-1}(v^*) \quad \blacksquare$$

Remark 210 (and main theorem) *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers with $m < q^n$ and $>$ a fixed monomial order. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$, $U \subseteq I(X)$ an arbitrary Gröbner basis of the vanishing ideal $I(X)$ and $f \in \mathbf{F}_q[\tau_1, ..., \tau_n]$ an arbitrary *polynomial. Then*

$$\overline{f}^U = \varphi^{-1}(v^*)$$

where v^* is the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$ and $\vec{b}$ is given by

$$b_i := \widetilde{f}(\vec{x}_i),\ i = 1, ..., m$$

Remark 211 *Let*

$$A := (\Phi_{\vec{X}}(g_{nq\alpha}))_{\alpha \in M_q^n} \in M(m \times q^n; \mathbf{F}_q)$$

be the matrix representing the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$ and S the matrix

$$S_{ij} := \left\langle g_{nq\alpha_i}, g_{nq\alpha_j} \right\rangle,\ i, j \in \{1, ..., q^n\}$$

representing the symmetric bilinear form with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$. Further let $\vec{y}_1, ..., \vec{y}_s \in \mathbf{F}_q^d$ be the coordinate vectors of $(u_1, ..., u_s)$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$. Then the above result states that the normal form $\overline{f}^U$ of f with respect to the Gröbner basis $U \subseteq I(X)$ can be calculated by solving the following system of inhomogeneous linear equations

$$\begin{aligned} A\vec{z} &= \vec{b} \\ \vec{y}_i^t S\vec{z} &= 0,\ i = 1, ..., s \end{aligned}$$

In some publications about applications of Gröbner bases (see, for instance, [97]) the so called *set of standard monomials* (see Definition 212) is introduced. Therefore, we finish this chapter including the relationship between the basis of $\ker(\Phi_{\vec{X}})^\perp$ and the set of standard monomials:

Definition 212 *Let K be a field, $n \in \mathbb{N}$ a natural number and $K[\tau_1, ..., \tau_n]$ the polynomial ring in n indeterminates over K. Further let $<$ be a monomial ordering and $I \subseteq K[\tau_1, ..., \tau_n]$ an ideal. Then $\langle LT(I) \rangle$ denotes the monomial ideal in $K[\tau_1, ..., \tau_n]$ generated by the leading terms of I and $O(\langle LT(I) \rangle)$ the set of all* monomials *not lying in the monomial ideal $\langle LT(I) \rangle$. The set $O(\langle LT(I) \rangle)$ is called the* set of standard monomials *associated to $<$ and I.*

Remark 213 *Note that the set $O(\langle LT(I) \rangle)$ has the property that all divisors of an element of $O(\langle LT(I) \rangle)$ are also in $O(\langle LT(I) \rangle)$.*

Theorem 214 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers with $m < q^n$ and $>$ a fixed monomial order. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$, $d := \dim(F_n(\mathbf{F}_q))$ and $s := \dim(\ker(\Phi_{\vec{X}}))$. In addition, let $(u_1, ..., u_s)$ be a cleaned *kernel basis of $\ker(\Phi_{\vec{X}}) \subseteq F_n(\mathbf{F}_q)$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$, $(u_1, ..., u_s, u_{s+1}, ..., u_d)$ an orthonormal basis of $F_n(\mathbf{F}_q)$ constructed* using the standard orthonormalization *and $I(X)$ the vanishing ideal of the set X. Then it holds*

$$O(\langle LT(I(X)) \rangle) = \left\{ \varphi^{-1}(u_{s+1}), ..., \varphi^{-1}(u_d) \right\}$$

Proof. It is important to notice that since the ideal $I(X)$ is a vanishing ideal, it contains the polynomials $\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n$ (see Theorem 241) and therefore

$$O(\langle LT(I(X))\rangle) \subset P_q^n(K)$$

Now consider a monomial $m \in O(\langle LT(I(X))\rangle)$. If we assume $\varphi(m) \in \ker(\Phi_{\vec{X}})$, then it follows $m \in I(X)$ and therefore $m \in \langle LT(I(X))\rangle$, which is a contradiction to the fact $m \in O(\langle LT(I(X))\rangle)$. Now the relationship (4.5) forces

$$\varphi(m) \in \ker(\Phi_{\vec{X}})^{\perp} \tag{4.10}$$

Since the standard orthonormalization chooses the vectors $u_{s+1}, ..., u_d$ among all canonical unit vectors, the only monomial functions in $\ker(\Phi_{\vec{X}})^{\perp}$ are exactly $u_{s+1}, ..., u_d$ and it follows from (4.10)

$$\varphi(m) \in \left\{\varphi^{-1}(u_{s+1}), ..., \varphi^{-1}(u_d)\right\}$$

and in general

$$O(\langle LT(I(X))\rangle) \subseteq \left\{\varphi^{-1}(u_{s+1}), ..., \varphi^{-1}(u_d)\right\}$$

For the other inclusion, consider any $\varphi^{-1}(u_i) \in \left\{\varphi^{-1}(u_{s+1}), ..., \varphi^{-1}(u_d)\right\}$. Assuming $\varphi^{-1}(u_i) \in \langle LT(I(X))\rangle$ would mean

$$\exists\ j \in \{1, ..., s\} \text{ s.t. } \varphi^{-1}(u_i) = hLT(\varphi^{-1}(u_j))$$

with an appropriate monomial $h \in P_q^n(K)$. The reason for this is that, according to Theorem 208,

$$\left(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n, \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\right)$$

is a Gröbner basis for $I(X)$ and $\varphi^{-1}(u_i) \in P_q^n(K)$. As a consequence, the (according to $>$) descending ordered polynomial

$$u := h(\varphi^{-1}(u_j)) = \varphi^{-1}(u_i) + R$$

where

$$R := h(\varphi^{-1}(u_j) - LT(\varphi^{-1}(u_j)))$$

would have the property

$$\varphi(u) = \varphi(h)u_j \in \ker(\Phi_{\vec{X}}) \tag{4.11}$$

Since u_i arises during the standard orthonormalization process as a canonical unit vector which is linearly independent from $u_1, ..., u_s$, the fact (4.11) would mean $\dim(\ker(\Phi_{\vec{X}})) > s$. Therefore $\varphi^{-1}(u_i) \in \langle LT(I(X))\rangle$ can not hold and we have

$$\varphi^{-1}(u_i) \in O(\langle LT(I(X))\rangle)$$

i.e.

$$\left\{\varphi^{-1}(u_{s+1}), ..., \varphi^{-1}(u_d)\right\} \subseteq O(\langle LT(I(X))\rangle)$$

■

Remark 215 *Since the dimension* $\dim(\ker(\Phi_{\vec{X}}))$ *of* $\ker(\Phi_{\vec{X}})$ *doesn't depend on the chosen monomial order* $>$*, the previous result shows that the number of elements in the set* $O(\langle LT(I(X))\rangle)$ *is an invariant among all monomial orderings. More generally, this statement is true for arbitrary fields* K *and arbitrary polynomial ideals* $I \subseteq K[\tau_1, ..., \tau_n]$ *with the property* $|O(\langle LT(I)\rangle)| < \infty$ *(see §3 of chapter 5 in [25]).*

Chapter 5

Reverse engineering of time discrete finite dynamical systems

5.1 Reverse engineering time discrete dynamical systems over a finite field

5.1.1 Definition of the reverse engineering problem

Reverse engineering is the attempt to infer the law governing a deterministic dynamical system based on successive observation or measurement of the system's evolution in time. Generally, if a real (physical or biological) system is being studied and measured, a modeling paradigm (i.e., type of mathematical model used to describe the system studied) has to be chosen by the modeler ([68]) before reverse engineering can be performed.

One well-known reverse engineering approach are the top-down methods, which try to infer network properties based on the observed global input-output-response. The observed input-output-response is usually only partially described by available experimental data. When the general structure of the law governing a deterministic dynamical system is known and only parameters of this law are undetermined, reverse engineering is also called parameter or system identification. Several methods for parameter identification have been developed, see, for instance, [68].

In the context of deterministic time discrete finite dynamical systems, the reverse engineering problem can be stated as follows: Given a time discrete finite dynamical system in n variables $F : X^n \to X^n$ and a data set $Y \subseteq X^n$ generated by iterating the function F starting at one or more initial values, can the function F be reconstructed from the observed time series Y?

[65] developed a top-down reverse engineering algorithm for deterministic time discrete dynamical systems over a finite field. Herein, we will refer to it as the LS-algorithm. The next subsection describes this specific type of reverse engineering problem and the LS-algorithm.

5.1.2 A short description of the LS-algorithm

In the modeling paradigm described by [65], a biological or biochemical system described by n varying quantities is studied by taking m consecutive measurements of each of the interacting quantities. This yields one time series

$$\tilde{\mathbf{s}}_1 = (s_{11}, s_{12}, ..., s_{1n}), ..., \tilde{\mathbf{s}}_m = (s_{m1}, s_{m2}, ..., s_{mn})$$

Such series of consecutive measurements are repeated t times starting from different initial conditions, where the length m_k of the series may vary. At the end of this experimental procedure,

several time series are obtained:

$$\begin{array}{c} \overrightarrow{\mathbf{s1}}_1, ..., \overrightarrow{\mathbf{s1}}_{m_1} \\ \vdots \\ \overrightarrow{\mathbf{sk}}_1, ..., \overrightarrow{\mathbf{sk}}_{m_k} \\ \vdots \\ \overrightarrow{\mathbf{st}}_1, ..., \overrightarrow{\mathbf{st}}_{m_t} \end{array}$$

Each point in a time series is a vector in $\mathbb{R}^n$. Time series are then discretized using a discretization algorithm that can be expressed as a map

$$D : \mathbb{R}^n \to S^n \tag{5.1}$$

where the set S is a finite field of cardinality $p := |S|$ (the cardinality of the field used is determined during the discretization process). The discretized time series can be written as

$$\overrightarrow{\mathbf{dk}}_1 := D(\overrightarrow{\mathbf{sk}}_1), ..., \overrightarrow{\mathbf{dk}}_{m_k} := D(\overrightarrow{\mathbf{sk}}_{m_k}),\ k = 1, ..., t$$

One fundamental assumption made in their paper is that the evolution in time of the discretized vectors obeys a simple rule, namely, that there is a function

$$F : S^n \to S^n$$

such that

$$\overrightarrow{\mathbf{dk}}_{i+1} = F(\overrightarrow{\mathbf{dk}}_i) \text{ for } i = 1, ..., m_k - 1,\ k = 1, ..., t \tag{5.2}$$

[65] call F the transition function of the system. One key ingredient in the LS-algorithm is the fact that the set S is endowed with the algebraic structure of a finite field. Under this assumption, the rule (5.2) reduces to a polynomial interpolation problem (see Corollary 28) in each component, i.e. for each $j \in \{1, ..., n\}$

$$\mathbf{dk}_{(i+1)j} = F_j(\overrightarrow{\mathbf{dk}}_i) \text{ for } k = 1, ..., t,\ i = 1, ..., m_k - 1 \tag{5.3}$$

The information provided by the equations (5.3) usually underdetermines the function $F_j : S^n \to S$, unless for all possible vectors $\vec{x} \in S^n$, the values $F_j(\vec{x})$ are established by (5.3). Indeed, any non-zero polynomial function that vanishes on all the data inputs

$$X := \{\overrightarrow{\mathbf{dk}}_i \mid k = 1, ..., t,\ i = 1, ..., m_k - 1\}$$

could be added to a function satisfying the conditions (5.3) and yield a different function that also satisfies (5.3). Among all those possible solutions, the LS-algorithm chooses the *most parsimonious* interpolating polynomial function $F_j : S^n \to S$ according to some chosen term order. To generate the most parsimonious function the algorithm first takes as input the discretized time series and generates functions f_j, $j = 1, ..., n$ that satisfy (5.3) for each $j \in \{1, ..., n\}$ correspondingly. Secondly, it takes a monomial order $<_j$ as input and generates the normal form of f_j with respect to the vanishing ideal $I(X)$ and the given order $<$. For every $j \in \{1, ..., n\}$, this normal form is the output F_j of the algorithm.
We also refer to 2.1 in [52] for another rigorous description of the LS-algorithm.

5.2 Orthogonality and the reverse engineering algorithm

The mathematical framework presented here is based on a general result stated in Chapter 4. This framework will allow us to study the LS-algorithm as well as a generalized algorithm that does not depend on the choice of term orders.

We start with the original problem: Given a time-discrete dynamical system over a finite field S in n variables

$$F : S^n \to S^n$$

and a data set $X \subseteq S^n$ generated by iterating the function F starting at one or more initial values, what are the chances of reconstructing the function F if the LS-algorithm or a similar algorithm is applied using X as input time series?[1] Since the algorithms studied here generate an output model $G : S^n \to S^n$ by calculating every single coordinate function $G_i : S^n \to S$ separately, we will focus on the reconstruction of a single coordinate function F_i which we will simply call f. We will use the notation $\mathbf{F}_q$ for a finite field of cardinality $q \in \mathbb{N}$. In what follows, we briefly review the main definitions and results stated and proved in Chapter 4:
We denote the q^n-dimensional vector space of functions $g : \mathbf{F}_q^n \to \mathbf{F}_q$ with $F_n(\mathbf{F}_q)$. A basis for $F_n(\mathbf{F}_q)$ is given by all the monomial functions $\overrightarrow{x}^{\alpha} := x_1^{\alpha_1} \cdot ... \cdot x_1^{\alpha_n}$ where the exponents α_i are non-negative integers satisfying $\alpha_i < q$. The set of all those monomial functions is denoted with $(g_{nq\alpha})_{\alpha \in M_q^n}$, where $M_q^n := \{\alpha \in (\mathbb{N}_0)^n \mid \alpha_j < q \ \forall \ j \in \{1, ..., n\}\}$. We call those monomial functions *fundamental monomial functions.*

Theorem 216 (and Definition) *Let $\mathbf{F}_q$ be a finite field and $n, m \in \mathbb{N}$ natural numbers with $m \leq q^n$. Further let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m ***different*** *n-tuples with entries in the field $\mathbf{F}_q$. Then the mapping*

$$\begin{array}{rcl} \Phi_{\vec{X}} & : & F_n(\mathbf{F}_q) \to \mathbf{F}_q^m \\ f & \mapsto & \Phi_{\vec{X}}(f) := (f(\vec{x}_1), ..., f(\vec{x}_m))^t \end{array}$$

is a surjective linear operator. $\Phi_{\vec{X}}$ is called the evaluation epimorphism of the tuple $\vec{X}$.

For a given set $X \subseteq \mathbf{F}_q^n$ of data points, the interpolation problem of finding a function $g \in F_n(\mathbf{F}_q)$ with the property

$$g(\vec{x}_i) = b_i \ \forall \ i \in \{1, ..., m\}, \ x_i \in X$$

can be expressed using the evaluation epimorphism as: Find a function $g \in F_n(\mathbf{F}_q)$ with the property

$$\Phi_{\vec{X}}(g) = \vec{b} \tag{5.4}$$

Since a basis of $F_n(\mathbf{F}_q)$ is given by the fundamental monomial functions $(g_{nq\alpha})_{\alpha \in M_q^n}$, the matrix

$$A := (\Phi_{\vec{X}}(g_{nq\alpha}))_{\alpha \in M_q^n} \in M(m \times q^n; \mathbf{F}_q)$$

representing the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$ has always the full rank $m = \min(m, q^n)$. That also means, that the dimension of the $\ker(\Phi_{\vec{X}})$ is

$$\dim(\ker(\Phi_{\vec{X}})) = \dim(F_n(\mathbf{F}_q)) - m = q^n - m \tag{5.5}$$

In the case $m < q^n$ where m is strictly smaller than $q^n = \left|\mathbf{F}_q^n\right|$ we have $\dim(\ker(\Phi_{\vec{X}})) > 0$ and the solution of the interpolation problem is not unique. There are exactly $q^{\dim(\ker(\Phi_{\vec{X}}))}$ different solutions which constitute an affine subspace of $F_n(\mathbf{F}_q)$. Only in the case $m = q^n$, that means,

[1] From an experimental point of view the following question arises: What is the function F in an experimental setting? Contrary to the situation when models with an infinite number of possible states are reverse engineered (see 1.2 in [68]), there is a finite number of experiments that could be, at least theoretically, performed to *completely* characterize the system studied. In this sense, even in an experimental setting, there is an underlying function F. The components of this function is what [52] called h_{true}.

when for all elements of $\mathbf{F}_q^n$ the corresponding interpolation values are given, the solution is unique. If the problem is underdetermined and no additional information about properties of the possible solutions is given, any algorithm attempting to solve the problem has to provide a selection criterion to pick a solution among the affine space of possible solutions. The LS-algorithm chooses the most parsimonious interpolating polynomial function according to some chosen term order. A more geometric approach to pick one solution would be to select the solution that is perpendicular (or orthogonal) to the affine space of solutions. As stated in Remark and Theorem 210, the solution selected by the LS-algorithm is precisely the orthogonal solution. For orthogonality to apply, a generalized inner product has to be defined on the space $F_n(\mathbf{F}_q)$. We finish this subsection reviewing this concepts (cf. Chapter 4).

The space $F_n(\mathbf{F}_q)$ is endowed with a symmetric bilinear form $\langle \cdot, \cdot \rangle : F_n(\mathbf{F}_q) \times F_n(\mathbf{F}_q) \to \mathbf{F}_q$, i.e. a generalized inner product. Orthogonality and orthonormality are defined as in an Euclidean vector space.

For a given set $X \subseteq \mathbf{F}_q^n$ of data points, consider the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$ and its kernel $\ker(\Phi_{\vec{X}})$. Now, let $(u_1, ..., u_s)$ be a basis of $\ker(\Phi_{\vec{X}}) \subseteq F_n(\mathbf{F}_q)$. By the basis extension theorem, we can extend the basis $(u_1, ..., u_s)$ to a basis

$$(u_1, ..., u_s, u_{s+1}, ..., u_d)$$

of the whole space $F_n(\mathbf{F}_q)$. (There are many possible ways this extension can be performed. See more details below). As in Example 192, we can construct a generalized inner product on $F_n(\mathbf{F}_q)$ by setting

$$\langle u_i, u_j \rangle := \delta_{ij} \ \forall \ i, j \in \{1, ..., d\}$$

The orthogonal solution of (5.4) is the solution $v^* \in F_n(\mathbf{F}_q)$ that is orthogonal to $\ker(\Phi_{\vec{X}})$, i.e. it holds $\Phi_{\vec{X}}(v^*) = \vec{b}$ and for an arbitrary basis $(w_1, ..., w_s)$ of $\ker(T)$ the following orthogonality conditions hold

$$\langle w_i, v^* \rangle = 0 \ \forall \ i \in \{1, ..., s\}$$

The way we extend the basis $(u_1, ..., u_s)$ of $\ker(\Phi_{\vec{X}})$ to a basis

$$(u_1, ..., u_s, u_{s+1}, ..., u_d)$$

of the whole space $F_n(\mathbf{F}_q)$ determines crucially the generalized inner product we get by setting

$$\langle u_i, u_j \rangle := \delta_{ij} \ \forall \ i, j \in \{1, ..., d\} \tag{5.6}$$

Consequently, the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$ may vary according to the chosen extension $u_{s+1}, ..., u_d \in F_n(\mathbf{F}_q)$. In Chapter 4 a systematic way to extend the basis $(u_1, ..., u_s)$ to a basis for the whole space is introduced. With the basis obtained, the process of defining a generalized inner product according to (5.6) is called the *standard orthonormalization.* This is because the basis $(u_1, ..., u_s, u_{s+1}, ..., u_d)$ is orthonormal with respect to the generalized inner product defined by (5.6).

As shown in Section 4.5 of Chapter 4, using the generalized inner product obtained by applying the standard orthonormalization, the functions generated by the LS-algorithm are orthogonal solutions of the polynomial interpolation problem as formulated in (5.4). Under these assumptions the orthogonal solution is also unique (see Theorem 196).

