Flutter and Post-Flutter Constraints in Aircraft Design Optimization

Eirikur Jonsson\textsuperscript{1}, Cristina Riso\textsuperscript{1}, Christopher A. Lupp\textsuperscript{1}, Carlos E. S. Cesnik\textsuperscript{1}, Joaquim R. R. A. Martins\textsuperscript{1}, and Bogdan I. Epureanu\textsuperscript{2}

\textsuperscript{1}Department of Aerospace Engineering, University of Michigan, Ann Arbor, MI, 48109
\textsuperscript{2}Department of Mechanical Engineering, University of Michigan, Ann Arbor, MI, 48109

Abstract

Flutter is a dynamic aeroelastic instability driven by the interaction of inertial, elastic, and aerodynamic forces. It is an undesirable phenomenon in aircraft because it causes divergent oscillations that may lead to structural damage or failure, performance and ride comfort degradation, or loss of control. If flutter is discovered at the aircraft certification stage, costly redesign is required. Performing flutter analysis early in the design process can mitigate this problem. Furthermore, including flutter analysis as a constraint in multidisciplinary design optimization reduces the risk of costly modifications late in the design cycle. We review the methods for flutter analysis in the context of aircraft design optimization. We also include methods for predicting post-flutter limit cycle oscillations due to the increasing impact of nonlinear effects on future aircraft. While there has been extensive work in flutter and post-flutter analyses, developing design optimization constraints associated with these analyses has additional requirements, such as acceptable computational cost, function smoothness, robustness, and derivative computation. We discuss these requirements and review efforts in the development, implementation, and application of flutter and post-flutter constraints in aircraft design optimization. We conclude the paper by summarizing the current state of this field and the main open problems. Flutter constraints have been included in structural optimizations, but optimizing both the structural sizing and the aerodynamic shape remains a challenge due to the need to recompute the aerodynamic properties at each design iteration. Additional difficulties arise in the presence of large structural deflections and transonic flow conditions due to the dependency of the flutter point on the equilibrium state and the high cost of nonlinear computations. Post-flutter constraints have rarely been included into design optimization, but they
are crucial in the prevention of undesirable limit cycle oscillations. Implementing such constraints requires the development of more efficient and robust prediction methods that can handle realistic configurations. While this paper focuses on flutter and post-flutter constraints for aircraft design optimization applications, the considerations and challenges are broadly applicable to the optimization of engineering systems including stability and post-critical dynamic constraints.

Contents

1 Introduction ................................................................. 3
2 Background on Multidisciplinary Design Optimization ............... 5
3 Background on Flutter and Post-Flutter Modeling .................... 7
   3.1 Flutter Modeling ....................................................... 8
   3.2 Post-Flutter Modeling ............................................... 16
4 Flutter Analysis in Aircraft Design Optimization .................... 18
   4.1 Prediction Methods ................................................... 18
      4.1.1 Eigenvalue Analysis Methods .................................. 19
      4.1.2 Direct Methods .................................................. 26
   4.2 Applications to Optimization ..................................... 28
   4.3 Open Problems ....................................................... 35
5 Post-Flutter Analysis in Aircraft Design Optimization ............. 37
   5.1 Model-Based Prediction Methods ................................... 37
      5.1.1 Time-Domain Methods ......................................... 38
      5.1.2 Frequency-Domain Methods .................................... 41
      5.1.3 Nonlinear Perturbation Methods ............................... 42
      5.1.4 Continuation Methods ......................................... 44
   5.2 Model-Free Prediction Methods ..................................... 45
   5.3 Applications to Optimization ..................................... 48
   5.4 Open Problems ....................................................... 52
6 Conclusions ................................................................. 53
7 Acknowledgments .................................................................. 55
1 Introduction

Flutter is a dynamic aeroelastic instability that causes divergent oscillatory vibrations [1]. It is an undesirable phenomenon in aircraft because it can cause structural damage or failure, performance and ride comfort degradation, or loss of control. Flutter computations are typically performed only after an initial detailed design of the aircraft is completed, because they require the vehicle stiffness, mass, and aerodynamic models to be available [2]. If the design does not satisfy the flutter requirements at this stage, a redesign is necessary, which adds costs and delays to the aircraft development cycle. Thus, it is desirable to consider flutter concurrently with the aircraft design and the wing design in particular. Such a process would not only shorten the design cycle, but also allow for advantageous design trade-offs between the flutter requirements, the other constraints, and the aircraft performance.

