Data-Driven Modeling with Hybrid Dynamical Systems

Bora S. Banjanin

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2019

Reading Committee:

Samuel A. Burden, Chair

Katherine M. Steele

Eli Shlizerman

Program Authorized to Offer Degree:
Department of Electrical and Computer Engineering

University of Washington

Abstract

Data-Driven Modeling with Hybrid Dynamical Systems

Bora S. Banjanin

Chair of the Supervisory Committee:
Professor Samuel A. Burden
Department of Electrical and Computer Engineering

Hybrid dynamical systems are used to describe systems that can instantaneously change state and dynamics. At small timescales, continuous electrodynamics govern the interaction of rigid bodies. Simulating the corresponding stiff differential equation introduces unnecessary complexity when the restitution of velocities post-impact is the phenomenon of interest. Although classical mathematics and physics deals primarily with smooth physical processes, the dynamics of real-world systems can and does abruptly change. We can learn from data to inform the structure and fit the parameters of hybrid dynamical models for such systems. These data-driven methods leverage developments in sensing and computation and are a natural progression in the study of modeling and controlling systems. Continuously collecting data can yield interactive systems that adapt towards a target behavior. An accurate computational model can also verify the safety and efficacy of engineered systems.

This thesis seeks to further the practical application of data-driven hybrid dynamical systems – to control robotic systems and assistive devices. In the first aim, hybrid dynamical systems are commonly used to model mechanical systems subject to unilateral constraints, *e.g.* legged locomotion. We demonstrated that nonsmoothness can cause standard optimization techniques to lose convergence guarantees and contribute to poor performance for the resulting control policy. The second aim seeks to predict rhythmic human locomotion with a motive to improve the clinical prescription of Ankle Foot Orthoses (AFO). We created

subject-specific models that can predict how an individual will respond to an untested AFO torque profile. These aims tie together advancements in data science with the inherent ability of hybrid dynamical systems to represent phenomena of interest in the real world.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

This thesis is a product of the contributions of many. Foresmostly, I thank Sam Burden, my advisor, for his mentorship in research and staunch support throughout the duration of my PhD. I thank Zoran Banjanin and Snezana Ivkovic, my parents, for their dedication and ceaseless encouragement. Marina Banjanin, my sister, for her trustworthy advice and loyalty. Bosa Ivkovic, my grandmother, for pampering me for countless summers. Emily Graham, my partner, for her enthusiasm and faith in me. Meeka, my cat, for her unyielding affection. I would also like to extend my sincerest thanks to all my friends and relatives.

The second aim resulted from a fruitful collaboration with Michael Rosenberg, Momona Yamagami, and Kat Steele. I thank the committee members who served on the various exams required by the program: Eli Shlizerman, Steve Shen, Howard Chizeck, and Maryam Fazel.

I cherished the community and sense of belonging at the University of Washington for the past ten years. Finally, I thank Vijeth Rai, Astrini Sie, Andrew Pace, Joey Sullivan, Ben Chasnov, Ben Ferleger, and other members of the graduate program for their camaraderie.

DEDICATION

to Zoran Banjanin and Snezana Ivkovic

Chapter 1

INTRODUCTION

Hybrid dynamical systems are used to describe systems that discretely jump between continuous domains of behavior. A popular example is the bouncing ball. While *aerial* (not in contact with the ground), its dynamics can be described by the force of gravity acting on it. Upon contact with the *ground*, there is a normal force applied, preventing ground penetration and rapidly changing its velocity. By defining two modes, *aerial* and *ground* for the bouncing ball, we can model it as hybrid dynamical system and simulate its behavior. This differs from the real *continuous* time evolution of the bouncing ball which would be subject to a stiff differential equation as it approaches the ground and is repulsed by electromagnetic fields. At this small of a timescale, it is difficult to computationally simulate the interaction between the ball and the ground, in addition to likely not being accurate at predicting the bouncing motion of the ball [31].

An accurate computational model can be used to verify the safety and efficacy of robotic [23] or rehabilitative devices [2], termed engineered systems in this section. This is especially valuable when designing systems that tend to be costly to prototype and build. Early in the engineering design process, it can be beneficial to quickly iterate over various designs using a computational model, even before building basic prototypes. Furthermore, prototypes of engineered systems can be used to inform parameter choices or train a data driven-model, improving its veracity.

Although much of classical mathematics and physics deals with smoothness observed in the natural world, the dynamics of real-world systems can abruptly change. This is explained by chaos theory (catastrophe theory more specifically), where small variations in parameters can lead to bifurcations and changes in stability [123]. Hybrid dynamical systems provide

a model structure that matches the intrinsic behavior frequently observed in physical processes, *e.g.* the abrupt change in velocities of colliding bodies. Data-driven methods give us the ability to learn from experience [42, 109]. Through the incorporation of data-driven methods into hybrid dynamical systems, this thesis seeks to further their practical application – specifically modeling and controlling robotic systems and by optimizing rehabilitative outcomes associated with AFO prescription.

1.1 Summary of contributions

My thesis consists of findings from two research projects and their unified contributions to data-driven modeling for hybrid dynamical systems. The first centers on the control of mechanical systems subject to unilateral constraints (a modeling framework commonly used to represent legged locomotion). This project is in conjunction with Prof. Sam Burden. This work was motivated by the observation that many popular gradient–based optimization approaches for the control of robots assume smoothness of the dynamics for a given maneuver [120] [117] [97]. We demonstrated that nonsmoothness can cause standard optimization techniques to lose convergence guarantees and in practice contribute to poor performance for the resulting control policy. We have shown that the number of unique contact sequences increases with the number of periodic cycles for a mechanical system operating near a limit cycle. This effects propagates the nonsmoothness of trajectory outcomes, breaking up previously smooth sections in the cost landscape, stifling attempts to locally optimize.

The second project seeks to predict rhythmic, human locomotion with a motive to improve the clinical prescription of Ankle Foot Orthoses (AFO). This work is in collaboration with Michael Rosenberg, Momona Yamagami, Prof. Sam Burden, and Prof. Kat Steele. Using data-driven techniques, we created subject-specific models that can predict how an individual will respond to an untested AFO torque profile. To train and test the models selected for comparison, we collected a dataset of individuals walking on a treadmill with bilateral AFOs. The participants were subjected to a range of AFO stiffnesses, providing clinically viable torque profiles. The prediction task was to interpolate for a given stiffness (a value not

contained within the training range). The content in this document concerns able-bodied individuals. Our initial results lay the groundwork for future research to extend these results to patient populations, which are likely to be more challenging to accurately predict as they display greater inter-individual gait variability [57].

1.2 Hybrid dynamical systems

The hybrid dynamical systems framework, defined later in this section, can be applied to model the legged locomotion of rigid and compliant legged robots, animals, and humans [67] [47] [50]. For such systems to transverse terrain, they must interact dynamically with the ground. A popular way to model contact-rich dynamics, legged locomotion, is to assign a distinct mode, continuous domain with associated dynamics, for each contact configuration [84] [112]. To familiarize the reader with hybrid dynamical systems, I will briefly cover the framework used throughout this dissertation.

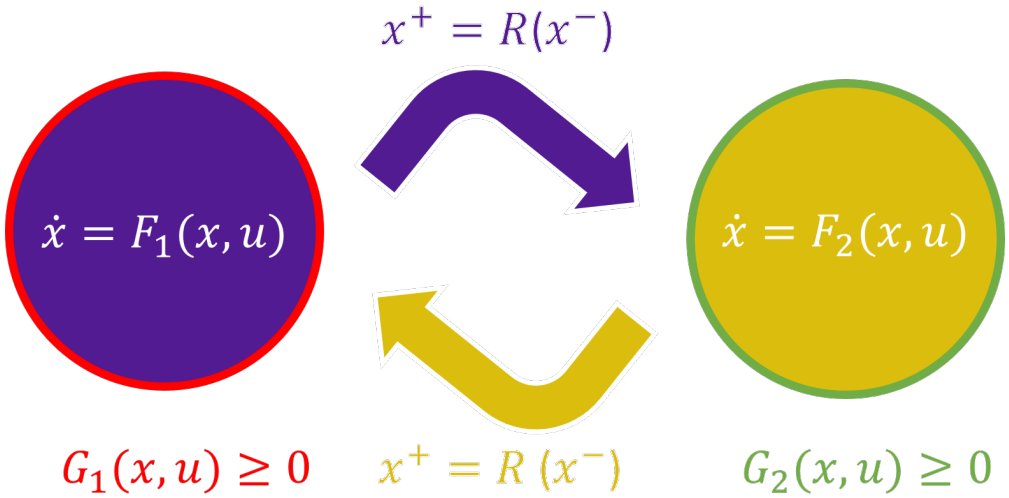


Figure 1.1: An example hybrid dynamical system with two modes.

This thesis utilizes a hybrid dynamical systems framework developed for the study of legged locomotion [67]. The framework defines a hybrid dynamical system as a tuple of mathematical objects $H = (D, U, F, G, R)$. $D = \amalg_{j \in J} D_j$ are the continuous domains where

each $D_j = \mathbb{R}^n$. $F_j : D \times U \to TD$ gives the vector field for each domain, $\dot{x} = F(x, u)$, defining how the system's state will evolve over continuous time. The space of control inputs is $u \in U = \mathbb{R}^m$. $G \subset D$ is the guard set, which indicates a trajectory has reached the boundary of the continuous domain and initiates the application of the reset function. The reset function is defined as $R : G \to D$, mapping a trajectory from the guard to another point in any of the modes. The reset function can take trajectories between modes. This is an overview of the class of hybrid dynamical systems that will be considered throughout the remainder of the dissertation.

Hybrid dynamical systems is an active area of research, particularly with respect to understanding the stability, asymptotic behavior, and control of a system, either along a trajectory or globally. In addition to modeling contact-rich dynamics, hybrid dynamical systems have been used in biology and medicine to model the behavior of neurons, genes, and cells [2]. Although hybrid dynamical systems appear effective at modeling real-world systems there are challenges pertaining to their study and implementation. Hybrid dynamical systems builds upon the study of dynamical systems, but many findings have still not been extended [53]. This means that there is less developed theoretical understanding and simulation software is tailored to a specific area of interest. This may explain why relaxing unilateral constraints, assuming universal smoothness, and then solving the resulting more tractable optimization problems is popular in graphics and robotics [93]. As will be discussed later, although these methods display desired behavior in simulation, they may introduce pathologies through this pernicious approximation giving rise to the gap in performance for simulated versus actual systems [27].

1.3 Data-driven modeling

Improvements in wireless networks, low-power sensors, and chips allow us to now ubiquitously record data [7]. We can now sense and track individuals and robots with a practical and off the shelf approach [114]. This gives us the capability to quickly collect rich data-sets that may train models which generalize better than ones trained on highly constrained experiments.

The proliferation of data collected on physical systems in tandem with our ever-growing computational power lays the foundation for the advancement of data-driven modeling and control [26] [117]. There are new methods of aggregation and computation on extensive and rich data-sets, driven by advancements in the fields of distributed and networked computing giving way to a new and accessible scale of computational resources to tackle a variety of problems. These tools have advanced the design of mechanical systems, especially those with a large prototype cost, *e.g.* nuclear reactors [86] and airplanes [25].

Deriving predictive models from first principles is a tedious ad-hoc process of trial and error that largely relies on human ingenuity. Phenomenological data-driven modeling techniques impose minimal pre-determined analytical structure and instead leverage large empirical datasets collected from the physical system. Since the resulting data-driven modeling tools rely less on domain-specific features or prior knowledge than other approaches to modeling, they may more readily generalize across disciplinary boundaries although their predictions are limited by the richness of the data-set. Hybrid dynamical systems can bridge the gap between a control algorithms simulated behavior and realization when implemented on a physical system. Opportunities for data-driven modeling extend to rehabilitative medicine, as current practice relies on observing and tuning rehabilitative devices in laboratory or clinical settings versus at home usage. Perhaps in the future, collecting data on subjects throughout the day and their evolving gait patterns may yield continuously adaptive devices that target and remedy pathologies [36].

Chapter 2
BACKGROUND

2.1 Introduction

In this chapter, I will cover prior research relevant to my thesis contributions. Due to the interdisciplinary nature of my research, the background section draws broadly upon control theory with a focus on robotics and biomechanics – particularly human locomotion. The first subsection of the background covers concepts in robotics and simulation of mechanical systems subject to unilateral constraints (Section 2.2). Prior research on biologically-inspired robotics, (Section 2.3), and reduced order modeling, (Section 2.4), are both topics from the first half of my PhD studies which influenced research directions and approaches later pursued. In the latter half of the background section, the focus shifts towards experiments from biomechanics, (Section 2.6), and ankle foot orthoses, exoskeleton, and prosthetic research and development, (Section 2.7). These areas all present opportunities for the application of data-driven hybrid dynamical systems, which can be applied to model existing systems and enhance control in the following applications: robots and their physical interaction with their environment, learning from biology to better engineer systems, control by reducing complexity, and rehabilitating or augmenting the human body.

2.2 Rigid & compliant robotics

In 2015 the DARPA Robotics Challenge inspired sensational videos of robots struggling to open doors and turn valves while occasionally falling down for no apparent reason [68]. The robotics community grappled with the outcome of the DARPA robotics challenge with some viewing the competition as an explicit failure [9]. Substantial time and money was invested mainly leading to the proliferation of online memes (satirical humor), deftly exposing the

limitations of a field. Many others in the robotics field, myself included, believe it served as a public benchmark for a host of sensing, estimations, and control techniques [73]. The slow and measured movement of the rigid robots, where they cautiously executed each maneuver, indicated they did not harness their dynamics to achieve task performance. The challenge provided insight into current capabilities and opportunities for improvements, the disaster relief tasks were navigated semi-autonomously, but slowly and with many failures. The challenge recalibrated expectations, stressed the value of software and hardware reliability, and underlined the challenge in transferring control from simulation to real [68].

Traditional research in the control of robotics focuses on forward and inverse kinematics, path planning, collision avoidance, and trajectory planning [121] [16]. These tools represent the foundations of robotics research. Recent research in learning, biologically-inspired robotics, medical robotics, locomotion, and manipulation bridges the gap between a variety of fields including statistics, computer science, biology, medicine, and mechanical design. Robotics is truly an integrative subject of study, drawing upon the advancements of other fields along with its own foundational research. One complementary area of study is legged terrestrial locomotion where models of humans and otherwise, can naturally be expressed with hybrid dynamical systems. Each contact configuration specifies a domain of dynamics, enumerating over all possible configurations [55]. Classical, calculus based, techniques apply to a single domain, it remains an open question on how to optimize hybrid dynamical systems for an unknown sequence of domains.

Videos of unpiloted UAVs performing acrobatics [1] and towel folding PR2s [80] showcased newfound abilities to perform dynamic maneuvers without explicitly coded control laws. The robots were trained and then allowed to perform autonomously. Although offering stunning demos that in turn generated excitement, there has been a failure to robustly extend these techniques [9]. Failure to generalize these successes can be due to a multitude of reasons — testing in highly favorable and controlled conditions, cherry picking successful instances over many failed experimental setups, hand-tuning by engineers resulting in specialized solutions, and a lack of fundamental understanding of the dynamics leading to a misapplication of RL

techniques. One response, still rudimentary and not covered in my document, is to make these methods more general and adaptable, deemed meta-learning or learning to learn [129]. An example of meta-learning research is to make deep reinforcement learning learn policies for multiple tasks in tandem without supervision [43].

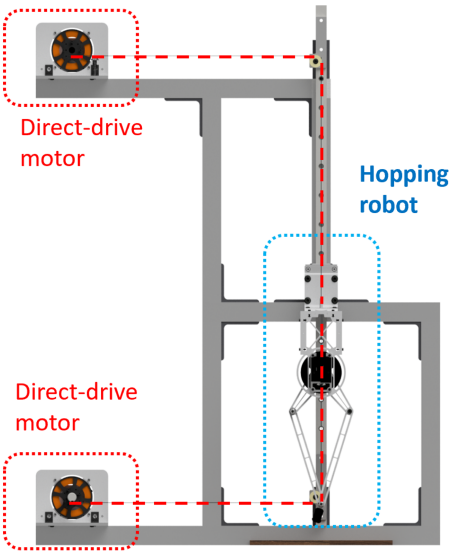


Figure 2.1: Constrained one-legged hopper.

The study of rigid body impacts can directly translate to the control of robots. Various factors determine the post impact state of a rigid body, among them friction and elasticity [132]. Simplifying assumptions on the impact makes control less complicated and more theoretically tractable. Frictionless contacts and plastic impacts are some of the modeling assumptions made due to the available theory for rigid systems. By better understanding the impact, *i.e.* a foot hitting the ground, we can develop models which predict motion in contact-rich regimes *e.g.* locomotion and manipulation. Some analytical contact models consider geometry and angle of impact [42]; new data-driven methods can allow us to fit models which may generalize better [88]. Limiting the possible contact configurations between the grasper and object and then utilizing learning techniques can overcome challenges posed by uncertainty in pose and contact dynamics proving performant in picking tasks [30].

A similar vein of research undertaken by Prof. Sam Burden and student researchers is to better understand, quantify, and predict the contact forces that arise from impact on the ground. The constrained one-legged hopper, as seen in Figure 2.1, serves as a scientific testbed. Motion is restricted to the vertical plane and the angle of contact is allowed to vary. After using a motion capture to collect data on the impact transition undergone by the hopper we can then develop data-driven hybrid dynamical systems models of contact for a single leg. Future work would then investigate multi-leg impact scenarios.

Compliant legged robotics is an active area of research in both academia and industry. StarletETH, developed at ETH Zurich, is a quadrupedal robot with series elastic actuators and serves a platform for the study of dynamic maneuvers [60]. Ghost robotics is a start-up, the individual legs from their robots are the basis of our lab research platform, seen in Figure 2.1. The most recent version of this robot, Minitaur, introduced compliance through backdrivability by using direct drive motors. Rigidity in robots simplifies positional control; while adding compliance introduces additional control complexity [131]. Compliance endows the robot with robustness to disturbances and makes its interaction with their environment safer for itself and surroundings [61]. Biological systems leverage compliance, offloading portions of feedback control to the feedforward mechanics [47]. This contrasts with the rigidity of contemporary industrial robots. We can introduce compliance through multiple mechanisms, backdrivable motors, softening contacts (*i.e.* end-effectors and feet) and changes in the structural mechanical design, introducing springiness or damping into an otherwise rigid system. A data-driven approach may offer a way to model the dynamics introduced by compliance and in design of the corresponding controller. Ultimately, adding compliance may help make robots more ubiquitous by improving robustness to errors in perception and planning [113], lowering power requirements [128], and by making robots safer to operate in the vicinity of humans [4].

2.3 Biologically-inspired robotics

The design of engineering systems can draw upon biology for existing and effective solutions. Many possible, unconventional morphologies seen in biological organisms can transfer to robots, including snakes [134] and undulating sandfish lizards [87]! X-RHex is a legged robot inspired by the cockroach [48]. Simplifying active control and instead relying on the system dynamics, feedforward control, to achieve task performance [59] [6]. As researchers, we would like to the harness and translate to robots the mechanisms of biological systems that makes them performant and versatile.

Locomotive capabilities observed in organisms are an aggregated result of the rapid mechanical preflexes, cyclic feedforward neural control, and slower feedback control mediated by sensed deviation from a desired behavior [35]. The interplay between these different mechanisms of control is complex and nuanced but gives rise to the performant and versatile behaviors seen in biological systems. In the case of terrestrial locomotion, measuring the ground reaction forces experienced when a leg comes into contact with the surface of the ground and understanding the impact the system experiences is quite challenging. The almost instantaneous impact and resulting rapid change in system state, makes it difficult to accurately observe and model. Contact forces experienced by a biological or robot system stem from the interaction between its intrinsic properties and the media its coming into contact with. Different dynamics dominate based on the media and interaction. Moving through granular media can be described with terradynamics [82] and swimming in water predicted by solving the Navier-Stokes equations [24]. Both are examples where the locomotor-substrate dynamics can be described by a set of equations and solved numerically if we overcome the challenge of observing and updating the state of hybrid systems as they switch contact configurations during legged locomotion.

2.4 Control using reduced-order models

One modeling paradigm applicable to biological systems and robotics involves *templates* and *anchors* [47]. Templates are simple models that represent a system's target behavior. They naturally embed in anchors, models that more faithfully represent the complexity observed in the actual mechanical (joints, muscles and tendons) and neural (preflexes, reflexes, task specification) systems [82]. Passing between templates and anchors entails moving up and down a hierarchy of physical fidelity, achieving different aims. At the abstract end, templates encode mechanistic understanding and can be used to synthesize new control strategies [115], facilitating translation between application domains, *e.g.* from simulation to a physical system. In a domain of interest, abstract mechanisms encoded in a template can be instantiated in an application-appropriate anchor, yielding concrete control techniques [83].

Spring loaded inverted pendulum (SLIP) is a fundamental template model in locomotion research, both in humans, robots and otherwise [49]. It has been used to model the center of mass motion for walking, running, and other legged gaits [51] [115]. Work has demonstrated stability properties for both a walking and running SLIP, an unintuitive result for such a rudimentary model [49]. In addition to modeling bipedal motion, SLIP is used for the time evolution of the center of mass for legged systems by combing the effect of the leg contacts with the ground into a single spring leg [52]. Instead of applying a single controller to a system, another approach decomposes control effort over multiple-levels. The low-level controller acts to make a component, the leg, behave like a target template, *e.g.* SLIP. Then a high-level controller is applied to the resulting virtual system; engineers can then utilize already developed controllers for the template giving greater modularity to control while also preventing duplication of effort [62]. 3D bipedal walkers can also be modeled as an hybrid dynamical systems, an analytical control law was found by using energy-based feedback control, decoupling the sagittal and lateral elements of walking. The stability of the resulting gait is then verified numerically [6]. Hybrid dynamical systems have also been used to model and design a passive 3D walker, where energy was injected into the system

through gravitational forces (the walker went downhill) [125]. The system was built for a set of initial conditions whose resulting trajectories converge towards a stable walking limit cycle. Principled design utilizing theoretical analysis and computational simulations was used to create a robust and stable system contrasting with typical ad-hoc and iterative design tactics. These results suggest that robots entirely governed by feedforward dynamics can achieve certain desirable maneuvers. Stabilizing feedforward dynamics are attractive as they can be more efficient, stable at small timescales, and can simplify feedback control. In **Aim 1**, we show the challenge in applying gradient based optimization techniques to reduced-order models for the legged locomotion of mechanical systems.

