Modeling and Approximation of Nonlinear Dynamics of Flapping Flight

Shirin Dadashi

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Mechanical Engineering

Andrew J. Kurdila, Chair
Javid Bayandor
Alexander Leonessa
Craig Woolsey
Rolf Müller

February 15, 2017
Blacksburg, Virginia

Keywords: Flapping Flight, Adaptive Control, Learning, Discrete Mechanics

Modeling and Approximation of Nonlinear Dynamics of Flapping Flight

Shirin Dadashi

Abstract

The underlying governing equations of biologically inspired robotic systems consist of a set of coupled nonlinear differential equations. Using analytical mechanics, we can derive dynamic equations of flapping wing robots that are similar to serial kinematic chains. However, the contribution of aerodynamic forces in these equations are not in the ordinary form. In fact, CFD simulations show that these forces are best described by history dependent profiles. Therefore the usual ordinary differential equations turn to functional differential equations. We construct these functional differential equations using integral operators and distributed parameter systems. We define a set of fixed history dependent integral operators and set an unknown parameter to reflect the weight in which a particular operator contributes to the aerodynamic forces. In the first part of this work we present an algorithm to identify the history dependent profile of lift and drag forces and prove the existence of solutions for such differential equations. Then we present sufficient conditions for the convergence of online estimation methods. We show that the approximation error improves by increasing the resolution of the estimation space and calculate the rate of convergence to the true model. Also, we propose an sliding mode adaptive controller strategy for the history dependent functional differential equation.

In the second part of this work we use camera observations for generating discrete equations of discrete mechanics that evolve near submanifols of configuration space. The idea comes from the analysis of animal motion data which suggests the existence of low dimensional sub manifolds of motion in the configuration manifold of the system of interest. The discrete camera observations are concentrated near the motion submanifold by a probability measure ρ. We derive error bounds for the discrete Lagrangian formulation in terms of the difference between the empirical Lagrangian L_z which is constructed using the camera observations and the limiting Lagrangian L_ρ. We derive a probabilistic error bound on discrete Lagrangian exploiting techniques from learning theory on reproducing kernel Hilbert spaces (RKHS). We validate our algorithm using a simple numerical example and present an open problem for future research direction.

Modeling and Approximation of Nonlinear Dynamics of Flapping Flight

Shirin Dadashi

General Audience Abstract

In this work we focus on modeling flapping flight mechanics by focusing our attention in two aspects of modeling. We first model the behavior of aerodynamic forces in charge of keeping the flying animal airborn. We present a mathematical model for history dependent profile of these forces. Also, we propose a novel adaptive controller to compensate these unknown forces in the dynamic model of the system. We also propose an algorithm to derive dynamic equations of the animal motion by using video data. We expect the model derived by this novel method to emulate the animal motion closely.

Dedication

To my best friend and husband *Mehdi*

Acknowledgments

First and foremost, I would like to thank my advisor Dr. Andrew Kurdila. He is going to be my role model and mentor for the rest of my life. I cannot thank him enough for supporting me through my PhD in different aspects. I still wonder how lucky I was to get a chance of working with him and to be exposed to his deep knowledge in addition to his academic manners. He is the best teacher and mentor in my life.

I am also very thankful to my friend and partner in crime Mathew Bender. I will miss our precious coffee times when we could talk about our research problems and anything else. I wish him the best throughout his life.

I am thankful to Dr. Alexander Leonessa for many many things that I have learned from him from adaptive control to Nonlinear Systems. His presence in the Lab had been a source of energy and high spirits for me and other students and his bright and sharp research comments had been source of inspiration for us.

I am also enourmously thankful to my best friend and husband Mehdi who has been always there for me in the journey of life. His love and support has empowered me through the ups and downs of my studies and my life.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

The first and most imperative step when designing a biologically inspired robot is to identify the underlying mechanics of the system or animal of interest. It is most common, perhaps, that this process generates a set of coupled nonlinear ordinary or partial differential equations. For this class of systems, the models derived from morphologyof the skeleton are usually very high dimensional, nonlinear, and complex. This is particularly true if joint and link flexibility are included in the model. In addition to complexities that arise from morphology of the animal, some of the external forces that influence the dynamics of animal motion are very hard to model. A very well-established example of these forces is the unsteady aerodynamic forces applied to the wings and the body of insects, birds, and bats. These forces result from the interaction of the flapping motion of the wing and the surrounding air. These forces generate lift and drag during flapping flight regime. As a result, they play a significant role in the description of the physics that underlies such systems. In this research we focus on dynamic and kinematic models that govern the motion of ground based robots that emulate flapping flight. The restriction to ground based biologically inspired robotic systems is predicated on two observations. First, it has become increasingly popular to design and fabricate bioinspired robots for wind tunnel studies. Second, by restricting the robotic systems to be anchored in an inertial frame, the robotic equations of motionare well understood, and we can focus attention on flapping wing aerodynamics for such nonlinear systems. We study nonlinear modeling, identification, and control problems that feature the above complexities. This document summarizes research progress and plans that focuses on two key aspects of modeling, identification, and control of nonlinear dynamics associated with flapping flight. The research carried out in this dissertation has been disseminated at top notch peer reviewdconferences including the American Control Conference (ACC), The Conference on Desision and Control (CDC) as follows

- Peer reviewed Conference Papers:
  - Shirin Dadashi, Hunter McClelland, Andrew Kurdila, **Learning Empirical Potentials for Modeling Discrete Mechanics**, to be presented at 2017 American Control Conference
  - Shirin Dadashi, Parag Bobade, Andrew J. Kurdila, **Error Estimates for Multiwavelet Approximations of a Class of History Dependent Operators**, IEEE Conference on Decission and Control 2016, Las Vegas
  - S. Dadashi, J. Feaster, G. Bledt, J. Bayandor, F. Battaglia and A. Kurdila, **Adaptive control for Flapping Wing Robots with History Dependent, Unsteady Aerodynamics**, IEEE American Control Conference 2015, Chicago
  - Shirin Dadashi, Jessica Gregory, Yu Lei , Matt Bender, Andrew Kurdila, Javid Bayandor, Rolf Muller, **Adaptive Control of a FlappingWing Robot Inspired by Bat Flight**, AIAA SciTech, AIAA Guidance, Navigation, and Control Conference 13-17 January 2014, National Harbor, Maryland
  - Ian P. Murphy, Shirin Dadashi, Jessica Gregory, Yu Lei, Javid Bayandor, Andrew

Kurdila, **Modeling and Adaptive Control for Tracking Wing Trajectories**, IMECE 2013, Nov 15-21, 2013, San Diego, USA

– J. Bayandor, G. Bledt, S. Dadashi, A. Kurdila, I. Murphy and Y. Lei, **Adaptive Control for Bioinspired Flapping Wing Robots**, IEEE American Control Conference 2013, Washington DC

Furthermore the research in this dissertation has been published or submitted to the following exceptional refereed journals:

- S. Dadashi, J. Feaster, J. Bayandor, F. Battaglia, A. J. Kurdila, **Identification and adaptive control of history dependent unsteady aerodynamics for a flapping insect wing**, Nonlinear Dynamics, 08 April 2016
- Shirin Dadashi, Parag Bobade, Andrew J. Kurdila, **Online Estimation and Adaptive Control for a Class of History Dependent Functional Differential Equations**
- Shirin Dadashi, Hunter McClelland, Andrew Kurdila, **Error Estimates for Discrete Lagrangian Mechanics with Empirical Potentials**

1.2 Introduction

Flapping flight achieved by insects, bats and birds is unsteady, nonlinear and is not as well-characterized as conventional fixed wing aerodynamic theory. A steady or quasi-steady model based on fixed wing aerodynamics is unlikely to provide adequate explanation of flapping flight for animals with articulated flapping wings. The richness of the problem of characterizing such dynamic systems has motivated the interest of the flapping flight research community to explain the physics underlying flapping flight. There have been significant efforts to model flapping flight aerodynamics in the last twenty years. References [1, 2, 3, 4, 5] are representative examples of such attempts. It has been generally acknowledged that an unsteady aerodynamic model is better suited for modeling the dynamics of flapping flight. Unsteady aerodynamics also plays a central, important role in the study of stability and control of flapping flight[6, 7].

Flapping flight aerodynamic forces originate due to a complex interaction. A full detailed representation of the flapping flight dynamic system evolves according to coupled differential equations that model the dynamics of the flapping animal and partial differential equations obtained from the incompressible Navier-Stokes equations. Motion models for ground based robotic systems that feature articulated wings can be constructed using techniques from robotics. The equation of robotics that govern the motion dynamics in a flapping experiment in a wind tunnel are written as

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}}(t) + \mathbf{C}\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\dot{\mathbf{q}}(t) + \frac{\partial V}{\mathbf{q}(t)} = \mathbf{Q}_a\left(\mathbf{q}(t), \dot{\mathbf{q}}(t), \mathbf{p}(t)\right) + \mathbf{B}\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)\tau(t),$$
$$\mathbf{q}(\mathbf{0}) = \mathbf{q}_0, \quad \dot{\mathbf{q}}(\mathbf{0}) = \dot{\mathbf{q}}_0.$$

where $\mathbf{q}$ contains the generalized coordinates of the robotic system, $\mathbf{M}(\mathbf{q})$ is the configuration dependent mass matrix, $\mathbf{C}\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)$ represents the centripetal and Coriolis terms, $V(\mathbf{q})$ is the potential energy term, $\mathbf{B}\left(\mathbf{q}(t), \dot{\mathbf{q}}(t)\right)$ is the control influence matrix, $\mathbf{Q}_a\left(\mathbf{q}(t), \dot{\mathbf{q}}(t), \mathbf{p}(t)\right)$ is the aerodynamic generalized force generated by the pressure field $\mathbf{p}(t)$, and $\tau(t)$ are the joint force and torques that derive the system. The Navier-Stokes equations with moving boundaries are

$$\partial\mathbf{v}(t)/\partial t + (\mathbf{v}(t)\cdot\nabla)\mathbf{v}(t) + \nabla\mathbf{p}(t) - \kappa\nabla\mathbf{v}(t) = \mathbf{f}(t) \tag{1.1}$$

$$\nabla\cdot\mathbf{v}(t) = 0, \tag{1.2}$$

$$\mathbf{v}(0) = \mathbf{v}_0, \tag{1.3}$$

$$\mathbf{v}(t)|_{\partial\Omega_\omega} = \mathbf{v}_\omega(\mathbf{q}(t), \dot{\mathbf{q}}(t)), \tag{1.4}$$

$$\Omega = \Omega(\mathbf{q}(t)). \tag{1.5}$$

where $\mathbf{v}(t)$ is the fluid velocity, $\mathbf{p}(t)$ is the fluid pressure, $f(t)$ is density of body forces per unit mass, $\mathbf{v}_\omega$ is the prescribed velocity on the boundry, and κ is the kinematic viscosity coefficient. The boundaries of the Navier-Stokes equations, $\Omega(\mathbf{q}(t))$, are a function of the robotics generalized coordinates. Also, the aerodynamic forces $\mathbf{Q}_a\left(\mathbf{q}(t), \dot{\mathbf{q}}(t), \mathbf{p}(t)\right)$ that act on flapping wings are function of the pressure distribution on the wings and kinematics of the animal. There are two issues that make these equations exceptionally difficult to model and employ in practice. The first challenge stems from the coupled nature of the above equations. Not only are the coupled differential equations very complicated to solve, but also there are few results in the literature that guarantee such a system is well-posed. The second issue arises from very high dimensionality of the unknowns in the model that makes it computationally infeasible to be used in a real time feedback control. In contrast, some averaged conventional models of aerodynamic forces are very simplified, and they do not always reflect some of the physical characteristics of these systems. For example, such conventional models may not distinguish between the aerodynamics in the up-stroke and down-stroke regimes. Such complexities can be relevant particularly when we try to model the articulated flapping motion of bats, one of flying animals that we study. Therefore, we seek to find a model that reflects these complexities, and yet is computationally feasible for control applications.

One approach to investigating the aerodynamic forces is to employ experimental methods that actuate flapping wings in a wind tunnel. The experiment is designed so that the flapping motion is close to that observed in wind tunnel. Aerodynamic forces are measured by installing pressure sensors on the wing and also by making force and moment measurements on the supports of the robotic hardware [8, 9]. Researchers at Brown University and the University of Illinois at Urbana-Champaign have carried out such experiments over the years. Another approach is to estimate the aerodynamic forces during each stroke with CFD simulations of the flapping motion for a certain wing model. This method has been pursued by several researchers such as [10, 11]. However, it is important to ensure the flapping motion is as close as possible to the actual flapping motion so that the measured aerodynamic forces match the physical phenomenon.

Therefore, we consider two different challenging subproblems in this research that arise from the general coupled system in equations (1.1-5). In the first challenge we propose a

history dependent formulation for the equations of motion of a ground based articulated flapping wing and we derive a novel adaptive control formulation to track observed flapping motions. In this approach, lift and drag aerodynamic forces are represented in terms of history dependent integral operators to model the unmeasurable and unknown aerodynamic forces on the robotic wing.

1.2.1 History Dependent Formulations of Aerodynamics

Insect wings are chosen in this dissertation as an iconic example to study the flapping flight aerodynamic forces using CFD simulations. Many results are not specific to this case, but they can be employed as a solid foundation for the study of more complicated cases such as the flapping flight of bats. In these studies flapping frequency and physical dimensions of the insect are matched to those of an actual insect to create dynamics that are close to the actual physics of the insect during flapping flight. We refer the reader to [11, 35] for further information about the CFD simulations and analysis that we have employed as the raw data in our identification of a history dependent model of flapping wing aerodynamics.

Given the history dependent behavior of the lift and drag functions from CFD simulations, we fit a low dimensional history dependent function by integration of history dependent kernels. In this approach, the lift and drag profiles are described in a linear in parameters standard form suitable for adaptive control strategies. The novelty of the formulation is that the regressors are history dependent functionals in this case. Convergence properties of low dimensional models to the infinite dimensional operator for a class of history dependent operators is investigated. A rate of convergence has also be derived as a function of the resolution level of the approximation space used in estimates.

In the next step, we prove existence and uniqueness of solutions for history dependent functional differential equations. A passivity based adaptive control strategies is derived for flapping wing prototypes that are used in a wind tunnel experiment. Lyapunov analysis guarantees stability of the closed-loop system. Furthermore, Barbalat's lemma is applied to show the tracking error and its derivative converge to zero.

1.2.2 Flight Submanifold Identification

The second primary thrust in this work has been the study of approximation and identification of low dimensional manifold that underly flapping flight.

Manifold Assumption

Motion analysis of typical bats, birds, or insects suggests that in order to construct biologically inspired robots, a large number of degrees of freedom may be needed for articulated, multibody systems composed of rigid bones interconnected by ideal joints. In addition, the modeling task is even more challenging if complexities such as joint compliance and wing

flexibility are considered. The most common approach to constructing a low dimensional model is to analyze and post-process the motion data of the system of interest to generate the reduced order model. One popular example of such an analysis method employs principal orthogonal decomposition (POD). In this method, the motion data captured by high speed cameras is analyzed after applying motion reconstruction algorithms to include the data from different cameras. In the POD analysis an eigenvector basis of the correlation matrix of the experimental data is computed. The calculated eigenvectors are called proper orthogonal modes. The eigenvalue associated with each POD mode corresponds to a measure of how much the specified mode contributes to the overall motion. These eigenvalues constitute a measure that assigns importance to the modes. For example, Riskin et. al. [13] uses POD and discovers that there are three groups of joints in the bat flight regime that move together as functional units. They show that more than 95 percent of the energy in a motion can be reconstructed by using around 13 modes. There are at least three reasons that are cited that support the use of POD analysis from a physical viewpoint: (1) multiple joints are activated by a common signal from the animal's nervous system; (2) in order to generate lift, regions of the wings need to move together; (3) different parts of the body which move together are physically connected. Therefore, it is possible to justify the idea that the motion of some parts of a body are a function of the motion of other parts. The above analysis strongly suggests that data samples captured by high speed video cameras for bat flapping motion live on a relatively low dimensional sets, or even manifolds, embedded within the higher dimensional configuration space that supports the full order dynamic model. One approach to effectively restrict the dynamics to evolve on a submanifold and reduce the complexity of the model is to develop a learning algorithm to estimate the motion submanifold from empirical data. Learning the motion submanifold involves using the high dimensional camera data to identify a much lower dimensional submanifold that can effectively describe the motion. Then the dynamics can be rigorously or approximately constrained to lie on the identified manifold. The approach proposed in this research develops a technique that steers the dynamics of the mechanical system to evolve near or on an identified submanifold using techniques from learning theory. In other words, instead of learning the submanifold in an explicit form, we construct approximate models such that their dynamics evolve in its close vicinity. In this method we combine methodologies used to derive discrete Lagrangian mechanics and data-driven methods of learning theory to obtain the complex dynamic models of biologically inspired motion.

Discrete Lagrangian Mechanics and Geometric Integration

Discrete Lagrangian formulations of mechanics and geometric integration methods have been an active area of research over the last few years. In order to improve the qualitative behavior of simulation and also to minimize the long time error, a geometric numerical integrator preserves geometric properties of the flow. In the differential equations for a system, the first integrals are functions that are constant along the solution curves. Many mechanical systems have total energy, angular momentum, or linear momentum that are first integrals of motion. These first integrals are not conserved using conventional numerical integrators, which can lead qualitative dynamics that are physically unreasonable.

Geometric integrators have been studied and developed to generate a discrete flow that respects these geometric properties. They also generate approximate solutions with excellent long-time global errors and numerical stability properties. The geometric integration methods of greatest interest to this research are the variational integrators of discrete Lagrangian or Hamiltonian mechanics. The literature on variational integration is too extensive to cover here in its entirety. We only review the topic briefly as needed to explain our approach. Numerical implementation of variational integrators in general is discussed in [16]. Reference [14] develops an integration algorithm for Hamiltonian systems that conserves the Hamiltonian or the total energy of the mechanical system, in addition to preserving the symmetry of the flow. Moreover, this method also exhibits invariant sets and preserves stability properties of the physical system. In [15] the author further considers mechanical systems subject to holonomic constraints and develops discrete integrators that preserve geometric properties such as total energy, linear and angular momentum, while enforcing the holonomic constraints.

Leyendecker in [17] further studies energy conserving integration methods for Hamiltonian systems subject to holonomic constraints. She reviews three classes of methods: the Lagrange multiplier method, the penalty method, and the augmented Lagrange method. Leyendecker shows that the Lagrange multiplier method, the penalty method (when the penalty weight goes to infinity or $\epsilon \to 0$) and, the Augmented Lagrangian method (for infinitely many iterations) are theoretically equivalent. It should be noted that all of these formulations require explicit, analytic representations of the holonomic constraints for implementation.

Thus, one method to steer the dynamics to satisfy the holonomic constraints, at least approximately, is to use the penalty method to"make motion costl" when it is distant from the set of states that satisfies the constraints. In the penalty method, a weighted norm of the constraint is included in the Hamiltonian, or equivalently the Lagrangian, of the system as a pseudo-potential function. This stiff potential term causes highly oscillatory solutions that approach the constrained solution of the dynamical system of interest provided that a collection of numerical hypotheses hold. The study of such systems starts with the work of [20] and has motivated by techniques in molecular dynamics in [21, 22, 23]. An extensive review of such systems can be found in [19] as well. The crucial issue in the numerical integration of systems having slow and fast subsystem arises from the required temporal resolution of the fast subsystem. Three classes of integration methods have been proposed in the literature for such systems: the impulse method, the mollified impulse method, and the projected impulse method.

The Impulse method is proposed by Grubmmuller et. al. [26] and Tuckerman et. al. [22]. In this method, the Hamiltonian is split into a fast and slow part $H(q,p) = H^{fast}(q,p) + U(q)$ and decomposes the exact flow of the Hamiltonian into

$$\phi_h^H \approx \phi_{h/2}^{slow} \circ \phi_h^{fast} \circ \phi_{h/2}^{fast}.$$

Each step of integration is composed substeps described as of a kick, an oscillation, and a kick. The mollified step method [24, 25] improves the error and stability characteristics of the impulse method by replacing the slow potential $U(q)$ by $U_{mollified}(q) = U(\alpha(q))$ where $\alpha(q)$ is an averaged or a projected value of q. Both impulse and mollified impulse methods

are symplectic and symmetric, and therefore they are suitable for our application. The projected impulse method preserves the convergence properties of the mollified method, but includes a simpler algorithm. Lubich et. al. [18] discuss how modifying the impulse method into the mollified or projected impulse methods improves the accuracy of integration.

In order to achieve a higher rate of convergence, we exploit the Galerkin variational integrator method that has emerged within the field of discrete mechanics developed by Hall and Leok [27] . In this method an $(n+1)$-dimensional function space $\mathbb{M}^n([0,h],Q) \subset C^2([0,h],Q)$ is used to approximate the action integral over $[0,h]$. The quadrature rule $h\sum_{j=1}^{m} w_j f(a_j h) \approx \int_0^h f(t)dt$ is used to compute the discrete action integral so that the discrete Lagrangian L_d satisfies

$$L_d(q_0,q_1;h) := \underset{\substack{q\in C^2([0,h];Q)\\ q(t_0)=q_0, q(t_1)=q_1}}{\text{ext}} \int_{t_0}^{t_1} L(q(\tau),\dot{q}(\tau))d\tau. \tag{1.6}$$

Contrary to the conventional methods where the accuracy is improved by increasing the temporal resolution of the method, i.e. by decreasing h, in this method the accuracy of the approximated action integral also depends on approximation the properties of the function space $\mathbb{M}^n$ from which it is constructed. That is, a better accuracy can be achievable by increasing the dimension of the function space n. Hall shows that Galerkin variational integrators, and in general spectral methods, accomplish geometric rates of convergence. That is, they converge faster than any polynomial order. It is this analysis on which our approach is based.

Empirical Potential Function

The essence of the proposed approach seeks to derive a discrete Lagrangian evolution equation that incorporates the result of a learning algorithm to construct a potential function from empirical observations. The proposed potential function has its extremum, approximately, on the subset or submanifold that supports the motion. We use a norm of this empirical potential function as the pseudo-constraint in the penalty method [17, 30] as we have briefly discussed earlier. Therefore, it is guaranteed that the discrete dynamics evolves on or near the motion submanifold that has been observed in experiments. This approach can be viewed as similar to the potential field method that are so popular in robotic navigation applications. In these methods, the obstacles are replaced with soft repelling fields to gently guide mobile robots away from obstacles as they navigate in a cluttered environment. In this method we exploit concepts from learning theory as it is cast in reproducing kernel Hilbert spaces (RKHS) [28, 29]. Since the learning operator proposed in this method is not computationally feasible, we use a sampling operator and a spectral filtering algorithm to achieve a computationally realizable algorithm. We also include the effect of the proposed algorithm in a novel one step error bound between the exact and the discrete Galerkin Lagrangian.

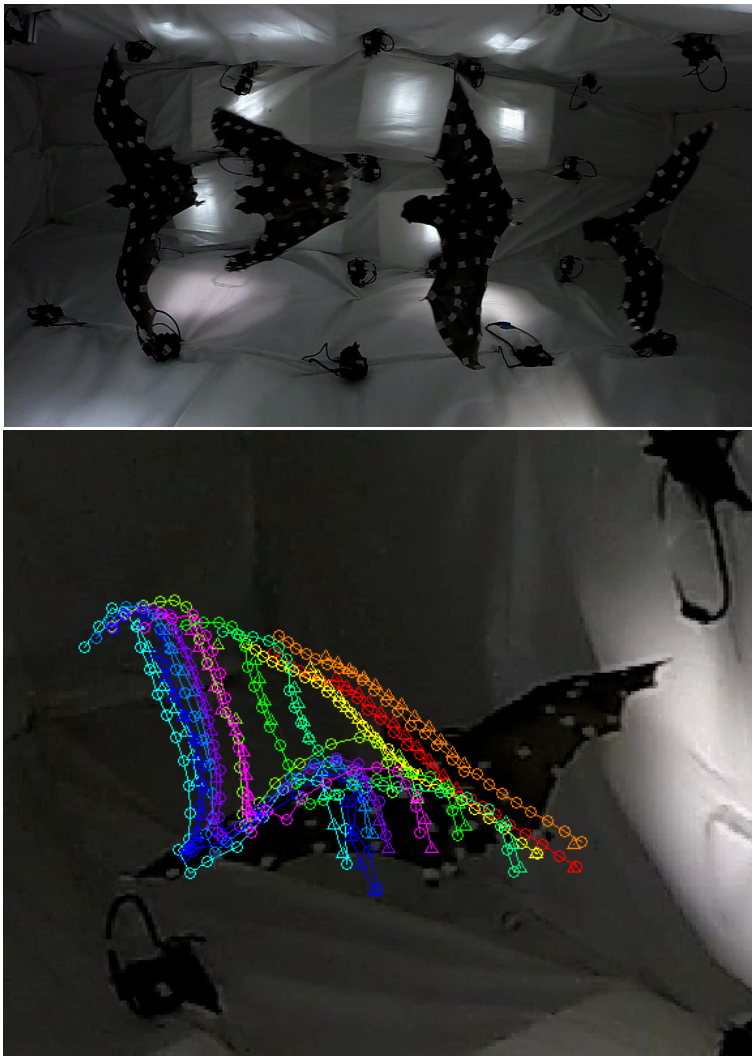


Figure 1.1: Bat Flight Experiments, SDU/VT international laboratory at Shandong University, courtesy of Matt Bender

Bibliography

[1] Haithem E. Taha, Muhammad R. Hajj ,Philip S. Beran, "State-spacerepresentationoftheunsteadyaerodynamicsofflappingflight", Aerospace Science and Technology 34(2014)111

[2] M.H.Dickinson,F.O.Lehmann,S.P.Sane,"Wingrotationandtheaerodynamicbasisofinsectflight" ,Science284(5422)(1999)19541960.

[3] Sanjay P. Sane, "The aerodynamics of insect flight", The Journal of Experimental Biology 206, 4191-4208, 2003

[4] S A Ansari, R Z bikowski, and K Knowles, "Non-linear unsteady aerodynamic model for insect-like flapping wings in the hover. Part 2: implementation", Proc. IMechE Vol. 220 Part G: J. Aerospace Engineering, 2006 and validation

[5] Pradeep Gopalakrishnan, "Unsteady Aerodynamic and Aeroelastic Analysis of Flapping Flight", PhD dissertation, Virginia Tech, 2008

[6] S. Chung and M. Dorothy, "Neurobiologically Inspired Control of Engineered Flapping Flight, *Journal of Guidance, Control, and Dynamics*, Vol. 33, No. 2, (2010)

[7] C. Orlowski and A. Girard, "Longitudinal Flight Dynamics of Flapping-Wing Micro Air Vehicles," *Journal of Guidance, Control, and Dynamics,* Vol. 35, No. 4, pp. 1115-1131 (2012)

[8] Derrick Yeo, Ella M. Atkins, and Wei Shyy. "Aerodynamic Sensing as Feedback for Ornithopter Flight Control",49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition 4 - 7 January 2011, Orlando, Florida

[9] Z. J. Wang, J. M. Birch, and M. H. Dickinson, "Unsteady forces and flows in low Reynolds number hovering ight: two-dimensional computations vs robotic wing experiments," J.Exp. Biology,", Vol. 207(3), pp. 449-460 (2004).

[10] K. Viswanath, K. Nagendra, J. Cotter, M. Frauenthal and D. K. Tafti, "Straight-line climbing flight aerodynamics of a fruit bat", Physics of Fluids, Vol 26, No. 2, 2014

[11] Jeffrey Feaster, Francine Battaglia and Javid Bayandor, "A Computational Study of Insect Wing Cross-Sectional Geometry on Flight Performance", ASME/JSME/KSME Joint Fluids Engineering Conference, Seoul, South Korea, July 2631, 2015

[12] S. Dadashi, J. Feaster, J. Bayandor, F. Battaglia , A. J. Kurdila, "Identification and adaptive control of history dependent unsteady aerodynamics for a flapping insect wing", Nonlinear Dynamics, 2016

[13] Riskin DK, Willis DJ, Iriarte-Daz J, Hedrick TL, Kostandov M, Chen J, Laidlaw DH, Breuer KS, Swartz SM, "Quantifying the complexity of bat wing kinematics", Journal of Theoretical Biology, 254(3):604-15, 2008 Oct 7

[14] O. Gonzalez ,"Time Integration and Discrete Hamiltonian Systems", J. Nonlinear Sci. Vol. 6: pp. 449467, 1996

[15] Oscar Gonzalez, "Mechanical systems subject to holonomic constraints: Differentialalgebraic formulations and conservative integration", Physica D 132 165174, 1999

[16] J.E. Marsden and M. West, "Discrete Mechanics and Variational Integrators", Acta Numerica, pp. 1158, 2001

[17] S. Leyendecker, P. Betsch, P. Steinmann, "Energy-conserving integration of constrained Hamiltonian systems a comparison of approaches", Computational Mechanics , Volume 33, Issue 3, pp 174-185, February 2004

[18] Christian Lubich and Daniel Weiss, "Numerical Integrators for Motion under a Strong Constraining Force", Multiscale Model. Simul., 12(4), 15921606, 2014

[19] E. Hairer, C. Lubich, and G. Wanner, "Geometric Numerical Integration. StructureP-reserving Algorithms for Ordinary Differential Equations", Springer, Berlin, 2006

[20] H. Rubin and P. Ungar, "Motion under a strong constraining force", Comm. Pure Appl. Math., 10, pp. 6587, 1957

[21] F. Takens, Motion under the influence of a strong constraining force, in Global theory of dynamical systems, Proc. Int. Conf., Evanston/Ill. 1979, vol. 819 of Lecture Notes in Mathematics, Springer, pp. 425445, 1980

[22] M. Tuckerman, B.J. Berne, and G.J. Martyna, "Reversible multiple time scale molecular dynamics", J. Chem. Phys., 97 , pp. 19902001, 1992

[23] F. Bornemann, "Homogenization in Time of Singularly Perturbed Mechanical Systems", vol. 1687 of Lecture Notes in Mathematics, Springer, 1998

[24] B. Garca-Archilla, J.M. Sanz-Serna, and R.D. Skeel, "Long-time-step methods for oscillatory differential equations", SIAM J. Sci. Comput., 20, pp. 930963, 1998

[25] J.A. Izaguirre, S. Reich, and R.D. Skeel, "Longer time steps for molecular dynamics", J. Chem. Phys., 110, pp. 98539864, 1999

[26] H. Grubmuller, H. Heller, A. Windemuth, and K. Schulten , "Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions", Mol. Sim., 6, pp. 121142, 1991

[27] James Hall, Melvin Leok, "Spectral Variational Integrators", Journal Numerische Mathematik archive Volume 130 Issue 4, pp. 681-740, August 2015

[28] Lorenzo Rosasco, Mikhail Belkin, Ernesto De Vito, "On Learning with Integral Operators", Journal of Machine Learning Research 11, 905-934, 2010

[29] Ernesto De Vito, Lorenzo Rosasco, Alessandro Toigo, "Learning Sets with Separating Kernels", Appl. Comput. Harmon. Anal. 37 185-217, 2014

[30] A. J. Kurdila t and Francis J. Narcowich , "Sufficient conditions for penalty formulation methods in analytical dynamics", Computational Mechanics 12, 81-96, 1993

Chapter 2

Identification and Adaptive Control of History Dependent Unsteady Aerodynamics for a Flapping Insect Wing

All of the computational Fluid Dynamics in this paper has been developed by J. Feaster, J. Bayandor, and F. Battaglia and I am endebted to their partnership on this material.

Identification and Adaptive Control of History Dependent Unsteady Aerodynamics for a Flapping Insect Wing [1]

S. Dadashi, J. Feaster, J. Bayandor, F. Battaglia, and Andrew Kurdila

Abstract

A history dependent formulation for the equations of motion of a ground-based robotic flapping wing, derived for an associated adaptive control strategy, was constructed to track observed flapping motions of insect flight. A general methodology was introduced in which lift and drag forces were represented in terms of history dependent integral operators to model and identify the unknown and unmeasurable aerodynamic loading on the flapping robot wing. First, computational fluid dynamics was used to predict the drag and lift forces for a flapping insect wing. The simulation data was then used to describe the adaptive control history. The resulting closed loop system constitutes an abstract Volterra integral equation whose state consists of the finite-dimensional vector of generalized coordinates for the robotic system and an infinite dimensional unknown function characterizing the kernel of the history dependent integral operator. Finite-dimensional approximations of the state equations were derived via quadrature formula and finite element methods. An adaptive control scheme based on passivity principles was derived for the approximate history dependent system. Lyapunov analysis guarantees stability of the closed loop system and that the tracking error and its derivative converge to zero. The novel control strategy introduced in this paper is noteworthy in that by introducing a history

[1]

- All of the computational Fluid Dynamics in this paper has been developed by J. Feaster, J. Bayandor, and F. Battaglia and I am endebted to their partnership on this material.

dependent adjoint operator in the state estimate equation, the analysis for convergence of the closed loop history dependent equations closely resembles the analysis used for conventional ordinary differential equation (ODE) systems.

2.1 Introduction and Motivation

It is well-known that there has been a renaissance in the study of flapping wing aerodynamics, flight dynamics, flight stability, and flight control [1, 2]. The last several decades have seen the continued study and refinement of fixed wing flight mechanics, computational aerodynamics, and flight robotics. As autonomous flight vehicles have become more prevalent, there has been a well documented trend to design new micro-air-vehicles (MAVs) that have wing spans that measure from a few millimeters to a few feet in length. Researchers and laboratories have looked to the interface of biology and engineering to find inspiration for alternative, groundbreaking approaches to flight that are pragmatic for this new class of small flight vehicles. Hundreds of publications on flapping flight have appeared in the literature, and it is important when reviewing the relevant research to understand the role various groups have had. Of course, it is impossible to be comprehensive here, and we will restrict consideration to some of the work most pertinent to this paper.

Reviews of experimental and computational studies on the aerodynamics of flapping flight, for a wide variety of flight conditions and including articulated flight, have appeared in [4, 3, 1, 5]. Currently, most of the computational research on flapping wing flight mechanics is limited to the case in which wings are structurally modeled as continuous membranes or plates, and in which the motion of the wing is a prescribed boundary condition on the flow. Representative results can be found in [34]. These studies focus on the formulation, implementation and computational performance of numerical methods for the simulation of 'single-body' wings as observed in many insects or idealizations of bird flight. These publications also seek to understand the flows that underlie various flapping wing strategies, and they have noted the importance of vortex formation and shedding in lift and drag generation during flapping flight.

On the other hand, research studies in the field of biomechanics or biology have invested considerable effort in studying the qualitative properties of articulated flapping flight, including the study of the biomechanics of articulated multibody flapping. These investigations seek to understand the relationship among biological structures such as bones, muscles and tendons, the joint morphology, and the overall flapping flight geometry. See [7], Sect. 1.6 for such a discussion of bat flight. In these early studies, the general qualitative nature of trajectories is discussed, for example, they describe how wing geometry relates to functions such as predation or maneuverability. As technologies, and in particular high resolution imaging sensors, and wind tunnel measurement systems have evolved

over the past few years, publications on flapping flight by biologists and engineers have become substantially more detailed. The work in [8, 9, 10, 11, 12], in some fashion, relies on the use of high speed, visual spectrum, charge-coupled device (CCD) cameras to track bat motion.

Researchers in the fields of flight stability, dynamics and control have also studied the mechanics of flapping flight in detail. The work in [13, 10, 16, 14, 15, 17, 18] is typical of this class. Excellent overviews of these fields can be found in [1, 2, 19]. For example, insects or birds are usually analyzed or modeled as rigid bodies, driven by a pair of wings that are each a single body. One to three degrees of freedom are used to represent the orientation of a single wing, while up to six degrees of freedom are used to represent the body orientation and translation. By virtue of symmetry arguments, if flight is restricted to longitudinal or lateral motions, existing studies of flapping flight stability or control typically consider from 2-6 degrees of freedom. See the summary in Tables 1.1, 1.2 and 1.3 of [2] for a succinct review. To the researchers studying flight stability and control, it is most important to guarantee performance and stability of a flapping flight control strategy. These works build on one another, and are similar in a few respects. They are all low dimensional: a few degrees of freedom are used to characterize the body, and from one to three degrees of freedom describe the motion of a wing. Symmetry arguments corresponding to assumptions of lateral or longitudinal flight motion result in system of equations varying from two to four degrees of freedom. One significant difference among the formulations is the choice of representation or model for the aerodynamic terms appearing in the governing equations. A comparison of several recent reduced order, unsteady, aerodynamic models for flapping flight is considered in [2].

Even though the substantial body of research has been carried out in recent years in the study of flapping flight kinematics, dynamics and control, it is surprising that the existing literature does not treat the problem of deriving models and associated controllers for the replication of flapping motions in ground experiments such as those that are carried out in wind tunnel tests. Because of the difficulty, cost, and complexity of studying flapping wing aerodynamics of animals in a wind tunnel environment, the construction of ground-based, articulating flapping experiments has been pursued recently [8]. This is the central goal of our paper: the derivation of adaptive control strategies for flapping wing prototypes that are used in wind tunnel experiments. The flapping wing experiment synthesize [8] provides excellent motivation for the control strategy developed herein.

Considering all of the above contributions to the state of the art in flapping flight, it is now acknowledged among researchers that the choice of the unsteady aerodynamic model is extremely important in the dynamics, stability, and control of flapping flight. A novel approach is introduced in this paper for the representation of unsteady flapping aerodynamics and the subsequent control of the flapping wing motion. The approach is general and is applied to flapping

wing ground robotic experiments that have the form of a kinematic chain. The key novelty of the approach is the introduction of a history dependent integral operator for the representation of unsteady flapping wing aerodynamics and the derivation of associated adaptive tracking controllers for the resulting history dependent equations. Computational fluid dynamics (CFD) is used to predict the drag and lift forces for the flapping flight of an insect wing, and the data is used to design and evaluate the adaptive control strategy. Preliminary results show that the class of proposed reduced order aerodynamic models provide excellent unsteady approximations of the lift and drag in flapping flight. Lyapunov analysis guarantees that the class of derived adaptive controllers, based on an extension of passivity arguments to the case of history dependent evolution laws, are stable and that the tracking error converges to zero. Section 2.2 provides a brief discussion of the class of unsteady aerodynamic phenomena that is modeled using CFD, including the methodology, numerical error, and results for drag and lift. Section 2.3 presents the approach to modeling unsteady aerodynamics using control theory. In section 3.2 is the formulation of the history dependent aerodynamic model. A formulation of an offline identification problem for a truth model representing aerodynamic contributions is presented in Sect. 2.3.2. Section 2.3.3 uses the integral operators introduced in Sect. 3.2 to derive the evolution equations for the open loop robotic system. The closed loop control strategy based on passivity principles is derived in Sect. 2.4, and numerical analysis to validate the model and control strategy are covered in Sect. 3.6.

2.2 Modeling Unsteady Aerodynamics using CFD

2.2.1 Flight Kinematics

The flapping flight aerodynamics of interest to this paper are described in [20] and [21]. The motion of a flapping wing is based on using data collected by Dudley and Ellington [20] using a high speed camera of the path traced by the wingtip of an insect in forward flight at a variety of flight velocities. A schematic of a wing is shown in Fig. 2.1 and identifies three rotational degrees of freedom (θ_1, θ_2 and θ_3) that describe the flight of the insect identified in [20]. Each of the angles can be expressed in terms of an amplitude θ_A and a sinusoidal function. The experimental data [20] was most sensitive to the θ_{2A} component, as this component varied in a nearly linear fashion with flight speed. The flapping frequency f, θ_{1A}, and θ_{3A} showed little correlation to the flight velocity and were considered constant at 150 Hz, 110° and 24°, respectively [20]. The resulting simplification is summarized in Table 2.1, where only θ_{2A} is varied with flight velocity.

In order to use the flight kinematics model in a two-dimensional CFD simulation, the three-dimensional angular amplitudes (θ_{1A}, θ_{2A}, and θ_{3A}) are used to compute

Table 2.1: The simplified kinematics of the insect wing, with $f = 150$ Hz, $\theta_{1A} = 110°$, $\theta_{3A} = 24°$ from [20]

V_∞ (m/s)	$\theta_{2A}(°)$
1	10
2	15
3	20
4	24
5	26

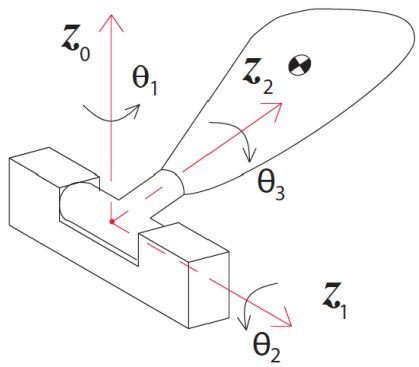


Figure 2.1: Schematic of the left wing of the robotic insect

the prescribed coordinates (x, y) of the aerodynamic center of the wing cross section. The wing angle of incidence θ_3 is selected to be the harmonic function

$$\theta_3 = \frac{\theta_{3A}}{2} \sin(\omega t + \varphi_3) \tag{2.1}$$

where f is the frequency, $\omega = 2\pi f$ is the angular frequency, t is the current time and φ is the phase shift of the wing rotation based on the high speed cinematography [20]. The remaining two rotations, θ_1 and θ_2, are applied as translational motions in the x and y directions in two-dimensional space. The translational positions are approximated as

$$x = \frac{r_0 \theta_{1A}}{2} \sin(\omega t + \varphi_1) \tag{2.2}$$

$$y = \frac{r_0 \theta_{2A}}{2} \sin(\omega t + \varphi_2) \tag{2.3}$$

where r_0 is the distance from the wing base to the wing midpoint. The time derivatives of Eqs. 2.1-2.2 are used to determine the velocity of the wing relative to the insect body. The angular velocity and translational velocities, respectively, are:

$$\dot{\theta}_3 = \frac{\omega \theta_{3A}}{2} \cos(\omega t + \varphi_3) \tag{2.4}$$

$$V_x = \frac{\omega r_0 \theta_{1A}}{2} \cos(\omega t + \varphi_1) \tag{2.5}$$

$$V_y = \frac{\omega r_0 \theta_{2A}}{2} \cos(\omega t + \varphi_2) \tag{2.6}$$

where $\dot{\theta}_3$, V_x and V_y are the time derivatives of Eqs. 2.1, 2.2 and 2.3, respectively. These two-dimensional kinematic equations represent the up-and-down strokes of the wing and is repeated for each beat of the wing in the analysis.

