

A Fully Discrete Adjoint Method for Optimization of Flow Problems on Deforming Domains with Time-Periodicity Constraints

M. J. Zahr^{a,1,*}, P.-O. Persson^{b,2}, J. Wilkening^{b,2}

^a*Institute for Computational and Mathematical Engineering, Stanford University, Stanford, CA 94035.*

^b*Department of Mathematics and Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720-3840.*

Abstract

A variety of shooting methods for computing fully discrete time-periodic solutions of partial differential equations, including Newton-Krylov and optimization-based methods, are discussed and used to determine the periodic, compressible, viscous flow around a 2D flapping airfoil. The Newton-Krylov method uses matrix-free GMRES to solve the linear systems of equations that arise in the nonlinear iterations, with matrix-vector products computed via the linearized sensitivity evolution equations. The adjoint method is used to compute gradients for the gradient-based optimization shooting methods. The Newton-Krylov method is shown to exhibit superior convergence to the optimal solution for these fluid problems, and fully leverages quality starting data.

The central contribution of this work is the derivation of the adjoint equations and the corresponding adjoint method for fully discrete, time-periodically constrained partial differential equations. These adjoint equations constitute a linear, two-point boundary value problem that is provably solvable. The periodic adjoint method is used to compute gradients of quantities of interest along the manifold of time-periodic solutions of the discrete partial differential equation, which is verified against a second-order finite difference approximation. These gradients are then used in a gradient-based optimization framework to determine the energetically optimal flapping motion of a 2D airfoil in compressible, viscous flow over a single cycle, such that the time-averaged thrust is identically zero. In less than 20 optimization iterations, the flapping energy was reduced nearly an order of magnitude and the thrust constraint satisfied to 5 digits of accuracy.

1. Introduction

Cyclic steady-state motion of a system, i.e. stable time-periodic behavior, is of central importance in bio-locomotion and many branches of engineering. Examples include flapping flight [1, 2, 3], swimming at low or high Reynolds number [4, 5, 6, 7], helicopter aerodynamics [8, 9, 10, 11], turbomachinery [12, 13], wind turbines [14, 15] and vehicle tires with treads [16, 17, 18], to name a few. A number of sophisticated algorithms have recently been developed to compute cyclic steady-states of systems governed by partial differential equations. However, in applications, one often wishes to optimize a quantity of interest over a cycle, such as minimizing energy subject to lift and thrust constraints. The goal of the present paper is to develop adjoint-based optimization techniques for such systems, focusing on the challenges that arise due to time-periodicity constraints.

A prerequisite to optimization is being able to accurately compute time-periodic solutions. For low-dimensional systems (including time-dependent PDEs with only one spatial dimension), orthogonal collocation and (temporal) Fourier collocation algorithms such as implemented in the software package AUTO [19, 20] have proven to be robust and widely applicable [21, 22, 23, 24, 25]. In the aerodynamic optimization

*Corresponding author

Email addresses: `mzahr@stanford.edu` (M. J. Zahr), `persson@berkeley.edu` (P.-O. Persson), `wilken@math.berkeley.edu` (J. Wilkening)

¹Graduate Student, Institute for Computational and Mathematical Engineering, Stanford University

²Associate Professor, Department of Mathematics, University of California, Berkeley.

community, these methods are known as harmonic balance [26], time spectral [27], or nonlinear frequency [28, 29] techniques. While these types of approaches can realize spectral convergence in time, they quickly lead to extremely large-scale computations as all time instances become coupled and the unknown state vector includes all spatial degrees of freedom at every collocation point, i.e., a tensor product between space and time. At the other extreme, shooting methods [30, 31] treat only the initial conditions as unknowns and use a numerical timestepping scheme to determine the state of the system at later times. These methods are very effective for computing stable or nearly stable time-periodic solutions in which nearby trajectories do not diverge wildly from each other over the timescale of the periodic solution and have a much smaller memory footprint than the spectral collocation approaches. Examples of systems exhibiting this non-chaotic behavior include mode-locked lasers [32], water waves [33, 34, 35], viscoelastic fluid flows [36], and rolling vehicle tires [17]. For chaotic systems, multiple-shooting methods [31] strike a balance between the robustness of a temporal collocation method and the efficiency of a shooting method for limiting the number of unknowns.

Shooting methods can further be classified by the method used to solve for the unknown initial conditions. If the periodic solution is stable with fairly large decay rates relative to the period of the solution, a simple and effective method is fixed point iteration. However, if a high degree of accuracy is desired, the slowest decaying modes often obstruct convergence in a reasonable amount of simulation time. For example, Thomases and Shelley observed “persistent oscillations” in a Stokesian viscoelastic fluid for which “a simulation up to $t = 10000$ reveals no decrease in their amplitude,” but could not conclude for certain that the limiting oscillations were time-periodic. Lust and Roose [37] devised a hybrid method in which some of the degrees of freedom are solved for by a shooting method while others converge via fixed point iteration. This works well if the number of unstable or mildly stable modes is small. For neutrally stable problems such as computing standing water waves, fixed point iteration cannot be used to solve for any of the modes, and a genuinely large-scale nonlinear solver has to be used. Mercer and Roberts developed a Newton-Raphson algorithm for computing large-amplitude standing water waves [38]. Ambrose and Wilkening devised an adjoint-based minimization algorithm for such problems based on the limited memory BFGS algorithm [39, 40, 41, 42]. This method was also used by Williams et. al. for mode-locked lasers [32], and by Isaacson for computing time-periodic solutions of the Oldroyd-B equations [36]. Wilkening and Yu [33, 34] later developed an overdetermined shooting method based on the Levenberg-Marquardt method that takes advantage of consolidation of work by computing multiple columns of the Jacobian matrix in parallel. This approach was also used by Wilkening and Rycroft for computing standing water waves in 3D. For studying transitions to turbulence in Couette flow [43, 44, 45] or pipe flow, Viswanath developed a Newton-Krylov algorithm with a locally constrained optimal hook step [46]. Similar work was done by Schneider and collaborators in the context of turbulent pipe flow [47, 48]. Building on these ideas, Govindjee, Potter and Wilkening developed a Newton-Krylov approach for computing cyclic steady states in rolling vehicle tires with treads [17].

Here we also adopt a Newton-Krylov approach, but we bring back adjoint methods to optimize various quantities of interest over a cycle of the (parameter-dependent) time-periodic solution. Our motivation comes from the goal of designing an energetically optimal flapping motion of a 2D airfoil in a compressible, viscous fluid subject to a time-averaged thrust constraint. Early methods toward computing energetically optimal flapping flight used derivative-free optimization solvers [49] and finite differences [50] or the sensitivity method to compute gradients for first-order optimizers [3]. The derivative-free approach is limited to small parameter spaces and coarse flow discretizations due to the large number of iterations required by such solvers. The finite difference and sensitivity method also require small parameter spaces as each entry of the gradient requires the solution of a nonlinear or linearized, forward evolution equation, respectively.

A number of fully discrete, time-dependent, adjoint-based methods have recently been introduced [51, 52, 53, 54] that enable high-dimensional parameter spaces to be efficiently searched, which leads to improved design and control. Continuous adjoint methods have also been introduced for the same purpose [55, 56], although they may result in inexact gradients and slow optimization convergence if the spatial discretization employed is not adjoint consistent. The method presented here improves on these existing adjoint-based methods by incorporating time-periodicity constraints that ensure a *representative, in-flight* cycle is considered during the optimization. The existing methods initialize the flow from the steady-state or uniform

flow and run several cycles to let non-physical transients die out. These approaches have the disadvantage of requiring many cycles to fully suppress the initial transients and will substantially increase the cost of each flow simulation. Stanford and Beran [57] – who considered structural optimization of a dry flapping wing, i.e., without a surrounding fluid – implicitly incorporated time-periodicity constraints by using a spectral time discretization that only supports time-periodic solutions. This approach, referred to as temporal Fourier collocation above, has the disadvantage of coupling all timesteps, which was mitigated in that work by considering the *dry* structure and using proper orthogonal decomposition-based model reduction methods.

The remainder of the paper is organized as follows. Section 2 discusses the numerical discretization of partial differential equations and reviews Newton-Krylov and optimization-based shooting methods for computing time-periodic solutions. A discussion on the stability of periodic orbits is included. The fully discrete framework is emphasized throughout as this will lead naturally to the corresponding fully discrete adjoint equations in Section 3. The derivation of the adjoint equations corresponding to time-periodically constrained fully discrete partial differential equations is provided in Section 3.1. These equations constitute a linear, two-point boundary value problem; existence and uniqueness of solutions is proved in Appendix A. The corresponding adjoint method for computing gradients of quantities of interest along the manifold of solutions of the time-periodically constrained, fully discrete partial differential equations is also provided in Section 3.1. A matrix-free Krylov shooting method is introduced in Section 3.2 for computing solutions of the periodic adjoint equations. In Section 4.1, the various primal and dual shooting methods are compared side-by-side on a flapping airfoil in compressible, viscous flow. Finally, in Section 4.2, the primal and dual shooting methods are used to compute optimization functionals and gradients, respectively, to determine the energetically optimal flapping motion in compressible, viscous flow, subject to a constraint on the time-averaged thrust.

2. Computing Time-Periodic Solutions of Partial Differential Equations

This section is devoted to the discretization and solution of partial differential equations with time-periodicity constraints. This will largely be a review of existing work on the topic [38, 46, 40, 42, 34, 17], although emphasis will be placed on equations that are *parametrized*. This will lead to the main contribution of this work, the fully discrete adjoint equations corresponding to time-periodic solutions of partial differential equations and their use in computing gradients of quantities of interest along the manifold of time-periodic solutions.

Consider the general, nonlinear, time-periodically constrained system of partial differential equations, parametrized by the vector $\boldsymbol{\mu} \in \mathbb{R}^{N_\mu}$,

$$\begin{aligned} \frac{\partial \mathbf{U}}{\partial t} &= \mathcal{L}(\mathbf{U}, \boldsymbol{\mu}, t) & \text{in } \Omega(\boldsymbol{\mu}, t) \times (0, T] \\ \mathbf{U}(\mathbf{x}, 0) &= \mathbf{U}(\mathbf{x}, T), \end{aligned} \tag{1}$$

where $\mathcal{L}(\cdot, \boldsymbol{\mu}, t)$ is a spatial differential operator on the parametrized, time-dependent domain $\Omega(\boldsymbol{\mu}, t) \subset \mathbb{R}^{n_{sd}}$. The boundary conditions have not been explicitly stated for brevity. This work will only consider temporally first-order partial differential equations, or those that have been recast as such. Without loss of generality, consider a quantity of interest of the form

$$\mathcal{F}(\mathbf{U}, \boldsymbol{\mu}) = \int_0^T \int_{\Gamma(\boldsymbol{\mu}, t)} f(\mathbf{U}, \boldsymbol{\mu}, t) dS dt, \tag{2}$$

where $\Gamma(\boldsymbol{\mu}, t) \subseteq \partial\Omega(\boldsymbol{\mu}, t)$. The generalization to other types of quantities of interest, such as volumetric integrals and instantaneous or pointwise quantities of interest, is immediate as the specific form of the quantity of interest will be abstracted away at the fully discrete level. The form in (2) will be used in the physical setup of the applications in Section 4. In subsequent sections, this quantity of interest will correspond to either the objective function or a constraint of an optimization problem governed by a partial differential equation and subject to a time-periodicity requirement. The remainder of this section will be concerned with the numerical discretization and solution of (1).

2.1. Numerical Discretization

As the form of the spatial differential operator in the time-periodically constrained system of partial differential equations in (1) was not specified, an unspecified semi-discretization is applied to yield a system of ordinary differential equations

$$\begin{aligned} \mathbb{M} \frac{\partial \mathbf{u}}{\partial t} &= \mathbf{r}(\mathbf{u}, \boldsymbol{\mu}, t) \\ \mathbf{u}(0) &= \mathbf{u}(T), \end{aligned} \quad (3)$$

where $\mathbf{u}(t) \in \mathbb{R}^{N_u}$ is the state vector of the semi-discretization, \mathbf{r} is the spatial discretization of \mathcal{L} , and \mathbb{M} is the mass matrix arising from the discretization of $\frac{\partial \mathcal{U}}{\partial t}$.

In Section 4, a high-order discontinuous Galerkin method is used to semi-discretize the compressible Navier-Stokes equations. In general, a partial differential equation with a parametrized, time-dependent spatial domain will lead to ordinary differential equations with a parametrized, time-dependent mass matrix. However, Section 4 will discuss an Arbitrary-Lagrangian-Eulerian formulation of general systems of conservation laws that solves a transformed set of equations on a *fixed* domain, leading to a constant mass matrix. Therefore, attention will be restricted to the case of a fixed mass matrix.

As the motivating applications for this work are viscous fluid dynamics problems, an implicit temporal discretization is used since timestep sizes tend to be stability-limited. Specifically, Diagonally Implicit Runge-Kutta (DIRK) schemes – Runge-Kutta schemes with a lower triangular Butcher tableau – are used since stable, high-order discretizations are possible without incurring the large cost of coupling all stages. An s -stage DIRK discretization of (3) leads to the fully discrete, nonlinear evolution equations

$$\begin{aligned} \mathbf{u}^{(n)} &= \mathbf{u}^{(n-1)} + \sum_{i=1}^s b_i \mathbf{k}_i^{(n)} \\ \mathbb{M} \mathbf{k}_i^{(n)} &= \Delta t_n \mathbf{r} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right), \end{aligned} \quad (4)$$

where

$$\mathbf{u}_i^{(n)} = \mathbf{u}^{(n-1)} + \sum_{j=1}^i a_{ij} \mathbf{k}_j^{(n)}. \quad (5)$$

From (4), it is clear that each timestep requires the solution of s nonlinear systems of equations of size N_u . Time-periodicity may then be expressed as the constraint

$$\mathbf{u}^{(0)} = \mathbf{u}^{(N_t)}, \quad (6)$$

where N_t is the time index of the cycle period.

While the form of the fully discrete time-periodically constrained partial differential equations are specific to a DIRK temporal discretization, this is not a fundamental restriction of this work. Extension of the analysis and derivations in Sections 2.2 and 3 to other classes of temporal discretizations, whether they are implicit or explicit, is possible.

With the full discretization of the partial differential equation addressed, the quantity of interest, \mathcal{F} , must also be discretized to be computable. While any numerical quadrature formula can be used to perform the discretization of the space-time integral, this may lead to a truncation error of different orders for the governing equations and the quantity of interest. This implies there is wasted effort since the largest order will dominate. A solver-consistent discretization of the quantities of interest [54] – where the spatial and temporal discretization of the partial differential equation are also used for the quantity of interest – can be used to circumvent this wasted effort. In this work, the discontinuous Galerkin shape functions will be used to discretize the spatial surface integral and the DIRK scheme to discretize the temporal integral. Regardless of the discretization chosen for the quantity of interest, the fully discrete version takes the form

$$F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}). \quad (7)$$

Note that the aforementioned solver-consistent discretization leads to the dependence of the fully discrete quantity of interest on the Runge-Kutta stages, $\mathbf{k}_i^{(n)}$, combined in such a way that the expected truncation order is attained, see [54] for details. A standard numerical quadrature, such as midpoint rule or Simpson's rule, would *not* use the stages, which are only low-order solutions of (3) [58].

With the full numerical discretization of the system of partial differential equations complete, the next section will discuss methods for solving the fully discrete, time-periodically constrained partial differential equations. The periodicity constraint, i.e. $\mathbf{u}^{(0)} = \mathbf{u}^{(N_t)}$, turns the problem into a nonlinear two-point boundary value problem, which eliminates the possibility of using traditional evolution methods (since the initial conditions are unknown).

2.2. Numerical Solvers: Shooting Methods

This section provides a brief, non-exhaustive review of methods which have been introduced for solving time-periodic partial differential equations. A distinguishing feature of this work is that we directly consider the fully discrete form of the governing equations, whereas previous work has focused on the continuous [59] or semi-discrete [60] levels. The section will conclude with a discussion of a Newton-Krylov shooting method using a purely matrix-free Krylov solver to solve the linear systems of equations that arise, which extends the work in [17].

Define $\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu})$ as the solution of the following initial-value problem

$$\begin{aligned} \mathbf{u}^{(0)} &= \mathbf{u}_0 \\ \mathbf{u}^{(n)} &= \mathbf{u}^{(n-1)} + \sum_{i=1}^s b_i \mathbf{k}_i^{(n)} \\ \mathbb{M} \mathbf{k}_i^{(n)} &= \Delta t_n \mathbf{r} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right), \end{aligned} \quad (8)$$

which can be solved using a traditional evolution algorithm that advances the solution from timestep n to $n + 1$. Notice that this overloads the notation introduced in Section 2.1, which defines $\mathbf{u}^{(N_t)}$ as the discrete approximation of the time-periodic solution of the system of partial differential equations at the final time. Here, it is a nonlinear function that maps a state $\mathbf{u}_0 \in \mathbb{R}^{N_u}$ to the state $\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu})$. From (4) and (8), it is clear that \mathbf{u}_0 is the time-periodic initial condition of the fully discrete partial differential equation, $\mathbf{u}^{(0)}$, if it is a fixed point of $\mathbf{u}^{(N_t)}(\cdot; \boldsymbol{\mu})$, namely

$$\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) = \mathbf{u}_0. \quad (9)$$

Then, provided the mapping $\mathbf{u}_0 \rightarrow \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu})$ is a contraction mapping, the Banach Fixed Point Theorem implies the existence of the fixed point and provides a convergent algorithm for finding it, see Algorithm 1. This is a convenient algorithm as it only relies on solution of the nonlinear evolution equation (8), but is known to suffer from poor convergence rates and lack of convergence if the mapping under consideration is not a contraction.