The standard orthonormalization process depends on the way the elements of the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of fundamental monomial functions are ordered. If they are ordered according to a term order, the calculation of the orthogonal solution of (5.4) yields the same result as the LS-algorithm. If more general linear orders are allowed, a more general algorithm emerges that is not restricted to the use of term orders. This algorithm can be seen as a generalization of the LS-algorithm. We call it the *term-order-free reverse engineering method.* The precise definition of the standard orthonormalization procedure is stated in Section 4.4 of Chapter 4. In the appendix we summarize the steps of the term-order-free reverse engineering method.

5.3 Performance of the reverse engineering method

5.3.1 Questions studied

The mathematical framework developed in the previous subsection will allow us to answer the following questions regarding the LS-algorithm and its generalization, the term-order-free reverse engineering method:

Problem 217 *Given a function $f \in F_n(\mathbf{F}_q)$, what are the minimal requirements on a set $X \subseteq \mathbf{F}_q^n$, such that the LS-algorithm reverse engineers f based on the knowledge of the values that it takes on every point in the set X?*

Problem 218 *Are there sets $X \subseteq \mathbf{F}_q^n$ that make the LS-algorithm more likely to succeed in reverse engineering a function $f \in F_n(\mathbf{F}_q)$ based* only *on the knowledge of the values that it takes on every point in the set X?*[2]

Problem 219 *Given a function $f \in F_n(\mathbf{F}_q)$ and an optimal set $X \subseteq \mathbf{F}_q^n$ (in the sense of the previous problem). If the term order used by the LS-algorithm is chosen randomly, can the probability of success be calculated? If the linear order used by the term-order-free method is chosen randomly, can the probability of success be calculated?*

Problem 220 *What is the asymptotic behavior of the probability for a growing number of variables n?*

It is pertinent to emphasize that, contrary to the scenario studied in [52], we do not necessarily assume that information about the number of variables actually affecting f is available. We will give further comments on this issue at the end of the conclusions.

5.3.2 Results

Basic definitions and facts

For what follows recall that $M_q^n = \{\alpha \in (\mathbb{N}_0)^n \mid \alpha_j < q \ \forall\ j \in \{1, ..., n\}\}$. The easy proof of the following two propositions is left to the reader.

Lemma 221 (and Definition) *Let K be a field, $n, q \in \mathbb{N}$ natural numbers and $K[\tau_1, ..., \tau_n]$ the polynomial ring in n indeterminates over K. Then the set of all polynomials of the form*

$$\sum_{\alpha \in M_q^n} a_\alpha \tau_1^{\alpha_1} ... \tau_n^{\alpha_n} \in K[\tau_1, ..., \tau_n]$$

with coefficients $a_\alpha \in K$ is a vector space over K. We denote this set with $P_q^n(K) \subset K[\tau_1, ..., \tau_n]$.

Theorem 222 *Let $\mathbf{F}_q$ be a finite field and $n \in \mathbb{N}$ a natural number. Then the vector spaces $P_q^n(\mathbf{F}_q)$ and $F_n(\mathbf{F}_q)$ are isomorphic via the mapping*

$$\begin{aligned} \varphi &: P_q^n(\mathbf{F}_q) \to F_n(\mathbf{F}_q) \\ g &= \sum_{\alpha \in M_q^n} a_\alpha \tau_1^{\alpha_1} ... \tau_n^{\alpha_n} \mapsto \varphi(g)(\vec{x}) := \sum_{\alpha \in M_q^n} a_\alpha \overrightarrow{x}^\alpha \end{aligned}$$

[2] A solution to this problem would provide criteria for the design of experiments.

Definition 223 *Let K be a field, $n, m \in \mathbb{N}$ natural numbers and $K[\tau_1, ..., \tau_n]$ the polynomial ring in n indeterminates over K. Furthermore, let $g_1, ..., g_m \in K[\tau_1, ..., \tau_n]$ be polynomials. The set*

$$\langle g_1, ..., g_m \rangle := \{h_1 g_1 + ...h_m g_m \mid h_1, ..., h_m \in K[\tau_1, ..., \tau_n]\}$$

is called the ideal *generated by $g_1, ..., g_m$.*

For a tuple $\vec{x} = (x_1, ..., x_n)$ we write $x := \{x_1, ..., x_n\}$ for the set containing all the entries in the tuple $\vec{x}$.

Conditions on the data set

Definition 224 *Let $f \in F_n(\mathbf{F}_q)$ be a polynomial function. The subset of $\mathbf{F}_q^n$ containing all values on which the polynomial function f vanishes is denoted by*

$$V(\varphi^{-1}(f))$$

where φ is the mapping defined in Theorem 222.

The following result tells us that if we are using the LS-algorithm to reverse engineer a nonzero function we necessarily have to use a data set X containing points where the function does not vanish.

Theorem 225 *Let $f \in F_n(\mathbf{F}_q)\backslash\{0\}$ be a nonzero polynomial function. Furthermore let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$, $\vec{b} \in \mathbf{F}_q^m$ be the vector defined by

$$b_i := f(\vec{x}_i),\ i = 1, ..., m$$

and $v^*$the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$. Then if $v^* = f$ it follows[3]

$$V(\varphi^{-1}(f))^c \cap X \neq \emptyset$$

Proof. If $V(\varphi^{-1}(f))^c \cap X = \emptyset$ then by definition of $V(\varphi^{-1}(f))$, the vector $\vec{b}$ would be equal to the zero vector $\vec{0}$. From Corollary 198 we know that the orthogonal solution $v^*$of $\Phi_{\vec{X}}(g) = \vec{0}$ is the zero function, thus $v^* \neq f$. ■

Theorem 226 *Let $f \in F_n(\mathbf{F}_q)\backslash\{0\}$ be a nonzero polynomial function. Furthermore let*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

be a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$, $\vec{b} \in \mathbf{F}_q^m$ be the vector defined by

$$b_i := f(\vec{x}_i),\ i = 1, ..., m$$

and $v^*$the orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$. In addition, assume $V(\varphi^{-1}(f))^c \cap X \neq \emptyset$. Then it holds

$$v^* = f \Leftrightarrow f \in span(u_{s+1}, ..., u_d)$$

Proof. The claim follows directly from the definition of orthogonal solution and its uniqueness. ■

[3] If A is a set, A^c denotes its complement

Remark 227 *From the necessary and sufficient condition*

$$f \in span(u_{s+1}, ..., u_d) \tag{5.7}$$

it becomes apparent, that if the function f is a linear combination of more than $d - s = m$ fundamental monomial functions, f can not be found as an orthogonal solution v^* of $\Phi_{\vec{X}}(g) = \vec{b}$. In particular, if f is a linear combination containing all d fundamental monomial functions in $(g_{nq\alpha})_{\alpha \in M_q^n}$, no proper *subset $X \subset \mathbf{F}_q^n$ of $\mathbf{F}_q^n$ will allow us to find f as orthogonal solution of $\Phi_{\vec{X}}(g) = \vec{b}$ (where $b_i := f(\vec{x}_i)$, $\vec{x}_i \in X$).*

Remark 228 *It follows from condition (5.7), that it is* necessary *that a monomial function appearing in f is linearly independent of the basis vectors $u_1, ..., u_s$ of $\ker(\Phi_{\vec{X}})$. For this reason, the set X should be chosen in such a way that no fundamental monomial function $(g_{nq\alpha})_{\alpha \in M_q^n}$ is linearly dependent on the basis vectors $u_1, ..., u_s$ of $\ker(\Phi_{\vec{X}})$. Otherwise, some of the terms appearing in f might vanish on the set X and wouldn't be detectable by any reverse engineering method, (as stated in [65]). This problem introduces a more general question about the existence of vector subspaces in "general position":*

Definition 229 *Let W be a finite dimensional vector space over a finite field $\mathbf{F}_q$ with $\dim(W) = d > 0$. Furthermore, let $(w_1, ..., w_d)$ be a fixed basis of W and $s \in \mathbb{N}$ a natural number with $s < d$. A vector subspace $U \subset W$ with $\dim(U) = s$ is said to be in* general position *with respect to the basis $(w_1, ..., w_d)$ if for any basis $(v_1, ..., v_s)$ of U and any injective mapping*

$$\pi : \{1, ..., (d - s)\} \to \{1, ..., d\}$$

the vectors

$$v_1, ..., v_s, w_{\pi(1)}, ..., w_{\pi(d-s)}$$

are linearly independent.

It can be shown, that if the cardinality q of the finite field $\mathbf{F}_q$ is sufficiently large, proper subspaces in general position of any positive dimension always exist. The proof is provided in the appendix.

Now assume that $\ker(\Phi_{\vec{X}})$ is in general position with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of $F_n(\mathbf{F}_q)$. Following the basis extension theorem and due to the general position of $\ker(\Phi_{\vec{X}})$, we can extend the basis $(u_1, ..., u_s)$ of $\ker(\Phi_{\vec{X}})$ to a basis

$$(u_1, ..., u_s, u_{s+1}, ..., u_d)$$

of the whole space $F_n(\mathbf{F}_q)$, where $\{u_{s+1}, ..., u_d\} \subset \{g_{nq\alpha}\}_{\alpha \in M_q^n}$ is *any* subset with $d - s$ elements of $\{g_{nq\alpha}\}_{\alpha \in M_q^n}$. Now we can construct a generalized inner product on $F_n(\mathbf{F}_q)$ by setting

$$\langle u_i, u_j \rangle := \delta_{ij} \; \forall \; i, j \in \{1, ..., d\}$$

The advantage in this situation is that there is no bias imposed by the data on the monomial functions that can be used to extend the basis $(u_1, ..., u_s)$ to a basis of $F_n(\mathbf{F}_q)$, i.e. there are no restrictions on the structure of $\ker(\Phi_{\vec{X}})^{\perp}$. In addition, having this degree of freedom, it is possible to calculate the exact probability of success of the method based on the number of fundamental monomial functions actually contained in f. We will give an explicit probability formula in the next Subsection. For our further analysis we need the following intermediate result, whose proof is left to the reader:

Lemma 230 (and Definition) *Let $\mathbf{F}_q$ be a finite field, $n, s \in \mathbb{N}$ natural numbers with $s \leq \dim(F_n(\mathbf{F}_q))$. Furthermore, let $U \subset F_n(\mathbf{F}_q)$ be an s-dimensional subspace. Then the set*

$$V(U) := V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s) \rangle) \subseteq \mathbf{F}_q^n$$

where $(u_1, ..., u_s)$ is any basis of U is independent on the choice of basis and it's called the variety of the subspace U.

Now the following question arises: How should the set X be chosen in order to have $\ker(\Phi_{\vec{X}})$ in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$? For a given natural number $s < d := \dim(F_n(\mathbf{F}_q))$ the idea is to start from a basis $(u_1, ..., u_s)$ of a vector subspace $U \subset F_n(\mathbf{F}_q)$ in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$. The next step is to calculate the variety

$$Y := V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\rangle) \subseteq \mathbf{F}_q^n$$

We assume $Y \neq \emptyset$ and order its elements arbitrarily to a tuple $\vec{Y} := (\vec{y}_1, ..., \vec{y}_m) \in (\mathbf{F}_q^n)^m$, where $m := |Y|$. We know from Remark 200 that $\dim(\ker(\Phi_{\vec{Y}})) = \dim(F_n(\mathbf{F}_q)) - |Y| = d - m$. Now, in general, for the kernel $\ker(\Phi_{\vec{Y}})$ of the corresponding evaluation epimorphism $\Phi_{\vec{Y}}$ it holds

$$U \subseteq \ker(\Phi_{\vec{Y}})$$

and therefore $s \leq \dim(\ker(\Phi_{\vec{Y}})) = d - m$, i.e. $m \leq d - s$. Now, the ideal scenario would be the case $\ker(\Phi_{\vec{Y}}) = U$, i.e. $m = d - s$. A less optimistic scenario is given when $U \subset \ker(\Phi_{\vec{Y}})$ is a proper subspace of $\ker(\Phi_{\vec{Y}})$. In such a situation, ideally we would wish for $\ker(\Phi_{\vec{Y}})$ to be itself in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$. This issues raise the following question: When does there exist a subspace $U \subset F_n(\mathbf{F}_q)$ in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ with $\dim(U) < \dim(F_n(\mathbf{F}_q))$ that in addition satisfies

$$\left|V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\rangle)\right| = \dim(F_n(\mathbf{F}_q)) - \dim(U) \tag{5.8}$$

This is an interesting question that requires further research. It is related to whether the subspace U is an ideal of $F_n(\mathbf{F}_q)$, when $F_n(\mathbf{F}_q)$ is seen as an algebra with the multiplication of polynomial functions as the multiplicative operation. In the Appendix we provide examples in which two subspaces, both in general position, show a different behavior regarding the condition (5.8). We formalize this property:

Definition 231 *Let $U \subset F_n(\mathbf{F}_q)$ be a subspace and $(u_1, ..., u_s)$ an arbitrary basis of U. U is said to satisfy the* codimension condition *if it holds*

$$codim(U) = \left|V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\rangle)\right|$$

where $codim(U) := \dim(F_n(\mathbf{F}_q)) - \dim(U)$.

A subspace $U \subset F_n(\mathbf{F}_q)$ in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ that satisfies the codimension condition allows for the construction of an optimal set for use with the LS-algorithm. The set $Y := V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\rangle)$ (where $u_1, ..., u_s$ is a basis of U) has namely the property $\ker(\Phi_{\vec{Y}}) = U$, i.e. $\ker(\Phi_{\vec{Y}})$ is in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$. In other words, subspaces in general position that satisfy the codimension condition provide a basic component for a constructive method for generating optimal data sets. More generally we define:

Definition 232 *A set $X \subseteq \mathbf{F}_q^n$ such that $\ker(\Phi_{\vec{X}})$ is in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$ is referred to as* optimal.

Remark 233 (and Definition) *Additional study is required to prove whether optimal data sets exist in general. (See the Appendix for concrete examples.) However, if no optimal sets can be determined, it is still advantageous to work with a data set X that was obtained as $V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_s)\rangle)$, where $(u_1, ..., u_s)$ is a basis for a subspace U in general position with respect to the basis $(g_{nq\alpha})_{\alpha\in M_q^n}$. In this case, at least $U \subseteq \ker(\Phi_{\vec{Y}})$ still holds and it might be that the dimensional difference between U and $\ker(\Phi_{\vec{Y}})$ is small. We call such data sets* pseudo-optimal.

Probabilities of finding the original function as the orthogonal solution

Theorem 234 *Let $\mathbf{F}_q$ be a finite field, $n, m \in \mathbb{N}$ natural numbers with $m < \dim(F_n(\mathbf{F}_q)) =: d$. Furthermore, let $f \in F_n(\mathbf{F}_q)\backslash\{0\}$ be a nonzero function consisting of a linear combination of exactly t fundamental monomial functions and*

$$\vec{X} := (\vec{x}_1, ..., \vec{x}_m) \in (\mathbf{F}_q^n)^m$$

a tuple of m different n-tuples with entries in the field $\mathbf{F}_q$ such that X is optimal. Now let $\vec{b} \in \mathbf{F}_q^m$ be the vector defined as

$$b_i := f(\vec{y}_i),\ i = 1, ..., m$$

$s := \dim(\ker(\Phi_{\vec{X}})) = d - m$ (cf. (5.5)), $(u_1, ..., u_s)$ a basis for $\ker(\Phi_{\vec{X}})$ and $\{u_{s+1}, ..., u_d\} \subset \{g_{nq\alpha}\}_{\alpha \in M_q^n}$ an arbitrary subset containing $d - s$ elements. Then the probability P that the orthogonal solution g^* of $\Phi_{\vec{X}}(g) = \vec{b}$ with respect to the generalized inner product

$$\langle u_i, u_j \rangle := \delta_{ij}\ \forall\ i, j \in \{1, ..., d\}$$

fulfills $f = g^*$ is given by

$$P = \frac{\binom{q^n - t}{q^n - m}}{\binom{q^n}{m}} \text{ if } t \leq m \tag{5.9}$$

and

$$P = 0 \text{ if } t > m$$

Proof. Due to the definition of general position, there are exactly

$$(d-s)! \binom{\dim(F_n(\mathbf{F}_q))}{\dim(F_n(\mathbf{F}_q)) - s} = (d-s)! \binom{d}{d-s} = (d-s)! \binom{q^n}{m}$$

different ways to extend a basis $(u_1, ..., u_s)$ of U to a basis of $F_n(\mathbf{F}_q)$ using $m = d - s$ fundamental monomial functions. If $t \leq m$, among such extensions, only

$$(d-s)! \binom{d-t}{d-s-t} = (d-s)! \binom{q^n - t}{s} = (d-s)! \binom{q^n - t}{q^n - m}$$

use the t fundamental monomial functions appearing in f. Now (5.9) follows immediately. If, on the other hand, $t > m$, the number of fundamental monomial functions usable to extend a basis $(u_1, ..., u_s)$ of $\ker(\Phi_{\vec{X}})$ to a basis of $F_n(\mathbf{F}_q)$ is too small and $\ker(\Phi_{\vec{X}})^\perp$ is not big enough to generate f. ■

Remark 235 *If the elements in the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ are ordered in a decreasing way according to a term order (the biggest element is at the left end, the smallest at the right end and position t means counting t elements from the right to the left) an analogous probability formula would be*

$$P = \frac{\textit{Number of arrangements that place the mon. functions in } f \textit{ after position } s}{\textit{Total number of arrangements}} \tag{5.10}$$

where an arrangement is an order of the elements of $(g_{nq\alpha})_{\alpha \in M_q^n}$ that obeys a term order. (Two different term orders could generate the same arrangement of the elements in the finite set $\{g_{nq\alpha}\}_{\alpha \in M_q^n}$). So, for instance, if f contains a term involving the monomial function $x_1^{q-1} \cdot ... \cdot x_n^{q-1}$, then the above probability (5.10) would be equal to zero, since every arrangement of the elements in $\{g_{nq\alpha}\}_{\alpha \in M_q^n}$ that obeys a term order would make that monomial function biggest. (It is inherent to term orders to make some monomial functions always *biggest). In more general terms, it is difficult to make estimates about the numbers involved in (5.10). This shows some of the disadvantages of using term orders.*

Remark 236 *Since for relatively small n and q the number $d := q^n$ is already very large, it is obvious that one should calculate the asymptotic behavior of the probability formula (5.9) for $d \to \infty$. Indeed, we have with $t \leq m$*

$$\begin{aligned} 0 &\leq \frac{\binom{d-t}{d-m}}{\binom{d}{m}} = \frac{\frac{(d-t)!}{(d-m)!(m-t))!}}{\frac{d!}{m!(d-m)!}} \\ &= \frac{(d-t)!m!}{(m-t)!d!} \leq \frac{(d-t)!m!}{d!} \\ &= \frac{m!}{d(d-1)...(d-t+1)} \longrightarrow 0 \text{ for } d \longrightarrow \infty \end{aligned}$$

If we write the amount of data used in proportion to the size $d = q^n$ of the space $\mathbf{F}_q^n$, and the number of terms displayed by f relative to the size q^n of the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$, it becomes apparent how quickly the probability formula converges to 0 for $d \longrightarrow \infty$. So let $r := m/d$ and $\gamma := d - t$. Then we would have

$$\begin{aligned} \frac{P}{r^d} &= \frac{\binom{d-t}{d-m}}{r^d \binom{d}{m}} = \frac{(d-t)!m!}{r^d(m-t)!d!} \\ &= \frac{m(m-1)...(m-t+1)}{r^d d(d-1)...(d-t+1)} = \frac{rd(rd-1)...(rd-t+1)}{r^d d(d-1)...(d-t+1)} \\ &= \frac{rdrd(1-\frac{1}{rd})...rd(1-\frac{t-1}{rd})}{r^d dd(1-\frac{1}{d})...d(1-\frac{t-1}{d})} = \frac{r^t d^t(1-\frac{1}{rd})...(1-\frac{t-1}{rd})}{r^d d^t(1-\frac{1}{d})...(1-\frac{t-1}{d})} \\ &= \frac{r^t(1-\frac{1}{rd})...(1-\frac{t-1}{rd})}{r^d(1-\frac{1}{d})...(1-\frac{t-1}{d})} \\ &= \frac{r^{-\gamma}(1-\frac{1}{rd})...(1-\frac{t-1}{rd})}{(1-\frac{1}{d})...(1-\frac{t-1}{d})} \longrightarrow r^{-\gamma} \text{ for } d \longrightarrow \infty \end{aligned}$$

In particular, it holds

$$\frac{\binom{d-t}{d-rd}}{\binom{d}{rd}} \approx r^t \text{ for big } d$$

This expression shows in a straightforward way how big the proportional amount of data should be in order to have an acceptable confidence in the obtained result. It also shows that for t close to d the probability is very low and the reverse engineering not feasible. Usually no information about t is available, so it is advisable to work with the maximal t, namely $d-1$ or with an average value for t.