2.2.2 Numerical Methodology

The commercial software ANSYS Fluent (v. 15) [22] is employed as the CFD tool to solve the unsteady, incompressible formulation of the Navier-Stokes equations. The solutions provide the velocity and pressure fields for the two-dimensional model of the flapping wing. The equations are briefly presented, beginning with the incompressible continuity equation:

$$\nabla \cdot \overrightarrow{v} = 0 \tag{2.7}$$

where $\overrightarrow{v}$ is the velocity vector. Neglecting gravitational effects, the momentum equations are:

$$\rho \left[\frac{\partial \overrightarrow{v}}{\partial t} + \overrightarrow{v} \cdot \nabla \overrightarrow{v} \right] = -\nabla p + \nabla \cdot \overline{\overline{\tau}} \tag{2.8}$$

where t is time, ρ is the fluid density, p is the pressure and $\overline{\overline{\tau}}$ is the fluid stress tensor. Based on the flight velocity V_∞ (Table 2.1), the equations are solved assuming laminar flow.

The segregated pressure-based Navier-Stokes formulation is used to efficiently solve the incompressible flow [23]. The methodology uses the semi-implicit method for pressure linked equations (SIMPLE) algorithm to solve for the pressure-velocity coupling. The gradients are discretized using the least squares cell based (LSCB) method. The convective terms are discretized using a second-order upwind scheme, and the temporal terms are discretized using a first-order implicit method. The absolute convergence criteria are set as 10^{-7}. The time step is determined using a modified version of the Courant-Fredrichs-Levy (CFL) number:

$$CFL = \frac{\sqrt{V_x^2 + V_y^2}\Delta t}{\Delta X} \tag{2.9}$$

where ΔX is the smallest cell size, and for the problem under consideration, Δt = 0.025 ms.

2.2.3 Initial and Boundary Conditions

The full domain used in the CFD simulations is 4.8 cm $\times$ 4.4 cm in the x- and y-directions, respectively. The chord length of the wing is c = 4 mm, whereby the

Figure 2.2: Tetrahedral unstructured mesh surrounding the wing cross-section.

nondimensional domain is $12c \times 11c$. A relatively large domain is used to ensure that the boundaries do not adversely affect the flow around the wing as it moves up and down and translates left and right, based on the kinematic equations (Eqs. 2.4-2.6). Figure 2.2 shows a portion of the 2D mesh (approximately 1/6th of the full domain) surrounding the wing profile to provide a closer view of the wing details. The geometry of the insect wing is based on the experimental cross-section in [21]. The wing cross-section is analogous to the chord of an airfoil and captures the physical corrugation of the insect vascular system. The fluid domain is initially specified as quiescent air, and a velocity V_∞ (which represents the speed of the flying insect) is specified at the right boundary. The upper and lower boundaries of the fluid domain use a slip condition at the same velocity as the incoming flow, V_∞. The pressure at the left side of the fluid domain is specified as ambient pressure (0 gauge). An unstructured finite-volume mesh is generated around the cross-section and is combined with a dynamic remeshing procedure. The mesh is updated with each time step to allow for proper mesh motion, using a combination of Laplacian smoothing and remeshing functions to retain mesh density around the wing at all points in time during the flapping cycle.

2.2.4 Numerical Uncertainty

The grid convergence index (GCI) methodology is used to determine the numerical accuracy of the solution [24], the results of which are given in Table 2.2. The GCI is based on the Richardson extrapolation, where the discrete solutions are assumed to have a series representation of the discretization error. The GCI converts the Richardson extrapolation error estimate into an uncertainty by using absolute values to provide an uncertainty bound. The complete details are

Table 2.2: Discretization using the GCI method.

parameters	C_D	C_L
N_3, N_2, N_1	274k, 506k, 1M	274k, 506k, 1M
r_{21}	1.6313	1.6313
r_{32}	1.3590	1.3590
ϕ_1	-0.0115	0.0685
ϕ_2	-0.0116	0.0680
ϕ_3	-0.0117	0.0674
ap	3.232	3.415
GCI_{ext}^{21}	0.012%	0.50%
GCI_{ext}^{32}	0.807%	0.62%

not presented here, but the parameter r is the grid refinement factor, p is the order of accuracy and ϕ is any variable, for which we have selected the predicted coefficient of drag (C_D) and lift (C_L). The CFD results are used to calculate the spatial average drag and lift coefficients:

$$C_D = \frac{D}{\frac{1}{2}\rho V_{mag}^2 A} \tag{2.10}$$

$$C_L = \frac{L}{\frac{1}{2}\rho V_{mag}^2 A} \tag{2.11}$$

where D is the drag force, L is the lift force, V_{mag} is the maximum translational velocity magnitude and A is the wing area.

The grid resolution study was performed using a static mesh but with similar cell distributions as would be expected in the dynamic mesh simulation. Three grid resolutions (N) are used, with meshes composed of 274,000, 506,000, and 1,000,000 cells, where the coarse grid is denoted using the subscript 3 and the fine grid as 1 in Table 2.2; comparisons made between two grids use two subscripts. The grid resolution study examined lift and drag coefficients and a maximum error of 1% can be expected with an apparent third-order numerical accuracy (see ap in Tab. 2.2). The CFD results presented for the remainder of the paper will use the grid resolution of 506k cells.

2.2.5 Validation

The solution methodology is validated using the experimental data reported by Wang et al. [40] for flow over a flat plate using the motion of a hoverfly. In the experiment, a robotic arm with three rotational degrees of freedom was actuated and the kinematics used in the experiment are given in Tab. 2.3 for $\rho = 880$ kg/m^3 and a kinematic viscosity of 1.15×10^{-4} m^2/s. It should be noted that the only non-zero phase shift corresponds with Eq. 2.5, as reported by Wang et al. [40].

Table 2.3: Validation kinematics from Wang et al. [40] for a flat plate

Parameter	Value
f	0.25 Hz
r_0	16.25 cm
θ_{1a}	23.5°
θ_{2a}	0°
θ_{3a}	45°
φ_1	$\pi/2$
φ_2	0
φ_3	0
c	2.385 cm

All boundaries were modeled as no-slip walls with a moving rectangular wing cross-section located at the center.

A two-dimensional force sensor was attached to the base of the wing to capture the transient C_D and C_L experienced by the wing [40]. Figure 2.3 presents the coefficients for drag and lift versus nondimensional time $\tau = ft$ for one flap cycle. The transient results shown in Fig. 2.3 are found to be extremely similar between the experiment and simulation with the largest differences manifesting at the stroke transition (e.g., $\tau = 2.05$ and 2.55). These differences are attributed to three-dimensional effects in the experiment that cannot be captured in a two-dimensional simulation. Quantitatively the predicted time-averaged C_D and C_L values of 1.33 and 0.82 agree well with the experimental values given in Wang et al. [40] of 1.34 and 0.86, respectively. The validation provides additional confidence with the 2D computational modeling to further pursue the wing profile analysis presented next.

2.2.6 Flow Field Characteristics

The study begins with a comparison of a fixed flight speed at $V_\infty = 3$ m/s using the 2D wing model to demonstrate the time-dependent changes in the flow field around the wing. Four instants in time are shown in Fig. 2.4 using pressure contours to identify regions of high and low pressure; the wing cross-section is the thin black corrugated profile. The pressure can be directly related to the lift and drag forces that the wing experiences (positive drag denotes thrust). Beginning with the top image, the wing is positioned with a zero angle of incidence (located towards the bottom left corner). As the wing begins the flapping cycle (moving towards the upper right corner), a low pressure region forms below the leading edge, which corresponds to negative lift and minimal drag during the upstroke. Another low pressure region is noticeable above the wing and corresponds to a leading edge vortex that formed during the previous cycle. The next image is approximately 70 time steps later, and the wing is at a $24°$ angle of incidence

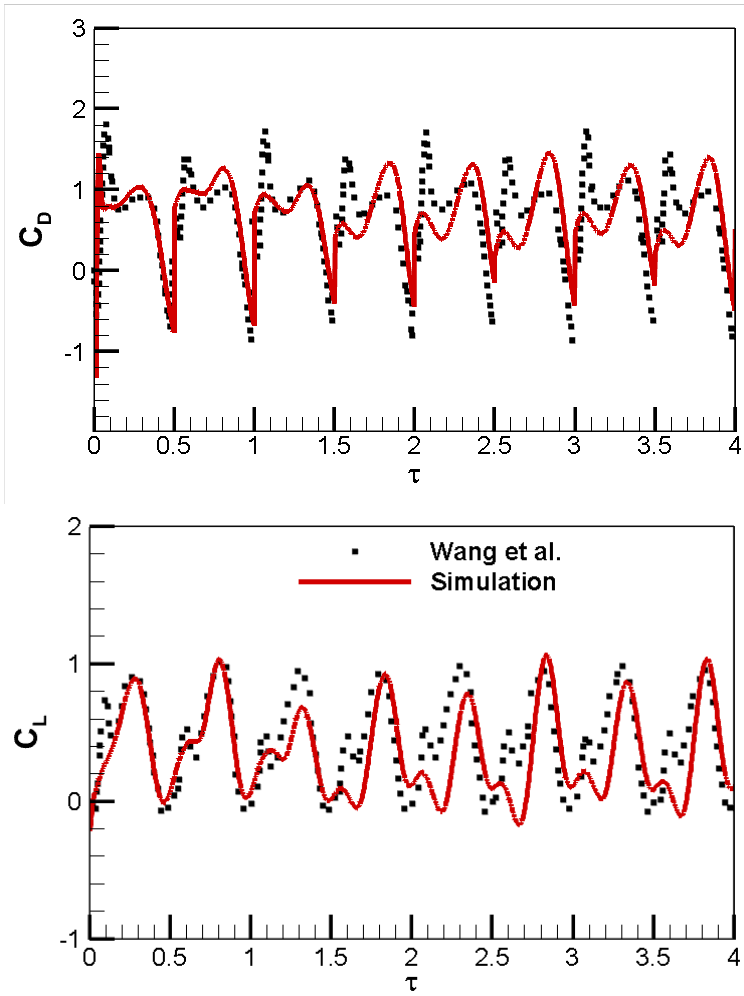
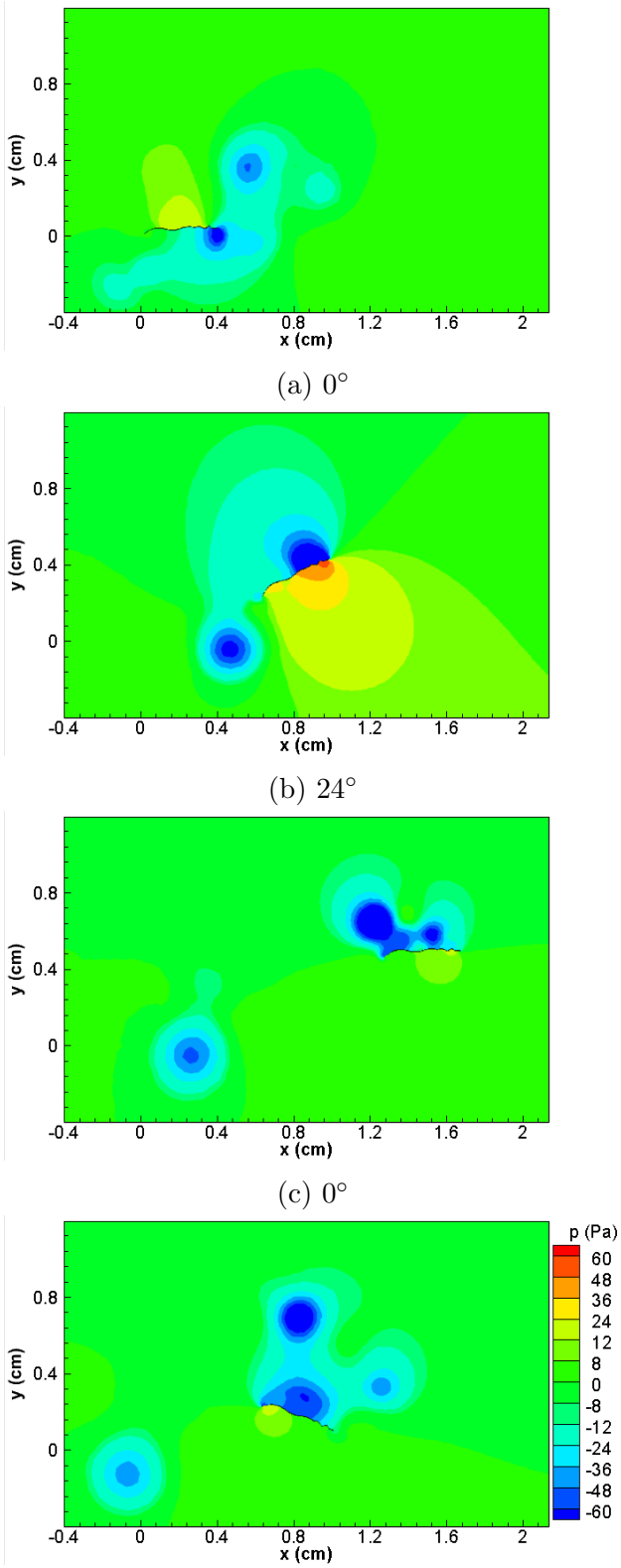


Figure 2.3: C_D and C_L versus τ for the simulation and experimental data [40].

as the wing ascends during the upstroke. At this instant, a low pressure regime forms on the upper surface as the wing captures the vortex from the previous cycle, with a high pressure regime forming below that corresponds to significant lift and drag. The last two images show the wing descending during the downstroke from $0°$ to $-24°$ angle of incidence, and a low pressure region forms above the wing (positive lift and thrust), with remnants of an earlier vortex in the downstream flow.

To further emphasize the transient nature of the flow, the time-dependent values for C_D and C_L are shown in Figs. 2.5 and 2.6 for 4 flapping cycles at flight velocities of 1, 3 and 5 m/s. The upstroke is the loop tracing θ_3 from $0°$ to $24°$ to $0°$ and the downstroke is from $0°$ to $-24°$ to $0°$. At the lowest velocity, the time-dependent nature is subtle, and the 4 flapping cycles repeat the same pattern without much deviation. At $V_\infty = 3$ and 5 m/s, the flow is more susceptible to instabilities, which are numerically induced due to the dynamic remeshing procedure but are similar to flow instabilities randomly caused in nature. Thus,



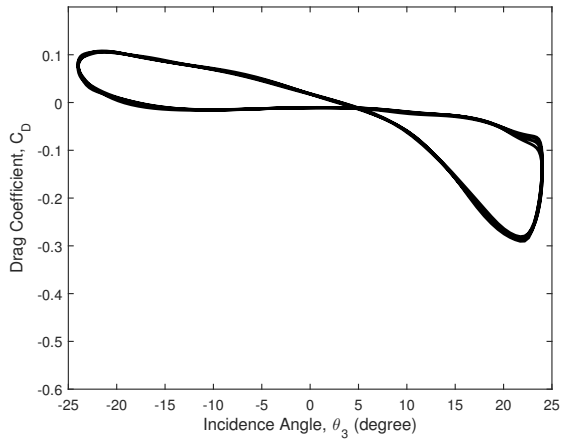
(a) 0°

(b) 24°

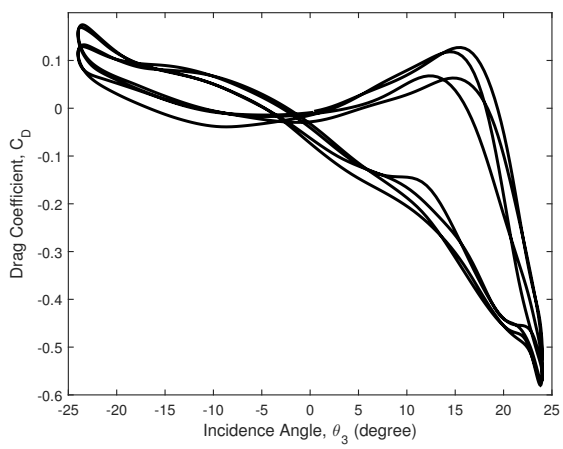
(c) 0°

(d) −24°

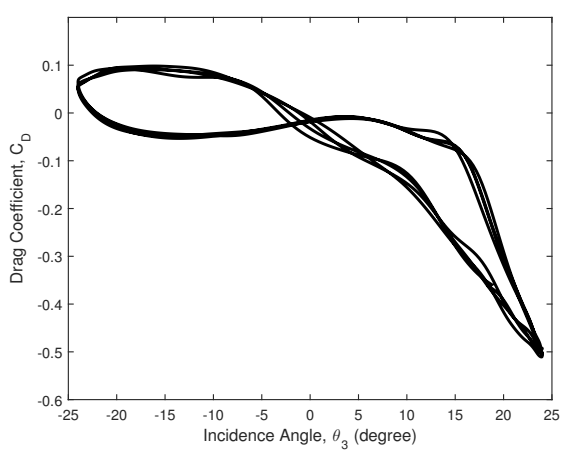
Figure 2.4: Predictions of pressure using CFD for the wing profile at 3 m/s for various angles through the upstroke and downstroke.



(a) 1 m/s

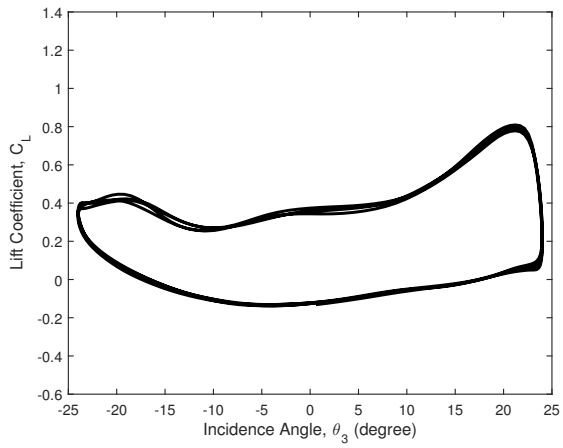


(b) 3 m/s

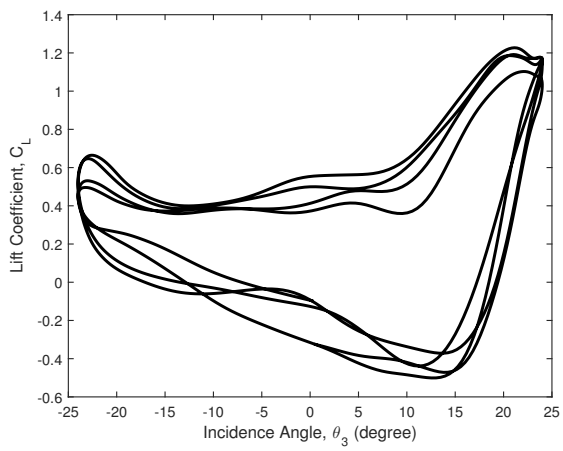


(c) 5 m/s

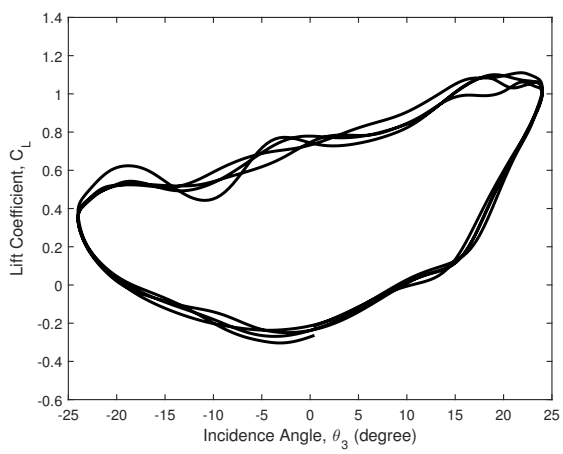
Figure 2.5: Predictions of the drag coefficient using CFD for the wing profile at various V_∞ for 4 flapping cycles.



(a) 1 m/s



(b) 3 m/s



(c) 5 m/s

Figure 2.6: Predictions of the lift coefficient using CFD for the wing profile at V_∞ for 4 flapping cycles.

the drag and lift coefficients change with each flapping cycle, although the overall patterns overlap. These figures demonstrate that the pressure distribution along the wing is affected by vortex eddies formed during the upstroke and downstroke of the flapping motion, which is inherently an unsteady phenomenon.

2.3 Modeling Unsteady Aerodynamics using Controls

2.3.1 History Dependent Aerodynamic Model

The last section has made clear the importance of unsteady and nonlinear mechanisms in the generation of drag and lift during flapping insect flight. Unfortunately, it is not feasible to employ the full three-dimensional models using the Navier-Stokes equations in a feedback control regime. This section introduces a general methodology to obtain reduced order models suitable for feedback synthesis for systems subject to unsteady, history dependent aerodynamics. The formulation casts aerodynamic contributions as a history dependent integral operator. General considerations regarding this class of integral operators can be found in [25, 15, 27]. The particular form of the integral operator used in this formulation is based on [28], a generalized play operator.

We assume that the unsteady value of the drag and lift forces as a function of time are defined as:

$$D(t) = \frac{1}{2}\rho\|V_{0,a}(t) - V_{0,w}(t)\|^2 AC_D(t)$$

$$L(t) = \frac{1}{2}\rho\|V_{0,a}(t) - V_{0,w}(t)\|^2 AC_L(t).$$

where $V_{0,a}(t)$ is the velocity of the aerodynamic center in the ground frame, $V_{0,w}(t)$ is the wind velocity in the inertial frame, and $C_D(t)$ and $C_L(t)$ are the history time-dependent drag and lift coefficients. Our aim is to reconstruct the history dependent lift and drag functions by a sum of weighted history dependent kernel functions. There are a plethora of approaches for generating history dependent input-output operators. These techniques include the definition of history dependent integral operators, methods based on ordinary and partial differential equations, and approaches based on differential inclusions. The interested reader may consult [25] and [15] for detailed discussion of the theory. In this paper we employ a relatively simple representation of the history dependent response. We define the lift and drag coefficients in terms of history dependent integral operators that have the form bellow:

$$C_D(t) := (K_D\alpha)(t) \circ \mu_D := \int_S (k_D(s,\alpha))(t)\mu_D(s)\mathrm{d}s \tag{2.12}$$

$$C_L(t) := (K_L\alpha)(t) \circ \mu_L := \int_S (k_L(s,\alpha))(t)\mu_L(s)\mathrm{d}s \tag{2.13}$$

In these equations $(k_D(s,\alpha))(t)$ and $(k_L(s,\alpha))(t)$ are the output at time t of elementary history dependent kernels defined on the half plane $S : \{(s_1, s_2) \in \mathbb{R}^2; s_1 \leq s_2\}$ and μ_D, μ_L are unknown functions that weight the contribution of the kernels to the overall response.

Each elementary kernel function is defined in terms of two functions $\bar{r}_1(s) := \bar{r}(s - s_1)$ and $\bar{r}_2(s) := \bar{r}(s - s_2)$ representing the output for increasing input and decreasing input, respectively, that are shifts of a single ridge function $\bar{r}(s)$ for $s := (s_1, s_2) \in S$. As shown in Fig. 2.7, we require that each ridge function is uniformly bounded. An elementary history dependent kernel constructed from the two functions $\bar{r}_1$ and $\bar{r}_2$ is depicted in Fig. 2.7. The output of the elementary kernel is defined by specifying its action on piecewise monotone input functions and subsequently extending the kernel to all continuous functions by a density argument. Suppose that the angle of attack α is a piecewise monotone function on $[0, T]$, and let $\tau \in [0, t)$ be the most recent time prior to t at which that input changes from a decreasing to an increasing, or from an increasing to a decreasing, function. Define the output for such a time τ as $k_\tau := k(s,\alpha)(\tau)$ for $\tau \leq t \leq T$. The output of the history dependent kernel at time $t > \tau$ is then defined as:

$$k(s,\alpha)(t) = \begin{cases} \max(\bar{r}_1(\alpha(t)), k_\tau) & \text{if } \alpha(t) \text{ is decreasing} \\ \min(\bar{r}_2(\alpha(t)), k_\tau) & \text{if } \alpha(t) \text{ is increasing} \end{cases} \tag{2.14}$$

where $s := (s_1, s_2)$.

Therefore, $k_L(s,\alpha)(t)$ and $k_D(s,\alpha)(t)$ are the kernels associated with the lift and drag integral operators that follow the definition described in Eq. 2.14. If the ridge function is sufficiently smooth, it is possible to show that $k_D, k_L : S \times C[0,T] \mapsto C[0,T]$ and $K_D, K_L : C[0,T] \mapsto C([0,T], \mathcal{L}(H,\mathbb{R}))$. Each of the these operators is causal since the output of each operator at time t only depends on present and past inputs. By virtue of these properties

$$k_D(s,\alpha) \in C[0,T] \quad \textbf{and} \quad k_D(s,\alpha)(t) \in \mathbb{R} \quad \textbf{for} \quad \forall s \in S$$

$$k_L(s,\alpha) \in C[0,T] \quad \textbf{and} \quad k_L(s,\alpha)(t) \in \mathbb{R} \quad \textbf{for} \quad \forall s \in S$$

$$K_D\alpha \in C([0,T], \mathcal{L}(H,\mathbb{R})) \quad \textbf{and} \quad (K_D\alpha)(t) \in \mathcal{L}(H,\mathbb{R})$$

$$K_L\alpha \in C([0,T], \mathcal{L}(H,\mathbb{R})) \quad \textbf{and} \quad (K_L\alpha)(t) \in \mathcal{L}(H,\mathbb{R})$$

where $H := L^2(S)$ is the space of Lebesgue square integrable functions. The values of the lift and drag coefficients at time t are consequently written as

$$\begin{Bmatrix} C_D(t) \\ C_L(t) \end{Bmatrix} := (\mathcal{K}\alpha)(t) \circ \mu = \begin{pmatrix} (K_D\alpha)(t) & 0 \\ 0 & (K_L\alpha)(t) \end{pmatrix} \circ \begin{Bmatrix} \mu_D \\ \mu_L \end{Bmatrix}$$

where $\mathcal{K}\alpha \in C([0,T], \mathcal{L}(H \times H, \mathbb{R}^2))$. In the discussion that follows, the symbol k denotes either the elementary drag or lift kernels k_D or k_L. Similarly, the symbol $K\alpha$ denotes either the weighted integral operators $K_D\alpha$ or $K_L\alpha$, and C represents either the drag coefficient C_D or the lift coefficient C_L. In this definition the integral operator K_D or K_L acts on the history of the angle of attack up to time t in determining its output at time t.

2.3.2 History Dependent Operator Identification

The kernel $k(\cdot,\cdot)$ is fixed but the weight μ_D or μ_L in Eqs. 2.12 or 2.13 that characterize the lift and drag operators are not known *a priori*. The adaptive control strategy to be derived in Sect. 2.4 relies on online estimation of the weight functions, and consequently provides important information regarding the aerodynamic lift and drag data. However, it is also important to be able to generate estimates of the weight functions by an offline identification procedure. An offline identification problem based on CFD data (Sect. 2.2.6) can provide a truth model to gauge the performance, validity and range of confidence of the adaptive control method. Additionally, offline identification from CFD data can be used to build an initial state for the weight function estimate in the adaptive control strategy, thereby improving performance and rate of convergence.

Consider the family of flow simulations for the flapping wing (Figs. 2.5-2.6) and subsequently collect from the simulations the angle of attack and aerodynamic coefficients $\{\overline{\alpha}(t_k), \overline{C}(t_k)\}_{k=0}^{m}$, e.g., either C_D or C_L. The identification problem seeks to find the function $\mu \in H := L^2(S)$, again corresponding to either μ_D or μ_L, that in some sense best fits the history dependent model

$$C(t) := (K\alpha)(t) \circ \mu := \int_S (k(s,\alpha))(t)\mu(s)\mathrm{d}s$$

To obtain a problem that is computationally tractable, we create a finite element discretization of the unknown weight function $\mu \in L^2(S)$ in the form

$$\mu(s) = \sum_{n=1}^{n_\mu} \mu_n \psi_n(s)$$

where $\psi_n(s)$ are basis functions for $n = 1 \ldots n_\mu$. The aerodynamic coefficient is now given in terms of the n_μ unknown coefficients $\{\mu_n\}_{n=1}^{n_\mu}$:

$$\begin{aligned} C(t) :&= \sum_{n=1}^{n_\mu} \int_S k(s,\alpha))(t)\psi_n(s)\mathrm{d}s \ \mu_n \\ &= \sum_{e=1}^{n_e} \sum_{n=1}^{n_\mu} \int_{S_e} k(s,\alpha))(t)\psi_n(s)\mathrm{d}s \ \mu_n \end{aligned}$$

and the finite element discretization induces the partition $S = \bigcup_{n_e} S_e$, where n_e is the number of elements. Mapping each triangular element to the reference triangle $\overline{S}_e$, the integral can be calculated using quadrature formula:

$$\begin{aligned} C(t) :&= \sum_{e=1}^{n_e} \sum_{l=1}^{n_p} \int_{S_e} k(s,\alpha))(t)\psi_{n(l)}(s)\mathrm{d}s \ \mu_{n(l)} \\ &= \sum_{e=1}^{n_e} \sum_{l=1}^{n_p} \sum_{j=1}^{n_q} \int_{\overline{S}_e} k_L(s(\xi_j),\alpha))(t)\psi_{n(l)}(s(\xi_j)) \left|\frac{\partial s}{\partial \xi}\right| \mathrm{d}\xi \ \mu_{n(l)} \end{aligned}$$

In this equation, n_p is the number of shape functions per element. We can approximate the aerodynamic coefficient by summing over n_q quadrature points $\{\xi_j\}_{j=1}^{n_q}$.

Given the time series from CFD simulations, $\{(\bar{\alpha}(t_k), \bar{C}(t_k))\}_{k=0}^{m}$, it is possible to express an approximate error between the CFD and history dependent model. The error of estimation at time t_k is:

$$E_k = \bar{C}(t_k) - \sum_{n=1}^{n_\mu} \int_S k(s, \overline{\alpha})(t_k)\psi_n(s)\mathrm{d}s\mu_n$$

A vector of error at all time steps can be assembled as follows:

$$\mathbf{E} = \bar{\mathbf{C}} - \mathbf{A}(\overline{\alpha})\mu$$

where

$$\bar{\mathbf{C}} = \begin{bmatrix} \bar{C}(t_0) & \bar{C}(t_1) & \cdots & \bar{C}(t_k) \end{bmatrix}^T$$

$$\mu = \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_{n_\mu} \end{bmatrix}^T$$

$$\mathbf{A}(\overline{\alpha}) = \begin{bmatrix} \int_S k(s, \overline{\alpha})(t_0)\psi_1(s)\mathrm{d}s & \cdots & \int_S k(s, \overline{\alpha})(t_0)\psi_{n_\mu}(s)\mathrm{d}s \\ \int_S k(s, \overline{\alpha})(t_1)\psi_1(s)\mathrm{d}s) & \cdots & \int_S k(s, \overline{\alpha})(t_1)\psi_{n_\mu}(s)\mathrm{d}s \\ \vdots & \ddots & \vdots \\ \int_S k(s, \overline{\alpha})(t_k)\psi_1(s)\mathrm{d}s) & \cdots & \int_S k(s, \overline{\alpha})(t_k)\psi_{n_\mu}(s)\mathrm{d}s \end{bmatrix}$$

Thus, the initial optimization problem over the infinite dimensional space $L^2(S)$ is approximated by a finite dimensional optimization problem where we seek the set of coefficients $\{\mu_n\}_{n=1}^{n_\mu}$ that minimize the error functional:

$$\begin{aligned} J(\mu) &= \frac{1}{2}\mathbf{E}^T\mathbf{W}\mathbf{E} = \frac{1}{2}(\bar{\mathbf{C}} - \mathbf{A}(\bar{\alpha})\mu)^T\mathbf{W}\frac{1}{2}(\bar{\mathbf{C}} - \mathbf{A}(\bar{\alpha})\mu) \\ &= \frac{1}{2}\mu^{\mathbf{T}}\mathbf{A}(\bar{\alpha})^{\mathbf{T}}\mathbf{W}\mathbf{A}(\bar{\alpha})\mu - \overline{\mathbf{C}}^{\mathbf{T}}\mathbf{W}\mathbf{A}(\bar{\alpha})\mu + \frac{\mathbf{1}}{\mathbf{2}}\overline{\mathbf{C}}^{\mathbf{T}}\mathbf{W}\overline{\mathbf{C}} \end{aligned}$$

where W is a symmetric positive definite real matrix selected by the analyst. Numerical solutions of this identification problem are described in Sect. 3.6. This offline identification procedure has been used in the adaptive control simulations as a truth model representing the Navier-Stokes flow.

2.3.3 Robotic Equations of Motion

The overall goal of this paper is to derive a model and an associated control method that drives a ground-based, flapping wing robot such that it tracks observed insect flapping motions. Such models are exploited to study aerodynamic forces, power inputs and outputs, and parameters such as stroke plane, wing shape, wing beat frequency, wing beat amplitude, etc. An example of such a

design can be found in [8] for studying bat flight. The equations of motion can be derived using Lagrange's equations, which for a fully actuated robotic system powered by three servomotors, forming a kinematic chain take the form:

$$\begin{aligned}\mathbf{M}(\mathbf{q}(\mathbf{t}))\ddot{\mathbf{q}}(\mathbf{t})+\mathbf{C}(\mathbf{q}(\mathbf{t}),\dot{\mathbf{q}}(\mathbf{t}))\dot{\mathbf{q}}(\mathbf{t})+\frac{\partial \mathbf{V}}{\partial \mathbf{q}}(\mathbf{q}(\mathbf{t}))\\ =\mathbf{Q}_a(t)+\mathbf{B}(\mathbf{q}(\mathbf{t}),\dot{\mathbf{q}}(\mathbf{t}))\tau(\mathbf{t})\end{aligned} \tag{2.15}$$

where $\mathbf{q}(t) \in \mathbb{R}^N$ is the vector of generalized coordinates, $\mathbf{M}(\mathbf{q}(\mathbf{t})) \in \mathbb{R}^{N\times N}$ is the generalized inertia matrix, $\mathbf{C}(\mathbf{q}(\mathbf{t}),\dot{\mathbf{q}}(\mathbf{t})) \in \mathbb{R}^{N\times N}$ is the generalized damping matrix, V is the potential energy, $\mathbf{B}(\mathbf{q}(\mathbf{t}),\dot{\mathbf{q}}(\mathbf{t})) \in \mathbb{R}^{N\times N}$ is the nonlinear control influence matrix that is assumed to be uniformly invertible, $\tau(t) \in \mathbb{R}^N$ is the vector of actuation forces and torques, and $\mathbf{Q}_a(t) \in \mathbb{R}^N$ is the vector of generalized forces due to aerodynamic loads acting on the robot.It should be noted that in this problem $\tau(t)$ represents the torque inputs that derive the joints. The robotic system is fully actuated that is the number of control input equals the number of generalized coordinates. Note that the gravity potential term has been neglected since the mass of the insect wing is very small.

2.3.4 Unsteady Aerodynamics Model

The aerodynamic forces acting on the wing are defined in terms of the wind frame W whose orientation relative to the body fixed B frame is defined by the change of basis

$$\begin{Bmatrix}\mathbf{w}_1\\ \mathbf{w}_2\\ \mathbf{w}_3\end{Bmatrix}=\underbrace{\begin{bmatrix}\cos\beta & \sin\beta & 0\\ -\sin\beta & \cos\beta & 0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\cos\alpha & 0 & \sin\alpha\\ 0 & 1 & 0\\ -\sin\alpha & 0 & \cos\alpha\end{bmatrix}}_{\mathbf{R}_B^W(\alpha,\beta)}\begin{Bmatrix}\mathbf{b}_1\\ \mathbf{b}_2\\ \mathbf{b}_3\end{Bmatrix}$$

where $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ is the basis for W, $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$ is the basis for B, α is the angle of attack, and β is the sideslip angle. With respect to wind axes, the aerodynamic force acting at the aerodynamic center a of the wing is $\mathbf{F}_a = -\mathbf{D}\mathbf{w_1} - \mathbf{L}\mathbf{w_3}$, and the components of this force relative to the wind axes are written as

$$\mathbf{F}_a^W = \bar{\mathbf{S}}\begin{Bmatrix}\mathbf{D}\\ \mathbf{L}\end{Bmatrix}$$

where

$$\bar{\mathbf{S}} = \begin{bmatrix}-1 & 0\\ 0 & 0\\ 0 & -1\end{bmatrix}.$$

The components relative to the ground frame of the aerodynamic forces can therefore be written as

$$\mathbf{F}_a^0(t) = \frac{1}{2}\rho||\mathbf{V}_{0,a}-\mathbf{V}_{0,w}||^2 A_w \mathbf{R}_k^0(\mathbf{q})\mathbf{R}_B^k\mathbf{R}_W^B(\alpha,\beta)\bar{\mathbf{S}}(\mathcal{K}\alpha\circ\mu)(t) \tag{2.16}$$

where link k of the kinematic chain that makes up the flapping wing robot is the wing. In Eq. 2.16, $\mathbf{R}_k^0$ is the rotation matrix mapping the frame fixed in link k into the ground frame and $\mathbf{R}_B^k$ is the rotation mapping from the body fixed B frame to the frame fixed in link k. Since the virtual displacement of the aerodynamic center is written as $\delta\mathbf{r}_{0,a}^0 = \mathbf{J}_{v_a}^0(\mathbf{q})\delta\mathbf{q}$ where $\mathbf{J}_{v_a}^0(\mathbf{q})$ is the Jacobian matrix of the robot [29, 30], generalized forces due to the aerodynamic loads are written as

$$\mathbf{Q_a}(t) = \Psi_a(t)(\mathcal{K}\alpha)(t) \circ \mu$$

where $\mu = [\mu_D \quad \mu_L]^T$, $(\mathcal{K}\alpha(t)) \circ \mu = [K_D\alpha(t) \circ \mu_D \quad K_L\alpha(t) \circ \mu_L]^T$, and

$$\Psi_a = \frac{1}{2}\rho||\mathbf{v}_{0,a}||^2 A_w \mathbf{J}_{v_a}^{0\,T}\mathbf{R}_k^0(\theta_1, ..., \theta_k)\mathbf{R}_B^k\mathbf{R}_W^B(\alpha, \beta)\bar{\mathbf{S}}$$

The above modeling strategy is discussed in [38]. We refer the reader to [39] for detailed derivation and implementation of such models.

2.4 Adaptive Control

The equations in Eq. 2.15 are not a system of ordinary differential equations since they depend on the history of the motion. They are examples of abstract Volterra equations as discussed in [33, 32, 31]. A general analysis of the existence and uniqueness of solutions of classes of abstract Volterra equations can be found in [35, 34]. Specific discussion of the existence and uniqueness of solutions for special cases of these equations can be found in several places, such as in [36] or more recently in Sect. 2.5 of [37]. Similar to the case for ordinary differential equations, the local existence of the solutions can be proven assuming a local Lipschitz condition. In contrast to the study of ordinary differential equations, however, the appropriate Lipschitz condition is expressed in terms of the norm on $C[0, T]$. We employ the following special case of Theorem I in [36] which establishes the existence of a local solution to a general class of history dependent or functional differential equations.

Theorem 1 ***Let*** B ***be a Banach space,*** $F : [0, T] \times C([0, t]; B) \to B$ ***and*** $-\infty < \alpha \leq \beta < \gamma < T < \infty$***. Suppose that***

1. $t \mapsto F(t, \eta)$ ***is continuous for*** $t \in [\beta, T)$ ***and*** $\eta \in C([\alpha, T); B)$***;***

2. ***for arbitrary*** γ ***and bounded set*** $B_0 \subseteq B$ ***there is a*** $L(\gamma, B_0)$ ***such that***

$$||F(t, \eta) - F(t, \phi)||_B \leq L(\gamma, B_0)||\eta - \phi||_{C[\alpha,t]}$$

for all $(t, \eta), (t, \phi) \in [\beta, \gamma] \times C([\alpha, \gamma]; B_0)$***.***

Then for every pair of initial data, there is an $h > 0$ ***such that unique solution of the evolution equation***

$$\dot{X}(t) = F(t, X); \quad X(t_0) = X_0.$$

is defined for $t \in [t_0, t_0 + h)$***.***

The application of this specific theorem to our problem is lengthy, so we outline the proof of local existence for our open loop equations. The local existence theorem for the closed loop system discussed later in this paper is similar, but much longer, so we omit it because of space limitations.