Dead-beat control, discontinuous changes in the input applied, can correct for deviations from a desired gait within a single stride. Dead-beat control may allow systems to rapidly reject disturbances. Research has shown that a 3-D walker can move stably across terrain, without sensing ground height, by controlling the length of its leg [135]. For a system moving through uneven or rough terrain simple dead-beat control reduces the reliance on accurate sensing and endowing a level of inherent robustness and maneuverability. Overall, searching for and optimizing gaits remains an active area of research [136] [118]. Methods for control in a single domain have not been extended to hybrid dynamical systems if the sequence of domain changes is unknown.

2.5 Reinforcement learning

Reinforcement learning (RL) algorithms, a branch of optimal control where the agent is capable of learning a policy by experiencing its environment, are a popular contemporary approach to controlling robotic systems [1, 80]. Leveraging the ever-improving ability to collect data and perform computational analysis, RL in practice utilizes concepts from control theory and information contained in data to improve performance and decrease reliance on human designers. In **Aim 1**, we will focus on policy gradient methods, a form of RL particularly attractive in robotics due to a large state and action space. Policy gradient methods minimize the cost function by varying parameters of the policy [120].

2.6 Modeling biomechanics

Robotics has still not replicated the locomotive capabilities of humans. This is likely due to three reasons: (1) limitations in the capabilities of the torque to weight ratio of actuators combined with the fidelity of the structure and its ability to withstand stress, (2) contemporary control strategies are not at the level of performance of biological counterparts, and (3) inefficiency in engineered systems which is most likely due to a lack of effective compliance usage [4, 70, 118]. There are many approaches to finding gaits for systems: tinkering [29], analytical analysis [50], and computer optimization [122] are all employed. Detailed, anchored models of the human musculoskeletal system are used to simulate walking and running [8]. Based on the geometric layout of the skeleton, individual Hill-type muscle models are used to generate forces that are based on empirical measurements and models of muscle activation and tendon properties [133]. Numerical simulations of reduced order models such as SLIP take into account the hybrid dynamics of legged locomotion, when optimized for energy efficiency at a specified speed, converging to a straight legged rolling center of mass motion similar to walking at low speeds and an impulsive bounce comparable to running at high speeds [122]. Although there are many possible solutions, there is only one optimal solution from a cost function perspective. Escaping local minima and moving towards a better solution is challenging in a nonlinear cost landscape and requires active intervention in humans [127].

In the experiment described in **Aim 2**, AFOs are treated as perturbations, altering the limit cycle of an individual pushing them towards a more optimal behavior. For a rhythmic task, a mechanical system may evolve according a lower dimension dynamical system [20]. This likely applies to terrestrial locomotion, walking and running, and suggests that simple models can capture the dominant dynamics of this reduced-order behavior. Phaser estimate the phase of a system engaged in a rhythmic task or motion along a stable limit cycle, takes in multiple noisy measurements and estimates their phase individually [108]. It then reduces the estimated phase variance across all measurements, returning an estimate of the

instantaneous phase for the whole system [107]. This method differs from conventional event-driven phase estimation techniques in locomotion, which linearly interpolate between contact events, and more sensitively accounts for perturbations to the system that occur within the gait cycle.

SLIP is used to model running for a variety of systems: humans, animals, and robots [50] [47] [62]. Stable gaits have been found for these systems using fixed-leg reset policies, which are instantaneous changes in leg length after lift-off [52]. Follow-up work shows that for biologically plausible parameter ranges the aforementioned control yield stable gaits [50]. The characteristics of mechanical components give rise to the emergent dynamics of locomotion for a system. Various parameters of the SLIP model, stiffness and leg-length reset policy, attempt to represent these characteristics [51]. By varying the parameters of SLIP, leg stiffness and angle of touchdown, there has been work into understanding the principles of running in humans. It was found that running is self-stabilizing and work has established a set of parameter values, experimentally validated in humans [119]. Research also shows that are many available solutions available for gaits in SLIP, alluding to the notion that in more sophisticated systems, *e.g.* humans, there may be multiple stable gaits available as well. These results validate the use of reduced-order models in the analysis of biological organisms with an eventual aim of transferring findings to engineered systems.

Passive, dynamic walkers can achieve periodic gaits with no feedback control, the walkers move down a ramp around a stable limit cycle. One of the goals of building these systems is to find elements of mechanical design, witnessed in ourselves and in other organisms, that enable passive locomotion. A successful example was first demonstrated in 2D [90] and then extended to 3D [29]. Both illustrate advancements in designing the dynamics of mechanical systems to achieve task performance versus assuming quasi-static motion, a common assumption in system design and control. Also this examines, the degree to which humans rely on neural feedback and thus corrective actuation, versus our natural feedforward dynamics partially mediated by our mechanics. This is relevant towards building more energetically efficient systems particularly due to the limitations of energy storage

techniques. There is an effort to better understand the stabilizing mechanisms of walking. Simple controller choices such as fore-aft foot placement have a large effect on lateral stability [13]. Although humans harness passive dynamics for walking, ultimately a degree of active control is still necessary to realize human gait. Simple models, such as 2D walker, can help contrast differences in performance between various methods of actuation. For the 2D walker, energy can be injected just before toe-off by the stance leg or by applying a torque through the stance hip [76]. The 2D walker can serve as a testbed to reduce a dynamic and refined movement to base principles, with an aim to then eventually extend findings to a more complex one. One motivating application for the study of 2D walker models is predicting AFO interventions. By changing the gait dynamics and therefore altering the gait pattern of a subject with an AFO, we can inject stabilizing and efficiency properties.

The Lateral Leg Spring (LLS) is similar to the SLIP model except that it models the motion of the center of mass in the horizontal plane. The LLS displays a variety of stabilizing behaviors and has been applied to the study of cockroaches [115]. LLS folds the contributions of all the individual legs, for a multilegged stance phase, into a single virtual leg. This reduction is representative when applied to cockroaches, since the net force applied by the legs is zero upon touchdown and again at liftoff for each tripod stance. The feedforward dynamics and the resultant stability of the cockroach is a product of its mechanical properties, which allows them to rapidly reject perturbations as they move through rough terrain. The walking motion of the individuals is similarly modeled by a version of SLIP [89]. In **Aim 2** we only consider steady-state walking, where the perturbation is applied by the device. Further research would broaden the scope of experimental consideration to a variety of AFO walking conditions and perturbations.

2.7 Ankle foot orthoses, exoskeletons, & prosthetics

A segment of the population would benefit from advances in prostheses, artificial limbs, and orthotics (or exoskeletons a relatively synonymous term that refers to an augmentative or more encompassing design) [37]. The development of active orthosis first started in 1935,

simultaneously in the United States and Yugoslavia, and has started in many other countries and institutions [37]. Modern prosthetic and orthotic research continues with rehabilitative, assistive, and augmentative aims in addition to their enabling technologies; energy storage, actuation, fabrication and control. Amputees with passive ankle-foot prostheses walk less efficiently than their able bodied counterparts and display gait asymmetry [10]. There are various trade-offs between powered and passive prostheses. The ankle applies positive work when in contact with the ground something a powered protheses can mimic [28]. Powered prosthetics have the ability to adapt more effectively to a variety of terrains and can help subjects address abnormalities – most likely due to previous use of prosthetics – in their gaits [40]. Currently though, available hardware still lacks the power density seen in the human ankle making the powered prosthetic ankle heavier; also the dynamics and control of the human ankle are still not fully understood making it not feasible to make a fully biomimetic prosthetic [58]. Passive AFO are more affordable and accessible, making them the current standard in clinical care. At slower walking speeds, a passive AFOs can mimic the energy storage and release of the Achilles tendon, at higher speeds the gait becomes asymmetric and energy efficient, individuals likely compensate for the lack of positive work done by the ankle with greater hip extension [40]. Recent results have excitingly reduced the energetic cost of walking in healthy individuals. For powered, ankle exoskeletons this improved efficiency was quickly achieved for a variety of gaits through an optimization procedure that iterated across various unimodal torque profiles [137]. The unpowered, ankle exoskeleton garnered efficiency improvements, by utilizing a mechanical clutch that engaged while the foot was in contact with the ground overcoming the limitation of muscles, which require metabolic energy for isometric and eccentric contractions [28].

Chapter 3

AIMS

3.1 Objective & Aims

The following aims outline the research direction and motivation for their respective research projects. The overarching goal of this research, and by extension thesis, is to further the study of hybrid dynamical systems in this new data-rich landscape, and explore new potential applications by leveraging theoretical results.

3.2 Statement & Aims

Thesis Statement: Further the study and application of data-driven hybrid dynamical systems

3.2.1 ***Aim 1*** *- Study the regularity of optimal policy and value functions for simulated mechanical systems subject to unilateral constraints*

Our work has shown that value and policy functions inherit the underlying regularity of trajectory outcomes for a given maneuver. This proves problematic for policy gradient methods, or any gradient-based optimization approach, as they rely upon smoothness assumptions in order to guarantee performance [120]. We created two computational examples to illustrate the formal findings; in one a box is dropped from a fixed height with plastic impact and frictionless contacts. This is a simple mechanical system subject to unilateral constraints that illustrates the phenomenon of nonsmooth trajectory outcomes. In the other computational example, a simple legged system is subject to a contact-mode dependent control law. This causes the system to also experience nonsmoothness: discontinuous, continuously-differentiable or piecewise-differentiable and not continuously-differentiable state outcomes.

These findings have consequences when performing optimization. We a computed a nonzero gradient step, using the Deterministic Policy Gradient algorithm, along the optimal policy for a maneuver with nonsmooth trajectory outcomes (a true gradient would be zero all along the optimal policy) [120]. Policy gradient methods operating in this regime lose convergence guarantees and may in practice move erratically and away from the optimum. Our findings indicate frequent misapplication of optimization approaches that make smoothness assumptions, constituting a first step towards altering and developing control strategies that will more successfully transfer from simulation to real systems. Lastly, we touch on the more specific case of mechanical systems subject to unilateral constraints operating near a limit cycle. We find that the amount of contact sequences increases with the number of cycles completed by the system. This indicates that local optimal optimization approaches would further struggle with an expanded control horizon, the fragmented cost landscape would decrease the reliability of computed gradients.

3.2.2 ***Aim 2*** *- Develop and survey models to predict the effect of an ankle foot orthosis from subject-specific data*

Currently, the prescription of AFOs is a tedious process of ad-hoc trial and error. Although there are clinical procedures available for evaluation and prescription, no standard protocol exists, and outcomes are inconsistent, even within a single clinical center [109]. The aim of my research is to provide a predictive tool that can help improve the prescription of AFOs and help clinicians optimize AFO interventions. Due to the cost of fabrication and clinical time required for each device iteration, many patients end up with sub-optimal solutions. This can prevent patient adoption and efficacy of the devices.

My work in this aim first establishes a baseline of predictive capabilities for a variety of models. Our prediction framework is as follows, given the current state of an individual rhythmically walking, predict a future state — the state is comprised of the kinematics, estimated kinetics, myoelectric response, and AFO torque profiles. The models are trained on clinically-practical amounts of data, a facet that led us to survey models whose predictive

abilities scale differently with training set sizes. We then predicted changes to gait kinematics and myoelectric response as the spring stiffness of the AFO was varied. To test the models, we predicted for a new AFO stiffness value contained within the range of training stiffnesses (interpolation), versus stiffnesses beyond the training range (extrapolation). The metrics of predictive performance, along with qualitative analysis, indicate which models may be suitable for future application in either a research or clinical setting. By varying the amount of data available for the models to train on, we sought to explore what training set size is necessary for the surveyed models. Higher data requirements make subject-specific models less viable in a clinical (expensive) setting. We discuss the functional complexity and generalizability of the surveyed data-driven models, and their predictive ability within the constraints of our experimental setup [46]. Finally, we examined the effects of the AFO torque profile, by separating the relevant coefficients in the linear model and their effect on the predicted values.

Part I

AIM 1

Chapter 4

NONSMOOTH OPTIMAL VALUE AND POLICY FUNCTIONS FOR MECHANICAL SYSTEMS SUBJECT TO UNILATERAL CONSTRAINTS

4.1 Introduction

In the optimization approach to robot control, a policy is sought that extremizes a given performance criterion; the performance achieved by this *optimal policy* is the *optimal value* of the problem. Two widely-applied frameworks for solving such problems are reinforcement learning [63, 80, 93] and trajectory optimization [99, 100, 104]. Many scalable algorithms in either framework leverage local approximations—gradients of values and policies—to iterate towards optimality. Mechanical models, that assume frictionless contacts and plastic impacts, balance fidelity with their physical counterparts versus analytical and computational tractability. Trajectories of the aforementioned class of models, over a continuously-differentiable range of initial conditions, can yield discontinuous, continuously-differentiable or piecewise-differentiable and not continuously differentiable state outcomes. This thesis contributes to controls, learning, and robotics with analytical findings and a computational example, that show value and policy functions inherit the underlying regularity of trajectory outcomes for a given maneuver. We further show that in cases of nonsmoothness these gradients can fail to exist for contact-rich robot dynamics, precluding application of state–of–the–art algorithms for optimal control.

We begin in Section 4.2 by modeling contact–rich robot dynamics using mechanical systems subject to unilateral constraints, and describe how nonsmoothness—discontintinuity or piecewise–differentiability—manifests in trajectory outcomes and (hence) trajectory costs. Then in Section 4.3 we provide mathematical derivations that show nonsmoothness in tra-

jectory outcomes and costs gives rise to nonsmoothness in optimal value and (hence) policy functions. Subsequently in Section 4.4 we present numerical simulations that demonstrate discontinuous or merely piecewise–differentiable optimal value and policy functions in a mechanical system subject to unilateral constraints. Finally in Section 4.5 we discuss the prevelance of nonsmoothness and how the lack of classical differentiability prevents gradient–based algorithms from converging to optimality.

4.2 Mechanical systems subject to unilateral constraints

In this section, we formalize a class of models for contact–rich dynamics in robot locomotion and manipulation as *mechanical systems subject to unilateral constraints* and formulate an optimal control problem for these systems.

4.2.1 Dynamics

Consider the dynamics of a mechanical system with configuration coordinates $q \in Q = \mathbb{R}^d$ subject to unilateral constraints $a(q) \geq 0$ specified by a differentiable function $a : Q \to \mathbb{R}^n$ where $d, n \in \mathbb{N}$ are finite. We are primarily interested in systems with $n > 1$ constraints, whence we regard the inequality $a(q) \geq 0$ as being enforced componentwise.

Given any $J \subset \{1, \ldots, n\}$, and letting $|J|$ denote the number of elements in the set J, we let $a_J : Q \to \mathbb{R}^{|J|}$ denote the function obtained by selecting the component functions of a indexed by J, and we regard the equality $a_J(q) = 0$ as being enforced componentwise.

It is well–known (cf. [11, Sec. 3] or [66, Sec. 2.4, 2.5]) that, with
$J = \{j \in \{1, \ldots, n\} : a_j(q) = 0\}$ denoting the *contact mode*, the system's dynamics take the form

$$M(q)\ddot{q} = f_J(q, \dot{q}, u) + c(q, \dot{q})\dot{q} + Da_J(q)^\top \lambda_J(q, \dot{q}, u), \tag{4.1a}$$

$$\dot{q}^+ = \Delta_J(q, \dot{q}^-), \tag{4.1b}$$

where $M : Q \to \mathbb{R}^{d \times d}$ specifies the mass matrix for the mechanical system in the q coordi-

nates, $f_J : TQ \to \mathbb{R}^d$ is termed the *effort map* [11] and specifies[1] the internal and applied forces, $u \in \mathcal{U}$ is an external input, $c : TQ \to \mathbb{R}^{d\times d}$ denotes the *Coriolis matrix* determinedby M, $Da_J : Q \to \mathbb{R}^{|J|\times d}$ denotes the (Jacobian) derivative of the constraint function a_J with respect to the coordinates, $\lambda_J : TQ \to \mathbb{R}^{|J|}$ denotes the reaction forces generated in contact mode J to enforce $a_J(q) \geq 0$, $\Delta_J : TQ \to \mathbb{R}^{d\times d}$ specifies the collision restitution law that instantaneously resets velocities to ensure compatibility with the constraint $a_J(q) = 0$, and $\dot{q}^+$ (resp. $\dot{q}^-$) denotes the right– (resp. left–)handed limits of the velocity with respect to time.

4.2.2 Regularity of dynamics

The seemingly benign equations in (4.1) can yield dynamics with a range of regularity properties. This issue has been thoroughly investigated elsewhere [11,95,96]; here we focus specifically on how design choices in a robot's *mechanical* and *control* systems affect regularity of its dynamics.

In what follows, we will frequently refer to the concept of a control system's *flow*, so we briefly review the concept before proceeding. Given a control system (e.g. (4.1) or (4.2)) with state space $\mathcal{X}$ and input space $\mathcal{U}$, a *flow* is a function $\phi : [0,t]\times\mathcal{X}\times\mathcal{U}^{[0,t]} \to \mathcal{X}$ such that for all initial states $x \in \mathcal{X}$ and inputs $(u : [0,t] \to \mathcal{U}) \in \mathcal{U}^{[0,t]}$, the function $\phi^{x,u} : [0,t] \to \mathcal{X}$ defined for all $s \in [0,t]$ by $\phi^{x,u}(s) = \phi(s,x,u)$ is a trajectory for the control system. Intuitively, the flow "bundles" all trajectories into a single function. Mathematically, the flow is useful for studying how trajectories vary with respect to initial states and inputs. So long as trajectories exist and are unique for every $x \in \mathcal{X}$ and $u \in \mathcal{U}^{[0,t]}$, the flow is a well–defined function.

It is common to assume that the functions in (4.1) are continuously–differentiable $(M, f, a, \gamma \in C^r)$; however, as illustrated by [11, Ex. 2], this assumption alone does not ensure existence or uniqueness of trajectories. This case contrasts starkly with that of classical

[1]We let $TQ = \mathbb{R}^d \times \mathbb{R}^d$ denote the *tangent bundle* of the configuration space Q; an element $(q,\dot{q}) \in TQ$ can be regarded as a pair of generalized configurations $q \in \mathbb{R}^d$ and velocities $\dot{q} \in \mathbb{R}^d$; we write $\dot{q} \in T_qQ$.

control systems, where the equation

$$\dot{x} = F(x, u) \tag{4.2}$$

yields unique trajectories whose regularity matches the vector field's: if F is continuously differentiable, then there exists a flow for (4.2) that is continuously differentiable to the same order.

To ensure trajectories for (4.1) exist uniquely, restrictions must be imposed; we refer the interested reader to [11, Thm. 10] for a specific result and [66] for a general discussion of this issue. Since we are chiefly concerned with how properties of the dynamics in (4.1) affect properties of optimal value and policy functions, we will assume in what follows that conditions have been imposed to ensure (4.1) has a flow for states, inputs, and time horizons of interest.

Assuming that a flow ϕ exists for (4.1) does not provide any regularity properties on the function ϕ; these properties are determined by the design of a robot's mechanical and control systems and their closed–loop interaction with the environment. For instance: when limbs are inertially coupled (e.g. by rigid struts and joints), so that one limb's constraint activation instantaneously changes another's velocity, ϕ is discontinuous at configurations where these two limbs activate constraints simultaneously [105, Table 3] [64]; when limbs are force coupled (e.g. by nonlinear damping), so that one limb's constraint (de)activation instantaneously changes the force on another, ϕ can be piecewise–differentiable at configurations where these two limbs (de)activate constraints simultaneously [96, Fig. 1]. In both instances, mechanical design choices lead to nonsmooth dynamics; Figure 4.1 provides examples where control design choices lead to nonsmooth dynamics (piecewise–differentiable in Figure 4.1(a,c,e), discontinuous in Figure 4.1(b,d,f)). Other nonsmooth phenomena can arise, e.g. *grazing*[2] and *Zeno*[3] trajectories; in what follows we will focus on the case of simultaneous constraint (de)activations due to its prevalence in robot gaits and maneuvers (see Section 4.5.1 for a

[2]where a constraint function a_j decreases to and then increases from zero without activating constraint j

[3]where a constraint is activated an infinite number of times on a finite time horizon

discussion of when this phenomena prevails).

4.2.3 Regularity of optimal value and policy functions

A broad class of optimal control problems for the dynamics in (4.1) can be formulated in terms of *final* ($\ell : \mathcal{X} \to \mathbb{R}$) and *running* ($\mathcal{L} : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$) costs:

$$\nu(x) = \min_{u \in \mathcal{U}^{[0,t]}} \ell(x^u(t)) + \int_0^t \mathcal{L}(x^u(s), u(s))\, ds, \tag{4.3}$$

where $x^u : [0,t] \to \mathcal{X}$ denotes the unique trajectory obtained from initial state $x^u(0) \in \mathcal{X}$ when input $u \in \mathcal{U}^{[0,t]}$ is applied; in terms of the flow, $x^u(s) = \phi(s, x(0), u)$ for all $s \in [0,t]$. To expose the dependence of the cost in (4.3) on the flow ϕ, we transcribe the problem in (4.3) to a simpler form using a standard state augmentation technique (cf. [98, Ch. 4.1.2]):

$$\nu(x) = \min_{u \in \mathcal{U}^{[0,t]}} c\left(\phi(t, x, u)\right) \tag{4.4}$$

As discussed in Section 4.2.2, the continuity and differentiability properties of ϕ are partly determined by a robot's design: it is possible for ϕ and hence $c \circ \phi$ to be discontinuous ($\phi \notin C^0$), continuously–differentiable ($\phi \in C^r$), or piecewise–differentiable and not continuously–differentiable ($\phi \in PC^r \setminus C^r$), depending on the properties of the robot's mechanical and control systems. In the next section, we study how continuity and differentiability properties of $c \circ \phi$ affect the corresponding properties of ν in (4.4).

4.3 Continuity and differentiability of optimal value and policy functions

Consider minimization of the cost function $c : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ with respect to an input $u \in \mathcal{U}$:

$$\nu(x) = \min_{u \in \mathcal{U}} c(x, u); \tag{4.5}$$

so long as $\mathcal{X}$ and $\mathcal{U}$ are compact and c is continuous, the function $\nu : \mathcal{X} \to \mathbb{R}$ indicated in (4.5), termed the *optimal value function*, is well–defined. We let $\pi : \mathcal{X} \to \mathcal{U}$ denote an *optimal policy* for (4.5), i.e.

$$\forall x \in \mathcal{X} : \pi(x) \in \arg\min_{u \in \mathcal{U}} c(x, u) \tag{4.6}$$

or, equivalently,

$$\forall x \in \mathcal{X} : \nu(x) = c(x, \pi(x)). \tag{4.7}$$

In this section we study how regularity properties (continuity, differentiability) of the cost function (c) relate to regularity properties of optimal value (ν) and policy (π) functions.

4.3.1 Discontinuous cost functions

If the cost ($c : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$) is discontinuous with respect to its first argument, then the optimal policy ($\pi : \mathcal{X} \to \mathcal{U}$) and value ($\nu : \mathcal{X} \to \mathbb{R}$) are generally discontinuous as well. This observation is clear in the trivial case that the cost only depends on its first argument, but manifests more generally.