5.3.3 Conclusions

The results we have obtained in the previous section provide guidelines on how to design experiments to generate data to be used with the LS-algorithm for the purpose of reverse engineering a biochemical network.
The following are minimal requirements on a set $X \subseteq \mathbf{F}_q^n$, such that the LS-algorithm reverse engineers f based on the knowledge of the values that it takes on every point in the set X :

1. If the LS-algorithm is used to reverse engineer a nonzero function $f \in F_n(\mathbf{F}_q)\backslash\{0\}$, necessarily the data set X used must contain points were the function does not vanish. In other words, not all the interpolation conditions must be of the type $\vec{x}_i \mapsto 0$ (Theorem 225).

2. If the LS-algorithm is used to reverse engineer a function $f \in F_n(\mathbf{F}_q)\backslash\{0\}$ displaying t different terms, it requires **at least** t different data points to *completely* reverse engineer f (Remark 227).

3. If $f \in F_n(\mathbf{F}_q)\backslash\{0\}$ is a polynomial function containing all p^n possible fundamental monomial functions, no *proper* subset $X \subset \mathbf{F}_q^n$ of $\mathbf{F}_q^n$ will allow the LS-algorithm to find f (Remark 227).

Our results also make possible the identification of optimal sets $X \subseteq \mathbf{F}_q^n$ that make the LS-algorithm more likely to succeed in reverse engineering a function $f \in F_n(\mathbf{F}_q)$ based only on the knowledge of the values that it takes on every point in the set X. Optimal data sets $X \subset \mathbf{F}_q^n$ are characterized by the property that $\ker(\Phi_{\vec{X}})$ is in general position with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ (see Definitions 232 and 229). Their advantage is given by the fact that they do not impose constraints on the set of candidate terms that can be used to construct a solution. Summarizing we can say:

1. Even though such sets can be constructed in particular examples (see Appendix), further research is required to prove their existence in general terms.

2. If no optimal sets can be determined, it is still advantageous to work with pseudo-optimal data sets (see Remark and Definition 233).

Since the identified optimal data sets are sets $X \subset \mathbf{F}_q^n$ of discretized vectors, in a real application, the optimal data set X has to be transformed back to a corresponding set $\widetilde{X} \subset \mathbb{R}^n$ of real vectors. This transformation can be performed using an "inverse" function of the discretization mapping (5.1). This "inverse" function has to be defined by the user, given the fact that discretization mappings are highly non-injective and by definition map entire subsets $Z \subset \mathbb{R}^n$ into a single value $\vec{z} \in \mathbf{F}_q^n$.

Having characterized optimal data sets, the next step in our approach was to provide an exact formula for the probability that the LS-algorithm will find the correct model under the assumption that an optimal data set is used as input. As stated in Remark 235, we weren't able to find such a formula for the LS-algorithm. The biggest difficulty we face is related to the use of term orders inherent to the LS-algorithm. We overcome this problem by considering a generalization of the LS-algorithm which we call the term-order-free reverse engineering method (see Appendix). This method not only allows for the calculation of the success probability but it also eliminates the issues and arbitrariness linked to the use of term orders (see Remark 235). In conclusion, our results on this issue are:

1. It is still an open problem how to derive a formula for the success probability of the LS-algorithm when optimal data sets are used as an input and the term order is chosen randomly. As stated in Remark 235, one of the main problems here is related to the use of term orders inherent to the LS-algorithm.

2. Let $f \in F_n(\mathbf{F}_q)\backslash\{0\}$ be a nonzero function consisting of the linear combination of exactly t fundamental monomial functions. If the linear order used by the term-order-free method is chosen randomly, the probability of successfully retrieving f using an optimal data set X of cardinality $|X| = m$ is given by (see Theorem 234)

$$P = \frac{\binom{q^n - t}{q^n - m}}{\binom{q^n}{m}} \text{ if } t \leq m \tag{5.11}$$

and

$$P = 0 \text{ if } t > m$$

3. Let $d = q^n$ be the cardinality of the space $\mathbf{F}_q^n$. Furthermore, let X be an optimal data set with cardinality $|X| = m$ and $r := m/d$ (note that $0 < r < 1$). Then the asymptotic behavior of the probability formula (5.11) for $d \to \infty$ (i.e. for $n \to \infty$) satisfies (see Remark 236)

$$\frac{\binom{d-t}{d-rd}}{\binom{d}{rd}} \approx r^t \text{ for big } d$$

As a consequence of the latter, we conclude that even if an optimal data set is used and the restrictions imposed by the use of term orders are overcome, the reverse engineering problem remains unfeasible, unless experimentally impracticable amounts of data are available.

Finally, we comment on one scenario identified in [52]. Specifically, in Conclusion 4(a), [52] makes the assumption that the wiring diagram of each of the underlying functions is known, i.e. the variables that actually affect the function f are known. Under this assumption, let k be the number of variables affecting f. If one could perform specific experiments such that for all possible values that the k variables can take the response of the network is measured, the function f would be uniquely determined. In this situation, reverse engineering f wouldn't imply making any choices among possible solutions. This raises the question of how many measurements are needed and how big this data set would be in proportion to the size q^n of the space $\mathbf{F}_q^n$ of all possible states the network can theoretically display. The number of measurements needed is q^k and therefore the proportion is equal to

$$\frac{q^k}{q^n} = \frac{1}{q^{n-k}}$$

If k is small compared to n (which is generally assumed by [52]), then the proportion would be conveniently small. In other words, in *relative* terms, it is worth performing the q^k specific experiments. However, performing q^k measurements might still be beyond experimental feasibility.

5.4 Issues related to the discretization of time series

It would go beyond the scope of this thesis to study all the critical issues related to the discretization[4] of real valued time series (see equation (5.1)). However, we finish this chapter with a short study of the effects of a finer discretization on the output of the LS-algorithm. We start providing a rigorous analytical one-dimensional (one variable case) counterexample to the statement, "It follows from results in Green (2003) that for p large enough the result of our reverse-engineering algorithm does not depend on p; in the sense that the terms in the polynomials remain the same, possibly with different coefficients." made by [65]. Since the polynomial ring in one indeterminate differs from polynomial rings in more than one indeterminate in some algebraic properties, we performed Monte Carlo simulations in the two variable case. These simulations showed no type of stabilization of the LS-algorithm's output as the number p of possible states increases. In reality, the total degree of the polynomial functions seems to grow unrestrainedly.
In the one variable case we begin with a realistic looking source of data: A continuous quantity that grows monotonically and then stabilizes after a certain period of time (a phenomenon widely observed in chemical and biochemical reactions). The time series is obtained by periodic sampling over a time interval, that is sufficiently large to allow for stabilization. In order to simulate an

[4] Also called *quantization*, see 1-3 in [83].

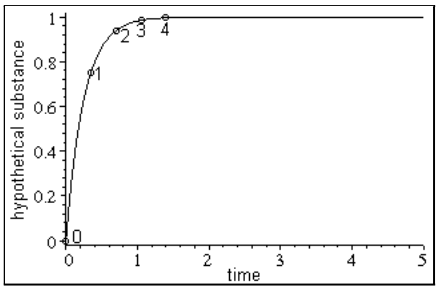

Figure 5.1: The continuous entity and 5 sampled points.

increasing number of possible states, we progressively shorten the sampling period. The period of the sampling is also chosen in a way such that the total number of sampled points is a prime number. (If the set of possible states contains a prime number p of elements, it can be endowed with the algebraic structure of the finite field $\mathbf{F}_p$, which is required by the LS-algorithm). For any given prime number p the corresponding sampled points can be labeled increasingly from 0 to $p-1$. (See Figure 5.1.)

The interpolation conditions for the transition function

$$F : \mathbf{F}_p \to \mathbf{F}_p$$

are very simple in this situation, characterized by the transitions

$$0 \mapsto 1 \mapsto 2 \mapsto ... \mapsto p-1 \mapsto p-1$$

Since no transition is missing, the interpolating function is uniquely determined (see also the Lagrange Interpolation Formula, Theorem 1.71 in [67]). As the reader can easily verify,

$$F(x) = (p-1)\left(\prod_{k=0}^{p-2}((p-1)-k)^{-1}(x-k)\right) + x + 1 \ \forall \ x \in \mathbf{F}_p$$

and obviously

$$\deg(F) = p-1$$

Now, the LS-algorithm must return F, given the fact that there is a unique interpolating function. This shows that, in general, the output of the LS-algorithm does indeed depend on p, no matter how large p is.

In the two variables case, we use the example of a two dimensional flow converging to a point in the plane. We attempt to capture the dynamics using the LS-algorithm. The surface

$$z = h(x,y) := -e^{(-\frac{3}{2}y^2+2yx^2-x^4-\frac{1}{2}x^2)}$$

has a local minimum at $(0,0)$. We chose this surface because it is reasonably simple without being overly symmetric. We modelled flow on this surface using the map $(x,y) \mapsto f(x,y)$, where[5]

$$\begin{aligned} f &: \mathbb{R}^2 \to \mathbb{R}^2 \\ (x,y) &\mapsto (x,y) - \frac{1}{10}\nabla h(x,y) \end{aligned}$$

(See left part of figure 5.2.) To use the LS-algorithm we consider the closed square $[-1,1]\times[-1,1]$ and an equidistant rectangular grid of $p \times p$ squares on it. Now, the evolution of a square in the grid is defined by the square which contains the image under f of the middle point of the square in question (See figure 5.2).

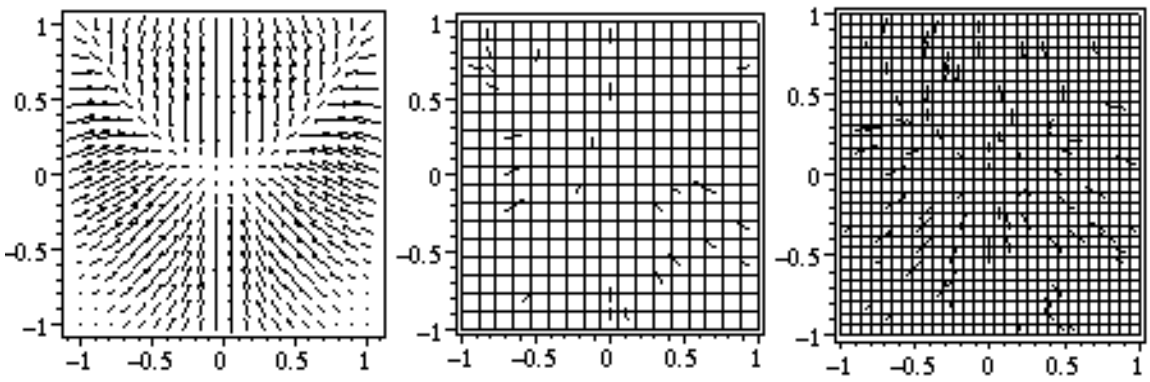

Figure 5.2: Left: Arrow field representing the flow towards the origin. Center: Discretization grid for $p = 17$. Right: Discretization grid for $p = 29$. In both, 10% of the squares display an arrow pointing to the square to which it is mapped by the flow.

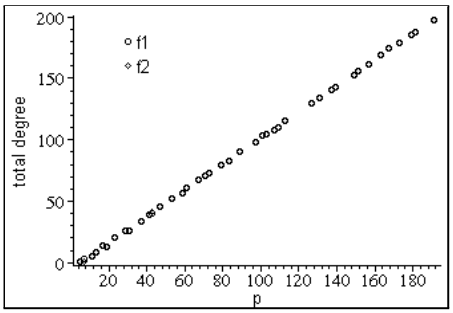

Figure 5.3: Dependency of the output functions' total degree on the cardinality p of the finite field used.

For a given prime number p and the corresponding equidistant grid, we ran the LS-algorithm allowing for a random sampling of 10% of the squares, i.e. the transition of 10% of the squares is used as time series for the algorithm. Iterating this process for an increasing prime number p shows that the total degree of the polynomial functions in two variables generated by the LS-algorithm keeps growing, showing that the terms in the polynomials do not remain the same (see figure 5.3). This behavior is also observed if a lower percentage of sampled squares is used as input time series. The highest prime number p considered ($p = 191$) was dictated by the value at which the actual run time for the LS-algorithm becomes too large for any practical application.
As with any mathematical algorithmic method based on discretization, some type of convergence as the discretization gets finer and finer (i.e. the step size gets smaller) is highly desirable, in the sense that after a certain degree of resolution, the method is capable of catching essential properties which won't vary significantly if the resolution is further increased. Contrary to their claim, the LS-algorithm doesn't generally show convergence at the level of the polynomial functions generated. It might show convergence at the level of the qualitative dynamic properties of the systems generated, but we have not explored this feature, neither was it scrutinized in [65].

[5] ∇h is the gradient of h, therefore
$\nabla h(x,y) = -\left((4(y-x^2)x - x)e^{(-(y-x^2)^2 - \frac{x^2-y^2}{2})}, (-3y + 2x^2)e^{(-(y-x^2)^2 - \frac{x^2-y^2}{2})}\right)$

Part III

The biological backstory of this thesis

Introductory remarks

PathSim is a stochastic agent-based model and computer simulation that attempts to model the interaction between the Epstein-Barr virus (EBV) and the human immune system. As mentioned in the introduction, in order to better analyze and understand the dynamics of the output of PathSim, Dr. Laubenbacher suggested to use the average output of PathSim as data to reverse engineer a deterministic, time discrete dynamical system over a suitable finite field. To this end, he and some of his graduate students developed the reverse engineering method that we described and analyzed in Chapter 5. The results obtained after analyzing the reverse engineering method (see Chapter 5) discouraged the PathSim team from using it. However, in the era of exponentially growing computer power this is of course not the end of the story. Indeed, multiple computer simulations of the model and statistical analysis of their output represent a common and powerful method in scientific research. This type of use of computational tools is not expected to change in the future, although the rigorous analytical study to assess the dynamical properties of the model is certainly more reliable than the results of simulations.

Given that this dissertation was written within the academic framework of the PathSim project, we consider it pertinent to present a brief description of it here. Chapter 6 is devoted to this description. The author wants to emphasize that the biological model, as well as the agent-based model described in Chapter 6, were already developed when he joined the PathSim team.

Moreover, since the author was directly involved in the *parameter space exploration* of PathSim, its resulting biological interpretations and potential biologically relevant insights are also presented. These results and their discussion constitute Chapter 7. This represents a co-contribution of this thesis to both the biological and biomedical sciences. However, the author emphatically acknowledges that the material presented in Chapter 7 is the result of a joint effort within the PathSim team.

The exposition of this material is intentionally brief and we refer the reader to our publications [36] and [104] for further details. For the understanding of Part III of this thesis, some basic knowledge about immunology is required. The interested reader can find an excellent introduction to this fascinating field in [105].

Chapter 6

The agent-based model/simulation PathSim

6.1 A brief description of the biological model of Epstein-Barr virus infection

Epstein-Barr virus (EBV) is a common human pathogen which infects greater than 90% of all people by the time they are adults [96],[111]. It is associated with several important diseases, including cancer. EBV is most commonly transmitted by saliva [48]. After accessing the pharynx, it starts the infection process on the surface of Waldeyer's ring, which consists of the tonsils and the adenoid. Here EBV infects the epithelium and is consequently amplified. It then infects naive B cells in the underlying lymphoid tissue. EBV uses a series of distinct latent gene transcription programs, which mimic a normal B cell response to antigen, to drive the differentiation of the newly infected B cells. During this stage, the infected cells are vulnerable to attack by cytotoxic T cells (CTLs) [57]. The differentiation process takes place within so called germinal centers that are formed inside tiny ellipsoidal structures called follicles. Follicles are numerous and distributed more or less uniformly throughout the tonsils and the adenoid. (For a more detailed anatomical description of Waldeyer's ring, see [2], [89].) Eventually, the latently infected B cells enter the peripheral circulation, the site of viral persistence, as resting memory cells that express no viral proteins [49] and thus are invisible to the immune response. The latently infected memory cells circulate between the periphery and the lymphoid tissue [62]. When they return to Waldeyer's ring, they are occasionally triggered to terminally differentiate into plasma cells. This is the signal for the virus to start the lytic program and virus replication [63]. These lytically infected B cells are again vulnerable to CTL attack [57]. Newly released virions may infect new B cells or be shed into saliva to infect new hosts, but are also the target of neutralizing antibody.

Primary EBV infection in adults and adolescents is usually symptomatic and referred to as infectious mononucleosis (AIM). It is associated with an initial acute phase in which a large fraction (up to 50%) of circulating memory B cells may be latently infected [50]. This induces the broad T lymphocyte immune response characteristic of acute EBV infection. Curiously, primary infection prior to adolescence is usually asymptomatic. In immunocompetent hosts, infection resolves over a period of months into a lifelong persistent phase in which ~ 1 in 10^5 B cells carry the virus [56]. Exactly how persistent infection is sustained is unclear. It is even unclear whether the virus actually establishes a steady state during persistence or continues to decay, albeit at an ever slower rate [50]. A diagrammatic version of the biological model is presented in the left panel of Figure 6.1.

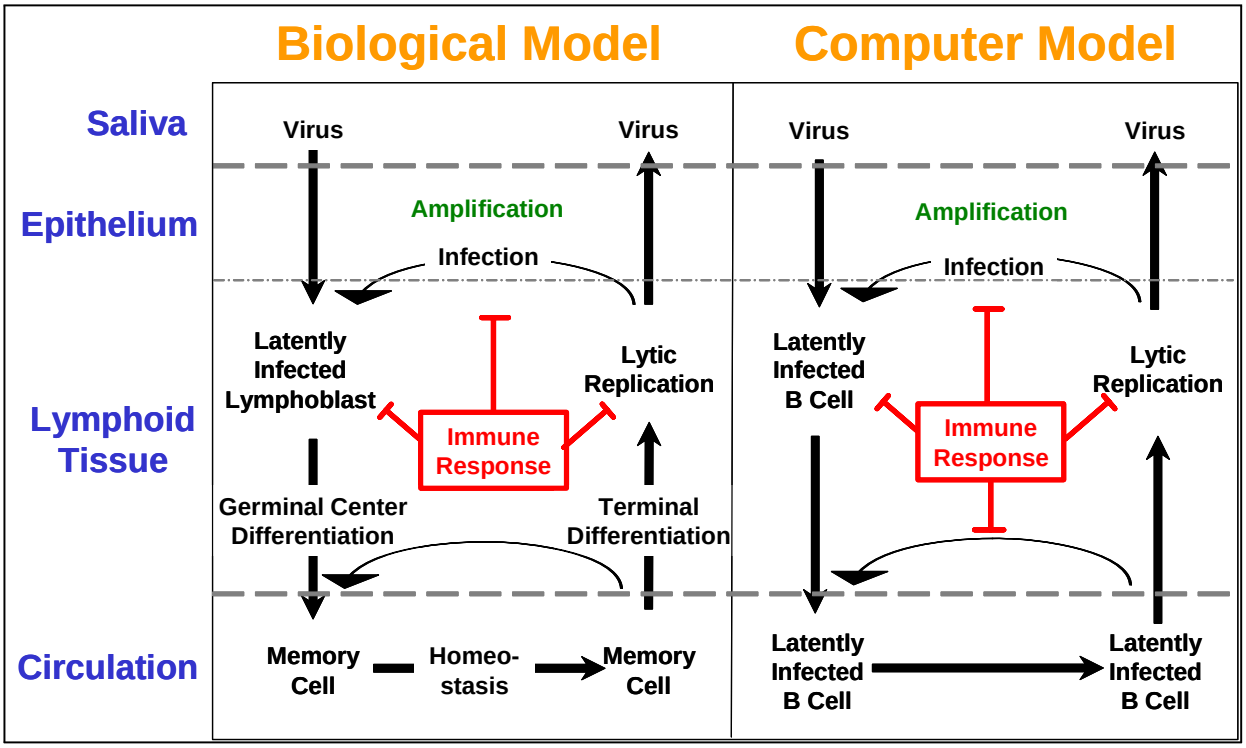

Figure 6.1: Comparison of the EBV Biological Model and the Computational Model. EBV co-opts normal B-cell biology to drive infected B-cells into the resting memory state in the peripheral circulation where they are not subject to immunosurveillance. This process is simplified in the model by omission of the germinal center differentiation. Upon return to Waldeyer's ring, infected memory B-cells may become lytic cells actively producing infectious virus that can either infect new B cells or be shed into saliva to infect new hosts. The immune system attacks latently infected lymphoblasts, lyticaly infected B-cells and free virus. In the model, B_{Lat}s are the target of "immune" response whenever they sojourn in the "lymphoid tissue". Virs and B_{Lyt}s are also engaged. Reproduced from [36] with permission.

