Vishwesh V. Kulkarni · Guy-Bart Stan
Karthik Raman *Editors*

A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations

Springer

A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations

Vishwesh V. Kulkarni
Guy-Bart Stan · Karthik Raman
Editors

A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations

Springer

Editors
Vishwesh V. Kulkarni
Electrical and Computer Engineering
University of Minnesota
Minneapolis, MN
USA

Guy-Bart Stan
Department of Bioengineering
Imperial College
London
UK

Karthik Raman
Department of Biotechnology
Bhupat and Jyoti Mehta School
of Biosciences
Indian Institute of Technology Madras
Chennai
India

ISBN 978-94-017-9040-6 ISBN 978-94-017-9041-3 (eBook)
DOI 10.1007/978-94-017-9041-3
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2014942544

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

श्रेयो हि ज्ञानमभ्यासाज्ज्ञानाद्ध्यानं विशिष्यते ।
ध्यानात् कर्मफलत्यागस्त्यागाच्छान्तिरनन्तरम् ॥१२-१२॥
– श्रीमद्भगवद्गीता

Understanding is superior to mere practice
Union with the subject matter supersedes that
Dispassion towards all results is better still
And manifests peace immediately

Bhagavat Gita (12:12)

To my father Vasant, brother Vinay, sister Ketki, and Prof. Peter Falb
To my mother in memory

Vishwesh V. Kulkarni

To my parents, Florentina and Stephan
To my wife Cristina and my daughter Eva-Victoria

Guy-Bart Stan

To my parents and teachers
In memory of Sunder Mama and Prof. E. V. Krishnamurthy

Karthik Raman

Foreword

There is no design template more versatile than DNA. Nor are any designs more consequential than those whose blueprints DNA encodes. This exquisite substance has been shaped over billions of years by the creative combination of mutation and selection. Yet in the very long history of this template, it is only during our times that complex living organisms are beginning to understand and manipulate the very template whose sequences define them. But how should we go about this understanding? And how can we use this understanding to more effectively and responsibly alter the DNA template?

The complexity and diversity of living organisms are daunting. *Systems biology* aims at reverse engineering biological complexity for the purpose of understanding their design principles. By measuring and characterizing interactions of key biological molecules in response to stimuli and perturbations, systems biology aims to construct models that capture the complexity of endogenous biological networks. Through the systematic understanding of such models, it is hoped that one will achieve a holistic understanding of biological networks and the way they achieve biological function.

At the same time, the versatility of DNA and the dramatic decrease in the cost of DNA synthesis is making it possible to economically design and test new complex genetic circuits. This has given impetus to a new field: *Synthetic biology*. In our quest to understand biological complexity, we have examined endogenous biological subsystems and ascribed functions and design principles to their components. But a true understanding of these biological design principles is demonstrated only when one can build such systems *de novo* and demonstrate their function. When these circuits do not exhibit behavior consistent with our models, further investigations will lead to a deeper understanding of the underlying biology. Synthetic biology, therefore, serves as an important testbed for our understanding of biological principles. But the promise of synthetic biology extends beyond scientific understanding. Whether it be the detection and interference with the course of disease through the introduction of designer circuits, the cost-effective synthesis of new bio-substances, or the development of improved food products, synthetic biology provides a tremendous opportunity to alleviate suffering and improve the quality of our lives.

In both systems and synthetic biology, challenges abound. Quantitative modeling, analysis, and design of biological networks must contend with difficulties arising from the inescapable fact that at its most basic level, biology involves complex dynamic interactions among nonlinear stochastic components, taking place at multiple temporal and spatial timescales. The complexity of network interconnections of such components and the crosstalk between them adds another level of difficulty.

System theory has emerged as a field to deal with the challenges and complexities emerging from the interconnection of engineered systems, many of which are shared with biological systems. Notions from system theory such as nonlinearity, stochasticity, feedback, loading, modularity, robustness, identifiability, etc., are needed for a deeper understanding of biological complexity and for a more reliable design of biological circuits. These concepts are now being utilized to help us expand our understanding of endogenous biological circuits and to design novel ones. The articles in this book make significant strides in this direction.

While system theory will undoubtedly aid our understanding and design of biological systems, there is no doubt that the study of biological designs that have evolved over billions of years will also shape the future of system theory. For example, evolution and development are two central themes in biology that have little analogy with engineered man-made systems. Through the study of these and other biological themes, new systems notions and insights will undoubtedly emerge, enriching system theory in the process. One need only look at the history of *feedback*, a predominant concept in system theory, to imagine what is possible. While its human discovery can be traced back a little over one millennium, it is likely that feedback was invented by nature more than three billion years earlier. Since then, it has been wildly successful as a biological design principle, as evidenced by its prevalence at every level of biological organization. One wonders if an early systematic understanding of this concept in its biological context could have sped up the course of our own technological development.

As the physical sciences helped us understand the physical world around us over the last few centuries, so will quantitative biological science help us understand who we are, how we function, and how we can effectively and responsibly synthesize this most consequential of substances, the DNA. I believe that system theory will be central to this understanding.

Zürich, September 2013 Mustafa Khammash

Preface

Underlying every living cell are billions of molecules interacting in a beautifully concerted network of pathways such as metabolic, signalling, and regulatory pathways. The complexity of such biological systems has intrigued scientists from many disciplines and has given birth to the highly influential field of *systems biology* wherein a wide array of mathematical techniques, such as flux balance analysis, and technology platforms, such as next generation sequencing, is used to understand, elucidate, and predict the functions of complex biological systems. This field traces its roots to the general systems theory of Ludwig von Bertalanffy and effectively started in 1952 with a mathematical model of the neuronal action potential for which Alan Hodgkin and Andrew Huxley received the Nobel Prize in 1963. More recently, the field of *synthetic biology*, i.e., *de novo* engineering of biological systems, has emerged. Here, the phrase 'biological system' can assume a vast spectrum of meanings: DNA, protein, genome, cell, cell population, tissue, organ, ecosystem, and so on. Scientists from various fields are focusing on how to render this *de novo* engineering process more predictable, reliable, scalable, affordable, and easy. Systems biology and synthetic biology are essentially two facets of the same entity. As was the case with electronics research in the 1950s, a large part of synthetic biology research, such as the *BioFab* project, has focused on reusable macromolecular "parts" and their standardization so that composability can be guaranteed. Recent breakthroughs in DNA synthesis and sequencing combined with newly acquired means to synthesize plasmids and genomes have enabled major advances in science and engineering and marked the true beginning of the era of synthetic biology. Significant industrial investments are already underway. For example, in 2009, Exxon Mobil set up a collaboration worth $600 million with Synthetic Genomics to develop next generation biofuels.

Recent advances in systems and synthetic biology clearly demonstrate the benefits of a rigorous and systematic approach rooted in the principles of systems and control theory—not only does it lead to exciting insights and discoveries but it also reduces the inordinately lengthy trial-and-error process of wet-lab experimentation, thereby facilitating significant savings in human and financial resources. So far, state-of-the-art systems-and-control-theory-inspired results in systems and synthetic biology have been scattered across various books and journals from various disciplines. Hence, we felt the need for an edited book that provides a

panoramic view and illustrates the potential of such systematic and rigorous mathematical methods in systems and synthetic biology.

Systems and control theory is a branch of engineering and applied sciences that rigorously deals with the complexities and uncertainties of interconnected systems with the objective of characterizing fundamental systemic properties such as stability, robustness, communication capacity, and other performance metrics. Systems and control theory also strives to offer concepts and methods that facilitate the design of systems with rigorous guarantees on these fundamental properties. For more than 100 years, the insights and techniques provided by systems and control theory have enabled outstanding technological contributions in diverse fields such as aerospace, telecommunication, storage, automotive, power systems, and others. Notable examples include Lyapunov's theorems, Bellman's theory of dynamic programming, Kalman's filter, H^∞ control theory, Nyquist-Shannon sampling theorem, Pontryagin's minimum principle, and Bode's sensitivity integral. Can systems and control theory have, or evolve to have, a similar impact in biology? The chapters in this book demonstrate that, indeed, systems and control theoretic concepts and techniques can be useful in our quest to understand how biological systems function and/or how they can be (re-)designed from the bottom up to yield new biological systems that have rigorously characterized robustness and performance properties.