The state space representation of robotic equations of our open loop motion, can be written as

$$\dot{\mathbf{X}} = \mathbf{F_0}(\mathbf{X}(t)) + (\mathbf{F_1}\mathbf{X})(t) + \mathbf{G}(\mathbf{X}(\mathbf{t}))\tau(t)$$

$$(\mathbf{F_1}\mathbf{X})(t) = \begin{Bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}(\mathbf{X}(t))\Psi_a(t)(\mathcal{K}\alpha(\mathbf{X}))(t) \circ \mu \end{Bmatrix}$$

choosing $\mathbf{X}(t) = \{\mathbf{q}^T(t) \quad \dot{\mathbf{q}}^T(t)\}^T$**. Clearly our governing equations correspond to the special case in Theorem 1 where** $F(t, X) = \mathbf{F_0}(\mathbf{X}(t)) + (\mathbf{F_1}\mathbf{X})(t) + \mathbf{G}(\mathbf{X}(\mathbf{t}))\tau(t)$**. Since the sum and product of two locally Lipschitz functions is locally Lipschitz, we focus on the history dependent term in the equations of motion and show that it satisfies the local Lipschitz condition.**

Theorem 2 ***The function*** $C(t) = (\mathcal{K}\alpha(\mathbf{X}))(t) \circ \mu$ ***is locally Lipchitz in*** $||.||_{C[0,T]}$

$$||[\mathcal{K}\alpha(\mathbf{X})](t) \circ \mu - [\mathcal{K}\alpha(\mathbf{Y})](t) \circ \mu||_{\mathbb{R}^2} \leq L||\mathbf{X} - \mathbf{Y}||_{C[0,T]}. \tag{2.17}$$

Proof 1 ***We have***

$$\begin{aligned} &||[\mathcal{K}\underbrace{\alpha(\mathbf{X})}_{\eta}](t) \circ \mu - [\mathcal{K}\underbrace{\alpha(\mathbf{Y})}_{\phi}](t) \circ \mu||_{\mathbb{R}^2} \\ &\leq ||(\mathcal{K}\eta)(t) - (\mathcal{K}\phi)(t)||_{\mathcal{L}(H^2,\mathbb{R}^2)}||\mu||_{H^2} \end{aligned}$$

where the operator norm is given by

$$||(\mathcal{K}\eta)(t)||_{\mathcal{L}(H^2,\mathbb{R}^2)} \equiv \sup_{\mu \in H^2} \frac{||\int_S [k(s,\eta)](t) \circ \mu \, ds||_{\mathbb{R}^2}}{||\mu||_{H^2}}$$

By our previous definition the integral operator can be expressed as

$$\begin{aligned} &||(\mathcal{K}\eta)(t) \circ \mu - (\mathcal{K}\phi)(t) \circ \mu||^2_{\mathbb{R}^2} \\ &= ||\int_S \begin{bmatrix} [k_D(s,\eta)](t) - [k_D(s,\phi)](t) & 0 \\ 0 & [k_L(s,\eta)](t) - [k_L(s,\phi)](t) \end{bmatrix} \begin{Bmatrix} \mu_D(s) \\ \mu_L(s) \end{Bmatrix} ds||_{\mathbb{R}^2} \\ &= ||\begin{Bmatrix} \int_S \{[k_D(s,\eta)](t) - [k_D(s,\phi)](t)\}\mu_D(s)ds \\ \int_S \{[k_L(s,\eta)](t) - [k_L(s,\phi)](t)\}\mu_L(s)ds \end{Bmatrix} ||_{\mathbb{R}^2} \end{aligned}$$

Applying Cauchy-Schwartz inequality on each element of the above vector

$$\Big|\int_S \Big([k(s,\eta)](t) - [k(s,\phi)](t)\Big)\mu(s)ds\Big|_{\mathbb{R}} \\ \leq \int_S |[k(s,\eta)](t) - [k(s,\phi)](t)|\ |\mu_{D/L}(s)|ds.$$

Using lemma 2.1 of [25] and considering the construction of the history dependent basis functions used in this problem, it can be shown that $\sup\big|[k(s,\eta)](t) - [k(s,\phi)](t)\big| \leq C||\eta-\phi||_{C[0,T]}$***. Therefore, it follows***

$$\begin{aligned} &\int_S \big|[k(s,\eta)](t) - [k(s,\phi)](t)\big|\ \big|\mu(s)\big|ds \\ &\quad \leq C||\eta-\phi||_{C[0,T]}\int_S 1|\mu(s)|ds \\ &\quad \leq C||\eta-\phi||_{C[0,T]}\ ||\mu||_H. \end{aligned} \tag{2.18}$$

$$\begin{aligned} &||(\mathcal{K}\eta)(t)\circ\mu - (\mathcal{K}\phi)(t)\circ\mu||^2_{\mathbb{R}^2} \\ &\quad \leq C^2||\eta-\phi||_{C[0,T]}\Big(||\mu_D||^2_H + ||\mu_L||^2_H\Big) \\ &\quad = C^2||\eta-\phi||_{C[0,T]}||\mu||^2_{H^2} \end{aligned} \tag{2.19}$$

If we assume that the angle of attack function $\alpha(\mathbf{X})$ ***is locally Lipschitz,***

$$||\eta-\phi||_{C[0,T]} = ||\alpha(\mathbf{X}) - \alpha(\mathbf{Y})||_{C[0,T]} \leq L_\alpha||\mathbf{X}-\mathbf{Y}||_{C[0,T]}$$

we can finally conclude

$$\big|\big|[\mathcal{K}\alpha(\mathbf{X})](t) - [\mathcal{K}\alpha(\mathbf{Y})](t)\big|\big|_{\mathcal{L}(H^2,\mathbb{R}^2)} \leq \big|\big|\mathbf{X}-\mathbf{Y}\big|\big|_{C[0,T]} \tag{2.20}$$

which induces Eq. 2.17.

With this locally Lipschitz condition for the history dependent operator, it is now straightforward to show that the Lipschitz condition in $C[0,T]$ applies for the function $F(t,X)$.

While the control of robotic systems as above without the generalized forces $\mathbf{Q}_a$ have been studied in detail, the control of a flapping wing robot with history dependent, unsteady aerodynamics has not been studied. In this paper we show that the common adaptive control strategy based on passivity principles for second order ODE systems can be extended to the case including history dependent terms. As we will see below, the critical step in the derivation relies on the introduction of the history dependent adjoint operator $(\mathcal{K}\alpha)^* : \mathbb{R}^2 \mapsto H^2$ in the parameter estimation equation. Define in the usual way [29] the auxiliary signals corresponding to the tracking error in joint space, filtered tracking error, modified desired velocity error and modified desired acceleration error as

$\mathbf{e} = \mathbf{q} - \mathbf{q_d}$, $\mathbf{r} = \dot{\mathbf{e}} + \Lambda\mathbf{e}$, $\mathbf{v} = \dot{\mathbf{q}}_\mathbf{d} - \Lambda\mathbf{e}$, and $\mathbf{a} = \ddot{\mathbf{q}}_\mathbf{d} - \Lambda\dot{\mathbf{e}}$. We choose the control law in the form

$$\begin{aligned}\tau(t) = \mathbf{M}(\mathbf{q}(t))\mathbf{a}(t) + \mathbf{C}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\mathbf{v}(t) + \mathbf{K}\mathbf{r}(t) \\ -\mathbf{\Psi_a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))(\mathcal{K}\alpha(\mathbf{q}, \dot{\mathbf{q}})(t) \circ \hat{\mu}(t)),\end{aligned} \tag{2.21}$$

and employ the estimate update law according to

$$\dot{\tilde{\mu}}(t) = -P^{-1}(\mathcal{K}\alpha)^*(t)\mathbf{\Psi_a}^T(\mathbf{q}(t), \dot{\mathbf{q}}(t))\mathbf{r}(t) \tag{2.22}$$

where $\tilde{\mu} = \mu - \hat{\mu}$ and $P : H \to H$ is a self adjoint, positive definite operator. Note that the update law for $\tilde{\mu}$ is defined in terms of the adjoint $(\mathcal{K}\alpha)^*$ of the integral operator $\mathcal{K}\alpha$. Also, since the true parameter vector μ is a constant, the estimation update law can be written by using $\dot{\hat{\mu}} = -\dot{\tilde{\mu}}$. The history dependent adjoint operator $(\mathcal{K}\alpha)^*(t) : \mathbb{R}^2 \mapsto H^2$ can be shown to be given by the mapping

$$(\mathcal{K}\alpha)^*(t) : \begin{Bmatrix} X_1 \\ X_2 \end{Bmatrix} \mapsto \begin{Bmatrix} (k_D(., \alpha))(t) \\ (k_L(., \alpha))(t) \end{Bmatrix} \in H^2.$$

Finally, the closed loop equations are obtained in the form

$$\begin{aligned}\mathbf{M}(\mathbf{q}(t))\dot{\mathbf{r}} + \mathbf{C}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\mathbf{r}(t) + \mathbf{K}\mathbf{r}(t) \\ = \mathbf{\Psi_a}(\mathbf{q}(t), \dot{\mathbf{q}}(t))(\mathcal{K}\alpha(\mathbf{q}, \dot{\mathbf{q}})(t) \circ (\mu - \hat{\mu}(t))\end{aligned} \tag{2.23}$$

The following theorem extends the passivity-based control law familiar from robotics to systems having history dependent terms.

Theorem 3 ***Suppose that the local solution for the closed loop equations can be extended to a maximal domain of definition corresponding to $h = \infty$ in Theorem 1. The tracking error $[e(t)^T \dot{e}(t)^T]^T$ and the estimation error $\tilde{\mu}(t)$ of the closed loop system of Eqn 2.23 are bounded and $e(t)$, $\dot{e}(t)$ converge to zero as $t \mapsto \infty$.***

Proof 2 ***We choose the Lyapunov function as***

$$\mathcal{V} = \frac{1}{2}\mathbf{r}^T\mathbf{M}(\mathbf{q})\mathbf{r} + \mathbf{e}^T\Lambda\mathbf{K}\mathbf{e} + \frac{1}{2}(P\tilde{\mu}, \tilde{\mu})_{H^2}. \tag{2.24}$$

If we differentiate with respect to time along system trajectories, we obtain

$$\dot{\mathcal{V}} = \mathbf{r}^T\mathbf{M}\dot{\mathbf{r}} + \frac{1}{2}\mathbf{r}^T\dot{\mathbf{M}}\mathbf{r} + 2\mathbf{e}^T\Lambda\mathbf{K}\dot{\mathbf{e}} + (P\dot{\tilde{\mu}}, \tilde{\mu})_{H^2}. \tag{2.25}$$

We exploit the properties of the history dependent adjoint operator $(\mathcal{K}\alpha)^*(t)$ to follow the same essential steps in the well-known proof of stability based on

passivity that is found in standard text such as [29].

$$
\begin{aligned}
\dot{\mathcal{V}} &= \mathbf{r}^T(-\mathbf{C}\mathbf{r} - \mathbf{K}\mathbf{r} + \mathbf{\Psi_a}\mathcal{K}\alpha \circ \tilde{\mu}) + \frac{1}{2}\mathbf{r}^T\dot{\mathbf{M}}\mathbf{r} + 2\mathbf{e}^T\Lambda\mathbf{K}\dot{\mathbf{e}} + (P\dot{\tilde{\mu}}, \tilde{\mu})_{H^2} \\
&= -\mathbf{r}^T\mathbf{C}\mathbf{r} + \frac{1}{2}\mathbf{r}^T(\dot{\mathbf{M}} - 2\mathbf{C})\mathbf{r} + (\mathbf{r}, \mathbf{\Psi_a}\mathcal{K} \circ \tilde{\mu})_{\mathbb{R}^n} + (P\dot{\tilde{\mu}}, \tilde{\mu})_{H^2} \\
&\quad + 2\mathbf{e}^T\Lambda\mathbf{K}\dot{\mathbf{e}} \\
&= -\dot{\mathbf{e}}^T\mathbf{K}\dot{\mathbf{e}} - \mathbf{e}^T\Lambda\mathbf{K}\Lambda\mathbf{e} + (P\dot{\tilde{\mu}} + (\mathcal{K}\alpha)^*\mathbf{\Psi_a}^T\mathbf{r}, \tilde{\mu})_{H^2} \\
&= -\dot{\mathbf{e}}^T\mathbf{K}\dot{\mathbf{e}} - \mathbf{e}^T\Lambda\mathbf{K}\Lambda\mathbf{e} \leq 0
\end{aligned}
$$

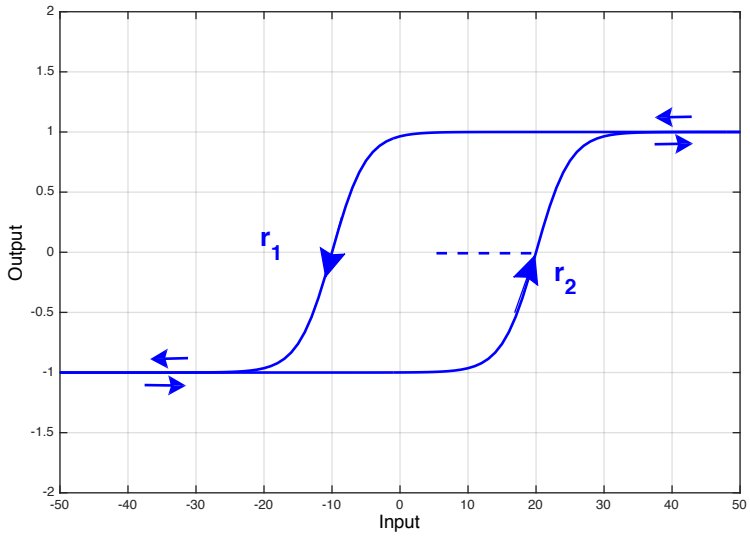
The remainder of the proof is based on Barbalat's lemma and is essentially identical to the approach taken for ordinary differential equations. Since $\mathcal{V}$ is positive definite and non increasing, it has a finite limit. We use a Lyapunov function and the uniform coercivity of the generalized mass matrix to conclude that e, r and, $\tilde{\mu}$ are bounded. From the equations of motion, by again using the uniform coercivity of $\mathrm{M}(\mathrm{q})$ and expansions of $\mathrm{C}(\mathrm{q}, \dot{\mathrm{q}})$ in terms of Christofel symbols, we conclude that $\ddot{e}$ is uniformly bounded. The only complication in this step is to show that the output of the history dependent operator $(\mathcal{K}\alpha(X))(t)$ is bounded, but this follows from the construction of the hysteretic kernels in terms of uniformly bounded ridge functions. It then follows that $\ddot{\mathcal{V}}$ is uniformly bounded and $\dot{\mathcal{V}} \to 0$. Considering the fact that $\mathcal{V}$ is lower bounded, we conclude from the Barbalat's lemma that $e \to 0$ and $\dot{e} \to 0$.

2.5 Computational Analysis and Results

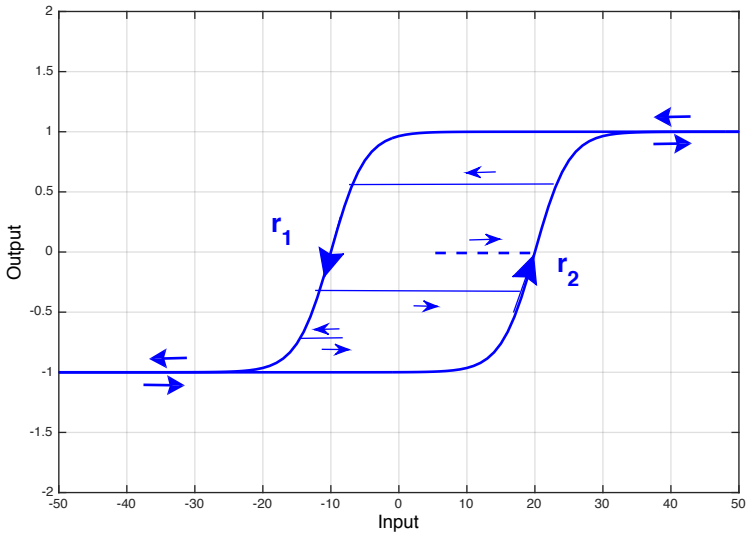
The kinematic model for the insect-inspired, flapping wing, robotic system studied in this paper has been implemented using the standard Denavit-Hartenberg convention. The stability and convergence properties of the adaptive controller introduced in Sect. 2.4 have been validated in several case studies. Figure 2.8 compares the CFD results and the output of the identified history dependent integral operator obtained by solving the optimization problem explained in Sect. 2.3.2. This identified model is used during control synthesis as the truth model representing the aerodynamic forces in the equations of motion. The adaptive controller is such that the unknown vector $\hat{\mu}(t)$ evolves through time. In addition, Fig. 2.9 depicts the identified $\mu_D(s)$ functions in two-dimensional space (i.e., the Preisach plane) for three different numbers of nodes consisting of $n_\mu = 66$, 231 and 496. The surface of the function gets smoother with increasing number of nodes in the Preisach plane as expected. The tracking performance and stability of the adaptive controller has also been validated by applying the following desired trajectories to joint angles: $\theta_{r1}(t) = \frac{\pi}{2}$, $\theta_{r2}(t) = \frac{\pi}{2} + \frac{\pi}{6}\sin(\omega t)$,and $\theta_{r3}(t) = 0$. As shown in Fig. 2.10, the trajectories of the joint angles converge asymptotically to their corresponding desired trajectories as time increases.

2.6 Conclusions

A general model was derived for a ground-based flapping robotic wing that has the form of a kinematic chain. Controllers for such ground based, articulated flapping wings are needed for the control of wind tunnel experimental prototypes as developed in [8]. To compute the critical unsteady mechanisms that contribute to flapping flight, a novel history dependent formulation of the unsteady aerodynamics in term of an integral operator was introduced. CFD was used to predict the drag and lift forces for the time-dependent history functions. The resulting evolution laws included a state consisting of a finite dimensional vector of flapping degrees of freedom and an infinite dimensional weighting function characterizing the integral operator. The analysis suggested that the proposed integral operator provided excellent predictions of lift and drag coefficients. The control strategy was derived by extending adaptive control laws for conventional robots to the history dependent case, and numerical experiments validate the stability guarantees derived via Lyapunov analysis. The contributions of the work include the introduction of a history dependent integral operator for the representation of unsteady flapping wing aerodynamics and the derivation of the associated adaptive tracking controllers for the resulting history dependent equations. While the results using a history dependent integral operator have proven to be promising for a ground-based flapping wing robot, the extensions of these principles to an MAV in free flight presents numerous additional difficulties. Most importantly, the passivity-based controller, which has been modified to account for history dependent terms, would not be directly appropriate in the case of free flight. We feel that this extension to free flight presents a challenging and important extension of this formulation.

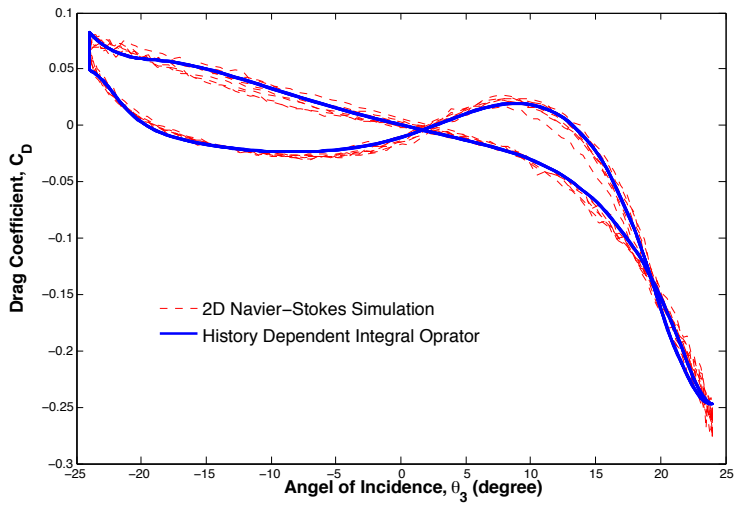


(a) Ridge Functions

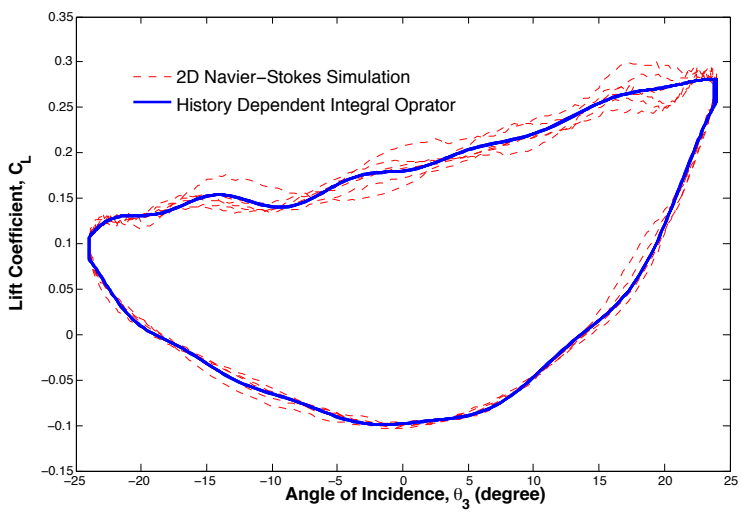


(b) Response

Figure 2.7: The history dependent kernel function

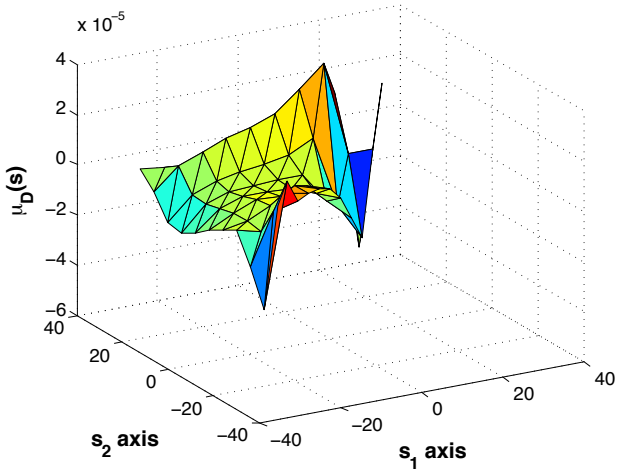


(a) Drag Coefficient C_D

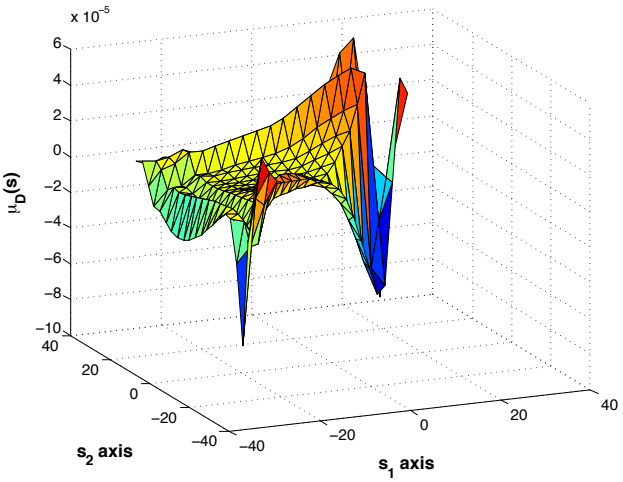


(b) Lift Coefficient C_L

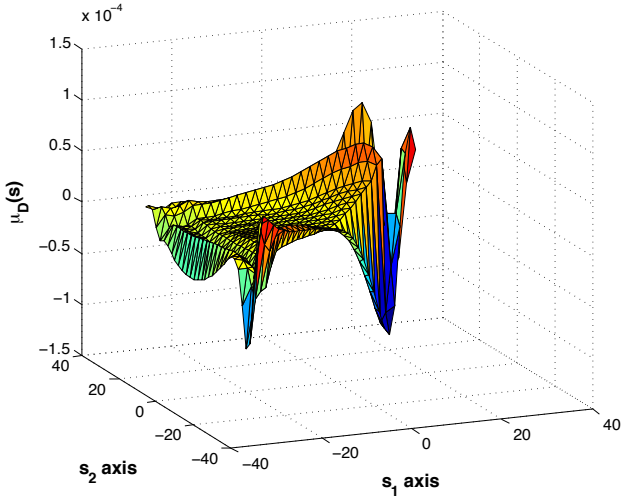
Figure 2.8: Comparison of drag and lift coefficients between CFD predictions and identification procedure



(a) $n_\mu = 66$

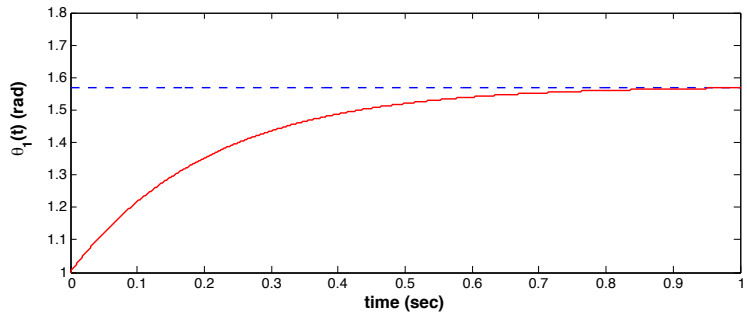


(b) $n_\mu = 231$

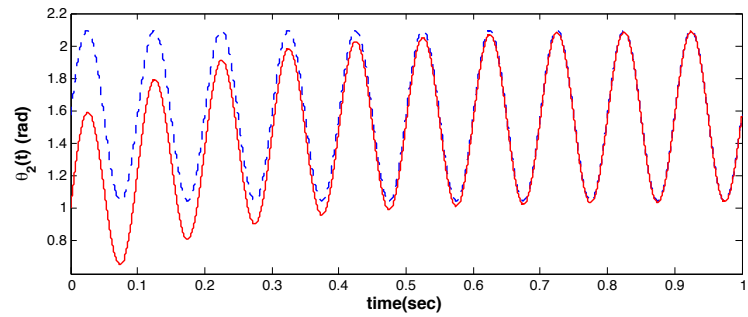


(c) $n_\mu = 496$

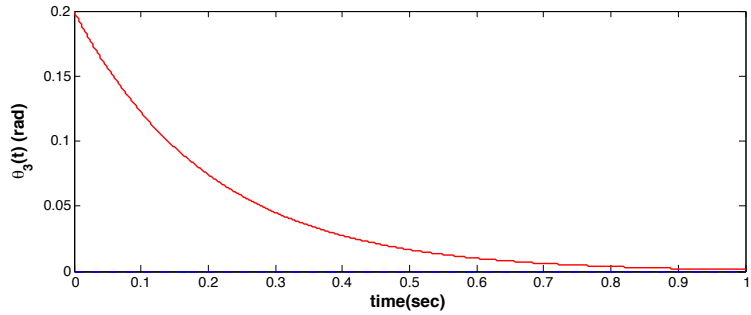
Figure 2.9: $\mu_D(s)$ for different number of nodes



(a) $\theta_1(t)$



(b) $\theta_2(t)$



(c) $\theta_3(t)$

Figure 2.10: Tracking performance for the joint angles.

Bibliography

[1] W. Shyy, H. Aono, S.K. Chimakurthi, P. Trizila, C.-K Kang, C.E.S. Cesnik, and H. Liu, "Progress in Flapping Wing Aerodynamics and Aeroelasticity," *Progress in Aerospace Sciences,* V. 46, No. 7, pp. 284-327, October, 2010.

[2] C. Orlowski, "Flapping Wing Micro-Air-Vehicles: An Analysis of the Importance of the Mass of the Wings to Flight Dynamics, Stability, and Control," *PhD Dissertation*, Department of Aerospace Engineering, University of Michigan, 2011.

[3] M.F. Platzer, K.D. Jones, J. Young, and C.S. Lai, "Flapping-Wing Aerodynamics: Progress and Challenges," *AIAA Journal*, Vol. 46, No. 9, pp. 2136-2149 (2008).

[4] K. Mazaheri and A. Ebrahimi, "Experimental Investigation on Aerodynamic Performance of a Flapping Wing Vehicle in Forward Flight," *Journal of Fluids and Structures,* Vol. 27, pp. 586-595 (2011).

[5] H. E.Taha, M. R.Hajj, P. S. Beran, "State-space Representation of the Unsteady Aerodynamics of Flapping Flight," *Aerospace Science and Technology,* Vol. 34, pp. 1-11 (2014).

[6] P. Gopalakrishnan and D. Tafti, "Effect of Wing Flexibility on Lift and Thrust Production in Flapping Flight," *AIAA Journal*, pp. 865-877, Vol. 48, No. 5, (2010).

[7] Cerhard Neuweiler, *The Biology of Bats*, Oxford University Press, 2000.

[8] Bahlman, J.W., Swartz, S.M., and Breuer, K.S. "Design and Charaterization of a Multi-Articulated Robotic Bat Wing", *Bioinspiration & Biomimetics*, Vol. 8, 016009 (2013).

[9] A.J. Bergou, S.M. Swartz, K.S. Breuer, and G. Taubin, "3D Reconstruction of Bat Flight Kinematics from Sparse Multiple Views," *proceedings of the IEEE ICCV: Int. Conf. on Computer Vision and Workshop on Dynamic Shape Capture*, 2011.

[10] S. Chung and M. Dorothy, "Neurobiologically Inspired Control of Engineered Flapping Flight, ***Journal of Guidance, Control, and Dynamics***, Vol. 33, No. 2, (2010).

[11] M. Wolf, L.C. Johansson, R. von Busse, Y. Winter, and A. Hedenstrom, "Kinematics of Flight and the Relationship to the Vortex Wake of a Pallas' Long Tongued Bat (***Glossophaga soricina***)," ***J. Exp. Biol.***, Vol. 213, pp. 2142-2153 (2010).

[12] S.M. Swartz and K.M. Middleton, "Biomechanics of the Bat Limb Skeleton: Scaling, Material Properties and Mechanics," ***Cells Tissues Organs***, Vol. 187, pp. 59-84 (2008).

[13] H. Rifai, N. Marchand, and G. Poulin, "Bounded Control of a Flapping Wing Micro Drone in Three Dimensions," ***Proceedings of the 2008 International Conference on Robotics and Automation,*** May 19-23, 2008, Pasadena, CA, IEEE.

[14] Y.-L. Park, B. Chen, N. O. Prez-Arancibia, D. Young, L. Stirling, R. J. Wood, E. C. Goldfield, and R. Nagpal, "Design and Control of a Biologically Inspired Active Soft Orthotic Device for Treating Ankle-Foot Pathologies," to appear in Bioinspiration and Biomimetics.

[15] P.-E.J. Duhamel, N.O. Prez-Arancibia, G.L. Barrows, and R.J. Wood, "Biologically Inspired Optical-Flow Sensing for Altitude Control of Flapping-Wing Microrobots," IEEE/ASME Transactions on Mechatronics, Vol. 18, No. 2, pp. 556-568 (2013).

[16] C. Orlowski and A. Girard, "Longitudinal Flight Dynamics of Flapping-Wing Micro Air Vehicles," ***Journal of Guidance, Control, and Dynamics,*** Vol. 35, No. 4, pp. 1115-1131 (2012).

[17] S. Dadashi, J. Gregory, Y. Lei, A. Kurdila, J. Bayandor, R. Mueller, "Adaptive Control of a Flapping Wing Robot Inspired by Bat Flight" ***AIAA SciTech***, 13-17 January 2014, National Harbor, Maryland.

[18] H.E. Taha, A.H. Nayfeh, M.R. Hajj, "Effect of the aerodynamic-induced parametric excitation on the longitudinal stability of hovering MAVs/insects," ***Nonlinear Dynamics***, Vol. 78, pp. 2399-2408 (2014).

[19] H.E. Taha, M.R. Hajj, A.H. Nayfeh, "Flight dynamics and control of flapping-wing MAVs: a review", ***Nonlinear Dyn***, Vol. 70, pp. 907-939 (2012).

[20] R. Dudley and C. P. Ellington, "Mechanics of Forward Flight in Bumblebees," ***The Journal of Experimental Biology***, Vol. 148, pp. 19-52 (1989).

[21] D. L. S. R. Jongerius, "Structural Analysis of a Dragonfly Wing," ***Experimental Mechanics***, Vol. 50, No. 9, pp. 1323-1334 (2010).

[22] **ANSYS Inc., Fluent 15.0 Theory Guide, 2014.**

[23] **ANSYS Inc., Fluent 15.0 Users Guide, 2014.**

[24] **Celik, I. B., Ghia, U., Roache, P. J., "Procedure for estimation and reporting of uncertainty due to discretization in CFD applications," *ASME J. Fluids Eng.*, Vol. 130, No. 7, 2008.**

[25] **A. Visintin, *Differential Models of Hysteresis*, Springer-Verlag, 1994.**

[26] **M.A. Krasnoselskii and A.V. Pokrovskii, *Systems with Hysteresis*, Springer, Berlin, 1989.**

[27] **R.J. Prazenica, A.J. Kurdila, "Multiwavelet Constructions and Volterra Kernel Identification", *Nonlinear Dyn*, Vol. 43, pp. 277-310 (2006).**

[28] **J.Koutney, M.Kruzik, A.Kurdila, and T.Roubicek, "Identification of Preisach-type Hysteresis Operators," *Numer. Funct. Anal. Optim.*, Vol. 29, pp. 149-160 (2008).**

[29] **M. Spong, S. Hutchinson and M. Vidyasagar, *Robot Modeling and Control*, John Wiley & Sons, 2006.**

[30] **F.L. Lewis, D.M. Dawson and C.T. Abdallah, *Robot Manipulator Control: Theory and Practice,* Marcel Dekker, Inc., 2004.**

[31] **C Corduneanu, "Abstract Volterra Equations: A Survey", *Mathematical and Computer Modelling*, Vol. 32, Issues 11-13, pp. 1503-1528 (2000).**

[32] **W. Silva, "Identification of Nonlinear Aeroelastic Systems Based on the Volterra Theory: Progress and Opportunities", *Nonlinear Dyn*, Vol. 39, pp. 25-62 (2005).**

[33] **C.M. Cheng, Z.K. Peng, W.M. Zhang, G. Meng, "Wavelet basis expansion-based Volterra kernel function identification through multilevel excitations", *Nonlinear Dyn*, Vol. 76, pp. 985999 (2014).**

[34] **C. Corduneanu, I Sandberg, "Volterra Equations and Applications", CRC Press, 2000.**

[35] **R. Lind, R.J. Prazenica, M.J. Brenner, "Estimating Nonlinearity Using Volterra Kernels in Feedback with Linear Models", *Nonlinear Dyn*, Vol. 39, pp. 3-23 (2005).**

[36] **V. P. Rudakov, "Qualitative Theory in a Banach Space, Lyapunov-Krasovskii Functionals, and Generalization of Certain Problems", *Ukrainian Mathematical Journal*, Vol. 30, No. 1, pp.101-103 (1978).**

[37] **A. Ilchmann, E. P. Ryan, and C. J. Sangwin, "Systems Of Controlled Functional Differential Equations and Adaptive Tracking," *SIAM J. CONTROL OPTIM*, Vol. 40, pp. 1746-1764 (2002).**

[38] **J. Bayandor, G. Bledt, S. Dadashi, A. Kurdila, and I. Murphy, "Adaptive Control for Bioinspired Flapping Wing Robots," *Proceedings of the 2013 American Control Conference,* pp. 609 - 614, October, 2012.**

[39] **I. P. Murphy, Modeling and control of flapping wing robots, Masters Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia, February 2013.**

[40] **Z. J. Wang, J. M. Birch, and M. H. Dickinson, "Unsteady forces and flows in low Reynolds number hovering flight: two-dimensional computations vs robotic wing experiments," *J.Exp. Biology*,", Vol. 207(3), pp. 449-460 (2004).**

Chapter 3

Online Estimation and Adaptive Control for a Class of History Dependent Functional Differential Equations

Online Estimation and Adaptive Control for a Class of History Dependent Functional Differential Equations [1]

Shirin Dadashi, Parag Bobade, Andrew J. Kurdila

Abstract

This paper presents sufficient conditions for the convergence of online estimation methods and the stability of adaptive control strategies for a class of history dependent, functional differential equations. The study is motivated by the increasing interest in estimation and control techniques for robotic systems whose governing equations include history dependent nonlinearities. The functional differential equations in this paper are constructed using integral operators that depend on distributed parameters. As a consequence the resulting estimation and control equations are examples of distributed parameter systems whose states and distributed parameters evolve in finite and infinite dimensional spaces, respectively. Well-posedness, existence, and uniqueness are discussed for the class of fully actuated robotic systems with history dependent forces in their governing equation of motion. By deriving rates of approximation for the class of history dependent operators in this paper, sufficient conditions are derived that guarantee that finite dimensional approximations of the online estimation equations converge to the solution of the infinite dimensional, distributed parameter system. The convergence and stability of a sliding mode adaptive control strategy for the history dependent, functional differential equations is established using Barbalat's lemma.

3.1 Introduction

It is typical in texts that introduce the fundamentals of modeling, stability, and control of robotic systems to assume that the underlying governing equations

[1] Sections 3.3 and 3.4 are contributions of my coworker Parag Bobade and I am endebted to his partnership on this material. I am also thankful for his help in developing the dynamic model of the wing section and the integrator for functional differential equations used in the numerical results. Sections 3.1 and 3.2 are result of our mutual work.

consist of a set of coupled nonlinear ordinary differential equations. This is a natural assumption when methods of analytical mechanics are used to derive the governing equations for systems composed of rigid bodies connected by ideal joints. A quick perusal of the textbooks [30], [29], or [22], for example, and the references therein gives a good account of the diverse collection of approaches that have been derived for this class of robotic system over the past few decades. Theses methods have been subsequently refined by numerous authors. Over roughly the same period, the technical community has shown a continued interest in systems that are governed by nonlinear, functional differential equations. These methods that helped to define the direction of initial efforts in the study of well-posedness and stability include [23], [20],[21], and their subsequent development is expanded in [12], [26], [25]. More recently, specific control strategies for classes of functional differential equations have appeared in [27], [16], and [14]. The research described in some cases above deals with quite general plant models. These can include classes of delay equations and general history dependent nonlinearities. One rich collection of history dependent models includes hysteretically nonlinear systems. General discussions of nonlinear hysteresis models can be found in [31] or [4], and some authors have studied the convergence and stability of systems with nonlinear hysteresis. For example, a synthesis of controllers for single-input / single-output functional differential equations is presented in [27] and [14], and these efforts include a wide class of scalar hysteresis operators.

The success of adaptive control strategies in classical manipulator robotics, as exemplified by [30], [22], [29], can be attributed to a large degree to the highly structured form of the governing system of nonlinear ordinary differential equations. As is well-known, much of the body of work in adaptive control for robotic systems relies on traditional linear-in-parameters assumptions.

The purpose of this paper is to explore the degree to which the approaches that have been so fruitful in adaptive control of robotic manipulators can be extended to robotic systems governed by certain history dependent, functional differential equations. Emulating the strategy used for robotic systems modeled by ordinary differential equations, we restrict attention to a class of hysteresis operators that satisfy a linear in *distributed* parameters condition. That is, the contribution to the functional differential equations takes the form of a nonlinear, history dependent operator that acts linearly on an infinite dimensional and unknown distributed parameter.

We illustrate the class of models that are considered in this paper by outlining a variation on two familiar problems encountered in robotic manipulator dynamics, estimation, and control. Consider the task of developing a model and synthesizing a controller for a flapping wing, test robot that will be used to study aerodynamics in a wind tunnel. See [1] for such a system that has been developed by researchers at Brown University over the past few years. Dynamics for a ground based flapping wing robot can be derived using analytical mechanics

in a formulation that is tailored to the structure of a serial kinematic chain [22], [30], [29]. The equations of motion take the form

$$M(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + \frac{\partial V}{\partial q} = Q_a(t, \mu) + \tau(t) \tag{3.1}$$

where $M(q)(t) \in \mathbb{R}^{N\times N}$ is the generalized inertia or mass matrix, $C(q(t), \dot{q}(t)) \in \mathbb{R}^{N\times N}$ is a nonlinear matrix that represents Coriolis and centripetal contributions, V is the potential energy, $Q_a(t, \mu) \in \mathbb{R}^N$ is a vector of generalized aerodynamic forces, and $\tau(t) \in \mathbb{R}^N$ is the actuation force or torque vector. The generalized forces $Q(t, \mu)$ due to aerodynamic loads are assumed to be expressed in terms of history dependent operators that are carefully discussed below in Section 3.2, and μ is the ***distributed parameter*** that defines the specific history dependent operator. For the current discussion, it suffices to note that the aerodynamic contributions are unknown, nonlinear, unsteady, and notoriously difficult to characterize.