6.2 A brief description of the stochastic agent-based model PathSim

The stochastic agent-based model and computer simulation (PathSim) is a representation of the biological model described in the previous section. This model omits many features of the immune system and simplifies those it does contain. A schematic version of both is shown in Figure 6.1.

PathSim consists of a simulation engine together with user friendly interface that allows for two- and three-dimensional data display and analysis. The simulation is performed on a graph that represents the anatomy of Waldeyer's ring together with abstract compartments for blood and lymph. Each vertex of this graph represents a small volume of tissue and is connected by edges to neighboring vertices. Agents can only interact when they are located at the same vertex. See Figure 6.2.

The graph is three dimensional and models the tissue within Waldeyer's ring as well as the geometry of the ring. For ease of construction and manipulation the graph was constructed with a repeating unit. This repeating unit represents one follicle (see above) and the germinal center with adjacent interfollicular tissue contained inside it. The unit is made of concentric ellipsoids (with hexagonal cross-section) covered by hexagonal layers representing the histologic structure. Additionally, in each unit, there is one vertex representing the high endothelial venule and one vertex representing the efferent lymph. These are the connections to the "circulatory" and the "lymphatic" system, respectively. Figure 6.3B provides a three dimensional magnification of this unit compared to a cross-section of the tissue it represents.

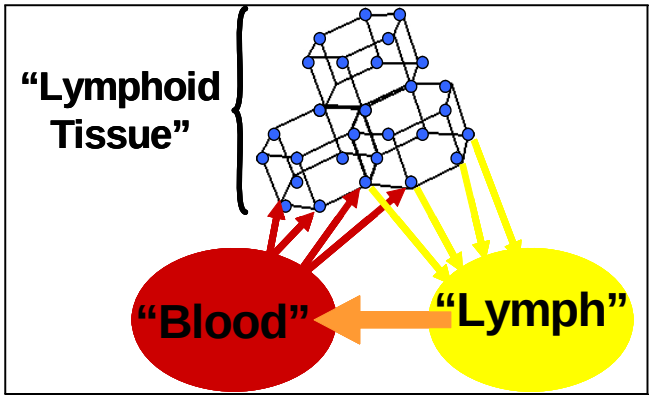

Figure 6.2: Schematic representation of the grid that models the lymphoid tissue and the two virtual compartments of blood and lymph. The arrows represent the possible flow directions for virtual B lymphocites.

Each "tonsil" and the "adenoid" is built on a roughly elliptical floor plan, which is tessellated with the base units representing the follicles (Figure 6.3A, right). The "tonsils" and the "adenoid" are interconnected according to their position in the ring.

The anatomical micro-structure of Waldeyer's ring, i.e. the follicles, is modeled very accurately in PathSim. However, this level of detail seems disproportionate, since the functional characteristics of follicles, namely, antigen presentation, T cell help and the subsequent germinal center maturation are omitted in PathSim.

PathSim's agent types are Vir, B_{Naive}, B_{Lat}, B_{Lyt}, T_{Naive}, T_{Lat} and T_{Lyt}, corresponding to virus, naive B cells, latently infected B cells, lytically infected B cells, naive T cells, and two types of activated cytotoxic T lymphocytes (CTLs), one directed against each kind of infected B cell (CTL latent and CTL lytic). The vertices act as containers for these agents. In the course of a simulation, these agents undergo creation, aging, interaction, motion, and death.

At startup, population of B_{Naive}s and T_{Naive}s are created and used to populate the underlying graph. These "cells" are distributed randomly throughout the entire "ring". Whenever an agent is created, it becomes a stochastic individual lifespan. (The numbers of agents created and their lifespans are governed by parameters which are set at initialization. These and other controlling parameters are discussed in detail in [104].) A virtual infection is started by creating a population of Vir and distributing these only on the surface of the "Waldeyer's ring". The simulation then advances in discrete time steps. At each step, agents age, interact, move to neighboring vertices, and undergo certain life cycle events that may be triggered by aging, interaction, or motion. The population of B_{Naive}s in the blood is also replenished whenever it drops. The time step represents six minutes of real time to accurately reflect interaction and motion rates.

One of the life cycle events induced by aging is death, which happens to Virs and B_{Naive}s at the end of their life-spans. The life cycles of virtual T cells are handled using a simplifying heuristic. T_{Naive}s are immortal. These may become virtual CTLs if they are appropriately triggered (discussed below). Virtual CTLs become again T_{Naive}s at the end of their life-spans.

The life cycle of the B_{Lat} is slightly more complicated. Its possible fates are to die due to passage of time, be killed by a T_{Lat}, or to become a B_{Lyt}. (As described below, they are not subject to "CTL" regulation while in the "blood" compartment.) The biological signal responsible for turning latently infected B cells lytic is unknown [63]. PathSim provides three methods by which B_{Lat}s may become B_{Lyt}s. Given the lack of viremia, all three of these methods are associated with presence in "Waldeyer's ring" and do not operate while the B_{Lat} is in the "blood". One method is based on the passage of time and is the default setting. The other two (which are similar) are based on return from the "blood" to "Waldeyer's ring". These two methods are optional.

When B_{Lyt}s reach the end of their cycle, they die and burst Virs. The number of Virs in this burst is determined stochastically within a pre-set range based on laboratory estimates of burst size. In vivo, virus can also enter the cells of the epithelium and reproduce within them [84]. While this has not been a major focus of the simulation, the PathSim team has allowed for continuing production of Vir in the epithelium for some runs. This turns out to have very little effect on the course of the simulation [36].

The "blood" compartment contains both B_{Naive}s and B_{Lat}s. B_{Naive}s are continually supplied to the "blood" to maintain homeostasis. The infected B cells in the blood compartment in vivo are resting memory B cells [5], [75] that do not express antigenic proteins [49] and thus escape immune surveillance. Accordingly, virtual T cells are excluded from the "blood" compartment. In sum, therefore, aging, "CTL" predation, and initiation of lytic replication for B_{Lat} occur in the "Waldeyer's ring"; they are not allowed to proceed in the "blood" compartment.

Interactions take place between agents located at the same vertex. Only certain pairs of agents interact. Vir and B_{Naive}s can interact, resulting in replacement of both with one or more B_{Lat}s. (More than one new infected B_{Lat} would arise as a consequence of proliferation of the freshly infected B cell post-infection. This proliferation feature is optional and not part of the default run.) Any infected virtual B cell (B_{Lat} or B_{Lyt}) and a virtual naive T cell (T_{Naive}) can interact, thereby converting the T_{Naive} into a virtual CTL (a T_{Lat} or T_{Lyt}, respectively). Virtual CTLs can interact with their cognate infected virtual B cells by killing them. T cell memory, antigen presentation and explicit virus neutralization by antibody are not modeled in PathSim, although the relatively short lifespan assigned to Virs implicitly reflects the action of neutralizing antibody. See [104] for more details concerning lifespans.

Each interaction is stochastically governed by two probabilities: the probability that these two agents encounter each other and the probability that they then interact. The probability that a Vir infects a B_{Lat} is near certainty. On the other hand, there is a rather low probability that a B_{Lat} will activate a T_{Naive}. The encounter probabilities depend on the motion of the agents within the small volume represented by the vertex at which they reside. These probabilities were calculated with the help of Monte Carlo simulations based on Brownian motion and neglecting any chemotaxis effects. See [104] for more details on encounter and interaction probabilities.

Motion takes place between adjacent vertices in the graph and is carried out stochastically. The probabilities depend on tissue locations, agent types, and populations. Some of them are zero, thus preventing certain movements entirely. These probabilities are also used to mimic short-range effects of chemotaxis. In the absence of chemotaxis, a baseline probability governs the likelihood of motion to an adjacent vertex. These probabilities were computed by a similar Monte Carlo simulation to the one used for encounter probabilities.

There is a one-way flow of B_{Naive}s and B_{Lat}s from "lymphatic tissue" to the "efferent lymph" to the "lymphatic system" to the "blood" compartment to "high endothelial venule" (HEV) and back to the "tonsil tissue". In particular, the "blood" and the "lymph" are adjacent to, and therefore, accessible from every section; (see Figure 6.2.) Motion probabilities are used to restrict Vir to the surface and epithelial layer. Virtual T cells may move through "Waldeyer's ring" freely, but are given an incentive to move towards higher concentrations of virtual B cells in a neighboring vertex. Infected virtual B cells may enter the germinal centers, but B_{Naive}s may not. A more detailed description of allowed motions and their corresponding probabilities is given in [104].

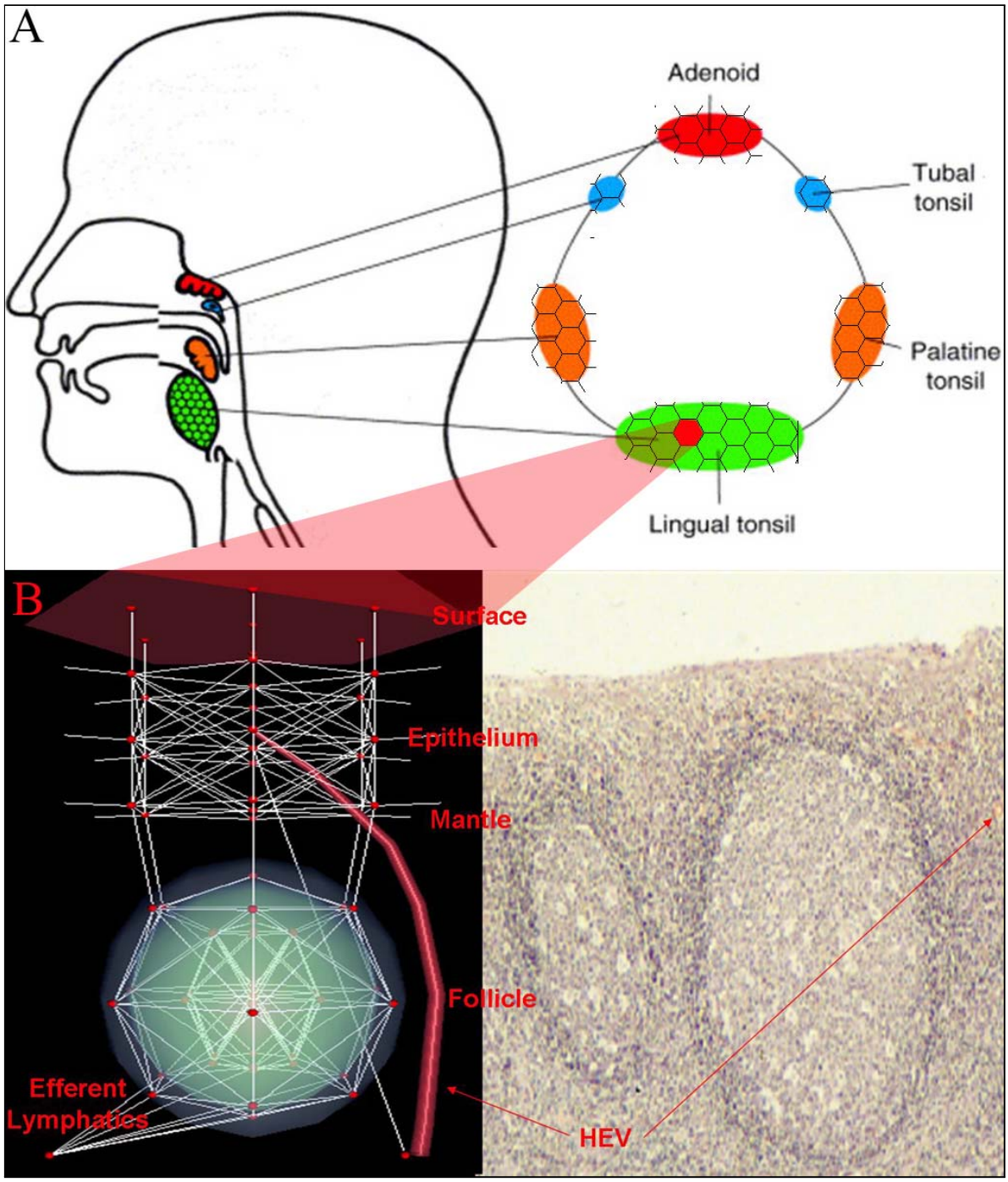

Figure 6.3: Panel A: Left: Anatomical structure of Waldeyer's ring. Right: Geometric representation with hexagonal grid structure of the tonsils and adenoid. Panel B: Left: Three dimensional magnification of the base unit (with hexagonal cross-section) modelling the follicle and germinal center. Right: Histologic cross-section of two follicles and germinal centers. Panel A partially taken from [89]. Panel B reproduced from [104] with permission.

Chapter 7

Parameter space exploration of PathSim and its biological interpretations

7.1 Results of the parameter space exploration

The results presented below were obtained with a default set of parameters (which is described in detail in [104]) or with certain variations of it. Additional biological implications of these outcomes and, in particular, the validation of PathSim are discussed in [36]. Here we exhibit the results of varying either the random seed controlling the stochastic choices during the simulation or values of the parameters themselves. Figures 7.2-7.6 at the end of this section display graphs of times series resulting from running the simulation. Figure 7.2, which examines stability with respect to stochasticity, is the only graph displaying multiple runs for a single parameter set. In all other figures a single line represents a single simulation run.

7.1.1 Stability and overall behavior

In PathSim, stochastic choices are made by referring to the successive values produced by a pseudo-random number generator. We claim that a minimum requirement for a simulation such as PathSim is that it should exhibit stability with respect to stochastic variation in most regimes. This kind of dependence should only show up in the case where the system is finely balanced between two differing outcomes. Thus, when the generic situation is simulated, we expect the overall course of the simulated infection to be independent of initial random seed.

Here we show the results of running PathSim with the default parameter set and twenty different random seeds. Figure 7.2 illustrates the total population for six of the seven agent types in multiple runs. (While the total population of T_{Naive}s is not shown, the pattern is identical to the other agents.) The runs are virtually superimposable for all seeds and show the characteristic biphasic behavior expected of a primary EBV infection with a peak of acute infection followed by long term low level persistence. Clearly, in the regime which we think best represents the course of infection, there are no critical dependencies on stochastic variation. In subsequent runs analyzing sensitivity to parameter variation we assumed that this stability would hold through the range of parameters tested.

7.1.2 Parameter variation and parameter sensitivity

Our sensitivity analysis focused around variations in the input parameters. We varied these parameters individually or in related pairs, (e.g., minimum and maximum viral burst size). We have not explored the parameter space in any systematic way. Our main goal was to bracket

physiologically reasonable parameter values and to extend them in either direction to test their effects on the virtual host-virus system. Our choices were dictated by curiosity about how the viral burst size, the strength and speed of the adaptive immune response, the roles (if any) of initial viral dose or ongoing epithelial re-infection and the rate and manner of lytic activation affect outcomes in EBV infection. We were particularly interested in those factors that contribute to long-term persistence. Here we focus specifically on burst size, viral response, and activation of naive cells by lytically infected B-cells.

Burst size

Burst size is controlled by a pair of parameters giving the minimum and maximum number of Virs produced, with the average burst size found at the mid-point in the range. When a B_{Lyt} bursts, the number of Virs it produces is determined stochastically according to a uniform distribution on the interval between minimum and maximum burst size. As the average burst size increases, peak levels of Virs (Figure 7.3A) and B_{Lat}s (Figure 7.3B) both increase. Above a certain burst size, the persistent phase after the peak is virtually unchanged except for an amplification of stochastic effects. Only at very low burst sizes (8 to 10 Virs per bursting cell) clearance is observed. A burst size of 40 to 60 Virs seems very near the level at which stochastic effects could make the difference between Virs clearance and persistence. In contrast to B_{Lat}s, the peak numbers of B_{Lyt}s decrease with increasing Virs burst size (See Figure 7.3C). This result is consistent with the notion that increasing numbers of B_{Lat}s engender a more aggressive "immune response", thereby shortening their lifespan. Fewer B_{Lat}s live long enough to become B_{Lyt}s.

The default parameter set uses a minimum value of 600 and a maximum value of 1,000.

Proliferation of newly infected B-cells

The biological model described in the previous chapter posits that newly infected naïve B-cells enter the germinal centers of Waldeyer's ring where they differentiate into resting memory B-cells and exit into the peripheral circulation (See Figure 6.1 and [111], [112]). It is not known to what extent they undergo cell division while in the germinal center, but we believe any proliferation must be quite limited. Extensive proliferation would likely be detrimental to the survival of the host. This sort of uncontrolled proliferation is seen in X-linked proliferative disease (XLP), an X chromosome linked predisposition to fatal acute EBV infection [99] and in patients who are immunosuppressed who are susceptible to tumors arising from the unregulated proliferation of EBV infected B cells. That these tumors are rare suggests that uncontrolled proliferation is a rare event, even in the immunosuppressed. (For a more detailed discussion of this issue see [111], [112].) To examine the implications of allowing newly infected B_{Lat} cells to proliferate, we tested three different amplification factors for newly infected "cells" in the simulation (1, 2 or 3 rounds of cell division, resulting in 2, 4 or 8 daughters) before allowing them to exit "Waldeyer's ring" and enter the "blood" compartment. There were some changes in the pattern of resolution of the acute phase (small oscillations, especially at the highest amplification), but qualitatively the overall dynamics were not different from those of the default case where there is no amplification of newly infected "cells". As the degree of proliferation increased, not surprisingly, we observed higher numbers of B_{Lat}s throughout the simulation and a shift in the peak of the B_{Lat}s population to later times, likely resulting from a delay in the ability of "CTLs" to catch up with the more rapidly expanding B_{Lat}s (Figure 7.4A and 7.4B). Not surprisingly perhaps, the overall trend suggested that more extensive proliferation would overwhelm the virtual host consistent with our idea, expressed above, that extensive proliferation could not be tolerated. This is a particularly interesting area to investigate in the context of EBV associated cancers.

As the degree of proliferation of newly infected B_{Lat}s rises, the number of B_{Lyt}s drops (Figure 7.4C). T_{Lyt}s follow an identical pattern of dropping as proliferation increases (Figure 7.4D). Again, this result is consistent with the notion that increasing numbers of B_{Lat}s engender a more

aggressive "immune response", thereby shortening their lifespan. Fewer B_{Lat}s live long enough to become B_{Lyt}s and the lower numbers of B_{Lyt}s lead to the creation of few T_{Lyt}s.

Immune response to infected B-cells

PathSim uses a very simplified model of T-cell activation. When a T_{Naive} encounters either a B_{Lat} or a B_{Lyt} there is a chance that it is activated to become a T_{Lat} or a T_{Lyt} respectively. The probability governing this outcome is the activation rate. Clearly, this process subsumes a large number of actual biological events, all of which have their own kinetics. The default activation rates are 0.015 for T_{Lat} and 0.035 for T_{Lyt}. When a T_{Lat} or a T_{Lyt} encounters its cognate infected "B-cell", there is a probability that this encounter results in killing that infected "cell". We refer to this probability as its kill rate. The activation and kill rates are set separately for latents and lytics. Default values for the kill rate are 0.3 and 0.6, respectively. Varying either the activation rate for T_{Lat}s (Figure 7.5A) or the kill rate for T_{Lat}s (Figure 7.5B) results in a monotonic change in peak infection, with maximal peak infection corresponding to minimal activation or kill rate. At very high levels of either activation or kill, clearance is observed. At the lowest level of each, the populations of virtual infected cells do not appear to go down to low level persistence observed in the default setting. This state of the simulation can be interpreted as long lasting acute illness that normally would cause the patient's death. Figure 7.5 illustrates trends when these rates are varied independently.