4.3.2 Continuously–differentiable cost functions

This section contains straightforward calculations based on standard results in classical (smooth) Calculus and nonlinear programming; it is provided primarily as a rehearsal for the more general setting considered in the subsequent section.

If c is continuously–differentiable, denoted $c \in C^1(\mathcal{X} \times \mathcal{U}, \mathbb{R})$ or simply $c \in C^1$, then necessarily [98, Ch. 1.1.1]

$$D_2 c(x, \pi(x)) = 0. \tag{4.8}$$

If c is two times continuously–differentiable (denoted $c \in C^2$) and the second–order sufficient condition [98, Ch. 1.1.2] for strict local optimality for (4.5) is satisfied at $\pi(x) \in \mathcal{U}$,

$$D_2^2 c(x, \pi(x)) > 0, \tag{4.9}$$

then the C^1 Implicit Function Theorem (IFT) [78, Thm. C.40] can be applied to (4.7) to choose π as a C^1 function near x. Note that

IFT specifically required the invertibility tacit in (4.9):

$$\text{the linear function } D_2^2 c(x, \pi(x)) : T_u\mathcal{U} \to T_u\mathcal{U} \text{ is invertible.} \tag{4.10}$$

If (4.8) and (4.9) are satisfied, then applying the C^1 Chain Rule [78, Prop. C.3] to (4.8) yields

$$D\pi(x) = -D_2^2 c(x, \pi(x))^{-1} \left(D_{12} c(x, \pi(x))\right), \tag{4.11}$$

and applying the C^1 Chain Rule to (4.7) yields

$$\begin{aligned} D\nu(x) &= D_x c(x, \pi(x)) \\ &= D_1 c(x, \pi(x)) + D_2 c(x, \pi(x)) D\pi(x), \end{aligned} \tag{4.12}$$

whence we obtain derivatives of the optimal value and policy functions in terms of derivatives of the cost function.

We conclude that if the cost function is two times continuously–differentiable ($c \in C^2$) and first–order necessary (4.8) and second–order sufficient (4.9), (4.10) conditions for optimality and stability of solutions to (4.5) are satisfied at $u = \pi(x)$, then the optimal policy and value functions are continuously–differentiable at x ($\pi, \nu \in C^1$) and their derivatives at x can be computed using (4.11), (4.12).

Theorem 1 *Consider a mechanical system with dynamics given in* (4.1) *and with a value function as defined in* (4.4). *If* $c\,(\phi) \in C^2(\mathcal{X} \times \mathcal{U}, \mathbb{R})$ *satisfies* (4.8), (4.9), *and* (4.10) *at* $(\xi, \mu) \in \mathcal{X} \times \mathcal{U}$, *then there exist neighborhoods* $X \subset \mathcal{X}$ *of* ξ *and* $U \subset \mathcal{U}$ *of* μ *and a function* $\pi \in C^1(X, U)$ *such that* $\pi(\xi) = \mu$ *and, for all* $x \in X$, $\pi(x)$ *is the unique minimizer for* (4.5)

the derivative of π *is given by* (4.11), *and the derivative of* ν *is given by* (4.12).

4.3.3 Piecewise–differentiable cost functions

If c is piecewise–differentiable,[4] denoted $c \in PC^1(\mathcal{X} \times \mathcal{U}, \mathbb{R})$ or simply $c \in PC^1$, then necessarily

$$\forall w \in T_u\mathcal{U} : D_2 c(x, \pi(x); w) \geq 0. \tag{4.13}$$

[4]We use the notion of piecewise–differentiability from [116, Ch. 4.1]: a function is piecewise–differentiable if it is everywhere locally a continuous selection of a finite number of continuously–differentiable functions.

Here and below, $D_2c(x,\pi(x)) : T_u\mathcal{U} \to \mathbb{R}$ denotes a continuous and piecewise–linear first–order approximation termed the *Bouligand* (or *B*–)derivative [116, Ch. 3] that exists by virtue of the cost being PC^1 [116, Lem. 4.1.3]; $D_2c(x,\pi(x);w)$ denotes the evaluation of $D_2c(x,\pi(x))$ at $w \in T_u\mathcal{U}$.

If c is two times piecewise–differentiable (denoted $c \in PC^2$), for (4.5) is satisfied at $\pi(x) \in \mathcal{U}$,

$$\begin{aligned} \forall w \in \{w \in T_u\mathcal{U} \mid w \neq 0,\ D_2c(x,\pi(x);w) = 0\} \\ : D_2^2c(x,\pi(x);w,w) > 0, \end{aligned} \tag{4.14}$$

and if the piecewise–linear function

$$D_2^2c(x,\pi(x)) : T_u\mathcal{U} \to T_u\mathcal{U} \text{ is invertible,} \tag{4.15}$$

then a PC^1 Implicit Function Theorem can be applied to choose $\pi \in PC^1$ near x [111, Cor. 3.4].[5] Applying the PC^1 Chain Rule [116, Thm. 3.1.1] to (4.13) yields (cf. [111, § 3])

$$\begin{aligned} \forall v \in & T_x\mathcal{X} : D\pi(x;v) = \\ & - D_2^2c(x,\pi(x))^{-1} \left(D_{12}c(x,\pi(x);v)\right), \end{aligned} \tag{4.16}$$

and applying the PC^1 Chain Rule to (4.7) yields

$$\begin{aligned} \forall v \in & T_x\mathcal{X} : D\nu(x;v) = D_xc(x,\pi(x);v) \\ & = D_1c(x,\pi(x);v) + D_2c(x,\pi(x);D\pi(x;v)), \end{aligned} \tag{4.17}$$

whence we obtain B–derivatives of the optimal value and policy functions in terms of B–derivatives of the cost.

We conclude that if the cost function is two times piecewise–differentiable ($c \in PC^2$) and first–order necessary (4.13) and second–order sufficient (4.14), (4.15) conditions for optimality and stability of solutions to (4.5) are satisfied at $u = \pi(x)$, then the optimal policy and value functions are piecewise–differentiable at x ($\pi, \nu \in PC^1$) and their B–derivatives at x can be computed using (4.16), (4.17).

---

[5] This Implicit Function Theorem requires D_2c be *strongly* B–differentiable; the costs considered here are not generally strongly B–differentiable, but they are generally PC^r–equivalent to strongly B–differentiable functions [75, Thm. 3.1], whence [111, Cor. 3.4] can be applied indirectly.

Theorem 2 *Consider a mechanical system with dynamics given in* (4.1) *and with a value function as defined in* (4.4). *If* $c(\phi) \in PC^2(\mathcal{X} \times \mathcal{U}, \mathbb{R})$ *satisfies* (4.13), (4.14), *and* (4.15) *at* $(\xi, \mu) \in \mathcal{X} \times \mathcal{U}$, *then there exist neighborhoods* $X \subset \mathcal{X}$ *of* ξ *and* $U \subset \mathcal{U}$ *of* μ *and a function* $\pi \in PC^1(X, U)$ *such that* $\pi(\xi) = \mu$ *and, for all* $x \in X$, $\pi(x)$ *is the unique minimizer for* (4.5)

the B–derivative of π *is given by* (4.16), *and the B–derivative of* ν *is given by* (4.17).

4.3.4 Conclusions regarding regularity of optimal value and policy functions

The results in Sections 4.3.1, 4.3.2, and 4.3.3 suggest that we should generally expect regularity of optimal value and policy functions to match that of the cost function: they should be discontinuous when the cost is discontinuous, or piecewise–differentiable when the cost is piecewise–differentiable. In Section 4.4 we provide instances of the class of models described in Section 4.2 that exhibit these effects.

4.4 Optimal value and policy functions for a mechanical system subject to unilateral constraints

We showed in the previous section that optimal value and policy functions for contact–rich robot dynamics inherit nonsmoothness from the underlying dynamics. To instantiate this result, we crafted the simplest mechanical system subject to unilateral constraints that exhibits the nonsmooth phenomena of interest (piecewise–differentiable or discontinuous trajectory outcomes), yielding the *touchdown* and *liftoff* maneuvers shown in Figure 4.1(a,b). For the touchdown maneuver, we seek the optimal (constant) force to exert in the left leg (u_1) when the left foot is in contact and the right foot is not; similarly, we seek the optimal choice of force in the right leg (u_2) when the right foot is in contact and the left foot is not: with $a_1, a_2 > 0$ as input penalty parameters,

$$c_{\text{touchdown}}(\theta_{\text{nadir}}, u_1, u_2) = (\theta_{\text{nadir}} - \theta_{\text{desired}})^2 + a_1 u_1^2 + a_2 u_2^2. \tag{4.18}$$

For the liftoff maneuver, we seek the optimal (constant) torque (u_{12}) to apply to the body while both feet are in contact: with $a_{12} > 0$ as an input penalty parameter,

$$c_{\text{liftoff}}(\theta_{\text{apex}}, u_{12}) = (\theta_{\text{nadir}} - \theta_{\text{desired}})^2 + a_{12}u_{12}^2. \tag{4.19}$$

We implemented numerical simulations of these models[6] and applied a scalar minimization algorithm[7] to compute optimal policies as a function of initial body rotation.[8]

As expected, the optimal value and policy functions computed for the touchdown and liftoff maneuvers are nonsmooth (Figure 4.3(c,d,e,f)). This result does not depend sensitively on the problem data; nonsmoothness is preserved after altering parameters of the model and/or cost function. We emphasize that the nonsmoothness in Figure 4.3 arises from the nonsmoothness in the underlying system dynamics (4.1); the functions in (4.18) and (4.19) are smooth.

4.5 Discussion

We conclude by discussing how often we expect to encounter the nonsmooth phenomena described above in models of robot behaviors (Section 4.5.1) and what our results imply about the use of smooth tools in this nonsmooth setting (Section 4.5.2).

4.5.1 Prevalence of nonsmooth phenomena

In Section 4.4, we presented two simple optimal control problems where the dynamics of a mechanical system subject to unilateral constraints gave rise to a nonsmooth cost: one where the cost was piecewise–differentiable, and another where it was discontinuous. The reader may have noticed that the nonsmoothness occurred along trajectories that underwent simultaneous constraint (de)activation. This peculiarity was not accidental: the cost is gen-

[6] using the modeling framework in [66] and simulation algorithm in [21]

[7] `SciPy v0.19.0 minimize_scalar`

[8] We plan to release the software used to generate these results as an `environment` in `OpenAI Gym` [18].

erally continuously–differentiable along trajectories that (de)activate constraints at distinct instants in time.[9]

If the constraint surfaces intersect transversely [78, Ch. 6], then the nonsmoothness presented in Section 4.4 is confined to a subset of the state space with zero Lebesgue measure. In light of this observation, intuition may lead one to ignore these states in practice. However, we believe this intuition will lead the practitioner astray as the complexity of considered behaviors increases. Indeed, since the number of contact mode sequences increases factorially with the number of constraints and exponentially with the number of constraint (de)activations, then the region where the cost function is continuously–differentiable is "carved up" into a rapidly increasing number of disjoint "pieces" as behavioral complexity[10] increases.

Although we cannot at present comment in general on how these smooth pieces fit together, we note that some important behaviors will reside near a large number of pieces. For instance, periodic behaviors with (near–)simultaneous (de)activation of $n \in \mathbb{N}$ constraints as in [5] could yield up to $(n!)^k$ pieces after $k \in \mathbb{N}$ periods. The combinatorics are similar for tasks that involve intermittently activating (a subset of) n constraints k times as in [93]. Since the dimension of the state space is independent of n and k, these pieces must be increasingly tightly packed as n and/or k increase.

4.5.2 Justifying the use of gradient–based algorithms

Suppose a (possibly non–optimal) policy $\pi : \mathcal{X} \to \mathcal{U}$ has an associated value $\nu^\pi : \mathcal{X} \to \mathbb{R}$. If this value admits a first–order approximation with respect to π, then it is natural to improve the policy using steepest descent: with $\alpha > 0$ as a stepsize parameter,

$$\pi^+ = \pi + \alpha \arg\min_{\|\delta\|=1} D_\pi \nu^\pi(\delta). \tag{4.20}$$

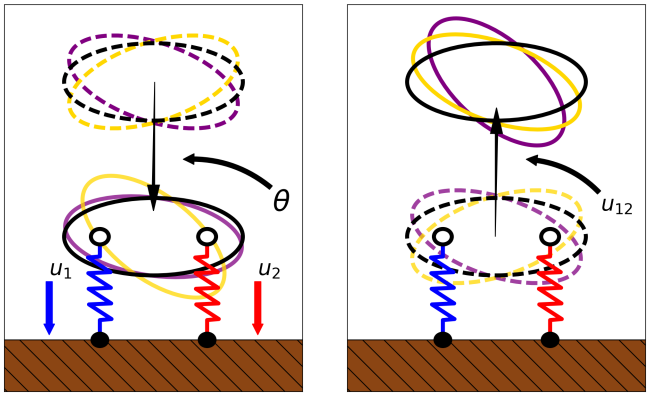
---

[9] This follows from [3, Eqn. 2.3] so long as the constraint (de)activations are *admissible* [96, Def. 3, Lem. 1].

[10] as measured by the number of constraints and/or constraint (de)activations

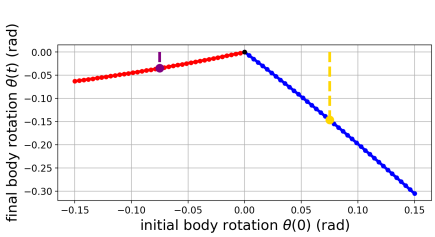
The update in (4.20) is a *direct policy gradient–based* algorithm [15, 124], and can be interpreted as a *natural* [69] or *trust region* [117] algorithm depending on the norm chosen. In practice, the derivative $D_\pi \nu^\pi$ is not generally available and must be estimated, e.g. using function approximation [39, 72] or sampling [15, 120]. This practice is justified for smooth control systems whose value functions are smooth; it is not generally justified for the mechanical systems subject to unilateral constraints considered here since the value of (optimal or non–optimal) policies can be nonsmooth. To see how nonsmoothness can prevent a gradient–based algorithm from converging to an optimal policy, consider the result of applying one step of the policy gradient algorithm in (4.20) to the optimal policies in Figure 4.3(c,d) when ν^π is merely piecewise–differentiable. Since the policy is optimal, the first–order necessary condition for optimality (4.13) implies that the arg min in (4.20) evaluates to zero, and therefore the optimal policy is a fixed point of the update in (4.20) when the true (Bouligand) derivative $D_\pi \nu^\pi$ is available. However, an estimate of $D_\pi \nu^\pi$ obtained via sampling or function approximation would be nonzero near $\theta(0) = 0$, causing one step of the policy gradient algorithm in (4.20) to *diverge* from the optimal policy. This can be seen in Figure 4.4, where central finite differences was used to compute the change in policy determined by naive application of the *deterministic policy gradient* algorithm [120]. Near discontinuities in the optimal policy, the change from the optimal policy can be arbitrarily large.

Recent work employs smooth approximations of the contact–rich robot dynamics in (4.1) to enable application of gradient–based learning [74, 79, 81] and optimization [41, 92, 93] algorithms. This approach leverages established scalable algorithms, but does not ensure that policies optimized for the smoothed dynamics are (near–)optimal when applied to the original system's nonsmooth dynamics, since the dynamics of the smooth system being optimized differ from those of the original system. As an alternative approach, the framework we introduced in [96] provides design conditions that ensure trajectories of (4.1) depend continuously–differentiably on initial conditions. Thus in future work it may be possible to justify applying established algorithms for optimal control directly on some mechanical systems subject to unilateral constraints.

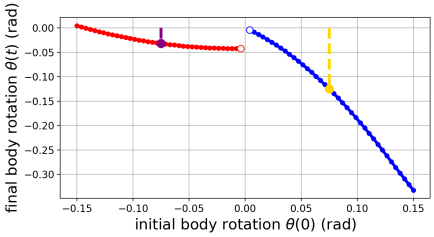


(a) *touchdown* maneuver illustration

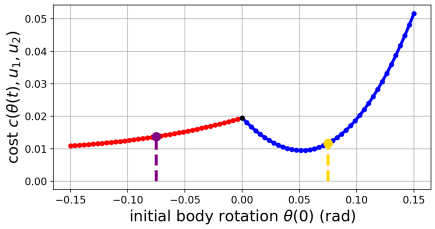
(b) *liftoff* maneuver illustration



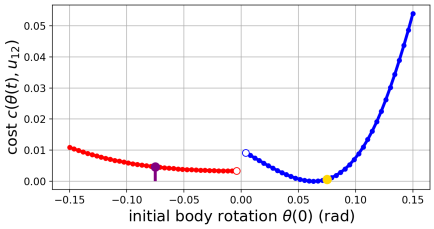
(c) touchdown trajectory outcomes



(d) liftoff trajectory outcomes

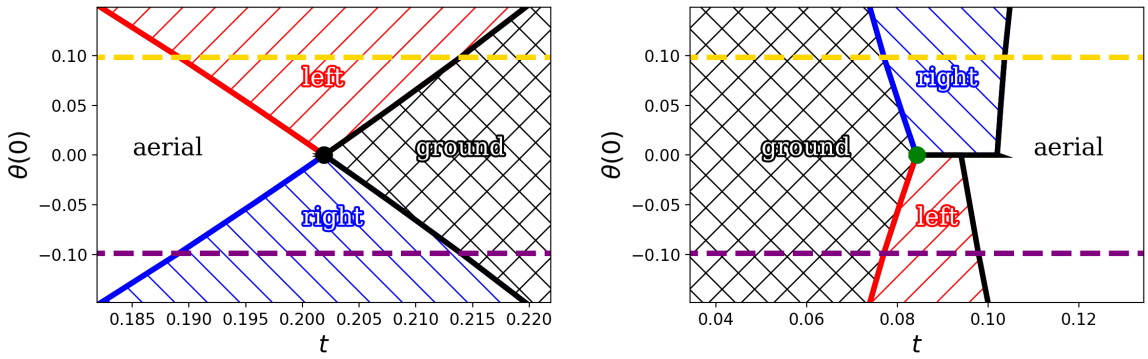


(e) touchdown value



(f) liftoff value

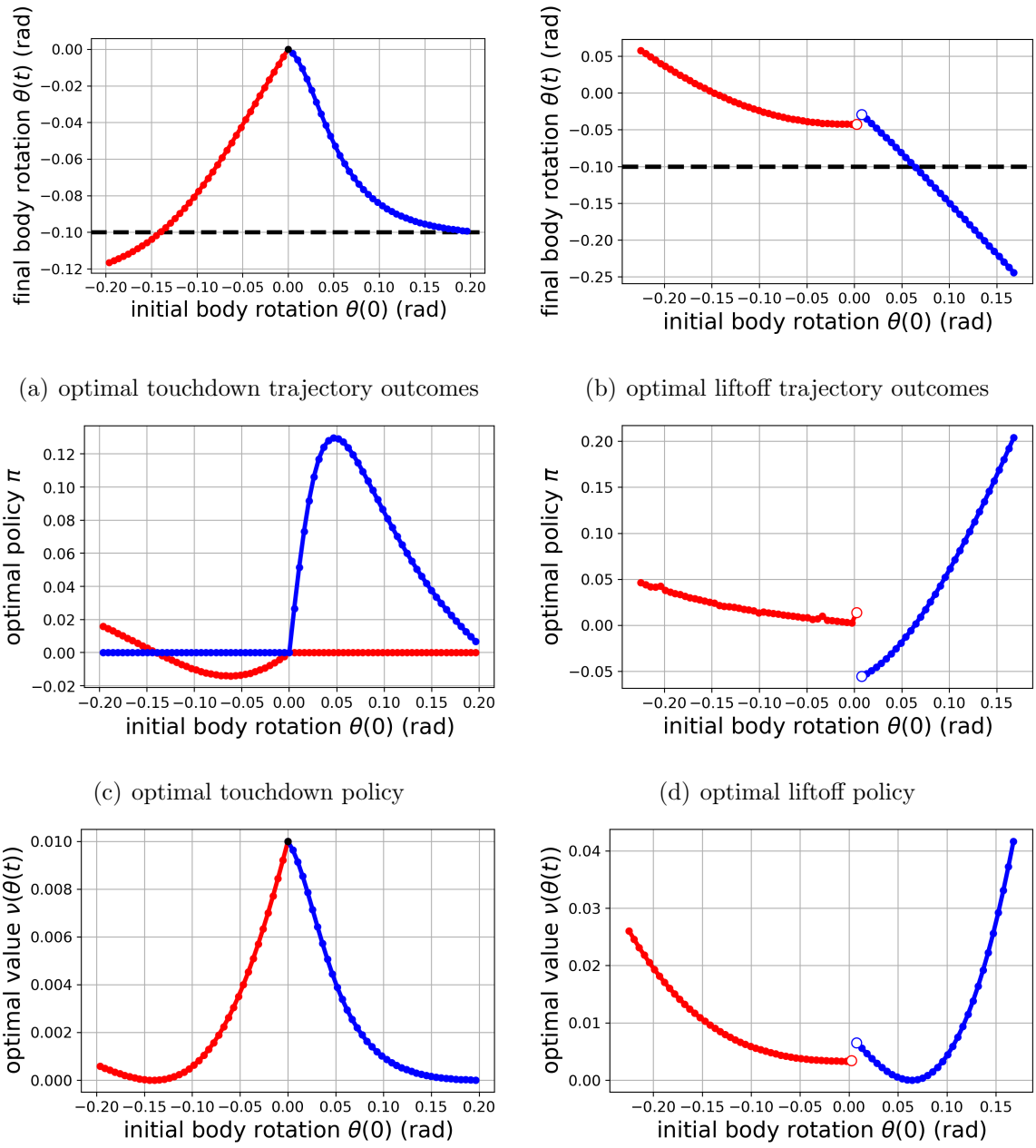
Figure 4.1: *Piecewise–differentiable and discontinuous trajectory outcomes in saggital–plane biped.* (a,b) Illustration of two maneuvers—*touchdown* and *liftoff*—performed under non–optimal policies that exert different forces depending on which feet are in contact with the ground. In the *touchdown* maneuver, feet are initially off the ground and trajectories terminate when the body height reaches nadir; in the *liftoff* maneuver, feet are initially on the ground and trajectories terminate when the body height reaches apex. (c,d) Trajectory outcomes (final body angle $\theta(t)$) as a function of initial body angle $\theta(0)$. (e,f) Performance of trajectories as measured by the cost functions in (4.18), (4.19). Dashed colored vertical lines on (c–f) indicate corresponding colored outcomes on (a,b).