Several barriers must be overcome to contribute significantly in this exciting journey. One of these is the language barrier, e.g., what a systems theorist means by the word *sensitivity* is different from what a biologist means by it. Another one is the knowledge barrier as, traditionally, systems and control theorists and biologists are not well versed with each other's knowledge base (although that scenario is now fast changing for the better with the introduction of bioengineering courses in systems and control theory at the undergraduate and graduate levels). A third barrier is due to the sheer volume of *big data*: the European Bioinformatics Institute in Hixton, UK, which is one of the world's largest biological data repositories, currently stores 20 petabytes of data and backups about genes, proteins and small molecules, and this number is more than doubling every year. Finally, a fourth barrier comes from the effort required to produce timely contributions based on currently available models. As an example of this last barrier, the systems and control theory community could have played a greater role than it did in two of the most significant technological advances of the last 50 years: VLSI and Internet. In retrospect, besides the fact that the systems and control theorists caught on the Internet too late, by which time infrastructures based on TCP/IP were already in place, the main difficulty posed by the Internet for the systems and control theory community was a lack of *good* models of the underlying networked system. This lack-of-good-models barrier is even more daunting in biology since some of the currently available *big data* are not guaranteed to be reproducible. As Prof. M. Vidyasagar illustrates and observes in the September 2012 issue of IEEE Lifesciences, one of the major challenges to the application of systems and control theory concepts in biology comes from "the fact that many biological experiments are not fully repeatable, and thus the resulting data sets are not readily amenable to

the application of methods that people like us [i.e., systems and control theorists] take for granted."

The chapters in this book serve to propose ways to overcome such barriers and to illustrate that biologists as well as systems and control theorists can make deep and timely contributions in life sciences by collaborating with each other to solve important questions such as how to devise experiments to obtain models of biological systems, how to obtain predictive models using information extracted from experimental data, how to choose components for (re-)engineering biological networks, how to adequately interconnect biological systems, and so on. Furthermore, and as Prof. Mustafa Khammash observes in his foreword, this research will fundamentally enrich systems and control theory as well by forcing it to investigate currently open questions that are specific to living biological systems, e.g., Why do biological systems naturally evolve the way they do? Can the evolvability of biological systems be consciously exploited for (re-)design and optimization purposes?

This book is intended for (1) systems and control theorists interested in molecular and cellular biology, and (2) biologists interested in rigorous modeling, analysis, and control of biological systems. We believe that research at the intersection of these disciplines will foster exciting discoveries and will stimulate mutually beneficial developments in systems & control theory and systems & synthetic biology.

The book consists of 12 chapters contributed by leading researchers from the fields of systems and control theory, systems biology, synthetic biology, and computer science. Chapters 1–6 focus on general mathematical concepts, methods, and tools that are currently used to answer important questions in biology. Chapters 7–12 describe various biological network modeling approaches used to untangle biological complexity and reverse-engineer biological networks from data.

- **Part I—Mathematical Analysis:** Chapters 1–6 present core mathematical concepts and methods that can be used and further adapted for solving specific problems in biology. As an example, consider the law of mass action. It has been widely used in chemistry since Guldberg and Waage formulated it in 1864. But does it have a deeper significance that is applicable outside chemistry? Likewise, reaction-diffusion systems feature in all pattern formation problems which, in turn, are significant in neuronal networks and disease phenotypes. Under which conditions is spatial uniformity guaranteed? The chapters in this part provide rigorous mathematical foundations that can be used to resolve such questions. A brief summary of each chapter is as follows.

  – Chapter 1: The law of mass action is used in (bio-)chemistry to characterize and predict the behavior of interacting (bio-)chemical species. Guldberg and Waage formulated it in 1864 and it has since been built upon and widely used in (bio-)chemistry and cellular biology. To make it available for consideration by researchers in areas other than chemistry, Adleman et al. present it in a new form, viz., in the context of event systems, after solidifying its mathematical foundations.

- Chapter 2: Molecular systems often have a mathematical representation with uncertainties embedded in it. These uncertainties make predictions of the system's behavior harder. Nonetheless, it is still possible, in some scenarios, to obtain certain qualitative behavioral results that are fairly parameter independent and, instead, are a property of the system structure. Blanchini and Franco use a parameter-free qualitative modeling framework and show under which conditions behaviors such as oscillations and multi-stability are only structure dependent.
- Chapter 3: Reaction-diffusion systems are of central importance in all applications that feature pattern formations. Aminzare et al. present conditions that guarantee the spatial uniformity of the solutions of reaction-diffusion partial differential equations. They demonstrate that these conditions can be verified using linear matrix inequalities and outline the applicability of these results in analyzing biological oscillations and enzymatic signalling pathways.
- Chapter 4: Biologists often rely on linearized models to examine stability and on phase-plane analysis to understand the effect of parameter variations. Although useful, phase-plane analysis cannot be used to address simultaneous variations in more than two parameters. Kulkarni et al. show how multiplier theory can be used to overcome these limitations and illustrate its use via a case study of the celebrated Elowitz–Leibler oscillator.
- Chapter 5: Modularity possibly emerged at the cellular level through natural selection and evolution. But do modules make sense in the context of metabolic networks? Goelzer and Fromion present a framework that allows a modular decomposition of steady-state metabolic networks, and show how this framework can also be used for a qualitative predictive modeling based on omics datasets.
- Chapter 6: Biological network modeling often encounters the problem of how to deal with hidden state dynamics. Santiello et al. address the problem of predicting hidden state transitions from temporal sequential datasets (for example, EEG, EMG, MER) by developing a Bayesian detection paradigm that combines optimal control and Markov processes.

- **Part II—Biological Network Modeling:** Chapters 7–12 focus on certain techniques that can be used to obtain predictive models of biological networks. Here, the limitations of the perturbation methods used to generate the data, the vast amount of available data (which does not necessarily correlate with the amount of *information* they contain), hidden states, measurement noise, and other factors combine to render this broad area of research one of the greatest scientific and technological challenges of today. The chapters in this section summarize some of these challenges and present architectures that constitute an important step in arriving at a definitive solution. Somewhat similar, but less complex system identification problems have been encountered and resolved in systems theory and computer science over the last decades. Can these techniques and the insight they provide be useful in biology? To answer this

question, it is crucial to understand the advantages and limitations of each particular technique. The set of chapters collated in this part aim to highlight the current state of the art for biological network modeling and the advantages and limitations of the presented approaches. A brief summary of each chapter is as follows.

- Chapter 7: In metabolic networks, the metabolite dynamics evolve on much shorter timescales than their catalytic enzymes. Kuntz et al. show how such timescale separation can be exploited using Tikhonov's theorem for singularly perturbed systems to derive reduced models whose behaviors are guaranteed to remain quantifiably close to those of the non-reduced models. They illustrate this approach by applying it to an example of genetic feedback control for branched metabolic pathways.
- Chapter 8: A central theme in complex network theory, popularized by the study of small-world and scale-free networks at the turn of the last century, is the study of biological networks using various metrics. In this chapter, Roy discusses the utility of various network metrics as well as the need to go beyond fundamental metrics, such as node degree, to better understand how an organism's phenotype is encoded by its network topology.
- Chapter 9: Even though most of the complex real-world systems exhibit nonlinearities, linear models serve as a useful first order approximation. Carignano et al. present a detailed exposition on how linear system identification techniques can be used to obtain causal relationships between biomolecular entities.
- Chapter 10: Fisher and Piterman discuss how ideas from computer science can be useful for *model checking* in systems biology. They present a methodology to analyze biochemical networks, and specifically a method to test for a faithful reproduction of biological interactions that are known a priori as well as to identify interactions that are not known a priori.
- Chapter 11: Bussetto et al. discuss objective-specific strategies for designing informative experiments in systems biology. Following a formal description of the task of experimental design, they illustrate the use of Bayesian and information-theoretic approaches to design experiments in systems biology.
- Chapter 12: Today, there is a critical need for new methods that rapidly transform high-throughput genomics, transcriptomics, and metabolomics data into predictive network models for metabolic engineering and synthetic biology. In this chapter, Chandrasekaran describes the state of the art of these methods and explains an approach for this purpose called Probabilistic Regulation of Metabolism (PROM).