We consider two specific sets of equations in this paper that are derived from the robotic Equations 3.1, both of which have similar form. We are interested in online identification problems in which we seek to find the final state and distributed parameters from observations of the states of the evolution equation. We are also interested in control synthesis where we choose the input to drive the system to some desired configuration, or to track a given input trajectory. To simplify our discussion, and following the standard practice for many control synthesis problems for robotics, we choose the original control input to be a partial feedback linearizing control that that reformats the control problem in a standard form. In the case of online identification, we choose the input $\tau = M(q)(u - G_1\dot{q} - G_0 q) - (C(q, \dot{q})\dot{q} + \frac{\partial V}{\partial q}(q))$ so that the governing equations take the form

$$\frac{d}{dt}\begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -G_0 & -G_1 \end{bmatrix}\begin{bmatrix} q(t) \\ \dot{q}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix}(M^{-1}(q)Q_a(t, \mu) + u(t)). \tag{3.2}$$

in terms of a new input u. The goal in the online identification problem is to learn the parameters μ and limiting values $q_\infty, \dot{q}_\infty$ from knowledge of the inputs and states $(u, q, \dot{q})$. We are also interested in tracking control problems. When the desired trajectory is given by q_d, we choose the input $\tau = M(q)(u + \ddot{q}_d - G_1\dot{e} - G_0 e) - (C(q, \dot{q})\dot{q} + \frac{\partial V}{\partial q}(q))$,and the equations governing the tracking error $e := q - q_d$ take the form

$$\frac{d}{dt}\begin{bmatrix} e(t) \\ \dot{e}(t) \end{bmatrix} = \begin{bmatrix} 0 & I \\ -G_0 & -G_1 \end{bmatrix}\begin{bmatrix} e(t) \\ \dot{e}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix}(M^{-1}(e + q_d)Q_a(t, \mu) + u(t)) \tag{3.3}$$

In either of the above two cases, we will show in the next section that the equations can be written in the general form

$$\dot{X}(t) = AX(t) + B((\mathcal{H}X)(t) \circ \mu + u(t)). \tag{3.4}$$

where $A \in \mathbb{R}^{m\times m}$ is the system matrix, $B \in \mathbb{R}^{m\times q}$ is the control input matrix, $u(t) \in \mathbb{R}^q$ is the corresponding input, and $(\mathcal{H}X)(t)$ is a history dependent operator that acts on the distributed parameter μ.

3.2 History Dependent Operators

There is a significant body of research to model and study the unsteady aerodynamic phenomena in flapping flight. Many different models have been presented in the last twenty years to study the aerodynamics and control of flapping flight. Numerically intensive computational fluid dynamics (CFD) presents a precise method to simulate and study the unsteady lift and drag aerodynamic forces. Generally CFD methods exploit high dimensional models that incorporate computationally expensive moving boundary techniques for the Navier-Stokes equations. They are powerful tools to explain some of the characteristics of the aerodynamic forces. One of the characteristics that has inspired the approach here is the history dependence of the aerodynamic lift and drag functions. We refer the interested reader to [35] to study this phenomena in detail. Although CFD methods are advantages in several aspects, they suffer from curse of dimensionality which makes them a very unfavorable choice for online control applications. In this section, we model the unsteady aerodynamics using history dependent operators. Moreover, we present a method that provides an alternative to a high dimensional aerodynamic model that typically evolves in a much lower dimensional space. We also study the accuracy of the presented method with respect to the resolution level of the lower dimensional model.

3.2.1 A Class of History Dependent Operators

Methods for modeling history dependent nonlinearities can be formulated using a wide array of approaches. Analytical methods for the study of such systems can be based on ordinary or partial differential equations, differential inclusions, functional differential equations, delay differential equations, or operator theoretic approaches. See references [15],[31],[30]. This paper treats evolution equations that are constructed using a specific class of history dependent operators $\mathcal{H}$ that are defined in terms of integral operators constructed from history dependent kernels. These operators are studied in general in [15] and [31]. In this paper the history dependent operators are mappings

$$\mathcal{H} : C([0,T),\mathbb{R}^m) \rightarrow C([0,T),P^*)$$

where the T is the final time of an interval under consideration, m is the number of input functions, q is the number of output functions, P is a Hilbert space of distributed parameters and its topological dual space P^*. We limit our consideration to input$-$output relationships that take the form

$$y(t) = (\mathcal{H}X)(t) \circ \mu \tag{3.5}$$

for each $t \in [0,T)$ where $y(t) \in \mathbb{R}^q$,$(\mathcal{H}X)(t) \in P^*$, and $\mu \in P$.

The definition of $\mathcal{H}$ in this paper is carried out in several steps. All of our history dependent operators $\mathcal{H}$ are defined by a superposition or weighting of elementary

hysteresis kernels κ_i that are continuous as mappings $\kappa_i : \Delta \times [0,T) \times C[0,T) \to C[0,T)$ for $i = 1, \ldots, \ell$. We first define the operator $h_i : C[0,T) \to C([0,T), P^*)$

$$(h_i f)(t) \circ \mu_i := \iint_\Delta \kappa_i(s,t,f)\mu_i(s)ds \tag{3.6}$$

for $\mu_i \in P_i$ and $P = P_1 \times \cdots \times P_\ell$. When we consider problems such as in our motivating examples and numerical case studies, we must construct vectors H of history dependent operators where we define the diagonal matrix

$$(HX)(t) := \begin{Bmatrix} h_1(a(X))(t) & & 0 \\ & \ddots & \\ 0 & & h_\ell(a(X))(t) \end{Bmatrix}$$

for each $t \in [0,T)$ where $a : \mathbb{R}^m \to \mathbb{R}$ is some nonlinear smooth map. Finally, our applications to robotics require that we consider

$$(\mathcal{H}X)(t) = b(X(t))(HX)(t), \tag{3.7}$$

where $b : \mathbb{R}^m \to \mathbb{R}^{q\times\ell}$ is some nonlinear, smooth map. In terms of our entrywise definitions of the input–output mappings, we have

$$y_i(t) := \sum_{j=1}^{l} b_{ij}(X(t))[h_j(a(X))](t) \circ \mu_j \tag{3.8}$$

for $i = 1, \cdots, q$.

In the following discussion, let κ be a generic representation of any of the kernels κ_i for $i = 1, \ldots, \ell$. We choose a typical kernel $\kappa(s,t,f)$ to be a special case of a generalized play operator [31]. We suppose that f is a piecewise linear function on $[0,t]$ with breakpoints $0 = t_0 < t_1 < \cdots < t_N = t$. The output function $t \mapsto \kappa(s,t,f)$, for a fixed $s = (s_1, s_2) \in \Delta \subset \mathbb{R}^2$ and piecewise linear $f : [0,t] \to \mathbb{R}$, is defined by the recursion where $\kappa^{n-1} := \kappa(s, t_{n-1}, f)$ and for $t \in [t_{n-1}, t_n]$ we have

$$\kappa(s,t,f) := \begin{cases} \max\{\kappa^{n-1}, \gamma_{s_2}(f(t))\} & f \text{ increasing on } [t_{n-1}, t_n], \\ \min\{\kappa^{n-1}, \gamma_{s_1}(f(t))\} & f \text{ decreasing on } [t_{n-1}, t_n]. \end{cases}$$

The recursion above depends on the choice of the left and right bounding functions $\gamma_{s_1}, \gamma_{s_2}$ that are depicted in Figure 3.1. There are given in terms of a single ridge function $\gamma : \mathbb{R} \to \mathbb{R}$ with

$$\begin{aligned} \gamma_{s_2}(\cdot) &:= \gamma(\cdot - s_2), \\ \gamma_{s_1}(\cdot) &:= \gamma(\cdot - s_1). \end{aligned} \tag{3.9}$$

As noted in [31], the definition of κ is extended for any $f \in C[0,T)$ by a continuity and density argument.

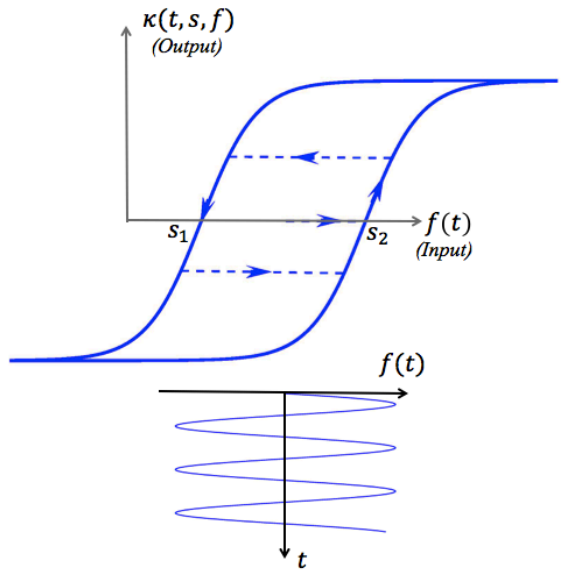


Figure 3.1: Elementary hysteresis kernel $t \to \kappa(s,t,f)$ for fixed $s = (s_1, s_2) \in \mathbb{R}^2$ and piecewise continuous $f : [0,t) \to \mathbb{R}$.

3.2.2 Approximation of History Dependent Operators

The integral operator introduced in Equation 3.6 allows for the representation of complex hysteretic response via the superposition or weighting of fundamental kernels κ_i. These fundamental kernels, each of which has simple input-output relationships, play the role of building blocks for modeling much more complex response characteristics. See [32] for studies of history dependent active materials, [35] for applications that represent nonlinear aerodynamic loading, or Section 3.6 of this paper to see an example of richness of this class of models. In this section we emphasize another important feature of this particular class of history dependent operators. We show that relatively simple approximation methods yields bounds on the error in approximation of the history dependent operator that are uniform in time and over the class of functions $\mu \in P$.

3.2.3 Approximation Spaces $\mathcal{A}_2^\alpha$

The approximation framework we follow in this paper is based on a straightforward implementation of approximation spaces discussed in detail in [11] or [10], and further developed by Dahmen in [6]. We will see that approximation of the class of history dependent operators under consideration exploit a well-known connection between the class of Lipschitz functions and certain approximation spaces as described in [10].

3.2.4 Wavelets and Approximation Spaces

Multiresolution Analysis (MRA) techniques use results from wavelet theory to model multiscale phenomena. To motivate our discussion, we begin by constructing Haar wavelets in one spatial dimension and subsequently discuss how the process can be easily extended to piecewise constant functions over triangulations in two dimensions. The Haar scaling function is defined as follows:

$$\phi(x) = \begin{cases} 1 & \textbf{if} \quad x \in [0,1) \\ 0 & \textbf{otherwise.} \end{cases}$$

The dilates and translates $\phi_{j,k}$ of ϕ are defined over $\mathbb{R}$ as

$$\phi_{j,k}(x) = 2^{j/2}\phi(2^j x - k) = 2^{j/2} 1_{\Delta_{j,k}}(x)$$

for $j = 0, \cdots, \infty$ and $k \in \mathbb{Z}$. It is important to note that with this normalization the functions $\{\phi_{j,k}\}_{k=0}^{2^j-1}$ are $L^2[0,1]$ orthonormal so that

$$\langle \phi_{j,k}, \phi_{j,l} \rangle = \int_{\mathbb{R}} \phi_{j,k}(x)\phi_{j,l}(x)\mathrm{d}x = \delta_{kl}.$$

In this equation $\Delta_{j,k} = \{x | x \in [2^{-j}k, 2^{-j}(k+1)]\}$ and $1_{\Delta_{j,k}}$ is the characteristic function of $\Delta_{j,k}$. For any fixed integer j, the $\phi_{j,k}$ form a orthonormal basis that spans the space of piecewise constants over $\{\Delta_{j,k}\}_{k=0}^{2^j-1}$. We let V_j denote space of piecewise constant functions

$$V_j = \underset{k=0,\cdots,2^j-1}{\text{span}} \{\phi_{j,k}\}. \tag{3.10}$$

Corresponding to Haar scaling function, the Haar wavelet ψ is defined as

$$\psi(x) = \begin{cases} 1 & x \in [0, \frac{1}{2}), \\ -1 & x \in [\frac{1}{2}, 1). \end{cases}$$

Again, the translates and dilates of $\psi_{j,k}$ of ψ are given by

$$\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k),$$

and the complement spaces W_j are defined by $W_j = \underset{k}{\text{span}}\{\psi_{j,k}\}$. It is straightforward to verify that the spaces V_{j-1} and W_{j-1} form an orthogonal direct sum of V_j. That is, we have

$$V_j = V_{j-1} \bigoplus W_{j-1}.$$

This process is well-known and standard in the literature as a means of constructing multiscale bases for $L^2[0,1]$. We will follow an analogous strategy to construct multiscale bases over the triangular domain depicted in Figure 3.2.

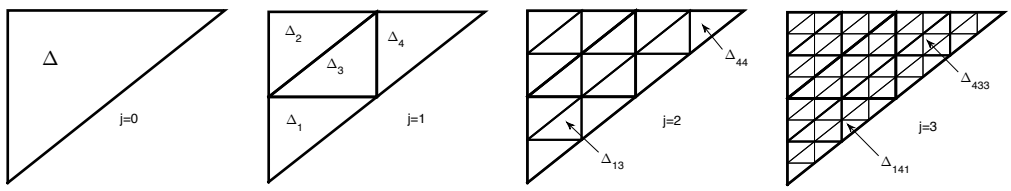


Figure 3.2: Regular refinement process for domain Δ

We first denote the characteristic functions over the triangular domain as shown in the Figure 3.2

$$1_{\Delta_s}(x) = \begin{cases} 1 & x \in \Delta_s \\ 0 & \textbf{otherwise.} \end{cases}$$

We next consider the regular refinement shown in Figure 3.2 where $\Delta_{i_1 i_2}$ is the i_2 child of Δ_{i_1}. In general $\Delta_{i_1 i_2 \ldots i_m i_{m+1}}$ is the $(m+1)^{st}$ child of $\Delta_{i_1 i_2 \ldots i_m}$. The multiscaling function $\phi_{j,k}$ is defined as

$$\phi_{j,k}(x) = \frac{1_{\Delta_{i_1 i_2 \ldots i_j}}(x)}{\sqrt{m(\Delta_{i_1 i_2 \ldots i_j})}}$$

where j refers to the level of refinement in the grid.

Since the history dependent operators $(\mathcal{H}X)(t)$ act on the infinite dimensional space $P = P_1 \times \cdots \times P_\ell$ of functions $\mu = (\mu_1, \ldots, \mu_\ell)$, we need approximations of these operators for computations and applications. In the discussion that follows we choose each function $\mu_i \in P_i := L^2(\Delta)$ where the domain $\Delta \subset \mathbb{R}^2$ is defined as

$$\Delta := \left\{ (s_1, s_2) \in \mathbb{R}^2 \middle| \underline{s} \leq s_1 \leq s_2 \leq \overline{s} \right\}.$$

The modification of the construction that follows for different domains Δ_i for the functions $\mu_i \in L^2(\Delta_i)$ is trivial, but notationally tedious, and we leave the more general case to the reader. Given the domain Δ we introduce a regular refinement depicted in Figure 3.2 and disscused in more detail in Appendix A. The set Δ is subdivided into $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ as shown, and each Δ_i is subdivided into $\Delta_{i1}, \Delta_{i2}, \Delta_{i3}, \Delta_{i4}$. Further subdivision recursively introduces the sets $\Delta_{i_1 \ldots i_j}$ for $i_j = 1, \ldots, 4$ that are the children of $\Delta_{i_1 \ldots i_{j-1}}$.

The characteristic functions $1_{\Delta_1}, \ldots, 1_{\Delta_4}$ define a collection of multiscaling functions $\phi^1, \ldots, \phi^4$ as defined in [19]. We define the space of piecewise constant functions V_j on grid refinement level j to be the span of the characteristic functions of the sets $\Delta_{i_1, \ldots, i_j}$, so that the dimension of V_j is 4^j. We denote by $\{\phi_{j,k}\}_{k=1,\ldots,4^j}$ the orthonormal basis obtained from these characteristic functions on a particular grid level, each normalized so that $(\phi_{j,k}, \phi_{j,\ell})_{L^2(\Delta)} = \delta_{k,l}$. Each of the basis functions $\phi_{j,k}$ will be proportional to the characteristic function $\phi^\ell(2^j x + d)$ for some $\ell \in, 1, 2, 3, 4$ and displacement vector d. It is straightforward in this case [19] to define 3 piecewise constant multiwavelets ψ^1, ψ^2, ψ^3 that that are used

to define functions $\psi_{j,m}$ for $m = 1, \ldots, 3 \times 4^j$ that span the complement spaces $W_j = \text{span}\,\{\psi_{j,m}\}_{m=1,\ldots,3\cdot 4^j}$ that satisfy

$$\underbrace{V_j}_{4^j\,functions} = \underbrace{V_{j-1}}_{4^{j-1}\,functions} \bigoplus \underbrace{W_{j-1}}_{3\times 4^{j-1}\,functions}.$$

It is a straightforward exercise to define $L^2(\Delta)-$orthonormal wavelets that span W_j for each $j \in \mathbb{N}_0$, but the nomenclature is lengthly. Since we do not use the wavelets specifically in this paper, the details are omitted. Each function $\psi_{j,m}$ is proportional to one of the three scaled and translated multiwavelet functions and satisfies the orthonormality conditions

$$\begin{array}{ll} (\psi_{j,k}, \psi_{m,\ell}) = \delta_{j,m}\delta_{k,\ell} & \text{for all } j,k,m,\ell, \\ (\psi_{j,k}, \phi_{m,\ell}) = 0 & \text{for } j \geq m \text{ and all } k,\ell. \end{array}$$

In the next step, we denote the orthogonal projection onto the span of the piecewise constants defined on a grid of resolution level j by Π_j so that

$$\Pi_j : P \to V_j.$$

Finally, we define the approximation space $\mathcal{A}_2^\alpha$ in terms of the projectors Π_j as

$$\mathcal{A}_2^\alpha := \left\{ f \in P \middle| \|f\|_{\mathcal{A}_2^\alpha} := \left(\sum_{j=0}^{\infty} 2^{2\alpha j} \|(\Pi_j - \Pi_{j-1})f\|_P^2 \right)^{1/2} \right\}.$$

Note that this is a special case of the more general analysis in [6]. We define our approximation method in terms of one point quadratures defined over the triangles $\Delta_{i_1\ldots i_j}$ that constitute the grid of level j that defines V_j. For notational convenience, we collect all triangles at a fixed level j in the singly indexed set

$$\{\Delta_{j,k}\}_{k\in\Lambda_j} := \left\{\Delta_{i_1\ldots i_j}\right\}_{i_1,\ldots,i_j\in 1,2,3,4}$$

where $\Lambda_j := \{k \in \mathbb{N} \quad |1 \leq k \leq 4^j\}$, and the quadrature points are chosen such that $\xi_{j,k} \in \Delta_{j,k}$ for $k = 1, \cdots, \Lambda_j$. We now can state our principle approximation result for the class of history dependent operators in this paper.

Theorem 1 ***Suppose that the function γ that defines the history dependent kernel in Equation 3.9 is a bounded function in $C^\alpha(\mathbb{R})$, and define the approximation h_j associated with the grid level j of the history dependent operator h to be***

$$(h_j f)(t) \circ \mu := \iint_\Delta \left(\sum_{\ell\in\Gamma_j} 1_{\Delta_{j,\ell}}(s)\kappa(\xi_{j,\ell}, t, f) \right) \mu(s)ds.$$

Then there is a constant $C > 0$ such that

$$|(h_j f)(t) \circ \mu - (hf)(t) \circ \mu| \leq C2^{-(\alpha+1)j} \tag{3.11}$$

for all $f \in C[0,T]$***,*** $t \in [0,T]$***, and*** $\mu \in P$***. If in addition*** $\mu \in \mathcal{A}_2^{\alpha+1}$***, there is a constant*** $\tilde{C} > 0$ ***such that***

$$|(h_j f)(t) \circ \Pi_j \mu - (hf)(t) \circ \mu| \leq \tilde{C} 2^{-(\alpha+1)j} \tag{3.12}$$

for all $f \in C[0,T]$ ***and*** $t \in [0,T]$***.***

Proof 3 ***We first prove the inequality in Equation 3.11. By definition of the operator*** h***, we can write***

$$\begin{aligned}
|(h_j f)(t) \circ \mu - (hf)(t) \circ \mu| &\leq \iint_\Delta \left| \left(\sum_{k \in \Lambda_j} 1_{\Delta_{j,k}}(s) \kappa(\xi_{j,k}, t, f) - \kappa(s,t,f) \right) \mu(s) \right| ds \\
&\leq \iint_\Delta \left| \sum_{k \in \Lambda_j} 1_{\Delta_{j,k}}(s) \left(\kappa(\xi_{j,k}, t, f) - \kappa(s,t,f) \right) \right| |\mu(s)| \, ds
\end{aligned}$$

Since the ridge function γ ***is a bounded function in*** $C^\alpha(\mathbb{R})$***, the output mapping*** $s \mapsto \kappa(s,t,f)$ ***is also a bounded function in*** $C^\alpha(\Delta)$ ***where the Lipschitz constant is independent of*** $t \in [0,T]$ ***and*** $f \in C[0,T]$***. Using Proposition 2.5 of [31], we have***

$$\begin{aligned}
|(h_j f)(t) \circ \mu - (hf)(t) \circ \mu| &\leq \iint_\Delta \left| \sum_{k \in \Lambda_j} 1_{\Delta_{j,k}}(s) L \|\xi_{j,k} - s\|^\alpha \right| |\mu(s)| \, ds \\
&\leq L \sum_{k \in \Lambda_j} \left(m(\Delta_{j,k}) \left(\frac{\sqrt{2}(\overline{s} - \underline{s})}{2^j} \right)^\alpha \right) \iint_{\Delta_{j,k}} |\mu(s)| ds \\
&\leq L \sum_{k \in \Lambda_j} \left(m(\Delta_{j,k}) \left(\frac{\sqrt{2}(\overline{s} - \underline{s})}{2^j} \right)^\alpha \right) m^{1/2}(\Delta_{j,k}) \|\mu\|_P \\
&\leq L 2^{2j} \left(\frac{1}{2} \frac{(\overline{s} - \underline{s})^2}{2^{2j}} \left(\frac{\sqrt{2}(\overline{s} - \underline{s})}{2^j} \right)^\alpha \right) \left(\frac{1}{2} \frac{(\overline{s} - \underline{s})^2}{2^{2j}} \right)^{1/2} \|\mu\|_P = C 2^{-(\alpha+1)j} \|\mu\|_P
\end{aligned}$$

Since we have

$$|(h_j f)(t) \circ \Pi_j \mu - (hf)(t) \circ \mu| \leq |(h_j f)(t) \circ \Pi_j \mu - (h_j f)(t) \circ \mu| + |(h_j f)(t) \circ \mu - (hf)(t) \circ \mu| \,,$$

the second inequality in Equation 3.12 follows from the first Equation 3.11 provided we can show that

$$|(h_j f)(t) \circ (\mu - \Pi_j \mu)| \leq C 2^{-(\alpha+1)j}$$

for some constant C***. But it is a standard feature of the approximation spaces that if*** $\mu \in \mathcal{A}_2^{\alpha+1}$***, then*** $\|\mu - \Pi_j \mu\|_P \leq 2^{-(\alpha+1)j} \|\mu\|_{\mathcal{A}_2^{\alpha+1}}$***. To see why this is so, suppose***

that $\mu \in \mathcal{A}_2^{\alpha+1}$***. We have***

$$\begin{aligned}\|\mu - \Pi_j\mu\|_P^2 &= \sum_{k=j+1}^{\infty} \|(\Pi_k - \Pi_{k-1})\mu\|_P^2 \\ &\leq \sum_{k=j+1}^{\infty} 2^{-2(\alpha+1)k} 2^{2(\alpha+1)k} \|(\Pi_k - \Pi_{k-1})\mu\|_P^2 \\ &\leq 2^{-2(\alpha+1)j} \sum_{k=j+1}^{\infty} 2^{2(\alpha+1)k} \|(\Pi_k - \Pi_{k-1})\mu\|_P^2 \leq 2^{-2(\alpha+1)j} \|\mu\|_{\mathcal{A}_2^{\alpha+1}}^2.\end{aligned}$$

When we apply this to our problem, the upper bound follows immediately

$$\begin{aligned}|(h_j f)(t) \circ (\mu - \Pi_j\mu)| &\leq \sup_{t\in[0,T]} \|(h_j f)(t)\|_{P^*} \|\mu - \Pi_j\mu\|_P \\ &\leq C 2^{-(\alpha+1)j},\end{aligned}$$

since the boundedness of the ridge function γ ***implies the uniform boundedness of the history dependent operators*** $(h_j f)(t)$ ***over*** $[0, T]$***.***

Theorem 1 can now be used to establish error bounds for input-output maps that have the form in Equation 3.8.

Theorem 2 ***Suppose that the hypotheses of Theorem 1 hold. Then we have***

$$\|(\mathcal{H}X)(t) - (\mathcal{H}_j X)(t)\Pi_j\| \lesssim 2^{-(\alpha+1)j}.$$

Where $\mathcal{H}$ is defined in Equations 3.5,3.8, and $\mathcal{H}_j$ is defined in Equations 3.13, 3.14 and 3.15 below.

Proof 4 ***Recall that for*** $i = 1 \ldots q$ ***we had***

$$y_i(t) = \sum_{\ell=1\ldots l} b_{i\ell}(X(t))(h_\ell(a(X))(t) \circ \mu_\ell.$$

In matrix form this equation can be expressed as

$$\begin{bmatrix} y_1(t) \\ \vdots \\ y_q(t) \end{bmatrix} = \underbrace{\begin{bmatrix} b_{11}(X(t)) & \ldots & b_{1\ell}(X(t)) \\ \vdots & \ddots & \vdots \\ b_{q1}(X(t)) & \ldots & b_{q\ell}(X(t)) \end{bmatrix}}_{\mathbb{R}^{q\times l}} \underbrace{\begin{bmatrix} h_1(a(X))(t) \circ \mu_1 \\ \vdots \\ h_\ell(a(X))(t) \circ \mu_\ell \end{bmatrix}}_{(HX)(t)\circ\mu}.$$

It follows that,

$$y(t) = \underbrace{(\mathcal{H}X)(t) \circ \mu}_{\in\ \mathbb{R}^q} = \underbrace{b(X(t))}_{\in\mathbb{R}^{q\times\ell}} \underbrace{(HX)(t) \circ \mu}_{\mathbb{R}^\ell}.$$

By assumption $X \in C([0,T],\mathbb{R}^m)$***. The construction of*** H ***and*** $\mathcal{H}$ ***guarantees that*** $(HX)(t) \in \mathcal{L}(P,\mathbb{R}^l)$,

$H : C([0,T],\mathbb{R}^m) \to C\left([0,T],\mathcal{L}(P,\mathbb{R}^l)\right)$, ***and*** $\mathcal{H} : C([0,T],\mathbb{R}^m) \to C([0,T],\mathbb{R}^q)$. ***In this proof we denote by*** $(\mathbb{R}^l, \|.\|_u)$ ***the norm vector space that endows*** $\mathbb{R}^l$ ***with the*** l^m ***norm*** $\|v\|_u := \left(\sum_{i=1}^l |v_i|^u\right)^{\frac{1}{u}}$ ***for*** $1 \le u \le \infty$***. The normed vector space*** $(\mathbb{R}^{q\times l}, \|.\|_{s,u})$ ***denotes the induced operator norm on matrices that map*** $(\mathbb{R}^l, \|.\|_u)$ ***into*** $(\mathbb{R}^q, \|.\|_s)$***. Now we define an approximation on the mesh level*** j ***of*** $\mathcal{H}$ ***to be***

$$(\mathcal{H}_j X)(t) = b(X(t)((H_j X)(t)\Pi_j, \tag{3.13}$$

$$(H_j X)(t) = \begin{bmatrix} h_{1,j}(a(X))(t) & 0 & \cdots & 0 \\ 0 & h_{2,j}(a(X))(t) & 0 & \vdots \\ \vdots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & h_{l,j}(a(X))(t) \end{bmatrix} \tag{3.14}$$

and

$$h_{i,j}(t) \circ \nu = \iint_\Delta \sum_{k\in\Lambda_j} 1_{\Delta_{k,j}}(s)\kappa(\xi_{j,k},t,f)\nu(s)\mathrm{d}s \tag{3.15}$$

for $i = 1\dots\ell$ ***and*** $\nu_i \in P_i$***. To simplify the derivation or an error bound for approximation of*** $\mathcal{H}X(t)\circ\mu$***, let*** $(HX)(t)\circ\mu$ ***be denoted by*** $g(t)$***. We have assumed that*** $X \to b(X)$ ***and*** $t \to X(t)$ ***are continuous mappings. There fore*** $t \mapsto b(X(t))$ ***is continuous and on a compact set*** $[0,T]$***, and*** $b(X(\cdot)) \in C([0,T],\mathbb{R}^{q\times l})$***. We therefore by definition have***

$$\begin{aligned} \|b(X(t))\|_{(\mathbb{R}^{q\times l},\|\cdot\|_{s,u})} &\le \sup_{\tau\in[0,T]} \|b(X(\tau))\|_{(\mathbb{R}^{q\times l},\|\cdot\|_{s,u})}, \\ &= \|b(X(\cdot))\|_{C([0,T],(\mathbb{R}^{q\times l},\|\cdot\|_{s,u}))}, \\ \|b(X(t))g(t)\|_{(\mathbb{R}^q,\|\cdot\|_q)} &\le \|b(X(\cdot))\|_{C([0,T],(\mathbb{R}^{q\times l},\|\cdot\|_{s,u}))}\|g(t)\|_{(\mathbb{R}^l,\|\cdot\|_u)}. \end{aligned}$$

with the norms explicitly denoted in the subscript. For $t \in [0,T]$***, and applying these definitions,***

$$\begin{aligned} \|b(X(t))\left((HX)(t) - (H_jX)(t)\Pi_j\right)\mu\|_{(\mathbb{R}^q,\|\cdot\|_s)} &\le \|b(X(t))\|_{(\mathbb{R}^{q\times l},\|\cdot\|_{s,u})}\|((HX)(t) - (H_jX)(t)\Pi_j)\circ\mu\|_{(\mathbb{R}^l,\|\cdot\|_u)} \\ &\le \|b(X(\cdot))\|_{(C([0,T],(\mathbb{R}^{q\times l},\|\cdot\|_{s,u}))}\|((HX)(t) - (H_jX)(t)\Pi_j)\circ\mu\|_{(\mathbb{R}^l,\|\cdot\|_u)} \end{aligned}$$

with

$$\|((HX)(t) - (H_jX)(t)\Pi_j)\circ\mu\|_{(\mathbb{R}^l,\|\cdot\|_u)} = \left\| \begin{bmatrix} h_{1,j}(a(X))(t) - h_{1,j}(a(X))(t)\Pi_j & & & 0 \\ & h_{2,j}(a(X))(t) - h_{2,j}(a(X))(t)\Pi_j & & \\ & & \ddots & \\ 0 & & & h_{l,j}(a(X))(t) - h_{l,j}(a(X))(t)\Pi_j \end{bmatrix} \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_l \end{bmatrix} \right\|_{(\mathbb{R}^l,\|\cdot\|_u)}.$$

Therefore we can now write

$$\|((HX)(t) - (H_jX)(t)\Pi_j)\circ\mu\|_{\left(\mathbb{R}^l,\|\cdot\|_u\right)} \le \|((HX)(t) - (H_jX)(t)\Pi_j)\|_{(\mathcal{L}(P,(\mathbb{R}^l,\|\cdot\|_u)))}\|\mu\|_P.$$

Hence, recalling Theorem 1 we can now derive the convergence rate

$$\begin{aligned}\|((HX)(t)-(H_jX)(t)\Pi_j)\|_{\left(\mathcal{L}\left(P,\left(\mathbb{R}^l,\|\cdot\|_u\right)\right)\right)} &= \sup_{\|\mu\|<1} \|\left((HX)\left(t\right)-\left(H_jX\right)\left(t\right)\Pi_j\right)\circ\mu\|_{(\mathbb{R}^l,\|\cdot\|_u)} \\ &\leq \sup_{\|\mu\|<1} |((h_{i,j}(a(X)))(t)-(h_{i,j}(a(X)))(t)\Pi_j)\circ\mu| \\ &\leq \hat{C}2^{-(\alpha+1)j}.\end{aligned}$$

Therefore we obtain the final bound

$$\|(\mathcal{H}X)(t)-(\mathcal{H}_jX)(t)\Pi_j)\|_{(\mathbb{R}^l,\|\cdot\|_u)} \lesssim 2^{-(\alpha+1)j}, \tag{3.16}$$

for all $t\in[0,T]$***.***

3.3 Well-Posedness: Existence and Uniqueness

The history dependent governing equations studied in this paper are a special case of the more general class of abstract Volterra equations or functional differential equations. A general treatise on abstract Volterra equations can be found in [5], while various generalizations of theory for the existence and uniqueness of functional differential equations have been given in [12], [25], [16]. We have noted in Section ?? that the general form of the governing equations we consider in this paper have the form

$$\dot{X}(t)=AX(t)+B((\mathcal{H}X)(t)\circ\mu+u(t)) \tag{3.17}$$

where the state vector $X(t)\in\mathbb{R}^m$**, the control inputs** $u(t)\in\mathbb{R}^q$**,** $A\in\mathbb{R}^{m\times m}=\mathbb{R}^{2n\times 2n}$ **is a Hurwitz matrix, and** $B\in\mathbb{R}^{m\times q}$ **is the control input matrix. We make the following assumptions about the history dependent operators** $\mathcal{H}$**:**

H1) $\mathcal{H}:C([0,\infty),\mathbb{R}^m)\mapsto C([0,\infty),P^*)$

H2) $\mathcal{H}$ **is causal in the sense that for all** $x,y\in C([0,\infty);\mathbb{R}^m)$**,**

$$x(\cdot)\equiv y(\cdot) \text{ on } [0,\tau] \implies (\mathcal{H}x)(t)=(\mathcal{H}y)(t) \quad \forall\, t\in[0,\tau].$$

H3) Define the closed set consisting of all continuous functions f **that remain within radius** r **of the initial condition** X_0 **over the closed interval** $[t,t+h]$**,**

$$\overline{\mathcal{B}}_{[t,t+h],r}(X_0):=\left\{f\in C([0,h),\mathbb{R}^m)\middle| f(0)=X_0 \text{ and } \|f(s)-X_0\|_{\mathbb{R}^m}\leq r \text{ for } s\in[t,t+h]\right\},$$

for a fixed $X_0\in\mathbb{R}^m$**. For each** $t\geq 0$**, we assume that there exist** $h,r,L>0$ **such that**

$$\|(\mathcal{H}X)(s)-(\mathcal{H}Y)(s)\|_{P^*}\leq L\|X-Y\|_{[t,t+h]} \qquad s\in[t,t+h] \tag{3.18}$$

for all $X,Y\in\overline{\mathcal{B}}_{[t,t+h],r}(X_0)$**.**

Our first result guarantees the existence and uniqueness of a local solution to Equation 3.4, and also describes an important case when such local solutions can be extended to $[0, \infty)$. This theorem can be proven via the existence and uniqueness Theorem 2.3 in [16] for functional delay-differential equations. However, since we are not interested in delay differential equations in this paper, but rather on a highly structured class of integral hysteresis operators, the proof can be much simplified.

Theorem 3 ***Suppose that the history dependent operator $\mathcal{H}$ satisfies the hypotheses (H1),(H2),(H3). Then there is a $\delta > 0$ such that Equation 3.17 has a solution $X \in C([0, \delta), \mathbb{R}^m)$. Suppose the interval $[0, \delta)$ is extended to the maximal interval $[0, \omega) \subset [0, \delta)$ over which such a solution exists. If the solution is bounded, then $[0, \omega) = [0, \infty)$.***

Corollary 1 ***Suppose that the history dependent operator $\mathcal{H}$ in Equation 3.17 is defined as in Equation 3.7 and 3.8 in terms of a globally Lipschitz, bounded continuous ridge function $\gamma : \mathbb{R} \to \mathbb{R}$ in Equation 3.9. Then Equation 3.17 has a unique solution $X \in C([0, \infty), \mathbb{R}^m)$ for each $\mu \in P$.***

Proof 5 ***For completeness, we outline a simplified version the proof of Theorem 3 for our class of history and parameter dependent equations. As a point of comparison, the reader is urged to compare the proof below to the conventional proof for systems of nonlinear ordinary differential equations, such as in [17]. If we integrate the equations of motion in time, we can define an operator $T : C([0, h), \mathbb{R}^m) \to C([0, h), \mathbb{R}^m)$ from***

$$X(t) = X_0 + \int_0^t AX(\tau) + B((\mathcal{H}X)(\tau) \circ \mu + u(\tau))\mathrm{d}\tau,$$
$$X(t) = (TX)(t),$$

for all $t \in [0, h]$. As introduced in hypothesis (H3), we select $h, r > 0$ and define

$$\overline{\mathcal{B}}_{[0,h],r}(X_0) := \left\{ X \in C([0, h), \mathbb{R}^m) \middle| X(0) = X_0, \|X_0 - X\|_{[0,\delta]} \leq r \right\}.$$

such that the local Lipschitz condition in Equation 3.18 holds. Now we consider restricting the equation to a subinterval $[0, \delta] \subseteq [0, h]$, and investigate conditions on T that enable the application of the contraction mapping theorem. We first study what conditions on $\delta > 0$ are sufficient to guarantee that $T : \overline{\mathcal{B}}_{[0,\delta],r}(X_0) \to$

$\overline{\mathcal{B}}_{[0,\delta],r}(X_0)$. ***We have***

$$\begin{aligned}
\|TX(t) - X_0\|_{\mathbb{R}^m} &\leq \int_0^t \|AX(s) + B((\mathcal{H}X(s) \circ \mu + u(s))\|_{\mathbb{R}^m} \mathrm{d}s \\
&\leq \int_0^t \Bigg(\|A\| \|X(s) - X_0\|_{\mathbb{R}^m} + \underbrace{\|AX_0\|_{\mathbb{R}^m}}_{\leq \|A\|\|X_0\| = M_A} \\
&\qquad + \|B\| \underbrace{(\|(\mathcal{H}X)(s) - (\mathcal{H}X_0)(s)\|_{P^*}}_{\leq L\|X - X_0\|_{[0,\delta]}} \underbrace{\|\mu\|_P}_{M_\mu} + \underbrace{\|(\mathcal{H}X_0)(s)\|_{P^*}}_{\leq M_H = \|\mathcal{H}X_0\|_{C([0,\delta],P^*)}} + \underbrace{\|u\|_{C([0,h),\mathbb{R}^p)}}_{\leq M_u = \|u\|_{C([0,\delta],\mathbb{R}^q)}}) \Bigg) \\
&\leq ((\|A\| + \|B\| M_\mu L)r + M_A + \|B\| M_T)t \\
&\leq ((\|A\| + \|B\| M_\mu L)r + M_A + \|B\| M_T)\delta
\end{aligned}$$

where $M_T = M_H + M_u$. ***Now we restrict*** δ ***so that***

$$\Big((\|A\| + \|B\| M_\mu L)r + M_A + \|B\| M_T\Big)\delta \leq r,$$

which implies

$$\delta < \frac{r}{(\|A\| + \|B\| M_\mu L)r + M_A + \|B\| M_T)}.$$

We thereby conclude that

$$\|TX(t) - X_0\|_{C([0,h),\mathbb{R}^p)} \leq r \quad \textbf{\textit{for}} \quad t \in [0, \delta],$$

and it follows that $T : \overline{\mathcal{B}}_{[0,\delta],r} \to \overline{\mathcal{B}}_{[0,\delta],r}$. ***Next we study conditions on*** δ ***that guarantee that*** $T : \overline{\mathcal{B}}_{[0,\delta],r} \to \overline{\mathcal{B}}_{[0,\delta],r}$ ***is a contraction. We compute directly a bound on the difference of the output as***

$$\begin{aligned}
\|(TX)(t) - (TY)(t)\|_{\mathbb{R}^m} &\leq \int_0^t \|AX(s) - AY(s) + B((\mathcal{H}X)(s) - (\mathcal{H}Y(s)) \circ \mu)\|_{\mathbb{R}^m} \mathrm{d}s \\
&\leq (\|A\| + \|B\| M_\mu L_\mu)\|X - Y\|_{\mathbb{R}^m} \delta.
\end{aligned}$$

If we choose

$$\delta < min\left\{h, \frac{r}{(\|A\| + \|B\| M_\mu L)r + M_A + \|B\| M_T)}, \frac{1}{\|A\| + \|B\| M_\mu L}\right\},$$

it is apparent that T ***is a contraction that maps the closed set*** $\overline{\mathcal{B}}_{[0,\delta],r}$ ***into itself. There is a unique solution in*** $\overline{\mathcal{B}}_{[0,\delta],r}$ ***on*** $[0, \delta]$.

3.4 Online Identification

A substantial literature has emerged that treats online estimation problems for linear or nonlinear plants governed by systems of ordinary differential equations. Approaches for these finite dimensional systems that are based on variants of

Lyapunov's direct method can be found in any of a number of good texts including, for instance, [24], [28], or [18]. The general strategies that have proven fruitful for such finite dimensional systems have often been extended to classes of systems whose dynamics evolve in an infinite dimensional space: distributed parameter systems. A discussion of the general considerations for identification of distributed parameter systems can be found in [2], for example, while studies that are specifically relevant to this paper include [7], [8], [9], and [3].

In this section we adapt the framework introduced in [3] to our class of history dependent, functional differential equations. The approach in [3] assumes that the state equations for the distributed parameter system have first order form, and they are cast in terms of a nonlinear, parametrically dependent bilinear form that is coercive. The resulting equations that govern the error in state and in distributed parameter estimates is a nonlinear function of the state trajectory of the plant. In contrast, a similar strategy in this paper yields error equations that depend nonlinearly on the history of the state trajectory.