In contrast to T_{Lat}s, varying the activation and kill rates for T_{Lyt}s has no appreciable effect on either B_{Lat}s (Figures 7.5C and 7.5D) or B_{Lyt}s (not shown). Perhaps this is because the numbers of B_{Lyt}s are generally so small that few of them are actually killed by T_{Lyt}s before they disappear at burst.

Lytic reactivation of B_{Lat}s

Little is known about the signal which causes latently infected B-cells to exit the memory state and become lytic [63],[11]. In PathSim's default parameter set, this state change occurs due to the passage of time in the latent state, i.e., at the end of its life, there is a 60% probability that a B_{Lat} simply dies but a 40% probability that it will initiate lytic replication of Virs. However, we have experimented with two additional methods of triggering lytic replication of Vir designed to mimic this unknown biological signal and applied them to a small fraction of the returning B_{Lat}s. In the simplest version, returning from the blood to Waldeyer's ring causes a small, user-determined fraction of B_{Lat} cells to automatically turn into B_{Lyt}s. This fraction is termed the alpha parameter. The second version is inspired by Thorley-Lawson's application of Lanzavecchia's ideas about homeostasis of the memory compartment to EBV [63],[11]. In this case, a small, user-defined fraction of returning B_{Lat}s spontaneously divides, each one in this fraction producing one B_{Lat} and one B_{Lyt}. We term this fraction the Lanzavecchia parameter. (In the default run, both the alpha parameter and the Lanzavecchia parameter are set to zero.) When either simulated signal is "received" by as few as 0.15% of returning B_{Lat}s, a significant change emerges in simulated disease progression. Both panels in Figure 7.6 illustrate that at values above 0.050% (red), the number of virtual infected B cells in the low level persistence phase begins to increase steadily. By 0.16% (orange), the initial peak phase does not resolve normally. Increasing this value to as little as 0.26% causes the virtual infected cells to overwhelm the virtual uninfected B-cell population.

We also observe that these two simulated signals produce indistinguishable results (Figure 7.6). The equivalence of these two parameters is not unexpected, given that the critical factor is the sharp rise in B_{Lat} numbers, while the small increase in B_{Lyt}s is negligible. This behavior could provide insight into rare fatal EBV infections.

7.2 Discussion

PathSim is the first simulation of EBV infection that is accurate enough to describe many aspects of the infection. Implemented at the level of cells and lymphoepithelial tissue, it contains sufficient detail to generate new biological insights and allow further investigation of the mechanism of infection [36]. Moreover, it sharpens the understanding of specific issues and suggests new experimental investigations.

The exploration of the parameter space of PathSim helps us understand what features of the simulation are critical for the observed outcome. Consequently, we can identify features that are very robust and features that are actually fine-tuned. If the real biological counterpart of such a feature is also robust (as suggested by the simulation), this robustness could contribute to the homeostasis observed in many biological systems. On the other hand, if the biological counterpart of such a feature is also fine-tuned, the simulation could be unveiling a very powerful intervention mechanism.

The homeostasis observed in biological systems protects them from (bounded) random perturbation, at least in physiological regimes. PathSim is a stochastic simulation and thus also experiences random fluctuation. The exact course of each simulation depends on the values produced by a random number generator. For this to be a usable simulation of living processes, it should also be stable in most regimes. Figure 7.2 demonstrates that this sort of stability is indeed observed in the simulation. However, it is not hard to imagine unstable situations, both in vivo and in silico. For example, when there are very few infected cells, small random fluctuations may mean the difference between persistence and clearance. Such a state lives on or near a stochastic border between differing outcomes. In a deterministic system one can find a well-defined border between the basins of attraction for different outcomes. In the presence of random fluctuations, this border may become a region in which different outcomes are likely and unpredictable. We consider these boundaries interesting from a therapeutic viewpoint since they represent the states in which a small intervention could produce a large change in outcome.

We have identified one parameter whose value has a very dramatic impact on outcome, namely the fraction of B_{Lat}s which are turned lytic immediately upon return from the "blood" to "Waldeyer's ring" (alpha parameter, Figure 7.6). The simulation suggests that the value of this fraction in vivo is a fine-tuned constant. One possible explanation of this fine-tuned value could be of evolutionary nature: Consider what happens if a mutation in the EBV virus arises in which this rate drops. A virion with this mutation produces fewer copies of itself within the infected individual and thereby decreases its opportunities for transmission. This will act to take the mutant gene out of the gene pool. On the other hand, a mutation which raises this rate might result in host death, removing both that EBV genome and the host from further procreation. While it has been suggested that this strategy may be quite effective for an acutely replicating virus that jumps quickly from host-to-host [40], it would be counter productive for a virus like EBV that uses persistence as a way to maximize infectious spread. Thus, overall, there is evolutionary pressure on the virus to prevent this rate from either rising or falling, and there is pressure on the host to keep it from rising. In this way, selection acts as a fine calibration mechanism of this rate.

We would like PathSim to give us insight into therapeutic targets for drug development. Consider a drug which suppresses viral replication. How effective would such a drug need to be to induce clearance? Our investigation of the effects of Virs burst size may provide some guidance about the drug efficacy required for complete clearance. Interpretation of these results, however, depends on information about EBV which we do not yet possess. Figure 7.3B shows persistence at a burst size around 40 and clearance at a burst size around 10. At some value between 40 and 10, we should see runs on the stochastic boundary described above. Let us assume this occurs at a burst size of 25. Now let y be the average number of viral particles that a lytically infected B-cell produces. Not every virus produced is a working virus. In addition, some of the viruses are likely to be neutralized by circulating antibodies. So, from each viral burst, only an average

effective number $z \leq y$ of viruses is viable. The drug's job is to reduce the number of viable virions produced to a maximum of 25 per burst. Therefore, the proportion D_{eff} of the effective viral burst that the drug must actually kill to reduce the value z to 25 obeys the inequality

$$D_{eff} \geq 1 - \frac{25}{z}$$

The plot of this formula (see Figure 7.1) reveals that D_{eff} grows very quickly and that for values of $z \geq 350$, the drug must already be more than 90% effective.

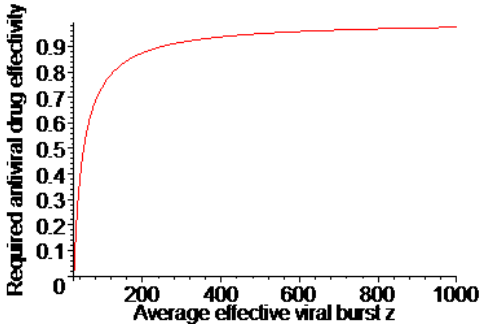

Figure 7.1: Dependency of the required antiviral drug effectivity on the average effective viral burst. It has been estimated that a lytically infected B-cell produces approximately 1,000 viral particles [1]. Thus, the plot range is $25 \leq z \leq 1000$.

These numbers are not conclusive. We have set the burst size at the beginning of each run. In reality, the drug would be administered after onset of symptoms, say, around day 50 (see [36]). It is hard to know whether this fact makes it easier to achieve clearance because the immune system has already mounted a response, or harder because the virus has established itself in the peripheral circulation.

We doubt that PathSim represents an optimal balance of the opposing quests for simplicity and accuracy that any scientific model faces. Nevertheless, as we have argued in [36], PathSim models many clinically observed features. Given PathSim's simplifications, this resemblance is striking and invites the question of what factors are responsible. We believe it is generic features of the rule set that produce the overall dynamics.

At the simplest level, PathSim exhibits a mix of positive and negative feedback between the different agent populations. We have a two-step process (latents and lytics) in which the infected virtual B-cells act to increase the populations of virtual CTLs and the virtual CTLs then act to decrease the populations of infected B-cells. Linked to this process are the mechanisms by which the Vir acts to increase its own population by going through the Vir- B_{Lat} - B_{Lyt} –Vir cycle with amplification at the last step.

We regard this dynamic as a sort of generalized predator-prey system: The agents Vir, B_{Lat} and B_{Lyt} can be seen as three different developmental forms of one prey. The B_{Naive}s can be seen as another prey, that the Virs need to feed on in order to grow into the "adolescent" form of a B_{Lat}. The B_{Lyt}s are the mature individuals capable of producing offspring. The T_{Naive}s are the predators. A T_{Naive} determines its life-long "eating habits" according to its first successful "tasting" of a B_{Lat} or a B_{Lyt}. In this view, PathSim is a spatially distributed, stochastic predator-prey system with the blood compartment acting to conceal part of the prey population. (This analogy is not exact. For example, production of "predators" is not proportional to their population.)

Mathematical study of predator-prey systems goes back to Lotka and Volterra [47]. Their model is a mean field model, that is, it does not take into account spatial distribution and those effects can be profound, as has been demonstrated in systems that are literally composed of predators and prey that modify their behaviors in response to various prevailing conditions [90].

Mobilia et al. [76][72] have studied spatially distributed, stochastic predator-prey systems. This rich area of study provides a general context for understanding some of the features of PathSim.

In summary, we do not believe that PathSim has yet evolved to the point where it reliably produces answers, but it is already quite effective at framing questions. Its major contribution is its ability to generate global insights and motivate further experimentation, while suggesting new avenues for future development.

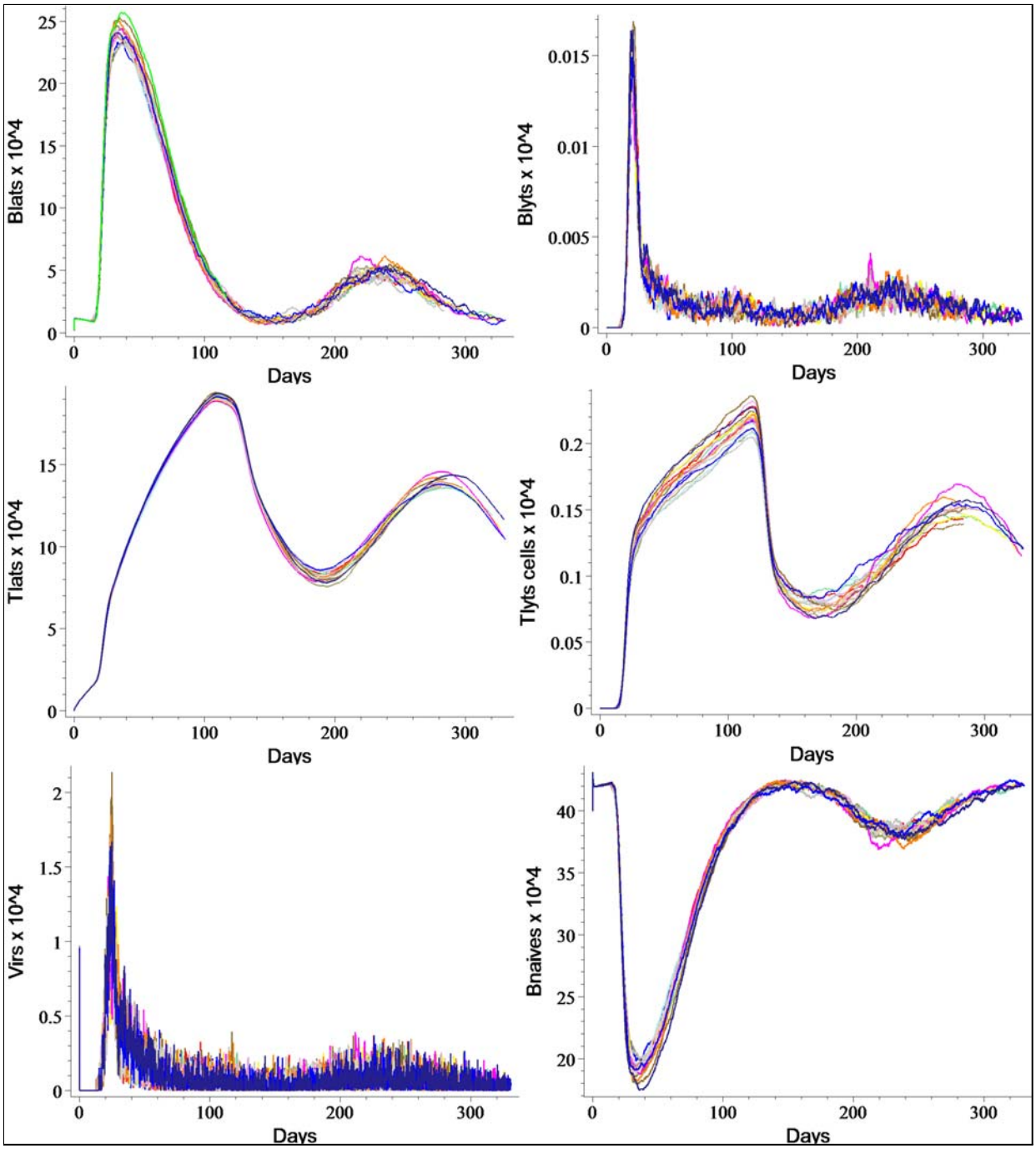

Figure 7.2: PathSim Stability with Respect to Stochastic Variation. Here we illustrate that multiple values of the random seed (here $n = 20$) yield nearly identical results in terms of total agent numbers. Six of the seven agent types are plotted. Virtual Naive T-cells (omitted) exhibit the same behavior. Reproduced from [104] with permission.

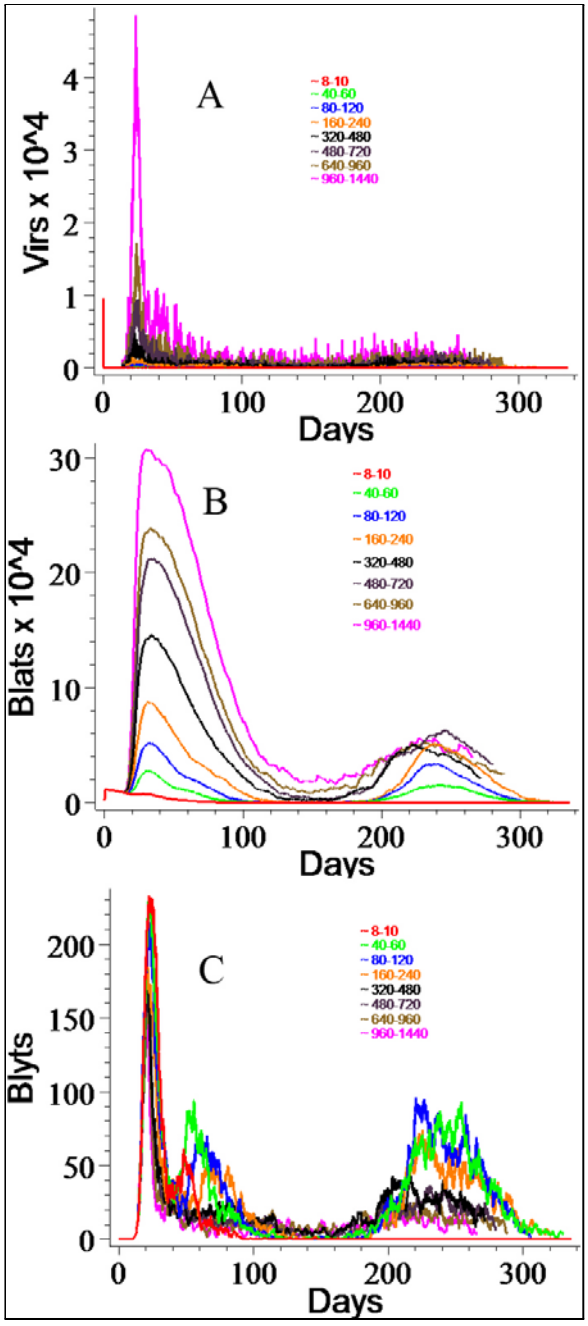

Figure 7.3: Effects of Variation of Viral Burst Size. (A) Peak levels of free Virs rise in response to increases in Virs burst size. Each line represents a single run for each parameter value. The range shown in the legend indicates the minimum and maximum Virs burst size, with the average burst size at the mid-point. The initial Virs dose was the same for all runs. (B) The number of B_{Lat}s rises in response to increasing the Virs burst size. Only with the lowest burst size (8-10 viruses, red) is actual clearance observed. Above about 120 Virs per burst, the low level persistence phase is virtually identical except for random fluctuations. (C) In contrast, B_{Lyt}s numbers drop in response to increases in Virs burst size, likely due to fewer B_{Lat}s surviving to become lytic. Reproduced from [104] with permission.

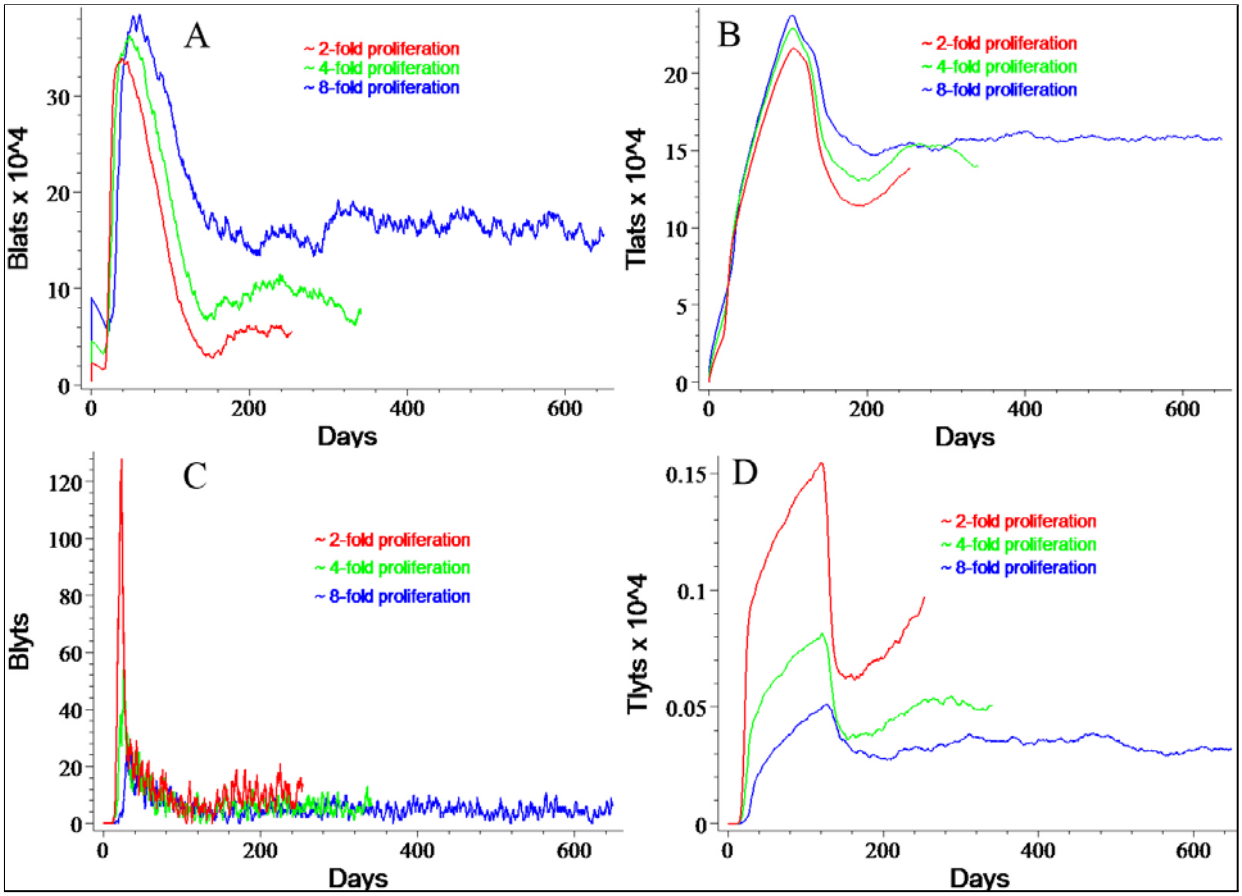

Figure 7.4: Effects of Cell Division after Initial Infection. Panels A and C illustrate the effects of allowing 1 (red), 2 (green), or 3 (blue) rounds of "cell" division immediately following initial infection on the numbers of B_{Lat} and B_{Lyt}, respectively. While B_{Lat} populations grow, B_{Lyt} populations actually shrink. Cognate virtual CTLs show parallel behavior. Reproduced from [104] with permission.