The burgeoning fields of systems biology and synthetic biology have thrown up a very large number of interesting research problems. As the pre-eminent computer scientist Donald Knuth put it, "biology easily has 500 years of exciting problems to work on." The chapters in this book address but a small fraction of these interesting challenges. Nevertheless, we believe this book can serve as a

good introduction on some of the currently open problems and on some of the state-of-the-art concepts and techniques available to propose solutions to such problems.

We are very grateful to all authors for their invaluable time and contributions and to Prof. Mustafa Khammash (ETH Zürich) for his stimulating foreword. We are also grateful to our institutions: University of Minnesota (Minneapolis, USA), Imperial College (London, UK), and Indian Institute of Technology Madras (Chennai, India) for their support and for providing a stimulating work environment. Finally, we thank and acknowledge the financial support of our respective funding agencies: the National Science Foundation, the UK Engineering and Physical Sciences Research Council, and the Ministry of Human Resource and Development of the Government of India.

Minneapolis, MN, USA, September 2013 Vishwesh V. Kulkarni
London, UK Guy-Bart Stan
Chennai, India Karthik Raman

Contents

Part I
Mathematical Analysis

Chapter 1
On the Mathematics of the Law of Mass Action

Leonard Adleman, Manoj Gopalkrishnan, Ming-Deh Huang, Pablo Moisset and Dustin Reishus

Abstract In 1864, Waage and Guldberg formulated the "law of mass action." Since that time, chemists, chemical engineers, physicists and mathematicians have amassed a great deal of knowledge on the topic. In our view, sufficient understanding has been acquired to warrant a formal mathematical consolidation. A major goal of this consolidation is to solidify the mathematical foundations of mass action chemistry—to provide precise definitions, elucidate what can now be proved, and indicate what is only conjectured. In addition, we believe that the law of mass action is of intrinsic mathematical interest and should be made available in a form that might transcend its application to chemistry alone. We present the law of mass action in the context of a dynamical theory of sets of binomials over the complex numbers.

Keywords Law of mass action · Mass action kinetics · Event systems · Binomials · String theory · Differential equations · Flow-invariant affine subspaces · Natural event systems · Lyapunov function

L. Adleman · M.-D. Huang
University of Southern California, Los Angeles, CA, USA
e-mail: adleman@usc.edu

M.-D. Huang
e-mail: huang@usc.edu

M. Gopalkrishnan (✉)
Tata Institute of Fundamental Research, Mumbai, India
e-mail: manojg@tifr.res.in

P. Moisset
Universidad de Chile, Santiago, Chile
e-mail: pablo.moisset@gmail.com

D. Reishus
University of Colorado Boulder, Boulder, CO, USA
e-mail: reishus@colorado.edu

V. V. Kulkarni et al. (eds.), *A Systems Theoretic Approach to Systems and Synthetic Biology I: Models and System Characterizations*, DOI: 10.1007/978-94-017-9041-3_1,

1.1 Introduction

The study of mass action kinetics dates back at least to 1864, when Guldberg and Waage [7] formulated the "law of mass action." Since that time, a great deal of knowledge on the topic has been amassed in the form of empirical facts, physical theories and mathematical theorems by chemists, chemical engineers, physicists and mathematicians. In recent years, Horn and Jackson [10], and Feinberg [5] have made significant mathematical contributions, and these have guided our work.

It is our view that a critical mass of knowledge has been obtained, sufficient to warrant a formal mathematical consolidation. A major goal of this consolidation is to solidify the mathematical foundations of this aspect of chemistry—to provide precise definitions, elucidate what can now be proved, and indicate what is only conjectured. In addition, we believe that the law of mass action is of intrinsic mathematical interest and should be made available in a form that might transcend their application to chemistry alone.

To make the law of mass action available for consideration by researchers in areas other than chemistry, we present mass action kinetics in a new form, which we call event-systems. Our formulation begins with the observation that systems of chemical reactions can be represented by sets of binomials. This gives us an opportunity to extend the law of mass action to arbitrary sets of binomials. Once this extension is made, there is no reason to restrict ourselves to binomials with real coefficients. Hence, we are led to a dynamical theory of sets of binomials over the complex numbers. Possible mathematical applications of this theory include:

1. Binomials are objects of intrinsic mathematical interest [4]. For example, they occur in the study of toric varieties, and hence in string theory. With each set of binomials over the complex numbers, we associate a corresponding system of differential equations. Ideally, this dynamical viewpoint will help advance the theory of binomials, and enhance our understanding of their associated algebraic sets.
2. When we extend the study of the law of mass action to sets of binomials over the complex numbers, we can consider reactions that involve complex rates, complex concentrations, and move through complex time. Extending to the complex numbers gives us direct access to the powerful theorems of complex analysis. Though this clearly transcends conventional chemistry, it may have applications in pure mathematics.

   For example, in ongoing work, we seek to exploit an analogy between number theory and chemistry, where atoms are to molecules as primes are to numbers. We associate a distinct species with each natural number. Then each multiplication rule $m \times n = mn$ is encoded by a reaction where the species corresponding to the number m reacts with the species corresponding to the number n to form the species corresponding to the number mn. With an appropriate choice of specific rates of reactions the resulting event-system has the property that the sum of equilibrium concentrations of all species at complex temperature s is the value of the

Riemann zeta function at s. We hope to pursue this approach to study questions related to the distribution of the primes.

3. Systems of linear differential equations are well understood. In contrast, systems of ordinary non-linear differential equations can be notoriously intractable. Differential equations that arise from event-systems lie somewhere in between—more structured than arbitrary non-linear differential equations, but more challenging than linear differential equations. As such, they appear to be an important new class for consideration in the theory of ordinary differential equations.

In addition to their use in mathematics, event-systems provide a vehicle by which ideas in algebraic geometry may be made readily available to the study of mass action kinetics. As such, they may help solidify the foundations of this aspect of chemistry. We expand on this in Sect. 1.7.

Part of our motivation for this research comes from the emerging field of nanotechnology. To quote from [1], "Self-assembly is the ubiquitous process by which objects autonomously assemble into complexes. Nature provides many examples: Atoms react to form molecules. Molecules react to form crystals and supramolecules. Cells sometimes coalesce to form organisms. Even heavenly bodies self-assemble into astronomical systems. It has been suggested that self-assembly will ultimately become an important technology, enabling the fabrication of great quantities of small objects such as computer circuits… Despite its importance, self-assembly is poorly understood." Hopefully, the theory of event-systems is a step towards understanding this important process.

The chapter is organized as follows:

In Sect. 1.2, we present the basic mathematical notations and definitions for the study of event-systems.

In Sect. 1.3, and all of the sections that follow, we restrict to finite event-systems. Theorem 3 demonstrates that the stoichiometric coefficients give rise to flow-invariant affine subspaces—"conservation classes."

In Sect. 1.4, and all of the sections that follow, we restrict to "physical event-systems." Though we have defined event-systems over the complex numbers, in this chapter we focus on consolidating results from the mass action kinetics of reversible chemical reactions. Physical event-systems capture the idea that the specific rates of chemical reactions are always positive real numbers. The main result of this section is Theorem 4, which demonstrates that for physical event-systems, if initially all concentrations are non-negative, then they stay non-negative for all future real times so long as the solution exists. Further, the concentration of every species whose initial concentration is positive, stays positive.

In Sect. 1.5, and all the sections that follow, we restrict to "natural event-systems." Natural event-systems capture the concept of detailed balance from chemistry. In Theorem 5, we give four equivalent characterizations of natural event-systems; in particular, we show that natural event-systems are precisely those physical event-systems that have no "energy cycles." In Theorem 7, following Horn and Jackson [10], we show that natural event-systems have associated Lyapunov functions. This theorem is reminiscent of the second law of thermodynamics. The main result

of this section is Theorem 10, which establishes that for natural event-systems, given non-negative initial conditions:

1. Solutions exist for all forward real times.
2. Solutions are uniformly bounded in forward real time.
3. All positive equilibria satisfy detailed balance.
4. Every conservation class containing a positive point also contains exactly one positive equilibrium point.
5. Every positive equilibrium point is asymptotically stable relative to its conservation class.