The general online estimation problem discussed in this section assumes that we observe the value of the state $X(t) \in \mathbb{R}^m$ at each time $t \geq 0$ that depends on some unknown distributed parameter $\mu \in P$, and subsequently use the observed state to construct estimates $\hat{X}$ of the states and $\hat{\mu}$ of the distributed parameters. We construct online estimates that evolve on the state space $\mathbb{R}^m \times P$ according to the time varying, distributed parameter system equations

$$\begin{aligned} \dot{\hat{X}}(t) &= A\hat{X}(t) + B\left((\mathcal{H}X)(t) \circ \hat{\mu}(t) + u(t)\right), \\ \dot{\hat{\mu}}(t) &= -\left(B(\mathcal{H}X)(t)\right)^* \hat{X}(t), \end{aligned} \tag{3.19}$$

for $t \geq 0$ where the initial conditions are $\hat{X}_0 := X_0$, $\hat{\mu}(0) := \mu_0$. In these equations, we denote the adjoint operator L^* for any bounded linear operator L. These equations can be understood as incorporating a natural choice of a parameter update law. The learning law above can be interpreted as generalization of the conventional gradient update law that features prominently in approaches for finite dimensional systems [18] and that has been extended to distributed parameter systems in [3]. It is immediate that the error in estimation of the states $\tilde{X} := X - \hat{X}$ and in the distributed parameters $\tilde{\mu} := \mu - \hat{\mu}$ satisfy the homogeneous system of equations

$$\begin{Bmatrix} \dot{\tilde{X}}(t) \\ \dot{\tilde{\mu}}(t) \end{Bmatrix} = \begin{bmatrix} A & B(\mathcal{H}X)(t) \\ -\left(B(\mathcal{H}X)(t)\right)^* & 0 \end{bmatrix} \begin{Bmatrix} \tilde{X}(t) \\ \tilde{\mu}(t) \end{Bmatrix}.$$

3.4.1 Approximation of the Estimation Equations

The governing system in Equations 3.19 constitute a distributed parameter system since the functions $\hat{\mu}(t)$ evolve in the infinite dimensional space P. In practice these equations must be approximated by some finite dimensional system. We

define $\tilde{X}_j = \hat{X} - \hat{X}_j$ and $\tilde{\mu}_j = \hat{\mu} - \hat{\mu}_j$ where $\tilde{X}_j$ and $\tilde{\mu}_j$ express approximation errors due to projection of solutions in $\mathbb{R}^m \times P$ to a finite dimensional approximation space. We construct a finite dimensional approximation of the the online estimation equations using the results of Section 3.2.2 and obtain

$$\dot{\hat{X}}_j(t) = A\hat{X}_j(t) + B\left((\mathcal{H}_j X)(t)\Pi_j \circ \hat{\mu}_j(t) + u(t)\right), \tag{3.20}$$

$$\dot{\hat{\mu}}_j(t) = -\left(B(\mathcal{H}_j X)(t)\Pi_j\right)^* X(t). \tag{3.21}$$

Theorem 4 ***Suppose that the history dependent operator $\mathcal{H}$ in Equation 3.17 is defined as in equation 3.7 and 3.8 in terms of a globally Lipschitz, bounded continuous ridge function $\gamma : \mathbb{R} \to \mathbb{R}$ in Equation 3.9. Then for any $T > 0$, we have***

$$\|\hat{X} - \hat{X}_j\|_{C([0,T],\mathbb{R}^m)} \to 0,$$
$$\|\hat{\mu} - \hat{\mu}_j\|_{C([0,T],P)} \to 0,$$

as $j \to \infty$.

Proof 6 ***Define the operators $G(t) : P \to \mathbb{R}^m$ and $G_j(t) : P \to \mathbb{R}^m$ for each $t \geq 0$ as***

$$G(t) := B(\mathcal{H}X)(t),$$
$$G_j(t) := B(\mathcal{H}_j X)(t)\Pi_j.$$

The time derivative of the error in approximation can be expanded as follows:

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\left((\tilde{X}_j, \tilde{X}_j)_{\mathbb{R}^m} + (\tilde{\mu}_j, \tilde{\mu}_j)_P\right) &= (\dot{\tilde{X}}_j, \tilde{X}_j)_{\mathbb{R}^m} + (\dot{\tilde{\mu}}_j, \tilde{\mu}_j)_P \\
&= (A\tilde{X}_j + G\hat{\mu} - G_j\hat{\mu}_j, \tilde{X}_j)_{\mathbb{R}^m} + (-(G - G_j)^* X, \tilde{\mu}_j)_P \\
&= (A\tilde{X}_j, \tilde{X}_j)_{\mathbb{R}^m} + \left((G - G_j)\hat{\mu}, \tilde{X}_j\right)_{\mathbb{R}^m} + \left(G_j(\hat{\mu} - \hat{\mu}_j), \tilde{X}_j\right)_{\mathbb{R}^m} - ((G - G_j)\tilde{\mu}_j, X)_{\mathbb{R}^m} \\
&\leq c(\tilde{X}_j, \tilde{X}_j)_{\mathbb{R}^m} + \|(G - G_j)\hat{\mu}\|_{\mathbb{R}^m}\|\tilde{X}_j\|_{\mathbb{R}^m} + \\
&\qquad \|G_j\|_{\mathcal{L}(P,\mathbb{R}^m)}\|\tilde{\mu}_j\|_P\|\tilde{X}_j\|_{\mathbb{R}^m} + \\
&\qquad \|G - G_j\|_{\mathcal{L}(P,\mathbb{R}^m)}\|\tilde{\mu}_j\|_P\|X\|_{\mathbb{R}^m}.
\end{aligned}$$

We will next use a common inequality that can be derived from two applications of the triangle inequality. We have

$$(a + b, a + b) = (a, a) + 2(a, b) + (b, b) \geq 0,$$
$$(a - b, a - b) = (a, a) - 2(a, b) + (b, b) \geq 0.$$

We conclude from this pair of inequalities that

$$|(a, b)| \leq \frac{1}{2}\left(\|a\|^2 + \|b\|^2\right).$$

The specific form that we apply this theorem is written as

$$|(a,b)| = |(\sqrt{\epsilon}a, \frac{1}{\sqrt{\epsilon}}b)| \leq \epsilon\frac{\|a\|^2}{2} + \frac{1}{\epsilon}\frac{\|b\|^2}{2}. \tag{3.22}$$

We apply the inequality in Equation 3.22 to each term in which $\tilde{\mu}_j$ and $\tilde{X}_j$ appear in a product.

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\left(\|\tilde{X}_j\|^2_{\mathbb{R}^m} + \|\tilde{\mu}_j\|^2_{P)}\right) \leq c\|\tilde{X}_j\|^2_{\mathbb{R}^m} &+ \frac{1}{2a}\|(G-G_j)\hat{\mu}\|^2_{\mathbb{R}^m} + \frac{a}{2}\|\tilde{X}_j\|^2_{\mathbb{R}^m} \\ &+ \frac{1}{2b}\|G_j\tilde{\mu}_j\|^2_{\mathbb{R}^m} + \frac{b}{2}\|\tilde{X}_j\|^2_{\mathbb{R}^m} + \frac{1}{2c}\|\tilde{\mu}_j\|^2_P \\ &+ \frac{c}{2}\|G-G_j\|^2_{\mathcal{L}(P,\mathbb{R}^m)}\|X\|^2_{\mathbb{R}^m}.\end{aligned}$$

Then

$$\begin{aligned}\frac{d}{dt}\left(\|\tilde{X}_j\|^2_{\mathbb{R}^m} + \|\tilde{\mu}_j\|^2_P\right) \leq c\|G-G_j\|^2_{\mathcal{L}(P,\mathbb{R}^m)}\|X\|^2_{\mathbb{R}^m} &+ (2c+a+b)\|\tilde{X}_j\|^2_{\mathbb{R}^m} \\ &+ \left(\frac{1}{c} + \frac{1}{b}\|G_j^*G_j\|\right)\|\tilde{\mu}_j\|^2_P + \frac{1}{a}\|G-G_j\|^2_{\mathcal{L}(P,\mathbb{R}^m)}\|\hat{\mu}\|^2_P.\end{aligned}$$

We integrate this inequality in time from 0 to t to obtain

$$\begin{aligned}\|\tilde{X}_j(t)\|^2_{\mathbb{R}^m} + \|\tilde{\mu}_j(t)\|^2_P \leq \|\tilde{X}_j(0)\|^2_{\mathbb{R}^m} &+ \|\tilde{\mu}_j(0)\|^2_P \\ &+ \int_0^t c\|G(s)-G_j(s)\|^2_{\mathcal{L}(P,\mathbb{R}^m)}\|X(s)\|^2_{\mathbb{R}^m}ds \\ &+ \int_0^t \left\{(2c+a+b)\|\tilde{X}(s)\|^2_{\mathbb{R}^m} + \left(\frac{1}{c} + \frac{1}{b}\|G_j^*(s)G_j(s)\|\right)\|\tilde{\mu}_j\|^2_P\right\}ds \\ &+ \int_0^t \frac{1}{a}\|G(s)-G_j(s)\|^2_{\mathcal{L}(P,\mathbb{R}^m)}\|\hat{\mu}\|^2_P ds.\end{aligned}$$

Choose $a, b > 0$ large enough so that $(2c+a+b) > 0$ and set $\gamma > 1$. If we define

$$\begin{aligned}\gamma &:= \max\left(2c+a+b, \frac{1}{c} + \frac{\eta}{b}\sup_{s\in[0,T]}\|G^*(s)G(s)\|, 1\right), \\ \lambda_j(t) &:= \|(I-\Pi_j)\hat{\mu}(0)\|_P + \int_0^t \|G(s)-G_j(s)\|^2_{\mathcal{L}(P,\mathbb{R}^m)}\left(c\|X\|^2_{\mathbb{R}^m} + \frac{1}{a}\|\hat{\mu}\|^2_P\right)ds,\end{aligned}$$

then the inequality can be written as

$$\|\tilde{X}_j(t)\|^2_{\mathbb{R}^m} + \|\tilde{\mu}_j(t)\|^2_P \leq \lambda_j(t) + \gamma\int_0^t \left(\|\tilde{X}_j(s)\|^2_{\mathbb{R}^m} + \|\tilde{\mu}_j(s)\|^2_P\right)ds.$$

Gronwall's Inequality now completes the proof of the theorem (see Appendix C).

We also further investigate $\lambda_j(t)$ to derive the convergence rate for the approximate states and parameters evolving associated with level j resolution. According to the convergence results obtained in Theorem 1 we have $\|G(s) - G_j(s)\|_{\mathcal{L}(P,\mathbb{R}^m)} = \|B(\mathcal{H}X)(t) - B(\mathcal{H}_j X)(t)\Pi_j\| \leq C_2 2^{-(\alpha+1)j}$. Therefore, $\|G(s) - G_j(s)\|^2 \leq C_2^2 2^{-(\alpha+1)2j}$. It then follows that

$$\begin{aligned}\lambda_j(t) &= \|(I-\Pi_j)\hat{\mu}(0)\|_P + \int_0^t 2^{-(\alpha+1)2j}\left(c\|X\|^2_{\mathbb{R}^m} + \frac{1}{a}\|\hat{\mu}\|^2_P\right)\mathrm{d}s \\ &\leq \|(I-\Pi_j)\|\|\hat{\mu}(0)\|_P + 2^{-(\alpha+1)2j}\left(c\|X\|^2_{\mathbb{R}^m} + \frac{1}{a}\|\hat{\mu}\|^2_P\right)t.\end{aligned}$$

If $t \simeq C_3 2^{(\alpha+1)j}$, then $\lambda_j(t) < \mathcal{O}(2^{-(\alpha+1)j})$ for $t \in [0, C_3 2^{(\alpha+1)j}]$.

3.5 Adaptive Control Synthesis

In order to estimate the function μ that weighs the contribution of history dependent kernels to the equations of motion, we first map it to an n-dimensional subspace of square integrable functions using a projection operator $\Pi^n : P \mapsto P^n$. Let

$$\dot{X} = AX + B((\mathcal{H}X)\circ(\mu - \hat{\mu}) + v) \tag{3.23}$$

be the governing equation of a robotic system after applying a feedback linearization control signal as mentioned in Equation 3.3 with $u = v - (\mathcal{H}X)\circ\hat{\mu}$. We substitute $\mu = \Pi^n\mu + (I - \Pi^n)\mu$ and write

$$\dot{X} = AX + B((\mathcal{H}X)\circ(\Pi^n\mu - \hat{\mu}) + v) + B((\mathcal{H}X)\circ(I - \Pi^n)\mu). \tag{3.24}$$

Finally, by replacing $d = \{(\mathcal{H}X)(I - \Pi^n)\circ\mu\}$ we obtain

$$\dot{X} = AX + B((\mathcal{H}X)\circ(\Pi^n\tilde{\mu}) + v + d), \tag{3.25}$$

where

$$\dot{\tilde{\mu}} = -((\mathcal{H}X)\Pi^n)^* B^T P X. \tag{3.26}$$

Theorem 5 ***Suppose the state equations have the form of Equation 3.3 and the matrix $\mathcal{P}$ is a symmetric positive definite solution of the Lyapunov equation $A^T\mathcal{P} + \mathcal{P}A = -Q$ where $Q > 0$. Then by employing the update law $\dot{\tilde{\mu}} = -((\mathcal{H}X)\Pi^n)^* B^T\mathcal{P}X$, the control signal***

$$v(t) = \begin{cases} -k\frac{B^T\mathcal{P}X}{\|B^T\mathcal{P}X\|}, & \textit{if } \|B^T\mathcal{P}X\| \geq \epsilon \\ -\frac{k}{\epsilon}B^T\mathcal{P}X, & \textit{if } \|B^T\mathcal{P}X\| < \epsilon \end{cases} \tag{3.27}$$

with $k > \|d\|$ drives the tracking error dynamics of the closed loop system is uniformly ultimately bounded and its norm is eventually $O(\epsilon)$.

Proof 7 ***We choose the Lyapunov function***

$$V = \frac{1}{2}X^T\mathcal{P}X + \frac{1}{2}\left(\tilde{\mu},\tilde{\mu}\right)_P \tag{3.28}$$

where $\mathcal{P}$ is the solution of the Lyapunov equation $A^T\mathcal{P}+\mathcal{P}A=-Q$. The derivative of the Lyapunov function V along the closed loop system trajectory is

$$\begin{aligned}
\dot{V} &= \frac{1}{2}(\dot{X}^T\mathcal{P}X + X^T\mathcal{P}\dot{X}) + \left(\dot{\tilde{\mu}},\tilde{\mu}\right)_p \\
&= \frac{1}{2}\big(AX + B((\mathcal{H}X)\circ(\Pi^n\tilde{\mu}) + v + d)\big)^T\mathcal{P}X + X^T\mathcal{P}(AX + B((\mathcal{H}X)\circ(\Pi^n\tilde{\mu}) + v + d)\big) + \left(\dot{\tilde{\mu}},\mu\right)_P \\
&= \frac{1}{2}X^T(A^T\mathcal{P}+\mathcal{P}A)X + X^TPB(v+d) + X^T\mathcal{P}B\big((\mathcal{H}X)\circ(\Pi^n\tilde{\mu})\big) + \left(\dot{\tilde{\mu}},\mu\right)_P \\
&= -\frac{1}{2}X^TQX + X^T\mathcal{P}B(v+d) + \left(\dot{\tilde{\mu}} + ((\mathcal{H}X)\Pi^n)^*B^T\mathcal{P}X,\tilde{\mu}\right)_P \\
&= -\frac{1}{2}X^TQX + X^T\mathcal{P}B(v+d).
\end{aligned}$$

Therefore we have

$$\begin{aligned}
\dot{V} &\leq -\frac{1}{2}X^TQX + X^T\mathcal{P}B(v+d), \\
&\leq -\frac{1}{2}X^TQX + \begin{cases} X^T\mathcal{P}B\left(-k\frac{B^T\mathcal{P}X}{\|B^T\mathcal{P}X\|} + d\right) & \textbf{\textit{if}}\ \|B^T\mathcal{P}X\| \geq \epsilon \\ X^T\mathcal{P}B\left(-\frac{k}{\epsilon}B^T\mathcal{P}X + d\right) & \textbf{\textit{if}}\ \|B^T\mathcal{P}X\| \leq \epsilon \end{cases}, \\
&\leq -\frac{1}{2}X^TQX + \begin{cases} -\left(k - \|d\|\right)\|B^T\mathcal{P}X\| & \textbf{\textit{if}}\ \|B^T\mathcal{P}X\| \geq \epsilon \\ \epsilon k & \textbf{\textit{if}}\ \|B^T\mathcal{P}X\| \leq \epsilon \end{cases} \\
&\leq -\frac{1}{2}X^TQX + \epsilon k.
\end{aligned}$$

By Theorem 4.18 in [17] we conclude that there is a $\bar{T}>0$ and $\tau>0$ such that $\|X(t)\| \leq \bar{C}\epsilon$ for all $t \geq \bar{T}$.

3.6 Numerical Simulations

Our principle approximation result, the proposed online identification, and adaptive control of systems with history dependent forces are verified in this section. In the first experiment, we validate the operator approximation error bound presented in Theorem 1. In the second experiment, we model a wind tunnel single wing section with a leading and trailing edge flaps and apply the proposed sliding mode adaptive controller presented in Theorem 4. We illustrate the stability of the closed loop system and convergence of the closed-loop system trajectories to the equilibrium point.

3.6.1 Operator Approximation Error

In this section we consider a collection of numerical experiments to validate the operator approximation rates derived in Theorem 1. In order to show that Equation 3.12 holds, we choose a function $\mu(s)$ over Δ and then calculate $(h_j f)(t) \circ \mu_j$ for different levels of refinement. Since the computation of $(hf)(t) \circ \mu$ exactly is numerically infeasible, we choose $J \gg j$ as the finest level of refinement in our simulation. According to Theorem 1, we have

$$|(h_J f)(t) \circ \mu_J - (hf)(t) \circ \mu| \leq C_J 2^{-(\alpha+1)J},$$

and for $j \ll J$ we see that

$$|(h_j f)(t) \circ \mu_j - (hf)(t) \circ \mu| \leq C_j 2^{-(\alpha+1)j}.$$

Assuming $C = \max\{C_j, C_J\}$ and using the triangle inequality, we obtain

$$\begin{aligned} &|(h_J f)(t) \circ \mu_J - (h_j f)(t) \circ \mu_j| \\ &\qquad \leq C(2^{-(\alpha+1)J} + 2^{-(\alpha+1)j}) \end{aligned} \tag{3.29}$$

Therefore, given the weights μ_J for the finest level of refinement J, we can evaluate $\mu_j = \Pi_j \mu_J$ and numerically verify Equation 3.29. Figure 3.3 shows the

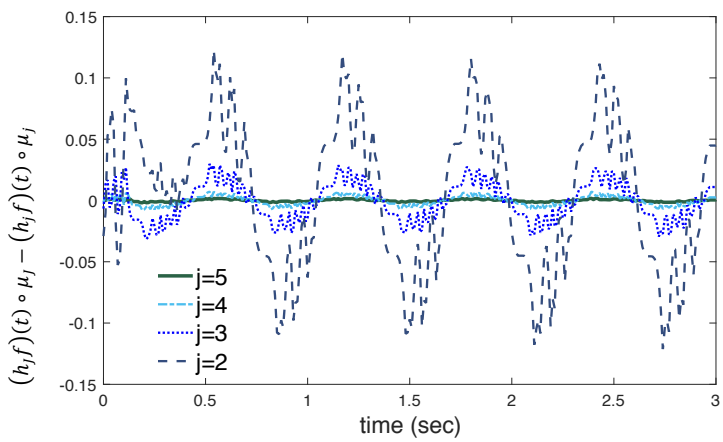


Figure 3.3: Error for different resolution simulations, $J = 7$

simulation results for $J = 7$ and $j = 2, 3, 4, 5$. The error term attenuates with increasing j. In order to investigate the rate of attenuation, we evaluate constant C for different levels of refinements. As shown in figure 3.4, C is approximately constant with respect to j which agrees with the result from Equation 3.29.

3.6.2 Online Identification of History Dependent Aerodynamics and Adaptive Control for a Simple Wing Model

The reformatted governing equations of the system take the form of Equation 3.2 where $Q_a(t, \mu)$ is the vector of generalized history dependent aerodynamic loads.

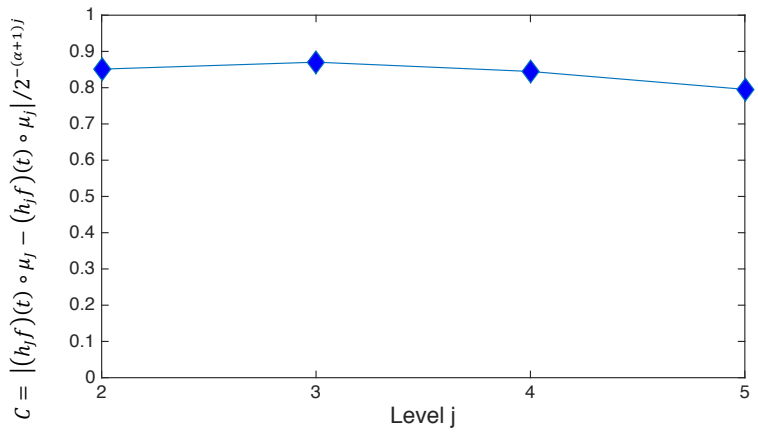


Figure 3.4: C for different level j refinement simulations

The dynamic equation of the system can be written in the form of Equation 3.4, where the history dependent term $M^{-1}(q)Q_a(t,\mu)$ is rewritten in terms of a history dependent operator $(\mathcal{H}X)(t)$ acting on the distributed parameter function μ. The history dependent operator includes a family of fixed history dependent kernels and the distributed parameters μ act as a weighting vector that determines the contribution of a specific history dependent kernel to the overall history dependent operator.

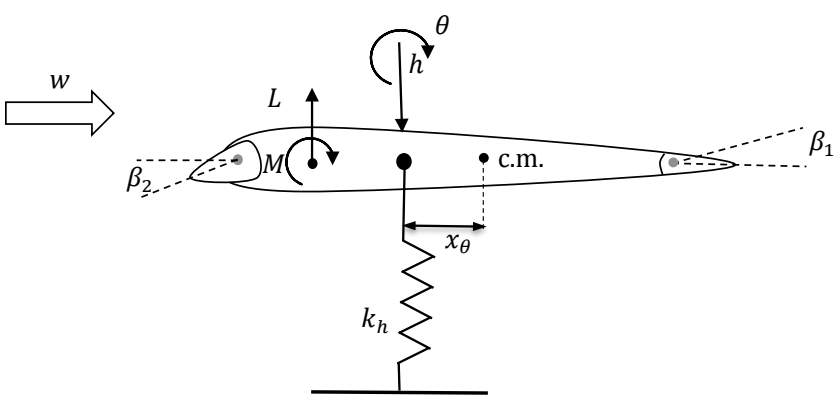


Figure 3.5: Prototypical model for a wing section

We perform an offline identification based on a set of experimental data collected from a wind tunnel experiments or CFDsimulations. These define a nominal model for the history dependent aerodynamic loads that appear in the governing equations of the system. We can exploit the model in the numerical simulations to perform an online estimation of the history dependent aerodynamics and adaptive control of a simple wing model. The details of offline identification of history dependent aerodynamics follow the steps explained in [35].

The model developed in Figure 3.5 is chosen to validate our proposed adaptive sliding mode controller where w is the velocity of wind, k_h is spring constant in plunge, k_θ is a spring constant in pitch, θ is the pitch angle, h is the plunge displacement, c_θ and c_h are viscous damping coefficients, m and I_θ are the mass and moment of inertia and, x_θ is the non-dimensionalized distance between center of mass and the elastic axis. Finally, L and M are lift and moment generated by the leading and trailing edge flaps. The angles β_1 and β_2 define the rotation of the trailing edge and leading edge flaps respectively. The dynamic equations of the wing model is derived in the appendix C as

$$\begin{bmatrix} m & mx_\theta \\ mx_\theta & mx_\theta^2 + I_\theta \end{bmatrix} \begin{Bmatrix} \ddot{h} \\ \ddot{\theta} \end{Bmatrix} + \begin{bmatrix} c_h & 0 \\ 0 & c_\theta \end{bmatrix} \begin{Bmatrix} \dot{h} \\ \dot{\theta} \end{Bmatrix} + \begin{bmatrix} k_h & 0 \\ 0 & k_\theta \end{bmatrix} \begin{Bmatrix} h \\ \theta \end{Bmatrix} = \begin{Bmatrix} L \\ 0 \end{Bmatrix} + \begin{Bmatrix} f_1(\beta_1, \beta_2) \\ f_2(\beta_1, \beta_2) \end{Bmatrix}. \quad (3.30)$$

We have assumed the aerodynamic moment M to be zero and the distance x_a between the aerodynamic center A and hinge point to be negligible to simplify the simulation. The unsteady aerodynamic lift is $L = Q_a(t, \mu)$ where $Q_a(t, \mu) = (\mathcal{H}X) \circ \mu$ reflects the history dependent nature of aerodynamic loads. We rewrite Equation 3.30 to achieve the standard form presented in Equation 3.2.

The adaptive controller presented in Theorem 4 is composed of two parts. The first part compensates for the flutter generated by the history dependent aerodynamic forces through online identification of the aerodynamics. The second part employs an sliding mode controller to compensate for modeling errors.

It is noteworthy that numerical time integration of the evolution equations must accommodate history dependent terms. Since the dynamics of such systems are given via functional differential equations, the ordinary integration rules are not directly applicable. We exploit the predictor-corrector integration rule that has been introduced first in [37]. We also refer the interested reader to our previous paper [39] for details of such integration rules.

Figure 3.6 Shows the simulation results for the case where $\epsilon = 0.01$ and $t_h = 0.001$. The system response eventually enters in a ϵ neighborhood of the sliding manifold. However, as depicted in the figure a chattering behavior occurs in the control signal and system trajectories. We trace this behavior back to the integration error induced by the size of time step. When we increase ϵ or reduce the integration time step, the control signal and system trajectories become smooth. The simulation results for $\epsilon = 0.01$ and $t_h = 0.0005$ are depicted in Figure 3.7. The system trajectories converge to a neighborhood of zero or the set $\mathcal{M}$ in Equation ?? with time and the control signals are relatively smooth. Also, Figure 3.8 shows the case when $\epsilon = 0.1$ and $t_h = 0.001$. The convergence rate of the signals to zero is slower but the results do not show any chattering. Therefore, the proposed smooth sliding mode adaptive controller proves to be effective to identify and compensate for the unknown history dependent aerodynamic forces.

3.7 Results and Conclusion

In this paper, we have derived an explicit bound for the error of approximation for certain history dependent operators that are used in construction of robotic FDE's in [35] and this paper. The numerical simulations presented validate our results. We establish uniform upper bounds on their accuracy of the approximations. The uniform $\mathcal{O}(2^{-(\alpha+1)j})$ rates of approximation for grid resolution j depend on the Holder coefficient α that describes the smoothness of the ridge functions that define the history dependent kernels. In Section 3.3 we prove the existence and uniqueness of a local solution for the special case of functional differential equations with history dependent terms shown in Equation 3.17. Since the functional differential equation of interest evolves in an infinite dimensional space, we construct finite dimensional approximations with grid resolution j. We further show that the solution of the finite dimensional distributed parameter system converges to the solution of the infinite dimensional FDE as the resolution is refined. Finally, we propose an adaptive control strategy to identify and compensate the unknown history dependent dynamics.

Appendix A: Wavelets and Approximation Spaces over the Triangular Domain

We define the multiscaling functions

$$\phi_{j,k}(x) = 1_{\Delta_{i_1,i_2,\ldots,i_j}}(x) \Big/ \sqrt{m(\Delta_{i_1,i_2,\ldots,i_j})}$$

in which

$$1_{\Delta_s}(x) = \begin{cases} 1 & x \in \Delta_s \\ 0 & \textbf{otherwise} \end{cases}$$

and $m(\Delta_{i_1,i_2,\ldots,i_j})$ **is the area of a triangle in the level** j **refinement. We have defined** $(hf)(t) \circ \mu = \iint_\Delta \kappa(s,t,f)\mu(s)ds$. **The approximation** $(h_j f)(t) \circ \mu$ **of this operator is given by**

$$(h_j f)(t) \circ \mu = \iint_\Delta \sum_{l \in \Gamma_j} 1_{\Delta_{j,l}}(s)\kappa(\xi_{j,l}, t, f)\mu(s)ds,$$

where $\xi_{j,l}$ is the quadrature point of number l triangle of grid level j. We approximate $\mu(s) \approx \sum_{m\in\Gamma_j} \mu_{j,m}\phi_{j,m}(s)$. Therefore,

$$\begin{aligned}
&(h_j f)(t) \circ \mu_j \\
&= \iint_S \left(\sum_{l\in\Gamma_j} 1_{\Delta_{j,l}}(s)\kappa(\xi_{j,l}, t, f) \sum_{m\in\Gamma_j} \mu_{j,m}\phi_{j,m}(s) \right) ds \\
&= \sum_{l\in\Gamma_j} \sum_{m\in\Gamma_j} \kappa(\xi_{j,l}, t, f) \left(\iint_S 1_{\Delta_{j,l}}(s)\phi_{j,m}(s) ds \right) \mu_{j,m} \\
&= \sum_{l\in\Gamma_j} \kappa(\xi_{j,l}, t, f)\sqrt{m(\Delta_{j,l})}\mu_{j,l}.
\end{aligned}$$

For an orthonormal basis $\{\phi_k\}_{k=1}^{\infty}$ of the separable Hilbert space P, we define the finite dimensional spaces for constructing approximations as $P_n :=$ span $\{\phi_k\}_{k=1}^{n}$. The approximation error E_n of P_n is given by

$$E_n(f) := \inf_{g\in P_n} \|f - g\|_P.$$

The approximation space $\mathcal{A}_2^\alpha$ of order α is defined as the collection of functions in P such that

$$\mathcal{A}_2^\alpha := \left\{ f \in P \middle| |f|_{\mathcal{A}_2^\alpha} := \left\{ \sum_{n=1}^{\infty} (n^\alpha E_n(f))^2 \frac{1}{n} \right\}^{1/2} < \infty \right\}.$$

For our purposes, the approximation spaces are easy to characterize: they consist of all functions $f \in P$ whose generalized Fourier coefficients decay sufficiently fast. That is, $f \in \mathcal{A}_2^\alpha$ if and only if

$$\sum_{k=1}^{\infty} k^{2\alpha} |(f, \phi_k)|^2 \leq C$$

for some constant C.

Appendix B: The Projection Operator $\Phi_{J\to j}$

The orthogonal projection operator $\Phi_{J\to j} : V_J \to V_j$ maps a distributed parameter μ_J to μ_j i.e. $\Phi_{J\to j} : \mu_J \mapsto \mu_j$. By exploiting the orthogonality property of the operator we have

$$\iint_\Delta \left(\sum_{m\in\Gamma_j} \mu_{j,m}\phi_{j,m}(s) - \sum_{l\in\Gamma_J} \mu_{J,l}\phi_{J,l}(s) \right) \phi_{j,n}(s) ds = 0.$$

Therefore, we can write

$$\sum_{m\in\Gamma_j}\left(\iint_\Delta \phi_{j,m}(s)\phi_{j,n}(s)ds\right)\mu_{j,m}=\sum_{l\in\Gamma_J}\left(\iint_\Delta \phi_{J,l}(s)\phi_{j,n}(s)ds\right)\mu_{J,l}.$$

Since orthogonality implies $\iint_\Delta \phi_{j,m}(s)\phi_{j,n}(s)ds=\delta_{m,n}$, **we conclude that**

$$\mu_{j,n}=\sum_{l\in\Gamma_j}\left(\iint_\Delta \phi_{j,n}(s)\phi_{J,l}(s)ds\right)\mu_{J,l}.$$

From Theorem 1 we have

$$|(h_jf)(t)\circ\Pi_j\mu-(hf)(t)\circ\mu|\le\tilde{C}2^{-\alpha j},$$

with

$$(hf)(t)\circ\mu=\iint_\Delta k(s,t,f)\mu(s)\mathrm{d}s,$$

$$(h_jf)(t)\circ\mu=\iint_\Delta \sum 1_{\Delta_j,l}(s)k(\zeta_{j,l},t,f)\mu(s)\mathrm{d}s,$$

where $\mu\in P=L^2(\Delta)$ **and we approximate** $\mu(s)\approx\sum_{l\in\Gamma_J}\mu_{J,l}\phi_{J,l}(s)\in V_J$**. To implement this for the finest grid** J**, we compute**

$$\begin{aligned}(h_Jf)(t)\circ\mu_J&=(h_jf)(t)\circ\Pi_J\mu_J,\\&=\iint\left(\sum 1_{\Delta_J,l}(s)k(\zeta_{J,l},t,f)\sum_{m\in\Gamma_J}\mu_{J,m}\phi_{J,m}(s)\right)\mathrm{d}s,\\&=\sum_{l\in\Gamma_J}\sum_{m\in\Gamma_J}k(\zeta_{J,l},t,f)\left(\iint 1_{\Delta_J,l}(s)\phi_{J,m}(s)\mathrm{d}s\right)\mu_{J,m},\\&=\sum_{l\in\Gamma_J}\frac{k(\zeta_{J,l},t,f)\mu_{J,l}}{\left(\sqrt{m(\Delta_{J,l})}\right)},\end{aligned}$$

when $\sqrt{m(\Delta_{J,l})}$ **is the area of the corresponding triangle** $\Delta_{J,l}$ **in the grid having resolution level** J**.**

Appendic C: Gronwall's Inequality

We employ the integral form of Gronwall's Inequality to obtain our final convergence result. Many forms of Gronwall's Inequality exist, and we will use a particularly simple version. See Section 3.3.4 in [18]. If the piecewise continuous function f **satisfies the inequality**

$$f(t)\le\alpha(t)+\int_0^t\beta(s)f(s)ds$$

with some piecewise continuous functions α,β **where** α **is nondecreasing, then**

$$f(t)\le\alpha(t)e^{\int_0^t\beta(s)ds}.$$

Appendix D: Modeling of a Prototypical Wing Section

Figure 3.5 shows a simplified model of the wing. In the figure we denote the center of mass by $c.m.$, A is the aerodynamic center, and O is the elastic axis of the wing. The constants K_h and K_θ are the linear and torsional stiffness, and h is the distance from origin to point O in the fixed reference frame. We denote by x_θ the distance between point O and center of mass, whereas x_a is the distance between O and A. Point O is the origin for the body fixed reference frame.

We employ The Euler-Lagrange technique to derive the equation of motion for the depicted wing model. The function $L(\theta, \dot{\theta})$ is the history dependent lift force acting at the aerodynamic center, and $M(\theta, \dot{\theta})$ is the history dependent aerodynamic moment about point A. The variables L_{β_1} and L_{β_2} are the actuating forces acting at point D, and β_1, β_2 are the angles between the midchord of the wing and the trailing edge and leading edge flaps, respectively.

The position vector of the mass center is given as

$$\mathbf{r}_{c.m.} = h\hat{n}_1 - x_\theta \hat{b}_2,$$

and therefore the corresponding velocity of point C is

$$\dot{\mathbf{r}}_{c.m.} = \dot{h}\hat{n}_1 + x_\theta \dot{\theta}\hat{b}_1.$$

The rotation matrix for transformation between inertial frame of reference to body fixed frame of reference is

$$\begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \hat{n}_1 \\ \hat{n}_2 \end{bmatrix}.$$

The kinetic energy is computed to be

$$T = \frac{1}{2} m(\mathbf{r}_{c.m.}.\mathbf{r}_{c.m.}) + \frac{1}{2} I_\theta \dot{\theta}^2,$$

$$T = \frac{1}{2} m(\dot{h}^2 + x_\theta^2 \dot{\theta}^2 + 2x_\theta \dot{h}\dot{\theta}\cos\theta) + \frac{1}{2} I_\theta \dot{\theta}^2,$$

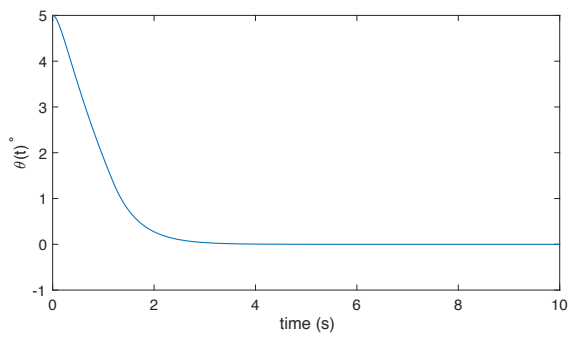
and the corresponding potential energy is

$$V = \frac{1}{2} K_h h^2 + \frac{1}{2} K_\theta \theta^2.$$

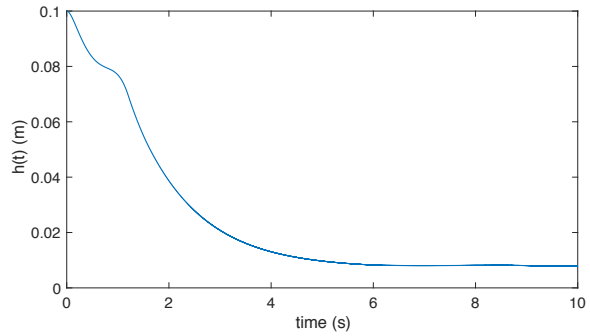
therefore we can write Lagrangian as $L = T - V$. We apply Euler-Lagrange equations to write the equation of motion as follows

$$\begin{bmatrix} m & mx_\theta \cos\theta \\ mx_\theta \cos\theta & mx_\theta^2 + J \end{bmatrix} \begin{bmatrix} \ddot{h} \\ \ddot{\theta} \end{bmatrix} + \begin{bmatrix} 0 & -mx_\theta \dot{\theta}\sin\theta \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{h} \\ \dot{\theta} \end{bmatrix} + \begin{bmatrix} K_h & 0 \\ 0 & K_\theta \end{bmatrix} \begin{bmatrix} h \\ \theta \end{bmatrix}$$
$$= \begin{bmatrix} L(\theta, \dot{\theta})\cos\theta \\ M(\theta, \dot{\theta}) + x_a L(\theta, \dot{\theta}) \end{bmatrix} + \begin{bmatrix} -L_{\beta_1}\cos(\theta + \beta_1) - L_{\beta_2}\cos(\theta + \beta_2) \\ -L_{\beta_1}(e_1 + d_1\cos\beta_1) + L_{\beta_2}(e_2 + d_2\cos\beta_2). \end{bmatrix} \quad (3.31)$$

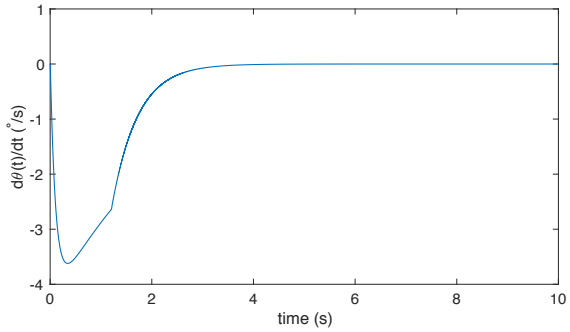
The above equation is written in the form of a standard robotic equations of motion $M(q(t))\ddot{q}(t) + C(q(t), \dot{q}(t))\dot{q}(t) + K(q(t)) = Q_a(t) + \tau(t)$, where $q = [h\ \theta]^T$. We have discussed control applications for such systems in detail in Section ??. In addition, we employ a simplified version of this equation to validate our online identification and adaptive control strategy in Section 3.6.2.