Algorithm 1 Fixed Point Iteration Time-Periodic Solutions of PDE

Input: Initial guess for periodic initial condition, \mathbf{u}_0 ; parameter configuration, $\boldsymbol{\mu}$

Output: Periodic initial condition, $\mathbf{u}^{(0)}$

- 1: **while** $\|\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{u}_0\|_2 > \epsilon$ **do**
- 2: Update

$$\mathbf{u}_0 \leftarrow \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu})$$

- 3: **end while**

- 4: Define periodic initial condition

$$\mathbf{u}^{(0)} = \mathbf{u}_0$$

Another class of solvers for time-periodically constrained partial differential equations rely on unconstrained, gradient-based optimization techniques. Define the function

$$j(\mathbf{u}_0) = \frac{1}{2} \|\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{u}_0\|_2^2 \quad (10)$$

and consider the unconstrained optimization problem

$$\underset{\mathbf{u}_0 \in \mathbb{R}^{N_u}}{\text{minimize}} \quad j(\mathbf{u}_0), \quad (11)$$

which can be solved using gradient-based optimization techniques such as steepest descent, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, or its limited-memory version, L-BFGS [61, 62, 63]. The gradient of (10), $\frac{dj}{d\mathbf{u}_0}$, is usually computed using the adjoint method since the large number of optimization variables, N_u , renders the finite differences method or the linearized forward method impractical [64]. Throughout this work, the notation $\frac{d(\cdot)}{d\boldsymbol{\mu}}$ will be used to denote the total derivative of a quantity of interest with respect to parameters – including the explicit dependence as well as the implicit dependence through the solution of the governing equation – and the partial derivative notation $\frac{\partial(\cdot)}{\partial\boldsymbol{\mu}}$ will be used elsewhere. The adjoint equations for the fully discrete evolution equations in (8) corresponding to the quantity of interest, $j(\mathbf{u}_0)$, with parameter \mathbf{u}_0 are

$$\begin{aligned} \boldsymbol{\lambda}^{(N_t)} &= \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{u}_0 \\ \boldsymbol{\lambda}^{(n-1)} &= \boldsymbol{\lambda}^{(n)} + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \boldsymbol{\kappa}_i^{(n)} \\ \mathbb{M}^T \boldsymbol{\kappa}_i^{(n)} &= b_i \boldsymbol{\lambda}^{(n)} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \boldsymbol{\kappa}_j^{(n)} \end{aligned} \quad (12)$$

for $n = 1, \dots, N_t$ and $i = 1, \dots, s$. The gradient of $j(\mathbf{u}_0)$ is reconstructed from the dual variables as

$$\frac{dF}{d\boldsymbol{\mu}} = \boldsymbol{\lambda}^{(0)T} + \mathbf{u}_0 - \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}). \quad (13)$$

See [54] for the derivation. These methods have been used with considerable success to solve a variety of time-periodic partial differential equations, including the Benjamin-Ono equation [40], a wave-guide array mode-locked laser system [32], and the vortex sheet with surface tension [42]. Unfortunately, the underlying optimization algorithms suffer from relatively slow convergence, requiring many line-searches before becoming superlinear, and never achieve quadratic convergence.

An attractive alternative is to recast the fixed point iteration as a nonlinear system of equations and use either the Newton-Raphson method or the Levenberg-Marquardt method to reap the benefits of quadratic convergence. To this end, define the nonlinear system of equations

$$\mathbf{R}(\mathbf{u}_0) = \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{u}_0 = 0 \quad (14)$$

with Jacobian matrix

$$\mathbf{J}(\mathbf{u}_0) = \frac{\partial \mathbf{R}}{\partial \mathbf{u}_0}(\mathbf{u}_0) = \frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{I} \quad (15)$$

where \mathbf{I} is the $N_u \times N_u$ identity matrix. The crucial component of the Newton-Raphson method is the solution of a linear system of equations with the Jacobian (15), i.e. the solution of $\mathbf{J}(\mathbf{u}_0)\mathbf{x} = \mathbf{b}$, given $\mathbf{u}_0 \in \mathbb{R}^{N_u}$ and $\mathbf{b} \in \mathbb{R}^{N_u}$. A *linear* evolution equation defining $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$, i.e. the sensitivity of the final state

with respect to perturbations in the initial state, is introduced by linearizing the fully discrete evolution equation in (8) about the primal state $\mathbf{u}^{(n)}$, $\mathbf{k}_i^{(n)}$ with respect to the initial state \mathbf{u}_0 . Direct differentiation of (8) with respect to \mathbf{u}_0 leads to the forward sensitivity equations

$$\begin{aligned}\frac{\partial \mathbf{u}^{(0)}}{\partial \mathbf{u}_0} &= \mathbf{I} \\ \frac{\partial \mathbf{u}^{(n)}}{\partial \mathbf{u}_0} &= \frac{\partial \mathbf{u}^{(n-1)}}{\partial \mathbf{u}_0} + \sum_{i=1}^s b_i \frac{\partial \mathbf{k}_i^{(n)}}{\partial \mathbf{u}_0} \\ \mathbb{M} \frac{\partial \mathbf{k}_i^{(n)}}{\partial \mathbf{u}_0} &= \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right) \left[\frac{\partial \mathbf{u}^{(n-1)}}{\partial \mathbf{u}_0} + \sum_{j=1}^i a_{ij} \frac{\partial \mathbf{k}_j^{(n)}}{\partial \mathbf{u}_0} \right].\end{aligned}\tag{16}$$

In general, $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$ is a large ($N_{\mathbf{u}} \times N_{\mathbf{u}}$), dense matrix that requires the solution of $N_{\mathbf{u}}$ linear evolution equations to form. While it is true that the columns of the matrix can be solved in parallel, formation and storage of this matrix may be impractical, particularly for the large-scale computational fluid dynamics problems that motivate this work. For non-dissipative problems such as standing waves in the free-surface Euler equations [34, 35], this is worth the expense since all perturbation directions have to be explored (as opposed to letting the evolution over a cycle damp out high frequency transients). But for viscous problems such as that explored in Section 4 below, solving the Newton-Raphson equations by Krylov subspace methods requires many fewer iterations than there are columns of the Jacobian.

Formation and storage of $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$ can be completely avoided if a matrix-free Krylov method [65] is used to solve the linear systems arising in the Newton-Raphson method, i.e. $\mathbf{J}(\mathbf{u}_0)\mathbf{x} = \mathbf{b}$. In this case, only matrix-vector products of the form

$$\mathbf{J}(\mathbf{u}_0)\mathbf{v} = \frac{\partial \mathbf{R}}{\partial \mathbf{u}_0}(\mathbf{u}_0)\mathbf{v} = \frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0; \boldsymbol{\mu})\mathbf{v} - \mathbf{v}\tag{17}$$

for any $\mathbf{v} \in \mathbb{R}^{N_{\mathbf{u}}}$, are required. For efficiency, these must be computed without explicitly forming the matrix $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$. This is accomplished by considering the forward sensitivity equations in (16) in the direction defined by \mathbf{v} . Multiplying (16) by the vector \mathbf{v} leads to the system of linear evolution equations

$$\begin{aligned}\frac{\partial \mathbf{u}^{(0)}}{\partial \mathbf{u}_0} \mathbf{v} &= \mathbf{v} \\ \frac{\partial \mathbf{u}^{(n)}}{\partial \mathbf{u}_0} \mathbf{v} &= \frac{\partial \mathbf{u}^{(n-1)}}{\partial \mathbf{u}_0} \mathbf{v} + \sum_{i=1}^s b_i \frac{\partial \mathbf{k}_i^{(n)}}{\partial \mathbf{u}_0} \mathbf{v} \\ \mathbb{M} \frac{\partial \mathbf{k}_i^{(n)}}{\partial \mathbf{u}_0} \mathbf{v} &= \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right) \left[\frac{\partial \mathbf{u}^{(n-1)}}{\partial \mathbf{u}_0} \mathbf{v} + \sum_{j=1}^i a_{ij} \frac{\partial \mathbf{k}_j^{(n)}}{\partial \mathbf{u}_0} \mathbf{v} \right].\end{aligned}\tag{18}$$

These can be solved for $\frac{\partial \mathbf{u}^{(n)}}{\partial \mathbf{u}_0} \cdot \mathbf{v}$ and $\frac{\partial \mathbf{k}_i^{(n)}}{\partial \mathbf{u}_0} \cdot \mathbf{v}$ directly, only requiring *one* linear evolution for each \mathbf{v} . Since the equations in (18) are linear, the underlying linear solver must be converged to high accuracy if accurate sensitivities are to be obtained. This mitigates the speedup with respect to the nonlinear, primal solves whose linear systems are usually solved to low precision. For the problems considered in Section 4, the primal equations were, on average, 2 times more expensive than the sensitivity equations, even though 5 nonlinear iterations were required for convergence. This implies the cost of evaluating $\mathbf{R}(\mathbf{u}_0)$ is approximately 2 times as expensive as a Jacobian-vector product $\mathbf{J}(\mathbf{u}_0)\mathbf{v}$. The Newton-Krylov method,

with Jacobian-vector products computed as the solution of (18), is summarized in Algorithm 2. If the linear system of equations arising at each iteration is solved to sufficient accuracy, this algorithm will converge quadratically. The starting guess can be obtained by fixed point iteration (Algorithm 1), or, if solutions come in families parametrized by an amplitude, by numerical continuation [21, 22, 20, 40, 42, 33, 35]. The latter approach is particularly useful if the system is not dissipative and externally driven, as fixed point iteration relies on transient modes being damped by the evolution equations, i.e. on the periodic solution being stable and attracting.

Given this exposition on methods for computing time-periodic solutions of partial differential equations, we turn our attention to determining the *stability* of the corresponding periodic orbit.

Algorithm 2 Newton-Krylov Shooting Method for Time-Periodic Solutions of PDE

Input: Initial guess for periodic initial condition, \mathbf{u}_0 ; parameter configuration, $\boldsymbol{\mu}$

Output: Periodic initial condition, $\mathbf{u}^{(0)}$

- 1: **while** $\|\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{u}_0\|_2 > \epsilon$ **do**
- 2: Solve *unsymmetric* linear system of equations

$$\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0; \boldsymbol{\mu}) \cdot \Delta \mathbf{u} = \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) - \mathbf{u}_0$$

using a *matrix-free* Krylov method with matrix-vector products

$$\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0; \boldsymbol{\mu}) \cdot \mathbf{v}$$

computed as the solution of the directional sensitivity equations (18)

- 3: Update solution

$$\mathbf{u}_0 \leftarrow \mathbf{u}_0 - \Delta \mathbf{u}$$

- 4: **end while**

- 5: Define periodic initial condition

$$\mathbf{u}^{(0)} = \mathbf{u}_0$$

2.3. Stability of Periodic Orbits of Fully Discrete Partial Differential Equations

In this section, the concept of stability of a periodic orbit of fully discrete partial differential equations is introduced and a method for determining the stability of a periodic solution presented. Fixed-point iteration will fail if the underlying solution is not stable and attracting. Moreover, the performance of the Newton-Krylov approach depends critically on whether “most” perturbations are damped out rapidly by the evolution itself. For some problems without dissipation or forcing, e.g. for standing water waves in an ideal fluid [33, 34, 35], the system is neutrally stable, with all Floquet multipliers lying on the unit circle. In that case, it pays to parallelize a full Jacobian calculation in a Newton-Raphson or Levenberg-Marquardt method since all perturbation directions have to be explored before convergence is achieved in practice using an iterative approach. However, for the flapping airfoil studied in Section 4, the Newton-Krylov approach converges in many fewer iterations than there are degrees of freedom. The purpose of this section and the results of Figure 9 below is to explore the stability properties that lead to this behavior for the flapping airfoil.

Recall the interpretation of $\mathbf{u}^{(N_t)}$ as a function that propagates an initial condition \mathbf{u}_0 to its final state $\mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu})$. Let $\mathbf{u}_0^*(\boldsymbol{\mu})$ be the time-periodic solution of the fully discrete partial differential equation in (4) at parameter configuration $\boldsymbol{\mu}$, i.e., $\mathbf{u}_0^*(\boldsymbol{\mu}) = \mathbf{u}^{(N_t)}(\mathbf{u}_0^*; \boldsymbol{\mu})$. A periodic orbit is stable if there is a $\delta > 0$ such that

$$\lim_{n \rightarrow \infty} \|\mathbf{u}^{(n \cdot N_t)}(\mathbf{u}_0^*(\boldsymbol{\mu}) + \Delta \mathbf{u}; \boldsymbol{\mu}) - \mathbf{u}_0^*(\boldsymbol{\mu})\| = 0 \quad (19)$$

if $\|\Delta \mathbf{u}\| < \delta$, where

$$\mathbf{u}^{(n \cdot N_t)}(\mathbf{u}_0; \boldsymbol{\mu}) = \mathbf{u}^{(N_t)}(\cdot; \boldsymbol{\mu}) \circ \dots \circ \mathbf{u}^{(N_t)}(\mathbf{u}_0; \boldsymbol{\mu}). \quad (20)$$

A Taylor expansion of $\mathbf{u}^{(N_t)}$ about the periodic solution leads to

$$\mathbf{u}^{(N_t)}(\mathbf{u}_0^*(\boldsymbol{\mu}) + \Delta \mathbf{u}; \boldsymbol{\mu}) = \mathbf{u}_0^*(\boldsymbol{\mu}) + \frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0^*(\boldsymbol{\mu}); \boldsymbol{\mu}) \cdot \Delta \mathbf{u} + \mathcal{O}(\|\Delta \mathbf{u}\|^2) \quad (21)$$

where time-periodicity of $\mathbf{u}_0^*(\boldsymbol{\mu})$ was used. Repeated application of (21) leads to

$$\mathbf{u}^{(n \cdot N_t)}(\mathbf{u}_0^*(\boldsymbol{\mu}) + \Delta \mathbf{u}; \boldsymbol{\mu}) = \mathbf{u}_0^*(\boldsymbol{\mu}) + \left[\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0^*(\boldsymbol{\mu}); \boldsymbol{\mu}) \right]^n \Delta \mathbf{u} + \mathcal{O}(\|\Delta \mathbf{u}\|^{n+1}). \quad (22)$$

Taking $\delta < 1$, the stability criteria in (19) is satisfied if all eigenvalues of $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}(\mathbf{u}_0^*(\boldsymbol{\mu}); \boldsymbol{\mu})$ have modulus strictly less than 1. In Section 4, the stability of the periodic flow around a flapping airfoil is verified using this method.

Given this exposition on solvers for time-periodically constrained partial differential equations, we turn our attention to deriving the corresponding fully discrete adjoint equations.

3. Fully Discrete Time-Periodic Adjoint Method

In this section, the adjoint equations corresponding to the fully discrete time-periodically constrained partial differential equations (4) and quantity of interest $F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu})$, will be derived. For the remainder of this section, $\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}$ will be taken as the time-periodic solution of the fully discrete partial differential equations (4) at parameter $\boldsymbol{\mu}$. The adjoint equations will be derived by linearizing the fully discrete equations about this periodic solution. This highlights the importance of an efficient periodic solver – the subject of Section 2.2 – as it is a prerequisite for the adjoint method.

Before proceeding to the derivation of the adjoint equations, the following definitions are introduced for the fully discrete time-periodic constraint and Runge-Kutta stage equations and state updates

$$\begin{aligned} \tilde{\mathbf{r}}^{(0)}(\mathbf{u}^{(0)}, \mathbf{u}^{(N_t)}) &= \mathbf{u}^{(0)} - \mathbf{u}^{(N_t)} = 0 \\ \tilde{\mathbf{r}}^{(n)}(\mathbf{u}^{(n-1)}, \mathbf{u}^{(n)}, \mathbf{k}_1^{(n)}, \dots, \mathbf{k}_s^{(n)}, \boldsymbol{\mu}) &= \mathbf{u}^{(n)} - \mathbf{u}^{(n-1)} - \sum_{i=1}^s b_i \mathbf{k}_i^{(i)} = 0 \\ \mathbf{R}_i^{(n)}(\mathbf{u}^{(n-1)}, \mathbf{k}_1^{(n)}, \dots, \mathbf{k}_i^{(n)}, \boldsymbol{\mu}) &= \mathbb{M} \mathbf{k}_i^{(n)} - \Delta t_n \mathbf{r}(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n) = 0 \end{aligned} \quad (23)$$

for $n = 1, \dots, n$ and $i = 1, \dots, s$.