For systems of reversible reactions that satisfy detailed balance, must concentrations approach equilibrium? We believe this to be the case, but are unable to prove it. In 1972, an incorrect proof was offered [10, Lemma 4C]. This proof was retracted in 1974 [9]. To the best of our knowledge, this question in mass action kinetics remains unresolved [14, p. 10]. We pose it formally in Open Problem 1, and consider it the fundamental open question in the field.

In Sect. 1.6, we introduce the notion of "atomic event-systems." As the name suggests, this is an attempt to capture mathematically the atomic hypothesis that all species are composed of atoms. The main theorem of this section is Theorem 11, which establishes that for natural, atomic event-systems, solutions with positive initial conditions asymptotically approach positive equilibria. Hence, Open Problem 1 is resolved in the affirmative for this restricted class of event-systems.

1.2 Basic Definitions and Notation

Before formally defining event-systems, we give a very brief, informal introduction to chemical reactions. All reactions are assumed to take place at constant temperature in a well-stirred vessel of constant volume.

Consider

$$A + 2B \underset{\tau}{\overset{\sigma}{\rightleftharpoons}} C.$$

This chemical equation concerns the reacting species A, B and C. In the forward direction, one mole of A combines with two moles of B to form one mole of C. The symbol "σ" represents a real number greater than zero. It denotes, in appropriate units, the rate of the forward reaction when the reaction vessel contains one mole of A and one mole of B. It is called the specific rate of the forward reaction. In the reverse direction, one mole of C decomposes to form one mole of A and two moles of B. The symbol "τ" represents the specific rate of the reverse reaction. Chemists typically determine specific rates empirically. Though irreversible reactions (those with $\sigma = 0$ or $\tau = 0$) have been studied, they will not be considered in this chapter.

Inspired by the law of mass action, we introduce a multiplicative notation for chemical reactions, as an alternative to the chemical equation notation. In our

notation, each chemical reaction is represented by a binomial. Consider the following examples. On the left are chemical equations. On the right are the corresponding binomials.

$$\begin{aligned} X_2 \underset{1/3}{\overset{1/2}{\rightleftharpoons}} X_1 &\rightarrow \frac{1}{3}X_2 - \frac{1}{2}X_1 \\ X_3 \underset{1/3}{\overset{1/2}{\rightleftharpoons}} X_1 + X_2 &\rightarrow \frac{1}{3}X_3 - \frac{1}{2}X_1X_2 \\ 2X_1 + 3X_6 \underset{\tau}{\overset{\sigma}{\rightleftharpoons}} 3X_1 + 2X_2 &\rightarrow \sigma X_1^2X_6^3 - \tau X_1^3X_2^2 \end{aligned}$$

Our notation leads us to view every set of binomials over an arbitrary field $\mathbb{F}$ as a formal system of reversible reactions with specific rates in $\mathbb{F}\backslash\{0\}$. For our present purposes, we will restrict our attention to binomials over the complex numbers. With this in mind, we now define our notion of event-system.

Notation 1 *Let $\mathbb{C}_\infty = \bigcup_{n=1}^{\infty} \mathbb{C}[X_1, X_2, \ldots, X_n]$. A monic monomial of $\mathbb{C}_\infty$ is a product of the form $\prod_{i=1}^{\infty} X_i^{e_i}$ where the e_i are non-negative integers all but finitely many of which are zero. We will write $\mathbb{M}_\infty$ to denote the set of all monic monomials of $\mathbb{C}_\infty$. More generally, if $S \subset \{X_1, X_2, \ldots\}$, we let $\mathbb{C}[S]$ be the ring of polynomials with indeterminants in S and we let $\mathbb{M}_S = \mathbb{M}_\infty \cap \mathbb{C}[S]$ (i.e. the monic monomials in $\mathbb{C}[S]$).*

If $n \in \mathbb{Z}_{>0}$, $p \in \mathbb{C}[X_1, X_2, \ldots, X_n]$, and $\boldsymbol{a} = \langle a_1, a_2, \ldots, a_n \rangle \in \mathbb{C}^n$ then, as is usual, we will let $p(a)$ denote the value of p on argument a.

Given two monic monomials $M = \prod_{i=1}^{\infty} X_i^{e_i}$ and $N = \prod_{i=1}^{\infty} X_i^{f_i}$ from $\mathbb{M}_\infty$, we will say M precedes *N (and we will write $M \prec N$) iff $M \neq N$ and for the least i such that $e_i \neq f_i$, $e_i < f_i$.*

It follows that 1 is a monic monomial of $\mathbb{C}_\infty$ and that each element of $\mathbb{C}_\infty$ is a $\mathbb{C}$-linear combination of finitely many monic monomials. We will be particularly concerned with the set of binomials $\mathbb{B}_\infty = \{\sigma M + \tau N \mid \sigma, \tau \in \mathbb{C}\backslash\{0\}$ and M, N are distinct monic monomials of $\mathbb{C}_\infty\}$.

Definition 2 (*Event-system*) An event-system $\mathscr{E}$ is a nonempty subset of $\mathbb{B}_\infty$.

If $\mathscr{E}$ is an event-system, its elements will be called "$\mathscr{E}$-events" or just "events." Note that if $\sigma M + \tau N$ is an event then $M \neq N$.

Our map from chemical equations to events is as follows. A chemical equation

$$\sum_i a_i X_i \underset{\tau}{\overset{\sigma}{\rightleftharpoons}} \sum_j b_j X_j \text{ goes to:}$$

$$1.\ \sigma \prod_i X_i^{a_i} - \tau \prod_j X_j^{b_j} \quad \text{if } \prod_i X_i^{a_i} \prec \prod_j X_j^{b_j}$$

$$\text{or } 2.\ \ \tau \prod_j X_j^{b_j} - \sigma \prod_i X_i^{a_i} \quad \text{if } \prod_j X_j^{b_j} \prec \prod_i X_i^{a_i}$$

For example:

$$X_1 \underset{1/2}{\overset{1/3}{\rightleftharpoons}} X_2 \to \frac{1}{3}X_2 - \frac{1}{2}X_1 \text{ (because } X_2 \prec X_1\text{)}$$

$$X_2 \underset{1/3}{\overset{1/2}{\rightleftharpoons}} X_1 \to \frac{1}{3}X_2 - \frac{1}{2}X_1$$

$$X_1 \underset{-1/2}{\overset{-1/3}{\rightleftharpoons}} X_2 \to -\frac{1}{3}X_2 + \frac{1}{2}X_1$$

$$X_1 \underset{-1/2}{\overset{1/3}{\rightleftharpoons}} X_2 \to \frac{1}{3}X_2 + \frac{1}{2}X_1$$

$$X_1 + X_2 \underset{1/2}{\overset{1/3}{\rightleftharpoons}} X_3 \to \frac{1}{3}X_3 - \frac{1}{2}X_1X_2$$

$$3X_1 + 2X_2 \underset{\tau}{\overset{\sigma}{\rightleftharpoons}} 2X_1 + 3X_6 \to \tau X_1^2 X_6^3 - \sigma X_1^3 X_2^2$$

Note that our order of monomials is arbitrary. Any linear order would do. The order is necessary to achieve a one-to-one map from chemical reactions to events.

Our definition of event-systems allows for an infinite number of reactions, and an infinite number of reacting species. Indeed, polymerization reactions are commonplace in nature and, in principle, they are capable of creating arbitrarily long polymers (for example, DNA molecules).

The next definition introduces the notion of systems of reactions for which the number of reacting species is finite.

Definition 3 (*Finite-dimensional event-system*) An event-system $\mathscr{E}$ is *finite-dimensional* iff there exists an $n \in \mathbb{Z}_{>0}$ such that $\mathscr{E} \subset \mathbb{C}[X_1, X_2, \ldots, X_n]$.

Definition 4 (*Dimension of event-systems*) Let $\mathscr{E}$ be a finite-dimensional event-system. Then the least n such that $\mathscr{E} \subset \mathbb{C}[X_1, X_2, \ldots, X_n]$ is the *dimension* of $\mathscr{E}$.