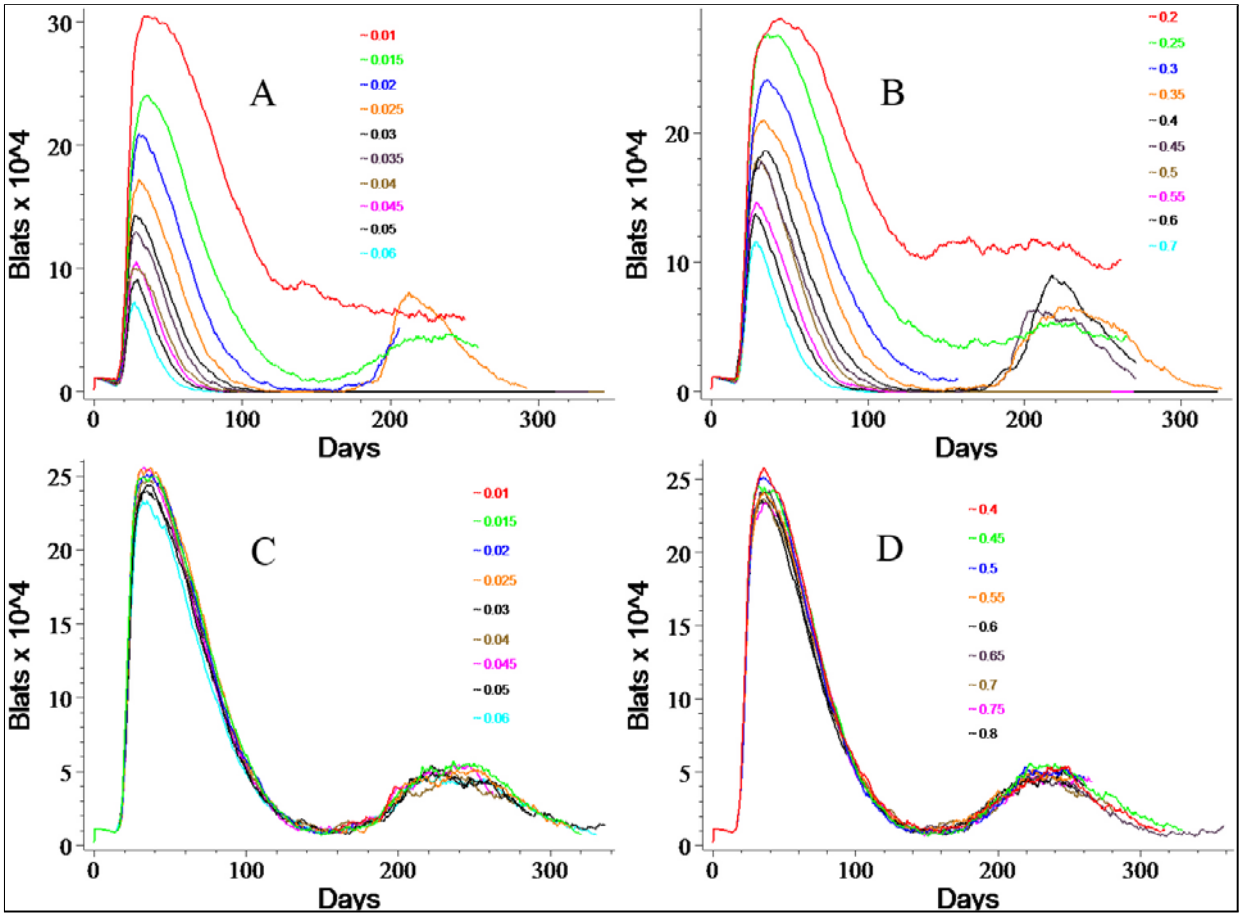

Figure 7.5: Responses of B_{Lat}s to Changes in "CTL" Parameters. Varying T_{Lat} activation rate (A) or kill rate (B) or T_{Lyt} activation rate (C) or kill rate (D) results in the changes illustrated above in B_{Lat} populations. Increasing either the activation or kill rate of T_{Lat}S decreases the B_{Lat}s ultimately resulting in clearance at the highest values of either of the two parameters, while for T_{Lyt}s increasing these two parameters has no significant effect on the B_{Lat} population. The legend indicates the probability of the event (activation or kill, respectively), where 1.0 corresponds to 100%. The default values for activation are T_{Lat}=0.015 and T_{Lyt}=0.035, while those for killing are T_{Lat}=0.30 and T_{Lyt}=0.60. Reproduced from [104] with permission.

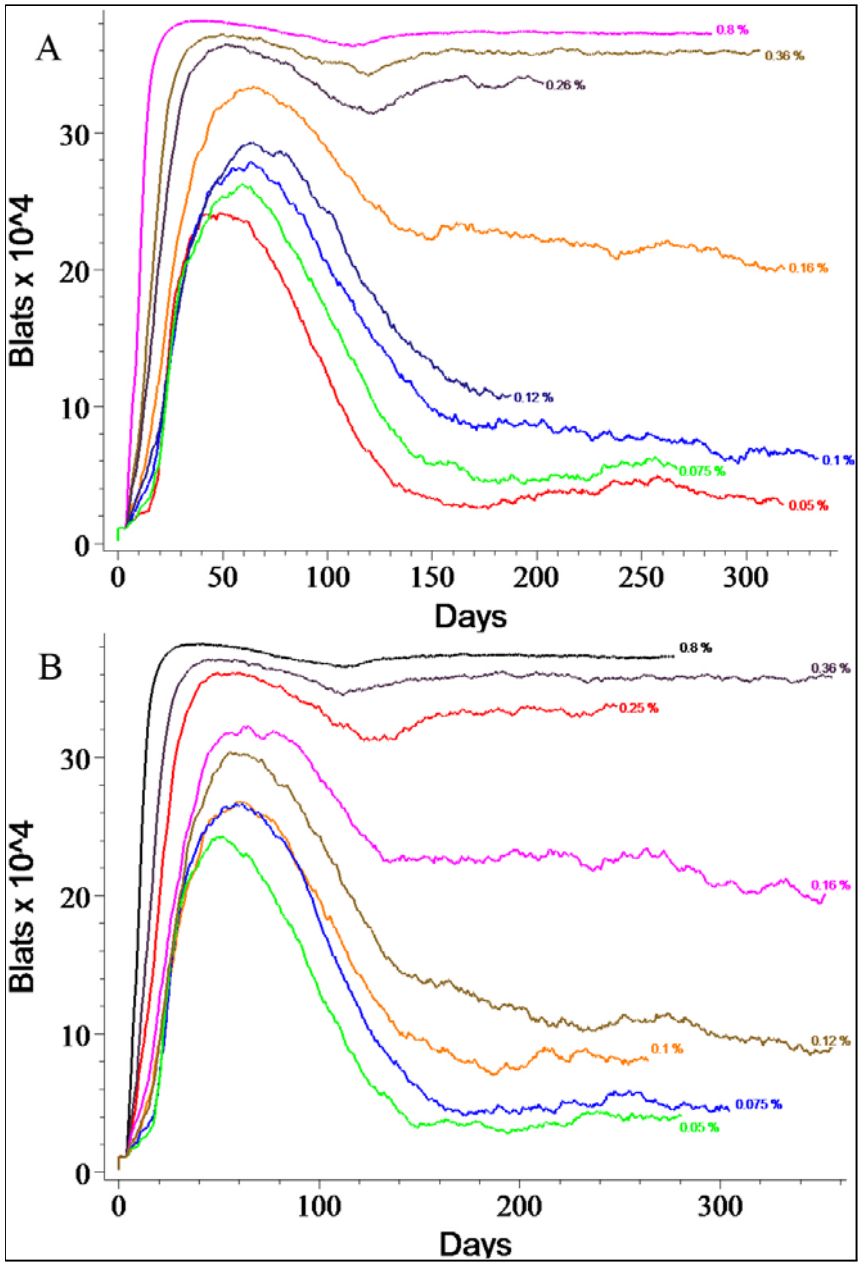

Figure 7.6: Lytic Reactivation Rate Strongly Determines Infection Outcomes. Very small changes in the reactivation rate lead to profoundly different outcomes. Panel A shows the effect of varying the alpha parameter which controls the fraction of B_{Lat}s that commence lytic replication immediately upon return to the tonsils. Panel B depicts the results of changing the Lanzavecchia parameter which determines what fraction of B_{Lat}s divide upon return to the tonsils, thereby producing one B_{Lyt} and one B_{Lat}. These two parameters produce nearly identical results and are the most sensitive in PathSim. Each curve is labeled with the value of the parameter that produced it and represents a single run. Reproduced from [104] with permission.

Appendix A

Appendix for Section 2.2

A.0.1 The Grothendieck group $G((E_q, \oplus))$ of the commutative monoid $(E_q, \oplus)$

The following construction appears in §7 of [64]. The idea is to construct an Abelian group $G(M)$ in the concisest possible way out of a commutative monoid M. According to §7 of [64], the so called group[1] $G((E_q, \oplus))$ of the commutative monoid $(E_q, \oplus)$ is defined as the set of congruence classes of $E_q \times E_q$ with respect to the following equivalence relation

$$(x, y) \sim (x', y') \Leftrightarrow \exists\ t \in E_q : x \oplus y' \oplus t = x' \oplus y \oplus t$$

This set is endowed with the binary operation

$$\overline{(x, y)} + \overline{(\tilde{x}, \tilde{y})} := \overline{(x \oplus \tilde{x}, y \oplus \tilde{y})} \tag{A.1}$$

For the sake of completeness, we provide here the proof that $G((E_q, \oplus))$ is indeed an Abelian group. The proof is written for the specific monoid $G((E_q, \oplus))$, although the arguments of course apply to any commutative monoid.

The operation (A.1) is well defined. To see this, consider $(x', y') \in \overline{(x, y)}$ and $(\tilde{x}', \tilde{y}') \in \overline{(\tilde{x}, \tilde{y})}$. By the definition of $\sim$ we have $\exists\ t \in E_q : x \oplus y' \oplus t = x' \oplus y \oplus t$ and $\exists\ \tilde{t} \in E_q : \tilde{x} \oplus \tilde{y}' \oplus \tilde{t} = \tilde{x}' \oplus \tilde{y} \oplus \tilde{t}$. Consequently,

$$x' \oplus y \oplus t \oplus \tilde{x}' \oplus \tilde{y} \oplus \tilde{t} = x \oplus y' \oplus t \oplus \tilde{x} \oplus \tilde{y}' \oplus \tilde{t}$$

which is equivalent to

$$(x' \oplus \tilde{x}') \oplus (y \oplus \tilde{y}) \oplus (t \oplus \tilde{t}) = (x \oplus \tilde{x}) \oplus (y' \oplus \tilde{y}') \oplus (t \oplus \tilde{t})$$

implying

$$(x' \oplus \tilde{x}', y' \oplus \tilde{y}') \sim (x \oplus \tilde{x}, y \oplus \tilde{y})$$

Then it follows

$$\overline{(x', y')} + \overline{(\tilde{x}', \tilde{y}')} = \overline{(x' \oplus \tilde{x}', y' \oplus \tilde{y}')} = \overline{(x \oplus \tilde{x}, y \oplus \tilde{y})}$$

Moreover, it is easy to see that $\overline{(0, 0)}$ is the neutral element and it holds

$$\overline{(0, 0)} = \overline{(x, x)}\ \forall\ x \in E_q$$

Consequently, the inverse element of any $\overline{(x, y)}$ is $\overline{(y, x)}$. Now we prove particular properties of the group $G((E_q, \oplus))$:

It holds

$$(0, 0) \sim (0, q-1) \text{ and } (0, 0) \sim (q-1, 0) \tag{A.2}$$

[1] Many authors refer to this group $G(M)$ as the Grothendieck group of a commutative monoid M. [64] defines the Grothendieck group using generators and relations within the free abelian group generated by M. It can be shown that both constructions are isomorphic.

To see this, consider any $t \in E_q \backslash \{0\}$. By 4. of Lemma 39 we have

$$red_q(q-1+t) = red_q(t)$$

and by definition

$$(q-1) \oplus t = t$$

which is equivalent to

$$0 \oplus (q-1) \oplus t = 0 \oplus 0 \oplus t$$

and to

$$0 \oplus 0 \oplus t = (q-1) \oplus 0 \oplus t$$

Furthermore, it is simple to verify that

$$\begin{aligned}
(1,0) &\sim (2,1) \sim (3,2) \sim ... \sim (q-1, q-2) \\
(2,0) &\sim (3,1) \sim (4,2) \sim ... \sim (q-1, q-3) \\
(3,0) &\sim (4,1) \sim (5,2) \sim ... \sim (q-1, q-4) \\
&\vdots \\
(q-2,0) &\sim (q-1,1)
\end{aligned}$$

and

$$\begin{aligned}
(0,1) &\sim (1,2) \sim (2,3) \sim ... \sim (q-2, q-1) \\
(0,2) &\sim (1,3) \sim (2,4) \sim ... \sim (q-3, q-1) \\
(0,3) &\sim (1,4) \sim (2,5) \sim ... \sim (q-4, q-1) \\
&\vdots \\
(0,q-2) &\sim (1,q-1)
\end{aligned}$$

as well as

$$\begin{aligned}
(k,0) &\nsim (j,0) \text{ for } k,j \in \{1,...,q-1\} \text{ with } k \neq j \\
(0,k) &\nsim (0,j) \text{ for } k,j \in \{1,...,q-1\} \text{ with } k \neq j
\end{aligned}$$

Now consider a monoid homomorphism

$$h : (E_q, \oplus) \to G((E_q, \oplus))$$

If $h(1) = \overline{(0,0)}$ then h is noninjective, since $h(0) = \overline{(0,0)}$ (by definition). If we assume $h(1) \neq \overline{(0,0)}$, then from the above considerations we can follow $\exists\ k \in \{1,...,q-2\}$ such that

$$h(1) = \overline{(k,0)} \text{ or } h(1) = \overline{(0,k)}$$

Thus, since h is a homomorphism, we have

$$h(x) = \overline{(red_q(xk), 0)} \text{ or } h(x) = \overline{(0, red_q(xk))}\ \forall\ x \in \{2,...,q-1\}$$

Consequently, if $red_q(xk) \neq q-1\ \forall\ x \in \{2,...,q-1\}$, h maps the set $\{1,...,q-1\}$ into a set of cardinality $|\{1,...,q-2\}| = q-2$, in other words, h is not injective. On the other hand, if $red_q(yk) = q-1$ for some $y \in \{2,...,q-1\}$, then by equation (A.2) we have

$$h(y) = \overline{(0,0)}$$

Summarizing, we can state that any monoid homomorphism $h : (E_q, \oplus) \to G((E_q, \oplus))$ must be *noninjective*.

Appendix B

Appendix for Section 2.4

B.0.2 The preprocessing algorithm

We start with the definition of monomial dynamical system according to [23] and [23]:

Definition 237 *Let $\mathbf{F}_q$ be a finite field. A map $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ is called a monomial dynamical system over $\mathbf{F}_q$ if for every $i \in \{1, ..., n\}$ there exists a tuple $(F_{i1}, ..., F_{in}) \in E_q^n$ and an element $a_i \in \{0, 1\} \subseteq \mathbf{F}_q$ such that*

$$f_i(x) = a_i x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n$$

In order to use the algorithm described in Section 2.4 to determine whether such a monomial dynamical system is a fixed point system we need to preprocess the system in the sense of Remark 31. To accomplish this task algorithmically, we add an element $-\infty$ to our exponents semiring $E_{q.}$ (See Definition 29 and Theorem 43.):

$$\overline{E_q} := E_q \cup \{-\infty\}$$

The arithmetic with this new element is as follows

$$\begin{aligned} a \oplus -\infty &= -\infty \oplus a = -\infty \ \forall \ a \in \overline{E_q} \\ a \bullet -\infty &= -\infty \bullet a = -\infty \ \forall \ a \in \overline{E_q} \backslash \{0\} \\ 0 \bullet a &= a \bullet 0 = 0 \ \forall \ a \in \overline{E_q} \end{aligned}$$

The addition is due to the additive "absorption property" of $-\infty$ obviously associative. The same holds for the multiplication, since both 0 and $-\infty$ show the multiplicative "absorption property" (although 0 wins over $-\infty$). With this rules we are already able to multiply pairs of matrices with entries in $\overline{E_q}$. With this extended exponents set we can represent the monomial dynamical systems defined above as follows:

Definition 238 *Let $\mathbf{F}_q$ be a finite field. A map $f : \mathbf{F}_q^n \to \mathbf{F}_q^n$ is called a monomial dynamical system over $\mathbf{F}_q$ if for every $i \in \{1, ..., n\}$ there exists a tuple $(F_{i1}, ..., F_{in}) \in E_q^n$ or a tupel $(F_{i1}, ..., F_{in}) \in \{-\infty\}^n$ such that*

$$f_i(x) = x_1^{F_{i1}} ... x_n^{F_{in}} \ \forall \ x \in \mathbf{F}_q^n$$

Now we describe the preprocessing algorithm: Given a monomial system $f \in MF_n^n(\mathbf{F}_q)$ and its representing matrix $F \in M(n \times n; \overline{E_q})$

1. Initialize $L_1 := 0$ and $L_2 := 0$ and an array v of length n to zero.
2. For k from 1 to n do $L_2 := L_2 + 1$ and $v[k] := 1$ if and only if $F_{k1} = -\infty$.

3. Compare L_1 and L_2. If $L_1 = L_2$ or $L_2 = n$, construct the matrix

$$F' \in M((n - L_2) \times (n - L_2); E_q)$$

by deleting the kth row and the kth column of F for all k s.t. $v[k] = 1$. Then return F' and stop. If $L_1 < L_2$ and $L_2 < n$, calculate the product $F^{\cdot 2}$, set $F := F^{\cdot 2}$ as well as $L_1 := L_2$ and go to step 2.

4. If the returned matrix F' is the empty matrix ($L_2 = n$) we can conclude that the system f is a fixed point system with $(0, ..., 0)^t \in \mathbf{F}_q^n$ as its unique fixed point (see Remark 31). If F' is not the empty matrix, the corresponding lower dimensional system $f' := \Psi(F')$ needs to be analyzed with the algorithm described in Section 2.4.

Step 1 implies $n + 2$ initializations. Step 2 of the algorithm requires n comparisons, at most n additions and at most n assignments. There are 2 comparisons in step 3. Each matrix multiplication in step 3 takes $2n^3 - n^2$ addition or multiplication operations[1] in $\overline{E_q}$. There is one initialization after each matrix multiplication. The worst case scenario is given when every time the algorithm performs step 3, the set L_1 grows by one element, forcing the algorithm to perform $n-1$ matrix multiplications. The construction of the matrix F' requires a number of comparisons and assignments that is obviously bounded above by $2n^2$. Summarizing, the worst case complexity of the algorithm is bounded above by

$$B(n) := (n + 2) + n(3n) + n2 + (n - 1)(2n^3 - n^2 + 1) + 2n^2 = 2n^4 - 3n^3 + 6n^2 + 4n + 1$$

Since

$$\lim_{n\to\infty} \frac{B(n)}{n^4} = 2$$

we can conclude $B(n) \in O(\ n^4)$.
It is pertinent to emphasize that this preprocessing algorithm represents a primitive first attempt. Since the matrix multiplications dominate the complexity of the algorithm, it seems meaningful to try to reduce the complexity of the multiplication. Indeed, the rows with entries $-\infty$ are preserved during the multiplication, i.e. those rows do not need to be calculated. In addition, if the first element of a row in the product matrix is equal to $-\infty$ we know that all the remaining elements of that row are going to be equal to $-\infty$ as well. As we can see, there are possibilities of improvement. However, for the purposes of this thesis, we are satisfied with a first working algorithm of polynomial complexity.

[1] See also the analysis of the arithmetic operations in the semiring E_q in Section 2.4.