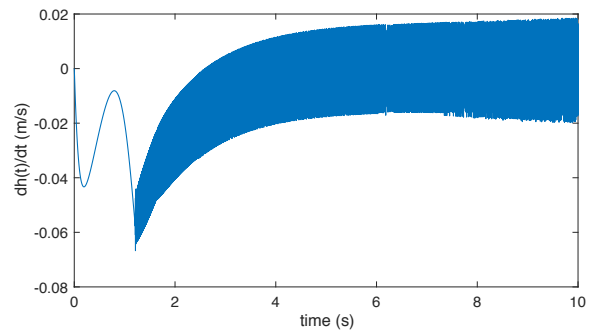
(a) State trajectory, $\theta(t)$



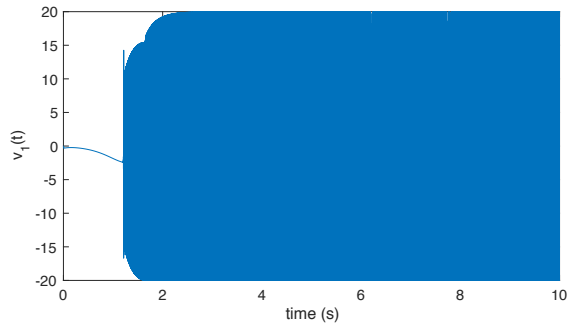
(b) State trajectory, $h(t)$



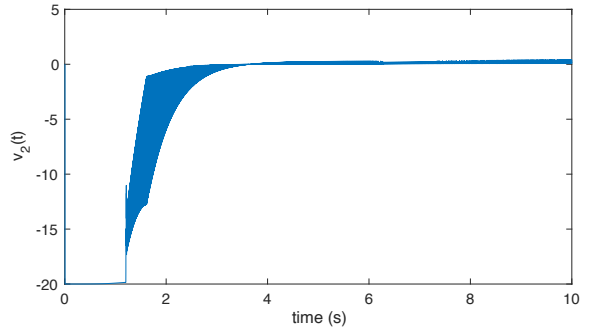
(c) State trajectory, $\dot{\theta}(t)$



(d) State trajectory, $\dot{h}(t)$

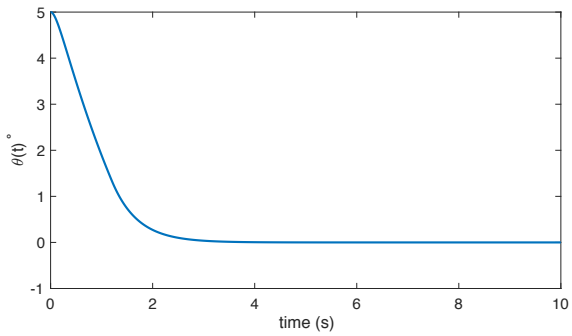


(e) Sliding mode control input signal, $v_1(t)$

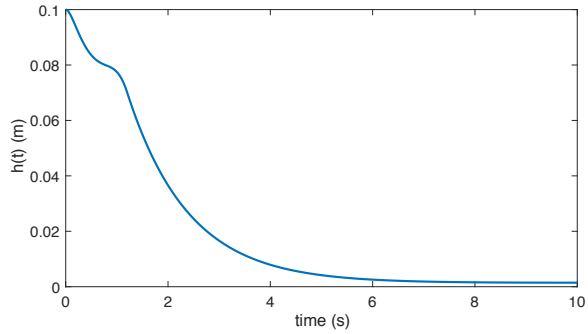


(f) Sliding mode control input signal, $v_2(t)$

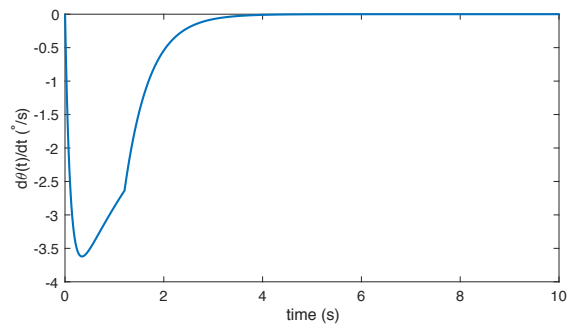
Figure 3.6: Time histories of the states and input signals for $\epsilon = 0.01$, $t_h = 0.001$(sec) and, $k = 20$



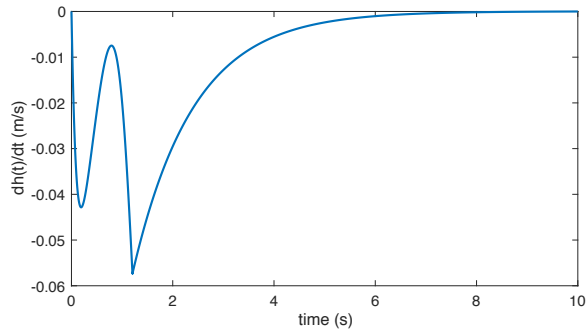
(a) State trajectory, $\theta(t)$



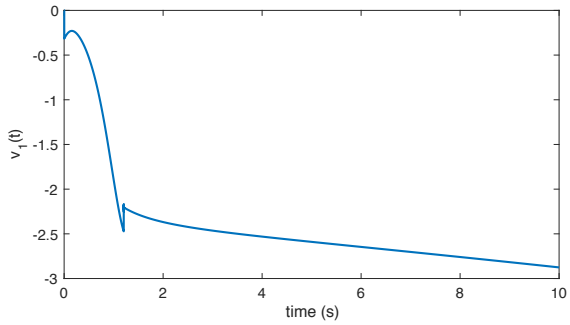
(b) State trajectory, $h(t)$



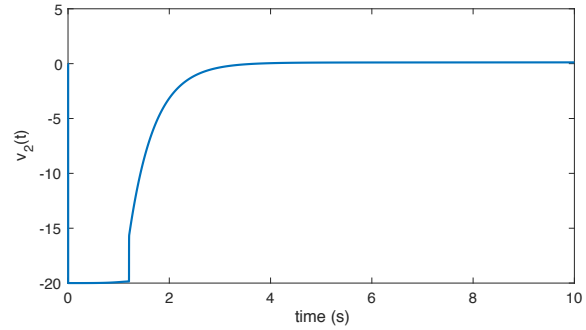
(c) State trajectory, $\dot{\theta}(t)$



(d) State trajectory, $\dot{h}(t)$

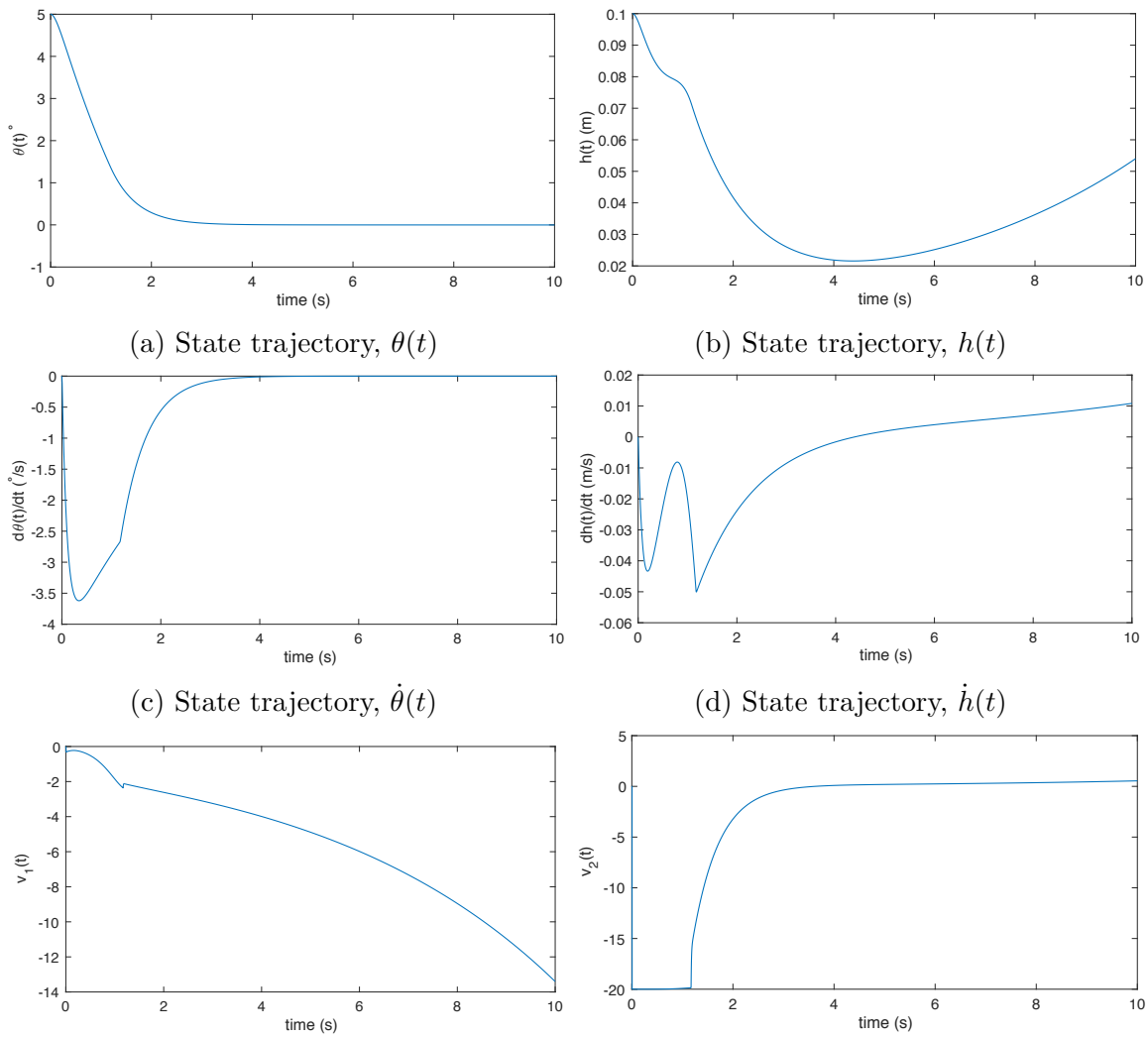


(e) Sliding mode control input signal, $v_1(t)$



(f) Sliding mode control input signal, $v_2(t)$

Figure 3.7: Time histories of the states and input signals for $\epsilon = 0.01$, $t_h = 0.0005$ (sec) and, $k = 20$



(a) State trajectory, $\theta(t)$ (b) State trajectory, $h(t)$

(c) State trajectory, $\dot{\theta}(t)$ (d) State trajectory, $\dot{h}(t)$

(e) Sliding mode control input signal, $v_1(t)$ (f) Sliding mode control input signal, $v_2(t)$

Figure 3.8: Time histories of the states and input signals for $\epsilon = 0.1$, $t_h = 0.001$ (sec) and, $k = 20$

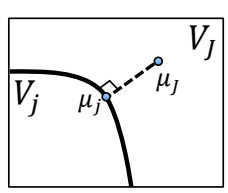


Figure 3.9: Projection Operator $\Phi_{J\to j} : V_J \to V_j$

Bibliography

[1] Joseph W. Bahlman, Sharon M. Swartz1, and Kenneth S. Breuer, *Design and Characterization of a Multi-Articulated Robotic Bat Wing*, *Bioinspiration & Biomimetics*, Vol. 8, No. 1, pp. 1-17, 2013.

[2] H.T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems,* Birkhauser, Boston, 1989.

[3] J. Baumeister, W. Scondo, M.A. Demetriou, and I.G. Rosen, *On-Line Parameter Estimation for Infinite Dimensional Dynamical Systems*, *SIAM J. Control. Optim.*, Vol. 35, No. 2, pp. 678-713, 1997.

[4] Martin Brokate and Jurgen Sprekels, *Hysteresis and Phase Transitions*, Springer-Verlag, 1996.

[5] C. Corduneaunu, *Integral Equations and Applications*, Cambridge University Press, Cambridge, 2008.

[6] Wolfgang Dahmen, Stability of Multiscale Transformations, *J. Fourier Anal. Appl.*, Vol. 4, 1996, 341-362.

[7] M.A. Demetriou, Adaptive Parameter Estimation of Abstract Parabolic and Hyperbolic Distributed Parameter Systems, *Ph.D. thesis,* Departments of Electrical-Systems and Mathematics, University of Southern California, Los Angeles, CA, 1993.

[8] M.A. Demetriou and I.G. Rosen, *Adaptive Identification of Second Order Distributed Parameter Systems, Inverse Problems*, Vol. 10, pp. 261-294, 1994.

[9] M.A. Demetriou and I.G. Rosen, *On the Persistence of Excitation in the Adaptive Identification of Distributed Parameter Systems, IEEE Trans. Automat. Control,* Vol. 39, pp. 1117-1123, 1994.

[10] Ronald A. DeVore, *Nonlinear Approximation*, *Acta Numerica*, Vol. 7, pp. 51-150, January, 1998.

[11] Ronald A. DeVore and George Lorentz,*Constructive Approximation*, Springer-Verlag, 1993.

[12] Rodney D. Driver, Existence and Stability of Solutions of a Delay-Differential System, ***Archive for Rational Mechanics and Analysis***, Vol. 10, No. 1, pp. 401-426, January, 1962.

[13] T.E. Duncan, B. Maslowski, and B. Pasik-Duncan, ***Adaptive Boundary and Point Control of Linear Stochastic Distributed Parameter Systems, SIAM J. Control Optim.***, Vol. 32, pp. 648-672, 1994.

[14] Achim Ilchmann, Hartmut Logemann, and Eugene P. Ryan, ***Tracking with Prescribed Transient Performance for Hysteretic Systems***, ***SIAM J. Control Optim.***, Vol. 48, No. 7, pp. 4731–4752, 2010.

[15] M.A. Krasnoselskii and A.V. Pokrovskii, ***Systems with Hysteresis***, Springer, Berlin, 1989.

[16] A. Ilchmann, E.P. Ryan, C.J. Sangwin, ***Systems of Controlled Functional Differential Equations and Adaptive Tracking***, ***SIAM J. Control. Optim.***, Vol. 40, No. 6, pp. 1746-1764, 2002.

[17] H. K. Khalil, Nonlinear Systems, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2003

[18] Petros Ioannou and Jing Sun, ***Robust Adaptive Control***, Dover, 2012.

[19] Fritz Keinert, ***Wavelets and Multiwavelets***, Chapman & Hall, CRC Press, 2004.

[20] N.N. Krasovskii, ***On the Application of the Second Method of A.M. Lyapunov to Equations with Time Delays***, ***Prikl. Mat. i Mekh.***, Vol. 20, pp. 315-327, 1956.

[21] N.N. Krasovskii, On the Asymptotic Stability of Systems with After-Effect, ***Prikl. Mat. i Mekh.***, 20, 513–518, 1956.

[22] Frank L. Lewis, Darren M. Dawson, and Chaouki T. Abdallah, ***Robot Manipulator Control: Theory and Practice***, Marcel Dekker, Inc., 2004.

[23] A.D. Myshkis, ***General Theory of Differential Equations with a Retarded Argument, Uspekhi Mat. Nauk***, Vol. 4, No. 5(33), pp. 99–141, 1949.

[24] Kumpati S. Narendra and Anuradha M. Annaswamy, ***Stable Adaptive Systems***, Dover, 2005.

[25] V.P. Rudakov, ***Qualitative Theory in a Banach Space, Lyapunov-Krasovskii Functionals, and Generalization of Certain Problems***, ***Ukrainskii Matematicheskii Zhurnal***, Vol. 30, No. 1, pp. 130-133, January-February, 1978.

[26] V.P. Rudakov, ***On Necessary and Sufficient Conditions for the Extendability of Solutions of Functional-Differential Equations of the Retarded Type***, ***Ukr. Mat. Zh.***, Vol. 26, No. 6, pp. 822-827, 1974.

[27] E.P. Ryan and C.J. Sangwin, ***Controlled Functional Differential Equations and Adaptive Tracking***, ***Systems and Control Letters***, Vol. 47, pp. 365–374, 2002.

[28] Shankar Sastry and Marc Bodson, ***Adaptive Control: Stability, Convergence and Robustness***, Dover, 2011.

[29] Bruno Siciliano, Lorenzo Sciavicco, Luigi Villani, and Guiseppe Oriolo, ***Robotics:Modeling, Planning, and Control***, Springer-Verlag, London, 2010.

[30] Mark W. Spong, Seth Hutchinson, and M. Vidyasagar, ***Robot Modeling and Control***, Wiley, 2005

[31] Augusto Visintin, ***Differential Models of Hysteresis***, Springer, 1994.

[32] Kurdila, A.J., Li, J., Strganac, T.W., and Webb, G., ***Nonlinear Control Methodologies for Hysteresis in PZT Actuated On-Blade Elevons***, Journal of Aerospace Engineering, Volume 16, Issue 4, pp. 167-176, October 2003.

[33] K. Viswanath and D. Tafti, ***Effect of Stroke Deviation on Forward Flapping Flight***,AIAA Journal, pp. 145-160, Vol. 51, No. 1, January, 2013.

[34] P. Gopalakrishnan and D. Tafti, ***Effect of Wing Flexibility on Lift and Thrust Production in Flapping Flight***, AIAA Journal, pp. 865-877, Vol. 48, No. 5, May, 2010.

[35] S. Dadashi, J. Feaster, J. Bayandor, F. Battaglia, A. J. Kurdila, ***Identification and adaptive control of history dependent unsteady aerodynamics for a flapping insect wing***, Nonlinear Dyn (2016) 85: 1405.

[36] Tobak, Murray, and Lewis B. Schiff. ***On the formulation of the aerodynamic characteristics in aircraft dynamics.*** National Aeronautics and Space Administration, 1976.

[37] Tavernini L, ***Linear multistep methods for the numerical solution of volterra functional differential equations*** Applicable Analysis 3(1973), 169-185.

[38] Jeonghwan Ko, Andrew J. Kurdila, and Thomas W. Strganac. ***Nonlinear Control of a Prototypical Wing Section with Torsional Nonlinearity***, Journal of Guidance, Control, and Dynamics, Vol. 20, No. 6 (1997), pp. 1181-1189.

[39] Shirin Dadashi, Parag Bobade, Andrew J. Kurdila, ***Error Estimates for Multiwavelet Approximations of a Class of History Dependent Operators*** 2016 IEEE 55th Conference on Decision and Control (CDC)

Chapter 4

Error Estimates for Discrete Lagrangian Mechanics Using Empirical Potentials

Error Estimates for Discrete Lagrangian Mechanics Using Empirical Potentials

Shirin Dadashi, Hunter McClelland, Andrew Kurdila

Abstract

This paper develops a novel method for generating discrete equations of Lagrangian mechanics that are constructed to evolve near subsets or submanifolds of the configuration space in which empirical observations are concentrated. When the discrete observations are distributed by a measure ρ on the configuration space Ω, error bounds for the discrete Lagrangian formulation are derived in terms of the difference between the empirical Lagrangian L_z and its limiting Lagrangian L_ρ. By using techniques from learning theory on reproducing kernel Hilbert spaces (RKHS), we derive a *probabilistic* error bound on the discrete Lagrangian error. This probabilistic framework generalizes a popular method for deriving bounds for discrete evolution laws via variational methods to account for empirical observations. The goal of this formulation is to obtain empirically based discrete Lagrangian formulations that approximately preserve submanifolds that are evident in observations of complex systems. It is noteworthy that the technique described in this paper does not require nor identify explicit representations or formulae, such as analytic expressions of constraints, that characterize an underlying manifold. The performance of the method is illustrated in a simple numerical example in which noisy data distributed around a submanifold in $\mathbb{R}^3$ is used to construct an evolution law that lives near the submanifold and approximately reconstructs the observed dynamics.

4.1 Introduction

Even though general frameworks for gauging the error of numerical methods for evolutionary equations, and in particular ordinary differential equations, have been available for decades, the last few years has seen a rapid proliferation of novel discrete approximation methods that preserve or otherwise respect important qualitative features of the underlying flow. These discrete evolution equations are constructed by formulating laws of discrete mechanics that are closely

associated with the continuous mechanics from which the underlying flow is derived. Depending on the application, the discrete evolution laws are derived so that they approximately preserve energy, angular momentum, linear momentum, or other quantities that are invariant under the flow. These guarantees of fidelity can be particularly important in simulations that are carried out over long time intervals and where the (near) invariance of a particularly relevant quantity is essential. General references such as [11] or [25] summarize the diverse collection of such approaches that have have been derived, and they provide compelling philosophical motivation for the general strategies involved. The literature is replete with numerical case studies in which a conventional numerical integration scheme, one that is not guaranteed *a priori* to conserve invariants over a long time frame, is outperformed by a method of discrete mechanics that has been designed expressly with conservation of such integrals in mind. Since some of the most impressive theoretical guarantees of the performance of approximation methods based on discrete mechanics are evident over long integration times, it is no surprise that many case studies of discrete mechanics consider problems of astronomy or classical mechanics. This paper is motivated by a class of problems that are in several respects qualitatively different from many of the usual types of systems considered in case studies of discrete formulations of mechanics. We have in mind he study of biomechanical models of animal motion.

These biomechanical systems are high dimensional in comparison to many examples in astronomy, celestial mechanics, or classical mechanics as in [11, 25]. Typically biomechanical models may have from $\mathcal{O}(10)-\mathcal{O}(100)$ degress of freedom. Their configuration space is quite similar under some simplifying hypotheses. If we neglect flexibility effects in the wings of a bat, it is possible in principle to model the bat so that its governing equations evolve on the configuration manifold

$$SE(3) \times \underbrace{\left(SO(3) \times S^1 \times \cdots \times S^1\right)^2}_{\text{left and right wings}}, \tag{4.1}$$

where $SE(3)$ represents the motion of the core body, $SO(3)$ represents the motion of the shoulder joint of the wing, and S^1 represents each of the remaining joints in the wing. Similarly, we can construct a model of human motion where the governing equations evolve on a configuration manifold

$$SE(3) \times \underbrace{\left(SO(3) \times S^1 \times \cdots \times S^1\right)^2}_{\text{left and right arms}} \times \underbrace{\left(SO(3) \times S^1 \times \cdots \times S^1\right)^2}_{\text{left and right legs}}, \tag{4.2}$$

where again $SE(3)$ represents the core body motion, $SO(3)$ represents either a shoulder or hip joint, and S^1 represents all the remaining joints in the bioskeletal system. The derivation of the evolution equations on the appropriate manifold for such nonlinear, multibody systems that have the connectivity of a topological tree has been carried out in many references, see for example [28]. In addition to the structural similarity of the configuration manifold, the two biomechanical systems are similar in other respects. Both are subject to external forces that are

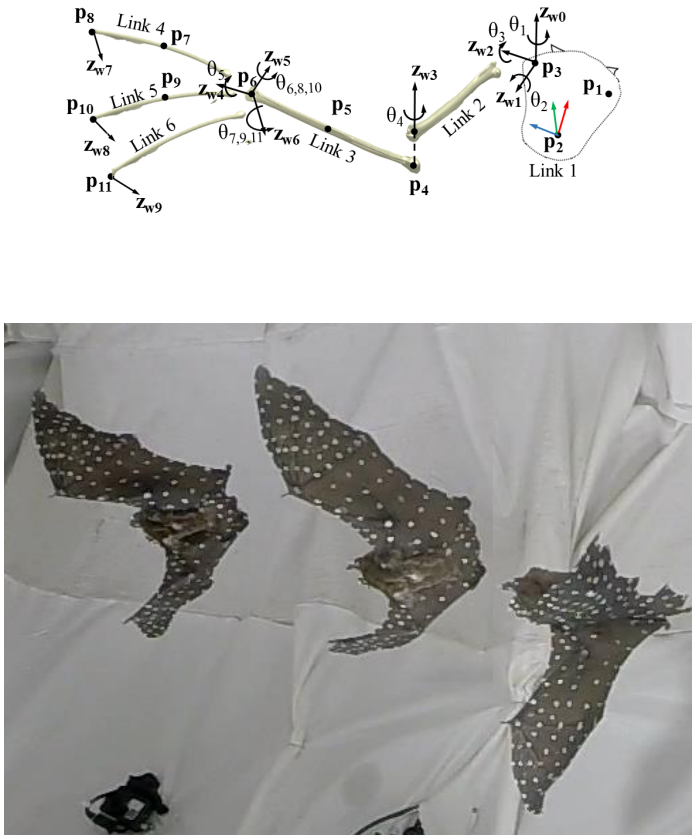


Figure 4.1: above: Body fixed frame definition of bat's wing skeleton, below: bat flight data collection experiment in a wind tunnel [2]

exceptionally difficult to measure or quantify. In the case of human locomotion, attempts to measure, model, and characterize the forces of contact have been studied extensively. [5, 26, 10] The aerodynamics forces that act on the wings of a bat during flight have likewise been the focus of careful study. [16, 17, 37, 31] While real-time measurement of contact forces is difficult, while real-time measurement of aerodynamic forces on bat wings is quite unlikely, if not impossible with existing hardware and technology. It is also worth noting that the *internal forces* delivered by the muscles that drive walking motion in a human or flapping flight in a bat are not now, nor likely will be in the near future, measurable in real-time *in vivo*. In summary then, the biomechanical systems that motivate this paper evolve on configuration manifolds of high dimension and are driven by uncertain or unknown inputs. In fact, the uncertainties are often so substantial that many authors in vision-based tracking, where real-time estimates of motion are generated, resort to the simplest of predictive models for humanoid walking or bat motion in typical filtering problems. Models that just use a random walk for joint variables, with no detailed model of system dynamics, often are used in the prediction steps of Kalman, Unscented Kalman, or particle filters for motion estimation. For example, see Equation 31 of [14] where a linear extrapolation of joint angles is used as the predictive, dynamic model for "conventional methods."

At the same time, a growing interest has evolved in identifying low dimensional dynamics that underly particular motion regimes (e.g. walking, hopping, level-flight) of complex systems. This trend is particularly evident in studies of human walking, and to a lesser degree in studies of bat flapping flight. For example, [15] investigates the choice of priors in methods that exploit manifold learning during vision-based motion estimation and filtering. Reference [13] derives an algorithm that generalizes classical unscented Kalman filtering and considers processes that evolve on manifolds, thereby enabling applications to human motion estimation.

Perhaps the most popular general approach for the generation of nonlinear, reduced dimension representations of human walking regimes employ versions of Gaussian process latent variable models introduced in [?]. In these methods a Gaussian process is used to define the prior information on some unknown function that represents the mapping from a low dimensional space to the the configuration manifold. References [14, 8, 9, 32, 21, 18, 13] are typical of this approach. While nonlinear dimension reduction methods such as manifold approximation do not seem to have been investigated in the context of bat flight as yet, linear order reduction methods such as principal orthogonal decomposition have been studied in [34]. Authors of this paper have observed while conducting the research in [2] that nonlinear reduced order models are also urgently needed for motion estimation of bat flight, just as in motion estimation for human walking. The large amplitude, nonlinear deformations among fiducial markers in successive images of the wings of bats during flight prove a significant challenge to the generation of automated, general, and highly redundant camera imaging for motion estimation of complex systems.

In view of these studies, it is evident that there is a clear need for systematic, general, and accurate reduced dimensional models for the prediction of motion regimes in complex systems. We refer to the general collection presented in this paper as being *empirical-analytical models*. These models use analytic expressions for the kinetic energy and *physically known* potential energy, and empirical models for an experimentally based potential energy that approximately concentrates the system trajectories near a subset or submanifold associated with observations of the system. It should be noted that our primary motivation for this paper is the creation of better models for use in the prediction step of recursive Bayesian estimation techniques as applied to complex multibody nonlinear systems. Our experience in this application area suggests that models that respect the underlying geometry of the problem at hand would be of great value, even given that the force and moments that drive such a system are unknown and inaccessible in real-time with the current generation of hardware and instrumentation.

This paper studies one approach to the generation of such models: a synthesis of certain manifold estimation methods that have been studied in the context of learning theory and discrete Lagrangian formulations of mechanics. In contrast to many conventional applications of discrete Lagrangian mechanics, the approach in this paper is not aimed at preserving some integral invariant per se, although the derived methods can indeed be constructed to exhibit such invariance. Rather, we seek to construct a discrete Lagrangian formulation of mechanics that incorporates empirical observations of motion from some underlying set or manifold to obtain a discrete evolution law that is guaranteed to evolve on or near the underlying set. In this way our goal is to formulate discrete Lagrangian equations that, at least approximately, "respect the underlying manifold or domain."

The approach studied in this paper derives a discrete Lagrangian evolution equation in terms of an empirical potential function that is constructed via techniques in manifold learning. The empirical potential has the effect of "shaping" the evolution that occurs under the discrete law so that the state evolves in proximity to the empirical observations. The result of the analysis is a Lagrangian discrete evolution law that depends on the approximation order of the finite dimensional subspaces used in the variational framework and on the number of samples used in the construction of the empirical potential function. The culmination of this paper is a rigorous error estimate of the order of approximation of the discrete Lagrangian formulation. In contrast to usual approaches in discrete Lagrangian mechanics, and analogous to approaches in statistical learning theory, the error in a single step of the discrete Lagrangian depends on a balance of deterministic and a probabalistic terms.

Our approach is ideologically similar to potential field methods, popularized via robotic navigation applications. In these methods, physical obstacles are replaced with soft repelling-fields in order to guide robots gently away from

collisions. The origin of these methods is credited to works summarized in [20] as well as to [27] and [23]. As theses theories have matured, their weaknesses were well studied [22] and provably correct extensions were suggested. [33] This direction of development is analogous to that suggested in this paper, as we prove convergence of the dynamic trajectory under "soft" empirical potential function converges to the manifold on which the underlying dynamics evolve.

The remainder of this paper is organized as follows. In Section 4.2 an overview of the relevant theory of discrete Lagrangians is presented. Section 4.2.1 focuses discrete mechanics for uncontrained, conservative systems. Section 4.2.3 introduces techniques adapted to the generation of empirical-analytical methods. Section 4.3 provides an overview of the relevant theory used in this paper from learning theory and approximation. Section 4.4 presents an analysis for discrete mechanics that culminates in terms of an error in probability for our class of variational integrators. Section 4.5 summarizes an open problem related to and motivated by this paper. Numerical simulations that illustrate the qualitative features of the derived approach are presented in Section 4.6.

4.2 Discrete Lagrangian Mechanics

Discrete Lagrangian and Hamiltonian mechanics has been studied in great detail over the past few decades, often in attempts to derive discrete temporal dynamics that preserve or respect properties of an underlying flow in continuous time. Good overviews of this topic can be found in [11, 25], while specific results on spectral rates of convergence that are relevant to this paper are given in [12].

4.2.1 Unconstrained, Conservative Systems

The classical statement of Hamilton's principle in continuous time for unconstrained, conservative systems states that of all the admissible motions of a mechanical system, the true motion renders the action functional $\mathcal{A}(q) := \int L(q(\tau), \dot{q}(\tau))d\tau$ stationary where $L : TQ \to \mathbb{R}$ is the Lagrangian that is defined on the tangent bundle TQ of the configuration manifold Q . As a problem of variational calculus, this principle requires that

$$\delta \int_{t_0}^{t_f} L(q(\tau), \dot{q}(\tau))d\tau = 0 \tag{4.3}$$

for all admissible variations of the true motion that satisfy $\delta q(t_0) = \delta q(t_f) = 0$. A host of problems of mechanics have been solved via this principle, often by establishing that motions that render the action integral stationary satisfy Lagrange's equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = 0 \tag{4.4}$$

subject to the conditions $q(t_0) = q_0$ and $q(t_f) = q_f$. One common framework for constructing discrete approximations of the such variational problems introduces the discrete Lagrangian $L_d : Q \times Q \to \mathbb{R}$ via

$$L_d(q_0, q_1; h) := \underset{\substack{q \in C^2([0,h];Q) \\ q(t_0)=q_0, q(t_1)=q_1}}{\mathbf{ext}} \int_{t_0}^{t_1} L(q(\tau), \dot{q}(\tau)) d\tau. \tag{4.5}$$

The original action functional $\mathcal{A}(q)$ is then approximated by the discrete action function $\mathcal{A}_d(\{q_k\}_{k=0}^n)$ where

$$\mathcal{A}_d(\{q_k\}_{k=0}^n) := \sum_{k=0}^{n-1} L_d(q_k, q_{k+1}; h) \approx \int_{t_0}^{t_f} L(q(\tau), \dot{q}(\tau)) d\tau, \tag{4.6}$$

Discrete evolution laws

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = 0 \tag{4.7}$$

for $k = 1, \ldots, n-1$ are derived by invoking stationarity of the discrete action function, and these discrete equations can be regarded as the discrete analogs of Lagrange's Equations 4.4. The operators $D_1(\cdot)$ and $D_2(\cdot)$ are simply the partial derivatives with respect to the first and second arguments of $L_d(\cdot, \cdot)$. If q_{k-1} and q_k are known, then the discrete law in Equation 4.7 is a nonlinear equation for the next state q_{k+1}. Providing this set of nonlinear equations admit solutions and can be solved, the discrete evolution equations determine a forward flow map $F_{L_d}^+ : Q \times Q \to Q \times Q$ where we have

$$F_{L_d} := F_{L_d}^+ := \begin{Bmatrix} q_{k-1} \\ q_k \end{Bmatrix} \mapsto \begin{Bmatrix} q_k \\ q_{k+1} \end{Bmatrix}. \tag{4.8}$$

4.2.2 Constrained, Conservative Systems

Many approaches have been pursued that modify Hamilton's principle, and thereby induce associated changes to Lagrange's equations, for mechanical systems subject to various types of constraints. The reader is referred to [4] or [3] for a detailed discussion of the case of evolution laws and variational calculus problems in continuous time. We will only consider systems that are subject to holonomic constraints in this paper. For such systems, Lagrange's equations take the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) - \frac{\partial L}{\partial q} = \langle \lambda, \nabla \phi \rangle \tag{4.9}$$

$$\phi(q) = 0 \tag{4.10}$$

Correspondingly, [25] derives a discrete Lagrangian approximation to these equations in the form

$$D_2 L_d(q_{k-1}, q_k) + D_1 L_d(q_k, q_{k+1}) = \langle \lambda_k, \nabla(\phi(q_k)) \rangle \tag{4.11}$$

$$\phi(q_k) = 0 \tag{4.12}$$

4.2.3 Empirical-Analytical Discrete Lagrangian Mechanics

The primary objective of this paper is the study of a class of discrete evolution laws which are derived from Lagrangians that depend on some empirical observations of a system under study. Specifically, we suppose that we must construct a model of a complex system whose full state q evolves in some compact subset $\Omega \subseteq Q$ of the configuration space $Q \equiv \mathbb{R}^d$. We make N observations $z = \left\{z_1, z_2, \cdots, z_N\right\}^T \in \Omega \times \cdots \times \Omega = \Omega^N \subseteq Q^N$ of the state q as the system evolves. We assume that the measurement process generates samples $z \in Q^N$ that are independent and identically distributed, each measurement having distribution defined in terms of a common probability measure ρ defined on $\Omega \subseteq Q$. The probability measure ρ is unknown. We construct an empirically generated Lagrangian $L_z := L_z(q, \dot{q}) = T(q, \dot{q}) - \mathcal{V}_z(q)$ from the set of observations $z := \left\{z_1, \cdots, z_N\right\}$, and subsequently seek to construct a discrete evolution law in terms of the empirical Lagrangian L_z. The asymptotic behavior of the discrete evolution law will be analyzed in terms of a limiting Lagrangian $L_\rho := L_\rho(q, \dot{q}) = T(q, \dot{q}) - \mathcal{V}_\rho(q)$ that depends on the unknown distribution ρ. That is, we will discuss in what sense the convergence $L_z \to L_\rho$ gurantees convergence of the discrete evolution law generated by L_z to that generated by L_ρ. In this paper we choose $\mathcal{V}_z = -\frac{1}{\epsilon^2} V_z$ and $\mathcal{V}_\rho = -\frac{1}{\epsilon^2} V_\rho$ where V_z and V_ρ are obtained using learning theory as discussed in Section 4.3 and ϵ is a small parameter.

The discrete evolution law will be obtained using the discrete empirical action functional $\mathcal{A}_{d,z}$

$$\mathcal{A}_{d,z}(q) := \int_0^h L_z(q(\tau), \dot{q}(\tau)) d\tau \tag{4.13}$$

and the m-point quadrature operator $\mathcal{Q}_m$

$$\mathcal{Q}_m \left(\int_0^h f(\tau) d\tau \right) = \sum_{j=1}^m w_j f(t_j) \tag{4.14}$$

where $\left\{(t_j, w_j)\right\}_{j=1}^m$ are the collection of quadrature points and weights. It is assumed that the quadrature operator is defined so that the quadrature error estimate

$$\left| \mathcal{Q}_m \left(\int_0^h f(\tau) d\tau \right) - \int_0^h f(\tau) d\tau \right| \leq C_q h^{n+1} \tag{4.15}$$

holds for a constant $C_q > 0$ and for all integrands f that are sufficiently smooth. Note that this is a standard assumption in the construction of variational integration rules [12]. We define the discrete empirical Lagrangian function as

$$L_{d,(m,n,z)}(q_0, q_1, h) := \underset{\substack{q \in \mathbb{M}^n([0,h];Q) \\ q(t_0)=q_0, q(t_1)=q_1}}{\textbf{ext}} \mathcal{Q}_m \mathcal{A}_{d,z}(q) := \underset{\substack{q \in \mathbb{M}^n([0,h];Q) \\ q(t_0)=q_0, q(t_1)=q_1}}{\textbf{ext}} \sum_{j=1}^m w_j L_z(q(t_j), \dot{q}(t_j)), \tag{4.16}$$

and a typical extremizer is denoted

$$q_{m,n,z} := \underset{\substack{q\in\mathbb{M}^n([0,h];Q)\\ q(t_0)=q_0, q(t_1)=q_1}}{\arg\operatorname{ext}} \mathcal{Q}_m\mathcal{A}_{d,z}(q) \tag{4.17}$$

with $\mathbb{M}^n([0,h];Q)$ an n dimensional subspace of $C^2([0,h];Q)$. Note that the extremizer $q_{m,n,z}$ depends on the number of quadrature points m, the dimension of $\mathbb{M}^n([0,h];Q)$, and empirical sample z. Just as for the empirical discrete action $\mathcal{A}_{d,z}$ and empirical discrete Lagrangian $L_{d,(m,n,z)}$, we also define the corresponding "exact discrete" quantities associated with the limiting Lagrangian L_ρ. The exact discrete action $\mathcal{A}_{d,\rho}$ and exact discrete Lagrangian $L^E_{d,\rho}$ are defined as

$$L^E_{d,\rho}(q_0,q_1,h) := \underset{\substack{q\in C^2([0,h];Q)\\ q(t_0)=q_0, q(t_1)=q_1}}{\operatorname{ext}} \mathcal{A}_{d,\rho}(q) := \underset{\substack{q\in C^2([0,h];Q)\\ q(t_0)=q_0, q(t_1)=q_1}}{\operatorname{ext}} \int_0^h L_\rho(q(\tau),\dot q(\tau))d\tau,$$

where a typical extremizer is denoted

$$q^E_{d,\rho} := \underset{\substack{q\in C^2([0,h];Q)\\ q(t_0)=q_0, q(t_1)=q_1}}{\arg\operatorname{ext}} \mathcal{A}_{d,\rho}(q). \tag{4.18}$$

Our first estimate generalizes Theorem 3.3 of [12].