Definition 5 (*Physical event, Physical event-system*) A binomial $e \in \mathbb{B}_\infty$ is a *physical event* iff there exist $\sigma, \tau \in \mathbb{R}_{>0}$ and $M, N \in \mathbb{M}_\infty$ such that $M \prec N$ and $e = \sigma M - \tau N$. An event-system $\mathscr{E}$ is *physical* iff each $e \in \mathscr{E}$ is physical.

Chemical reaction systems typically have positive real forward and backward rates. Physical event-systems generalize this notion.

Definition 6 Let $n \in \mathbb{Z}_{>0}$. Let $\alpha = \langle \alpha_1, \alpha_2, \ldots, \alpha_n \rangle \in \mathbb{C}^n$.

1. α is a *non-negative point* iff for $i = 1, 2, \ldots, n$, $\alpha_i \in \mathbb{R}_{\geq 0}$.
2. α is a *positive point* iff for $i = 1, 2, \ldots, n$, $\alpha_i \in \mathbb{R}_{>0}$.
3. α is a *z-point* iff there exists an i such that $\alpha_i = 0$.

In chemistry, a system is said to have achieved detailed balance when it is at a point where the net flux of each reaction is zero. Given the corresponding event-system, points of detailed balance corresponds to points where each event evaluates to zero, and vice versa. We call such points "strong equilibrium points."

Definition 7 (*Strong equilibrium point*) Let $\mathscr{E}$ be a finite-dimensional event-system of dimension n. $\alpha \in \mathbb{C}^n$ is a *strong $\mathscr{E}$-equilibrium point* iff for all $e \in \mathscr{E}$, $e(\alpha) = 0$.

In the language of algebraic geometry, when $\mathscr{E}$ is a finite-dimensional event-system, its corresponding algebraic set is precisely the set of its strong $\mathscr{E}$-equilibrium points.

It is widely believed that all "real" chemical reactions achieve detailed balance. We now introduce natural event-systems, a restriction of finite-dimensional, physical event-systems to those that can achieve detailed balance.

Definition 8 (*Natural event-system*) A finite-dimensional event-system $\mathscr{E}$ is *natural* iff it is physical and there exists a positive strong $\mathscr{E}$-equilibrium point.

Our next goal is to introduce atomic event-systems: finite-dimensional event-systems obeying the atomic hypothesis that all species are composed of atoms. Towards this goal, we will define a graph for each finite-dimensional event-system. The vertices of this graph are the monomials from $\mathbb{M}_\infty$ and the edges are determined by the events. If a weight r is assigned to an edge, then r represents the energy released when a reaction corresponding to that edge takes place. For the purpose of defining atomic event-systems, the reader may ignore the weights; they are included here for use elsewhere in the chapter (Definition 24).

Though graphs corresponding to systems of chemical reactions have been defined elsewhere (e.g. [5, 14, p. 10]), it is important to note that these definitions do not coincide with ours.

Definition 9 (*Event-graph*) Let $\mathscr{E}$ be a finite-dimensional event-system. The event-graph $G_{\mathscr{E}} = \langle V, E, w \rangle$ is a weighted, directed multigraph such that:

1. $V = \mathbb{M}_\infty$
2. For all $M_1, M_2 \in \mathbb{M}_\infty$, for all $r \in \mathbb{C}$,
   $\langle M_1, M_2 \rangle \in E$ and $r \in w(\langle M_1, M_2 \rangle)$ iff
   there exist $e \in \mathscr{E}$ and $\sigma, \tau \in \mathbb{C}$ and $M, N, T \in \mathbb{M}_\infty$ such that $e = \sigma M + \tau N$ and $M \prec N$ and either
   (a) $M_1 = TM$ and $M_2 = TN$ and $r = \ln\left(-\frac{\sigma}{\tau}\right)$ or
   (b) $M_1 = TN$ and $M_2 = TM$ and $r = -\ln\left(-\frac{\sigma}{\tau}\right)$

Notice that two distinct weights r_1 and r_2 could be assigned to a single edge. For example, let $\mathscr{E} = \{X_1X_2 - 2X_1^2, X_2 - 5X_1\}$. Consider the edge in $G_{\mathscr{E}}$ from the monomial X_1^2 to the monomial X_1X_2. Weight $\ln 2$ is assigned to this edge due to the event $X_1X_2 - 2X_1^2$, with $T = 1$. Weight $\ln 5$ is also assigned to this edge due to the event $X_2 - 5X_1$, with $T = X_1$.

Definition 10 Let $\mathscr{E}$ be a finite-dimensional event-system. For all $M \in \mathbb{M}_\infty$, the *connected component* of M, denoted $C_{\mathscr{E}}(M)$, is the set of all $N \in \mathbb{M}_\infty$ such that there is a path in $G_{\mathscr{E}}$ from M to N.

It follows from the definition of "path" that every monomial belongs to its connected component.

Definition 11 (*Atomic event-system*) Let $\mathscr{E}$ be a finite-dimensional event-system of dimension n. Let $S = \{X_1, X_2, \ldots, X_n\}$. Let $A_{\mathscr{E}} = \{X_i \in S \mid C_{\mathscr{E}}(X_i) = \{X_i\}\}$. $\mathscr{E}$ is *atomic* iff for all $M \in \mathbb{M}_S$, $C(M)$ contains a unique monomial in $\mathbb{M}_{A_{\mathscr{E}}}$.

If $\mathscr{E}$ is atomic then the members of $A_{\mathscr{E}}$ will be called *the atoms of* $\mathscr{E}$. It follows from the definition that in atomic event-systems, atoms are not decomposable, non-atoms are uniquely decomposable into atoms and events preserve atoms.

Since the set $\mathbb{M}_{\{X_1, X_2 \ldots, X_n\}}$ is infinite, it is not possible to decide whether $\mathscr{E}$ is atomic by exhaustively checking the connected component of every monomial in $\mathbb{M}_{\{X_1, X_2 \ldots, X_n\}}$. The following is sometimes helpful in deciding whether a finite-dimensional event-system is atomic (proof not provided).

Let $\mathscr{E}$ be an event-system of dimension n with no event of the form $\sigma + \tau N$. Let $B_{\mathscr{E}} = \{X_i \mid \text{For all } \sigma, \tau \in \mathbb{C}\backslash\{0\} \text{ and } N \in \mathbb{M}_\infty : \sigma X_i + \tau N \notin \mathscr{E}\}$. Then $\mathscr{E}$ is atomic iff there exist $M_1 \in C_{\mathscr{E}}(X_1) \cap \mathbb{M}_{B_{\mathscr{E}}}$, $M_2 \in C_{\mathscr{E}}(X_2) \cap \mathbb{M}_{B_{\mathscr{E}}}, \ldots, M_n \in C_{\mathscr{E}}(X_n) \cap \mathbb{M}_{B_{\mathscr{E}}}$ such that:

$$\text{For all } \sigma \prod_{i=1}^{n} X_i^{a_i} - \tau \prod_{i=1}^{n} X_i^{b_i} \in \mathscr{E}, \quad \prod_{i=1}^{n} M_i^{a_i} = \prod_{i=1}^{n} M_i^{b_i}. \tag{1.1}$$

We have shown (proof not provided) that if $\mathscr{E}$ and $B_{\mathscr{E}}$ are as above, and there exist $M_1 \in C_{\mathscr{E}}(X_1) \cap \mathbb{M}_{B_{\mathscr{E}}}$, $M_2 \in C_{\mathscr{E}}(X_2) \cap \mathbb{M}_{B_{\mathscr{E}}}, \ldots, M_n \in C_{\mathscr{E}}(X_n) \cap \mathbb{M}_{B_{\mathscr{E}}}$ and there exists $\sigma \prod_{i=1}^{n} X_i^{a_i} - \tau \prod_{i=1}^{n} X_i^{b_i} \in \mathscr{E}$ such that $\prod_{i=1}^{n} M_i^{a_i} \neq \prod_{i=1}^{n} M_i^{b_i}$, then $\mathscr{E}$ is not atomic. Hence, to check whether an event-system with no event of the form $\sigma + \tau N$ is atomic, it suffices to examine an arbitrary choice of $M_1 \in C_{\mathscr{E}}(X_1) \cap \mathbb{M}_{B_{\mathscr{E}}}$, $M_2 \in C_{\mathscr{E}}(X_2) \cap \mathbb{M}_{B_{\mathscr{E}}}, \ldots, M_n \in C_{\mathscr{E}}(X_n) \cap \mathbb{M}_{B_{\mathscr{E}}}$, if one exists, and check whether (1.1) above holds.