3.1. Derivation

The derivation of the fully discrete adjoint equations corresponding to the output functional, F , begins with the introduction of test variables

$$\boldsymbol{\lambda}^{(0)}, \boldsymbol{\lambda}^{(n)}, \boldsymbol{\kappa}_i^{(n)} \in \mathbb{R}^{N_u} \quad (24)$$

for $n = 1, \dots, N_t$ and $i = 1, \dots, s$. Since $\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}$ are taken as the solution of the fully discrete time-periodic problem in (23), the following identity holds, for any $\boldsymbol{\mu} \in \mathbb{R}^{N_\mu}$,

$$F = F + 0 = F - \boldsymbol{\lambda}^{(0)T} \tilde{\mathbf{r}}^{(0)} - \sum_{n=1}^{N_t} \boldsymbol{\lambda}^{(n)T} \tilde{\mathbf{r}}^{(n)} - \sum_{n=1}^{N_t} \sum_{i=1}^s \boldsymbol{\kappa}_i^{(n)T} \mathbf{R}_i^{(n)} \quad (25)$$

for any value of the test functions $\boldsymbol{\lambda}^{(n)}$ and $\boldsymbol{\kappa}_i^{(n)}$. In (25), arguments have been dropped for brevity; it is understood that all terms are evaluated at the periodic solution of (4) at parameter $\boldsymbol{\mu}$. Since (23) holds for

any $\boldsymbol{\mu} \in \mathbb{R}^{N_\mu}$, provided $\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}$ is the corresponding periodic solution, differentiation with respect to $\boldsymbol{\mu}$ leads to

$$\begin{aligned} \frac{dF}{d\boldsymbol{\mu}} &= \frac{\partial F}{\partial \boldsymbol{\mu}} + \sum_{n=0}^{N_t} \frac{\partial F}{\partial \mathbf{u}^{(n)}} \frac{\partial \mathbf{u}^{(n)}}{\partial \boldsymbol{\mu}} + \sum_{n=1}^{N_t} \sum_{i=1}^s \frac{\partial F}{\partial \mathbf{k}_i^{(n)}} \frac{\partial \mathbf{k}_i^{(n)}}{\partial \boldsymbol{\mu}} - \boldsymbol{\lambda}^{(0)T} \left[\frac{\partial \tilde{\mathbf{r}}^{(0)}}{\partial \boldsymbol{\mu}} + \frac{\partial \tilde{\mathbf{r}}^{(0)}}{\partial \mathbf{u}^{(0)}} \frac{\partial \mathbf{u}^{(0)}}{\partial \boldsymbol{\mu}} + \frac{\partial \tilde{\mathbf{r}}^{(0)}}{\partial \mathbf{u}^{(N_t)}} \frac{\partial \mathbf{u}^{(N_t)}}{\partial \boldsymbol{\mu}} \right] \\ &\quad - \sum_{n=1}^{N_t} \boldsymbol{\lambda}^{(n)T} \left[\frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \boldsymbol{\mu}} + \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \mathbf{u}^{(n)}} \frac{\partial \mathbf{u}^{(n)}}{\partial \boldsymbol{\mu}} + \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \mathbf{u}^{(n-1)}} \frac{\partial \mathbf{u}^{(n-1)}}{\partial \boldsymbol{\mu}} + \sum_{p=1}^s \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \mathbf{k}_p^{(n)}} \frac{\partial \mathbf{k}_p^{(n)}}{\partial \boldsymbol{\mu}} \right] \\ &\quad - \sum_{n=1}^{N_t} \sum_{i=1}^s \boldsymbol{\kappa}_i^{(n)T} \left[\frac{\partial \mathbf{R}_i^{(n)}}{\partial \boldsymbol{\mu}} + \frac{\partial \mathbf{R}_i^{(n)}}{\partial \mathbf{u}^{(n-1)}} \frac{\partial \mathbf{u}^{(n-1)}}{\partial \boldsymbol{\mu}} + \sum_{j=1}^i \frac{\partial \mathbf{R}_i^{(n)}}{\partial \mathbf{k}_j^{(n)}} \frac{\partial \mathbf{k}_j^{(n)}}{\partial \boldsymbol{\mu}} \right]. \end{aligned} \quad (26)$$

Re-arrangement of terms in (26) such that the state variable sensitivities are isolated leads to the following expression for $\frac{dF}{d\boldsymbol{\mu}}$

$$\begin{aligned} \frac{dF}{d\boldsymbol{\mu}} &= \frac{\partial F}{\partial \boldsymbol{\mu}} + \left[\frac{\partial F}{\partial \mathbf{u}^{(N_t)}} - \boldsymbol{\lambda}^{(N_t)T} \frac{\partial \tilde{\mathbf{r}}^{(N_t)}}{\partial \mathbf{u}^{(N_t)}} - \boldsymbol{\lambda}^{(0)T} \frac{\partial \tilde{\mathbf{r}}^{(0)}}{\partial \mathbf{u}^{(N_t)}} \right] \frac{\partial \mathbf{u}^{(N_t)}}{\partial \boldsymbol{\mu}} - \sum_{n=0}^{N_t} \boldsymbol{\lambda}^{(n)T} \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \boldsymbol{\mu}} - \sum_{n=1}^{N_t} \sum_{p=1}^s \boldsymbol{\kappa}_p^{(n)T} \frac{\partial \mathbf{R}_p^{(n)}}{\partial \boldsymbol{\mu}} \\ &\quad + \sum_{n=1}^{N_t} \left[\frac{\partial F}{\partial \mathbf{u}^{(n-1)}} - \boldsymbol{\lambda}^{(n-1)T} \frac{\partial \tilde{\mathbf{r}}^{(n-1)}}{\partial \mathbf{u}^{(n-1)}} - \boldsymbol{\lambda}^{(n)T} \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \mathbf{u}^{(n-1)}} - \sum_{i=1}^s \boldsymbol{\kappa}_i^{(n)T} \frac{\partial \mathbf{R}_i^{(n)}}{\partial \mathbf{u}^{(n-1)}} \right] \frac{\partial \mathbf{u}^{(n-1)}}{\partial \boldsymbol{\mu}} \\ &\quad + \sum_{n=1}^{N_t} \sum_{p=1}^s \left[\frac{\partial F}{\partial \mathbf{k}_p^{(n)}} - \boldsymbol{\lambda}^{(n)T} \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \mathbf{k}_p^{(n)}} - \sum_{i=p}^s \boldsymbol{\kappa}_i^{(n)T} \frac{\partial \mathbf{R}_i^{(n)}}{\partial \mathbf{k}_p^{(n)}} \right] \frac{\partial \mathbf{k}_p^{(n)}}{\partial \boldsymbol{\mu}}. \end{aligned} \quad (27)$$

The dual variables, $\boldsymbol{\lambda}^{(n)}$ and $\boldsymbol{\kappa}_i^{(n)}$, which have remained arbitrary to this point, are chosen such that the bracketed terms in (27) vanish

$$\begin{aligned} \frac{\partial \tilde{\mathbf{r}}^{(0)T}}{\partial \mathbf{u}^{(N_t)}} \boldsymbol{\lambda}^{(0)} + \frac{\partial \tilde{\mathbf{r}}^{(N_t)T}}{\partial \mathbf{u}^{(N_t)}} \boldsymbol{\lambda}^{(N_t)} &= \frac{\partial F}{\partial \mathbf{u}^{(N_t)}} \\ \frac{\partial \tilde{\mathbf{r}}^{(n)T}}{\partial \mathbf{u}^{(n-1)}} \boldsymbol{\lambda}^{(n)} + \frac{\partial \tilde{\mathbf{r}}^{(n-1)T}}{\partial \mathbf{u}^{(n-1)}} \boldsymbol{\lambda}^{(n-1)} &= \frac{\partial F}{\partial \mathbf{u}^{(n-1)}} - \sum_{i=1}^s \frac{\partial \mathbf{R}_i^{(n)T}}{\partial \mathbf{u}^{(n-1)}} \boldsymbol{\kappa}_i^{(n)} \\ \sum_{j=i}^s \frac{\partial \mathbf{R}_j^{(n)T}}{\partial \mathbf{k}_i^{(n)}} \boldsymbol{\kappa}_j^{(n)} &= \frac{\partial F}{\partial \mathbf{k}_i^{(n)}} - \frac{\partial \tilde{\mathbf{r}}^{(n)T}}{\partial \mathbf{k}_i^{(n)}} \boldsymbol{\lambda}^{(n)} \end{aligned} \quad (28)$$

for $n = 1, \dots, N_t$ and $i = 1, \dots, s$. These are the *fully discrete adjoint equations* corresponding to the time-periodic primal evolution equations in (23), discrete quantity of interest F , and parameter $\boldsymbol{\mu}$. Defining the dual variables as the solution of the adjoint equations in (28), the expression for $\frac{dF}{d\boldsymbol{\mu}}$ in (27) reduces to

$$\frac{dF}{d\boldsymbol{\mu}} = \frac{\partial F}{\partial \boldsymbol{\mu}} - \sum_{n=0}^{N_t} \boldsymbol{\lambda}^{(n)T} \frac{\partial \tilde{\mathbf{r}}^{(n)}}{\partial \boldsymbol{\mu}} - \sum_{n=1}^{N_t} \sum_{p=1}^s \boldsymbol{\kappa}_p^{(n)T} \frac{\partial \mathbf{R}_p^{(n)}}{\partial \boldsymbol{\mu}}. \quad (29)$$

This provides a means of computing the total derivative $\frac{dF}{d\boldsymbol{\mu}}$ without explicitly computing the large, dense state sensitivities since the expression in (29) is independent of them. Direct differentiation of $\tilde{\mathbf{r}}^{(n)}$ and $\mathbf{R}_i^{(n)}$ from their definitions in (23) leads to the final form of the adjoint equations of the fully discrete,

time-periodically constrained partial differential equations in (4)

$$\begin{aligned}
\boldsymbol{\lambda}^{(N_t)} &= \boldsymbol{\lambda}^{(0)} + \frac{\partial F}{\partial \mathbf{u}^{(N_t)}}{}^T \\
\boldsymbol{\lambda}^{(n-1)} &= \boldsymbol{\lambda}^{(n)} + \frac{\partial F}{\partial \mathbf{u}^{(n-1)}}{}^T + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \boldsymbol{\kappa}_i^{(n)} \\
\mathbb{M}^T \boldsymbol{\kappa}_i^{(n)} &= \frac{\partial F}{\partial \mathbf{k}_i^{(n)}}{}^T + b_i \boldsymbol{\lambda}^{(n)} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \boldsymbol{\kappa}_j^{(n)}
\end{aligned} \tag{30}$$

for $n = 1, \dots, N_t$ and $i = 1, \dots, s$. Similarly, the total derivative of F , independent of state sensitivities, takes the form

$$\frac{dF}{d\boldsymbol{\mu}} = \frac{\partial F}{\partial \boldsymbol{\mu}} + \sum_{n=1}^{N_t} \Delta t_n \sum_{i=1}^s \boldsymbol{\kappa}_i^{(n)T} \frac{\partial \mathbf{r}}{\partial \boldsymbol{\mu}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right). \tag{31}$$

From (30), it can be seen that the fully discrete adjoint equations take the form of a *linear, two-point boundary-value problem* and cannot be solved directly as an evolution equation. Appendix A proves existence and uniqueness of solutions to (30). The next section will discuss solvers for the discrete time-periodic adjoint equations in (30).

3.2. Numerical Solver: Matrix-Free Krylov Method

As the adjoint equations corresponding to the fully discrete time-periodic partial differential equation are linear, this section will consider matrix-free Krylov methods to solve them. Alternatively, any of the methods discussed in Section 2.2 could be used.

Define $\boldsymbol{\lambda}^{(0)}(\boldsymbol{\lambda}_{N_t}; \boldsymbol{\mu})$ as the solution of the linear, backward evolution equations

$$\begin{aligned}
\boldsymbol{\lambda}^{(N_t)} &= \boldsymbol{\lambda}_{N_t} \\
\boldsymbol{\lambda}^{(n-1)} &= \boldsymbol{\lambda}^{(n)} + \frac{\partial F}{\partial \mathbf{u}^{(n-1)}}{}^T + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \boldsymbol{\kappa}_i^{(n)} \\
\mathbb{M}^T \boldsymbol{\kappa}_i^{(n)} &= \frac{\partial F}{\partial \mathbf{k}_i^{(n)}}{}^T + b_i \boldsymbol{\lambda}^{(n)} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \boldsymbol{\kappa}_j^{(n)},
\end{aligned} \tag{32}$$

which can be directly evolved, backward-in-time. Similar to Section 2.2 this constitutes a notation overload since $\boldsymbol{\lambda}^{(0)} \in \mathbb{R}^{N_u}$ is the initial solution of the adjoint equations corresponding to the fully discrete periodic partial differential equations, as well as the linear function that takes a state $\boldsymbol{\lambda}_{N_t}$ to $\boldsymbol{\lambda}^{(0)}(\boldsymbol{\lambda}_{N_t}; \boldsymbol{\mu})$. Then, $\boldsymbol{\lambda}^{(0)}(\boldsymbol{\lambda}_{N_t}; \boldsymbol{\mu})$ is the initial solution of (30) if the following *linear* equation is satisfied

$$\boldsymbol{\lambda}^{(0)}(\boldsymbol{\lambda}_{N_t}; \boldsymbol{\mu}, t) = \boldsymbol{\lambda}_{N_t} - \frac{\partial F}{\partial \mathbf{u}^{(N_t)}}{}^T. \tag{33}$$

This is a linear system of equations of the form, $\mathbf{A}\mathbf{x} = \mathbf{b}$ where

$$\mathbf{A} = \frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} - \mathbf{I}. \tag{34}$$

The columns of the linear operator \mathbf{A} can be formed by considering perturbations of (32) with respect to

the final state $\boldsymbol{\lambda}_{N_t}$. Differentiation of (32) with respect to $\boldsymbol{\lambda}_{N_t}$ leads to the adjoint sensitivity equations

$$\begin{aligned}\frac{\partial \boldsymbol{\lambda}^{(N_t)}}{\partial \boldsymbol{\lambda}_{N_t}} &= \mathbf{I} \\ \frac{\partial \boldsymbol{\lambda}^{(n-1)}}{\partial \boldsymbol{\lambda}_{N_t}} &= \frac{\partial \boldsymbol{\lambda}^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \frac{\partial \boldsymbol{\kappa}_i^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} \\ \mathbb{M}^T \frac{\partial \boldsymbol{\kappa}_i^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} &= b_i \frac{\partial \boldsymbol{\lambda}^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \frac{\partial \boldsymbol{\kappa}_j^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}}.\end{aligned}\tag{35}$$

Similar to the situation for the primal problem, the matrix $\frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}}$ is an $N_{\mathbf{u}} \times N_{\mathbf{u}}$ dense matrix that requires $N_{\mathbf{u}}$ linear evolution equations to form. As this is impractical for large problems, a matrix-free Krylov method is used to solve (33), which only requires matrix-vector products of the form

$$\mathbf{A} \mathbf{v} = \frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} - \mathbf{v}.\tag{36}$$

The first term in this matrix-vector product can be computed directly by considering the adjoint sensitivity equations in a given direction \mathbf{v}

$$\begin{aligned}\frac{\partial \boldsymbol{\lambda}^{(N_t)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} &= \mathbf{v} \\ \frac{\partial \boldsymbol{\lambda}^{(n-1)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} &= \frac{\partial \boldsymbol{\lambda}^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \frac{\partial \boldsymbol{\kappa}_i^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} \\ \mathbb{M}^T \frac{\partial \boldsymbol{\kappa}_i^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} &= b_i \frac{\partial \boldsymbol{\lambda}^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \frac{\partial \boldsymbol{\kappa}_j^{(n)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v}.\end{aligned}\tag{37}$$

The equations in (37) can be solved for $\frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} \cdot \mathbf{v}$ at the cost of one linear evolution solution for each \mathbf{v} .

The adjoint sensitivity equations in (37) are *independent* of the quantity of interest, F . If there are multiple quantities of interest, fast multiple right-hand side solvers [66, 67, 68] could be used to solve $\mathbf{A} \mathbf{x} = \mathbf{b}$ as the matrix \mathbf{A} will be fixed and only the right-hand side varied. Furthermore, the adjoint sensitivity equations in (37) and the adjoint equations in (32) are identical, with the exception of the terms $\frac{\partial F}{\partial \mathbf{u}^{(n-1)}}$ and $\frac{\partial F}{\partial \boldsymbol{\kappa}_i^{(n)}}$.

Therefore, the adjoint sensitivities are less expensive to compute than the adjoint states and the savings become substantial if $\frac{\partial F}{\partial \mathbf{u}^{(n-1)}}$ and $\frac{\partial F}{\partial \boldsymbol{\kappa}_i^{(n)}}$ are expensive to compute. Algorithm 3 below details the use of a matrix-free GMRES method to solve (33) with matrix-vector products defined by (37).

With the solution of the fully discrete primal and dual time-periodic problems fully specified, from numerical discretization to solution algorithms, we close this section with an algorithm that uses the fully discrete adjoint method to compute the gradient of the quantity of interest *on the manifold of periodic solutions*. First, the fully discrete time-periodic solution (4) must be computed, i.e. using a matrix-free Newton-Krylov method, to yield $\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}$. Next, the corresponding fully discrete adjoint equations are defined about this periodic solution and solved, i.e. using a matrix-free Krylov method, for $\boldsymbol{\lambda}^{(0)}, \dots, \boldsymbol{\lambda}^{(N_t)}, \boldsymbol{\kappa}_1^{(1)}, \dots, \boldsymbol{\kappa}_s^{(N_t)}$. Finally, (31) is used to reconstruct the desired gradient $\frac{dF}{d\boldsymbol{\mu}}$. This procedure is summarized in Algorithm 4.