Theorem 6 ***Under the hypotheses of Theorem 3.3 of [12], we have the single step error bound***

$$\left|L_{d,(m,n,z)} - L^E_{d,\rho}\right| \le (2\ell C_a + C_q)h^{n+1} + h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)}. \tag{4.19}$$

Proof 8 ***We assume the finite dimensional subspaces $\mathbb{M}^n([0,h],Q)$ are constructed and equipped with approximation operators***

$$\Pi^n : C^2([0,h];Q) \to \mathbb{M}^n([0,h];Q)$$

that are guaranteed to satisfy the approximation error bounds

$$\|\Pi^n q - q\|_{C([0,h];Q)} + \left\|\frac{d}{dt}(\Pi^n q) - \dot q\right\|_{C([0,h];Q)} \le C_a h^n \tag{4.20}$$

for all $q \in C^2([0,h];Q)$ for some positive constant $C_a > 0$. Recall that the hypotheses in Theorem 3.3 of [12] include the assumption that there exists minimizers $q^E_{d,\rho}$ and $q_{m,n,z}$ of Equations 4.17 and 4.18. The desired estimate is obtained by decomposing the difference into two parts

$$\begin{aligned}\left|L^E_{d,\rho} - L_{d,(m,n,z)}\right| &= \left|\mathcal{A}_{d,\rho}(q^E_{d,\rho}) - \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})\right| \\ &\le \underbrace{\left|\mathcal{A}_{d,\rho}(q^E_{d,\rho}) - \mathcal{A}_{d,\rho}(\Pi^n q^E_{d,\rho})\right|}_{\textit{term 1}} + \underbrace{\left|\mathcal{A}_{d,\rho}(\Pi^n q^E_{d,\rho}) - \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})\right|}_{\textit{term 2}}.\end{aligned} \tag{4.21}$$

The first term on the right

$$\left|\mathcal{A}_{d,\rho}(q^E_{d,\rho}) - \mathcal{A}_{d,\rho}(\Pi^n q^E_{d,\rho})\right| \leq \int_0^h \left| L_\rho\left(q^E_{d,\rho}(\tau), \dot{q}^E_{d,\rho}(\tau)\right) - L_\rho(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{d\tau}\left(\Pi^n q^E_{d,\rho}\right)(\tau))\right| d\tau$$

is bounded using the Lipschitz continuity of the Lagrangian L_ρ. The Lipschitz continuity of the Lagrangian L_ρ (see Appendix 4.8.1), the approximation properties of the spaces $\mathbb{M}^n([0,h];Q)$, and the definition of the approximation operators Π^n enables us to write

$$\begin{aligned}
|\mathcal{A}_{d,\rho}(q^E_{d,\rho}) - \mathcal{A}_{d,\rho}(\Pi^n q^E_{d,\rho})| &\leq \int_0^h \left| L_\rho(q^E_{d,\rho}(\tau), \dot{q}^E_{d,\rho}(\tau) - L_\rho(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{d\tau}\left(\Pi^n q^E_{d,\rho}(\tau)\right)\right| d\tau, \\
&\leq \int_0^h \ell \left\{ \left\| q^E_{d,\rho}(\tau) - \Pi^n q^E(\tau)\right\|_{\ell^1(R^d)} + \left\| \dot{q}^E_{d,\rho}(\tau) - \frac{d}{d\tau}\left(\Pi^n q^E_{d,\rho}(\tau)\right)\right\|_{\ell^1(R^d)} \right\} d\tau, \\
&\leq h\ell \left\{ \left\| q^E_{d,\rho} - \Pi^n q^E_{d,\rho}\right\|_{C([0,h];Q)} + \left\| \dot{q}^E_{d,\rho} - \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)\right\|_{C([0,h];Q)} \right\}, \\
&\leq \ell C_a h^{n+1}.
\end{aligned}$$

We next turn to bounding the second term in Equation 4.21. Recall that since $q_{m,n,z}$ is assumed to be a minimizer, we know that $L_{d,(m,n,z)} = \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) \leq \mathcal{Q}_m\mathcal{A}_{d,z}(q)$ for any $q \in \mathbb{M}^n([0,h];Q)$. Using this observation, and assuming that the empirical potential L_z is smooth enough that the quadrature is accurate to order $\mathcal{O}(h^{n+1})$, we find that

$$\begin{aligned}
\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) &= \sum_{j=1}^m w_j L_z(q_{m,n,z}(t_j), \dot{q}_{m,n,z}(t_j) \leq \sum_{j=1}^m w_j L_z\left(\Pi^n q^E_{d,\rho}(t_j), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(t_j)\right), \\
&\leq \int_0^h L_z\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau + C_q h^{n+1} = \mathcal{A}_{d,z}(\Pi^n q^E_{d,\rho}) + C_q h^{n+1}.
\end{aligned}$$

An upper bound for $\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})$ is obtained subsequently by introducing a uniform bound $\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)}$ on the difference between the empirical potential $\mathcal{V}_z$ and the limiting potential $\mathcal{V}_\rho$.

$$\begin{aligned}
\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) &\leq \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau + C_q h^{n+1}, \\
&+ \int_0^h \left(L_z\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) - L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right)\right) d\tau, \\
&\leq \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau + C_q h^{n+1} \\
&+ h \sup_{\tau\in[0,h]} \left| L_z\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) - L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right)\right|, \\
&\leq \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau + h \sup_{q\in\Omega\subseteq Q} |\mathcal{V}_z(q) - \mathcal{V}_\rho(q)| + C_q h^{n+1}, \\
&= \mathcal{A}_{d,\rho}(\Pi^n q^E_{d,\rho}) + h\left\|\mathcal{V}_z - \mathcal{V}_\rho\right\|_{C(\Omega)} + C_q h^{n+1}.
\end{aligned}$$

We now construct a lower bound for $\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})$. By assuming that the discrete Lagrangian L_z is smooth enough that the quadrature error estimate for $\mathcal{Q}_m$ holds, we know that

$$\begin{aligned}
\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) &= \sum_{j=1}^{m} w_j L_z(q_{m,n,z}(t_j), \dot{q}_{m,n,z}(t_j)),\\
&\geq \int_0^h L_z\left(q_{m,n,z}(\tau), \dot{q}_{m,n,z}(\tau)\right) d\tau - C_q h^{n+1},\\
&= \int_0^h L_\rho\left(q_{m,n,z}(\tau), \dot{q}_{m,n,z}(\tau)\right) d\tau - C_q h^{n+1},\\
&+ \int_0^h \left(L_z\left(q_{m,n,z}(\tau), \dot{q}_{m,n,z}(\tau)\right) - L_\rho\left(q_{m,n,z}(\tau), \dot{q}_{m,n,z}(\tau)\right)\right) d\tau,\\
&\geq \int_0^h L_\rho\left(q_{m,n,z}(\tau), \dot{q}_{m,n,z}(\tau)\right) d\tau - h \sup_{q\in\Omega\subseteq Q} |\mathcal{V}_z(q) - \mathcal{V}_\rho(q)| - C_q h^{n+1}. \quad (4.22)
\end{aligned}$$

By assumption, $q^E_{d,\rho}$ is a minimizer of the exact discrete action, so that $\mathcal{A}_{d,\rho}(q) \geq \mathcal{A}_{d,\rho}(q^E_{d,\rho})$ for all $q \in C^2([0,h];Q)$. We apply this inequality to the last line of Equation 4.22, and subsequently employ the Lipchitz continuity of L_ρ as well as the approximation properties of $\mathbb{M}^n([0,h];Q)$, to find

$$\begin{aligned}
\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) &\geq \int_0^h L_\rho\left(q^E_{d,\rho}(\tau), \dot{q}^E_{d,\rho}(\tau)\right) d\tau - h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)} - C_q h^{n+1},\\
&\geq \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau,\\
&+ \int_0^h \left(L_\rho\left(q^E_{d,\rho}(\tau), \dot{q}^E_{d,\rho}(\tau)\right) - L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right)\right) d\tau\\
&\qquad - h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)} - C_q h^{n+1},\\
&\geq \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau - \ell C_a h^{n+1} - C_q h^{n+1} - h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)}.
\end{aligned}$$

In summary then, we have derived the pair of inequalities that bound $\mathcal{A}_{d,\rho}(\Pi^n q^E_{d,\rho}) - \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})$ from above and below. We have

$$(\ell C_a + C_q)\, h^{n+1} + h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)} \geq \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau - \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}),$$

and,

$$C_q h^{n+1} + h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)} \geq \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) - \int_0^h L_\rho\left(\Pi^n q^E_{d,\rho}(\tau), \frac{d}{dt}\left(\Pi^n q^E_{d,\rho}\right)(\tau)\right) d\tau.$$

We can combine this pair of inequalities into an upper bound on the second term in Equation 4.21.

$$\begin{aligned}
\left|\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) - \int_0^h L_\rho\left(\Pi^n q^E(\tau), \frac{d}{dt}\left(\Pi^n q^E\right)(\tau)\right) d\tau\right| &= \left|\mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z}) - \mathcal{A}_{d,\rho}(\Pi^n q^E)\right|,\\
&\leq (\ell C_a + C_q)\, h^{n+1} + h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)}.
\end{aligned}$$

Finally, the generalization of Theorem 3.3 in [12] follows from combining the final form of the bounds for terms 1 and 2 in Equation 4.21.

$$\begin{aligned}\left|L_d^E - L_{d,(m,n,z)}\right| &= \left|\mathcal{A}_{d,\rho}(q_{d,\rho}^E) - \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})\right|, \\ &\leq \left|\mathcal{A}_{d,\rho}(q_{d,\rho}^E) - \mathcal{A}_{d,\rho}(\Pi^n q_{d,\rho}^E)\right| + \left|\mathcal{A}_{d,\rho}(\Pi^n q_{d,\rho}^E) - \mathcal{Q}_m\mathcal{A}_{d,z}(q_{m,n,z})\right|, \\ &\leq \ell C_a h^{n+1} + (\ell C_a + \mathrm{C}_q)h^{n+1} + h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)}, \\ &= (2\ell C_a + C_q)\, h^{n+1} + h\|\mathcal{V}_z - \mathcal{V}_\rho\|_{C(\Omega)}.\end{aligned}$$

4.3 Learning Theory and Approximation

The construction of discrete mechanics in this paper utilizes learned, empirical potential functions that depend on a set of field observations. In this section we review recent relevant efforts in learning theory in reproducing kernel Hilbert spaces that will enable the construction of probabilistic error estimates for the empirical potential functions, and subsequently, for the discrete evolution laws. We begin with a discussion of the fundamentals of reproducing kernel Hilbert spaces.

4.3.1 Reproducing Kernel Hilbert Spaces

While our discrete dynamics evolve in the state space $Q \equiv \mathbb{R}^d$, we assume that the evolution law will be constructed such that the discrete state remains in some compact metric space (Ω, d) with $\Omega \subseteq Q$. A continuous kernel $K : \Omega \times \Omega \to \mathbb{R}$ is given and is assumed to generate a reproducing kernel Hilbert space $(H, (\cdot,\cdot)_H)$ over Ω. [36] The reproducing property of the kernel guarantees that

$$(K_q, f)_H = f(q) \tag{4.23}$$

for all $f \in H$ and $q \in \Omega$ where the function $K_q(\cdot) := K(q, \cdot)$. The field observations $Z = \{z_1, z_2, \cdots, z_N\} \subseteq \Omega \subseteq Q$ are assumed to be generated as random samples that are distributed in terms of the unknown probability measure ρ that is defined on Ω. In our application, the measure ρ describes how the samples are concentrated in $\Omega \subseteq Q$. We assume that the kernel K is sufficiently regular to continuously embed H in $L^2(\Omega, \rho)$. In other words the linear injection $I_K : H \to L^2(\Omega, \rho)$

$$I_K : f \mapsto I_K f = f \tag{4.24}$$

is bounded, and we have $\|f\|_{L^2(\Omega,\rho)} \leq \|I_K\|\|f\|_H$ for all $f \in H$. The adjoint operator $I_K^* : L^2(\Omega, \rho) \to H$ is given by

$$\begin{aligned}(I_K K_q, g)_{L^2(\Omega,\rho)} &= (K_q, I_K^* g)_H \\ &= (I_K^* g)(q) = \int_\Omega K_q(y)g(y)\rho(dy).\end{aligned}$$

$$\begin{tikzcd}
 & L^2(\Omega,\rho) \arrow[rr, "T_K"] \arrow[dr, "I_K^*"] & & L^2(\Omega,\rho) \\
H \arrow[ur, "I_K"] \arrow[rr, "T_\rho"] & & H \arrow[ur, "I_K"] &
\end{tikzcd}$$

Figure 4.2: Commutative diagram defining operators T_K and T_ρ in terms of I_K, I_K^*

We define the operator $T_K : L^2(\Omega,\rho) \to L^2(\Omega,\rho)$ as $T_K := I_K I_K^*$, and we see that

$$T_K g := I_K I_K^* g = I_K \underbrace{\int_\Omega K(\cdot,r)g(r)\rho(dr)}_{\in H} = \underbrace{\int_\Omega K(\cdot,r)g(r)\rho(dr)}_{\in L^2(\Omega,\rho)}.$$

Analogously, we set $T_\rho := I_K^* I_K$ so that $T_\rho : H \to H$. Since I_K^* is a compact operator, both T_ρ and T_K are compact and self-adjoint. The spectral theory for compact, self-adjoint operators is reviewed in Appendix 4.8.3. More extensive summaries can be found in [?, 39]. The eigenvalues of the operators T_K and T_ρ are identical and are arranged in an extended enumeration, including multiplicities, in nonincreasing order

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq 0.$$

Recall that the only accumulation point of this infinite sequence is 0. We denote by $\{\psi_j\}_{j=1}^\infty$ and $\{\phi_j\}_{j=1}^\infty$ the eigenvectors of T_ρ and T_K, respectively, associated with the eigenvalues $\{\lambda_j\}_{j=1}^\infty$. The spectral theory for compact, self-adjoint operators guarantees that the following expansions are norm-convergent:

$$T_\rho g = \sum_{j=1}^\infty \lambda_j (g,\psi_j)_H \psi_j \qquad \text{in } H,$$

$$T_K f = \sum_{j=1}^\infty \lambda_j (f,\phi_j)_{L^2(\Omega,\rho)} \phi_j \qquad \text{in } L^2(\Omega,\rho),$$

$$I_K^* f = \sum_{j=1}^\infty \sigma_j (f,\phi_j)_{L^2(\Omega,\rho)} \psi_j \qquad \text{in } H,$$

for each $f \in L^2(\Omega,\rho)$ and $g \in H$. In these equations $\sigma_j := \sqrt{\lambda_j}$ is the jth singular value of the operator I_K^*.

4.3.2 The Discrete Operator T_z

In our problem T_ρ and T_K are integral operators. They map

$$T_\rho : H \to H$$

$$T_K : L^2(\Omega,\rho) \to L^2(\Omega,\rho),$$

so they act between infinite-dimensional Hilbert spaces in general. Our goal is to construct a suitable approximation T_z of T_ρ. The construction of T_z takes several steps. Suppose we have a sampling operator

$$S_z : H \to \mathbb{R}^N, \qquad S_z(f) := \begin{Bmatrix} f(x_1) \\ \vdots \\ f(x_N) \end{Bmatrix} \in \mathbb{R}^N$$

We can easily compute the adjoint of the sampling operator by its definition $(S_z f, x)_{\mathbb{R}^N} = (f, S_z^* x)_H$, when H is a reproducing kernel Hilbert space. As shown in Appendix 4.8.2, the adjoint $S_z^* : \mathbb{R}^N \to H$ with

$$S_z^* \begin{Bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{Bmatrix} = \sum_{i=1}^{N} y_i K_{z_i}$$

when the samples $Z = \{z_1, z_2, \cdots, z_N\} \subseteq \Omega \subseteq Q$.The discrete operator T_z is defined as

$$T_z = \frac{S_z S_z^*}{N} = \frac{1}{N} K(z_i, z_j) = \frac{\mathbb{K}}{N}. \tag{4.25}$$

$\mathbb{K}$ is the learning matrix where $\mathbb{K}_{i,j} = K(z_i, z_j)$ and z_i represents the ith data point in the data set. More details on the construction of the sampling operator S_z and the discrete operator T_z can be found in in Appendix 4.8.2.

4.3.3 Empirical Potential Functions

The discrete evolution laws studied in this paper are constructed by using an empirical potential function V_z that drives the discrete state to evolve near the subset $\Omega \subset Q$. We define the empirical potential function V_z, its "infinite sample" limit V_ρ, and its approximation V_j due to spectral filtering via the expressions

$$\begin{aligned} V_\rho(q) &:= (T_\rho^\dagger T_\rho K_q, K_q)_H, \\ V_z(q) &:= (g_\lambda(T_z) T_z K_q, K_q)_H, \text{ and} \\ V_j(q) &:= (g_\lambda(T_\rho) T_\rho K_q, K_q)_H. \end{aligned}$$

for all $q \in \Omega \subseteq Q$ where the filter function g_λ (see [38]) is defined to be the spectral cutoff function. With these definitions, we can now derive an expression that can be used in conjunction with Theorem 1 to generate an error bound on the discrete Lagrangian.

Theorem 7 ***Assuming that the kernel*** K ***and probability measure*** ρ ***satisfy the*** **smoothness prior**

$$\sup_{q\in\Omega} \|T_\rho^{-s/2} P_\rho K_q\|_H^2 \leq C_s,$$

for some constant $C_s > 0$ ***and*** $s > 0$ ***, we have***

$$\mathbb{E}_{\rho^N}\left(\sup_{q\in\Omega}\left|V_j(q) - V_z(q)\right|\right) \leq \frac{C}{\sqrt{N}(\lambda_{M_j} - \lambda_{M_j+1})},$$

where $\mathbb{E}_{\rho^N}$ ***is the expectation over the*** N ***data points that are used to construct the empirical potential and*** λ_{M_j}***,*** λ_{M_j+1} ***are the singular values right before and after the cut-off frequency in the spectral cut-off function*** g_λ***. The index*** M_j ***is the number of retained singular values used in the construction of*** $V_j(q)$***,*** s ***is a smoothness parameter, and*** P_ρ ***(see Appendix 4.8.2) is a projection operator on the range of*** T_ρ***.***

Proof 9 ***A bound for the difference between the empirical and limiting potential function,*** V_z ***and*** V_ρ***, respectively, can be derived by writing***

$$\begin{aligned}\left|V_z(q) - V_\rho(q)\right| &\leq \left|V_z(q) - V_j(q)\right| + \left|V_j(q) - V_\rho(q)\right| \\ &= \underbrace{\left|\left((g_\lambda(T_z)T_z - g_\lambda(T_\rho)T_\rho)\, K_q, K_q\right)_H\right|}_{\textit{term 1}} + \underbrace{\left|\left(\left(g_\lambda(T_\rho)T_\rho - T_\rho^\dagger T_\rho\right) K_q, K_q\right)_H\right|}_{\textit{term 2}}. \end{aligned} \tag{4.26}$$

The term 1 ***on the right satisfies the inequality***

$$\begin{aligned}\left|\left(\left(g_\lambda(T_z)T_z - g_\lambda(T_\rho)T_\rho\right)K_q, K_q\right)_H\right| &\leq \left\|\left(g_\lambda(T_z)T_z - g_\lambda(T_\rho)T_\rho\right)K_q\right\|_H \|K_q\|_H, \\ &\leq \left\|\left(g_\lambda(T_z)T_z - g_\lambda(T_\rho)T_\rho\right)\right\|_{\textit{operator}} \|K_q\|_H \|K_q\|_H, \\ &\leq \left\|\left(g_\lambda(T_z)T_z - g_\lambda(T_\rho)T_\rho\right)\right\|_{HS}, \\ &\leq \frac{2}{\lambda_{M_j} - \lambda_{M_j+1}} \|T_z - T_\rho\|_{HS}, \end{aligned}$$

where the final inequality follows from Section 5.4 of [38]. The bound on term 1 ***in Inequality 4.26 is obtained when we note that we have***

$$\|T_\rho - T_z\|_{HS} \leq \frac{2(\delta \vee \sqrt{2\delta})}{\sqrt{N}}$$

with probability at least $1-2e^{-\delta}$ ***from Lemma 1 of [38]. This completes the proof for bounding term 1. We can bound term*** 2 ***by utilizing the Schmidt expansion of compact operators as summarized in Appendix 4.8.3. We have***

$$\begin{aligned}\left|\left(\left(g_\lambda(T_\rho)T_\rho - T_\rho^\dagger T_\rho\right)K_q, K_q\right)_H\right| &= \left|\left(\left(g_\lambda(T_\rho)T_\rho - T_\rho^\dagger T_\rho\right)P_\rho K_q, P_\rho K_q\right)_H\right|, \\ &= \left(\sum_{j=M_j+1}^{\infty}(P_\rho K_q, \psi_j)_H \psi_j, P_\rho K_q\right)_H, \\ &= \sum_{j=M_j+1}^{\infty}(P_\rho K_q, \psi_j)_H^2 \leq \lambda_{M_j+1}^s C_s.\end{aligned}$$

The last step in this string of inequalities follows from the assumption that

$$\sup_{q\in Q}\left\|T_\rho^{-s/2}P_\rho K_q\right\|_H^2 \leq C_s.$$

This inequality is known as the approximation assumption and is discussed in [35, 38].

Combining the estimates for terms 1 ***and*** 2 ***yields the desired bound on the difference between the empirical and limiting potential function***

$$\sup_{q\in\Omega}\left|V_z(q) - V_\rho(q)\right| \leq \frac{4(\delta \vee \sqrt{2\delta})}{\sqrt{N}(\lambda_{M_j} - \lambda_{M_j+1})} + C_s\lambda_{M_j+1}^s. \tag{4.27}$$

Ultimately, we wish to derive error bounds on discrete evolution laws that arise from discrete formulations Lagrangian or Hamiltonian mechanics. It will be convenient to re-express Equation 4.27 in terms of probabalistic concentration inequalities in the style discussed in references [1, 6]. Equation 4.27 can be understood as implying that

$$\sup_{q\in Q}|V_\rho(q) - V_z(q)| \leq C_s\lambda_{M_j+1}^s + \eta \tag{4.28}$$

for all samples $z \notin \Lambda_{j,N}(\eta)$ ***where***

$$\Lambda_{j,N}(\eta) := \left\{z: \quad \sup_{q\in Q}|V_j(q) - V_z(q)| > \eta,\right\}$$

and the measure of the set $\Lambda_{j,k}(\eta)$ ***of bad samples satisfies***

$$\rho^N\left\{z \in \Lambda_{j,N}(\eta)\right\} \leq 2e^{-\delta}.$$

The equivalence of Equations 4.27 and 4.28 is evident when we set

$$\eta = \frac{4(\delta \vee \sqrt{2\delta})}{\sqrt{N}(\lambda_{M_j} - \lambda_{M_j+1})}. \tag{4.29}$$

It is clear that we can write

$$\eta = \begin{cases} \frac{4\delta}{\sqrt{N}(\lambda_{M_j}-\lambda_{M_j+1})} & \textbf{\textit{if}}\ \delta \leq 2 \\ \frac{4\sqrt{2\delta}}{\sqrt{N}(\lambda_{M_j}-\lambda_{M_j+1})} & \textbf{\textit{if}}\ \delta > 2 \end{cases}$$

or equivalently,

$$\delta = \begin{cases} \frac{1}{4}\sqrt{N}(\lambda_{M_j}-\lambda_{M_j+1})\eta & \textbf{\textit{if}}\ \eta \leq \eta_{cr} \\ \frac{1}{32}N(\lambda_{M_j}-\lambda_{M_j+1})^2\eta^2 & \textbf{\textit{if}}\ \eta > \eta_{cr} \end{cases}$$

where $\eta_{cr} := 8\sqrt{N}\left(\lambda_{M_j}-\lambda_{M_j+1}\right)$***. It is therefore possible to summarize the probabilistic error bound in the form***

$$\rho^N\left\{\sup_{q\in Q}|V_j(q)-V_z(q)|>\eta\right\} \leq \begin{cases} 1 & \textbf{\textit{if}}\ \eta \leq \eta_{cr} \\ 2e^{-\frac{1}{32}N\left(\lambda_{M_j}-\lambda_{M_j+1}\right)^2\eta^2} & \textbf{\textit{if}}\ \eta > \eta_{cr} \end{cases} \tag{4.30}$$

As emphasized in [6], the utility in estimates of the form in Equation 4.30 is that this expression gives a succinct description of the interplay between the number of samples and approximation error, and it is quite useful in deriving bounds for the expectation of the error. We have

$$\begin{aligned} \mathbb{E}_{\rho^N}\left(\sup_{q\in Q}\left|V_j(q)-V_z(q)\right|\right) &= \int_Q \rho^N\left\{z : \sup_{q\in Q}|V_j(q)-V_z|>\eta\right\}d\eta \\ &= \int_0^{\eta_{cr}} 1 d\eta + \int_{\eta_{cr}}^{\infty} 2e^{-N(\lambda_{M_j}-\lambda_{M_j+1})^2\eta^2}d\eta \\ &\leq \frac{C}{\sqrt{N}(\lambda_{M_j}-\lambda_{M_j+1})} \end{aligned}$$

for some constant $C > 0$***.***

4.4 Error in Probability for the Galerkin Variational Integrator

Having established the bounds in theorems 1 and 2, it is now possible to derive a probabilistic error bound for the empirical, discrete Lagrangian $L_{d,(m,n,z)}$**.**

Theorem 8 ***Choose*** $\epsilon > 0$ ***and*** $h > 0$***. Calculate the modal truncation level*** j ***such that*** $\lambda_{M_j+1} \leq \epsilon^2 h^n$ ***and choose the number of samples*** N ***such that***

$$\frac{1}{\sqrt{N}} \leq \epsilon^2 h^n(\lambda_{M_j}-\lambda_{M_j+1}).$$

Then the discrete Lagrangian satisfies the error bound

$$\mathbb{E}_{\rho^N}\left\{|L_{d,(m,n,z)} - L^E_{d,\rho}|\right\} \lesssim h^{n+1}.$$

Proof 10 ***We know that***

$$\begin{aligned}&\left|L_{d,(m,n,z)} - L^E_{d,\rho}\right| \\ &\quad \leq C_1 h^{n+1} + \frac{h}{\epsilon^2}\|V_j - V_\rho\|_{C(\Omega)} + \frac{h}{\epsilon^2}\|V_z - V_j\|_{C(\Omega)}.\end{aligned}$$

This inequality implies that

$$\begin{aligned}&\mathbb{E}_{\rho^N}\left(L_{d,(m,n,z)} - L^E_{d,\rho}\right) \\ &\quad \leq C_1 h^{n+1} + C_s \frac{h\lambda^s_{M_j+1}}{\epsilon^2} + C_2 \frac{h}{\epsilon^2\sqrt{N}(\lambda_{M_j} - \lambda_{M_j+1})} \\ &\quad \lesssim h^{n+1}.\end{aligned}$$

Theorems 1,2, and 3 provide a rigorous framework for the construction of empirical potentials and for the determination of probabilistic error bounds for their associated variational integrators. As such the developed should be a fruitful source of combined analytical-empirical models for our applications such as Bayesian estimation for biomechanical motion models. As discussed in our introduction, such recursive estimators require a prediction step that uses the motion model, and the prediction is then updated using the recursive filter. We refer to these models as analytical-empirical models since they utilize an analytic expression for the kinetic energy and an empirical expression for the potential energy term that effectively constraints motion to concentrate as dictated by the unknown probability measure. The ultimate idea is to use biomechanical motion models from one specimen or motion regime to bootstrap the prediction step with better estimates than those that are generated by random walk models.

4.5 An Open Problem

Theorems 6, 7, and 8 provide a general framework for generating analytical-empirical equations of discrete mechanics using variational methods. We close our theoretical discussion with a description of an open problem. Suppose that the empirical data collected happens to be generated by a Lagrangian system that is rigorously constrained to lie on a submanifold that is defined in terms of constraints. For simplicity assume that the constraints are holonomic. It is well known that the trajectories of such as system can be described in terms of a Lagrangian L_λ that depends on Lagrange multipliers λ and the holonomic constraints. Although we have established criteria in which the discrete Lagrangian $L_{d,(m,n,z)}$ can be used to build an approximation of the flow generated by L_ρ, we have not established in this paper that such discrete trajectories approximate the trajectories of the *exactly constrained* system whose evolution is determined

by L_λ and the holonomic constraints. Such an analysis requires substantial additional analysis and in principle would follow from approaches as outlined in [40] or [11]. We briefly discuss the open question in more detail.

Let $L_{z,\epsilon}$ be an an empirical-analytical Lagrangian that has the form

$$L_{z,\epsilon}(q,\dot{q}) = T(q,\dot{q}) - (U(q) + \frac{1}{\epsilon^2}V_z(q)). \tag{4.31}$$

for some fixed small constant $\epsilon > 0$. We define ρ-limiting Lagrangian $L_{\rho,\epsilon}$ as

$$L_{\rho,\epsilon}(q,\dot{q}) = T(q,\dot{q}) - (U(q) + \frac{1}{\epsilon^2}V_\rho(q)). \tag{4.32}$$

We denote by $q_{z,\epsilon}$ and $q_{\rho,\epsilon}$ typical trajectories determined by the unconstrained Lagrangians in Equations 4.31 and 4.32, respectively. The approaches in this paper establish how discrete Lagrangian approximations of $L_{z,\epsilon}$ can be shown to approximate continuous solutions $q_{\rho,\epsilon}$ for a fixed $\epsilon > 0$. The approximation results in this paper say nothing about how those discrete equations of mechanics converge *as* $\epsilon \to 0$ to the continuous time trajectories generated from a Lagrangian L_λ with multipliers λ and its associated holonomic constraints.

To illustrate this subtle point more clearly, we summarize results from [11] and [40]. These references state their convergence results in a Hamiltonian framework which we briefly summarize. By applying Legendre transform $H = \dot{q}^T p - L$, we can define Hamiltonians $H_{z,\epsilon}$ and $H_{\rho,\epsilon}$ associated with $L_{z,\epsilon}$ and $L_{\rho,\epsilon}$, respectively. They are written in the form

$$H_{z,\epsilon}(q,p) = \frac{1}{2}p^T M(q)^{-1}p + U(q) + \frac{1}{\epsilon^2}V_z(q)$$

and

$$H_{\rho,\epsilon}(q,p) = \frac{1}{2}p^T M(q)^{-1}p + U(q) + \frac{1}{\epsilon^2}V_\rho(q),$$

To fix dimensions, we assume that $\Omega \subseteq Q := \mathbb{R}^d$, that the $d \times d$ mass matrix $M(q)$ is assumed to be uniformly positive definite, smooth, and symmetric. Suppose further that there exist constraint functions $\phi : \Omega \to \mathbb{R}^m$ for $m < d$ that satisfy

$$\{q \in \Omega : \phi(q) = 0\} = \{q \in \Omega : V_\rho(q) = 0\} \tag{4.33}$$

We now define the effective constrained Hamiltonian H_λ as

$$H_\lambda = \frac{1}{2}P^T M(Q)^{-1}P + U(Q) + W(I,Q),$$

and designate by Q a continuous time trajectory generated from H_λ and the holonomic constraints $\phi(Q) = 0$. Note that the effective Hamiltonian H_λ includes a correction term, $W(I,Q) = \sum_{k=1}^m I_k\omega_k(Q)$ in which ω_k for $k = 1,\ldots,m$ are frequencies and $I = (I_1,\cdots,I_m)$ are the associated actions. If (1) the initial conditions

are sufficiently close to the constrained submanifold, (2) the frequencies ω_k satisfy nonresonance conditions, and (3) the mass matrix satisfies certain tangency conditions defined in terms of the submanifold, references [40], [11] show that

$$Q(t) - q_{\rho,\epsilon}(t) = \mathcal{O}(\epsilon)$$

or $\|q_{\rho,\epsilon} - Q\| \leq C\epsilon$ for some constant C. Under these very restrictive assumptions, we can construct error estimates between the discrete evolution laws $q_{d,\epsilon,z}$ and the exactly constrained continuous time trajectory via

$$\|q_{d,\epsilon,z}(k) - Q(t_k)\| \leq \underbrace{\|q_{d,\epsilon,z}(k) - q_{\rho,\epsilon}(t_k)\|}_{\textbf{term 1}} + \underbrace{\|q_{\rho,\epsilon}(t_k) - Q(t_k)\|}_{\textbf{term 2}}$$

for discrete times t_k for $k \in \mathbb{N}$. Bounds on term 1 above follow from Theorems 6, 7, and 8, while bounds on term 2 follow from the analysis just summarized from [40] and [11].

Clearly, the assumptions (1), (2), and (3) are complex and difficult to guarantee in the simplest of cases ***when an analytic expression for the constraints are known.*** This paper is predicated on the assumption that we do not know the constraints. Additionally, the probability measure ρ that dictates concentration of samples in Ω is assumed unknown. This is a standard assumption that underlies the theory of distribution-free learning theory, one of the primary theoretical underpinnings of our approach. Finally, since V_ρ is expressed via an infinite summation of unknown eigenfunctions that depend on ρ, it is impossible to verify condition 4.33.

One of the implications of the above analysis is that the discrete trajectories will concentrate near the submanifold of interest, but the discrete dynamics ***restricted to the submanifold*** may not match the exactly constrained dynamics of the system with Lagrange multipliers. However, the evolution equations do reflect the geometry of the underlying submanifold if it exists. This point is clearly illustrated in the numerical examples in the next section.

4.6 Numerical Examples

In this section we explore the effectiveness of our method with some numerical examples. First we investigate the construction of the empirical potential function and explore how different variables affect the shape of the potential function. We next examine the performance of the method with a simple pendulum example. We then propose an additional numerical algorithm to construct a data-driven effective constraint to compensate for some of the numerical complexities that arise because the true constraints on the system are unknown.

4.6.1 Construction of the Empirical Potential Function when the Submanifold is a Circle

In this example we use a time series data $\{z_i\}_{i=1}^N$ of the position of a particle moving on a circle to construct an empirical potential function based on the method presented in Section . The physics of the problem is such that views of the empirical potential function can be shown in a three dimensional figure. We use a random distribution of data around the circle of $x^2 + y^2 = 1, z = 0$, where the distributions is defined to be a zero mean Gaussian noise with $SNR = 30dB$. A kernel function $K_{z_i}(q) := K(z_i, q)$ is centered at each data point to constitute a metric for the distance between any point in the configuration space and the data point. The exponential kernel function

$$K_q(z_i) = K(q, z_i) = e^{-\beta\|q-z_i\|^2} \tag{4.34}$$

centered at z_i is employed in this example with the parameter β determining the spread or variance of the function around z_i. The RKHS H is then $H := \overline{\{K_x(\cdot)|x \in \Omega\}}$. The matrix $\mathbb{K}_{i,j} = k(z_i, z_j)$ is constructed employing data points scattered around the submanifold of interest. We use the spectral cut-off filter function to calculate $g_\lambda(\frac{1}{N}\mathbb{K})$. The empirical data function is then

$$\begin{aligned} -V_z(q) &= \left(\frac{1}{N} g_\lambda\left(\frac{1}{N}\mathbb{K}\right) S_z K_q, S_z K_q\right), \\ &= \frac{1}{N} K_z(q)^* g_\lambda\left(\frac{1}{N}\mathbb{K}\right) K_z(q). \end{aligned}$$

Figure 4.3 depicts $-V_z(q)$ for different values of the spread β where the function attains its maximum value in the vicinity of the submanifold of interest. The empirical potential function has been constructed so that V_z obtains its minimum value on the support of the observed dynamics. As it is shown in this example, there is a trade-off between the convexity of the function and its smoothness. The parameter β can be regulated to obtain a desired smoothness in the empirical potential function.

4.6.2 The Simulation of the Pendulum

In the second example, we use the data derived from a mass attached to a rod moving in a plane. Our proposed algorithm has four main steps to derive the discrete evolution law of a system based on field observations:

1. **collect data from field observations,**
2. **create the empirical potential function,**
3. **derive the Lagrangian for the system,**

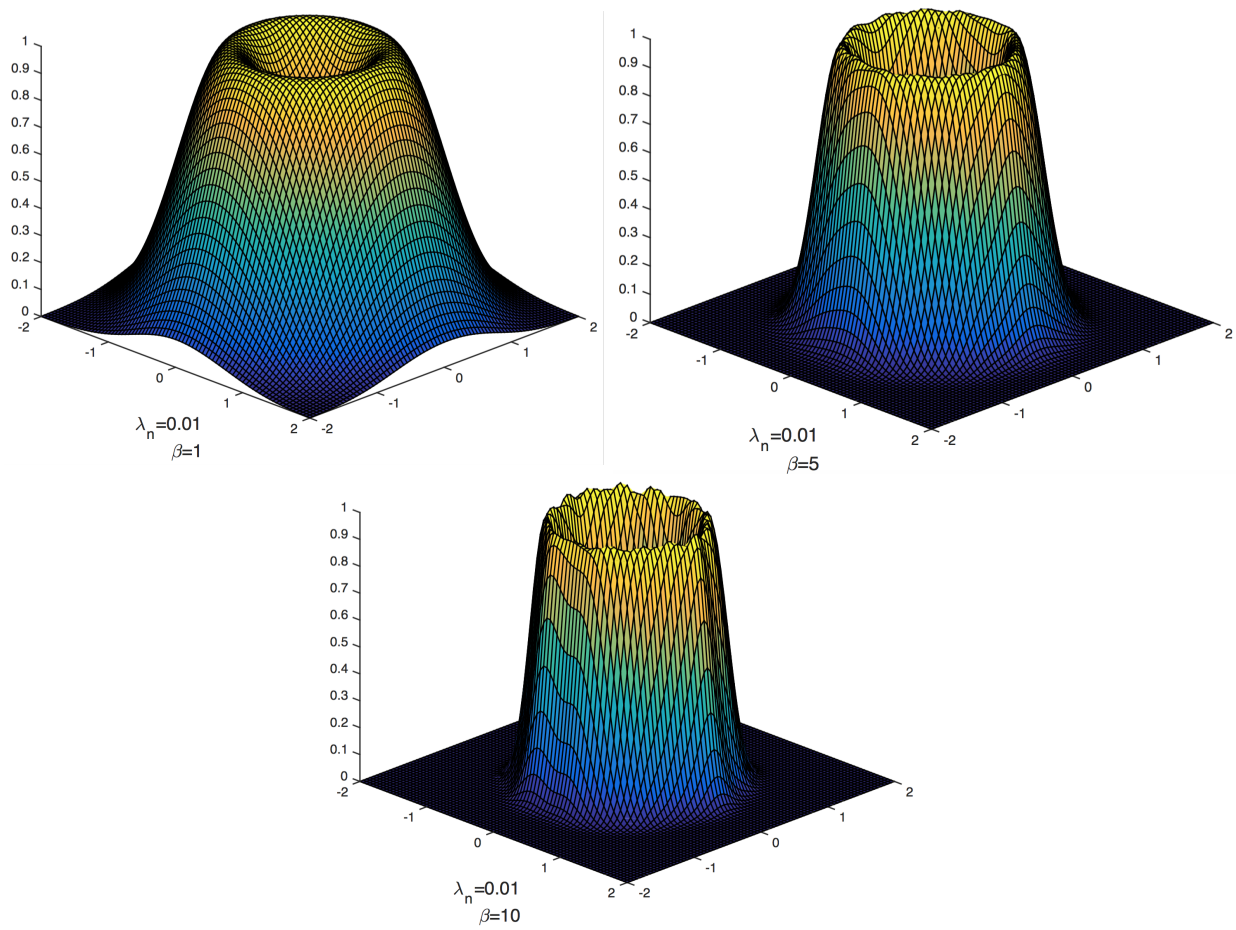


Figure 4.3: $F_n(q)$ for Different β

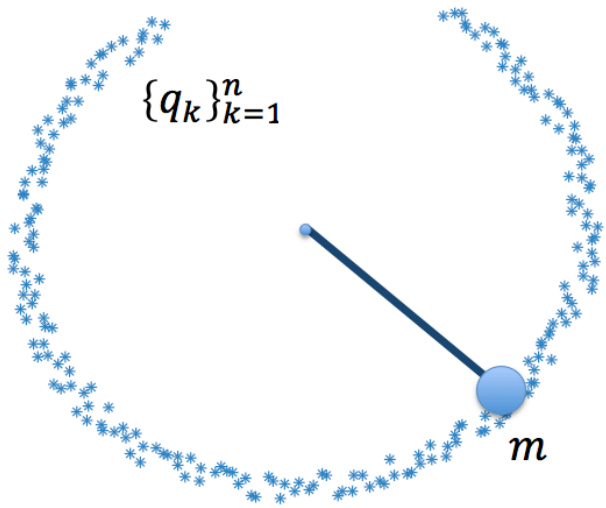


Figure 4.4: Noisy data collected from a pendulum

4 derive the discrete evolution law that concentrates trajectories near the embedded set or submanifold.

Figure 4.4 depicts the random samples generated with the distribution defined to be a zero mean white Gaussian noise with $SNR = 10dB$. In the second step, we use the parameters that we investigated in the previous example to generate the empirical potential function

$$V_z(q) = -\frac{1}{N} K_z(q)^* g_\lambda \left(\frac{1}{N}\mathbb{K}\right) K_z(q).$$

Next, we include any *a priori* information about the system in the Lagrangian that we might know. For instance, in this example the the mass of the particle is known and therefore a gravitational potential function can be added in the Lagrangian formulation. We could include any structural information in a more complicated dynamical systems. The weighted empirical potential function is constructed from observations while the gravitational potential term is called the physical potential. The empirical potential is added as a penalty term in the Lagrangian formulation as

$$\begin{aligned} L(q,\dot{q}) &= T(\dot{q}) - \left(U(q) + \frac{1}{\epsilon^2} V_z(q)\right), \\ &= \frac{1}{2} m\dot{q}^2 - \left(mgq_3 + \frac{1}{\epsilon^2} V_z(q)\right). \end{aligned}$$

The Störmer-Verlet method is used to calculate the discrete evolution law since it is a type of variational integrator that fits the mathematical structure we have outlined. Figure 4.5 shows the simulation results for three different values of β. In this simulation, the time step is $h = 10^{-5}$, the weight for the stiff potential

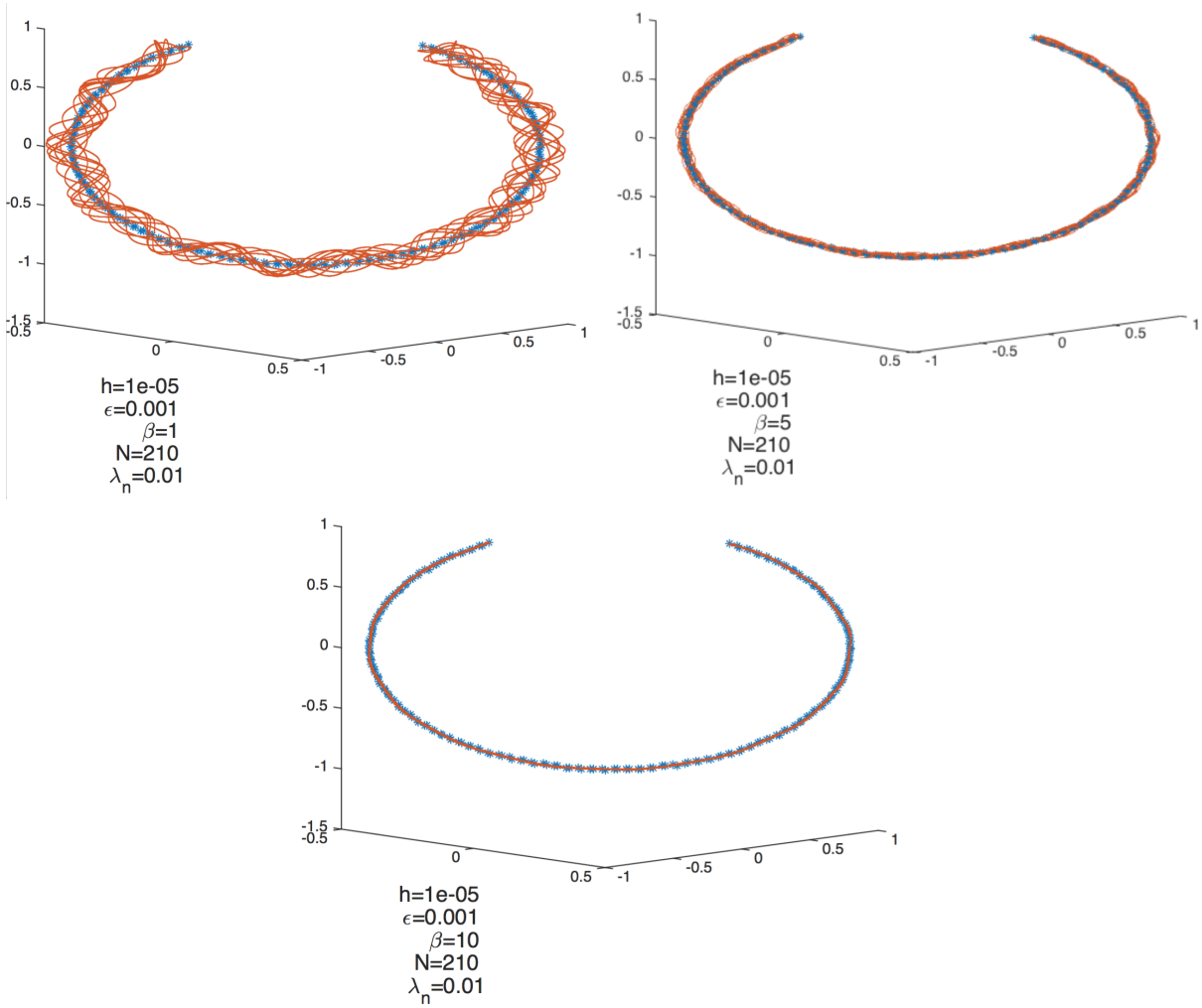


Figure 4.5: Simulation of motion on the submanifold for different values of β, blue dots: data, red line: the simulated motion

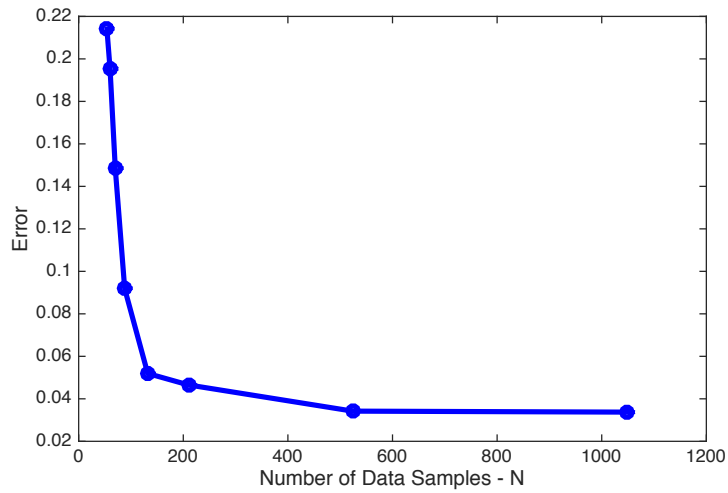


Figure 4.6: The error between simulated motion and the submanifold improves by increasing the number of data samples.

term is $1/\epsilon^2$ where $\epsilon = 0.001$, and the filtering frequency in the spectral filter is $\lambda_n = 0.01$. The convexity of empirical potential function influences the qualitative behavior of the motion close to the data submanifold. Higher values of beta are associated with sharper potential functions and therefore the motion stays closer to the submanifold of interest.