Example 1 Let $\mathscr{E} = \{X_2^2 - X_1^2\}$. Then $B_{\mathscr{E}} = \{X_1, X_2\}$. Let $M_1 = X_1$ and $M_2 = X_2$. Trivially, $M_1, M_2 \in \mathbb{M}_{B_{\mathscr{E}}}$, $M_1 \in C_{\mathscr{E}}(X_1)$ and $M_2 \in C_{\mathscr{E}}(X_2)$. Consider the event $X_2^2 - X_1^2$. Since $M_2^2 = X_2^2 \neq X_1^2 = M_1^2$, $\mathscr{E}$ is not atomic. Note that the event $X_2^2 - X_1^2$ does not preserve atoms.

Example 2 Let $\mathscr{E} = \{X_4^2 - X_2, X_5^2 - X_3, X_2X_3 - X_1\}$. Then $B_{\mathscr{E}} = \{X_4, X_5\}$. Let $M_1 = X_4^2X_5^2$, $M_2 = X_4^2$, $M_3 = X_5^2$, $M_4 = X_4$, $M_5 = X_5$. Clearly these are all in $\mathbb{M}_{B_{\mathscr{E}}}$. $X_5^2 - X_3 \in \mathscr{E}$ implies $M_3 \in C_{\mathscr{E}}(X_3)$. $X_4^2 - X_2 \in \mathscr{E}$ implies $M_2 \in C_{\mathscr{E}}(X_2)$. Since $(X_1, X_2X_3, X_2X_5^2, X_4^2X_5^2)$ is a path in $G_{\mathscr{E}}$, we have $M_1 \in C_{\mathscr{E}}(X_1)$. For the event $X_4^2 - X_2$, we have $M_4^2 = X_4^2 = M_2$. For the event $X_5^2 - X_3$, we have $M_5^2 = X_5^2 = M_3$. For the event $X_2X_3 - X_1$, we have $M_2M_3 = X_4^2X_5^2 = M_1$. Therefore, $\mathscr{E}$ is atomic.

Note that it is possible to have an atomic event-system where $A_{\mathscr{E}}$ is the empty set. For example:

Example 3 Let $\mathscr{E} = \{1 - X_1\}$. In this case, $S = \{X_1\}$ and $\mathbb{M}_S$ is the set $\{1, X_1, X_1^2, X_1^3, \ldots\}$. It is clear that $\mathbb{M}_S$ forms a single connected component C in $G_{\mathscr{E}}$. Hence, X_1 is not in $A_{\mathscr{E}}$, and $A_{\mathscr{E}} = \emptyset$. 1 is the only monomial in $\mathbb{M}_{A_{\mathscr{E}}}$. Since 1 is in C, $\mathscr{E}$ is atomic.

1.3 Finite Event-Systems

The study of infinite event-systems is embryonic and appears to be quite challenging. In the rest of this chapter only finite event-systems (i.e., where the set $\mathscr{E}$ is finite) will be considered. It is clear that all finite event-systems are finite-dimensional.

Definition 12 (*Stoichiometric matrix*) Let $\mathscr{E} = \{e_1, e_2, \ldots, e_m\}$ be an event-system of dimension n. Let $i \leq n$ and $j \leq m$ be positive integers. Let $e_j = \sigma M + \tau N$, where $M \prec N$. Then $\gamma_{j,i}$ is the number of times X_i divides N minus the number of times X_i divides M. The *stoichiometric matrix* $\Gamma_{\mathscr{E}}$ of $\mathscr{E}$ is the $m \times n$ matrix of integers $\Gamma_{\mathscr{E}} = (\gamma_{j,i})_{m \times n}$.

Example 4 Let $e_1 = 0.5X_2^5 - 500X_1X_2^3X_7$. Let $\mathscr{E} = \{e_1\}$. Then $\gamma_{1,1} = 1$, $\gamma_{1,2} = -2$, $\gamma_{1,7} = 1$ and for all other i, $\gamma_{1,i} = 0$, hence $\Gamma_{\mathscr{E}} = \begin{pmatrix} 1 & -2 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$.

Definition 13 Let $\mathscr{E} = \{e_1, \ldots, e_m\}$ be a finite event-system of dimension n. Then:

1. $P_{\mathscr{E}}$ is the column vector $\langle P_1, P_2, \ldots, P_n \rangle^T = \Gamma_{\mathscr{E}}^T \langle e_1, e_2, \ldots, e_m \rangle^T$.
2. Let $\alpha \in \mathbb{C}^n$. Then α is an $\mathscr{E}$-equilibrium point iff for $i = 1, 2, \ldots, n$: $P_i(\alpha) = 0$.

The P_i's arise from the Law of Mass Action in chemistry. For a system of chemical reactions, the P_i's are the right-hand sides of the differential equations that describe the concentration kinetics. Definition 13 extends the Law of Mass Action to arbitrary event-systems, and hence, arbitrary sets of binomials.

It follows from the definition that for finite event-systems, all strong equilibrium points are equilibrium points, but the converse need not be true.

Example 5 Let $e_1 = X_2 - X_1$ and $e_2 = X_2 - 2X_1$. Let $\mathscr{E} = \{e_1, e_2\}$. Then $\Gamma_{\mathscr{E}} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$ and $p_{\mathscr{E}} = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = \begin{pmatrix} 2X_2 - 3X_1 \\ 3X_1 - 2X_2 \end{pmatrix}$. Therefore $(2, 3)$ is an $\mathscr{E}$-equilibrium point. Since $e_1(2, 3) = 1$, $(2, 3)$ is not a strong $\mathscr{E}$-equilibrium point.

Example 6 Let $e_1 = 6 - X_1X_2$ and $e_2 = 2X_2^2 - 9X_1$. Let $\mathscr{E} = \{e_1, e_2\}$. Then $\Gamma_{\mathscr{E}} = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$ and $p_{\mathscr{E}} = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = \begin{pmatrix} 6 - X_1X_2 + 2X_2^2 - 9X_1 \\ 6 - X_1X_2 - 4X_2^2 + 18X_1 \end{pmatrix}$. The point $(2, 3)$ is a strong equilibrium point because $e_1(2, 3) = 0$ and $e_2(2, 3) = 0$. Since $P_1(2, 3) = e_1(2, 3) + e_2(2, 3) = 0$ and $P_2(2, 3) = e_1(2, 3) - 2e_2(2, 3) = 0$, the point $(2, 3)$ is also an equilibrium point.

The event-system in Example 5 is not natural, whereas the one in Example 6 is. In Theorem 8, it is shown that if $\mathscr{E}$ is a finite, natural event-system then all positive $\mathscr{E}$-equilibrium points are strong $\mathscr{E}$-equilibrium points.

Definition 14 (*Event-process*) Let $\mathscr{E}$ be a finite event-system of dimension n. Let $\langle P_1, P_2, \ldots, P_n \rangle^T = p_{\mathscr{E}}$. Let $\Omega \subseteq \mathbb{C}$ be a non-empty simply-connected open set. Let $f = \langle f_1, f_2, \ldots, f_n \rangle$ where for $i = 1, 2, \ldots, n$, $f_i : \mathbb{C} \to \mathbb{C}$ is defined on Ω. Then f is an $\mathscr{E}$-process on Ω iff for $i = 1, 2, \ldots, n$:

1. f_i' exists on Ω.
2. $f_i' = P_i \circ f$ on Ω.

Note that $\mathscr{E}$-processes evolve through complex time, and hence generalize the idea of the time-evolution of concentrations in a system of chemical reactions.

Definition 14 immediately implies that if $f = \langle f_1, f_2, \ldots, f_n \rangle$ is an $\mathscr{E}$-process on Ω, then for $i = 1, 2, \ldots, n$, f_i is holomorphic on Ω. In particular, for each i and all $\alpha \in \Omega$, there is a power series around α that agrees with f_i on a disk of non-zero radius.