Algorithm 3 GMRES for Solution of Fully Discrete, Time-Periodic Adjoint PDE

Input: Initial guess for periodic adjoint final condition, $\lambda_{N_t,0}$; parameter configuration, μ ; periodic primal solution, $\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}$

Output: Periodic adjoint final condition, $\lambda^{(N_t)}$

1: Compute

$$\mathbf{r}_0 = \lambda^{(0)}(\lambda_{N_t,0}, \mu) + \frac{\partial F}{\partial \mathbf{u}^{(N_t)}}{}^T - \lambda_{N_t,0}$$

2: Set $\beta = \|\mathbf{r}_0\|_2$, $\mathbf{v}_1 = \mathbf{r}_0/\beta$, and $\lambda_{N_t} = \lambda_{N_t,0}$

3: **while** $\|\lambda^{(0)}(\lambda_{N_t}, \mu) + \frac{\partial F}{\partial \mathbf{u}^{(N_t)}}{}^T - \lambda_{N_t}\|_2 > \epsilon$ **do**

4: **for** $j = 1, 2, \dots, m$ **do**

5: Compute

$$\mathbf{w}_j = \frac{\partial \lambda^{(0)}}{\partial \lambda_{N_t}} \mathbf{v}_j - \mathbf{v}_j$$

 as the solution of the adjoint sensitivity equations (37)

6: **for** $i = 1, \dots, j$ **do**

7: $h_{ij} = (\mathbf{w}_j, \mathbf{v}_i)$

8: $\mathbf{w}_j = \mathbf{w}_j - h_{ij}\mathbf{v}_i$

9: **end for**

10: $h_{j+1,j} = \|\mathbf{w}_j\|_2$

11: $\mathbf{v}_{j+1} = \mathbf{w}_j/h_{j+1,j}$

12: **end for**

13: Compute

$$\mathbf{y}_m = \arg \min \|\beta \mathbf{e}_1 - \mathbf{H}_m \mathbf{y}\|_2,$$

 where \mathbf{e}_1 is the first canonical unit vector in \mathbb{R}^{N_u} and $\mathbf{H} = \{h_{ij}\}_{1 \leq i \leq m+1, 1 \leq j \leq m}$

14: Update solution

$$\lambda_{N_t} = \lambda_{N_t,0} + \mathbf{V}_m \mathbf{y}_m$$

 where

$$\mathbf{V}_m = [\mathbf{v}_1 \quad \dots \quad \mathbf{v}_m]$$

15: **end while**

16: Define adjoint periodic final condition

$$\lambda^{(N_t)} = \lambda_{N_t}$$

Algorithm 4 Gradients on Manifold of Time-Periodic Solutions of PDEs

Input: Parameter configuration, $\boldsymbol{\mu}$, and fully discrete quantity of interest, $F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)})$

Output: Gradient, $\frac{dF}{d\boldsymbol{\mu}}$, on manifold of time-periodic solutions

- 1: For parameter $\boldsymbol{\mu}$, compute time-periodic solution of fully discrete PDE in (4), i.e. using the Newton-Krylov shooting method in Algorithm 2

$$\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}$$

- 2: For fully discrete functional $F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)})$, compute adjoint solution of fully discrete time-periodic PDE in (30), i.e. using GMRES shooting method in Algorithm 3 with matrix-vector products computed from the backward evolution of the adjoint sensitivity equations in (37)

$$\boldsymbol{\lambda}^{(0)}, \dots, \boldsymbol{\lambda}^{(N_t)}, \boldsymbol{\kappa}_1^{(1)}, \dots, \boldsymbol{\kappa}_s^{(N_t)}$$

- 3: Reconstruct $\frac{dF}{d\boldsymbol{\mu}}$ using dual variables according to (31)
-

3.3. Generalized Reduced-Gradient Method for PDE Optimization with Time-Periodicity Constraints

Consider the fully discrete time-dependent PDE-constrained optimization problem

$$\begin{aligned}
 & \underset{\substack{\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)} \in \mathbb{R}^{N\mathbf{u}}, \\ \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)} \in \mathbb{R}^{N\mathbf{u}}, \\ \boldsymbol{\mu} \in \mathbb{R}^{N\boldsymbol{\mu}}}}{\text{minimize}} & F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu}) \\
 & \text{subject to} & \mathbf{c}(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu}) \geq 0 \\
 & & \mathbf{u}^{(0)} = \mathbf{u}_0 \\
 & & \mathbf{u}^{(n)} = \mathbf{u}^{(n-1)} + \sum_{i=1}^s b_i \mathbf{k}_i^{(n)} \\
 & & \mathbb{M} \mathbf{k}_i^{(n)} = \Delta t_n \mathbf{r}(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n)
 \end{aligned} \tag{38}$$

where F is a fully discrete output functional of the partial differential equation and \mathbf{c} is a vector of such output functionals. The nested or Generalized Reduced-Gradient (GRG) approach to solve (38) explicitly enforces the PDE constraint at each optimization iteration. The implicit function theorem states that the solution of the discretized PDE, can be considered an implicit function of the parameter $\boldsymbol{\mu}$, i.e., $\mathbf{u}^{(n)} = \mathbf{u}^{(n)}(\boldsymbol{\mu})$ and $\mathbf{k}_i^{(n)} = \mathbf{k}_i^{(n)}(\boldsymbol{\mu})$. Strict enforcement of the discretized partial differential equation allows the PDE variables and equations to be removed from the optimization problem

$$\begin{aligned}
 & \underset{\boldsymbol{\mu} \in \mathbb{R}^{N\boldsymbol{\mu}}}{\text{minimize}} & F(\mathbf{u}^{(0)}(\boldsymbol{\mu}), \dots, \mathbf{u}^{(N_t)}(\boldsymbol{\mu}), \mathbf{k}_1^{(1)}(\boldsymbol{\mu}), \dots, \mathbf{k}_s^{(N_t)}(\boldsymbol{\mu}), \boldsymbol{\mu}) \\
 & \text{subject to} & \mathbf{c}(\mathbf{u}^{(0)}(\boldsymbol{\mu}), \dots, \mathbf{u}^{(N_t)}(\boldsymbol{\mu}), \mathbf{k}_1^{(1)}(\boldsymbol{\mu}), \dots, \mathbf{k}_s^{(N_t)}(\boldsymbol{\mu}), \boldsymbol{\mu}) \geq 0.
 \end{aligned} \tag{39}$$

To solve this optimization problem using gradient-based techniques, the terms $\frac{dF}{d\boldsymbol{\mu}}$ and $\frac{d\mathbf{c}}{d\boldsymbol{\mu}}$ – gradients of quantities of interest along the manifold of solutions of the PDE – are required. Depending on the relative number of variables in $\boldsymbol{\mu}$ to the number of constraints in \mathbf{c} , the direct or adjoint method can be efficiently used to compute these gradients *without relying on finite differences*.

Now consider the optimization problem in (38) with the time-periodicity constraint added

$$\begin{aligned}
& \underset{\substack{\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)} \in \mathbb{R}^{N\mathbf{u}}, \\ \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)} \in \mathbb{R}^{N\mathbf{u}}, \\ \boldsymbol{\mu} \in \mathbb{R}^{N\boldsymbol{\mu}}}}{\text{minimize}} & F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu}) \\
& \text{subject to} & \mathbf{c}(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu}) \geq 0 \\
& & \mathbf{u}^{(0)} = \mathbf{u}^{(N_t)} \\
& & \mathbf{u}^{(n)} = \mathbf{u}^{(n-1)} + \sum_{i=1}^s b_i \mathbf{k}_i^{(n)} \\
& & \mathbb{M} \mathbf{k}_i^{(n)} = \Delta t_n \mathbf{r} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right).
\end{aligned} \tag{40}$$

Strict enforcement of the time-periodic partial differential equations leads to an application of the implicit function theorem, similar to that above, i.e., $\mathbf{u}^{(n)} = \mathbf{u}^{(n)}(\boldsymbol{\mu})$ and $\mathbf{k}_i^{(n)} = \mathbf{k}_i^{(n)}(\boldsymbol{\mu})$, where $\mathbf{u}^{(n)}$ and $\mathbf{k}_i^{(n)}$ are the time-periodic solution of the discrete partial differential equations. This results in an optimization problem identical to that in (39) with this new definition of $\mathbf{u}^{(n)}(\boldsymbol{\mu})$ and $\mathbf{k}_i^{(n)}(\boldsymbol{\mu})$. The novel periodic adjoint method, derived in Section 3.1, can be used to compute gradients along the manifold of time-periodic solutions of the fully discrete PDE, i.e. $\frac{dF}{d\boldsymbol{\mu}}$ and $\frac{d\mathbf{c}}{d\boldsymbol{\mu}}$, for the use in gradient-based optimization.

4. Applications

This section will present two numerical experiments that study time-periodic solutions of the compressible Navier-Stokes equations, and the behavior of the methods discussion. The first experiment will solely consider solutions of the fully discrete primal and dual periodic two-point boundary-value problems, and study convergence of the methods introduced. The second will apply the novel, fully discrete, periodic adjoint method to solve an optimal control problem governed by a time-periodically constrained partial differential equation.

The compressible Navier-Stokes equations take the form

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i} (\rho u_i) = 0, \tag{41}$$

$$\frac{\partial}{\partial t} (\rho u_i) + \frac{\partial}{\partial x_i} (\rho u_i u_j + p) = \frac{\partial \tau_{ij}}{\partial x_j} \quad \text{for } i = 1, 2, 3, \tag{42}$$

$$\frac{\partial}{\partial t} (\rho E) + \frac{\partial}{\partial x_i} (u_j (\rho E + p)) = -\frac{\partial q_j}{\partial x_j} + \frac{\partial}{\partial x_j} (u_j \tau_{ij}), \tag{43}$$

in $\Omega(\boldsymbol{\mu}, t)$ where ρ is the fluid density, u_1, u_2, u_3 are the velocity components, and E is the total energy. The viscous stress tensor and heat flux are given by

$$\tau_{ij} = \mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right) \quad \text{and} \quad q_j = -\frac{\mu}{\text{Pr}} \frac{\partial}{\partial x_j} \left(E + \frac{p}{\rho} - \frac{1}{2} u_k u_k \right). \tag{44}$$

Here, μ is the viscosity coefficient and $\text{Pr} = 0.72$ is the Prandtl number which we assume to be constant. For an ideal gas, the pressure p has the form

$$p = (\gamma - 1) \rho \left(E - \frac{1}{2} u_k u_k \right), \tag{45}$$

where $\gamma = 1.4$ is the adiabatic gas constant. Furthermore, entropy of the system is assumed constant, i.e. isentropic, which is equivalent to the flow being adiabatic and reversible. For a perfect gas, the entropy is defined as

$$s = p / \rho^\gamma. \tag{46}$$

From the isentropic assumption, (46) explicitly relates the pressure and density of the flow, rendering the energy equation redundant. This effectively reduces the square system of PDEs of size $n_{sd} + 2$ to one of size $n_{sd} + 1$, where n_{sd} is the number of spatial dimensions. It can be shown, under suitable assumptions, that the solution of the isentropic approximation of the Navier-Stokes equations converges to the solution of the incompressible Navier-Stokes equations as the Mach number approaches zero [69, 70, 71].

The compressible Navier-Stokes equations can be written in conservation form as

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{U}, \nabla \mathbf{U}) = 0 \quad \text{in } \Omega(\boldsymbol{\mu}, t), \quad (47)$$

where $\Omega(\boldsymbol{\mu}, t)$ is, in general, a time-dependent, parametrized domain. The Arbitrary-Lagrangian-Eulerian (ALE) form of the conservation law are obtained – following the work in [72] – by introducing a mapping \mathcal{G} from the deforming domain $\Omega(\boldsymbol{\mu}, t)$ to a fixed, reference domain V (Figure 1). This mapping is used to transform the conservation law on $\Omega(\boldsymbol{\mu}, t)$ to one on V . The transformed system of conservation laws, defined on the reference domain V takes the form

$$\frac{\partial \mathbf{U}_{\mathbf{X}}}{\partial t} \Big|_{\mathbf{X}} + \nabla_{\mathbf{X}} \cdot \mathbf{F}_{\mathbf{X}}(\mathbf{U}_{\mathbf{X}}, \nabla_{\mathbf{X}} \mathbf{U}_{\mathbf{X}}) = 0 \quad (48)$$

where $\nabla_{\mathbf{X}}$ denotes spatial derivatives with respect to the reference variables, \mathbf{X} . The transformed state vector, $\mathbf{U}_{\mathbf{X}}$, and its corresponding spatial gradient with respect to the reference configuration take the form

$$\mathbf{U}_{\mathbf{X}} = g\mathbf{U}, \quad \nabla_{\mathbf{X}} \mathbf{U}_{\mathbf{X}} = g^{-1} \mathbf{U}_{\mathbf{X}} \frac{\partial g}{\partial \mathbf{X}} + g \nabla \mathbf{U} \cdot \mathbf{G}, \quad (49)$$

where $\mathbf{G} = \nabla_{\mathbf{X}} \mathcal{G}$, $g = \det(\mathbf{G})$, $\mathbf{v}_{\mathbf{G}} = \frac{\partial \mathbf{x}}{\partial t} = \frac{\partial \mathcal{G}}{\partial t}$, and the arguments have been dropped, for brevity. The transformed fluxes are

$$\begin{aligned} \mathbf{F}_{\mathbf{X}}(\mathbf{U}_{\mathbf{X}}, \nabla_{\mathbf{X}} \mathbf{U}_{\mathbf{X}}) &= \mathbf{F}_{\mathbf{X}}^{inv}(\mathbf{U}_{\mathbf{X}}) + \mathbf{F}_{\mathbf{X}}^{vis}(\mathbf{U}_{\mathbf{X}}, \nabla_{\mathbf{X}} \mathbf{U}_{\mathbf{X}}), \\ \mathbf{F}_{\mathbf{X}}^{inv}(\mathbf{U}_{\mathbf{X}}) &= g \mathbf{F}^{inv}(g^{-1} \mathbf{U}_{\mathbf{X}}) \mathbf{G}^{-T} - \mathbf{U}_{\mathbf{X}} \otimes \mathbf{G}^{-1} \mathbf{v}_{\mathbf{G}}, \\ \mathbf{F}_{\mathbf{X}}^{vis}(\mathbf{U}_{\mathbf{X}}, \nabla_{\mathbf{X}} \mathbf{U}_{\mathbf{X}}) &= g \mathbf{F}^{vis} \left(g^{-1} \mathbf{U}_{\mathbf{X}}, g^{-1} \left[\nabla_{\mathbf{X}} \mathbf{U}_{\mathbf{X}} - g^{-1} \mathbf{U}_{\mathbf{X}} \frac{\partial g}{\partial \mathbf{X}} \right] \mathbf{G}^{-1} \right) \mathbf{G}^{-T}. \end{aligned} \quad (50)$$

For details regarding the derivation of the transformed equations and modifications that ensure the Geometric Conservation Law [73] is satisfied, the reader is referred to [72, 54]. The system of equations (48) is discretized using a nodal discontinuous Galerkin (DG) method on unstructured meshes of triangles, with polynomial degrees 3 within each element. The viscous fluxes are chosen according to the compact DG method [74] method, and the our implementation is fully implicit with exact Jacobian matrices and a range of parallel iterative solvers [75]. The resulting semi-discrete system has the form of our general system of ODEs (3). All partial derivatives of the semi-discrete governing equations and corresponding quantities of interest, namely $\frac{\partial \mathbf{r}}{\partial \mathbf{u}}$, $\frac{\partial \mathbf{r}}{\partial \boldsymbol{\mu}}$, $\frac{\partial f_h}{\partial \mathbf{u}}$, $\frac{\partial f_h}{\partial \boldsymbol{\mu}}$ are computed via automatic symbolic differentiation at the element-level with the MAPLE software [76] and subsequent assembly. The semi-discrete quantity of interest f_h is defined as the approximation of $\int_{\Gamma(\boldsymbol{\mu}, t)} f(\mathbf{U}, \boldsymbol{\mu}, t) dS$ in (2) using the DG shape functions and required, along with the temporal discretization scheme, to compute the discrete output functional F in (7). Additional details regarding computation of the partial derivatives with respect to $\boldsymbol{\mu}$ in the case of a parametrized, deforming domain are provided in [54].

The remainder of this document will consider the time-periodic solution and optimization of a flapping NACA0012 airfoil, shown in Figure 1. Two quantities of interest that will be considered are the total work exerted by the fluid on the airfoil, \mathcal{W} , and the impulse in the x -direction imparted on the airfoil by the fluid, \mathcal{J}_x , which take the form

$$\mathcal{W}(\mathbf{U}, \boldsymbol{\mu}) = \int_0^T \int_{\Gamma} \mathbf{f}(\mathbf{U}, \boldsymbol{\mu}, t) \cdot \dot{\mathbf{x}} dS dt \quad \text{and} \quad \mathcal{J}_x(\mathbf{U}, \boldsymbol{\mu}) = \int_0^T \int_{\Gamma} \mathbf{f}(\mathbf{U}, \boldsymbol{\mu}, t) \cdot \mathbf{e}_1 dS dt \quad (51)$$

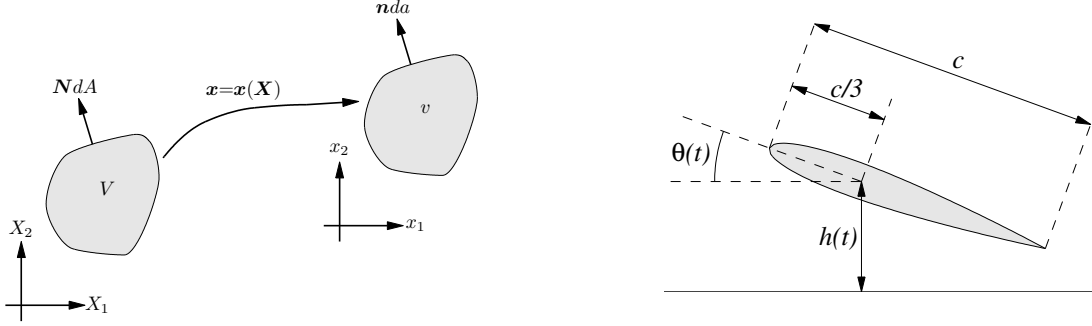


Figure 1: Time-dependent mapping between reference and physical domain (left) and kinematic description of body under consideration, NACA0012 airfoil (right).

In this case, Γ is the surface of the airfoil, $\mathbf{e}_1 \in \mathbb{R}^{n_{sd}}$ is the 1st canonical unit vector, $\mathbf{f}(\mathbf{U}, \boldsymbol{\mu}, t) \in \mathbb{R}^{n_{sd}}$ is the instantaneous force that the fluid exerts on the airfoil, and $\dot{\mathbf{x}}$ is the pointwise velocity of airfoil. The solver-consistent discretization, discussed in Section 2.1 and [54], of these quantities results in the fully discrete approximations $W(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu})$ and $J_x(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu})$. The *instantaneous* quantities of interest corresponding to those in (51) are the power and x -directed force the fluid exerts on the airfoil, which take the form

$$\mathcal{P}(\mathbf{U}, \boldsymbol{\mu}, t) = \int_{\Gamma} \mathbf{f}(\mathbf{U}, \boldsymbol{\mu}, t) \cdot \dot{\mathbf{x}} dS \quad \text{and} \quad \mathcal{F}_x(\mathbf{U}, \boldsymbol{\mu}, t) = \int_{\Gamma} \mathbf{f}(\mathbf{U}, \boldsymbol{\mu}, t) \cdot \mathbf{e}_1 dS.$$

Define $\mathcal{P}^h(\mathbf{u}, \boldsymbol{\mu}, t)$ and $\mathcal{F}_x^h(\mathbf{u}, \boldsymbol{\mu}, t)$ as the solver-consistent semi-discretization of these instantaneous quantities of interest.

4.1. Time-Periodic Solutions of the Compressible Navier-Stokes Equations

In this section, the various solvers discussed in this document for determining primal and dual time-periodic solutions of partial differential equations are compared for a flapping airfoil in an isentropic, viscous flow. The stability of the periodic orbit is verified by performing an eigenvalue analysis of $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$. The section closes with validation of the adjoint method, introduced for efficient gradient computation of quantities of interest, against a second-order finite difference approximation.