(a) touchdown contact modes

(b) liftoff contact modes

Figure 4.2: *Contact modes for touchdown and liftoff maneuvers.* The saggital–plane biped illustrated in Figure 4.1(a,b) can be in one of four *contact modes* corresponding to which subset $J \subset \{1, 2\}$ of the (two) limbs are in contact with the ground; each subset yields different dynamics in (4.1). (a,b) System contact mode at each time t for a given initial body rotation $\theta(0)$; the body torque input is zero ($u_{12} = 0$) and the leg forces are different ($u_1 \neq u_2$) in mode *left* ($\{1\}$) and *right* ($\{2\}$) than in *aerial* ($\emptyset$) or ground ($\{1, 2\}$). Dashed colored horizontal lines indicate corresponding colored trajectories in Figure 4.1. The increase in force during the transition to modes *left* and *right* in (b) changes the ground reaction force discontinuously, delaying liftoff and causing discontinuous trajectory outcomes in Figure 4.1(d).



(a) optimal touchdown trajectory outcomes

(b) optimal liftoff trajectory outcomes

(c) optimal touchdown policy

(d) optimal liftoff policy

(e) optimal touchdown value

(f) optimal liftoff value

Figure 4.3: *Optimal trajectories, values and policies for touchdown and liftoff maneuvers.* Optimizing (4.18), (4.19) for the biped in Figure 4.1 yields trajectory outcomes (a,b), policies (c,d), and values (e,f) that are nonsmooth (piecewise–differentiable or discontinuous). Asymmetries in trajectory outcomes are due to unequal input penalty parameters ($a_1 \neq a_2$) in (a) and unequal leg forces ($u_1 \neq u_2$) in (b).

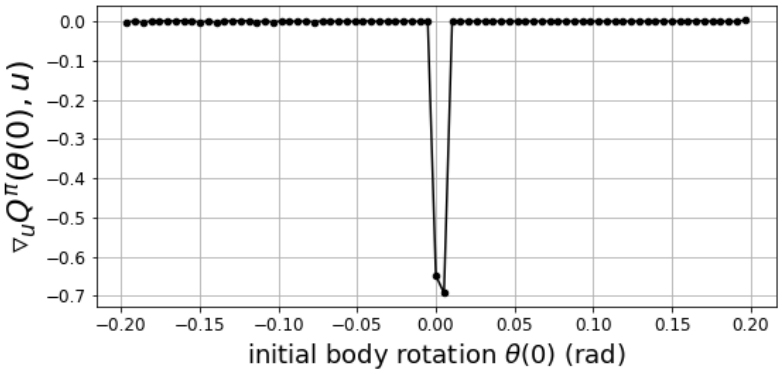


Figure 4.4: Computed the input derivative of the Q-function, the update for deterministic policy gradient. Taken around the optimal policy for the up maneuver.

Chapter 5

NONSMOOTH OPTIMAL VALUE AND POLICY FUNCTIONS FOR (ANOTHER) MECHANICAL SYSTEMS SUBJECT TO UNILATERAL CONSTRAINTS AND SIMULATION METHODS

5.1 Introduction

In Chapter 4, the regularity changes for trajectory outcomes of a hoping the saggital-plane biped are a consequence of the mode-dependent control policy. In the following chapter we will look at another system, Box, whose dynamics lead to discontinuous trajectory outcomes. Additionally, this chapter will cover simulation techniques that facilitated the implementations of both systems.

The motivation of the work in this aim was the growing popularity of reinforcement learning algorithms, specifically policy gradients, in the robotics community and the apparent mismatch between expected and actual performance, a discrepancy frequently referred to as the simulation to reality divide [71]. While optimizing a policy for a robot, policy gradient methods descend the cost landscape until performance achieves acceptable standards or a local cost minimum is reached. There are many methods to find or estimate the gradient, the most straightforward, if possible, is analytically differentiating the cost function with respect to the policy parameters. After obtaining a gradient, the policy can be updated to one that incurs a lower cost or higher reward [124]. Training a policy on a simulated model is cheaper and faster than its physical counterpart. Unfortunately, transferring these learned policies to physical systems leads to performance degradation [27] [12]. Transfer learning is primarily challenging due to model fidelity.

In practice, robot control frequently relies on reduced-order models and hand-tuned con-

trollers limiting task generalizability and flexibility. Although there are developed methods trying to bridge the simulation to robot learning performance gap, much work remains as the ability to control robots lags behind their physical capabilities [92]. Hybrid dynamical systems can be used to model contact-rich legged locomotion [67]. In general, contemporary RL techniques assume smoothness, oftentimes not explicitly, of trajectory outcomes. Our work, in conjunction with Prof. Sam Burden, shows that in the case of contact-rich dynamics where a system moves through different modes, policy gradient methods can fail to converge, due to nonsmoothness in trajectory outcomes. Expanding upon this claim, the regularity of the trajectory outcomes composed with the cost function is inherited by the value function and optimal policy for the maneuver. The nonsmoothness of the trajectory outcomes arises from the instantaneous changes in state and dynamics experienced by the system as it switches modes. The computational results later provided, further confirm our theoretical analysis, and provide an instantiation of nonsmooth outcomes. We would expect the challenge in optimizing a maneuver to grow with the number of mode transitions.

5.2 (Another) Mechanical systems subject to unilateral constraints

The rest of the section will delve into perhaps the simplest mechanical system, the box, that can exhibit the aforementioned regularities in its trajectory outcomes. The box is subject to plastic impacts, the normal velocity is zeroed instantaneously upon impact, and contact with the ground is frictionless mirroring assumptions made in Chapter 4.

$$a_l(q) = y - \frac{h}{2}cos(\theta) - \frac{w}{2}sin(\theta) \tag{5.1a}$$

$$a_r(q) = y - \frac{h}{2}cos(\theta) + \frac{w}{2}sin(\theta) \tag{5.1b}$$

The box has two unilateral constraints expressed in (5.1). The constraints prevent penetration of the ground by the corners of the box.

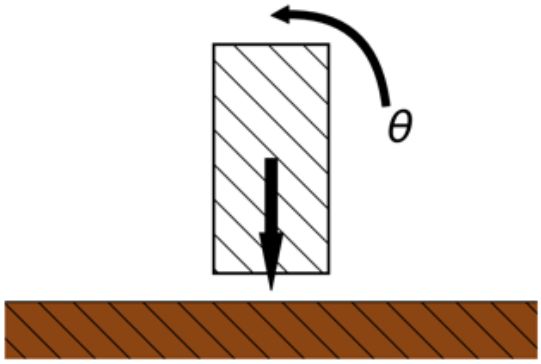


Figure 5.1: The box has three degrees of freedom which is represented by the configuration vector for this system, $q = [x, y, \theta]$.

$$\begin{aligned} J = \{&(aerial = (a_l > 0, a_r > 0)) \\ &, (left = (a_l = 0, a_r > 0)) \\ &, (right = (a_l > 0, a_r = 0)) \\ &, (ground = (a_l = 0, a_r = 0))\} \end{aligned} \tag{5.2}$$

The set of modes are all possible contact configurations of the box with the ground denoted by the set J and shown in (5.2).

5.3 Simulating mechanical systems subject to unilateral constraints

$$k_1 = hF(x_n) \tag{5.3a}$$

$$k_2 = hF(x_n + \frac{1}{2}k_1) \tag{5.3b}$$

$$k_3 = hF(x_n + \frac{1}{2}k_2) \tag{5.3c}$$

$$k_4 = hF(x_n + k_3) \tag{5.3d}$$

$$x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \tag{5.3e}$$

To simulate the hybrid system, we needed to convert the continuous-time system into a discrete-time system. The continuous dynamics of the hybrid system, $\dot{x} = F_j(x)$, can be discretized, $\bar{F}_j(x) = hF_j(x)$, with the $x^+ = F_j(x^-)$ describing the systems discrete-time evolution with a step size of h. We simulated the discretized system using the Runge-Kutta method, (5.2), a higher order variation of the Forward Euler method [65].

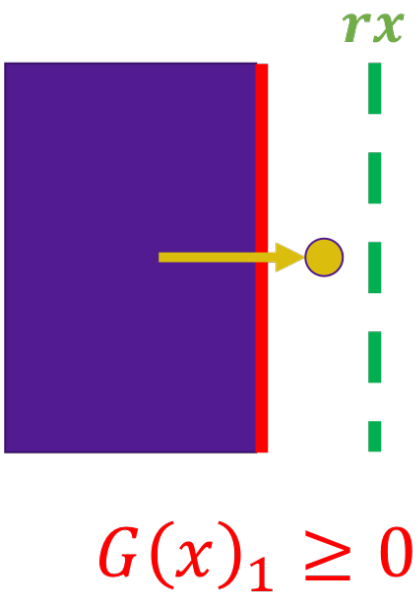


Figure 5.2: Relaxation solver for the simulated constraints.

At each iteration of the simulation, the guard function is evaluated. If a guard has been violated, the simulation software next checks to see if the point lies within the relaxation distance, rx, of the guard. The computational process for finding a point within the relaxation distance is depicted in Figure 5.2. If the point is not within the relaxation distance, the step size is halved, and the discretized vector field is evaluated again. The relaxation distance is chosen by the user prior to running the simulation and determines the fidelity of the guard function. A smaller parameter value improves adherence to the guards but requires more computations. The pseudocode is shown in (5.3). When running, the simulation checks for *Zeno* which is defined as an infinite number of discrete transitions in finite time [53]. This

pathology arises from the construction of hybrid systems and the interaction between the vector fields and guards of the various modes. The simulation software has a basic check to try to detect *Zeno* execution, once a predetermined number of discrete transitions (parameter choice) has been met the simulation is terminated.

$$1.a(q) \geq 0, \quad 2.\lambda \geq 0, \quad 3.a(q)^T\lambda = 0 \tag{5.4a}$$

$$1.\nabla a(q)^T\dot{q} \geq 0, \quad 2.\lambda \geq 0, \quad 3.\nabla a(q)^T\dot{q}\lambda = 0 \tag{5.4b}$$

Once a guard is reached, the reset function is called and the velocity vector of the box, $\dot{q} = [\dot{x}, \dot{y}, \dot{\theta}]$, is updated to prevent a corner from penetrating the ground. The complementarity test is shown in equation (5.5a). The velocity impulse scheme, shown in equation (5.5b), is an equivalent higher order formulation which projects the velocity of the box onto the direction normal to the surface and is the test used in our simulation. After test is applied, the mode of the system is updated instantaneously. The test is composed of three conditions: 1. none of the unilateral constraints should be violated, 2. all constraint forces should be positive or zero (there should not be adherence forces for this class of models), and 3. there should only be a constraint force in the case of contact. By running this test, only one of the modes will meet all three conditions for linear systems [19]. The sagittal-plane biped and box are not linear, the application of the complementarity test is the first step in determining mode changes in simulation [19].

$$\tilde{J} = 2^{\{1,..,n\}} \tag{5.5a}$$

$$J^* = argmin_{J\in\tilde{J}} \ (\dot{q}_J - \dot{q})^T M(q)(\dot{q}_J - \dot{q}) \tag{5.5b}$$

If the complementarity test as described above fails to yield a single mode, we then apply Gauss's principle of least constraint [103]. The principle assigns each mode a real scalar, giving us the ability to distinguish between multiple modes that may be valid or barely non-valid according to the complementarity test. The set of all available modes, see in equation (5.5a), is the powerset of possible unilateral constraints. We then select the mode with the

smallest value allowing the simulation to proceed. After applying the aforementioned steps, we were able to simulate the box in a manner that matched expected behavior.

5.4 Optimizing the Box Maneuver

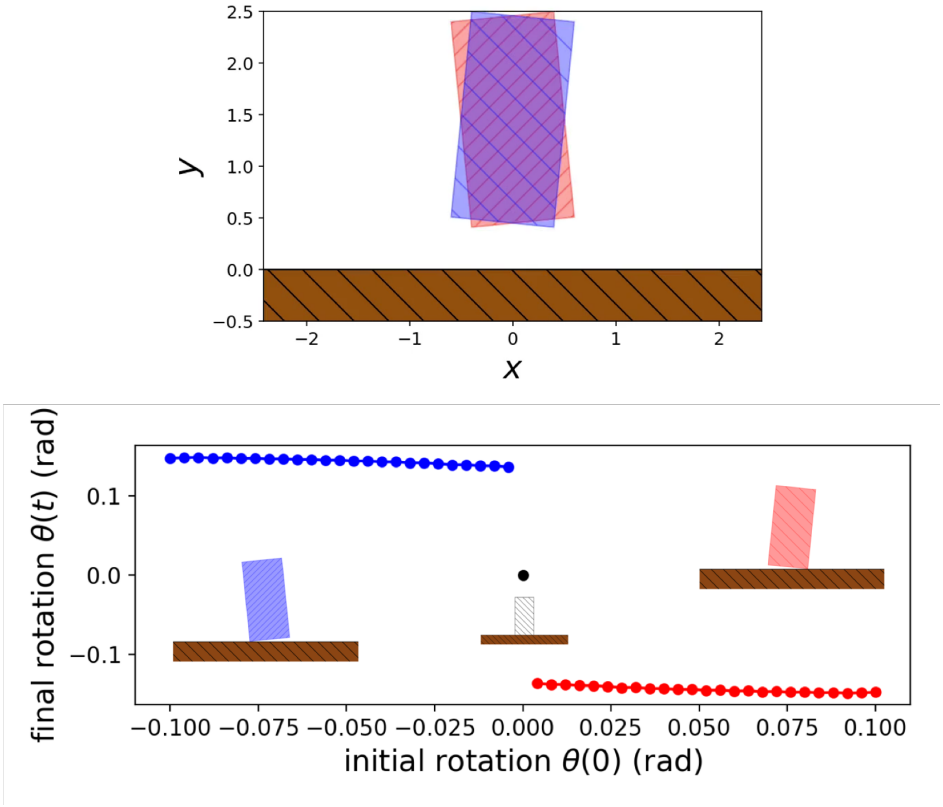


Figure 5.3: Box Trajectory Outcomes

Figure 5.3 shows the initial configuration of the box for the given maneuver. The initial cartesians positions, x, y, are held constant and the orientation of the box, $\theta(0)$, is varied. Around $\theta(0) = 0$, there are two distinct trajectory outcome branches. In this chapter, the discontinuities in the trajectory outcomes are due to the dynamics of the box. This differs from the sagittal-plane biped in Chapter 4 where the observed discontinuities are due to the mode-dependent control policy.

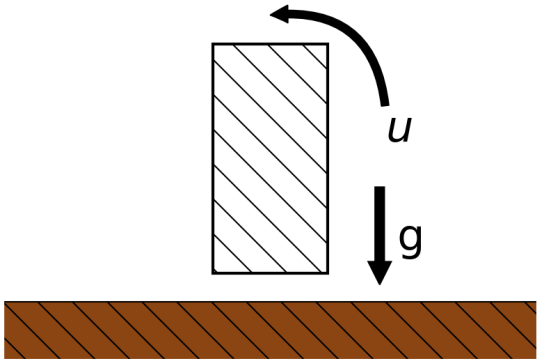


Figure 5.4: Box Maneuver

$$c(\theta, u) = (\theta_d - \theta(t))^2 + \alpha u^2 \tag{5.6}$$

Next, we add a rotational control input to the box in the clockwise direction, shown in Figure 5.4. We can then associate a cost function for the maneuver where the box is dropped from a fixed height and time horizon with the previously stated dynamics. In equation (5.6), θ_d refers to the desired final orientation of the box, $\theta(t)$ the final rotation of the box, α a scalar used to penalize control input, and θ the rotational trajectory of the system through the maneuver.

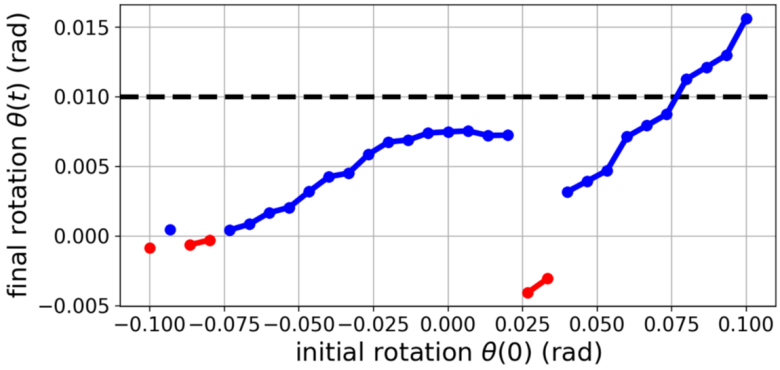


Figure 5.5: Optimized box trajectory outcomes.

In Figure 5.5 we can see trajectory outcomes of the system subject to the to the optimal policy. The dashed line indicates the desired final angle of the box. As per the previous figures, blue indicates a final *left* stance and red a *right* stance. Therefore, the outcomes are discontinuous for this maneuver. To find the optimal policy, a line search was conducted along the input values. A numerical scalar optimization function was used [1]. The problem is nonlinear, making popular gradient methods lose optimality guarantees. As expected, this analysis would becomes significantly more challenging as the number of variables increases, indicative of optimization challenges faced by the entire robotics community [97].

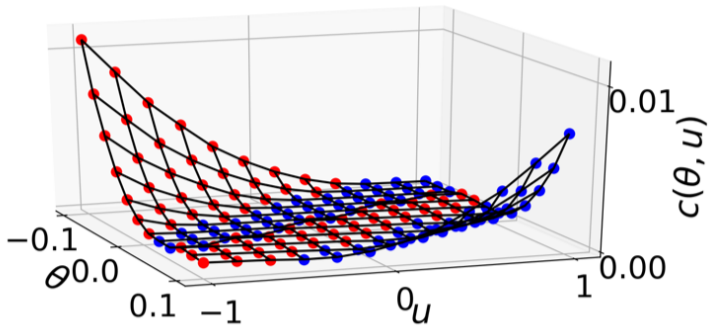
In Figure 5.6, we can see the cost landscape over a range of values for the initial orientation and control input. Depending on the two initial values, there are distinct regions for the final configuration. This indicates that as the initial value is smoothly varied, the final contact mode varies and the cost discontinuously changes, as seen in Figure 5.6.

Figure 5.7 shows the value function and optimal policy for each initialization of the system. Both vary discontinuously, mirroring the trajectory outcomes of the system and lining up with theoretical expectations.

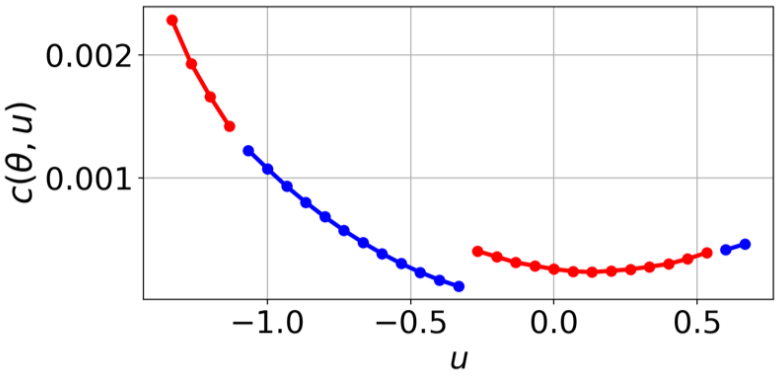
5.5 Conclusion

The challenge of going from simulation to real, faced by the robotics community, may be a result of smoothness assumptions we make of physical and simulated systems. In this chapter the box showcased this pathology. Instantaneous updates to state may be more representative when describing rigid-body impacts while operating at the time scale of interest for robotics. This chapter also outlined contributions on how to detect and assign mode changes when simulating hybrid dynamical systems.

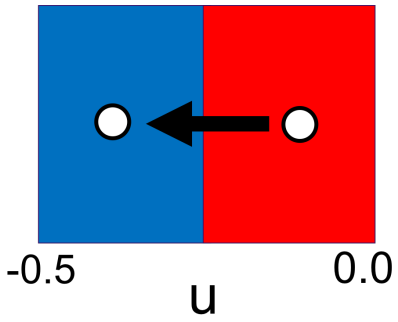
[1] `SciPy v0.19.0 minimize_scalar`



(a) 3D cost landscape

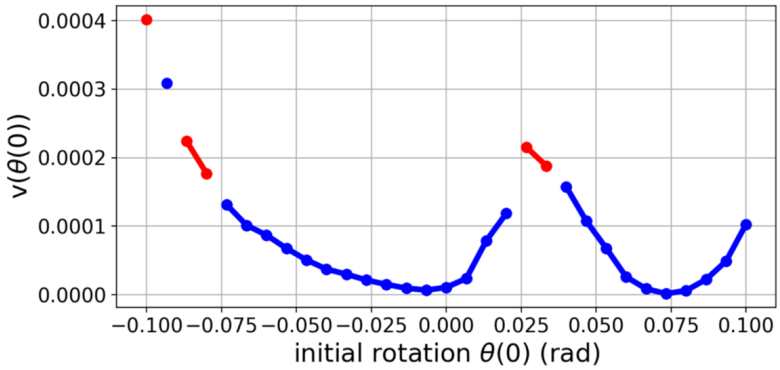


(b) 2D cost landscape

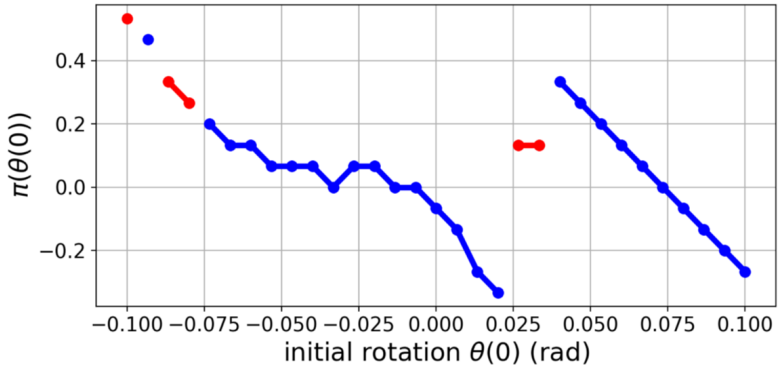


(c) Policy iteration and contact sequence change

Figure 5.6: Cost landscape for the box maneuver.



(a) Optimal Value



(b) Optimal Policy

Figure 5.7: Optimal box maneuvers.

Chapter 6

CONTACT SEQUENCES FOR MECHANICAL SYSTEMS NEAR PERIODIC ORBITS

6.1 Introduction

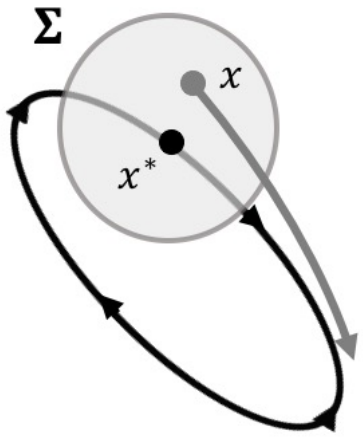


Figure 6.1: A mechanical system near limit cycle.