Systems of chemical reactions sometimes obey certain conservation laws. For example, they may conserve mass, or the total number of each kind of atom. Event-systems also sometimes obey conservation laws.

Definition 15 (*Conservation law, Linear conservation law*) Let $\mathscr{E}$ be a finite event-system of dimension n. A function $g: \mathbb{C}^n \to \mathbb{C}$ is a *conservation law of* $\mathscr{E}$ iff g is holomorphic on $\mathbb{C}^n$, $g(\langle 0, 0, \ldots, 0 \rangle) = 0$ and $\nabla g \cdot P_{\mathscr{E}}$ is identically zero on $\mathbb{C}^n$. If g is a conservation law of $\mathscr{E}$ and g is linear (i.e. $\forall c \in \mathbb{C}, \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{C}^n$, $g(c\boldsymbol{\alpha} + \boldsymbol{\beta}) = cg(\boldsymbol{\alpha}) + g(\boldsymbol{\beta})$), then g is a *linear conservation law of* $\mathscr{E}$.

The event-system described in Example 5 has a linear conservation law $g(X_1, X_2) = X_1 + X_2$. The next theorem shows that conservation laws of $\mathscr{E}$ are dynamical invariants of $\mathscr{E}$-processes.

Theorem 1 *For all finite event-systems $\mathscr{E}$, for all conservation laws g of $\mathscr{E}$, for all simply-connected open sets $\Omega \subseteq \mathbb{C}$, for all $\mathscr{E}$-processes f on Ω, there exists $k \in \mathbb{C}$ such that $g \circ f - k$ is identically zero on Ω.*

Proof Let n be the dimension of $\mathscr{E}$. Let $\langle P_1, P_2, \ldots, P_n \rangle^T = p_{\mathscr{E}}$. For all $t \in \Omega$, by Definition 14, for $i = 1, 2, \ldots, n$, $f_i(t)$ and $f_i'(t)$ are defined. Further, by Definition 15, g is holomorphic on $\mathbb{C}^n$. Hence, $g \circ f$ is holomorphic on Ω. Therefore, by the chain rule, $(g \circ f)'(t) = (\nabla g|_{f(t)}) \cdot \langle f_1'(t), f_2'(t), \ldots, f_n'(t) \rangle$. By Definition 14, for all $t \in \Omega$, $\langle f_1'(t), f_2'(t), \ldots, f_n'(t) \rangle = \langle P_1(f(t)), P_2(f(t)), \ldots, P_n(f(t)) \rangle$. From these, it follows that $(g \circ f)'(t) = (\nabla g \cdot P_{\mathscr{E}})(f(t))$. But by Definition 15, $\nabla g \cdot P_{\mathscr{E}}$ is identically zero. Hence, for all $t \in \Omega$, $(g \circ f)'(t) = 0$. In addition, Ω is a simply-connected open set. Therefore, by [2, Theorem 11], there exists $k \in \mathbb{C}$ such that $g \circ f - k$ is identically zero on Ω.

The next theorem shows a way to derive linear conservation laws of an event-system from its stoichiometric matrix.

Theorem 2 *Let $\mathscr{E}$ be a finite event-system of dimension n. For all $V \in \ker \Gamma_{\mathscr{E}}$, $V \cdot \langle X_1, \ldots, X_n \rangle$ is a linear conservation law of $\mathscr{E}$.*

Proof Let $\Gamma = \Gamma_{\mathscr{E}}$, then $\ker \Gamma$ is orthogonal to the image of Γ^T. By the definition of $P = P_{\mathscr{E}}$, for all $w \in \mathbb{C}^n$, $P(w)$ lies in the image of Γ^T. Hence, for all $V \in \ker \Gamma$, for all $w \in \mathbb{C}^n$, $V \cdot P(w) = 0$. But V is the gradient of $V \cdot \langle X_1, \ldots, X_n \rangle$. It now follows from Definition 15 that $V \cdot \langle X_1, \ldots, X_n \rangle$ is a linear conservation law of $\mathscr{E}$.

Definition 16 (*Primitive conservation law*) Let $\mathscr{E}$ be a finite event-system of dimension n. For all $V \in \ker \Gamma_{\mathscr{E}}$, the linear conservation law $V \cdot \langle X_1, X_2, \ldots, X_n \rangle$ is a *primitive conservation law*.

We can show that in physical event-systems all linear conservation laws are primitive and, in natural event-systems, all conservation laws arise from the primitive ones.

Definition 17 (*Conservation class, Positive conservation class*) Let $\mathscr{E}$ be a finite event-system of dimension n. A coset of $(\ker \Gamma_{\mathscr{E}})^{\perp}$ is a *conservation class of* $\mathscr{E}$. If a conservation class of $\mathscr{E}$ contains a positive point, then the class is a *positive conservation class of* $\mathscr{E}$.

Equivalently, $\alpha, \beta \in \mathbb{C}^n$ are in the same conservation class if and only if they agree on all primitive conservation laws. Note that if H is a conservation class of $\mathscr{E}$ then it is closed in $\mathbb{C}^n$. The following theorem shows that the name "conservation class" is appropriate.

Theorem 3 *Let $\mathscr{E}$ be a finite event-system. Let $\Omega \subset \mathbb{C}$ be a simply-connected open set containing* 0. *Let f be an $\mathscr{E}$-process on Ω. Let H be a conservation class of $\mathscr{E}$ containing $f(0)$. Then for all $t \in \Omega$, $f(t) \in H$.*

Proof Let $\mathscr{E}, \Omega, f, H$ and t be as in the statement of this theorem. For all $V \in \ker \Gamma_{\mathscr{E}}$, the primitive conservation law $V \cdot \langle X_1, X_2, \ldots, X_n \rangle$ is a dynamical invariant of f, from Theorem 2 and Theorem 1. Hence,

$$V \cdot \langle f_1(0), f_2(0), \ldots, f_n(0) \rangle = V \cdot \langle f_1(t), f_2(t), \ldots, f_n(t) \rangle$$

That is,

$$V \cdot \langle f_1(0) - f_1(t), f_2(0) - f_2(t), \ldots, f_n(0) - f_n(t) \rangle = 0$$

Hence, $f(t) - f(0)$ is in $(\ker \Gamma_{\mathscr{E}})^{\perp}$. By Definition 17, $f(t) \in H$.

1.4 Finite Physical Event-Systems

In this section, we investigate finite, physical event-systems—a generalization of systems of chemical reactions.

It is widely believed that systems of chemical reactions that begin with positive (respectively, non-negative) concentrations will have positive (respectively, non-negative) concentrations at all future times. This property has been addressed mathematically in numerous papers [6, p. 6], [5, Remark 3.4], [3, Theorem 3.2], [14, Lemma 2.1]. The notion of "system of chemical reactions" varies between papers. Several papers have provided no proof, incomplete proofs or inadequate proofs that this property holds for their systems. Sontag [14, Lemma 2.1] provides a lovely

proof of this property for the systems he considers—zero deficiency reaction networks with one linkage class. We shall prove in Theorem 4 that the property holds for finite, physical event-systems. Finite, physical event-systems have a large intersection with the systems considered by Sontag, but each includes a large class of systems that the other does not. We remark that our methods of proof differ from Sontag's, but it is possible that Sontag's proof might be adaptable to our setting.

Lemma 4 and Lemma 10 are proved here because they apply to finite, physical event-systems. However, they are only invoked in subsequent sections. Lemma 4 relates $\mathscr{E}$-processes to solutions of ordinary differential equations over the reals. Lemma 10 establishes that if an $\mathscr{E}$-process defined on the positive reals starts at a real, non-negative point, then its ω-limit set is invariant and contains only real, non-negative points.

The next lemma shows that if two $\mathscr{E}$-processes evaluate to the same real point on a real argument then they must agree and be real-valued on an open interval containing that argument. The proof exploits the fact that $\mathscr{E}$-processes are analytic, by considering their power series expansions.