Consider the NACA0012 airfoil in Figure 1 immersed in an isentropic, viscous flow with Reynolds number set to 1000. The Mach number is set to 0.2, in order to get a good compromise between approximate incompressibility and well-conditioned equations. The kinematic motion of the foil is parametrized with a single Fourier mode, i.e.,

$$\begin{aligned} h(\boldsymbol{\mu}, t) &= A_h \sin(\omega_h t + \phi_h) + c_h \\ \theta(\boldsymbol{\mu}, t) &= A_\theta \sin(\omega_\theta t + \phi_\theta) + c_\theta. \end{aligned} \quad (52)$$

The vector of parameters is fixed for the remainder of this section

$$\boldsymbol{\mu} = [A_h \ \omega_h \ \phi_h \ c_h \ A_\theta \ \omega_\theta \ \phi_\theta \ c_\theta] = [1.0 \ 0.4\pi \ 0.0 \ 0.0 \ \frac{\pi}{15} \ 0.4\pi \ \frac{\pi}{2} \ 0.0], \quad (53)$$

and corresponds to the motion in Figure 2 with period $T = 5$. The mapping $\mathcal{G}(\mathbf{X}, t)$ from the fixed reference domain V to the physical domain $\Omega(\boldsymbol{\mu}, t)$ takes the form of a parametrized rigid body motion

$$\mathcal{G}(\mathbf{X}, t) = \mathbf{v}(\boldsymbol{\mu}, t) + \mathbf{Q}(\boldsymbol{\mu}, t)(\mathbf{X} - \mathbf{x}_0) + \mathbf{x}_0, \quad (54)$$

where \mathbf{x}_0 is the location of pitching axis in the reference configuration (the 1/3 chord) and

$$\mathbf{Q}(\boldsymbol{\mu}, t) = \begin{bmatrix} \cos \theta(\boldsymbol{\mu}, t) & \sin \theta(\boldsymbol{\mu}, t) \\ -\sin \theta(\boldsymbol{\mu}, t) & \cos \theta(\boldsymbol{\mu}, t) \end{bmatrix} \quad \mathbf{v}(\boldsymbol{\mu}, t) = \begin{bmatrix} 0 \\ h(\boldsymbol{\mu}, t) \end{bmatrix}.$$

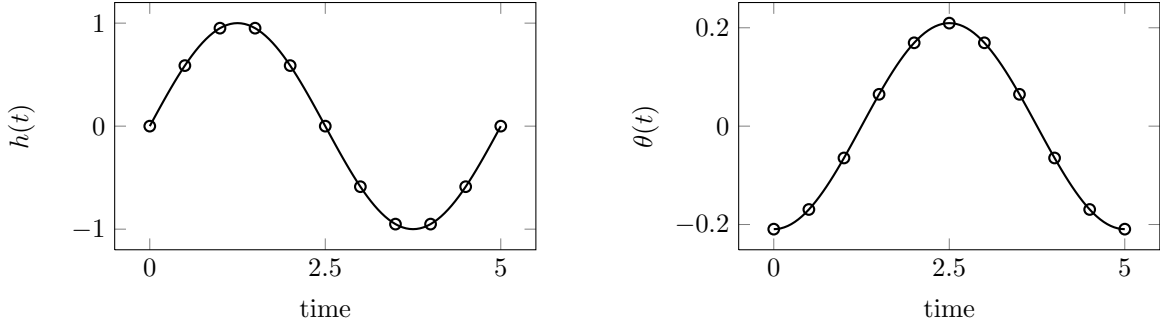


Figure 2: Trajectories of $h(t)$ and $\theta(t)$ that define the motion of the airfoil in Figure 1 and will be used to study primal and dual time-periodic solvers.

The isentropic Navier-Stokes equations are discretized with the discontinuous Galerkin scheme of Section 4 using 978 triangular $p = 3$ elements. No-slip boundary conditions are imposed on the airfoil wall and characteristic free-stream boundary conditions at the far-field. The temporal discretization uses a third-order diagonally implicit Runge-Kutta solver with 100 equally spaced steps to discretize a single period of the motion. The airfoil and surrounding fluid vorticity field are shown in Figures 3 and 4 with the flow field initialized from steady-state flow and the time-periodic initial condition, respectively. It is clear that the flow in Figure 4 will seamlessly transition between periods. The initialization from the steady-state solution in Figure 3 will introduce non-physical transients into the flow as discussed in the next section.

First, the solvers introduced in Section 2.2 are compared for different initial guesses for the time-periodic initial condition. In the absence of any a-priori information regarding the time-periodic solution, a reasonable initial guess is the steady-state flow. Since the problem under consideration is being forced by an input – the periodic motion of the foil – a mechanism for improving the initial guess is to simulate the flow field for m periods of the foil motion and use the final state of the final period as the initial guess. This corresponds to using m iterations of fixed point iteration (Algorithm 1) as a nonlinear preconditioner for the nonlinear system of equations (9) that enforces time-periodicity of the flow.

Figure 5 and Table 1 compare the solvers under consideration for different levels of nonlinear preconditioning. The iteration count is a good metric for comparison as the bulk of the time in any of these methods is timestepping the forward (primal) problem, the sensitivity equations, and/or the adjoint equations, which are the same from one method to the next. In our implementation, the linearized equations (sensitivity and adjoint) are about $2\times$ less expensive to solve than the nonlinear, primal equations – see discussion in Section 2.2. The linear algebra involved inside e.g. the L-BFGS and GMRES algorithms is negligible in comparison to these timestepping costs. Regardless of nonlinear preconditioning, the Newton-GMRES solver converges most rapidly for a range of linear system tolerances from 10^{-2} to 10^{-4} and the optimization algorithms (steepest descent and L-BFGS) converge most slowly. In fact, without any nonlinear preconditioning the optimization algorithms fail to make progress toward the optimal solution and were not included in the figure. Nonlinear preconditioning helps the Newton-GMRES algorithm most substantially, particularly with $m = 5$, as this appears to place the initial guess close enough to the solution that quadratic convergence is obtained from the outset. This causes the number of Newton iterations to be reduced from 8 or 9 to 3 or 4. From Table 1, this does not save many primal solves – since the nonlinear preconditioning requires primal solves – but requires far fewer linear system iterations and therefore fewer sensitivity solutions. Figure 6 isolates the Newton-GMRES solver (for $m = 0$, i.e., the case without preconditioning) to highlight convergence rates for different GMRES tolerances. It also shows the convergence of GMRES for each nonlinear iteration and each tolerance considered. As expected, more GMRES iterations are required near convergence as it becomes more difficult to reduce the linear residual the prescribed orders of magnitude.

The time history of the instantaneous quantities of interest in Figure 7 illustrate the non-physical transients that result from initializing the flow with the steady-state solution. While the transients mostly vanish after a single Newton iteration, the trajectories of these quantities of interest do not coincide with those of

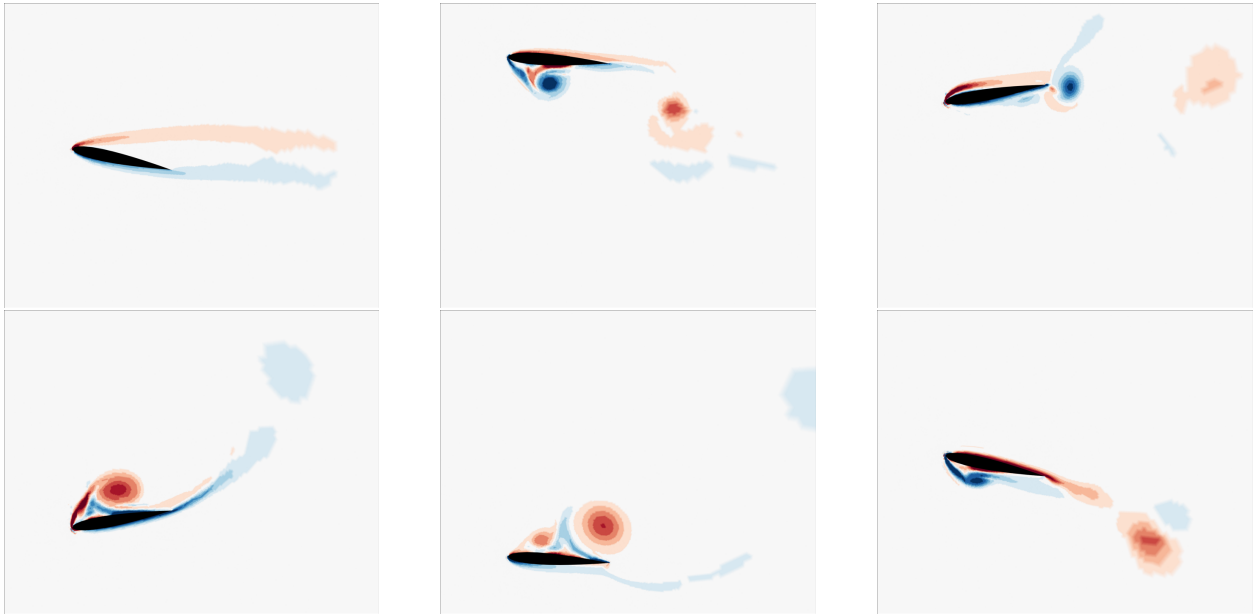


Figure 3: Flow vorticity around heaving/pitching airfoil for simulation initialized from steady state flow. Non-physical transients are introduced at the beginning of the time interval that result in non-trivial errors in integrated quantities of interests. Snapshots taken at times $t = 0.0, 1.0, 2.0, 3.0, 4.0, 5.0$.

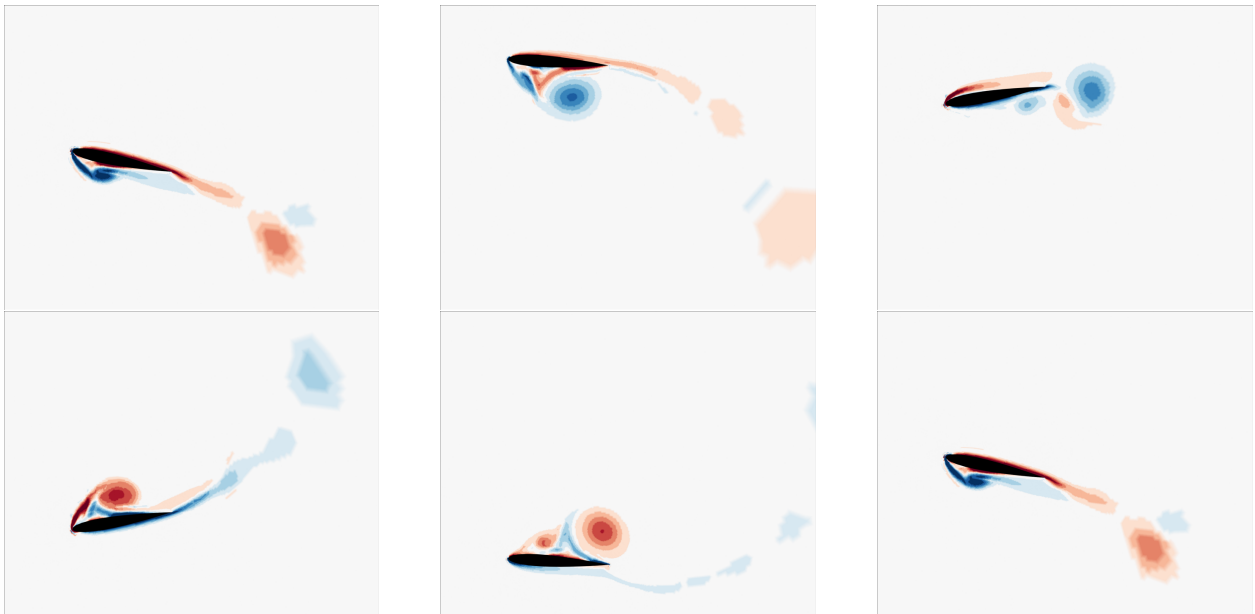


Figure 4: Time-periodic flow vorticity around heaving/pitching airfoil, i.e., initialized from periodic initial condition. The time-periodic initial condition ensures transients are not introduced at the beginning of the simulation; the result is a seamless transition between periods, as would be experienced in-flight, and trusted integrated quantities of interest. Snapshots taken at times $t = 0.0, 1.0, 2.0, 3.0, 4.0, 5.0$.

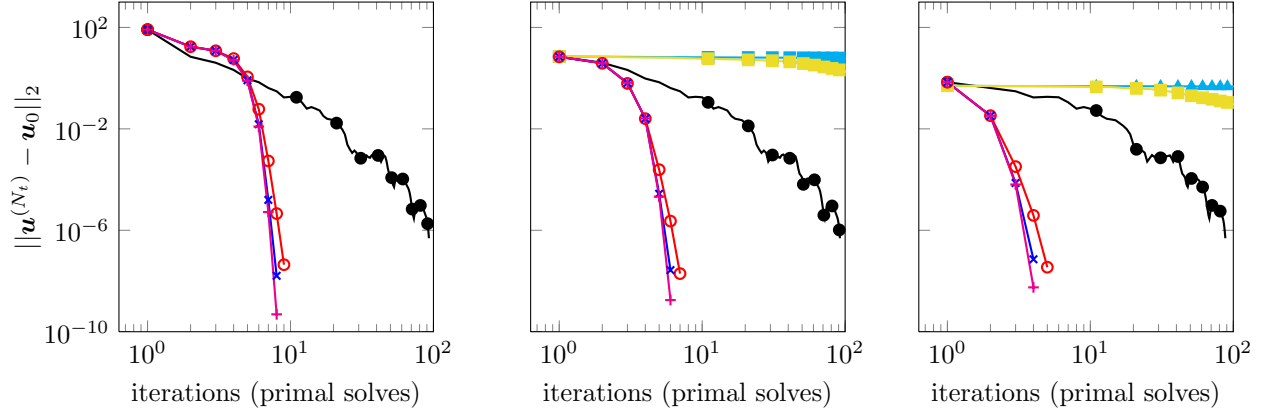


Figure 5: Convergence comparison for numerical solvers for fully discrete time-periodically constrained partial differential equations (4), nonlinearly preconditioned with m fixed point iterations. Left: $m = 0$, middle: $m = 1$, right: $m = 5$. Solvers: fixed point iteration (\bullet), steepest decent (\blacktriangle), L-BFGS (\blacksquare), Newton-GMRES: $\epsilon = 10^{-2}$ (\circ), $\epsilon = 10^{-3}$ (\ast), $\epsilon = 10^{-4}$ (\dagger). The optimization algorithms (steepest decent and L-BFGS) were not included in the $m = 0$ study due to lack of convergence issues.

$m = 0$	$\ \mathbf{u}^{(N_t)} - \mathbf{u}_0\ _2$	Primal Solves (8)	Sensitivity Solves (16)	Adjoint Solves (30)
Fixed Point Iteration	8.10e-07	90	0	0
Newton-Krylov (10^{-2})	4.41e-08	9	128	0
Newton-Krylov (10^{-3})	1.60e-08	8	156	0
Newton-Krylov (10^{-4})	4.85e-10	8	220	0
$m = 1$	$\ \mathbf{u}^{(N_t)} - \mathbf{u}_0\ _2$	Primal Solves (8)	Sensitivity Solves (16)	Adjoint Solves (30)
Fixed Point Iteration	8.10e-07	90	0	0
Steepest Decent	6.09e+00	121	0	121
L-BFGS	1.36e+00	121	0	121
Newton-Krylov (10^{-2})	1.96e-08	8	104	0
Newton-Krylov (10^{-3})	2.69e-08	7	116	0
Newton-Krylov (10^{-4})	1.77e-09	7	149	0
$m = 5$	$\ \mathbf{u}^{(N_t)} - \mathbf{u}_0\ _2$	Primal Solves (8)	Sensitivity Solves (16)	Adjoint Solves (30)
Fixed Point Iteration	8.10e-07	90	0	0
Steepest Decent	4.65e-01	125	0	125
L-BFGS	7.40e-02	125	0	125
Newton-Krylov (10^{-2})	3.50e-08	10	92	0
Newton-Krylov (10^{-3})	7.18e-08	9	88	0
Newton-Krylov (10^{-4})	5.61e-09	9	121	0

Table 1: Table summarizing performance of numerical solvers for fully discrete time-periodic partial differential equations, considering nonlinear preconditioning via m fixed point iterations.

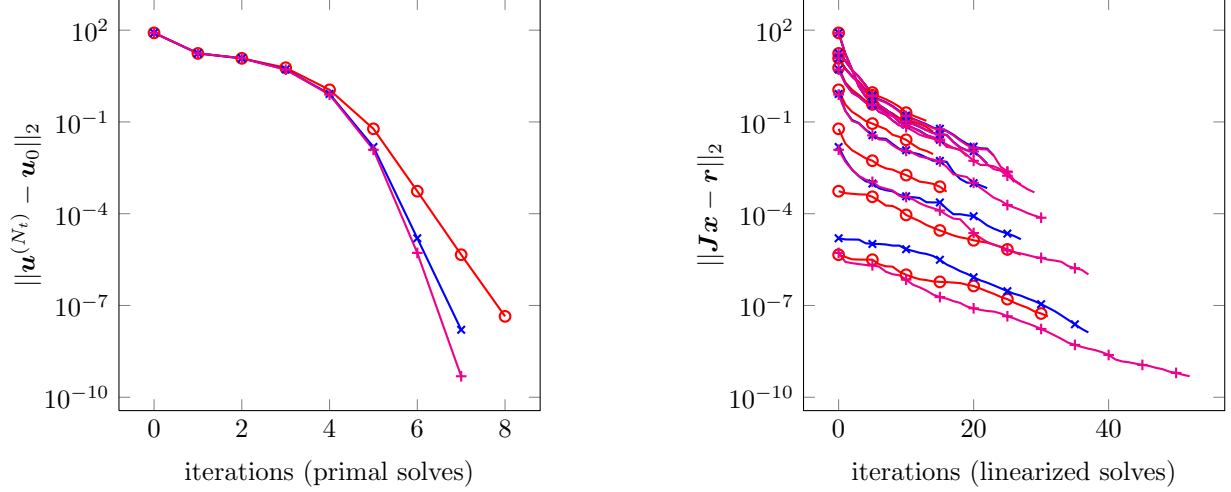


Figure 6: Linear and nonlinear convergence of Newton-GMRES method for determining fully discrete time-periodic solutions with various linear system tolerances, ϵ , i.e. $\|\mathbf{J}\mathbf{x} - \mathbf{R}\| < \epsilon$, where \mathbf{r} and \mathbf{J} are defined in (14) and (15). Tolerances considered: $\epsilon = 10^{-2}$ (\circ), $\epsilon = 10^{-3}$ (\times), $\epsilon = 10^{-4}$ ($+$).