A limit cycle is an isolated closed trajectory and by definition flowing along one results in periodic motion [123]. Near the limit cycle a Poincaré map, P, takes points, x, from Σ, a space of codimension one, and maps them back onto Σ, otherwise called the transversal section [126]. For P this defines a relation where $x^+ = P(x^-)$ and $x^+, x^- \in \Sigma$. The mapping is depicted in Figure 6.1. In this chapter, we will examine how this relation is impacted as you move a small distance, δ, from the periodic orbit for a mechanical system in a rhythmic motion that achieves distinct *aerial* (no limbs in contact with the ground) and *ground* modes (both limbs in contact with the ground). When the system is perturbed by

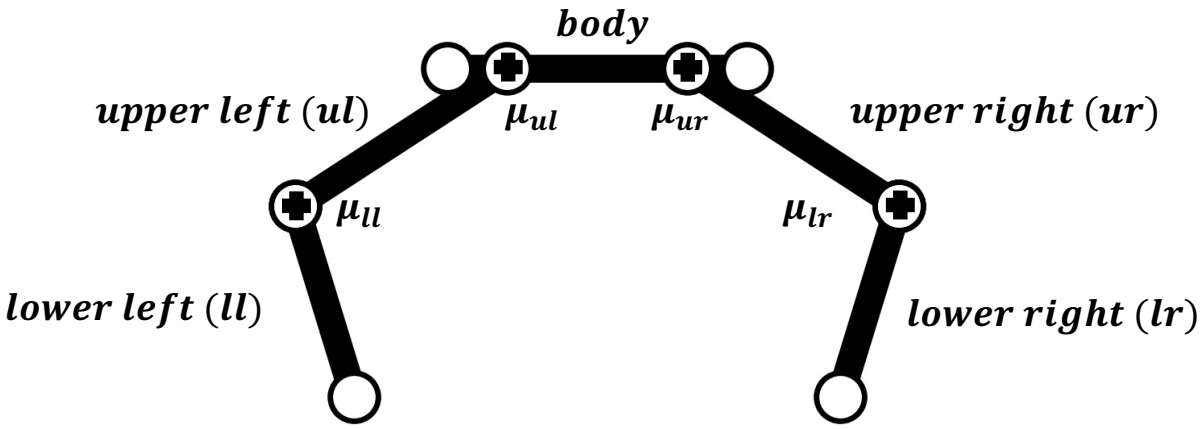


Figure 6.2: Hopper diagram. The mass is uniformly distributed across each of the links. The + indicates a pin joint and actuator.

a randomly chosen δ (small in magnitude), we expect this system to potentially experience arbitrary sequences of contact configuration.The set of contact sequences relates to the number of discrete transitions allowed (discrete time horizon), dynamics of the system, and the numerical effects incurred while simulating. In the remainder of this chapter, we will discuss the contact sequences experienced by a simulated mechanical system subject to unilateral constraints that was perturbed away from its limit cycle.

6.2 Hopper system

The system in Figure 6.2 was the mechanical system simulated, the Hopper. The Hopper has five rigid links, for each link there are two translational and one rotational degrees of freedom resulting in a total of fifteen. The links have the following naming convention as seen in Figure 6.2: body, upper left *(ul)*, upper right *(ur)*, lower left *(ll)* and lower right *(lr)*. The control input, $u = [\mu_{ll}, \mu_{ul}, \mu_{ur}, \mu_{ll}]$, is applied as rotational torques at the actuators. The rotational direction of the torques follows the right-hand convention.

6.2.1 Constraints

Matching assumptions from Chapter 4 and Chapter 5, the Hopper shown in Figure 6.2 has frictionless contacts and plastic impacts. Both of the feet cannot penetrate the ground (two unilateral constraints) and pin joints tie the links together (four bilateral constraints). Constraint forces follow the framework from [67] and constraint stabilization from [14].

6.2.2 Motor inputs

$$\alpha = \frac{1}{2}cos(\frac{2\pi t}{T}) + \frac{1}{2} \tag{6.1}$$

$$\bar{q}_\theta = \alpha q_{\theta_0} + (1-\alpha) q_{\theta_1} \tag{6.2}$$

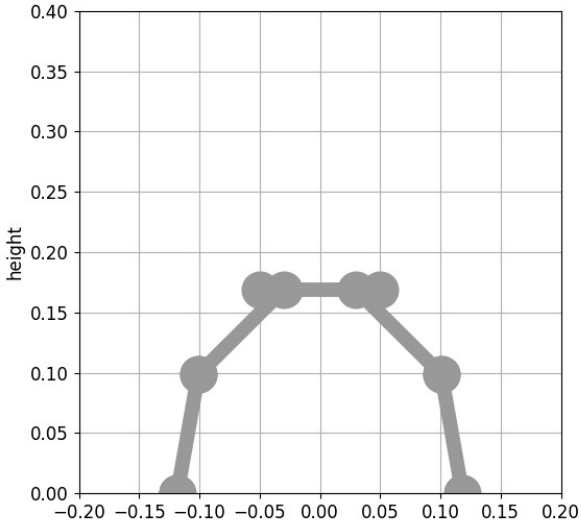
$$\mu = K_p(q_\theta - \bar{q}_\theta) - K_d \dot{q}_\theta \tag{6.3}$$

To achieve a hopping motion, a periodic control input was applied to the system. θ_1 is the angle between the body and upper links and θ_2 is the angle between the upper and lower links. The configuration vector defined as $q_\theta = [\theta_1, \theta_2]$. Two reference configurations were identified for both the *open*, $\bar{q}_{\theta_{open}}$, and *closed*, $\bar{q}_{\theta_{closed}}$ contact configurations as seen in Figure 6.3. The system minimizes error, $e = q_\theta(t) - \bar{q}_\theta(t)$, with the PD controller shown in (6.3). For periodic motion to happen, a linear combination of the configurations is the desired and is described according to function in (6.2). A smooth tracking signal was chosen to keep the reference tracking error small to prevent the simulation from becoming unstable due to large control inputs. The oscillations are driven by the periodic function shown in (6.1).

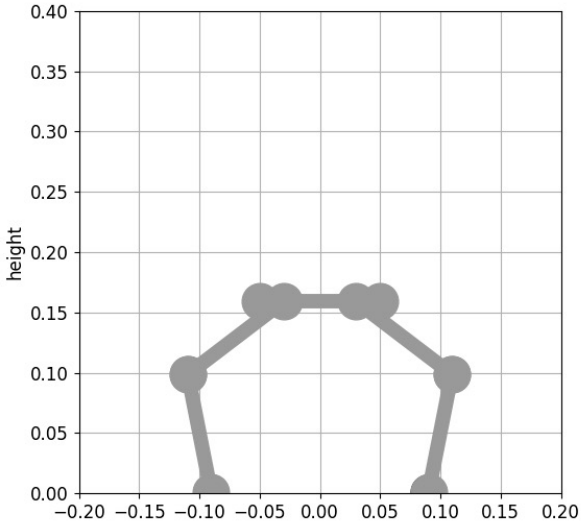
6.3 Contact Configurations

$$f_{Baumgarte} = \ddot{c} + 2\alpha\dot{c} + \beta^2 c \tag{6.4}$$

Baumgarte stabilization, seen in the function from (6.4), has two parameters, α and β, which determine the response of the stabilizer. c is the constraint vector, a is the unilateral



(a) *open* contact configuration



(b) *closed* contact configuration

Figure 6.3: The Hopper impact configurations.

constraints, and b the bilateral constraints so that $c = [a, b]$. The force exerted by the stabilizer corrects for error accumulation caused by the discretization of continuous dynamics when simulating the system. Incorporating Baumgarte stabilization into the computational framework enabled the simulation of the Hopper, which has a greater number of bilateral constraints than the systems in Chapter 4 (box) and Chapter 5 (hopper). Previously, the bilateral constraint error would continue to accumulate causing the constraints to become unstable.

6.4 Contact sequences

6.4.1 Definition

In the remainder of this chapter we will refer to $aerial := (0, 0)$, $left := (1, 0)$, $right := (0, 1)$, and $ground := (1, 1)$. The periodic motion of the Hopper when controlled according to Section 6.2.2 is symmetric and undergoes the following transitions $(0, 0) \rightarrow (1, 1) \rightarrow (0, 0)...$ indefinitely as both feet hit the ground simultaneously before sliding along the ground and pushing off back into the *aerial* mode. In the chapter, we will define the transverse section, Σ, as the apex of the hop when the vertical velocity of the body equals zero. We can further suppress notation, by defining a sequence as $(1, 1)_P := (0, 0) \rightarrow (1, 1) \rightarrow (0, 0)$. Due to the dynamics of system, both feet simultaneously lift-off, but either may come into contact with the ground first. Thus we define the other two contact sequences that can be observed: $(1, 0)_P := (0, 0) \rightarrow (1, 0) \rightarrow (1, 1) \rightarrow (0, 0)$ and $(0, 1)_P := (0, 0) \rightarrow (1, 0) \rightarrow (1, 1) \rightarrow (0, 0)$.

6.4.2 Perturbations applied

We sought to find all possible contact sequences for an N-cycle hop horizon of the the Hopper. We conducted this search for both impact configurations, *open* and *closed*. A configuration is classified as *open* when the horizontal position of the feet is outside the pin-joint linking the upper and lower link as the feet come into contact with the ground. The force of gravity acting upon the system causes the feet to splay further apart post impact. A configuration

is classified as *closed* when the horizontal position of the feet is inside the pin-joint for the upper and lower links upon impact, driving the feet closer together post impact. Since there are three possible outcomes for each cycle, the number of sequences is bounded from above by 3^N. We applied perturbations to the rotational and velocity components of the state. For the rotational perturbations, a random vector was generated from a normal distribution. The rotations were then applied in the following order: body, upper links, and lower links (the vector had a length of three). This was to ensure consistency within the kinematics (prevent violation of the bilateral constraints). A perturbation to the velocity components of the system was applied by projecting a randomly generated vector onto the constraint tangent plane to the bilateral constraints. The perturbations were randomly sampled from a normal distribution and scaled according to selected parameters, chosen to sufficiently perturb the system and sample the area around the limit cycle in the transverse section.

6.5 Results & discussion

Contact Sequence	Observed (Boolean)
$(0,1)_p \rightarrow (1,0)_p$	True
$(1,0)_p \rightarrow (0,1)_p$	True
$(1,0)_p \rightarrow (1,0)_p$	True
$(0,1)_p \rightarrow (0,1)_p$	True

Figure 6.4: Contact sequences for 2-cycle horizon for both *closed* and *open* configurations.

The Hopper was simulated 250 times for both the *closed* and *open* contact configurations. Figure 6.4, Figure 6.5, and Figure 6.6 display non-trivial results, $(1,1)_p$ occurs *ad infinitum* in the absence of perturbations per our choice of limit cycle. For a 2-cycle horizon, all of the contact sequences were observed for both contact configurations. For the 3-cycle horizon, both contact configurations had 6 out of the possible 8 contact sequences. For

Contact Sequence	Observed (Boolean)
$(0,1)_p \to (0,1)_p \to (0,1)_p$	False
$(0,1)_p \to (0,1)_p \to (1,0)_p$	True
$(0,1)_p \to (1,0)_p \to (0,1)_p$	True
$(0,1)_p \to (1,0)_p \to (1,0)_p$	True
$(1,0)_p \to (0,1)_p \to (0,1)_p$	True
$(1,0)_p \to (0,1)_p \to (1,0)_p$	True
$(1,0)_p \to (1,0)_p \to (0,1)_p$	True
$(1,0)_p \to (1,0)_p \to (1,0)_p$	False

Figure 6.5: Contact sequences for 3-cycle horizon for *closed* configuration.

Contact Sequence	Observed (Boolean)
$(0,1)_p \to (0,1)_p \to (0,1)_p$	True
$(0,1)_p \to (0,1)_p \to (1,0)_p$	True
$(0,1)_p \to (1,0)_p \to (0,1)_p$	False
$(0,1)_p \to (1,0)_p \to (1,0)_p$	True
$(1,0)_p \to (0,1)_p \to (0,1)_p$	True
$(1,0)_p \to (0,1)_p \to (1,0)_p$	False
$(1,0)_p \to (1,0)_p \to (0,1)_p$	True
$(1,0)_p \to (1,0)_p \to (1,0)_p$	True

Figure 6.6: Contact sequences for 3-cycle horizon for both *open* configuration.

the *open* configuration the alternating sequences, *i.e.* $(1,0)_p \to (0,1)_p \to (0,1)_p$, were not observed. Repeating sequences, *i.e.* $(1,0)_p \to (0,1)_p \to (1,0)_p$, were not observed for the *closed* configuration. The number of contact sequences stayed constant for a 4-cycle hori-

zon. The natural dynamics of the contact configuration paired with the perturbation applied make certain contact sequences possible and with varying levels of frequency. We believe the reported results are due to multiple effects: numerical imprecision from the simulation framework, the dampening effect of the PD controller, and the dynamics of the maneuvers. A *closed* configuration causes the Hopper to loosely mirror the previously discussed box, for each subsequent hop the Hopper frequently rockers (rotates around its center of mass) resulting in alternating contact sequences. In order for a a repeating sequence to occur, we suspect there must be a very precise perturbation onto a small area of the transverse section. The area likely shrinks or disappears for larger cycle horizons. This is due to the decay effects of the simulation, at each time step the discretized dynamics incur constraint error. The enforcement and stabilization of constraints causes the Hopper to dissipate high frequency energy and trend towards a $(1,1)_p$ contact sequence. The *open* contact configuration sees repeating sequences more frequently. While splaying, the Hopper dissipates rotational velocities therefor maintaining a consistent body rotation across contact sequences.

A potential future research direction would find periodic hoping gaits that display more than 6 contact sequences for 3-cycle or larger horizons. Increasing the precision of the simulation would aid the search for more contact sequences by reducing the dissipation of perturbations. In summary, this chapter has provided preliminary evidence that the number of contact sequences a mechanical systems subject to unilateral constraints experiences increases with the cycle horizon. These findings have ramifications on smooth optimization techniques and ties in with findings from Chapter 4 and Chapter 5, each contact sequence propagates nonsmoothness.

Part II

AIM 2

Chapter 7

SUBJECT-SPECIFIC MODELS FOR PREDICTING HUMAN LOCOMOTOR RESPONSE TO ANKLE FOOT ORTHOSES (AFO)

7.1 Introduction

AFOs can be used to partially restore walking economy for children with cerebral palsy [17]. The potential assistive and rehabilitative properties of AFOs inspired this project [37]. The AFO device we used is passive, a spring runs along the heel and is actuated through motion in the sagittal plane of the ankle. Since the device is passive, the spring stores energy as the foot comes into contact with the ground and goes into dorsiflexion. As the person toes-off and the foot moves towards plantar flexion, the spring releases energy. The net work applied by the device is negative because of mechanical dissipation. To motivate the application of this aim, currently the prescription of AFOs is largely subjective in nature and places a large reliance on the clinician [94]. Due to the cost of fabrication and clinical time required for each device iteration, many patients end up with sub-optimal solutions. This can prevent patient adoption and efficacy of the devices. Although there are clinical procedures available for evaluation and prescription, no standard protocol exists, and outcomes are inconsistent, even within a single clinical center [94] [110].

We hope to assist in the clinical prescription of AFOs and improve patient outcomes by creating an accurate subject-specific predictive model for kinematics. Its worth noting that the aim of this work is not to replace clinicians, but offer another tool, quantitative in nature, to facilitate prescription and improve outcomes. The model would predict changes to kinematics as the torque applied by the AFO is varied throughout the gait cycle. For our device design, the stiffness of the spring was the only parameter varied. Predicting gait kinematics

for those with cerebral palsy is particularly challenging because they display greater gait variation than healthy individuals [57]. Additionally, the models are only afforded clinically-feasible amounts of data preventing the straightforward adoption of methods developed in other data-rich environments, **e.g.** recurrent neural networks. To simplify the prediction challenge and make finding subjects easier, we decided to first develop and survey a collection of models on healthy individuals. Using motion capture and electromyography (EMG) sensors, collaborators collected kinematic and EMG data as subjects walked at a steady-state on a treadmill. Kinetic data was estimated by taking the gradient of the kinematics and then interpolating. The subjects had four treatments, one was a control and the rest spanned a range of biomechanically relevant stiffnesses [28]. Data was collected on twelve subjects walking at a steady state on a treadmill. Each subject first went through an adaptation session and in a later session data was collected. The data processing pipeline took in raw position measurements and used an inverse kinematics algorithm in tandem with filtering to estimate joint angles. In the remainder of this chapter, I will describe the methods we developed and used in Section 7.3 and their basis in previous research. Next in Section 7.6, we will cover results from our effort to predict joint angles and EMG. Lastly in Section 7.7, we will discuss the significance of our findings, outline research contributions, and overview potential future research directions.

7.2 Determining & using phase

A number of our proposed models involve predicting futures states from current state, where the input and output states are defined in relation to their position in the gait cycle. Consequently, we must identify an appropriate definition of the gait cycle, over which the dominant dynamics of human locomotion evolve. Most biomechanical studies evaluate the gait cycle over a domain defined by initial contact events. Generally, this involves linearly interpolating the data from heel strike to heel strike. There is evidence that biological systems operate in a rhythmic fashion due to internal neural mechanisms versus sensory feedback [33]. The basis of neural control for locomotion, especially in vertebrates, is likely an innate capability.

Learned adaptive behaviors are applied on top of centrally generated patterns [54]. These findings lend evidence to the school of thought that biological specimens operating in a rhythmic fashion operate with nominal feedforward neural control. The mechanics of specimens in turn mediate rapid perturbations giving time for neural feedback control to rectify slower deviations from their target rhythmic behavior [108]. This perspective supports the assumption that the dynamics of a biological system can be represented in a lower-dimensional subspace. This inference underlines modeling rhythmic behaviors as reduced-order dynamical systems near a stable limit cycle [108] [89]. This contrasts with another common approach where specific events, *i.e.* heel strike, are given a fixed phase and estimates of the phase are obtained through interpolation [107]. In this section we will refer to the collected dataset as D, a matrix which is an element of $\mathbb{R}^{m \times n}$. m is the number of samples, n is the dimensionality of the state, joint angles, EMG, and torque profiles, and $X_\phi \in \mathbb{R}^n$ a single state sample at a specific phase, ϕ, of the gait cycle. The remainder of this section will utilize the discussed prior research for the phase estimator (Section 7.2.1), set mapper (Section 7.2.2), and predictive models (Section 7.4).

7.2.1 Phaser

We are using Phaser to estimate the gait cycle phase of the subject. Phaser, an algorithm developed for nonlinear dynamical systems, uses multiple dimensions considered to be phase locked to estimate the phase of the *master oscillator* of the system [107]. This approach leverages our experimental design, as measurements from multiple noisy markers are used in tandem to achieve a better estimate of the underlying phase of a walking individual. Utilizing the configuration of the system and accounting for any changes in internal phase, allows for an estimate of the system state that is robust to perturbations. The use of kinematic measures to estimate the instantaneous phase is not limited to Phaser [130]. When applying Phaser to our dataset, we looked for a subset of states whose time-series trajectories are relatively consistent across gait cycles. Our choice was hip flexion in the pelvis (with respect to both the left and right legs), which are indexed from the full state vector, X_ϕ, by $I_{hip\_flexion}$. We

then apply the function F_{phaser} to the columns of D indexed by $I_{hip\_flexion}$, to receive an estimate of the phase at each sample, denoted by $\phi \in \mathbb{R}^M$.

7.2.2 Set mapper

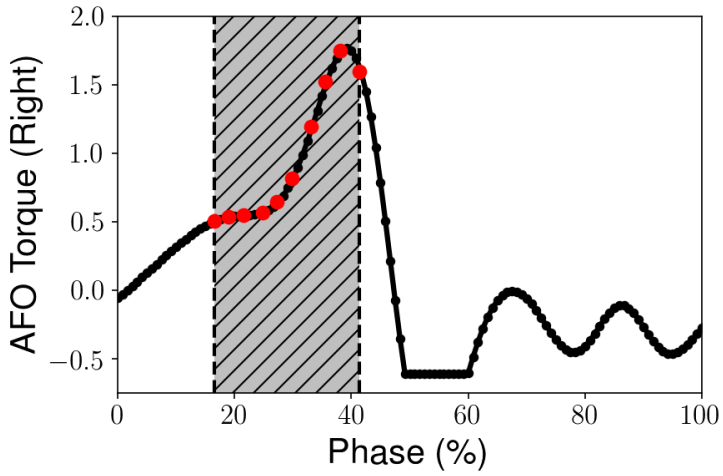


Figure 7.1: Sampling the estimated torque signal.

The torque was sampled uniformly between the input phase, ϕ, and output phase, $\phi + \theta$. These samples were then appended to the dataset before prediction. We treat the influence of the AFO on the subject as a perturbation and thus allow ourselves access to future information. We sampled the signal ten times, a hyper-parameter than can be adjusted. We selected a larger value, that yields slower run times but ensures maximal predictive accuracy (which saturates with an increasing number of samples) possible within the framework.

To construct a training set for the phase lookahead models we find state pairings offset by a parameter, θ, deemed the lookahead from the phase of the sample, ϕ. Because our data is sampled from a continuous space, the data corresponding to the lookahead phase, $\theta + \phi$, in the dataset matrix, D, is contained in the interval, $[\phi_1, \phi_2]$, defined by the unrolled phase estimate of two samples. Samples at phase lookahead, $X_{\phi+\theta} \in \mathbb{R}^n$, are then determined with linear interpolation whenever this condition is met.

7.3 Dataset & input-output structure

We collected data on four stiffness from twelve healthy participants. We had subjects walk on the treadmill for six minutes and conducted measurements at 120 *Hz*. The participants were subjected to a bilateral treatment, AFOs with the same stiffness were applied to both legs. Subjects self-selected their speeds, following instructions to maintain a natural walking pace. When choosing the appropriate granularity for the input-output relation we ended up deciding that subject-specific models are most appropriate when attempting to predict the gait kinematics and EMG response of cerebral palsy (CP) and stroke survivors. In contrast, when trying to predict joint angles and EMG for a homogenous population, you could potentially train on a large amount of subjects and then offer effective predictions for a separate group of test subjects. Due to the heterogenous nature of the injuries in our target subject population, we chose to train on a sample of stiffnesses representative of the range of viable values [57] [44]. We would then attempt to predict the subject's response to a novel stiffness and thus torque profile. Hopefully, this would allow clinicians to quickly evaluate the potential effectiveness of a stiffness, matching our aim to optimize clinical outcomes and not replace practitioners.