Appendix C

Appendix for Section 4.5

Lemma 239 *Let K be a field, $n \in \mathbb{N}$ a natural number, $K[\tau_1, ..., \tau_n]$ the polynomial ring in n indeterminates over K and $>$ an arbitrary monomial order. Then for each natural number $m \in \mathbb{N}$ and each $i \in \{1, ..., n\}$ it holds*

$$\tau_i^m > \tau_i^{m-1} > ... > \tau_i > \tau_i^0 \tag{C.1}$$

Proof. For each possible monomial order $>$ and for each $i \in \{1, ..., n\}$ it holds

$$\tau_i > \tau_i^0 = 1$$

Now, applying the translation invariance of the order $>$ we get that for each natural number $m \in \mathbb{N}$

$$\tau_i^m > \tau_i^{m-1}$$

Therefore using the transitivity of $>$ we conclude

$$\tau_i^m > \tau_i^{m-1} > ... > \tau_i \ \forall \ m \in \mathbb{N}, \ i \in \{1, ..., n\}$$

■

Theorem 240 *Let $\mathbf{F}_q$ be a finite field and $n \in \mathbb{N}$ a natural number. Then the family of polynomials*

$$(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n)$$

is a basis for the vanishing ideal

$$I(\mathbf{F}_q^n) \subseteq \mathbf{F}_q[\tau_1, ...\tau_n]$$

Proof. The inclusion

$$\langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n \rangle \subseteq I(\mathbf{F}_q^n) \tag{C.2}$$

is given by the fact that in the finite field $\mathbf{F}_q$ we always have

$$a^q = a \ \forall \ a \in \mathbf{F}_q$$

Now let $f \in I(\mathbf{F}_q^n)$ be a polynomial in the vanishing ideal of $\mathbf{F}_q^n$ and $>$ any monomial order. From the inequalities C.1 it follows for $>$

$$LT(\tau_i^q - \tau_i) = \tau_i^q \ \forall \ i \in \{1, ..., n\}$$

After division of f by $(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n)$ we can write f as

$$f = \sum_{i=1}^{n} h_i(\tau_i^q - \tau_i) + r$$

where $h_i \in \mathbf{F}_q[\tau_1, ...\tau_n]$, $i = 1, ..., n$ and $r \in \mathbf{F}_q[\tau_1, ...\tau_n]$ is the remainder. Assume $r \neq 0$. We know that no term in r is divisible by $(\tau_1^q, \tau_2^q, ..., \tau_n^q)$. As a consequence, each term of r must be of the form

$$a\tau_1^{\alpha_1}...\tau_n^{\alpha_n} \text{ with } \alpha_j < q \ \forall \ j \in \{1, ..., n\} \text{ and } a \in \mathbf{F}_q$$

and we can write r as

$$r = \sum_{\alpha \in M_q^n} a_\alpha \tau_1^{\alpha_1}...\tau_n^{\alpha_n}$$

with suitable $a_\alpha \in \mathbf{F}_q$, $\alpha \in M_q^n$. Now, since $f \in I(\mathbf{F}_q^n)$ and $r = f - \sum_{i=1}^n h_i(\tau_i^q - \tau_i)$ we have

$$r \in I(\mathbf{F}_q^n)$$

This means the polynomial function

$$\begin{aligned} \widetilde{r} \quad &: \quad \mathbf{F}_q^n \to \mathbf{F}_q \\ \vec{x} \quad &\mapsto \quad \widetilde{r}(\vec{x}) := \sum_{\alpha \in M_q^n} a_\alpha \overrightarrow{x}^\alpha \end{aligned}$$

vanishes on all points of $\mathbf{F}_q^n$. Since the fundamental monomial functions $(g_{nq\alpha})_{\alpha \in M_q^n}$ are linearly independent (see Lemma 26), it follows

$$a_\alpha = 0 \ \forall \ \alpha \in M_q^n$$

Therefore, the polynomial r must be zero and we have

$$f = \sum_{i=1}^n h_i(\tau_i^q - \tau_i)$$

As a consequence

$$I(\mathbf{F}_q^n) \subseteq \langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n \rangle$$

This result together with the inclusion (C.2) proofs the theorem. ■

Theorem 241 *Let $\mathbf{F}_q$ be a finite field and $n \in \mathbb{N}$ a natural number. Then the family of polynomials*

$$(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n)$$

is a universal Gröbner basis for the vanishing ideal

$$I(\mathbf{F}_q^n) \subseteq \mathbf{F}_q[\tau_1, ...\tau_n]$$

Proof. It follows from the inequalities C.1 for all possible monomial orders

$$LM(\tau_i^q - \tau_i) = \tau_i^q \ \forall \ i \in \{1, ..., n\}$$

As a consequence, for the least common multiple (lcm) of $LM(\tau_j^q - \tau_j)$ and $LM(\tau_i^q - \tau_i)$, $i \neq j$ it holds

$$\text{lcm}(LM(\tau_j^q - \tau_j), LM(\tau_i^q - \tau_i)) = \text{lcm}(\tau_j^q, \tau_i^q) = \tau_j^q \tau_i^q \ \forall \ i, j \in \{1, ..., n\} \text{ with } i \neq j$$

and for the S-polynomial of $\tau_j^q - \tau_j$ and $\tau_i^q - \tau_i$, $i \neq j$ we have

$$S(\tau_j^q - \tau_j, \tau_i^q - \tau_i) = \tau_i^q(\tau_j^q - \tau_j) - \tau_j^q(\tau_i^q - \tau_i) = \tau_j^q \tau_i - \tau_i^q \tau_j \ \forall \ i, j \in \{1, ..., n\} \text{ with } i \neq j$$

Let's divide $S(\tau_j^q - \tau_j, \tau_i^q - \tau_i) = \tau_j^q \tau_i - \tau_i^q \tau_j$ by $(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n)$. Let without loss of generality

$$\tau_j^q \tau_i > \tau_i^q \tau_j \Leftrightarrow LT(\tau_j^q \tau_i - \tau_i^q \tau_j) = \tau_j^q \tau_i$$

Then we get after the first division step the remainder

$$-\tau_i^q \tau_j + \tau_i \tau_j$$

Now we know from the inequalities (C.1) after translation by τ_j

$$\tau_i^q \tau_j > \tau_i \tau_j \Rightarrow LT(-\tau_i^q \tau_j + \tau_i \tau_j) = -\tau_i^q \tau_j$$

so we can continue the division process and we get the remainder

$$-\tau_i^q \tau_j + \tau_i \tau_j - (-\tau_j)(\tau_i^q - \tau_i) = 0$$

By the previous theorem

$$I(\mathbf{F}_q^n) = \langle \tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n \rangle$$

And so, by Buchberger's S-pair criterion (see Theorem 6 of chapter 2, §6 in [25]),

$$(\tau_1^q - \tau_1, \tau_2^q - \tau_2, ..., \tau_n^q - \tau_n)$$

is a universal Gröbner Basis for $I(\mathbf{F}_q^n)$. ■

Appendix D

Appendix for Sections 5.2 and 5.3

D.0.3 Examples of vector spaces in general position and the codimension condition

Example 242 *Let $n = 2$, $q = 2$ and consider the vector space $F_2(\mathbf{F}_2)$ an its basis $(g_{22\alpha})_{\alpha \in M_2^2} = (x_1x_2, x_1, x_2, 1)$ ordered according to the lexicographic order with $x_1 > x_2$. Furthermore let $U := span(x_1x_2 + x_1 + x_2 + 1)$. The basis vector $u_1 := x_1x_2 + x_1 + x_2 + 1$ has the coordinates $(1,1,1,1)^t$ with respect to the basis $(g_{22\alpha})_{\alpha \in M_2^2}$. Therefore, U is in general position with respect to $(g_{22\alpha})_{\alpha \in M_2^2}$. It is easy to verify*

$$\begin{aligned} \left|V(\langle \varphi^{-1}(u_1) \rangle)\right| &= \left|\{(x,y) \in \mathbf{F}_2^2 \mid xy + x + y + 1 = 0 \mod 2\}\right| \\ &= |\{(0,1),(1,0),(1,1)\}| = 3 \\ &= 2^2 - 1 = codim(U) \end{aligned}$$

As a consequence, the set $X := \{(0,1),(1,0),(1,1)\}$ constitutes an optimal data set to reverse engineer any function $f \in F_2(\mathbf{F}_2)$ displaying no more than 3 terms. If the term-order-free reverse engineering method is used, the probability of successfully retrieving a nonzero function displaying 1 term would be

$$P = \frac{\binom{2^2-1}{2^2-3}}{\binom{2^2}{3}} = \frac{\binom{3}{1}}{\binom{4}{3}} = \frac{3}{4} = 0.75$$

For a function displaying 2 terms $P = 0.5$ and 3 terms $P = 0.25$.

Example 243 *Let $n = 2$, $q = 3$ and consider the vector space $F_2(\mathbf{F}_3)$ an its basis $(g_{23\alpha})_{\alpha \in M_3^2} = (x_1^2x_2^2, x_1^2x_2, x_1x_2^2, x_1^2, x_1x_2, x_2^2, x_1, x_2, 1)$ ordered according to a total degree term order with $x_1 > x_2$. Furthermore let U be the 8-dimensional subspace of $F_2(\mathbf{F}_3)$ generated by*

$$U := span(x_1^2x_2^2 + x_1^2x_2, x_1^2x_2 + x_1x_2^2, x_1x_2^2 + x_1^2, x_1^2 + x_1x_2, x_1x_2 + x_2^2, x_2^2 + x_1, x_1 + x_2, x_2 + 1)$$

The coordinate vectors of the generating vectors are

$$\begin{aligned} \widehat{u}_1 &:= (1,1,0,...,0)^t \\ \widehat{u}_2 &:= (0,1,1,0,...,0)^t \\ &\vdots \\ \widehat{u}_8 &:= (0,...,0,1,1)^t \end{aligned}$$

By calculating the determinant of the matrices

$$A_j := \begin{pmatrix} \widehat{u}_1^t \\ \widehat{u}_2^t \\ \vdots \\ \widehat{u}_8^t \\ e_j^t \end{pmatrix}, \ j = 1, ..., 9$$

(where e_j is the jth canonical unit vector of $\mathbf{F}_3^9$), one can easily show that U is in general position with respect to $(g_{23\alpha})_{\alpha \in M_3^2}$. To determine the set $V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_8) \rangle)$, we start solving the three last equations given by

$$\begin{array}{ccc} x_2^2 + x_1 = 0 & & x_2^2 = -1 \\ x_1 + x_2 = 0 & \Leftrightarrow & x_1 = 1 \\ x_2 + 1 = 0 & & x_2 = -1 \end{array}$$

This system of equations has no solution in the set $\mathbf{F}_3^2$. Therefore

$$V(\langle \varphi^{-1}(u_1), ..., \varphi^{-1}(u_8) \rangle) = \emptyset$$

Consequently, U does not satisfy the codimension condition and thus does not yield an optimal data set.

D.0.4 Existence of vector subspaces in general position

The basic idea of the proof is to treat the problem over the real numbers and then construct a solution over finite fields based on the existence of a solution over the real numbers. This last step takes advantage of the density of the rational numbers in the set of real numbers.

We recall the definition of general position for vector spaces over a finite field:

Definition 244 *Let W be a finite dimensional vector space over a finite field $\mathbf{F}_q$ with $\dim(W) = d > 0$. Furthermore, let $(w_1, ..., w_d)$ be a fixed basis of W and $s \in \mathbb{N}$ a natural number with $s < d$. A vector subspace $U \subset W$ with $\dim(U) = s$ is said to be in* general position *with respect to the basis $(w_1, ..., w_d)$ if for any basis $(v_1, ..., v_s)$ of U and any injective mapping*

$$\pi : \{1, ..., (d-s)\} \to \{1, ..., d\}$$

the vectors

$$v_1, ..., v_s, w_{\pi(1)}, ..., w_{\pi(d-s)} \tag{D.1}$$

are linearly independent.

It can be easily shown that if the linear independence condition (D.1) holds for one basis of U, it holds for every other basis of U.

Now we will construct an s-dimensional subspace $U \subset W$ in general position with respect to a given basis of W, where s is an arbitrary natural number with $s < d$. For this purpose we will find the coordinates with respect to the basis $(w_1, ..., w_d)$ of a basis of U. We denote the coordinates sought as follows

$$\vec{\xi}_1 = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}, \vec{\xi}_2 = \begin{pmatrix} x_{d+1} \\ \vdots \\ x_{2d} \end{pmatrix}, \cdots, \vec{\xi}_s = \begin{pmatrix} x_{(s-1)d+1} \\ \vdots \\ x_{sd} \end{pmatrix}$$

The next step is to count all different injective mappings $\pi : \{1, ..., (d-s)\} \to \{1, ..., d\}$ as $\pi_1, ..., \pi_N$. For each π_i we consider the coordinate vectors $\vec{\xi}_1, ..., \vec{\xi}_s, \vec{e}_{\pi_i(1)}, ..., \vec{e}_{\pi_i(d-s)}$ with respect

to the basis $(w_1, ..., w_d)$, where $\vec{e}_j$ is the jth canonical unit vector of $\mathbf{F}_q^d$. Now, for $i = 1, ..., N$ we define the determinant functions

$$\begin{aligned} D_{\pi_i} &: \mathbb{R}^{sd} \to \mathbb{R} \\ \vec{x} &\mapsto \left| \vec{\xi}_1, ..., \vec{\xi}_s, \vec{e}_{\pi_i(1)}, ..., \vec{e}_{\pi_i(d-s)} \right| \end{aligned}$$

where $\vec{e}_j$ is seen as the jth canonical unit vector of $\mathbb{R}^d$. The linear independence condition (D.1) is equivalent to

$$D_{\pi_i}(\vec{x}) \neq 0$$

Due to the structure of $\left(\vec{\xi}_1, ..., \vec{\xi}_s, \vec{e}_{\pi_i(1)}, ..., \vec{e}_{\pi_i(d-s)} \right)$ and by the Leibniz determinant formula we know that D_{π_i} are nonzero polynomial functions in the variables $x_1, ..., x_{sd}$ and therefore nonzero analytic functions in $\mathbb{R}^{sd}$ with an infinite radius of convergence, (in particular, continuous functions). Consequently, no D_{π_i} can be identical to zero on any open subset of $\mathbb{R}^{sd}$. By the continuity of D_{π_1} we know that there is a non-empty open subset $O_1 \subseteq \mathbb{R}^{sd}$ such that $D_{\pi_1}|_{O_1} \neq 0$. Using the same argument we know that there is a non-empty set $O_2 \subseteq O_1$ open in $\mathbb{R}^{sd}$ such that $D_{\pi_2}|_{O_2} \neq 0$. After applying this argument N times we identify a non-empty open subset $O_N \subseteq \mathbb{R}^{sd}$ such that $D_{\pi_i}|_{O_N} \neq 0 \; \forall \; i \in \{1, ..., N\}$. Since the set $\mathbb{Q}^{sd}$ is a dense subset of $\mathbb{R}^{sd}$, there is a point $\vec{y} \in O_N$ with rational entries, i.e. $y_l \in \mathbb{Q} \; \forall \; l \in \{1, ..., sd\}$. Let

$$\vec{y} = (\frac{a_1}{b_1}, ..., \frac{a_{sd}}{b_{sd}})^t$$

and $c := \prod_{k=1}^{sd} b_k$. Since $\vec{y} \in O_N$, we know $D_{\pi_i}(\vec{y}) \neq 0 \; \forall \; i \in \{1, ..., N\}$. By the rules of determinants we also know

$$D_{\pi_i}(c\vec{y}) \neq 0 \; \forall \; i \in \{1, ..., N\}$$

Moreover, $c\vec{y}$ has integer entries, i.e. $cy_l \in \mathbb{Z} \; \forall \; l \in \{1, ..., sd\}$. For a sufficiently large prime number p, the entries cy_l can be seen as elements of the finite field $\mathbf{F}_p$ of integers modulo p. Therefore, the values $cy_l \in \mathbf{F}_p$, $l = 1, ..., sd$ can be used as the coordinates with respect to the basis $(w_1, ..., w_d)$ of a basis for an s-dimensional subspace $U \subset W$ in general position with respect to the basis $(w_1, ..., w_d)$ of W, a vector space over the finite field $\mathbf{F}_p$. ■

D.0.5 The term-order-free reverse engineering algorithm

The input of the term-order-free reverse engineering algorithm is a set $X \subseteq \mathbf{F}_q^n$ of $m \leq q^n$ different data points, a list of m interpolation conditions

$$\vec{x}_i \mapsto b_i, \; x_i \in X$$

and a linear order $>$ for the elements of the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of $F_n(\mathbf{F}_q)$, (i.e. the elements of the basis are ordered decreasingly according to $>$). The steps of the algorithm are as follows:

1. Calculate the entries of the matrix

   $$A := (\Phi_{\vec{X}}(g_{nq\alpha}))_{\alpha \in M_q^n} \in M(m \times q^n; \mathbf{F}_q)$$

   representing the evaluation epimorphism $\Phi_{\vec{X}}$ of the tuple $\vec{X}$ with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of $F_n(\mathbf{F}_q)$ and the canonical basis of $\mathbf{F}_q^m$.

2. Calculate a basis $\vec{y}_1, ..., \vec{y}_s \in \mathbf{F}_q^d$ of $\ker(A)$.

3. Extend the basis $\vec{y}_1, ..., \vec{y}_s$ of $\ker(A)$ to a basis $(\vec{y}_1, ..., \vec{y}_s, \vec{y}_{s+1}, ..., \vec{y}_d)$ of $\mathbf{F}_q^d$ using the standard orthonormalization procedure. (See Section 4.4 in Chapter 4).

4. Define a generalized inner product $\langle .,. \rangle : \mathbf{F}_q^d \to \mathbf{F}_q$ by setting

$$\langle \vec{y}_i, \vec{y}_j \rangle := \delta_{ij} \ \forall \ i,j \in \{1,...,d\}$$

and calculate the entries of the matrix S defined by

$$S_{ij} := \langle \vec{e}_i, \vec{e}_j \rangle \ , \ i,j \in \{1,...,q^n\}$$

where $\vec{e}_j$ is the jth canonical unit vector of $\mathbf{F}_q^d$.

5. The coordinate vector with respect to the basis $(g_{nq\alpha})_{\alpha \in M_q^n}$ of the output function is obtained by solving the following system of inhomogeneous linear equations

$$\begin{aligned} A\vec{z} &= \vec{b} \\ \vec{y}_i^t S\vec{z} &= 0, \ i = 1,...,s \end{aligned}$$

The steps described above represent an intelligible description of the algorithm and are not optimized for an actual computational implementation.

Bibliography

[1] Epstein-barr virus. In D.V. Ablashi and V. Dharam, editors, *Fourth International Symposium on Epstein-Barr Virus and Associated Malignant Diseases*, Experimental Biology and Medicine, page 455, Hualien, Taiwan, 1990. Humana Press.

[2] Kazuya Abbey and Isuzu Kawabata. Computerized three-dimensional reconstruction of the crypt system of the palatine tonsil. *Acta Otolaryngol*, 454 (Suppl):39–42, 1988.

[3] G. An. Agent-based computer simulation and sirs: Building a bridge between basic science and clinical trials. *Shock*, 16(4):266–273, 2001.

[4] G. An. In silico experiments of existing and hypothetical cytokine-directed clinical trials using agent-based modeling. *Critical Care Medicine*, 32(10):2050–2060, 2004.

[5] G. J. Babcock, L. L. Decker, M. Volk, and D. A. Thorley-Lawson. Ebv persistence in memory b cells in vivo. *Immunity*, 9(3):395–404, 1998.

[6] F. Barbieri and G. Facchinetti. Osservazioni sopra alcune definizioni di pseudo-prodotti interni. *Atti Sem. Mat. Fis. Univ. Modena*, 22:48–59 (1974), 1973.

[7] C. Beauchemin. Probing the effects of the well-mixed assumption on viral infection dynamics. *Journal of Theoretical Biology*, 242(2):464–477, 2006.

[8] T. Becker and V. Weispfenning. *Gröbner bases*, volume 141 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993. A computational approach to commutative algebra, In cooperation with Heinz Kredel.