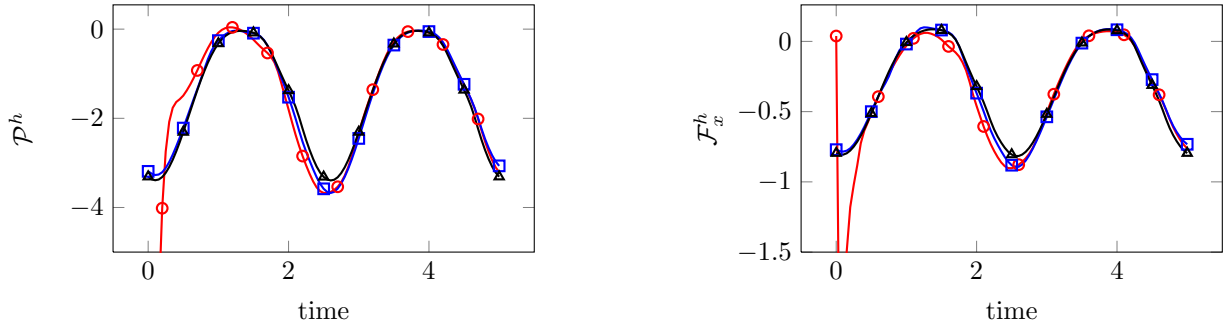


Figure 7: Time history of power, $\mathcal{F}_x^h(\mathbf{u}, \boldsymbol{\mu}, t)$, and x -directed force, $\mathcal{P}^h(\mathbf{u}, \boldsymbol{\mu}, t)$, after k Newton-GMRES iterations ($\epsilon = 10^{-2}$) starting from steady-state. Values of k : 0 (\circ), 1 (\square), and 8 (\blacktriangle).

the true time-periodic solution. The error between the integrated quantities of interest – W and J_x – at the time-periodic flow versus intermediate iterations is shown in Figure 8. Comparing Figures 5 and 8, it can be seen that a tolerance of 10^{-8} on $\|\mathbf{u}^{(N_t)} - \mathbf{u}^{(0)}\|_2$ leads to an accuracy of 10^{-6} in the integrated quantities of the time-periodic solution.

Next, the stability of the periodic orbit is verified by considering the eigenvalues of $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$, evaluated at the time-periodic solution. As discussed in Section 2.3 and many prior works [77, 78], the periodic orbit is stable if all eigenvalues of this matrix have modulus less than unity. Figure 9 shows that the 200 eigenvalues of largest modulus lie within the unit circle in the complex plane; thus, the periodic orbit is stable for this problem.

This completes the discussion of the primal time-periodic problem and attention is turned to the dual, or adjoint, problem. First, a brief comparison of two potential solvers – fixed point iteration and GMRES – for the periodic adjoint equation is provided. In contrast to the primal problem, there is a less pronounced difference between the convergence of fixed point iteration and the Krylov solver in the dual problem. Figure 10 shows the convergence history for two different right-hand sides of $\mathbf{A}\mathbf{x} = \mathbf{b}$, each corresponding to the adjoint method for a different quantity of interest. However, it should be noted that the iterations for

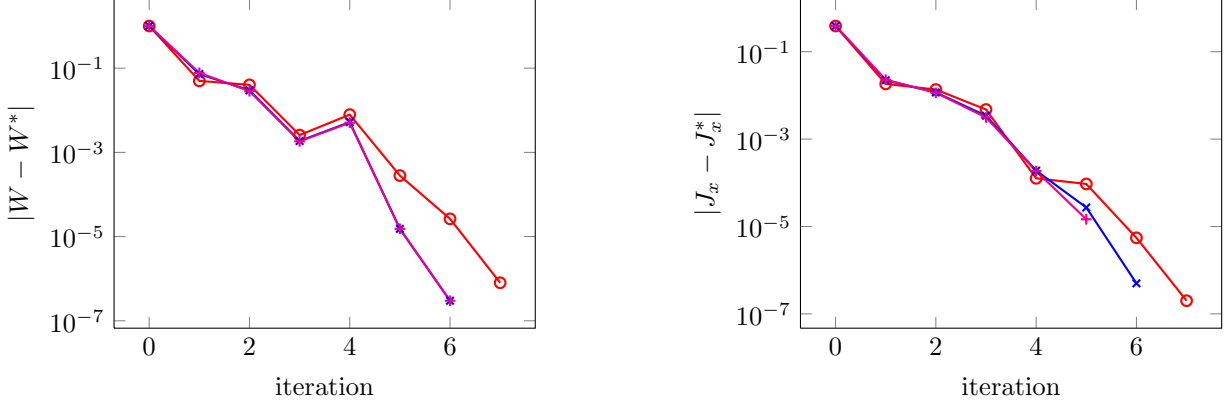


Figure 8: Convergence of fully discrete quantities of interest to their values at the time-periodic solution, W^* and J_x^* , for various solvers, without nonlinear preconditioning. Solvers: Newton-GMRES: $\epsilon = 10^{-2}$ (\circ), $\epsilon = 10^{-3}$ (\star), $\epsilon = 10^{-4}$ (\ast).

	$\frac{\partial W}{\partial A_h}$	$\frac{\partial W}{\partial \omega_h}$	$\frac{\partial W}{\partial A_\theta}$	$\frac{\partial W}{\partial c_\theta}$
Adjoint	-2.3091901647e+01	-2.59357909093e+01	-7.9956810720e+00	5.88159501747e-01
Finite difference	-2.3091901345e+01	-2.59357939542e+01	-7.9956815124e+00	5.88159491776e-01
	$\frac{\partial J_x}{\partial A_h}$	$\frac{\partial J_x}{\partial \omega_h}$	$\frac{\partial J_x}{\partial A_\theta}$	$\frac{\partial J_x}{\partial c_\theta}$
Adjoint	-1.8543679019e-01	-1.02983075340e-01	6.7297082256e+00	1.27010690706e-02
Finite difference	-1.8543677450e-01	-1.02983412630e-01	6.7297089105e+00	1.27011295647e-02

Table 2: Comparison of non-zero derivatives of total energy, W , and x -impulse, J_x , computed with the adjoint method and a second-order finite difference approximation with step size $\tau = 10^{-6}$.

the GMRES solver are cheaper than those of the fixed point solver as the terms $\frac{\partial F}{\partial \mathbf{u}^{(n-1)}}$ and $\frac{\partial F}{\partial \mathbf{k}_i^{(n)}}$ are not computed. Therefore, the GMRES algorithm is superior to fixed point iterations as there are fewer required iterations, each of which is cheaper.

Finally, the adjoint method for computing gradients of quantities of interest on the manifold of time-periodic solutions of the partial differential equations is verified against a second-order finite difference approximations. The finite difference approximation to gradients on the aforementioned manifold requires finding the *time-periodic* solution of the governing equations *at perturbations* about the nominal parameter configuration in (53). Figure 11 and Table 2 show the relative error between the gradients computed via the adjoint method in Algorithm 4 to this finite difference approximation for a sweep of finite difference intervals, τ . To realize the sub- 10^{-6} finite difference errors in the time-periodic gradient, tolerances of 10^{-12} were used for the primal and dual time-periodic solutions. As expected, the error starts to increase after τ drops too small due to the trade-off between finite difference accuracy and round-off error.

4.2. Energetically Optimal Flapping with Thrust and Time-Periodicity Constraints

In this section, the periodic adjoint method is used to solve an optimal control problem with *time-periodicity constraints* using gradient-based optimization techniques. The optimization problem is to determine the *energetically optimal* flapping motion of the NACA0012 airfoil in isentropic, viscous flow – over a single *representative*, in-flight period – such that the x -directed impulse on the body is identically 0. The

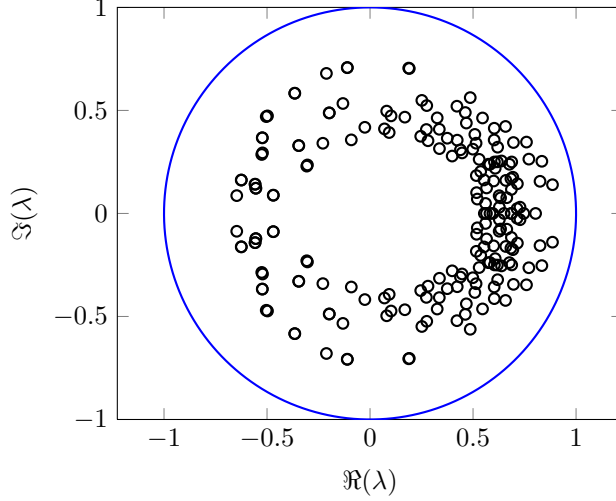


Figure 9: First 200 eigenvalues (\circ) of $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0}$ – evaluated at periodic solution – with largest magnitude. All eigenvalues lie in unit circle, thus the periodic orbit is stable.

continuous form of the optimal control problem is given as

$$\begin{aligned}
 & \underset{\mathbf{U}, \boldsymbol{\mu}}{\text{minimize}} && \mathcal{W}(\mathbf{U}, \boldsymbol{\mu}) \\
 & \text{subject to} && \mathcal{J}_x(\mathbf{U}, \boldsymbol{\mu}) = 0 \\
 & && \mathbf{U}(\mathbf{x}, 0) = \mathbf{U}(\mathbf{x}, T) \\
 & && \frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{U}, \nabla \mathbf{U}) = 0 \quad \text{in } \Omega(\boldsymbol{\mu}, t).
 \end{aligned} \tag{55}$$

After spatial and temporal discretization via the high-order discontinuous Galerkin and diagonally implicit Runge-Kutta schemes in Section 4, the continuous optimization problem in (55) is replaced with its fully discrete counterpart

$$\begin{aligned}
 & \underset{\substack{\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)} \in \mathbb{R}^{N_u}, \\ \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)} \in \mathbb{R}^{N_k}, \\ \boldsymbol{\mu} \in \mathbb{R}^{N_\mu}}}{\text{minimize}} && \mathcal{W}(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu}) \\
 & \text{subject to} && \mathcal{J}_x(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu}) = 0 \\
 & && \mathbf{u}^{(0)} = \mathbf{u}^{(N_t)} \\
 & && \mathbf{u}^{(n)} = \mathbf{u}^{(n-1)} + \sum_{i=1}^s b_i \mathbf{k}_i^{(n)} \\
 & && \mathbb{M} \mathbf{k}_i^{(n)} = \Delta t_n \mathbf{r} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right).
 \end{aligned} \tag{56}$$

The physical and numerical setup are identical to that in Section 4.1 with the exception of the kinematic parametrization. Instead of a single Fourier mode, the kinematic motion is parametrized by cubic splines with 5 equally spaced knots and boundary conditions that enforce

$$\begin{aligned}
 h(\boldsymbol{\mu}, t) &= -h(\boldsymbol{\mu}, t + T/2) \\
 \theta(\boldsymbol{\mu}, t) &= -\theta(\boldsymbol{\mu}, t + T/2)
 \end{aligned} \tag{57}$$

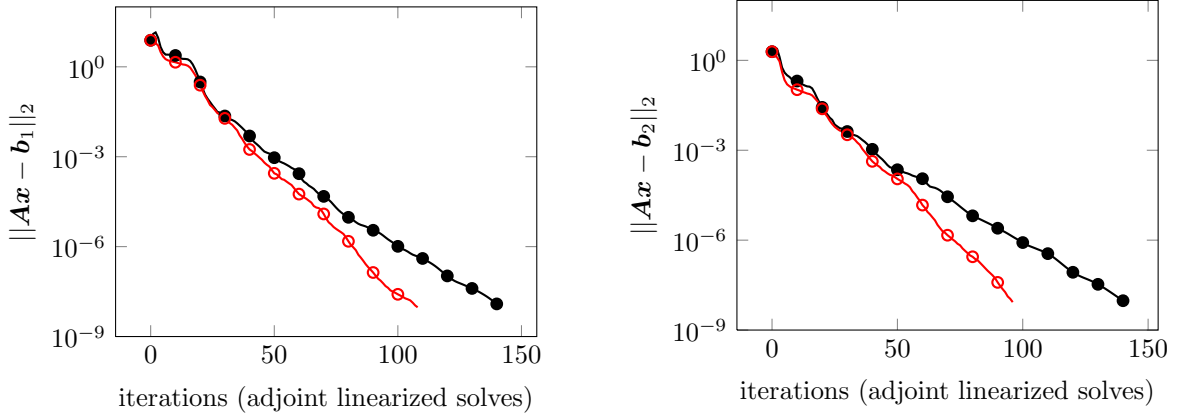


Figure 10: GMRES convergence for determining solution of adjoint equations corresponding to fully discrete time-periodic partial differential equation, i.e. a linear two-point boundary value problem. \mathbf{A} defined in (34), $\mathbf{b}_1 = \frac{\partial W}{\partial \mathbf{u}^{(N_t)}}$, and $\mathbf{b}_2 = \frac{\partial J_x}{\partial \mathbf{u}^{(N_t)}}$ from (33), where W is fully discrete approximation of the total work done by fluid on airfoil (Section 4) and J_x is the x-directed impulse. Solvers: fixed point iteration (\bullet) and GMRES (\circ). The linearization is performed about the time-periodic solution obtained with Newton-Krylov ($\epsilon = 10^{-4}$) method.

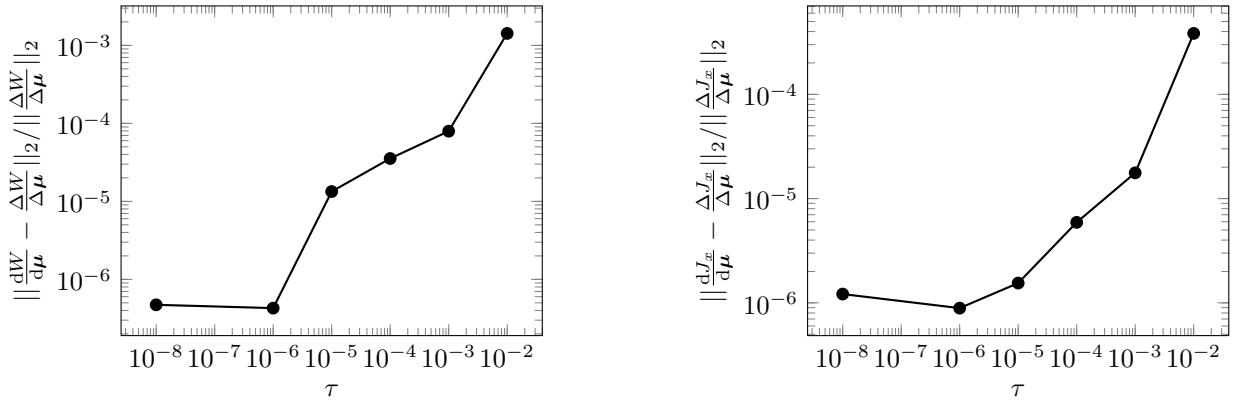


Figure 11: Verification of periodic adjoint-based gradient with second-order centered finite difference approximation, for a range of finite intervals, τ . The computed gradient match the finite difference approximation to nearly 7 digits before round-off errors degrade the accuracy.

where t is time and $T = 5$ is the fixed period of the flapping motion. The vector of parameters, $\boldsymbol{\mu}$ – used as optimization parameters – are the *knots* of the cubic splines. This leads to $N_{\boldsymbol{\mu}} = 8$ parameters; 4 knots for the motion of $h(\boldsymbol{\mu}, t)$ and $\theta(\boldsymbol{\mu}, t)$ ³. Notice that (57) enforces the trajectories of $h(\boldsymbol{\mu}, t)$ and $\theta(\boldsymbol{\mu}, t)$ in $[T/2, T]$ to be the mirror of those in $[0, T/2]$, which implicitly enforces periodicity with period T . The mapping \mathcal{G} from the reference to physical domain required for the DG-ALE formulation is defined in (54) with the new definition of $h(\boldsymbol{\mu}, t)$ and $\theta(\boldsymbol{\mu}, t)$ with periodic cubic splines.

The optimization problem in (56) is solved using the extension of the nested framework for PDE-constrained optimization, or generalized reduced-gradient method, introduced in Section 3.3. The solvers introduced in Section 2.2 will be used to determine the time-periodic flow around the airfoil. Given the results in the previous section, the Newton-GMRES method with a tolerance of $\Delta = 10^{-3}$, warm-started from $m = 5$ fixed-point iterations is employed. The flow is deemed to be periodic if

$$\|\mathbf{u}^{(0)} - \mathbf{u}^{(N_t)}\|_2 \leq 10^{-10}. \quad (58)$$

The periodic flow is used to compute quantities of interest – the total work and x -impulse. Then, the periodic adjoint method will be used to compute gradients of the quantities of interest along the manifold of *time-periodic* solutions of the governing equation. GMRES is used to solve the dual linear, periodic adjoint equations with a tolerance of $\Delta = 10^{-4}$. Since there are two quantities of interest, two periodic adjoint solves must be performed at each optimization iteration. Finally, the quantities of interest and their gradients are passed to an optimization solver – SNOPT [79] is used in this work – and progress is made toward a local minimum.

The initial condition for the optimization solver is shown in Figure 12; the heaving motion is a sinusoid with amplitude 1 and there is no pitch – pure heaving motion. The vorticity snapshots in Figure 15 show this motion induces a fairly violent flow with shedding vortices. The corresponding time history of the power, $\mathcal{P}^h(\mathbf{u}, \boldsymbol{\mu}, t)$, and x -directed force, $\mathcal{F}_x^h(\mathbf{u}, \boldsymbol{\mu}, t)$, imparted onto the airfoil by the fluid are shown in Figure 13. After 16 periodic optimization iterations, the first-order optimality conditions have been reduced by two orders of magnitude. From Figure 12, the optimal airfoil motion is a combination of heaving and pitching. From the initial guess, the amplitude of the heaving motion has been reduced by more than a factor of two and the pitching amplitude increased to 18.7° . The convergence history for the optimization solver is given in Figure 14. At the optimal solution, the total work required to perform the flapping motion is more than an order of magnitude smaller than at the initial guess (pure heaving). Figures 15 and 16 show snapshots of the flow in time at the initial, purely heaving motion and the optimal flapping motion. From these figures, it is clear that the flow corresponding to the optimal motion is relatively benign with no shedding vortices, which explains the reduction in required work. The efficiency of combined pitching and heaving has been repeatedly observed experimentally [80, 81, 82] and the phase angle of approximately 90° between pitching and heaving, as observed in Figure 12, has also been observed in experiments [80, 81, 82, 83]. The specific pitching and heaving amplitudes were determined by the optimizer such that the thrust constraint is satisfied; if the thrust requirement was increased, these magnitudes would increase and result in a more violent flow field, eventually leading to vortex shedding [80, 54].