A definition for the state of a system is the minimum set of dimensions necessary to predict future state of the system. Picking this representation when seeking to create data-driven models for actual systems is a trade-off; balancing between time, money, & effort collecting data and the expected predictive performance. The choice of dimensions collected follows current biomechanics conventions in addition to the measurement capabilities currently available in the Amplifying Movement & Performance (AMP) Lab [28] [34]. During the experiment, marker trajectories were collected using a motion capture system [1]. To convert marker trajectories to joint angles an inverse kinematics algorithm [2] was applied in addition to low-pass filtering, following already existing procedures [102]. The resultant 19 joint

[1] `Qualysis AB`

[2] `OpenSim 3.3`

angles represent pelvis, hip, ankle, knee, and lumbar rotations. A custom script, standard across AMP lab experiments, was used to process the EMG. We utilized 14 EMG measurements taking muscle activity measurements from muscle groups responsible for walking. The joint angles, EMG signals, and estimated torque applied by the AFO along with their estimated derivatives comprise the state. We estimated the derivatives using a higher-order finite differences scheme [3] [101].

The outputs are the predictions for the joint angle and EMG dimensions. We find this to be the most generalized prediction task for the experiment conducted. Metabolic energy cost, another popular metric, is not within the scope of this aim [38]. If other metrics are of interest, particularly those with clinical significance, the general scheme of our outputs may allow some to be computed later.

7.3.1 Studentization

$$X_s = \frac{X - \mu_x}{\sigma_x} \tag{7.1}$$

To improve numerical conditioning and allow for easier comparison between dimensions we studentized the data. In (7.1), X_s denotes the studentized dataset, μ_x and σ_x, the mean and standard deviation of the original dataset.

7.3.2 Phase lookahead

In this aim we try to predict the future state for a selected phase in the future. This can be seen in Figure 7.2, where the state, X_ϕ, at phase, ϕ, is used to predict a constant phase, $\phi + \theta$, in the future.

[3] `NumPy v1.13.3 gradient`

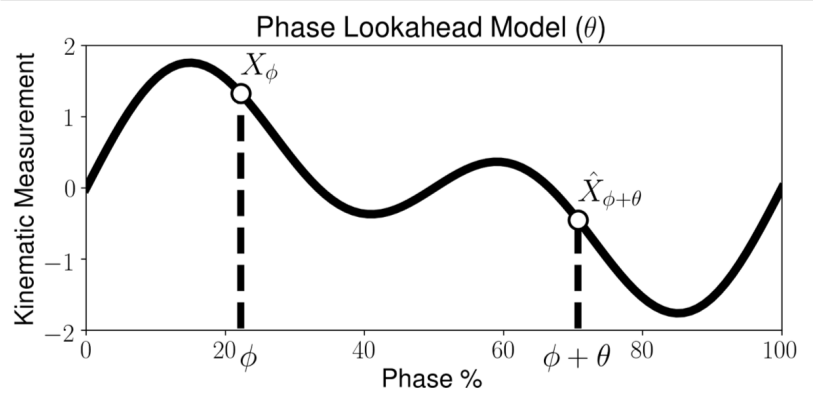


Figure 7.2: Phase based prediction framework

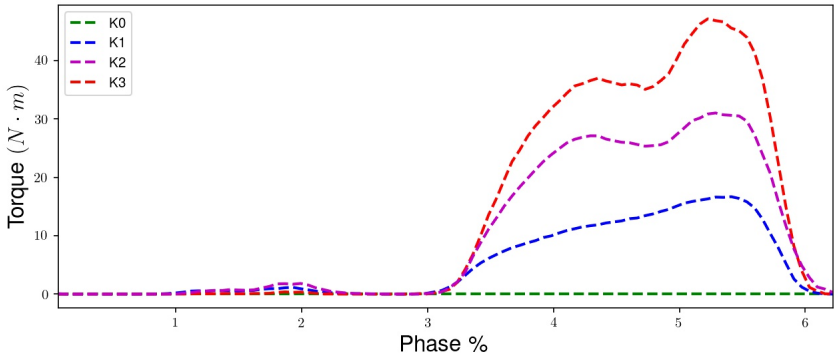
7.3.3 Training & testing procedure

Spring Name	Stiffness ($\frac{N}{m}$)
K0 (control)	0
K1	3200
K2	9800
K3	15200

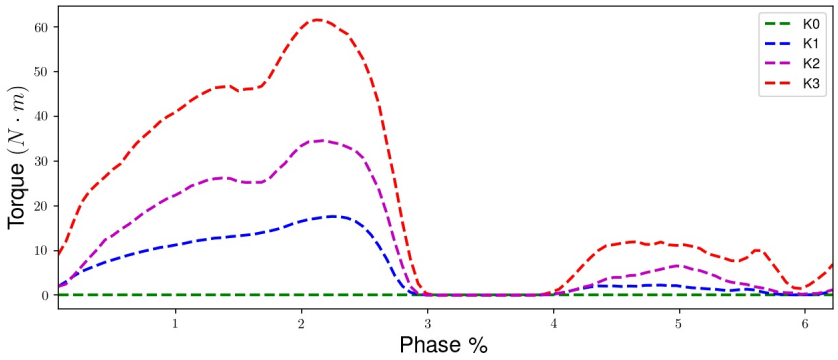
Figure 7.3: Spring stiffnesses

The springs used – K0, K1, K2 and K3 – range from the control, which has no spring causing the AFO to act as a hinge, to the stiffest spring in our experiment. The choice of spring stiffnesses used in our experiment is backed by prior research which has shown these values to measurably influence the measured quantities of interest [28]. We chose our prediction task to be an interpolation task, the values we are attempting to predict should be within the range of values trained on. This is likely easier than extrapolating to a spring stiffness beyond the range trained of the training set.

The estimated torque profile – torque applied by the AFO to the subject – for a single



(a) Mean Estimated Right AFO Torque.



(b) Mean Estimated Left AFO Torque.

Figure 7.4: Average torque profiles for each spring stiffness.

subject (*ID:5*) is shown in Figure 7.4. Matching our expectations, the torque applied by the AFO is largest while in stance phase. Additionally, we can see an ordering between the magnitude of the profile and the value of the spring stiffness. This reaffirms our approach to train on K0, K1, and K3 and test on K2, since we sought an interpolation prediction task. The profiles look qualitatively symmetric, although there are differences as expected between the right and left leg (even with the same treatment simultaneously applied to both legs).

7.4 Models surveyed

While selecting and devising the models for our survey of available techniques we wanted to explore tradeoffs along two spectrums, complexity and specificity. We chose to investigate three models that we deemed most applicable and representative: 1) Phase-Varying (PV), 2) Linear Phase-Varying (LPV), and 3) the Nonlinear Phase-Varying (NPV).

Model Name	Functional Form
Phase-Varying (PV)	$\hat{X}_\phi = F^w_{pv}(\phi)$
Linear Phase-Varying (LPV)	$\hat{X}_{\phi+\theta} = F^w_{lpv,\theta}(\phi)X_\phi$
Nonlinear Phase-Varying (NPV)	$\hat{X}_{\phi+\theta} = F^w_{npv,\theta}(\phi, X_\phi)$

Figure 7.5: Models surveyed

LPV and NPV take a state measurement at the phase, ϕ, and provide a state estimate, $\hat{X}_{\phi+\theta}$, a constant, θ, phase in the future. PV takes a phase, ϕ, and returns the estimated state mean, $\hat{X}_\phi$, at that phase. The models in the table are ordered in increasing functional complexity: the PV model is a function of phase, LPV is a function of phase and linearly state, and the NPV is a function of phase and state. Generally, the first two models (PV, LPV) would be considered specifically applicable to rhythmic phenomena while the NPV is a general supervised learning model.

7.4.1 Phase-Varying (PV)

$$F^w_{pv} : \phi \rightarrow \hat{X}_\phi \tag{7.2}$$

$$[F^w_{pv}]_i = \frac{1}{2}w_{0,i} + \sum_{h=1}^{H} w_{2h-1,i}\cos(h\phi) + \sum_{h=1}^{H} w_{2h,i}\sin(h\phi) \tag{7.3}$$

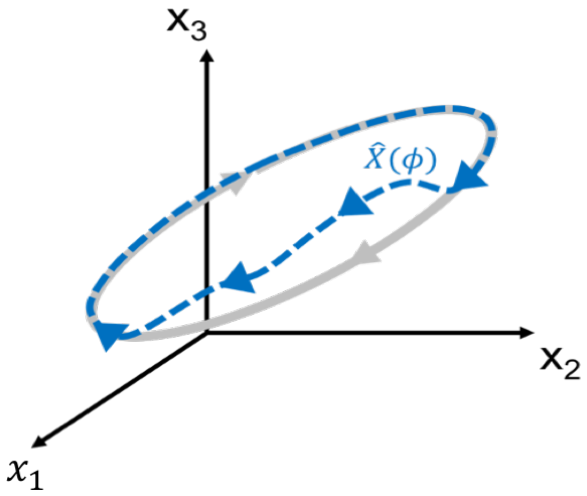


Figure 7.6: Depiction of the Phase-Varying (PV) model.

$$\arg\min_{w_i} \sum_{m=1}^{M} (F_{pv}^{w}(\hat{\Phi}_m)_i - D_{m,i})^2 \tag{7.4}$$

The PV model is relatively simple and specific to the task of rhythmic locomotion (a graphical depiction shown in Figure 7.8). The PV model draws on referenced work that estimates the instantaneous phase of a mechanical system near a stable limit cycle. It encodes the rhythmic nature of the system's limit cycle, the phase-varying mean, with a truncated Fourier Series, a signal composed out of the sum of sines and cosines. The function takes in phase, ϕ, and returns a prediction of the system's state, $\hat{X}_\phi \in \mathbb{R}^n$, at that moment. A matrix, $w \in \mathbb{R}^{n\times(2H+1)}$, parameterizes the truncated Fourier Series. Each component of the state, $i = 1, ..., n$, is predicted by a Fourier Series up to order H, a hyperparameter chosen to be 7. The periodic signal is fast enough to capture any rhythmic element 7 times faster than the base gait frequency. Accordingly, the PV function can be defined as the function in (7.3). The parameters, w, are chosen to minimize the error between the dataset, D, and the predicted value of the function as seen in (7.2).

7.4.2 Linear Phase-Varying (PV)

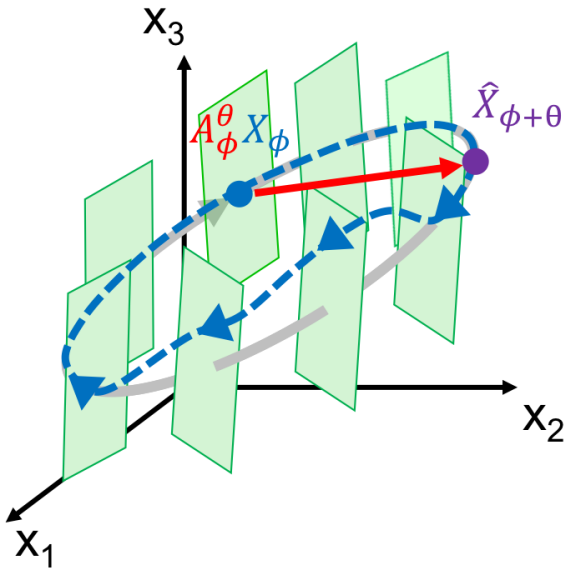


Figure 7.7: Depiction of the Linear Phase-Varying (LPV) model.

The LPV model builds upon the PV model with novel extensions, a graphical depiction is shown in Figure 7.7. Previous work has viewed rhythmic human locomotion as a behavior that can be modeled with a limit cycle oscillator [89]. The walking individual corrects for perturbations from a desired rhythmic behavior on the treadmill. This motivates increasing the complexity of the model, we attempt to model the stabilizing dynamics of the system along the limit cycle. More specifically, a behavior with discrete events, *i.e.* foot-falls and foot-liftoffs, can be described as a system near a stable limit cycle and thus we would expect the trajectories to exponentially decay towards the limit cycle [22].

The LPV function takes in a phase, ϕ, and returns a matrix, A_ϕ, of dimension $\mathbb{R}^{n\times n+1} \times \mathbb{R}^{n+1}$. The dimensions of the lookahead matrix is determined by the number of samples and the number of measurements plus one to make the operator affine. The function is parametrized by a tensor, w, of dimension $\mathbb{R}^{n\times n+1\times x\times(2H+1)}$. The second subscript, θ, specifies

the phase lookahead of the model.

$$F_{lpv}^{w} : \phi \rightarrow A_{\phi_k} \tag{7.5}$$

We first ran Phaser on the dataset to receive phase estimates for all the samples. Next, we use the PV model to subtract the phase mean from the dataset as the LPV model seeks to model the recovery dynamics of the subject in a rhythmic gait.

The first step in fitting the LPV models divides the phase space into L sections. At each of the sections we fit a discrete mapping, a matrix representative of an affine transformation. We weight the locality of all of the available data with the matrix, G_{ϕ_I}, indexed by $I = 1, ..., L$. The Gaussian weight scheme assigns greater weight to points closer in the entire phase space, $[0, 2\pi]$, to the section. The kernel width, a hyperparameter choice, was selected to be 0.5 making it roughly 16% of the gait cycle. We solved each of the posed weighted least squares problems with the Moore-Penrose inverse at each sample phase, ϕ_I. $\tilde{A}_{\phi_I}$ represents the discrete mapping for the section at ϕ_I. $X_\phi \in \mathbb{R}^{k \times n}$ is a matrix where each row is a single sample of the state, X_{ϕ_i}, and k is the number of samples that meet inclusion criteria for the set mapper, samples that have discrete mappings contained within the interval of phases sampled and thus can be interpolated.

$$G_{\phi_l} X_{\phi+\theta} = G_{\phi_l} X_\phi \tilde{A}_{\phi_l} \ for\ l = 1, ..., L, \phi_l = 2\pi(\frac{l-1}{L}) \tag{7.6a}$$

$$\tilde{A}_{\phi_l} = (X_\phi^T G_{\phi_l} X_\phi)^{-1} X_\phi^T G_{\phi_l} X_{\phi+\theta} \tag{7.6b}$$

Each of the entries in $F_{lpv,\theta}^{w}$ are represented by a 3rd-order Fourier Series, giving us a function that can continuously return a matrix with $n(n+1)$ distinct Fourier series. $[A_{\phi_I}^{w}]_{i,j}$ denotes i, j entry in the matrix returned by the function $F_{lpv,\theta}^{w}$.

$$[A_{\phi+\theta}^{w}]_{i,j} = \frac{1}{2} w_{i,j,0} + \sum_{h=1}^{H} w_{i,j,2h-1} \cos(h\phi) + \sum_{h=1}^{H} w_{i,j,2h} \sin(h\phi) \tag{7.7}$$

Lastly, we solve for the LPV model weights by fitting a truncated Fourier series to the corresponding entries in the discrete mapping matrices by minimizing deviation between the

evaluated continuous function, $A^w_{\phi_I}$, and the fitted discrete mapping, $\tilde{A}_{\phi_I}$. The optimization procedure is run independently for each Fourier series.

$$\arg\min_{w_{i.j}} \sum_{l=1}^{L} ([A^w_{\phi_l}]_{i,j} - [\tilde{A}^w_{\phi_l}]_{i,j})^2 \tag{7.8}$$

7.4.3 Nonlinear Phase-Varying (NPV)

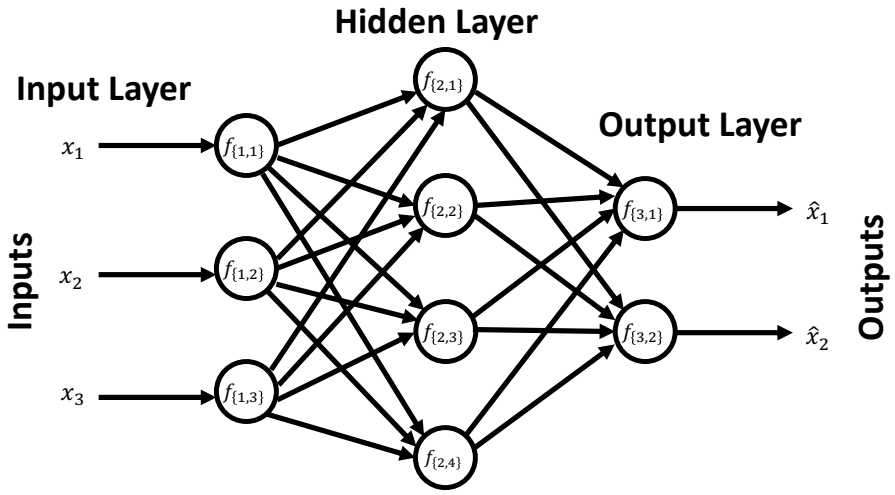


Figure 7.8: Representative configuration of the Nonlinear Phase-Varying (NPV) model.

Neural networks, particularly in a deep configuration, are considered state of the art predictors in a number of domains: image recognition, speech recognition, natural language processing, and robotics [77] [1]. There are many viable potential configurations and activation functions, for our model we chose a feed-forward neural network with three layers of depth to ensure the universal function approximation property [32]. Each layer is fully connected and 128 units wide, greater than or equal to the number of states to prevent encoding behaviors. Each unit used the sigmoid activation function. We implemented the model in the Keras Python framework and optimized with RMSprop gradient descent. To ensure a relevant comparison we picked an architecture that has the same input-output structure as

the LPV model. This constraint ruled out another competitive model, the recurrent neural network a state of the art predictor for time series data [85] [91]. The deep regressor takes in phase estimate, ϕ, and state, X_ϕ, and predicts the state at a fixed phase lookahead, $\hat{X}_{\phi+\theta}$. The interconnection weights and biases, the set of parameters of the used in the network are denoted by w.

7.5 Methods

Predicting human motion is challenging due to the high dimensionality of human locomotion, the variability in our periodic motion that in turn makes it rhythmic, and the nonlinearity of our dynamics [108].

7.5.1 Torque Estimation

$$\tau_{exo}(t) = \begin{cases} -k(\theta_{ankle}(t) - \theta^*), & \text{if } \theta_{ankle}(t) \geq \theta^* \\ 0, & \text{else if } \theta_{ankle}(t) < \theta^* \end{cases} \tag{7.9}$$

We currently view the torque applied by the AFO as a piecewise function, as seen in (7.9). Defining the variables used in the equation: k is the spring stiffness, θ_{ankle} the time-series of ankle angles, and θ^* is the equilibrium point for the device. The physical interpretation of this equation, the AFO applies a torque at the ankle when the foot's dorsiflexion angle is greater than the equilibrium angle. We augment the dataset with the estimated torque applied by the device, ten samples are uniformly taken from the initial to final phase of the lookahead window. Incorporating the torque profile in this manner assumes that the entire signal, not just impulse or final magnitude, influences future states.

7.5.2 Relative Remaining Variance

$$\mu = \mathbb{E} \tag{7.10a}$$

$$Var(X) = \mathbb{E}[(X - \mu)^2] \tag{7.10b}$$

$$RRV = \frac{Var(\hat{Y} - Y)}{Var(Y)} \tag{7.11}$$

The following results will use the relative remaining variance metric (RRV) shown in equation (7.9) and also used in closely related prior work [89]. The function in (7.10) defines the standard statistical variance computation, where the variance is defined as the square of the expected deviation from the mean. The timeseries from dataset, Y, and predictions given by the model, $\hat{Y}$. In order to obtain confidence intervals of the mean, the data is bootstrapped – a procedure where the residuals are repeatedly randomly sampled – and the mean is computed. This procedure can give us a distribution of computed means and a confidence that the mean is contained within certain intervals.

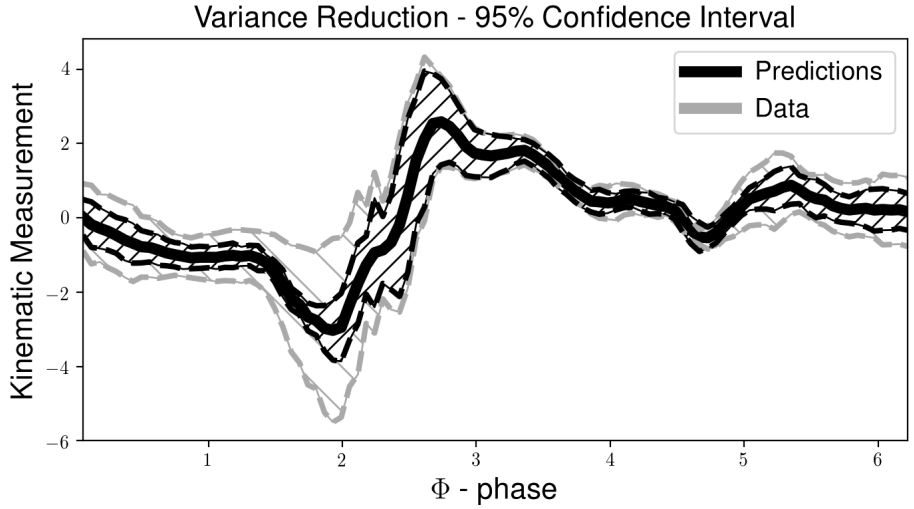


Figure 7.9: Reduction in relative remaining variance.

In Figure 7.9 we can see, for an example subject and kinematic measurement (joint angle) the observed reduction in variance of the residuals. For the RRV statistic, *0* indicates perfect prediction and *1* or anything greater points to a lack of any predictive power. Although unintuitive (since simply predicting the mean yields an RRV of *1*), RRV values greater than *1* are common with respect to predictions on real datasets. Figure 7.9 demonstrates the

reduction in the variance of the residuals varies with respect to phase indicating that the model's predictive performance varies within the gait cycle.

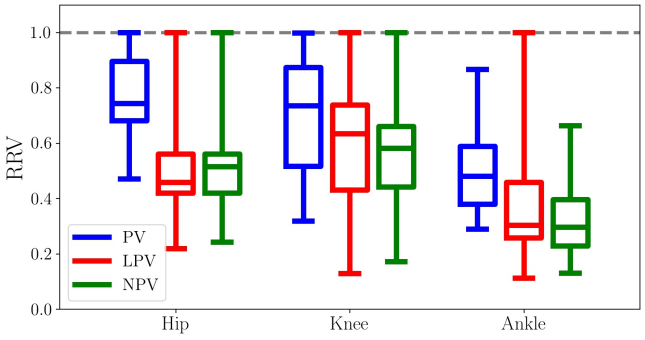
7.5.3 Predicting deviations from the nominal limit cycle

The surveyed models provide predictions for walking individuals at any given phase throughout the limit cycle. We adopt the assumption that the AFO acts as a perturbation, intermittently applied when the subject's foot is in contact with the ground. Deviations from the nominal limit cycle are induced by the AFO and natural variability of a walking individual. A PV model is trained on the control treatment for each subject. State estimates from the control PV model are then subtracted from all of the collected time series (including the control). The subsequent section, Section 7.6, reports the surveyed models ability to predict the remaining residuals. This prediction task is more challenging as a fraction of the variance is already explained by the control trained PV model. By removing the nominal limit cycle, represented by the control PV model, we aim to tease apart the ability of models surveyed in predicting motion relative to the limit cycle.