Lemma 1 *Let $\mathscr{E}$ be a finite, physical event-system of dimension n, let $\Omega, \Omega' \subseteq \mathbb{C}$ be open and simply-connected, let $f = \langle f_1, f_2, \ldots, f_n \rangle$ be an $\mathscr{E}$-process on Ω and let $g = \langle g_1, g_2, \ldots, g_n \rangle$ be an $\mathscr{E}$-process on Ω'. If $t_0 \in \Omega \cap \Omega' \cap \mathbb{R}$ and $f(t_0) \in \mathbb{R}^n$ and $f(t_0) = g(t_0)$, then there exists an open interval $I \subseteq \mathbb{R}$ such that $t_0 \in I$ and for all $t \in I$:*

1. $f(t) = g(t)$.
2. *For $i = 1, 2, \ldots, n$: if $\sum_{j=0}^{\infty} c_j (z - t_0)^j$ is the Taylor series expansion of f_i at t_0 then for all $j \in \mathbb{Z}_{\geq 0}$, $c_j \in \mathbb{R}$.*
3. $f(t) \in \mathbb{R}^n$.

Proof Let $k \in \mathbb{Z}_{\geq 0}$. By Definition 14, f and g are vectors of functions analytic at t_0. For $i = 1, 2, \ldots, n$, let $f_i^{(k)}$ be the kth derivative of f_i and let $f^{(k)} = \langle f_1^{(k)}, f_2^{(k)}, \ldots, f_n^{(k)} \rangle$. Define $g_i^{(k)}$ and $g^{(k)}$ similarly. To prove 1, it is enough to show that for $i = 1, 2, \ldots, n$, f_i and g_i have the same Taylor series around t_0. Let $V_0 = \langle X_1, X_2, \ldots, X_n \rangle$. Let $V_k = \mathrm{Jac}(V_{k-1}) P_{\mathscr{E}}$ (recall that if $H = \langle h_1(X_1, X_2, \ldots, X_m), h_2(X_1, X_2, \ldots, X_m), \ldots, h_n(X_1, X_2, \ldots, X_m) \rangle$ is a vector of functions in m variables then $\mathrm{Jac}(H)$ is the $n \times m$ matrix $(\frac{\partial h_i}{\partial x_j})$, where $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$). Let $\langle V_{k,1}, V_{k,2}, \ldots, V_{k,n} \rangle = V_k$. We claim that $f^{(k)} = V_k \circ f$ on Ω and $g^{(k)} = V_k \circ g$ on Ω' and for $i = 1, 2, \ldots, n$, $V_{k,i} \in \mathbb{R}[X_1, X_2, \ldots, X_n]$. We prove the claim by induction on k. If $k = 0$, the proof is immediate. If $k \geq 1$, on Ω:

$$
\begin{aligned}
f^{(k)} &= (f^{(k-1)})' \\
&= (V_{k-1} \circ f)' \quad \text{(Inductive hypothesis)} \\
&= (\mathrm{Jac}(V_{k-1}) \circ f) f' \quad \text{(Chain-rule of derivation)} \\
&= (\mathrm{Jac}(V_{k-1}) \circ f)(P_{\mathscr{E}} \circ f) \quad (f \text{ is an } \mathscr{E}\text{-process})
\end{aligned}
$$

$$= (\mathrm{Jac}(V_{k-1})P_{\mathscr{E}}) \circ f$$
$$= V_k \circ f$$

By a similar argument, we conclude that $g^{(k)} = V_k \circ g$ on Ω'. By the inductive hypothesis, V_{k-1} is a vector of polynomials in $\mathbb{R}[X_1, X_2, \ldots, X_n]$. It follows that $\mathrm{Jac}(V_{k-1})$ is an $n \times n$ matrix of polynomials in $\mathbb{R}[X_1, X_2, \ldots, X_n]$. Since $\mathscr{E}$ is physical, $P_{\mathscr{E}}$ is a vector of polynomials in $\mathbb{R}[X_1, X_2, \ldots, X_n]$. Therefore, $V_k = \mathrm{Jac}(V_{k-1})P_{\mathscr{E}}$ is a vector of polynomials in $\mathbb{R}[X_1, X_2, \ldots, X_n]$. This establishes the claim.

We have proved that $f^{(k)} = V_k \circ f$ on Ω and $g^{(k)} = V_k \circ g$ on Ω'. Since, by assumption, $t_0 \in \Omega \cap \Omega'$ and $f(t_0) = g(t_0)$, it follows that $f^{(k)}(t_0) = g^{(k)}(t_0)$. Therefore, for $i = 1, 2, \ldots, n$, f_i and g_i have the same Taylor series around t_0. For $i = 1, 2, \ldots, n$, let a_i be the radius of convergence of the Taylor series of f_i around t_0. Let $r_f = \min_{i\in\{1,2,\ldots,n\}} a_i$. Define r_g similarly. Let $D \subseteq \Omega \cap \Omega'$ be some non-empty open disk centered at t_0 with radius $r \leq \min(r_f, r_g)$. Since Ω and Ω' are open sets and $t_0 \in \Omega \cap \Omega'$, such a disk must exist. Letting $I = (t_0 - r, t_0 + r)$ completes the proof of 1.

By assumption, $f(t_0) \in \mathbb{R}^n$, and we have proved that $f^{(k)} = V_k \circ f$ and V_k is a vector of polynomials in $\mathbb{R}[X_1, X_2, \ldots, X_n]$. It follows that $f^{(k)}(t_0) \in \mathbb{R}^n$. Therefore, for $i = 1, 2, \ldots, n$, all coefficients in the Taylor series of f_i around t_0 are real. It follows that f_i is real valued on I, completing the proof of 3.

The next lemma is a kind of uniqueness result. It shows that if two $\mathscr{E}$-processes evaluate to the same real point at 0 then they must agree and be real-valued on every open interval containing 0 where both are defined. The proof uses continuity to extend the result of Lemma 1.

Lemma 2 *Let $\mathscr{E}$ be a finite, physical event-system of dimension n, let $\Omega, \Omega' \subseteq \mathbb{C}$ be open and simply-connected, let $f = \langle f_1, f_2, \ldots, f_n\rangle$ be an $\mathscr{E}$-process on Ω and let $g = \langle g_1, g_2, \ldots, g_n\rangle$ be an $\mathscr{E}$-process on Ω'. If $0 \in \Omega \cap \Omega'$ and $f(0) \in \mathbb{R}^n$ and $f(0) = g(0)$, then for all open intervals $I \subseteq \Omega \cap \Omega' \cap \mathbb{R}$ such that $0 \in I$, for all $t \in I$, $f(t) = g(t)$ and $f(t) \in \mathbb{R}^n$.*

Proof Assume there exists an open interval $I \subseteq \Omega \cap \Omega' \cap \mathbb{R}$ such that $0 \in I$ and $B = \{t \in I | f(t) \neq g(t) \text{ or } f(t) \notin \mathbb{R}^n\} \neq \varnothing$. Let $B_P = B \cap \mathbb{R}_{\geq 0}$ and let $B_N = B \cap \mathbb{R}_{<0}$. Note that $B = B_P \cup B_N$, hence, $B_P \neq \varnothing$ or $B_N \neq \varnothing$. Suppose $B_P \neq \varnothing$ and let $t_P = \inf(B_P)$. By Lemma 1, there exists an $\varepsilon \in \mathbb{R}_{>0}$ such that $(-\varepsilon, \varepsilon) \cap B = \varnothing$. Hence, $t_P \geq \varepsilon > 0$. By definition of t_P, for all $t \in [0, t_P)$, $f(t) = g(t)$ and $f(t) \in \mathbb{R}^n$. Since f and g are analytic at t_P, they are continuous at t_P. Therefore, $f(t_P) = g(t_P)$ and $f(t_P) \in \mathbb{R}^n$. By Lemma 1, there exists an $\varepsilon' \in \mathbb{R}_{>0}$ such that for all $t \in (t_P - \varepsilon', t_P + \varepsilon')$, $f(t) = g(t)$ and $f(t) \in \mathbb{R}^n$, contradicting t_P being the infimum of B_P. Therefore, $B_P = \varnothing$. Using a similar agument, we can prove that $B_N = \varnothing$. Therefore, $B = \varnothing$, and for all $t \in I$, $f(t_P) = g(t_P)$ and $f(t_P) \in \mathbb{R}^n$.

The next lemma is a convenient technical result that lets us ignore the choice of origin for the time variable.