³There are only 4 degrees of freedom since the mirror boundary condition in (57) prescribes the value of one of the knots given the other four.

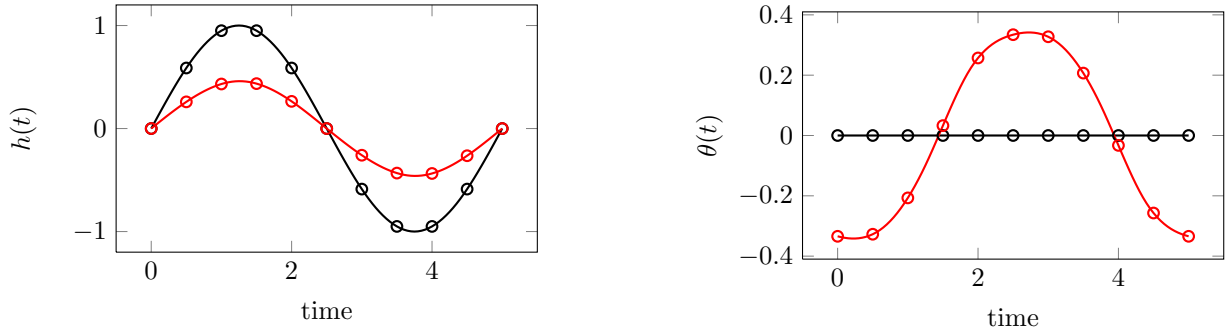


Figure 12: Trajectories of $h(t)$ and $\theta(t)$ at initial guess ($\text{---}\circ\text{---}$) and optimal solution ($\text{---}\circ\text{---}$) for optimization problem in (56).

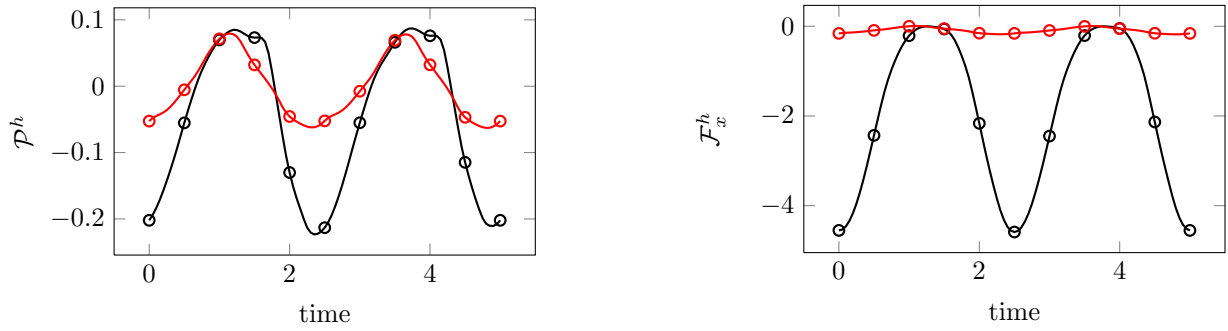


Figure 13: Time history of the power, $\mathcal{P}^h(\mathbf{u}, \boldsymbol{\mu}, t)$, and x -directed force, $\mathcal{F}_x^h(\mathbf{u}, \boldsymbol{\mu}, t)$, imparted onto foil by fluid at initial guess ($\text{---}\circ\text{---}$) and optimal solution ($\text{---}\circ\text{---}$) for optimization problem in (56).

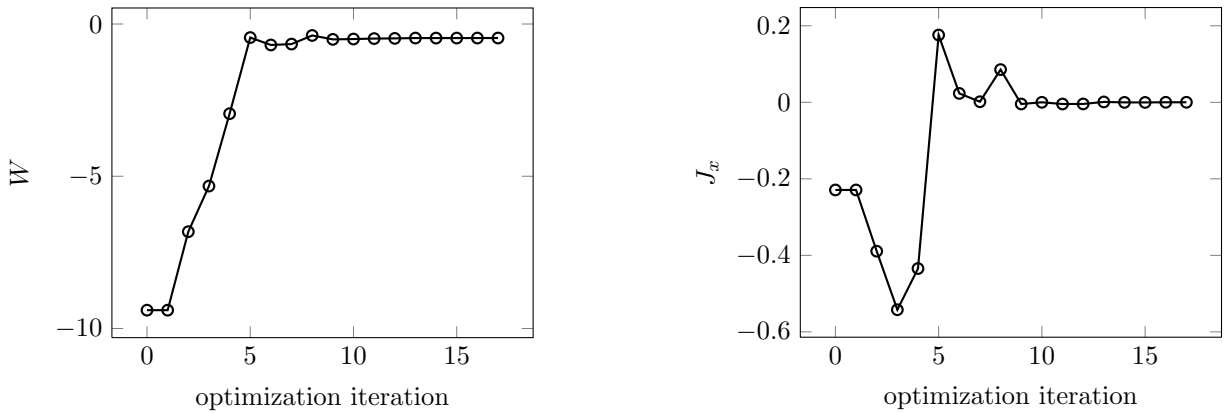


Figure 14: Convergence of quantities of interest, W and J_x , with optimization iteration. Each optimization iteration requires the a periodic flow computation and its corresponding adjoint to evaluate the quantities of interest and their gradients.

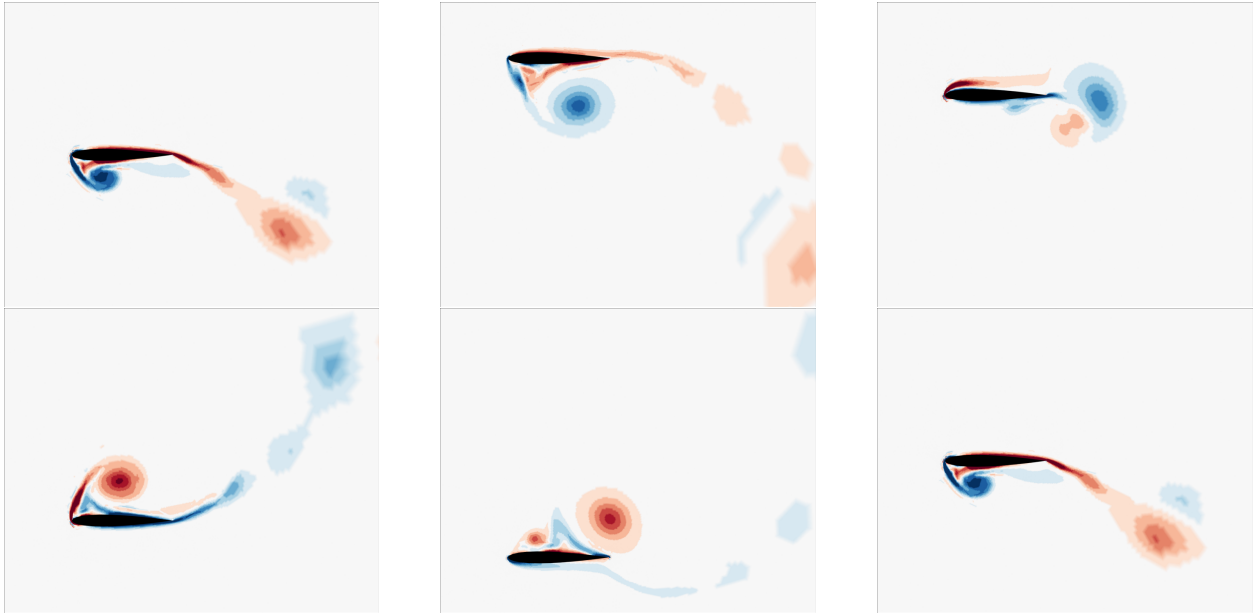


Figure 15: Trajectory of airfoil and flow vorticity at initial guess for optimization (pure heaving motion, see Figure 12). Snapshots taken at times $t = 0.0, 1.0, 2.0, 3.0, 4.0, 5.0$.

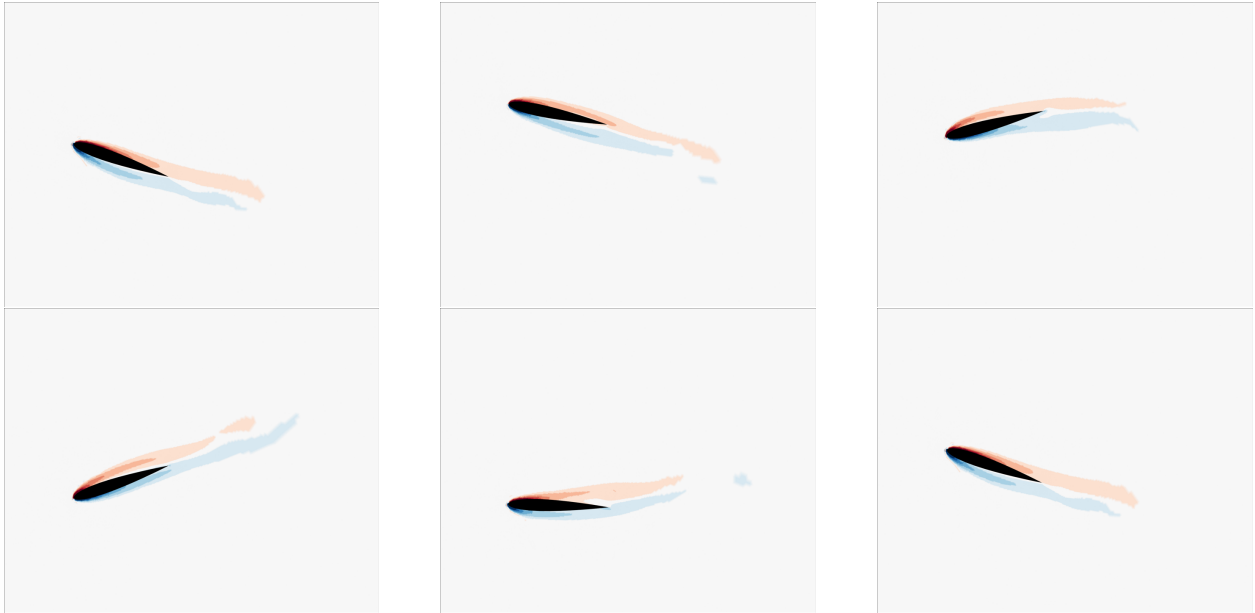


Figure 16: Trajectory of airfoil and flow vorticity at energetically optimal, zero-impulse flapping motion (see Figure 12). Snapshots taken at times $t = 0.0, 1.0, 2.0, 3.0, 4.0, 5.0$.

5. Conclusion

This document discussed a fully discrete framework for computing time-periodic solutions of partial differential equations. The discussion included the spatio-temporal discretization of the governing equations and a slew of time-periodic shooting solvers, including optimization-based and Newton-Krylov methods. These shooting methods consider the state at the final time to be a nonlinear function of the initial condition and solve $\mathbf{u}^{(N_t)}(\mathbf{u}_0) = \mathbf{u}_0$ using Newton-Raphson iterations or optimization techniques to minimize its norm. The linear system of equations, arising in the Newton-Raphson iterations, were solved using matrix-free GMRES with matrix-vector products computed as the solution of the linearized, sensitivity equations (with appropriate initial condition). The adjoint method was used to compute the gradients in the gradient-based optimization solvers. These periodic solvers were used to compute the time-periodic flow around a flapping airfoil in isentropic, compressible, viscous flow, and their performance compared. The Newton-Krylov solver exhibits superior convergence to the optimization-based shooting methods, even when inexact tolerances were used on the linear system solves, and fully leverages quality starting guesses. An eigenvalue analysis is provided to show the periodic orbit of the flapping problem is stable.

The main contribution of the document is the derivation of the adjoint equations corresponding to the fully discrete time-periodically constraint partial differential equations. As opposed to the backward-in-time evolution equations, these equations constitute a linear, *two-point boundary value problem* that is provably solvable. The corresponding adjoint method was introduced for computing *exact* gradients of quantities of interest along the manifold of time-periodic solutions of the discrete conservation law. The gradients were verified against a second-order finite difference approximation. These quantities of interest and their gradients were used in the context of gradient-based optimization to solve an optimal control problem with time-periodicity constraints, among others. In particular, the energetically optimal flapping motion of a 2D airfoil in *time-periodic*, isentropic, compressible, viscous flow that generates a prescribed time-averaged thrust is sought. The proposed framework improves the nominal flapping motion by reducing the flapping energy nearly an order of magnitude and exactly satisfies the thrust constraint.

While this work is an initial step toward problems of engineering and scientific relevance, additional development will be required to solve truly impactful problems. One extension of this work is the development of robust solvers for determining *nearly* time-periodic solutions of problems where a time-periodic solution does not exist, but exhibits quasi-cyclic behavior. An example of such a problem is the 3D turbulent flow around periodically driven bodies such as helicopter and windmill blades. Another extension will be the development of faster numerical solvers to reduce the cost of computing time-periodic solutions or solving optimization problems with time-periodicity constraints. For example, economical, matrix-free preconditioners could result in non-trivial speedups for the Newton-Krylov time-periodicity solver and Krylov solver for the periodic adjoint equations. Model order reduction techniques could dramatically reduce the cost of computing the solution of the primal partial differential equations, and consequently the entire time-periodic solver.

Appendix A. Existence and Uniqueness of Solutions of the Adjoint Equations of the Fully Discrete, Time-Periodically Constrained Partial Differential Equations

This section proves existence and uniqueness of solutions of the adjoint equations of the fully discrete, time-periodically constrained partial differential equation. The strategy is to show the linear operator that encapsulates them is the transpose of the linear operator that defines the fully discrete, sensitivity equations, which is assumed non-singular at a time-periodic solution.

Consider the initial-value problem (8), with the initial condition parametrized by $\boldsymbol{\mu}$,

$$\begin{aligned} \mathbf{u}^{(0)} &= \mathbf{u}_0(\boldsymbol{\mu}) \\ \mathbf{u}^{(n)} &= \mathbf{u}^{(n-1)} + \sum_{i=1}^s b_i \mathbf{k}_i^{(n)} \\ \mathbb{M} \mathbf{k}_i^{(n)} &= \Delta t_n \mathbf{r} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right). \end{aligned} \tag{A.1}$$

The fully discrete adjoint equations corresponding to the primal equation in (A.1) and the discrete quantity of interest, $F(\mathbf{u}^{(0)}, \dots, \mathbf{u}^{(N_t)}, \mathbf{k}_1^{(1)}, \dots, \mathbf{k}_s^{(N_t)}, \boldsymbol{\mu})$ are

$$\begin{aligned}\boldsymbol{\nu}^{(N_t)} &= \frac{\partial F}{\partial \mathbf{u}^{(N_t)}}{}^T \\ \boldsymbol{\nu}^{(n-1)} &= \boldsymbol{\nu}^{(n)} + \frac{\partial F}{\partial \mathbf{u}^{(n-1)}}{}^T + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \boldsymbol{\tau}_i^{(n)} \\ \mathbb{M}^T \boldsymbol{\tau}_i^{(n)} &= \frac{\partial F}{\partial \mathbf{k}_i^{(n)}}{}^T + b_i \boldsymbol{\nu}^{(n)} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \boldsymbol{\tau}_j^{(n)},\end{aligned}\tag{A.2}$$

and the gradient of the quantity of interest can be reconstructed as

$$\frac{dF}{d\boldsymbol{\mu}} = \frac{\partial F}{\partial \boldsymbol{\mu}} + \boldsymbol{\nu}^{(0)T} \frac{\partial \mathbf{u}_0}{\partial \boldsymbol{\mu}} + \sum_{n=1}^{N_t} \Delta t_n \sum_{i=1}^s \boldsymbol{\tau}_i^{(n)T} \frac{\partial \mathbf{r}}{\partial \boldsymbol{\mu}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right),\tag{A.3}$$

where $\boldsymbol{\nu}^{(n)}$ and $\boldsymbol{\tau}_i^{(n)}$ are the Lagrange multipliers. These equations can be obtained using an identical derivation to that in Section 3.1; see [54]. At this point, take $F = \mathbf{v}^T \mathbf{u}^{(N_t)}$ and $\boldsymbol{\mu} = \mathbf{u}_0$ for a fixed, arbitrary vector $\mathbf{v} \in \mathbb{R}^{N_u}$. For this selection of F and $\boldsymbol{\mu}$, the above equations reduce to

$$\begin{aligned}\boldsymbol{\nu}^{(N_t)} &= \mathbf{v} \\ \boldsymbol{\nu}^{(n-1)} &= \boldsymbol{\nu}^{(n)} + \sum_{i=1}^s \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_i^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_i \Delta t_n \right)^T \boldsymbol{\tau}_i^{(n)} \\ \mathbb{M}^T \boldsymbol{\tau}_i^{(n)} &= b_i \boldsymbol{\nu}^{(n)} + \sum_{j=i}^s a_{ji} \Delta t_n \frac{\partial \mathbf{r}}{\partial \mathbf{u}} \left(\mathbf{u}_j^{(n)}, \boldsymbol{\mu}, t_{n-1} + c_j \Delta t_n \right)^T \boldsymbol{\tau}_j^{(n)}\end{aligned}\tag{A.4}$$

and

$$\frac{dF^T}{d\boldsymbol{\mu}} = \frac{\partial \mathbf{u}^{(N_t)T}}{\partial \mathbf{u}_0} \mathbf{v} = \boldsymbol{\nu}^{(0)}.\tag{A.5}$$

The equations in (A.4) defining $\boldsymbol{\nu}^{(0)}$ are *identical* to those in (32) defining $\frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v}$, which leads to the relation

$$\frac{\partial \mathbf{u}^{(N_t)T}}{\partial \mathbf{u}_0} \mathbf{v} = \frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} \mathbf{v}\tag{A.6}$$

for any \mathbf{v} . Thus, it can be concluded that

$$\frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} = \frac{\partial \mathbf{u}^{(N_t)T}}{\partial \mathbf{u}_0}.\tag{A.7}$$

Since the Jacobian of the time-periodic residual, $\frac{\partial \mathbf{u}^{(N_t)}}{\partial \mathbf{u}_0} - \mathbf{I}$, is non-singular at a time-periodic solution, the matrix defining the linear, two-point boundary value problem, $\frac{\partial \boldsymbol{\lambda}^{(0)}}{\partial \boldsymbol{\lambda}_{N_t}} - \mathbf{I}$ must also be non-singular. Thus, a solution of the linear, two-point boundary value problem exists and is unique.