Another important parameter that affects the closeness of simulated motion to an embedded set or submanifold is the number of data points that are used to construct the empirical potential function. Figure 4.6 depicts the distance of the simulated motion to the motion submanifold as the number of samples N is varied.

4.6.3 The Empirical Potential and Data-driven Pseudo-constraint

In this section we summarize an algorithm that uses the empirical-potential to construct an constraint on the system. We refer to the constraint so defined as a pseudo-constraint. In theory we have developed an empirical potential function that attain its extremum in the close vicinity of a subset or submanifold that approximately supports the observed dynamics. However, our implementations of numerical examples have shown that the effectiveness of the proposed empirical potential is sensitive to initial conditions and to various hyperparameters such as ϵ or α used in the formulation. In this section we summarize how the empirical potential can be used to define an effective constraint, or pseudo-constraint. The pseudo-constraint can provide valuable insight into the nature of the resulting simulations. We begin by calculating the singular value decomposition (SVD) of the function $g_\lambda\left(\frac{\mathbb{K}}{N}\right)$. Since $g_\lambda\left(\frac{\mathbb{K}}{N}\right)$ is a real, symmetric matrix, we have $-V_z(q) = u\Sigma u^T$ where Σ is nonnegative diagonal matrix of singular values in de-

scending order, and u is an orthonormal matrix. We construct a partition of u according to

$$u = \begin{bmatrix} u_1 \cdots & u_s & u_{s+1} \cdots u_N \end{bmatrix},$$

define $u_s := \begin{bmatrix} u_1 \cdots u_s \end{bmatrix}$, and define Σ_s conformally. The pseudo-constraint ϕ_s is defined in terms of this decomposition as

$$\phi_s(q) = \frac{1}{\sqrt{N}} \Sigma_s^{1/2} u_s^T K_z(q). \tag{4.35}$$

We now define pseudo-constraint functions ϕ_s so that they satisfy the equations

$$-V_z(q) = \phi_s(q)^T \phi_s(q) \approx \frac{1}{N} K_z(q)^* g_\lambda \left(\frac{\mathbb{K}}{N} \right) K_z(q).$$

Assuming the pseudo-constraint function $\phi_s(q) \in C^2(Q)$, the approach considered in this paper can interpreted as a particular class of penalty methods for Lagrangian formulations. The references [41], [43], [42] are a good sources to review this class of methods. Using the pseudo-penalty function $\phi_s(q) \in C^2(Q)$ in the inertial penalty method in [42, 41], the motion is generated by the Lagrangian function $L = T(q) - U(q)$ and the pseudo-constraint function ϕ_s. Trajectories satisfy the equations

$$\frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}} L\right) - \frac{\partial}{\partial q} L + \frac{\partial \phi}{\partial q}\left(\frac{\gamma}{\epsilon^2}\ddot{\phi}_s + \frac{\mu}{\epsilon^2}\dot{\phi}_s + \frac{\alpha}{\epsilon^2}\phi_s\right) = 0. \tag{4.36}$$

When $\gamma = \mu = 0$, the above equations are of the form approximated by our formulation of discrete mechanics. Again, it should be noted that the convergence of trajectories of the above equations to those of the exactly constrained Lagrangian with multipliers is not guaranteed unless several additional conditions summarized in Section 4.5 are satisfied. The above Euler-Lagrange equation approximately simulates the motion of the mechanical system of interest as ϵ converges to zero and the weights of the penalty terms grow. The weighting matrices γ, μ, and α regulate the penalty terms. For the sake of simplicity we assume $\mu = 0$ but all the results will hold where μ is not zero. When the data-driven pseudo-constraint is constructed, it is straightforward to calculate the remaining terms using the chain rule as $\dot{\phi}_s = \frac{\partial \phi_s}{\partial q}\dot{q}$, $\ddot{\phi}_s = \frac{d}{dt}\left(\frac{\partial \phi_s}{\partial q}\right)\dot{q} + \frac{\partial \phi_s}{\partial q}\ddot{q}$. Following the steps explained in [41] it follows that

$$\frac{1}{2}\frac{\alpha}{\epsilon^2}\phi_s^T \phi_s + \frac{1}{2}\frac{\gamma}{\epsilon^2}\dot{\phi}_s^T \dot{\phi}_s + T(q, \dot{q}) + U(q) = T(q(0), \dot{q}(0)) + U(q(0)) \tag{4.37}$$

for any natural or T_2 system if the initial conditions exactly satisfy the pseudo-constraint $\phi_s(q(t)) \equiv 0$. We see that, even if the trajectories of the penalized system do not converge to those of the exactly constrained system as $\epsilon \to 0$, it is true that $\|\phi(q(t))\|^2 + \|\dot{\phi}_s(q(t), \dot{q}(t)\|^2 \leq C\epsilon^2$ for some constant C. The summary of this means of implementation of the discrete mechanics is as follows:

1 **collect data $z = \{z_1 \dots z_N\}$ from field observations,**

2 **create the data-driven pseudo-constraint using equation 4.35**

3 **derive the Lagrangian for the system and use equation 4.36 to derive discrete evolution law**

4 **choose a value for ϵ and the hyperparameters α, β, γ for construction of approximate trajectories. Repeat starting with Step 2 as needed.**

We implement the above algorithm for the pendulum example. Our aim is to show that the system trajectory converges to the motion submanifold as ϵ converges to zero. Since $\phi_s^T \phi_s$ and $\dot{\phi}_s^T \dot{\phi}_s$ are quadratic forms, $\alpha \phi_s^T \phi_s + \gamma \dot{\phi}_s^T \dot{\phi}_s$ will measure the proximity of the simulated motion to the motion submanifold.

Figure 4.7 shows the simulation results. We see that the above algorithm effectively constrains the dynamics to stay within and ϵ-neighborhood of the underlying set or submanifold. The error betwen the effective manifold and the system trajectores coverge to zero as ϵ converges to zero.

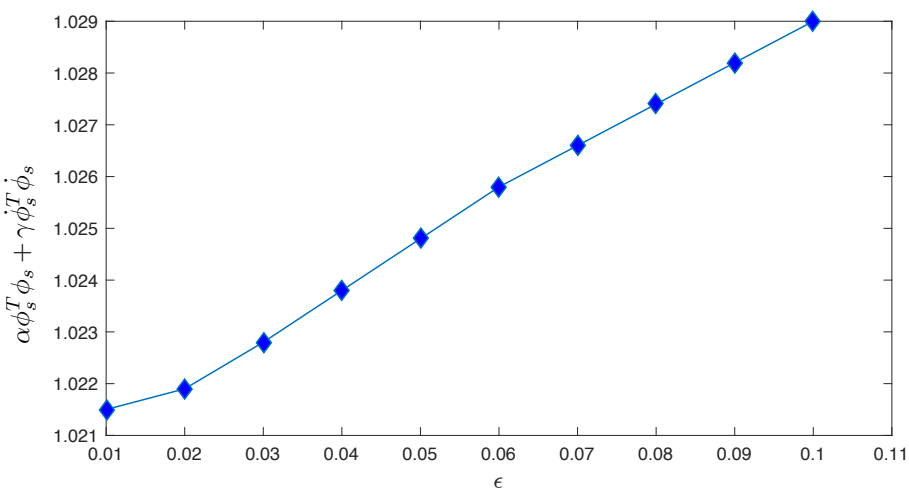


Figure 4.7: $\alpha \phi_s^T \phi_s + \gamma \dot{\phi}_s^T \dot{\phi}_s$ decreases as ϵ converges to zero

4.7 Conclusions

This paper has formulated equations of discrete Lagrangian mechanics that evolve so as to remain concentrated near subsets or submanifolds that are identified from empirical observations. By generalizing a technique from the variational theory of discrete Lagrangian dynamics to a probabilistic framework, we have derived novel probabilistic error bounds for the discrete evolution laws that approximate dynamics on observations of motion near the underlying subsets or submanifolds. In future work we seek to apply this technique to derive motion

equations for complex biological systems, such as bats in flight, that yield faithful representations of submanifolds observed in experiment.

4.8 Appendices

4.8.1 Lipschitz continuity

The Lipschitz continuity of L_ρ requires that

$$\begin{aligned}|L_\rho(x,y) - L_\rho(u,v)| &\leq \ell \, \|(x,y)-(u,v)\|_{\ell^1(\mathbb{R}^{2d})} \\ &= \ell \left\{ \|x-u\|_{\ell^1(\mathbb{R}^d)} + \|y-v\|_{\ell^1(\mathbb{R}^d)} \right\}\end{aligned}$$

with Lipschitz coeffient $\ell > 0$. Either of the following two expressions

$$\begin{aligned}\|f\|_{C([0,h];Q)} &\approx \sup_{t\in[0,h]} \sum_{i=1}^{d} |f_i(t)| \\ &\approx \sup_{t\in[0,h]} \max_{i=1,\dots,d} |f_i(t)|\end{aligned}$$

define equivalent norms for $C([0,h];Q)$ since we know that there are positive constants $c_1, c_2 > 0$ such that

$$c_1 \|v\|_{\ell^1(\mathbb{R}^d)} \leq \|v\|_{\ell^\infty(\mathbb{R}^d)} \leq c_2 \|v\|_{\ell^1(\mathbb{R}^d)}.$$

If x, y, u, v are functions of t on $[0,h]$, we can write

$$\begin{aligned}|L_\rho(x(t),y(t)) - L_\rho(u(t),v(t))| &\leq \ell \left\{ \|x(t)-u(t)\|_{\ell^1(\mathbb{R}^d)} + \|y(t)-v(t)\|_{\ell^1(\mathbb{R}^d)} \right\} \\ &\leq \ell \left\{ \sup_{\tau\in[0,h]} \|x(t)-u(t)\|_{\ell^1(\mathbb{R}^d)} + \sup_{\tau\in[0,h]} \|y(t)-v(t)\|_{\ell^1(\mathbb{R}^d)} \right\} \\ &= \ell \left\{ \|x-u\|_{C([0,h],Q)} + \|y-v\|_{C([0,h],Q)} \right\}\end{aligned}$$

4.8.2 Construction of the sampling operator S_z and the discrete operator T_z

We want to build approximations of T_ρ. In other words, we need a computable approximation T_z of T_ρ that depennds on samples $z = \{z_1, \cdots, z_n\}$ that live on the domain Ω.
The sampling operator S_z is defined as $S_z : H \to \mathbb{R}^N$ where

$$S_z f = \begin{bmatrix} f(z_1) \\ \vdots \\ f(z_N) \end{bmatrix}$$

Recall that we can find an analytical expression for the adjoint $E_q^* : \mathbb{R} \to H$ of the evaluation functional $E_q : H \to \mathbb{R}$ for each $q \in \Omega \subseteq Q$. We have

$$\begin{aligned}(E_q f, y)_{\mathbb{R}} &= (f, E_q^* y)_H = (f(q), y)_{\mathbb{R}}, \\ &= ((K_q, f)_H, y)_{\mathbb{R}} = (yK_q, f)_H.\end{aligned}$$

we see then that $E_q^* y = yK_q$ for $q \in \Omega \subseteq Q$. This identity can be used to show that

$$S_z^* : \mathbb{R}^N \to H,$$

$$S_z^* \begin{Bmatrix} y_1 \\ \vdots \\ y_N \end{Bmatrix} = \sum_{i=1}^N y_i K_{z_i}$$

Now it is a simple matter to compute

$$S_z S_z^* y = S_z \left(\sum_{i=1}^N y_i K_{z_i} \right) = \begin{Bmatrix} \sum_{i=1}^N K(z_i, z_1) y_i \\ \vdots \\ \sum_{i=1}^N K(z_i, z_N) y_i \end{Bmatrix} = K(z_j, z_i)\mathbf{y}.$$

Finally, we define

$$T_z = \frac{1}{N} S_z S_z^* = \frac{1}{N} K(z_i, z_j) = \frac{\mathbb{K}}{N}.$$

Note that $S_z S_z^*$ and $S_z^* S_z$ have the same eigenvalues, but $S_z S_z^* : \mathbb{R}^N \to \mathbb{R}^N$ and $S_z^* S_z : H \to H$. We now define the empirical potential function V_z where

$$\begin{aligned} V_z(q) &= \Big(g_\lambda(T_z) T_z K_q, K_q \Big)_H \\ &= \frac{1}{N} \left(g_\lambda \left(\frac{\mathbb{K}}{N} \right) S_z K_q, S_z K_q \right)_{\mathbb{R}^N} \\ &= \left(\frac{1}{N} S_z^* g_\lambda \left(\frac{\mathbb{K}}{N} \right) S_z K_q, K_q \right)_H \\ &= \left(\frac{1}{n} S_z^* g_\lambda \left(\frac{\mathbb{K}}{N} \right) S_z \begin{Bmatrix} K(z_1, q) \\ \vdots \\ K(z_N, q) \end{Bmatrix}, \begin{Bmatrix} K(z_1, q) \\ \vdots \\ K(z_N, q) \end{Bmatrix} \right)_H . \end{aligned}$$

Note that we need to show that the operators $g_\lambda \left(\frac{S_z^* S_z}{N} \right) \frac{S_z^* S_z}{N}$ and $\frac{1}{N} S_z^* g_\lambda \left(\frac{S_z^* S_z}{N} \right) S_z$ are equivalent. In the proof that follows let $u_k \in \mathbb{R}^N$ be the eigenvectors of $S_z S_z^*$ and $v_k \in H$ be the eigenvectors of $S_z^* S_z$. First, for $h_1 \in H$, we can write

$$S_z^* S_z h_1 = \sum_k \sigma_k^2 \left(v_k, h_1 \right)_H v_k,$$

and by the functional calculus we have

$$g_\lambda \left(\frac{S_z^* S_z}{N} \right) h_1 = \frac{1}{N} \sum_k g_\lambda \left(\sigma_k^2 \right) \left(v_k, h_1 \right)_H v_k.$$

When we choose for $h_1 = S_z^* S_z f$, we obtain

$$g_\lambda\left(\frac{S_z^* S_z}{N}\right) S_z^* S_z f = \frac{1}{N}\sum_k g_\lambda\left(\sigma_k^2\right)\left(v_k, S_z^* S_z f\right)_H v_k.$$

Also, since $S_z^* : \mathbb{R}^N \to H$, we have $S_z^* h_2 = \sum_k \sigma_k \left(u_k, h_2\right)_{\mathbb{R}^N} v_k$ for $h_2 \in H$. Then for $h_2 = g_\lambda\left(\frac{S_z S_z^*}{n}\right) S_z f$, we have

$$\begin{aligned}
S_z^* g_\lambda\left(\frac{S_z S_z^*}{N}\right) S_z f &= \frac{1}{N}\sum_k \sigma_k \left(u_k, \sum_j g_\lambda(\sigma_j^2)\left(u_j, S_z f\right)_{\mathbb{R}^N} u_j\right)_{\mathbb{R}^N} v_k \\
&= \frac{1}{N}\sum_k \sigma_k g_\lambda\left(\sigma_k^2\right)\left(\underbrace{u_k}_{\frac{1}{\sigma_k} S_z v_k}, S_z f\right)_{\mathbb{R}^N} v_k \\
&= \frac{1}{N}\sum_k g_\lambda\left(\sigma_k^2\right)\left(S_z v_k, S_z f\right)_{\mathbb{R}^N} v_k \\
&= \frac{1}{N}\sum_k g_\lambda\left(\sigma_k^2\right)\left(v_k, S_z^* S_z f\right)_H v_k
\end{aligned}$$

Then, we conclude that

$$g_\lambda\left(\frac{S_z^* S_z}{N}\right) S_z^* S_z f = S_z^* g_\lambda\left(\frac{S_z S_z^*}{N}\right) S_z f.$$

Finally, we note that

$$\begin{aligned}
\left(g_\lambda\left(T_z\right) T_z K_q, K_q\right)_H &= \frac{1}{N}\left(g_\lambda\left(\frac{1}{N} S_z S_z^*\right) S_z S_z^*, K_q\right)_H \\
&= \frac{1}{N}\Big(S_z^* g_\lambda\left(S_z^* S_z\right) S_z K_q, K_q\Big)_H \\
&= \frac{1}{N}\left(g_\lambda\left(\frac{\mathbb{K}}{N}\right) S_z K_q, S_z K_q\right)_{\mathbb{R}^N}.
\end{aligned}$$

4.8.3 Spectral Decomposition: Compact, Self-Adjoint Operators

In this section we review the fundamental properties of compact, self-adjoint operators and their spectral decompositions, which are used extensively in Section 4.3. We first review the construction of the singular value decomposition for matrices, and subsequently discuss the generalization to the Schmidt decomposition of compact operators acting between Hilbert spaces.

For any matrix $A \in \mathbb{R}^{m\times n}$, the symmetric and positive semidefinite matrix A^TA has a collection of real eigenvalues $\lambda_k := \lambda_k(A^TA)$ that can be arranged with multiplicities in nonincreasing order $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. When the corresponding eigenvectors $v_1, \ldots, v_n \in \mathbb{R}^{n\times n}$ are chosen to be of unit length and mutually orthogonal, we obtain the spectral factorization of A^TA in the form

$$A^TA = V\Sigma_n^2 V^T$$

where $\Sigma_n \in \mathbb{R}^{n\times n}$ is the diagonal matrix $\Sigma_n = \mathbf{diag}(\sigma_1, \ldots, \sigma_n)$ and $V \in \mathbb{R}^{n\times n}$ is an orthogonal matrix whose columns are the associated eigenvectors $V = [v_1, \ldots, v_n]$. If we carry out the same construction for the matrix $AA^T \in \mathbb{R}^{m\times m}$, we obtain

$$AA^T = U\Sigma_m^2 U^T$$

where $\Sigma_m^2 \in \mathbb{R}^{m\times m}$ is the diagonal matrix of non-increasing eigenvalues $\sigma_m := \sqrt{\lambda_m} := \sqrt{\lambda_m(AA^T)}$ and $U = [u_1, \ldots, u_m] \in \mathbb{R}^{m\times m}$ is an orthogonal matrix whose columns are the corresponding orthonormal eigenvectors of AA^T. It is a quick calculation to show that these decompositions are further related: we always have

$$u_k = \frac{1}{\sigma_k} A v_k$$

for $k = 1, \ldots, \mathbf{min}(m, n)$.

The singular value decomposition of the matrix A is then defined as

$$A = U\Sigma V^T$$

where $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal with entries $\sigma_1, \ldots, \sigma_{\mathbf{min}(m,n)}$ on the diagonal. It is convenient for comparison to the infinite dimensional operators to express these decompositions in terms of operators acting on arbitrary vectors. We have

$$\begin{aligned} Ax &= \sum_{k=1,\ldots,m} \sigma_k (v_k, x)_{\mathbb{R}^n} u_k, \\ AA^Ty &= \sum_{k=1,\ldots,m} \sigma_k^2 (u_k, y)_{\mathbb{R}^m} u_k, \\ A^TAx &= \sum_{k=1,\ldots,n} \sigma_k^2 (v_k, x)_{\mathbb{R}^n} v_k, \end{aligned}$$

for each $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

4.8.4 Functions of Operators

Suppose $A \in \mathbb{R}^{m\times n}, x \in \mathbb{R}^n$ and let

$$A = U\Sigma V^T$$

be the singular value decomposition. Then the matrix

$$T = AA^T \in \mathbb{R}^{m\times n}$$

is real, symmetric, and positive semi definite. It has the eigenvalue decomposition

$$T = U\Sigma_n^2 U^T = U\Lambda U^T = Udiag(\lambda_n)U^T$$

whose action can be written as

$$Ty = \sum_{i=1}^{n} \lambda_i (u_k, y)_{\mathbb{R}^n} u_k.$$

We usually define the matrix exponential by the power series

$$e^T := \sum_{k=0}^{\infty} \frac{T^k}{k!},$$

but a quick expansion of the series in terms of the eigenvector decomposition gives

$$e^T = Ue^{\Lambda}U^T.$$

We can rewrite the action of decomposition in in the form

$$e^T y = \sum_{k=1}^{n} e^{\lambda_k} (u_k, y)_{\mathbb{R}^n} u_k.$$

Now consider the case when U, V are Hilbert spaces and $A : V \to U$ is a compact operator. Since A^*A is a compact, self-adjoint operator, all of its eigenvalues are real and are contained in the interval $[0, \|A^*A\|]$. The number of eigenvalues greater than any given positive constant is finite, so the only possible accumulation point of the eigenvalues is zero. Each eigenspace corresponding to a nonzero eigenvalue is finite dimensional. The eigenvalues $\lambda_k(A^*A)$ are assumed to be arranged in an extended enumeration that includes multiplicities in nonincreasing order

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq 0.$$

The kth singular value $\sigma_k(A)$ of $A : V \to U$ is defined as

$$\sigma_k(A) := \lambda_k^{1/2}(A^*A).$$

There exist orthonormal collections of eigenvectors $\{u_k\}_{k=1}^{\infty}$ for AA^* and $\{v_k\}_{k=1}^{\infty}$ for A^*A such that

$$Af = \sum_{k=1}^{\infty} \sigma_k (v_k, f)_V u_k \qquad \textbf{convergence in } U,$$

$$AA^*g = \sum_{k=1}^{\infty} \sigma_k^2 (u_k, g)_U u_k \qquad \textbf{convergence in } U,$$

$$A^*Af = \sum_{k=1}^{\infty} \sigma_k^2 (v_k, f)_V v_k \qquad \textbf{convergence in } V,$$

for each $f \in V$ and $g \in U$.

It turns out that for any self-adjoint compact operator $T : H \to H$, the function $g(T)$ can be defined in an analogous way

$$(g(T))(f) := \sum_{k=1}^{\infty} g(\lambda_i)(u_k, f)_H u_k.$$

4.8.5 Approximation Spaces

Our error estimates for the empirical potential functions used in our discrete evolution laws will be based on either linear or nonlinear approximation methods as they arise in the construction of approximations spaces. See [29], or more recently [7], for a thorough theoretical discussion of these spaces. An approximation method is a pair $(X, \{X_n\}_{n=0}^{\infty})$ where X is a (quasi-)Banach space and $\{X_n\}$ is a sequence of subsets of X that satisfy the following conditions:

(i) $\{0\} = X_0 \subseteq X_1 \subseteq X_2 \subseteq \cdots \subseteq X$**,**

(ii) $aX_n \subset X_n$ **for all** $a \in \mathbb{R}$ **and** $n = 1, 2, \ldots$**,**

(iii) $X_m + X_n \subset X_m + X_n$ **for** $m, n = 1, \ldots$**,**

When we define the nth approximation number as

$$a_n(f, X) := E_{n-1}(f, X) := \inf_{a \in X_{n-1}} \|f - a\|_X \tag{4.38}$$

for each $n > 0$, the approximation space X_u^r is defined to be the collection of $f \in X$ for which the sequence $\left\{n^{r-1/u} a_n(f, X)\right\}_n \in \ell_u$. The approximation space X_u^r is a (quasi-)Banach space when we set

$$\|f\|_{X_u^r} := \left\|\left\{n^{r-1/u} a_n(f, X)\right\}_n\right\|_{\ell_u} = \left(\sum_{n=1}^{\infty} \left[n^r E_{n-1}(f, X)\right]^u \frac{1}{n}\right)^{1/u}.$$

Note that since the first term in the sequence above is $\|f\|_X$, so that if $\|f\|_{X_u^r} = 0$, we have $f = 0$. We obtain a (quasi-)seminorm $|f|_{X_u^r}$ on the space X_u^r if we omit the first term in the sequence above. Many of the basic properties of approximation spaces are described in [29, 30, 7].

One of the important properties that we will use in this paper is a theorem that guarantees the representation of elements in an approximation space in terms of quasi-geometric sequences. A sequence of integers $\{n_k\}_{k=0}^{\infty}$ is said to be quasi-geometric if $n_0 = 1$ and there are two positive constants a, b such that

$$1 < a \leq \frac{n_{k+1}}{n_k} \leq b < \infty$$

for all $k \geq 0$. The representation theorem [30] is stated below:

Proposition 1 ***Let*** $\{n_k\}_{k=0}^{\infty}$ ***be a quasi-geometric series. The expression***

$$\|f\|_{X_u^r} \approx \| \{n_k^r a_{n_k}\}_k \|_{\ell_u}$$

defines an equivalent quasi-norm on X_u^r***. A function*** $f \in X$ ***belongs to*** X_u^r ***if and only if there exists a representation of the form***

$$f = \sum_{k=0}^{\infty} x_{n_k}$$

such that $x_{n_k} \in X_{n_k}$ ***and***

$$\{n_k^r \|x_{n_k}\|_X\}_k \in \ell_u.$$

In this case we have

$$\|f\|_{X_u^r} \approx \inf_{\substack{f=\sum_k x_{n_k} \\ x_{n_k} \in X_{n_k}}} \left\| \{n_k^r x_{n_k}\}_k \right\|_{\ell_u}.$$

This last characterization of the approximation spaces X_u^r sometimes are presented in slightly different forms. As noted in [7], we can derive an equivalent expression for the seminorm $|f|_{X_u^r}$ choosing the geometric sequence where $n_k \equiv 2^k$ for $k = 0, 1, \ldots$.

$$|f|_{X_u^r} = \left(\sum_{n=0}^{\infty} [2^{nr} E_{2^n}(f, X)]^q \right)^{1/q}.$$

Approximation spaces have been studied for a wide range of choices of the underlying space X and sequence of approximating subsets $\{X_n\}$. In this paper we will restrict consideration to approximation spaces X is in fact a Hilbert space and the subsets $\{X_n\}$ are defined in terms of an orthonormal basis for X. See [19] Section (2) for discussion of this classical choice, where it is compared to the more general cases in which the family is generated from an unconditional or greedy basis for a (quasi-)Banach space X.

Linear Approximation Methods

Suppose that $\{\phi_k\}_{k=1}^{\infty}$ is an orthonormal basis for the Hilbert space X. A linear approximation method chooses the sets $X_n := \text{span}\,\{\phi_k\}_{k \le n}$. In this case we have

$$a_n = \inf_{g \in X_{n-1}} \|f - g\|_X = \|(I - \Pi_{n-1})f\|_X = E_{n-1}(f, X)$$

where Π_n is the orthogonal projection onto X_n. We always have

$$f = \sum_{k=0}^{\infty} \Phi_k f := \sum_k (\Pi_{2^k} - \Pi_{2^{k-1}}) f$$

when we define $\Pi_{-1} := 0$. It follows that we have the equivalent norm on X_u^r

$$\|f\|_{X_u^r} \approx \left\| \left\{ 2^{kr} \|\Phi_k f\|_X \right\}_k \right\|_{\ell_u}$$

In the special case that $u = 2$, it can be shown that $f \in X_2^r$ if and only if there is a constant $C > 0$ such that

$$\sum_{k=1}^{\infty} k^{2r} |(f, \phi_k)|^2 \leq C$$

In this case we have

$$|f|_{X_2^r} \approx \sum_{m=1}^{\infty} 2^{2mr} \sum_{k=2^{m-1}+1}^{2^m} |(f, \phi_k)|^2.$$

Proof 11 ***We claim that***

$$\sum_{k=1}^{\infty} k^{2r} |(f, \phi_k)|^2 \approx \sum_{m=1}^{\infty} 2^{2mr} \sum_{k=2^{m-1}+1}^{2^m} |(f, \phi_k)|^2.$$

Simply by re-grouping terms we see that

$$\begin{aligned}
\sum_{k=1}^{\infty} k^{2r} |(f, \phi_k)|^2 &= \sum_{m=1}^{\infty} \sum_{k=2^{m-1}+1}^{2^m} k^{2r} |(f, \phi_k)|^2, \\
&\leq \sum_{m=1}^{\infty} \sum_{k=2^{m-1}+1}^{2^m} (2^{mr})^2 |(f, \phi_k)|^2, \\
&\leq \sum_{m=1}^{\infty} 2^{2mr} \sum_{k=2^{m-1}+1}^{2^m} |(f, \phi_k)|^2.
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\sum_{m=1}^{\infty} \sum_{k=2^{m-1}+1}^{2^m} 2^{2mr} |(f, \phi_k)|^2 &\leq \sum_{m=1}^{\infty} \sum_{k=2^{m-1}+1}^{2^m} 2^{2r} 2^{2(m-1)r} |(f, \phi_k)|^2, \\
&\leq 2^{2r} \sum_{m=1}^{\infty} \sum_{k=2^{m-1}+1}^{2^m} k^{2r} |(f, \phi_k)|^2.
\end{aligned}$$

Bibliography

[1] Frank Bauer and Sergi Pereverzev and Lorenzo Rosasco, "On Regularization Algorithms in Learning Theory", Journal of Complexity, 23(1), pp. 52-72, 2007.

[2] Matt J Bender, Hunter G McClelland, Gerardo Bledt, Andrew J Kurdila, Tomonari Furukawa, and Rolf Mueller. "Trajectory estimation of bat flight using a multi-view camera system", In AIAA SciTech, page 1806, Kissimmee, FL, USA, January 2015. AIAA.

[3] Anthony M. Bloch. " Nonholonomic Mechanics and Control. Interdisciplinary Applied Mathematics." Springer, New York, NY, USA, 2003.

[4] Francesco Bullo and Andrew Lewis. "Geometric Control of Mechanical Systems". Springer, 2004.

[5] Peter R Cavanagh and Mario A Lafortune. Ground reaction forces in distance running. Journal of Biomechanics, 13(5):397406, June 1980.

[6] Ronald DeVore, Gerard Kerkyacharian, Dominique Picard, and Vladimir Temlyakov. Approxmation methods for supervised learning. Foundations of Computational Mechanics, 6(1):358, February 2006.

[7] Ronald A. DeVore and George G. Lorentz. Constructive Approximation, volume 303 of A Series of Comprehensive Studies in Mathematics. Springer-Verlag, 1993.

[8] Carl Henrick Ek, Phillip H. S. Torr, and Neil D. Lawrence. Gaussian process latent variable models for human pose estimation. In Andrei Popescu-Belis, Steve Renals, and Herve Bourlard, editors, Machine Learning for Multimodal Interaction, 4th Workshop, volume 4892 of LNCS, pages 132143, Brno, Czech Repblic, June 2007.

[9] Guoliang Fan, Xin Zhang, and Meng Ding. Gaussian processes for human motion modeling: A comparative study. In IEEE International Workshop on Machine Learning for Signal Processing, pages 1821, Beijing, China, 2011. IEEE.

[10] Hartmut Geyer and Hugh Herr. A muscle-reflex model that encodes principles of legged mechanics produces human walking dynamics and muscle activities. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 18(3):263273, June 2010.

[11] Ernst Hairer, Christian Lubich, and Gerhard Wanner, "Geometric Numerical Integration": StructurePreserving Algorithms for Ordinary Differential Equations. Springer, 2nd edition, 2006.

[12] James Hall and Melvin Leok. "Spectral variational integrators". Numerische Mathematik, pages 160, December 2014.

[13] Søren Hauberg, Francois Lauze, and Kim Steenstrup Pedersen. Unscented kalman filtering on riemannian manifolds. Journal of Mathematical Imaging and Vision, 46(1):103120, August 2013.

[14] Søren Hauberg and Kim Steenstrup Pedersen. Predicting articulated human motion from spatial processes. International Journal of Computer Vision, 94(3):317334, March 2011.

[15] Søren Hauberg, Stefan Sommer, and Kim Steenstrup Pedersen. Natural metrics and least-committed priors for articulated tracking. Image and Vision Computing, 30(6):453461, November 2012.

[16] Anders Hedenström, LC Johansson, Marta Wolf, R Von Busse, Y Winter, and Geoffrey R. Spedding. Bat flight generates complex aerodynamic tracks. Science, 316(5826):894897, May 2007.

[17] Tatjana Y Hubel, Nickolay I Hristov, Sharon M Swartz, and Kenneth S Breuer. Time-resolved wake structure and kinematics of bat flight. Experiments in Fluids, 46(5):933943, February 2009.

[18] Hyuk Kang and Frank C. Park. Humanoid motion optimization via nonlinear dimension reduction. In International Conference on Robotics and Automation, pages 14441449, St. Paul, Minnesota, USA, May 2012. IEEE.

[19] Gerard Kérkyacharian and Dominique Picard. Entropy, universal coding, approximation, and bases properties. Constructive Approximation, 20(1):137, November 2003.

[20] Oussama Khatib. Real-time obstacle avoidance for manipulators and mobile robots. International Journal of Robotics Research, 5(1):9098, March 1986.

[21] Jonathan Ko and Dieter Fox. Learning GP-BayesFilters via gaussian process latent variable models. Autonomous Robots, 30(1):323, January 2011.

[22] Yoram Koren and Johann Borenstein. Potential field methods and their inherent limitations for mobile robot navigation. In International Conference on Robotics and Automation, volume 2, pages 13981404, Sacramento, CA, USA, April 1991. IEEE.

[23] B. Krogh. A generalized potential field approach to obstacle avoidance control. In SME Conf. Proc. Robotics Research: The Next Five Years and Beyond, Bethlehem, Pennsylvania, August 1984.

[24] Neil D Lawrence. Gaussian process latent variable models for visualisation of high dimensional data. Advances in Neural Information Processing Systems, 16:329336, 2004.

[25] Jerrold E Marsden and Matthew West. Discrete mechanics and variational integrators. Acta Numerica, 10:357514, May 2001.

[26] Carolyn F Munro, Doris I Miller, and Andrew J Fuglevand. Ground reaction forces in running: A reexamination. Journal of Biomechanics, 20(2):147155, 1987.

[27] W. Newman and Neville Hogan. High speed robot control and obstacle avoidance using dynamic potential functions. In International Conference on Robotics and Automation, volume 4, pages 1424. IEEE, March 1987.

[28] F. C. Park and M. W. Kim. Lie theory, riemannian geometry, and the dynamics of coupled rigid bodies. Zeitschrift f ur angewandte Mathematik und Physik, 51(5):820834, September 2000.

[29] Albrecht Pietsch. Approximation spaces. Journal of Approximation Theory, 32(2):115134, June 1981.

[30] Pietsch, A. Arch, Tensor products of sequences, functions, and operators, Math (1982) 38: 335.

[31] Igor V Pivkin, E Hueso, R Weinstein, David H Laidlaw, Sharon Swartz, and G E Karniadakis. Simulation and visualization of air flow around bat wings during flight. In Vaidy S Sunderam, GeertDick van Albada, Peter M. A. Sloot, and Jack J Dongarra, editors, Computational Science ICCS 2005, volume 3515 of Lecture Notes in Computer Science, pages 689694. Springer Berlin Heidelberg, 2005.

[32] Shi Qu, Ronghuan Yu, Yingmei Wei, and Lingda Wu. Gaussian process latent variable models for inverse kinematics. Journal of Multimedia, 6(1):4855, February 2011.

[33] Elon Rimon and Daniel E. Koditschek. Exact robot navigation using artificial potential functions. IEEE Transactions on Robotics and Automation, 8(5):501518, October 1992.

[34] Daniel K. Riskin, David J. Willis, José Iriarte-Díaz, Tyson L. Hedrick, Mykhaylo Kostandov, Jian Chen, David H. Laidlaw, Kenneth S. Breuer, and Sharon M. Swartz. Quantifying the complexity of bat wing kinematics. Journal of Theoretical Biology, 254(3):604615, October 2008.

[35] **Steve Smale and Ding-Xuan Zhou. Learning theory estimates via integral operators and their approximations. Constructive Approximation, 26(2):153172, August 2007.**

[36] **Ingo Steinwart and Clint Scovel. Mercers theorem on general domains: On the interaction between measures, kernels, and RKHSs. Constructive Approximation, 35(3):363417, June 2012.**

[37] **Sharon M Swartz, Jose Iriarte-Diaz, Daniel K Riskin, Arnold Song, Xiaodong Tian, David J Willis, and Kenneth S Breuer. Wing structure and the aerodynamic basis of flight in bats. In 45th AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, January 2007. AIAA.**

[38] **Ernesto De Vito, Lorenzo Rosasco, and Alessandro Toigo. Learning sets with separating kernels. Applied and Computational Harmonic Analysis, 37(2):185217, April 2014.**

[39] **Joachim Weidmann. Linear Operations in Hilbert Spaces. Springer, New York, NY, USA, 1980.**

[40] **Christian Lubich and Daniel Weiss, "Numerical Integrators for Motion under a Strong Constraining Force," *Multiscale Model. Simul.*, 12(4), pp. 15921606, 2014.**

[41] **Andrew J. Kurdila and Francis Narcowich, "Sufficient conditions for penalty formulation methods in analytical dynamics", *Computational Mechanics*, January 1993, Volume 12, Issue 1, pp 8196.**

[42] **Javier Garca de Jaln and Eduardo Bayo, " Kinematic and dynamic simulation of multibody systems: the real-time challenge", Springer-Verlag, 1994.**

[43] **Hanan Rubin and Peter Ungar, " Motion Under a Strong Constraining Force", *Communications on Pure and Applied Mathematics*, Volume 10, Issue 1, Pages 6587, 1957**

Chapter 5

Summary and Conclusions

This dissertation addresses two challenging problems that arise in modeling of nonlinear flapping wing robotic systems. In the first problem, we propose and derive a class of history dependent models suitable for representation of aerodynamic loads during flapping flight. A set of weighted, relatively low dimensional, nonlinear history dependent operators has been employed to represent history dependent, unsteady aerodynamics of flapping flight. We also establish existence and uniquness of a local solution for such functional differential equations that feature history dependent aerodynamics. A novel adaptive control strategy has also been proposed to learn and compensate for the unknown history dependent aerodynamics. Lyapunov stability analysis guarantees stability of the closed loop system. It is also shown that the tracking error and its derivative converge to zero.

We also derive approximation methods and construct error bounds in approximation space for history dependent operators. We use the results of wavelet theory and multiresolution analysis to compute an error bound as a function of resolution level of the approximation space. We then propose sliding mode adaptive control strategies to identify and compensate for the unknown history dependent aerodynamic forces, and we validate our algorithm by numerical simulations.

In the second problem, we derive methods that inherently yield dynamics that lie, at least approximately, on or near observed subsets or low dimensional submanifolds. We derive three theorems to establish a general framework for creating analytical-empirical equations of discrete mechanics. Our main proposed algorithm uses experimental observations to construct discrete evolution laws that evolve on or near such subsets or low dimensional submanifolds. As part of our theory, we also derive a new single step error bound for the discrete Galerkin Lagrangian and the exact discrete Lagrangian. This bound consists of a term that is common to conventional variational discrete Galerkin methods, but it additionally includes a probabilistic term that arises from the empirical potential function that has been constructed using the camera observations. We have implemented simulations using noisy data generated in the vicinity of a submanifold of Euclidean space, featuring motion of a single constrained particle on a thin wire. The results show that the motion stays arbitrarily close to the submanifold of interest, and the motion reflects the geometry of the underlying submanifold. We discuss the theoretical ground behind this phenomena and discuss in detail an open problem for future directions of this research.