[9] M. Bernaschi and F.Castiglione. Design and implementation of an immune system simulator. *Computers in Biology and Medicine*, 31(5):303–331, 2001.

[10] M. Bernaschi, S. Succi, and F. Castiglione. Large-scale cellular automata simulations of the immune system response. *Physical Review E*, 61(2):1851–1854, 2000.

[11] N. L. Bernasconi, E. Traggiai, and A. Lanzavecchia. Maintenance of serological memory by polyclonal activation of human memory b cells. *Science*, 298:2199–2202, 2002.

[12] M. Bezzi, F. Celada, S. Ruffo, and P. E. Seiden. The transition between immune and disease states in a cellular automaton model of clonal immune response. *Physica*, 245:145–163, 1997.

[13] Sebastian Bonhoeffer, Robert M. May, George M. Shaw, and Martin A. Nowak. Virus dynamics and drug therapy. *Proceedings of the National Academy of Science*, 94:6971–6976, 1997.

[14] B. Buchberger. Ein algorithmisches Kriterium für die Lösbarkeit eines algebraischen Gleichungssystems. *Aequationes Math.*, 4:374–383, 1970.

[15] B. Buchberger. A theoretical basis for the reduction of polynomials to canonical forms. *ACM SIGSAM Bull.*, 10(3):19–29, 1976.

[16] F. Castiglione, K.A. Duca, J.S. Jarrah, R. Laubenbacher, D. Hochberg, and D.A. Thorley-Lawson. Simulating epstein-barr virus infection with c-immsim. *Bioinformatics*, submitted.

[17] F. Celada and P. E. Seiden. A computer model of cellular interactions in the immune system. *Immunology Today*, 13(2):56–62, 1992.

[18] F. Celada and P. E. Seiden. Modeling immune cognition. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 4, page 3787Ű3792, 1998.

[19] D.L. Chao, M.P. Davenport, S. Forrest, and A.S. Perelson. A stochastic model of cytotoxic t cell responses. *Journal of Theoretical Biology*, 228(2):227–240, 2004.

[20] M. Cohn, R. Langman, and J. Mata. A computerized model for self -non-self discrimination at the level of the th (th genesis) i. the origin of primer effector th cells. *International Immunology*, 14:1105–1112, 2002.

[21] M. Cohn, R. Langman, and J. Mata. A computerized model for the self-non-self discrimination at the level of the th (th genesis) ii. the behavior of the system upon encounter with non-self antigens. *International Immunology*, 15(5):593Ű609, 2003.

[22] Omar Colón-Reyes, Abdul Salam Jarrah, Reinhard Laubenbacher, and Bernd Sturmfels. Monomial dynamical systems over finite fields. *Complex Systems*, 16(4):333–342, 2006.

[23] Omar Colón-Reyes, Reinhard Laubenbacher, and Bodo Pareigis. Boolean monomial dynamical systems. *Ann. Comb.*, 8(4):425–439, 2004.

[24] Omar Colón-Reyes. *Monomial Dyanmical Systems over Finite Fields*. PhD thesis, Virginia Tech, Blacksburg, Virginia, 2005.

[25] D. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms, An introduction to computational algebraic geometry and commutative algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1997.

[26] Paul Cull. Linear analysis of switching nets. *Kybernetik*, 8(1):31–39, 1971.

[27] M.P. Davenport, C. Fazou, A.J. McMichael, and M.F.C. Callan. Clonal selection, clonal senescence, and clonal succession: The evolution of the t cell response to infection with a persistent virus. *Journal of Immunology*, 168:3309–3317, 2002.

[28] H. De Jong. Modeling and simulation of genetic regulatory systems: A literature review. *J. Comput. Biol.*, 9(1):67–103, 2002.

[29] F. Degani Cattelani and C. Fiocchi. Problema degli autovalori in spazi con prodotto pseudo-interno. *Atti Sem. Mat. Fis. Univ. Modena*, 23(1):55–69 (1975), 1974.

[30] F. Degani Cattelani and C. Fiocchi. Spettro simmetrico e rango numerico di un operatore non lineare in spazi con prodotto pseudo-interno. *Atti Sem. Mat. Fis. Univ. Modena*, 24(1):88–105 (1976), 1975.

[31] E. Delgado-Eckert. Canonical representatives for residue classes of a polynomial ideal and orthogonality. *Comm. Algebra*, under review. See temporary version at http://arxiv.org/abs/0706.1952v1.

[32] E. Delgado-Eckert. Reverse engineering time discrete finite dynamical systems: A feasible undertaking? *J. Comput. Biol.*, under review. See temporary version at http://arxiv.org/abs/0706.3234v1.

[33] Edgar Delgado-Eckert. An algebraic and graph theoretical framework to study monomial dynamical systems over a finite field. *Complex Systems*, under review. See temporary version at http://arxiv.org/abs/0711.1230v1.

[34] P. D'haeseleer, S. Liang, and R. Somogyi. Genetic network inference: From co-expression clustering to reverse engineering. *Bioinformatics*, 16:707–726, 2000.

[35] Elena S. Dimitrova, John J. McGee, and Reinhard C. Laubenbacher. Discretization of time series data, 2005.

[36] K. A. Duca, M. Shapiro, E. Delgado-Eckert, V. Hadinoto, A. S. Jarrah, R. Laubenbacher, K. Lee, K. Luzuriaga, N. F. Polys, and D. A. Thorley-Lawson. A virtual look at epstein-barr virus infection: biological interpretations. *PLoS Pathog*, 3(10):e137, 2007.

[37] R. Durrett and S. Levin. The importance of being discrete (and spatial). *Theoretical Population Biology*, 46:363–394, 1994.

[38] B. Elspas. The theory of autonomous linear sequential networks. *IRE Transactions on Circuit Theory*, CT-6:45–60, 1959.

[39] Drew Endy, Deyu Kong, and John Yin. Intracellular kinetics of a growing virus: A genetically structured simulation for bacteriophage t7. *Biotechnology and Bioengineering*, 55(2):376–389, 1997.

[40] Paul W. Ewald. *Evolution of Infectious Disease*. Oxford University Press, New York, NY, 1994.

[41] C. V. Forst. Host-pathogen systems biology. *Drug Discovery Today*, 11(5-6):220–227, 2006.

[42] T. S. Gardner and J. J. Faith. Reverse-engineering transcription control networks. *Phys. Life Rev.*, 2:65–88, 2005.

[43] D. J. Garrett-Dancik and K. Dorman. An agent-based model for leishmania infection. *International Journal of Complex Systems*, In Press.

[44] Frank Harary, Robert Z. Norman, and Dorwin Cartwright. *Structural models: An introduction to the theory of directed graphs*. John Wiley & Sons Inc., New York, 1965.

[45] René A. Hernández Toledo. Linear finite dynamical systems. *Comm. Algebra*, 33(9):2977–2989, 2005.

[46] Andreas V.M. Herz, Sebastian Bonhoeffer, Roy M. Anderson, Robert M. May, and Martin A. Nowak. Viral dynamics in vivo: Limitations on estimates of intracellular delay and virus decay. *PNAS*, 93:7247–7251, 1996.

[47] M. Hirsch and S. Smale. *Differential Equations, Dynamical Systems and Linear Algebra*. Academic Press (Elsevier, Inc.), Burlington, MA, 1974.

[48] R. J. Hoagland. The transmission of infectious mononucleosis. *American Journal of Medical Science*, 229:262–272, 1955.

[49] D. Hochberg, J. M. Middeldorp, M. Catalina, J. L. Sullivan, K. Luzuriaga, and D. A. Thorley-Lawson. Demonstration of the burkitt's lymphoma epstein-barr virus phenotype in dividing latently infected memory cells in vivo. *Proceedings of the National Academy of Science USA*, 101(1):239–244, 2004.

[50] D. Hochberg, T. Souza, M. Catalina, J. L. Sullivan, K. Luzuriaga, and D. A. Thorley-Lawson. Acute infection with epstein-barr virus targets and overwhelms the peripheral memory b-cell compartment with resting, latently infected cells. *J Virol*, 78(10):5194–204, 2004.

[51] D. Iber and P. K. Maini. A mathematical model for germinal centre kinetics and affinity maturation. *Journal of Theoretical Biology*, 219(2):153–175, 2002.

[52] W. Just. Reverse engineering discrete dynamical systems from data sets with random input vectors. *J. Comput. Biol.*, 13(8):1435–1456 (electronic), 2006.

[53] M. Kaplan. *Computeralgebra*. Springer-Verlag, Berlin, Heidelberg, 2005.

[54] S. Kasahara. Linear independency of linear space valued mappings and pseudo-inner-products. *Math. Japon.*, 25(3):321–325, 1980.

[55] S. A. Kauffman. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, Oxford, UK, May 1993.

[56] G. Khan, E. M. Miyashita, B. Yang, G. J. Babcock, and D. A. Thorley-Lawson. Is ebv persistence in vivo a model for b cell homeostasis? *Immunity*, 5(2):173–9, 1996.

[57] R. Khanna, D. J. Moss, and S. R. Burrows. Vaccine strategies against epstein-barr virus-associated diseases: lessons from studies on cytotoxic t-cell-mediated immune regulation. *Immunol Rev*, 170:49–64, 1999.

[58] S. Kleinstein and P. Seiden. Simulating the immune system. *Computing in Science and Engineering*, 2(4):69–77, 2000.

[59] S.H. Kleinstein and J.P. Singh. Toward quantitative simulation of germinal center dynamics: Biological and modeling insights from experimental validation. *Journal of Theoretical Biology*, 211(3):253–275, 2001.

[60] B. Kohler, R. Puzone, P. E. Seiden, and F. Celada. A systematic approach to vaccine complexity using an automaton model of the cellular and humoral iimmune system i. viral characteristics and polarized responses,. *Vaccine*, 19(7-8):862–876, 2000.

[61] J. U. Kreft, G. Booth, and J. W. Wimpenny. Bacsim, a simulator for individual-based modelling of bacterial colony growth. *Microbiology*, 144(12):3275–3287, 1998.

[62] L. L. Laichalk, D. Hochberg, G. J. Babcock, R. B. Freeman, and D. A. Thorley-Lawson. The dispersal of mucosal memory b cells: evidence from persistent ebv infection. *Immunity*, 16(5):745–54, 2002.

[63] L. L. Laichalk and D. A. Thorley-Lawson. Terminal differentiation into plasma cells initiates the replicative cycle of epstein-barr virus in vivo. *Journal of Virology*, 79(2):1296–1307, 2005.

[64] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[65] R. Laubenbacher and B. Stigler. A computational algebra approach to the reverse engineering of gene regulatory networks. *J. Theoret. Biol.*, 229(4):523–537, 2004.

[66] M. Lauer. Canonical representatives for residue classes of a polynomial ideal. In *SYMSAC '76: Proceedings of the third ACM symposium on Symbolic and algebraic computation*, pages 339–345, New York, NY, USA, 1976. ACM Press.

[67] R. Lidl and H. Niederreiter. *Finite fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, second edition, 1997. With a foreword by P. M. Cohn.

[68] L. Ljung. *System identification: theory for the user*. Prentice Hall information and system sciences series. Prentice Hall PTR, Upper Saddle River, New Jersey, second edition, 1999.

[69] G. Lumer. Semi-inner-product spaces. *Trans. Amer. Math. Soc.*, 100:29–43, 1961.

[70] Y. Mansury, M. Diggory, and T. S. Deisboeck. Evolutionary game theory in an agent-based brain tumor model: Exploring the Šgenotype-phenotypeŠ link. *Journal of Theoretical Biology*, 238(1):146–156, 2006.

[71] M. Meir-Schellersheim. *Simmune, a Tool for Simulating and Analyzing Immune System Behavior*. Phd, University of Hamburg, 1999.

[72] M. E. Meyer-Hermann, P. K. Maini, and D. Iber. An analysis of b cell selection mechanisms in germinal centers. *Mathematical Medical Biology*, 23(3):255–277, 2006.

[73] D. K. Milligan and M. J. D. Wilson. The behavior of affine boolean sequential networks. *Connection Sciences*, 5(2):153–167, 1993.

[74] M. Mininni and G. Muni. A degree theory with respect to a pseudo-inner product in a locally convex vector space. *Ricerche Mat.*, 28(2):365–374, 1979.

[75] E. M. Miyashita, B. Yang, G. J. Babcock, and D. A. ThorleyňLawson. Identification of the site of epstein-barr virus persistence in vivo as a resting b-cell. *Journal of Virology*, 71(7):4882–4891, 1997.

[76] M. Mobilia, I. T. Georgiev, and U. C. Taeuber. Fluctuations and correlations in lattice models for predator-prey interaction. *Physical Review E*, 73(4), 2006.

[77] D. Morpurgo, R. Serentha, P. E. Seiden, F. Celada, and E. Sercarz. Modelling thymic functions in a cellular automaton. *International Immunology*, 7(4):505–516, 1995.

[78] P.W. Nelson and A.S. Perelson. Mathematical analysis of delay differential equation models of hiv-1 infection. *Mathematical Biosciences*, 179(1):73–94, 2002.

[79] E. A. Neugebauer and T. Tjardes. New approaches to shock and trauma reearch: Learning from multidisciplinary exchange. *Journal of Trauma*, 56(5):1156–1165, 2004.

[80] A.U. Neumann, N.P. Lam, H. Dalari, D.R. Gretch, T.E. Wiley, T.J. Layden, and A.S. Perelson. Hepatitis c viral dynamics in vivo and the antiviral efficacy of interferon a therapy. *Science*, 282:103–107, 1998.

[81] M.A. Nowak and C.R. Bangham. Population dynamics of immune responses to persistent viruses. *Science*, 272:74–79, 1996.

[82] Martin A. Nowak, Sebastian Bonhoeffer, Andrew M. Hill, Richard Boehme, Howard C. Thomas, and Hugh McDade. Viral dynamics in hepatitis b virus infection. *PNAS*, 93:4398–4402, 1996.

[83] K. Ogata. *Discrete-Time Control Systems*. Prentice Hall, Inc., Upper Saddle River, New Jersey, second edition, 1995.

[84] D. M. Pegtel, J. Middeldorp, and D. A. Thorley-Lawson. Epstein-barr virus infection in ex vivo tonsil epithelial cell cultures of asymptomatic carriers. *Journal of Virology*, 78(22):12613–12624, 2004.

[85] A. Perelson. Modeling viral and immune system dynamics. *Nature Reviews: Immunology*, 2(1):28–36, 2002.

[86] A. Perelson and G. Weissbuch. Immunology for physicists. *Review of Modern Physics*, 69(4), 1995.

[87] A.S. Perelson, D.E. Kirschner, and R. DeBoer. Dynamics of hiv-infection of cd4+ t-cells. *Mathematical Biosciences*, 4(1):81–125, 1993.

[88] A.S. Perelson, A.U. Neumann, M. Markovitz, J.M. Leonard, and D.D. Ho. Hiv-1 dynamics in vivo: Virion clearance rate, infected cell life-span, and viral generation time. *Science*, 271(5255):1582–1586, 1996.

[89] M. Perry and A. Whyte. Immunology of the tonsils. *Immunology Today*, 19(9):414–421, 1998.

[90] Eric Post, Rolf O. Peterson, Nils Chr. Stenseth, and Brian E. McLaren. Ecosystem consequences of wolf behavioural response to climate. *Nature*, 401:905–907, 1999.

[91] R. Puzone, B. Kohler, P. E. Seiden, and F. Celada. Immsim, a flexible model for in machina experiments on immune system responses. *Future Generation Computer Systems*, 18(7):961–972, 2002.

[92] J. Reger and K. Schmidt. A finite field framework for modeling, analysis and control of finite state automata. *Mathematical and Computer Modelling of Dynamical Systems (MCMDS)*, 10(3):253–285, 2004.

[93] J. Reger and K. Schmidt. Modeling and analyzing finite state automata in the finite field f2. *Mathematics and Computers in Simulation*, 66(2-3):193–206, 2004.

[94] R.M. Ribeiro, A. Lo, and A.S. Perelson. Dynamics of hepatitus b virus infection. *Microbes and Infection*, 4(8):829–835, 2002.

[95] Ruy M. Ribiero and Sebastian Bonhoeffer. Production of resistant hiv mutants during antiretroviral therapy. *Proceedings of the National Academy of Science*, 97(14):7681–7686, 2000.

[96] Alan B. Rickinson and Elliot Kieff. Epstein-barr virus. In D. Knipe and P. Howley, editors, *Virology*, volume 2, pages 2575–2628. Lipincott, Williams, and Wilkins, New York, New York, fourth edition, 2001.

[97] L. Robbiano. Gröbner bases and statistics. In *Gröbner bases and applications (Linz, 1998)*, volume 251 of *London Math. Soc. Lecture Note Ser.*, pages 179–204. Cambridge Univ. Press, Cambridge, 1998.

[98] W. Scharlau. *Quadratic forms*, volume 22 of *Queen's papers in pure and applied mathematics*. Queen's University, Kingston, Ontario, 1969.

[99] T. A. Seemayer, T. G. Gross, R. M. Egeler, S. J. Pirruccello, J. R. Davis, C. M. Kelly, M. Okano, A. Lanyi, and J. Sumegi. X-linked lymphoproliferative disease: twenty-five years after the discovery. *Pediatr Res*, 38(4):471–478, 1995.

[100] Lee A. Segel. Spatio-temporal models in immunology. *Bifurcation and Chaos*, 12:2343–2347, 2002.

[101] Lee A. Segel and Ruth Lev Bar-Or. On the role of feedback in promoting conflicting goals of the adaptive immune system. *Journal of Immunology*, 163:1342–1349, 1999.

[102] Jose L. Segovia-Juarez, Suman Ganguli, and Denise Kirschner. Identifying control mechanisms of granuloma formation during m. tuberculosis infection using an agent-based model. *Journal of Theoretical Biology*, 231:357–376, 2004.

[103] P. E. Seiden and F. Celada. A model for simulating cognate recognition and response in the immune system. *Journal of Theoretical Biology*, 158(3):329–357, 1992.

[104] M. Shapiro, K.A. Duca, E. Delgado-Eckert, V. Hadinoto, A.S. Jarrah, R. Laubenbacher, K. Lee, N.F. Polys, and D.A. Thorley-Lawson. A virtual look at epstein-barr virus infection: Simulation mechanism. *J Theor Bio*, under revision, 2007.

[105] L. M. Sompayrac. *How the Immune System Works*. Blackwell Publishing, Malden, MA, second edition, 2003.

[106] Eduardo D. Sontag. *Mathematical control theory*, volume 6 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1998. Deterministic finite-dimensional systems.

[107] R. Srivastava, L. You, J. Summers, and J. Yin. Stochastic vs. deterministic modeling of intracellular viral kinetics. *Journal of Theoretical Biology*, 218(3):309–321, 2002.

[108] J. J. Stewart, H. Agosto, S. Litwin, J. D. Welsh, M. Shlomchik, M. Weigert, and P. E. Seiden. A solution to the rheumatoid factor paradox: Pathologic rheumatoid factors can be tolerized by competition with natural rheumatoid factors. *Journal of Immunology*, 159(4):1728–1738, 1997.

[109] J. Tay and A. Jhavar. Cafiss: A complex adaptive framework for immune system simulations. In *ACM Symposium for Applied Computing-Bioinformatics*, 2005.

[110] R. Thomas. Regulatory networks seen as asynchronous automata: A logical description. *J. Theoret. Biol.*, 153(1):1–23, 1991.

[111] David A. Thorley-Lawson. Epstein-barr virus: Exploiting the immune system. *Nature Reviews: Immunology*, 1(1):75–82, 2001.

[112] David A. Thorley-Lawson and A. Gross. Persistence of the epstein-barr virus and the origins of associated lymphomas. *New England Journal of Medicine*, 350(13):1328–1337, 2004.

[113] W. B. Vasantha and T. Johnson. New spectral theorem for vector spaces over finite fields Z_p. *Varāhmihir J. Math. Sci.*, 3(2):355–364, 2003.

[114] D. C. Walker, G. Hill, S. M. Wood, R. H. Smallwood, and J. Southgate. Agent-based computational modeling of wounded epithelial cell monolayers. *IEEE transactions on nanobioscience*, 3(3):153–163, 2004.

Index