Acknowledgments

This work was supported in part by the Department of Energy Computational Science Graduate Fellowship Program of the Office of Science and National Nuclear Security Administration in the Department of

Energy under contract DE-FG02-97ER25308 (MZ), and by the Director, Office of Science, Computational and Technology Research, U.S. Department of Energy under contract number DE-AC02-05CH11231 (PP and JW). The content of this publication does not necessarily reflect the position or policy of any of these supporters, and no official endorsement should be inferred.

References

- [1] Z. Wang, “Vortex shedding and frequency selection in flapping flight,” *Journal of Fluid Mechanics*, vol. 410, pp. 323–341, 2000.
- [2] 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,” *Journal of Experimental Biology*, vol. 207, no. 3, pp. 449–460, 2004.
- [3] H. Soueid, L. Guglielmini, C. Airiau, and A. Bottaro, “Optimization of the motion of a flapping airfoil using sensitivity functions,” *Computers & Fluids*, vol. 38, no. 4, pp. 861–874, 2009.
- [4] J. Wilkening and A. E. Hosoi, “Shape optimization of a sheet swimming over a thin liquid layer,” *J. Fluid Mech.*, vol. 601, pp. 25–61, 2008.
- [5] J. Lohéac, J.-F. Scheid, and M. Tucsnak, “Controllability and time optimal control for low reynolds numbers swimmers,” *Acta Applicandae Mathematicae*, vol. 123, no. 1, pp. 175–200, 2013.
- [6] R. Trouilloud, S. Y. Tony, A. Hosoi, and E. Lauga, “Soft swimming: Exploiting deformable interfaces for low reynolds number locomotion,” *Physical review letters*, vol. 101, no. 4, p. 048102, 2008.
- [7] C. Eloy and L. Schouveiler, “Optimisation of two-dimensional undulatory swimming at high reynolds number,” *International Journal of Non-Linear Mechanics*, vol. 46, no. 4, pp. 568–576, 2011.
- [8] J. Dugundji and J. H. Wendell, “Some analysis methods for rotating systems with periodic coefficients,” *AIAA journal*, vol. 21, no. 6, pp. 890–897, 1983.
- [9] S. Hwang, *Frequency domain system identification of helicopter rotor dynamics incorporating models with time periodic coefficients*. 1997.
- [10] D. A. Peters, “Fast floquet theory and trim for multi-bladed rotorcraft,” *Journal of the American Helicopter Society*, vol. 39, no. 4, pp. 82–89, 1994.
- [11] V. Verdult*, M. Lovera, and M. Verhaegen, “Identification of linear parameter-varying state-space models with application to helicopter rotor dynamics,” *International Journal of Control*, vol. 77, no. 13, pp. 1149–1159, 2004.
- [12] I. Bucher and O. Shomer, “Detecting asymmetry in rotating structures a combined acutation and signal processing approach,” in *23rd International Modal Analysis Conference (IMAC XXIII)*, 2005.
- [13] M. Allen and J. H. Ginsberg, “Floquet modal analysis to detect cracks in a rotating shaft on anisotropic supports,” in *24th International Modal Analysis Conference (IMAC XXIV)*, 2006.
- [14] G. Bir and K. Stol, “Operating modes of a teetered-rotor wind turbine,” in *SPIE proceedings series*, pp. 1586–1596, Society of Photo-Optical Instrumentation Engineers, 1999.
- [15] K. Stol, M. Balas, and G. Bir, “Floquet modal analysis of a teetered-rotor wind turbine,” *Journal of solar energy engineering*, vol. 124, no. 4, pp. 364–371, 2002.
- [16] J. Oden and R. Lin, “On the general rolling contact problem for finite deformations of a viscoelastic cylinder,” *Computer Methods in Applied Mechanics and Engineering*, vol. 57, pp. 297–367, 1986.

- [17] S. Govindjee, T. Potter, and J. Wilkening, “Cyclic steady states of treaded rolling bodies,” *Int. J. Numer. Meth. Engng.*, vol. 99, pp. 203–220, 2014.
- [18] S. Govindjee, T. Potter, and J. Wilkening, “Dynamic stability of spinning viscoelastic cylinders at finite deformation,” *Int. J. Solids Struct.*, vol. 51, pp. 3589–3603, 2014.
- [19] E. J. Doedel, H. B. Keller, and J. P. Kernévez, “Numerical analysis and control of bifurcation problems: (I) Bifurcation in finite dimensions,” *Int. J. Bifurcation and Chaos*, vol. 1, no. 3, pp. 493–520, 1991.
- [20] E. J. Doedel, H. B. Keller, and J. P. Kernévez, “Numerical analysis and control of bifurcation problems: (II) Bifurcation in infinite dimensions,” *Int. J. Bifurcation and Chaos*, vol. 1, no. 4, pp. 745–772, 1991.
- [21] W. J. F. Govaerts, *Numerical methods for bifurcations of dynamical equilibria*. Philadelphia: SIAM, 2000.
- [22] H. B. Keller, *Numerical methods in bifurcation problems*. New York: Springer, 1987.
- [23] H. S. Brown, I. G. Kevrekidis, A. Oron, and P. Rosenau, “Bifurcations and pattern formation in the “regularized” Kuramoto-Sivashinsky equation,” *Physics Letters A*, vol. 163, pp. 299–308, 1992.
- [24] C. P. Tsai and D. S. Jeng, “Numerical Fourier solutions of standing waves in finite water depth,” *Appl. Ocean Res.*, vol. 16, pp. 185–193, 1994.
- [25] M. Okamura, “Almost limiting short-crested gravity waves in deep water,” *J. Fluid Mech.*, vol. 646, pp. 481–503, 2010.
- [26] K. C. Hall, J. P. Thomas, and W. S. Clark, “Computation of unsteady nonlinear flows in cascades using a harmonic balance technique,” *AIAA journal*, vol. 40, no. 5, pp. 879–886, 2002.
- [27] A. Gopinath and A. Jameson, “Time spectral method for periodic unsteady computations over two-and three-dimensional bodies,” *AIAA paper*, vol. 1220, pp. 10–13, 2005.
- [28] M. McMullen, A. Jameson, and J. Alonso, *Application of a non-linear frequency domain solver to the Euler and Navier-Stokes equations*. American Institute of Aeronautics and Astronautics, 2002.
- [29] S. Nadarajah and A. Jameson, “Optimum shape design for unsteady three-dimensional viscous flows using a nonlinear frequency-domain method,” *Journal of Aircraft*, vol. 44, no. 5, pp. 1513–1527, 2007.
- [30] H. B. Keller, *Numerical Methods for Two-Point Boundary-Value Problems*. New York: Blaisdell, 1968.
- [31] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*. New York: Springer, third ed., 2002.
- [32] M. O. Williams, J. Wilkening, E. Shlizerman, and J. N. Kutz, “Continuation of periodic solutions in the waveguide array mode-locked laser,” *Physica D*, vol. 240, no. 22, pp. 1791–1804, 2011.
- [33] J. Wilkening, “Breakdown of self-similarity at the crests of large amplitude standing water waves,” *Phys. Rev. Lett*, vol. 107, p. 184501, 2011.
- [34] J. Wilkening and J. Yu, “Overdetermined shooting methods for computing standing water waves with spectral accuracy,” *Computational Science & Discovery*, vol. 5, p. 014017, 2012.
- [35] C. H. Rycroft and J. Wilkening, “Computation of three-dimensional standing water waves,” *J. Comput. Phys.*, vol. 255, pp. 612–638, 2013.
- [36] E. Isaacson, *Some periodic solutions of the two-dimensional Stokes–Oldroyd–B system with stress diffusion*. PhD thesis, University of California, Berkeley, May 2012.
- [37] K. Lust and D. Roose, “Newton-Picard methods with subspace iteration for computing periodic solutions of partial differential equations,” *ZAMM*, vol. 76, no. S2, pp. 605–606, 1996.

- [38] G. N. Mercer and A. J. Roberts, “Standing waves in deep water: Their stability and extreme form,” *Phys. Fluids A*, vol. 4, no. 2, pp. 259–269, 1992.
- [39] C. G. Broyden, “The convergence of a class of double-rank minimization algorithms, Parts I and II,” *J. Inst Maths Applies*, vol. 6, pp. 76–90, 222–231, 1970.
- [40] D. M. Ambrose and J. Wilkening, “Computation of time-periodic solutions of the Benjamin–Ono equation,” *J. Nonlinear Sci.*, vol. 20, no. 3, pp. 277–308, 2010.
- [41] D. M. Ambrose and J. Wilkening, “Global paths of time-periodic solutions of the Benjamin–Ono equation connecting pairs of traveling waves,” *Comm. App. Math. and Comp. Sci.*, vol. 4, no. 1, pp. 177–215, 2009.
- [42] D. M. Ambrose and J. Wilkening, “Computation of symmetric, time-periodic solutions of the vortex sheet with surface tension,” *Proc. Nat. Acad. Sci.*, vol. 107, no. 8, pp. 3361–3366, 2010.
- [43] J. M. Hamilton, J. Kim, and F. Waleffe, “Regeneration mechanisms of near-wall turbulence structures,” *J. Fluid Mech.*, vol. 287, pp. 317–348, 1995.
- [44] G. Kawahara and S. Kida, “Periodic motion embedded in plane Couette turbulence: regeneration cycle and burst,” *J. Fluid Mech.*, vol. 449, pp. 291–300, 2001.
- [45] P. Cvitanović, R. Artuso, R. Mainieri, G. Tanner, and G. Vattay, *Chaos: Classical and Quantum*. Copenhagen: ChaosBook.org, Niels Bohr Institute, 2005.
- [46] D. Viswanath, “Recurrent motions within plane Couette turbulence,” *J. Fluid Mech.*, vol. 580, pp. 339–358, 2007.
- [47] T. M. Schneider, B. Eckhardt, and J. A. Yorke, “Turbulence transition and the edge of chaos in pipe flow,” *Phys. Rev. Letters*, vol. 99, no. 3, p. 034502, 2007.
- [48] T. M. Schneider, B. Eckhardt, and J. Vollmer, “Statistical analysis of coherent structures in transitional pipe flow,” *Phys. Rev. E*, vol. 75(2), no. 6, p. 066313, 2007.
- [49] U. Pesavento and Z. J. Wang, “Flapping wing flight can save aerodynamic power compared to steady flight,” *Physical review letters*, vol. 103, no. 11, p. 118102, 2009.
- [50] M. Culbreth, Y. Allaneau, and A. Jameson, “High-fidelity optimization of flapping airfoils and wings,” *AIAA paper-Hawaii 2011*, 2011.
- [51] E. J. Nielsen, B. Diskin, and N. K. Yamaleev, “Discrete adjoint-based design optimization of unsteady turbulent flows on dynamic unstructured grids,” *AIAA Journal*, vol. 48, no. 6, pp. 1195–1206, 2010.
- [52] M. P. van Schroyen Lantman and K. Fidkowski, “Adjoint-based optimization of flapping kinematics in viscous flows,” in *21st AIAA Computational Fluid Dynamics Conference*, 2013.
- [53] M. Jones and N. K. Yamaleev, “Adjoint based shape and kinematics optimization of flapping wing propulsive efficiency,” 43rd AIAA Fluid Dynamics Conference. San Diego, CA, 2013. AIAA 2013-2472, 2013.
- [54] M. J. Zahr and P.-O. Persson, “Fully-discrete, time-dependent adjoint method for high-order discontinuous galerkin discretizations on deforming domains: Application to pde-constrained optimization,” *Journal of Computational Physics*, 2016.
- [55] S. K. Nadarajah and A. Jameson, “Optimum shape design for unsteady flows with time-accurate continuous and discrete adjoint method,” *AIAA Journal*, vol. 45, pp. 1478–1491, 2015/06/22 2007.

- [56] T. D. Economon, F. Palacios, and J. J. Alonso, “Unsteady continuous adjoint approach for aerodynamic design on dynamic meshes,” *AIAA Journal*, vol. 53, no. 9, pp. 2437–2453, 2015.
- [57] B. K. Stanford and P. S. Beran, “Cost reduction techniques for the design of non-linear flapping wing structures,” *International Journal for Numerical Methods in Engineering*, vol. 88, no. 6, pp. 533–555, 2011.
- [58] M. J. Zahr and P.-O. Persson, “High-order, time-dependent aerodynamic optimization using a discontinuous Galerkin discretization of the Navier-Stokes equations,” in *AIAA Science and Technology Forum and Exposition (SciTech 2016)*, (San Diego, California), January 4–8 (2016).
- [59] S. Vandewalle and R. Piessens, “Efficient parallel algorithms for solving initial-boundary value and time-periodic parabolic partial differential equations,” *SIAM journal on scientific and statistical computing*, vol. 13, no. 6, pp. 1330–1346, 1992.
- [60] R. Srzednicki, “Periodic and bounded solutions in blocks for time-periodic nonautonomous ordinary differential equations,” *Nonlinear Analysis: Theory, Methods & Applications*, vol. 22, no. 6, pp. 707–737, 1994.
- [61] P. E. Gill, W. Murray, and M. H. Wright, “Practical optimization,” 1981.
- [62] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, “Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 23, no. 4, pp. 550–560, 1997.
- [63] J. Nocedal and S. Wright, *Numerical optimization, series in operations research and financial engineering*. Springer, 2006.
- [64] M. D. Gunzburger, *Perspectives in flow control and optimization*, vol. 5. Siam, 2003.
- [65] D. A. Knoll and D. E. Keyes, “Jacobian-free Newton–Krylov methods: a survey of approaches and applications,” *J. Comput. Phys.*, vol. 193, pp. 357–397, 2004.
- [66] V. Simoncini and E. Gallopoulos, “An iterative method for nonsymmetric systems with multiple right-hand sides,” *SIAM Journal on Scientific Computing*, vol. 16, no. 4, pp. 917–933, 1995.
- [67] T. F. Chan and W. L. Wan, “Analysis of projection methods for solving linear systems with multiple right-hand sides,” *SIAM Journal on Scientific Computing*, vol. 18, no. 6, pp. 1698–1721, 1997.
- [68] M. H. Gutknecht, “Block krylov space methods for linear systems with multiple right-hand sides: an introduction,” 2006.
- [69] C.-K. Lin, “On the incompressible limit of the compressible navier-stokes equations,” *Communications in partial differential equations*, vol. 20, no. 3-4, pp. 677–707, 1995.
- [70] B. Desjardins, E. Grenier, P.-L. Lions, and N. Masmoudi, “Incompressible limit for solutions of the isentropic navier–stokes equations with dirichlet boundary conditions,” *Journal de Mathématiques Pures et Appliquées*, vol. 78, no. 5, pp. 461–471, 1999.
- [71] B. M. Froehle, *High-order discontinuous Galerkin fluid-structure interaction methods*. University of California, Berkeley, 2013.
- [72] P.-O. Persson, J. Bonet, and J. Peraire, “Discontinuous galerkin solution of the navier–stokes equations on deformable domains,” *Computer Methods in Applied Mechanics and Engineering*, vol. 198, no. 17, pp. 1585–1595, 2009.
- [73] P. Thomas and C. Lombard, “Geometric conservation law and its application to flow computations on moving grids,” *AIAA journal*, vol. 17, no. 10, pp. 1030–1037, 1979.

- [74] J. Peraire and P.-O. Persson, “The compact discontinuous galerkin (cdg) method for elliptic problems,” *SIAM Journal on Scientific Computing*, vol. 30, no. 4, pp. 1806–1824, 2008.
- [75] P.-O. Persson and J. Peraire, “Newton-GMRES preconditioning for discontinuous Galerkin discretizations of the Navier-Stokes equations,” *SIAM J. Sci. Comput.*, vol. 30, no. 6, pp. 2709–2733, 2008.
- [76] V. Maple, “Waterloo maple software,” *University of Waterloo, Version*, vol. 5, 1994.
- [77] E. A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*. Malabar, Florida: Krieger Publishing Company, 1984.
- [78] P. A. Kuchment, *Floquet theory for partial differential equations*, vol. 60. Birkhäuser, 2012.
- [79] P. E. Gill, W. Murray, and M. A. Saunders, “SNOPT: An SQP algorithm for large-scale constrained optimization,” *SIAM journal on optimization*, vol. 12, no. 4, pp. 979–1006, 2002.
- [80] I. H. Tuncer and M. Kaya, “Optimization of flapping airfoils for maximum thrust and propulsive efficiency,” *AIAA journal*, vol. 43, no. 11, pp. 2329–2336, 2005.
- [81] R. Ramamurti and W. Sandberg, “Simulation of flow about flapping airfoils using finite element incompressible flow solver,” *AIAA journal*, vol. 39, no. 2, pp. 253–260, 2001.
- [82] M. F. Platzer, K. D. Jones, J. Young, and J. S. Lai, “Flapping wing aerodynamics: progress and challenges,” *AIAA journal*, vol. 46, no. 9, pp. 2136–2149, 2008.
- [83] A. Oyama, Y. Okabe, K. Shimoyama, and K. Fujii, “Aerodynamic multiobjective design exploration of a flapping airfoil using a navier-stokes solver,” *Journal of Aerospace Computing, Information, and Communication*, vol. 6, no. 3, pp. 256–270, 2009.