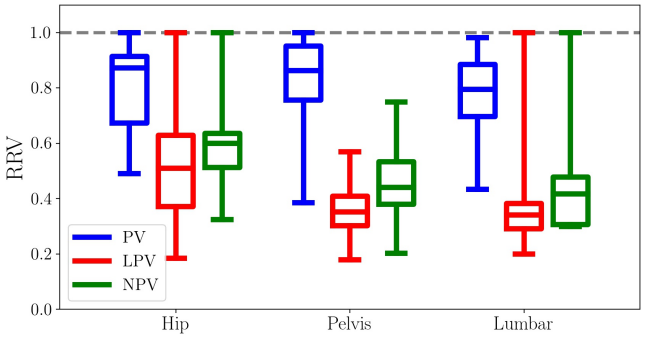
7.6 Results

This section covers the computational results for the surveyed models and their statistics for predicting joint angles and EMG. Results showing the changes in the predictive performance when varying the phase lookahead and reduction in training set size will also be presented. For all of the RRVs covered in this section the maximum value was capped at 1. An RRV with a value of 1 or greater suggests no predictive ability. The plots summarize across subjects and sometimes multiple joint angles and we did not want to skew the reported results with arbitrarily large RRV values. Lastly, the subjects had bilateral AFOs and we expect the participants to have roughly symmetric gaits. The results reported in this section pool predictions for the left and right joint angles when there are comparable measures, *eg* ankle, knee.

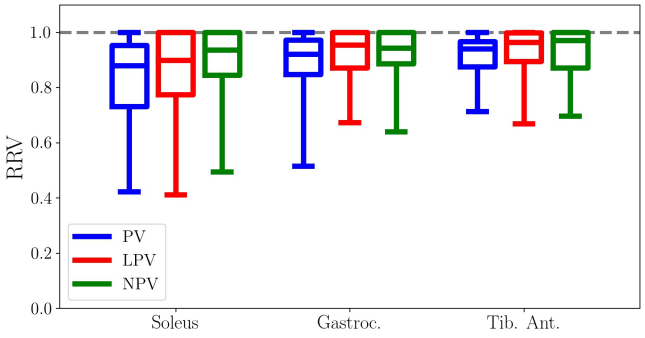
The sagittal and transversal plane joint angles shown in Figure 7.10 are a subset of all



(a) Transversal plane joint angles.



(b) Transversal plane joint angles.



(c) EMG RRV

Figure 7.10: RRV for joint angles and EMG for all subjects and surveyed models. The phase lookahead is $\frac{\pi}{4}$, an eighth of a stride.

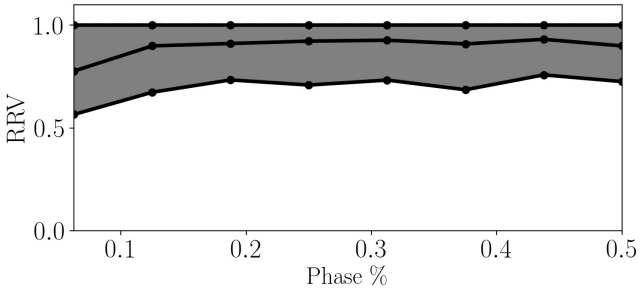
the available joint angles, we are reporting the dimensions most affected by the perturbation applied by the AFO. In general, the predictive performance of the LPV and NPV surpassed the PV. Usually, the joint angle RRVs were lower than the corresponding EMG values. We speculate the that EMG RRVs were higher because their dynamics are faster and more variable, making the prediction task more challenging.

The computed RRV values, shown in Figure 7.11, increase for the LPV model in tandem with the phase lookahead, *i.e.* percent of gait cycle. For the hip, knee, and ankle angles, the RRVs increase and then asymptote as the percent of gait cycle lookahead is increased. This effect is most pronounced in the knee angle.

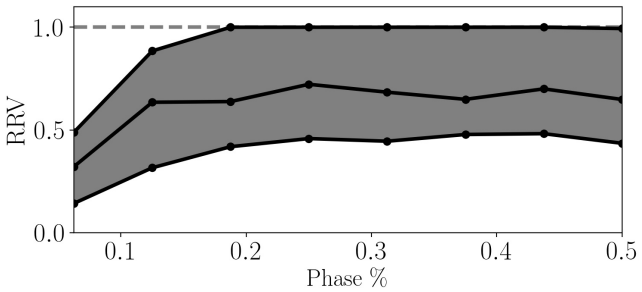
The positive relation between predictive ability and training set size is seen in Figure 7.12. The data is temporally organized, the first sample corresponds to the start of the trial and the last the finish. When reducing the size of the training set, the chosen interval starts from beginning of the available time series and continues sequentially. This allows us to gauge the effects of changing the duration of the trial, or the amount of clinical time, and how the predictive performance of the models varies. As the number of samples increases to 5000, shown in in Figure 7.12, all of the models displayed marked improvement. The NPV shows a large improvement for the full dataset, the largest sampled training set size. The NPV hyperparameters, along with the PV and LPV, were tuned on the full dataset. This may explain the sudden, large improvement observed to an extent in all of the models.

7.7 Discussion

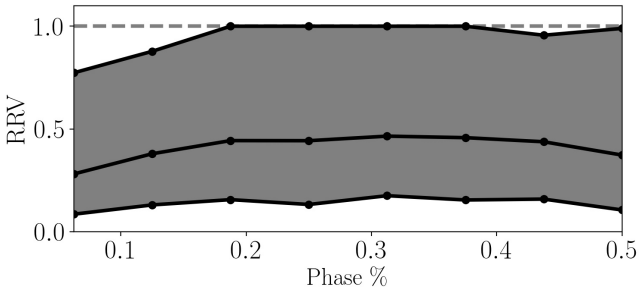
We have surveyed subject-specific models for predicting the response to AFOs. The models are a function of the estimated phase of the walking subjects along with joint angles, EMG measurements, and AFO torque profiles. Phaser, the algorithm we used, estimates the instantaneous phase from the configuration of the subject [106,107]. We leverage the regularity of steady-state walking to train the surveyed models and provide predictions on untrained stiffness values. Our modeling approach assumes that the subject undergo repeated perturbations from the AFO. This causes the individual to deviate from the unperturbed limit



(a) Sagittal hip angle.



(b) Sagittal knee angle.



(c) Sagittal ankle angle.

Figure 7.11: RRV for select joint angles for all subjects and surveyed models. The phase lookahead is varied for the LPV model. The upper bound is the .975 quantile and the lower bound is the .025 quantile.

cycle. The resultant shift is the basis of the proposed clinical intervention aiming to improve gait kinematics, metabolic efficiency, or other clinically relevant outcomes. Naturally, the applied perturbations needs to maintain or improve stability while retaining subject comfort.

Floquet theory says to a first-order dynamics are linear phase-varying near a limit cycle, *i.e.* exponentially stable periodic orbit [45, 56]. This inference underlines modeling the walking subjects with the LPV model. All of the surveyed models may apply to other rhythmic behaviors but the PV and LPV do not extend to aperiodic behaviors. The LPV and NPV are sensitive to the size of the training set size as seen in the eventual downward trend of RRV values as the number of samples increases in Figure 7.12. Requiring larger training set sizes reduces the viability of subject-specific models in a clinical setting. This is particularly relevant to the target patient population, *e.g.* cerebral palsy survivors, which display greater gait heterogeneity and are therefor more suited to subject-specific predictive models (versus population level models) [57]. For a more homogeneous target patient population, training models on a separate group of individuals before application in a clinical setting may reduce clinical time.

The results presented in the previous section build upon previous work where the Phaser algorithm was used on running individuals [89]. In this experiment, the individuals ran at a steady-state and were unperturbed. Additionally, the trials were significantly longer for each of the subjects, 24 minutes, and many were disqualified. In this paper, the authors were able to find discrete mappings which were stable. We speculate that the length of trials and more stringent exclusion criteria yielded a more consistent dataset to fit stable models *i.e.* discrete mappings.

All of the models surveyed are functions of phase. The PV model offers the best predictions at the lowest number of samples, seen in Figure 7.12, and offers a comparison baseline for the other two models (LPV, NPV). The PV model does not directly take into account state information and thus relies on deviations of the walking individuals to trend towards the mean. The PV model does not account for the magnitude of deviations from the estimated limit cycle and associated recovery dynamics. The LPV builds upon the previous

phase-based estimation techniques and theory while effectively matching the predictive performance of the NPV [56, 89]. At a given phase, the LPV returns an affine operator which allows it to predict first-order dynamics near a limit cycle. Finally, the NPV does not incorporate any *a priori* model structure for the task of predicting rhythmic locomotion. We suspect that the lack of in-built assumptions along with the universal function approximation property allows this model to scale with data. If future experiments were be run for longer periods of time we think that the inherent representational capabilities of the NPV would allow it to continue to scale. This potential capability comes with the trade off that the NPV may not have enough data to be effectively trained in a time constrained setting *e.g.* clinical.

This chapter reports on the phase lookahead modeling structure discussed in Section 7.3.2. An alternate structure we explored was predicting the state on a shorter time or phase scale. We first fitted discrete state to state mappings. In order to generate predictions over a longer time or phase scale, the given state would be propagated forward with the iterated mappings. Inspecting the eigenvalues of the fitted discrete mappings informed us of their stability. The mappings when fit to our experimental data had some eigenvalues outside of the unit circle, indicating instability. This posed problems when attempting to generate predictions as the iterated state would diverge from the limit cycle. As previously noted, this approach succeeded in prior work and we suspect our inability to find stable models was due to the length of trials [89]. Proposed future experiments would collect longer datasets, allowing the iterated models to provide additional insight into the dynamics of walking individuals and offer another tool in clinical interventions.

7.7.1 Generalizability

The results reported in Section 7.6 illustrate a tradeoff faced in model selection balancing between predictive performance and model simplicity [46]. The PV mode is simple and gives reasonable RRVs while the LPV and NPV model scale with additional data. This coincides with the notion that the quantity of data, to successfully train a model, increases

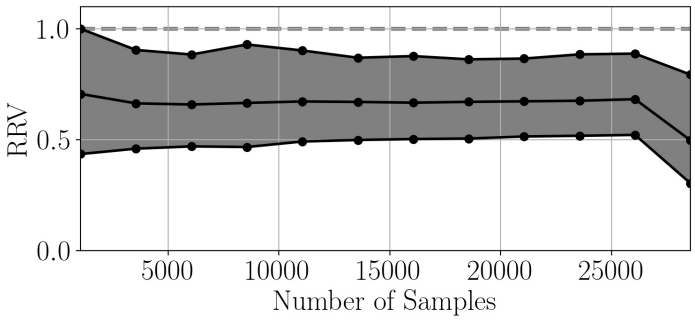
with the lack of assumptions built into the model. When we cannot *a priori* glean a model that predicts the desired task, collecting data and fitting a model presents an alternative. In a clinical setting collecting data is costly and time-consuming. We balance this with an attempt to provide the best possible clinical outcomes by finding an effective, reliable model which requires the least amount of data. Asurvey of available methods can help find the *Pareto* optimum where the ability to predict joint angles and EMG is not compromised by the choice of model.

7.7.2 Functional Complexity

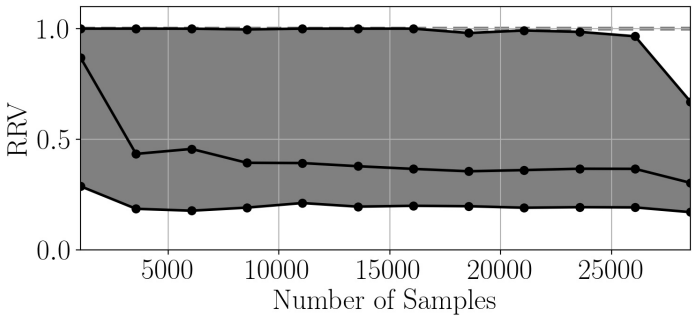
All of the models surveyed, as seen in Figure 7.5, display different functional complexities. The PV is nonlinearly a function of phase, LPV is nonlinearly a function of phase and linearly a function of state, and NPV is nonlinearly a function of phase and state. Although the functional forms do not provide any guarantees on the amount of data required for a chosen predictive performance, we suspect and Figure 7.12 serves as evidence, that their exists a relationship where a greater functional complexity offers a higher ceiling for goodness-of-fit provided the training data is representative of the testing data.

7.7.3 Conclusion

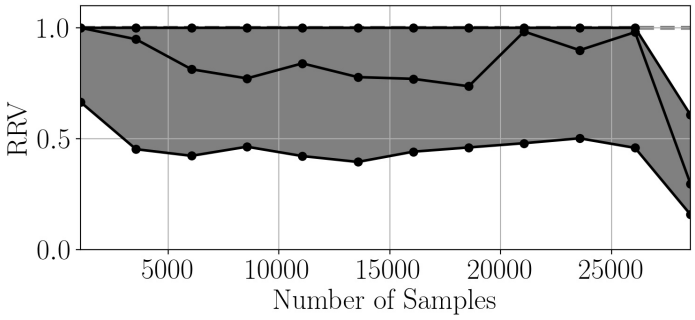
The surveyed models exhibited an ordering in predicting joint angles as the LPV and NPV surpassed the performance of the PV model, seen in Figure 7.10. No clear trend was observed in EMG, seen in Figure 7.10, as all of the models were almost non-predictive. Naturally, the computational time to train the models went in reverse order of their predictive performance. The NPV was the slowest, then LPV, and PV was the fastest. Once trained, the speed of predictions was relatively similar for all of the surveyed models. If the AFO was actively controlled - all the models surveyed could offer sufficiently fast predictions to enable real-time control. Overall, the results of this section indicate that the LPV approximately matches the predictive performance of the NPV, providing experimental evidence that individuals engaged in steady-state walking can be modeled with a linear phase-varying model.



(a) Training set size for the PV model.



(b) Training set size for the LPV model.



(c) Training set size for the NPV model.

Figure 7.12: The computed RRV across all joint angles and subjects. The size of the training set size is varied. The upper bound is the .975 quantile and the lower bound is the .025 quantile.

Chapter 8

CONCLUSION & FUTURE RESEARCH DIRECTIONS

8.1 Part I

In Chapter 4, we showed that nonsmoothness in trajectory outcomes is inherited in the resulting optimal policy and value functions. This nonsmoothness precludes the application of gradient-based optimization methods for controlling systems with contact-rich dynamics. A natural extension would be the subsequent development of theory and algorithms that enable optimization over discontinuous state transitions [95]. A hybrid dynamical system operating on or near a limit cycle can be described by a smooth reduced-order model [22]. Extending this result to finite time-horizon maneuvers would allow the application of optimization techniques that make smoothness assumptions. Alternatively, computational methods which can optimize over nonsmooth cost landscapes may be developed. A potential new optimization technique would smooth discontinuous trajectory outcomes with a kernel and then show that the resultant policy approximates the true optimal one. Both approaches outlined would further the ability to control mechanical systems subject to unilateral constraints and if successful would at least partially address the challenge associated with the transfer of control from simulation to real.

For the Hopper, described in Chapter 7, we suspect that for some parameters choices – reference configurations, controller parameters, and state initialization– all possible contact sequences could be observed for a given number of cycles. We have already found all contact sequences for a 2-cycle hop as discussed in Chapter 7. Matching the upper bound, n^m – where n is the number of contact sequences possible per cycle and m is the cycle horizon, for three or more hop cycles remains an open task. Increasing the sensitivity of the simulation may produce more contact sequences. Decreasing each time-step will yield a more accurate

discretization of the continuous dynamics and less constraint error will accumulate over time. This would decrease the decay of perturbations over time, a consequence of constraint stabilization, and would give more accurate simulation results. Second, we suspect that a hopping maneuver where both lower links are perpendicular to ground upon contact, splitting the *closed* and *open* contact configuration incidence angles, would display both alternating and repeating sequences when subjected to small perturbations away from the limit cycle. To find the parameters of this configuration, we could solve an optimization problem where the angle error from the desired trajectory would be minimized with respect to the parameter choices of the maneuver.

8.2 Part II

Now that we have run a pilot study on able-bodied subjects, a logical extension for the surveyed models is predicting locomotion for patient populations. Cerebral palsy survivors may be more challenging, due to greater gait heterogeneity, and therefor longer datasets or more stiffness training values may be required [57]. The subject *(ID:4)* has the largest reaction to the AFO torque applied and subsequently had the lowest RRV values for both joint angles and EMG. Additionally, increasing the range of spring stiffnesses may allow for more effective interventions by exploring a greater range of outcomes and thus clinical intervention.

All of the models surveyed follow the phase lookahead structure described in Chapter 6. When we attempted to fit small timescale models to the dynamics of the walking subject the resultant models were not stable, the linear models had eigenvalues which sat outside the unit circle. If we forward simulated the nonlinear models some of dimensions would exponentially diverge from the limit cycle. Prior work has indicated that a 24 minute experiment of steady-state running yielded stable linear models [89]. This suggests, that if we increase the length of the experiment we should be able to fit stable small timescale models. This would allow us to analyze and quantify the stability of the resultant AFO stiffness. Finally, attempting to fit a dynamical system to each contact configuration may prove more representative than

the input-output relation of the current surveyed models.

8.3 Data-driven hybrid dynamical systems

The contents of this thesis have sought to further the practical application of data-driven hybrid dynamical systems. The two applications in this thesis, legged locomotion and rehabilitative medicine, constitute a subset of all potential ones. Advancements in computing and sensing offer new tools to the engineering discipline, augmenting historical pen and paper analysis. Harnessing greater quantities and quality of data presents new avenues for the modeling and control of hybrid dynamical systems. The field of data-driven methods, fitting models and solving for control, contains many open problems with promise for substantial impact.

BIBLIOGRAPHY

[1] Pieter Abbeel, Adam Coates, Morgan Quigley, and Andrew Y Ng. An application of reinforcement learning to aerobatic helicopter flight. In *Advances in Neural Information Processing Systems*, 2007.

[2] Kazuyuki Aihara and Hideyuki Suzuki. Theory of hybrid dynamical systems and its applications to biological and medical systems. *Philosophical Transactions of The Royal Society A*, 368(1930):4893–4914, 2010.

[3] Mark A. Aizerman and Felix R. Gantmacher. Determination of stability by linear approximation of a periodic solution of a system of differential equations with discontinuous right–hand sides. *The Quarterly Journal of Mechanics and Applied Mathematics*, 11(4):385–398, 1958.

[4] Alin Albu-Schaffer, Oliver Eiberger, Markus Grebenstein, Sami Haddadin, Christian Ott, Thomas Wimbock, Sebastian Wolf, and Gerd Hirzinger. Soft robotics. *IEEE Robotics & Automation Magazine*, 15(3), 2008.

[5] Robert M. Alexander. The gaits of bipedal and quadrupedal animals. *The International Journal of Robotics Research*, 3(2):49–59, 1984.

[6] Aaron D. Ames, Robert D. Gregg, and Mark W. Spong. A geometric approach to three-dimensional hipped bipedal robotic walking. In *Decision and Control*. IEEE, 2007.

[7] Th. Arampatzis, John Lygeros, and Stamatis Manesis. A survey of applications of wireless sensors and wireless sensor networks. In *Mediterranean Conference on Control and Automation*. IEEE, 2005.

[8] Edith M. Arnold, Samuel R. Ward, Richard L. Lieber, and Scott L. Delp. A model of the lower limb for analysis of human movement. *Annals of Biomedical Engineering*, 38(2):269–279, 2010.

[9] Christopher G. Atkeson, B. P. W. Babu, N. Banerjee, D. Berenson, C. P. Bove, X. Cui, M. DeDonato, R. Du, S. Feng, and P. Franklin. What Happened at the Darpa Robotics Challenge, and Why? *Springer Tracts in Advanced Robotics*, 2018.

[10] Samuel Au, Max Berniker, and Hugh Herr. Powered ankle-foot prosthesis to assist level-ground and stair-descent gaits. *Neural Networks*, 21(4):654–666, 2008.

[11] Patrick Ballard. The dynamics of discrete mechanical systems with perfect unilateral constraints. *Archive for Rational Mechanics and Analysis*, 154(3):199–274, 2000.

[12] Samuel Barrett, Matthew E. Taylor, and Peter Stone. Transfer learning for reinforcement learning on a physical robot. In *International Conference on Autonomous Agents and Multiagent Systems*, 2010.

[13] Catherine E. Bauby and Arthur D. Kuo. Active control of lateral balance in human walking. *Journal of Biomechanics*, 33(11):1433–1440, 2000.

[14] Joachim Baumgarte. Stabilization of constraints and integrals of motion in dynamical systems. *Computer Methods in Applied Mechanics and Engineering*, 1(1):1–16, 1972.

[15] Jonathan Baxter, Peter L. Bartlett, and Lex Weaver. Experiments with infinite-horizon, policy-gradient estimation. *Journal of Artificial Intelligence Research*, 15:351–381, 2001.

[16] Michael Brady, John M. Hollerbach, Timothy L. Johnson, Tomás Lozano-Pérez, Matthew T. Mason, and Patrick Henry Winston. *Robot motion: Planning and control*. MIT press, 1982.

[17] Merel-Anne Brehm, Jaap Harlaar, and Michael Schwartz. Effect of ankle-foot orthoses on walking efficiency and gait in children with cerebral palsy. *Journal of Rehabilitation Medicine*, 40(7):529–534, 2008.

[18] Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. OpenAI Gym. *arXiv preprint arXiv:1606.01540*, 2016.

[19] Bernard Brogliato, A.A. ten Dam, Laetitia Paoli, Frank Genot, and Michel Abadie. Numerical simulation of finite dimensional multibody nonsmooth mechanical systems. *Applied Mechanics Reviews*, 55(2):107–150, 2002.

[20] Samuel Burden, Shai Revzen, and S. Shankar Sastry. Dimension reduction near periodic orbits of hybrid systems. In *Decision and Control and European Control Conference*. IEEE, 2011.

[21] Samuel A. Burden, Humberto Gonzalez, Ramanarayan Vasudevan, Ruzena Bajcsy, and S. Shankar Sastry. Metrization and simulation of controlled hybrid systems. *IEEE Transactions on Automatic Control*, 60(9):2307–2320, 2015.

[22] Samuel A. Burden, Shai Revzen, and S. Shankar Sastry. Model reduction near periodic orbits of hybrid dynamical systems. *IEEE Transactions on Automatic Control*, 60(10):2626–2639, 2015.

[23] Martin Buss, Markus Glocker, Michael Hardt, Oskar Von Stryk, Roland Bulirsch, and Günther Schmidt. Nonlinear hybrid dynamical systems: modeling, optimal control, and applications. In *Modelling, Analysis, and Design of Hybrid Systems*, pages 311–335. Springer, 2002.

[24] John Carling, Thelma L. Williams, and Graham Bowtell. Self-propelled anguilliform swimming: simultaneous solution of the two-dimensional Navier-Stokes equations and Newton's laws of motion. *Journal of Experimental Biology*, 201(23):3143–3166, 1998.

[25] Dean R. Chapman. Computational aerodynamics development and outlook. *AIAA Journal*, 17(12):1293–1313, 1979.

[26] Hsinchun Chen, Roger H.L. Chiang, and Veda C. Storey. Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 2012.

[27] Paul Christiano, Zain Shah, Igor Mordatch, Jonas Schneider, Trevor Blackwell, Joshua Tobin, Pieter Abbeel, and Wojciech Zaremba. Transfer from simulation to real world through learning deep inverse dynamics model. *arXiv preprint arXiv:1610.03518*, 2016.

[28] Steven H. Collins, M. Bruce Wiggin, and Gregory S. Sawicki. Reducing the energy cost of human walking using an unpowered exoskeleton. *Nature*, 522(7555):212, 2015.

[29] Steven H. Collins, Martijn Wisse, and Andy Ruina. A three-dimensional passive-dynamic walking robot with two legs and knees. *The International Journal of Robotics Research*, 20(7):607–615, 2001.

[30] Nikolaus Correll, Kostas E. Bekris, Dmitry Berenson, Oliver Brock, Albert Causo, Kris Hauser, Kei Okada, Alberto Rodriguez, Joseph M. Romano, and Peter R. Wurman. Analysis and observations from the first Amazon Picking Challenge. *IEEE Transactions on Automation Science and Engineering*, 15(1):172–188, 2016.

[31] C.F. Curtiss and Joseph O. Hirschfelder. Integration of stiff equations. *Proceedings of the National Academy of Sciences of the United States of America*, 38(3):235, 1952.

[32] George Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989.

[33] Fred Delcomyn. Neural basis of rhythmic behavior in animals. *Science*, 210(4469):492–498, 1980.

[34] Scott L. Delp, Frank C. Anderson, Allison S Arnold, Peter Loan, Ayman Habib, Chand T. John, Eran Guendelman, and Darryl G. Thelen. OpenSim: open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering*, 54(11):1940–1950, 2007.

[35] Michael H. Dickinson, Claire T. Farley, Robert J. Full, M.A.R. Koehl, Rodger Kram, and Steven Lehman. How animals move: an integrative view. *Science*, 288(5463):100–106, 2000.

[36] Ye Ding, Myunghee Kim, Scott Kuindersma, and Conor J. Walsh. Human-in-the-loop optimization of hip assistance with a soft exosuit during walking. *Science Robotics*, 3(15), 2018.

[37] Aaron M. Dollar and Hugh Herr. Lower extremity exoskeletons and active orthoses: challenges and state-of-the-art. *IEEE Transactions on Robotics*, 24(1):144–158, 2008.

[38] J. Maxwell Donelan, Rodger Kram, and Arthur D. Kuo. Mechanical work for step-to-step transitions is a major determinant of the metabolic cost of human walking. *Journal of Experimental Biology*, 205(23):3717–3727, 2002.

[39] Kenji Doya. Reinforcement learning in continuous time and space. *Neural Computation*, 12(1):219–245, January 2000.

[40] Michael F. Eilenberg, Hartmut Geyer, and Hugh Herr. Control of a powered ankle–foot prosthesis based on a neuromuscular model. *IEEE transactions on Neural Systems and Rehabilitation Engineering*, 18(2):164–173, 2010.

[41] Tom Erez and Emo Todorov. Trajectory optimization for domains with contacts using inverse dynamics. In *Intelligent Robots and Systems*. IEEE, 2012.

[42] Nima Fazeli, Samuel Zapolsky, Evan Drumwright, and Alberto Rodriguez. Learning data-efficient rigid-body contact models: Case study of planar impact. *arXiv preprint arXiv:1710.05947*, 2017.

[43] Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. *arXiv preprint arXiv:1703.03400*, 2017.

[44] Ulla-Britt Flansbjer, Anna Maria Holmbäck, David Downham, Carolynn Patten, and Jan Lexell. Reliability of gait performance tests in men and women with hemiparesis after stroke. *Journal of Rehabilitation Medicine*, 37(2):75–82, 2005.

[45] Gaston Floquet. Sur les équations différentielles linéaires à coefficients périodiques. *Annales Scientifiques de l'École Normale Supérieure*, 12:47–88, 1883.

[46] Malcolm R. Forster. Key concepts in model selection: Performance and generalizability. *Journal of Mathematical Psychology*, 44(1):205–231, 2000.

[47] Robert J. Full and Daniel E. Koditschek. Templates and anchors: neuromechanical hypotheses of legged locomotion on land. *Journal of Experimental Biology*, 202(23):3325–3332, 1999.

[48] Kevin C. Galloway, Galen Clark Haynes, B. Deniz Ilhan, Aaron M. Johnson, Ryan Knopf, Goran A. Lynch, Benjamin N. Plotnick, Mackenzie White, and Daniel E. Koditschek. X-RHex: A highly mobile hexapedal robot for sensorimotor tasks. 2010.

[49] Hartmut Geyer, Reinhard Blickhan, and Andre Seyfarth. Natural dynamics of spring-like running: Emergence of selfstability. In *International Conference on Climbing and Walking Robots*, 2002.

[50] Hartmut Geyer, Andre Seyfarth, and Reinhard Blickhan. Spring-mass running: simple approximate solution and application to gait stability. *Journal of Theoretical Biology*, 232(3):315–328, 2005.

[51] Hartmut Geyer, Andre Seyfarth, and Reinhard Blickhan. Compliant leg behaviour explains basic dynamics of walking and running. *Proceedings of The Royal Society B*, 273(1603):2861–2867, 2006.

[52] Raffaele M. Ghigliazza, Richard Altendorfer, Philip Holmes, and D. Koditschek. A simply stabilized running model. *SIAM Journal on Applied Dynamical Systems*, 2(2):187–218, 2003.

[53] Rafal Goebel, Ricardo G. Sanfelice, and Andrew R. Teel. Hybrid dynamical systems. *IEEE Control Systems*, 29(2):28–93, 2009.

[54] Sten Grillner. Neurobiological bases of rhythmic motor acts in vertebrates. *Science*, 228(4696):143–149, 1985.

[55] Jessy W. Grizzle, Christine Chevallereau, Ryan W. Sinnet, and Aaron D. Ames. Models, feedback control, and open problems of 3d bipedal robotic walking. *Automatica*, 50(8):1955–1988, 2014.

[56] John Guckenheimer and Philip Holmes. *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*. Springer, 1983.

[57] Geoffrey G. Handsfield, Craig H. Meyer, Mark F. Abel, and Silvia S. Blemker. Heterogeneity of muscle sizes in the lower limbs of children with cerebral palsy. *Muscle & Nerve*, 53(6):933–945, 2016.

[58] Andrew H. Hansen, Dudley S. Childress, Steve C. Miff, Steven A. Gard, and Kent P. Mesplay. The human ankle during walking: implications for design of biomimetic ankle prostheses. *Journal of Biomechanics*, 37(10):1467–1474, 2004.

[59] Ross L. Hatton and Howie Choset. Geometric motion planning: The local connection, Stokes' theorem, and the importance of coordinate choice. *The International Journal of Robotics Research*, 30(8):988–1014, 2011.

[60] Marco Hutter, Christian Gehring, Michael Bloesch, Mark A. Hoepflinger, C. David Remy, and Roland Siegwart. StarlETH: A compliant quadrupedal robot for fast, efficient, and versatile locomotion. In *International Conference on Climbing and Walking Robots*, 2012.

[61] Marco Hutter, Mark A. Hoepflinger, Christian Gehring, Michael Bloesch, C. David Remy, and Roland Siegwart. Hybrid operational space control for compliant legged systems. *Robotics*, page 129, 2013.

[62] Marco Hutter, C. David Remy, Mark A. Höpflinger, and Roland Siegwart. Slip running with an articulated robotic leg. In *Intelligent Robots and Systems*. IEEE, 2010.

[63] Jemin Hwangbo, Joonho Lee, Alexey Dosovitskiy, Dario Bellicoso, Vassilios Tsounis, Vladlen Koltun, and Marco Hutter. Learning agile and dynamic motor skills for legged robots. *Science Robotics*, 4(26), 2019.

[64] Yildirim Hürmüzlü and D. B. Marghitu. Rigid body collisions of planar kinematic chains with multiple contact points. *The International Journal of Robotics Research*, 13(1):82–92, 1994.

[65] Antony Jameson, Wolfgang Schmidt, and Eli Turkel. Numerical solution of the Euler equations by finite volume methods using Runge-Kutta time stepping schemes. In *Fluid and Plasma Dynamics Conference*, 1981.

[66] A. M. Johnson, S. A. Burden, and D. E. Koditschek. A hybrid systems model for simple manipulation and self-manipulation systems. *The International Journal of Robotics Research*, 35(11):1354–1392, 2016.

[67] Aaron M. Johnson and Daniel E. Koditschek. Legged self-manipulation. *IEEE Access*, 1:310–334, 2013.

[68] Matthew Johnson, Brandon Shrewsbury, Sylvain Bertrand, Tingfan Wu, Daniel Duran, Marshall Floyd, Peter Abeles, Douglas Stephen, Nathan Mertins, Alex Lesman, and others. Team IHMC's lessons learned from the DARPA robotics challenge trials. *Journal of Field Robotics*, 32(2):192–208, 2015.

[69] Sham Kakade. A natural policy gradient. In *Advances in Neural Information Processing Systems*, 2001.

[70] Daniel A. Kingsley, Roger D. Quinn, and Roy E. Ritzmann. A cockroach inspired robot with artificial muscles. In *Intelligent Robots and Systems*. IEEE, 2006.

[71] Nate Kohl and Peter Stone. Policy gradient reinforcement learning for fast quadrupedal locomotion. In *International Conference on Robotics and Automation*. IEEE, 2004.

[72] Vijay R. Konda and John N. Tsitsiklis. On Actor-Critic Algorithms. *SIAM Journal on Control and Optimization*, 42(4):1143–1166, January 2003.

[73] Scott Kuindersma, Robin Deits, Maurice Fallon, Andrés Valenzuela, Hongkai Dai, Frank Permenter, Twan Koolen, Pat Marion, and Russ Tedrake. Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot. *Autonomous Robots*, 40(3):429–455, 2016.

[74] Vikash Kumar, Emo Todorov, and Sergey Levine. Optimal control with learned local models: Application to dexterous manipulation. In *International Conference on Robotics and Automation*. IEEE, May 2016.

[75] L. Kuntz and S. Scholtes. Structural Analysis of Nonsmooth Mappings, Inverse Functions, and Metric Projections. *Journal of Mathematical Analysis and Applications*, 188(2):346–386, 1994.

[76] Arthur D. Kuo. Energetics of actively powered locomotion using the simplest walking model. *Journal of Biomechanical Engineering*, 124(1):113–120, 2002.

[77] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436, 2015.

[78] John M. Lee. *Introduction to Smooth Manifolds*. Springer, New York ; London, 2nd ed. edition, 2012.

[79] Sergey Levine and Pieter Abbeel. Learning neural network policies with guided policy search under unknown dynamics. In *Advances in Neural Information Processing Systems*, 2014.

[80] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End-to-end training of deep visuomotor policies. *The Journal of Machine Learning Research*, 17(1):1334–1373, 2016.

[81] Sergey Levine, Chelsea Finn, Trevor Darrell, and Pieter Abbeel. End–to–end training of deep visuomotor policies. *Journal of Machine Learning Research*, 17(1):1334–1373, 2016.

[82] Chen Li, Tingnan Zhang, and Daniel I. Goldman. A terradynamics of legged locomotion on granular media. *Science*, 339(6126):1408–1412, 2013.

[83] Thomas Libby, Aaron M. Johnson, Evan Chang-Siu, Robert J. Full, and Daniel E. Koditschek. Comparative Design, Scaling, and Control of Appendages for Inertial Reorientation. *IEEE Transactions on Robotics*, 32(6):1380–1398, 2016.

[84] Pei-Chun Lin, Haldun Komsuoglu, and Daniel E. Koditschek. Sensor data fusion for body state estimation in a hexapod robot with dynamical gaits. *IEEE Transactions on Robotics*, 22(5):932–943, 2006.

[85] Zachary C. Lipton, John Berkowitz, and Charles Elkan. A critical review of recurrent neural networks for sequence learning. *arXiv preprint arXiv:1506.00019*, 2015.

[86] Roger Y. Lu, Zeses Karoutas, and T.L. Sham. CASL virtual reactor predictive simulation: Grid-to-rod fretting wear. *JOM*, 63(8):53, 2011.

[87] Ryan D. Maladen, Yang Ding, Chen Li, and Daniel I. Goldman. Undulatory swimming in sand: Subsurface locomotion of the sandfish lizard. *Science*, 325(5938):314–318, 2009.

[88] Matthew T. Mason, Alberto Rodriguez, Siddhartha S. Srinivasa, and Andrés S. Vazquez. Autonomous manipulation with a general-purpose simple hand. *The International Journal of Robotics Research*, 31(5):688–703, 2012.

[89] Horst-Moritz Maus, Shai Revzen, John Guckenheimer, Christian Ludwig, Johann Reger, and Andre Seyfarth. Constructing predictive models of human running. *Journal of the Royal Society Interface*, 12(103):20140899, 2015.

[90] Tad McGeer. Passive dynamic walking. *The International Journal of Robotics Research*, 9(2):62–82, 1990.

[91] Gábor Melis, Chris Dyer, and Phil Blunsom. On the state of the art of evaluation in neural language models. *arXiv preprint arXiv:1707.05589*, 2017.

[92] Igor Mordatch, Kendall Lowrey, and Emo Todorov. Ensemble-CIO: Full-body dynamic motion planning that transfers to physical humanoids. In *Intelligent Robots and Systems*. IEEE, 2015.

[93] Igor Mordatch, Emanuel Todorov, and Zoran Popović. Discovery of complex behaviors through contact-invariant optimization. *ACM Transactions on Graphics*, 31(4):43, 2012.

[94] Elaine Owen. The importance of being earnest about shank and thigh kinematics especially when using ankle-foot orthoses. *Prosthetics and Orthotics International*, 34(3):254–269, 2010.

[95] Andrew M. Pace and Samuel A. Burden. Decoupled limbs yield differentiable trajectory outcomes through intermittent contact in locomotion and manipulation. In *International Conference on Robotics and Automation*. IEEE, 2017.

[96] Andrew M. Pace and Samuel A. Burden. Piecewise–differentiable trajectory outcomes in mechanical systems subject to unilateral constraints. In *Hybrid Systems: Computation and Control*, 2017.

[97] Jan Peters and Stefan Schaal. Policy gradient methods for robotics. In *Intelligent Robots and Systems*. IEEE, 2006.

[98] Elijah Polak. *Optimization: Algorithms and Consistent Approximations*. Springer–Verlag, 1997.

[99] Michael Posa, Cecilia Cantu, and Russ Tedrake. A direct method for trajectory optimization of rigid bodies through contact. *The International Journal of Robotics Research*, 33(1):69–81, 2014.

[100] Michael Posa, Scott Kuindersma, and Russ Tedrake. Optimization and stabilization of trajectories for constrained dynamical systems. In *International Conference on Robotics and Automation*. IEEE, 2016.

[101] Alfio Quarteroni, Riccardo Sacco, and Fausto Saleri. *Numerical Mathematics*. Springer, 2007.

[102] Apoorva Rajagopal, Christopher L. Dembia, Matthew S. DeMers, Denny D. Delp, Jennifer L. Hicks, and Scott L. Delp. Full-body musculoskeletal model for muscle-driven simulation of human gait. *IEEE Transactions on Biomedical Engineering*, 63(10):2068–2079, 2016.

[103] Stephane Redon, Abderrahmane Kheddar, and Sabine Coquillart. Gauss' least constraints principle and rigid body simulations. In *International Conference on Robotics and Automation*. IEEE, 2002.

[104] Jacob Reher, Eric A. Cousineau, Ayonga Hereid, Christian M. Hubicki, and Aaron D. Ames. Realizing dynamic and efficient bipedal locomotion on the humanoid robot DURUS. In *International Conference on Robotics and Automation*. IEEE, 2016.

[105] C. David Remy, Keith Buffinton, and Roland Siegwart. Stability analysis of passive dynamic walking of quadrupeds. *The International Journal of Robotics Research*, 29(9):1173–1185, 2010.

[106] Shai Revzen, Samuel A. Burden, Talia Y. Moore, Jean-Michel Mongeau, and Robert J. Full. Instantaneous kinematic phase reflects neuromechanical response to lateral perturbations of running cockroaches. *Biological Cybernetics*, 107(2):179–200, 2013.

[107] Shai Revzen and John M. Guckenheimer. Estimating the phase of synchronized oscillators. *Physical Review E*, 78(5):051907, 2008.

[108] Shai Revzen and John M. Guckenheimer. Finding the dimension of slow dynamics in a rhythmic system. *Journal of The Royal Society Interface*, 9(70):957–971, 2012.

[109] Andrew J. Ries, Tom F. Novacheck, and Michael H. Schwartz. A data driven model for optimal orthosis selection in children with cerebral palsy. *Gait & Posture*, 40(4):539–544, 2014.

[110] Andrew J. Ries, Tom F. Novacheck, and Michael H. Schwartz. The efficacy of ankle-foot orthoses on improving the gait of children with diplegic cerebral palsy: a multiple outcome analysis. *PM&R*, 7(9):922–929, 2015.

[111] Stephen M. Robinson. An Implicit-Function Theorem for a Class of Nonsmooth Functions. *Mathematics of Operations Research*, 16(2):292–309, 1991.

[112] Jacob Roll, Alberto Bemporad, and Lennart Ljung. Identification of piecewise affine systems via mixed-integer programming. *Automatica*, 40(1):37–50, 2004.

[113] Daniela Rus and Michael T. Tolley. Design, fabrication and control of soft robots. *Nature*, 521(7553):467, 2015.

[114] Alberto Sanfeliu, Norihiro Hagita, and Alessandro Saffiotti. Network robot systems. *Robotics and Autonomous Systems*, 56(10):793–797, 2008.

[115] John Schmitt and Philip Holmes. Mechanical models for insect locomotion: dynamics and stability in the horizontal plane I. Theory. *Biological Cybernetics*, 83(6):501–515, 2000.

[116] Stefan Scholtes. *Introduction to Piecewise Differentiable Equations*. Springer–Verlag, 2012.

[117] John Schulman, Sergey Levine, Philipp Moritz, Michael I. Jordan, and Pieter Abbeel. Trust region policy optimization. In *International Conference on Machine Learning*, 2015.

[118] Gerrit Schultz and Katja Mombaur. Modeling and optimal control of human-like running. *IEEE/ASME Transactions on Mechatronics*, 15(5):783–792, 2010.

[119] Andre Seyfarth, Hartmut Geyer, Michael Günther, and Reinhard Blickhan. A movement criterion for running. *Journal of Biomechanics*, 35(5):649–655, 2002.

[120] David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. Deterministic policy gradient algorithms. In *International Conference on Machine Learning*, 2014.

[121] Mark W. Spong, Seth Hutchinson, and Mathukumalli Vidyasagar. *Robot Modeling and Control*. Wiley New York, 2006.

[122] Manoj Srinivasan and Andy Ruina. Computer optimization of a minimal biped model discovers walking and running. *Nature*, 439(7072):72, 2006.

[123] Steven H. Strogatz. *Nonlinear Dynamics and Chaos: with Applications to Physics, Biology, Chemistry, and Engineering*. CRC Press, 2018.

[124] Richard S. Sutton, David McAllester, Satinder Singh, and Yishay Mansour. Policy gradient methods for reinforcement learning with function approximation. In *Advances in Neural Information Processing Systems*, 2000.

[125] Russ Tedrake, Teresa Weirui Zhang, Ming-fai Fong, and H. Sebastian Seung. Actuating a simple 3d passive dynamic walker. In *International Conference on Robotics and Automation*. IEEE, 2004.

[126] Gerald Teschl. *Ordinary differential equations and dynamical systems*, volume 140. American Mathematical Soc., 2012.

[127] Francisco J. Valero-Cuevas. *Fundamentals of Neuromechanics*. Springer, 2016.

[128] Bram Vanderborght, Björn Verrelst, Ronald Van Ham, Michaël Van Damme, Pieter Beyl, and Dirk Lefeber. Development of a compliance controller to reduce energy consumption for bipedal robots. *Autonomous Robots*, 24(4):419–434, 2008.

[129] Ricardo Vilalta and Youssef Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2):77–95, 2002.

[130] Dario J. Villarreal and Robert D. Gregg. A survey of phase variable candidates of human locomotion. In *Engineering in Medicine and Biology Society*. IEEE, 2014.

[131] Wei Wang, Robert N.K. Loh, and Edward Y. Gu. Passive compliance versus active compliance in robot-based automated assembly systems. *Industrial Robot: An International Journal*, 25(1):48–57, 1998.

[132] Yu Wang and Matthew Mason. Modeling impact dynamics for robotic operations. In *International Conference on Robotics and Automation*. IEEE, 1987.

[133] Jack M. Winters. Hill-based muscle models: a systems engineering perspective. In *Multiple Muscle Systems*, pages 69–93. Springer, 1990.

[134] Cornell Wright, Aaron Johnson, Aaron Peck, Zachary McCord, Allison Naaktgeboren, Philip Gianfortoni, Manuel Gonzalez-Rivero, Ross Hatton, and Howie Choset. Design of a modular snake robot. In *Intelligent Robots and Systems*. IEEE, 2007.

[135] Albert Wu and Hartmut Geyer. The 3-D Spring-Mass Model Reveals a Time-Based Deadbeat Control for Highly Robust Running and Steering in Uncertain Environments. *IEEE Transactions on Robotics*, 29(5):1114–1124, 2013.

[136] Weitao Xi and C. David Remy. Optimal gaits and motions for legged robots. In *Intelligent Robots and Systems*. IEEE, 2014.

[137] Juanjuan Zhang, Pieter Fiers, Kirby A. Witte, Rachel W. Jackson, Katherine L. Poggensee, Christopher G. Atkeson, and Steven H. Collins. Human-in-the-loop optimization of exoskeleton assistance during walking. *Science*, 356(6344):1280–1284, 2017.