Performing multidisciplinary design optimization (MDO) that considers both aero-dynamic shape and structural sizing simultaneously while enforcing flutter constraints is a way to address this issue [3, 4]. Structural optimization alone, even if including aerostructural analyses for enforcing flutter constraints, yields design solutions with suboptimal performance compared to the optimal designs resulting from MDO, where structural and aerodynamic sizing variables are optimized simultaneously [4, 5]. MDO can minimize structural weight, fuel consumption, or a combination of these two objectives with respect to wing shape, internal structure arrangements, and sizing, while accounting for the interactions between aerodynamics, structures, and other disciplines, and satisfying various constraints. MDO with flutter constraints results in designs with optimal aeroelastic tailoring. Omitting flutter constraints in the MDO process when minimizing fuel consumption tends to yield light-weight, high-aspect-ratio wing (HARW) designs that despite being highly efficient may not be feasible [6, 7].

After the aircraft has been designed and a prototype has been built, certification requires flight tests to demonstrate that the aircraft be free from flutter in the flight envelope with a 15% safety margin beyond the dive speed. If flutter is discovered at the flight test certification stage, it requires redesign to address it, incurring additional costs. The redesign effort typically increases the structural weight, reducing the performance originally anticipated for the aircraft.

The trend towards HARW aircraft is driven by better fuel efficiency, but their increased flexibility makes it all the more important to consider flutter accurately and early in the design process [8]. Another recent trend is the increasing use of control surfaces to suppress flutter. Active flutter suppression systems can be incorporated late in the design process when aeroelastic instabilities are encountered and a passive solution such as redesign is impractical and expensive [9, 10]. Alternatively, MDO provides a way to obtain the best possible configuration by co-designing the wing shape and internal structure, which contribute to passive flutter suppression that can then be augmented with an active flutter control system.

While there has been extensive work in methods for flutter analysis, integrating flutter constraints into design optimization requires additional considerations. Models used for flutter prediction should capture the relevant physics with adequate accuracy to correctly drive the optimizer and inevitably there is a compromise between model
fidelity and computational cost. To include flutter analysis in a numerical optimization cycle, speed of execution is particularly important to make sure that the overall optimization process is tractable.

Another important characteristic for integrating flutter analysis into the optimization process is the robustness of the flutter prediction method. Since the optimization process automatically samples the design space, it is likely to request for the analysis of designs that would normally not be chosen by a human designer. Thus, it is important that the flutter analysis converges for such designs so that the overall optimization process is not interrupted.

Gradient-based optimization algorithms are needed to optimize practical aircraft configurations parameterized with a large number of design variables [11]. When using gradient-based algorithms, it is important to consider the smoothness of the objective and constraint functions, as well as the accuracy and efficiency of the derivative computations.

Optimizing HARW configurations subject to flutter constraints is even more challenging because it requires capturing couplings between aeroelasticity and flight dynamics along with geometric nonlinearities that arise in the presence of low natural vibration frequencies and significant structural flexibility [12, 13].

Nonlinearities in the structure (large deflections, free-play of control surfaces, follower loading) or the aerodynamics (shock waves and flow separation) can cause self-sustained oscillations of limited amplitude that remain constant in time, known as limit cycle oscillations (LCOs). For certain types of nonlinearities, LCOs may exist at flight conditions below the flutter point [14]. When nonlinear effects become important, post-flutter analysis should be integrated into the design process in the form of constraints to make sure that the optimal design be feasible. In the context of aircraft design practice, LCO typically refers to oscillations due to control-surface free-play nonlinearities. From a certification perspective, these are typically addressed by computing free-play tolerances. In this paper, we use the term LCO to refer more generally to a post-flutter response due to structural or aerodynamic nonlinearities. While free-play LCOs are a particular case of post-flutter response, the paper focuses on LCOs involving vehicle components or the entire airframe. Taking into account these global LCOs is challenging in design optimization due to the high cost of nonlinear post-flutter computations and the difficulties in formulating post-flutter constraints.

There have been several review papers and textbooks on flutter and post-flutter analysis. Livne [15, 16] reviewed the state-of-the-art and future challenges in aeroelasticity of conventional and unconventional vehicles. A recent review by the same author focused on active flutter suppression control systems [10]. Friedmann [17] reviewed the general challenges in nonlinear aeroelasticity, where the applications focused on rotary wings. Dowell et al. [14] classified nonlinear aeroelastic behaviors and discussed theoretical, computational, and experimental analysis efforts. de C. Henshaw et al. [18] discussed traditional industrial linear flutter prediction and recent efforts for including nonlinear effects, particularly due to transonic flows. More recently, Afonso et al. [8] reviewed nonlinear aeroelasticity of HARWs. Dimitriadis [19] discussed nonlinear post-flutter behaviors in aeroelastic systems and the related analysis methods. Beran et al. [20] reviewed methods for uncertainty quantification in aircraft aeroelasticity and
their application to formulate nondeterministic optimization problems. However, the field lacks a review on the integration of flutter and post-flutter analysis as constraints in aircraft design optimization.

In this paper, we address this shortcoming by reviewing methods for flutter and post-flutter prediction, and we discuss their advantages and disadvantages in the context of aircraft design optimization. First, we provide a brief background on multidisciplinary design optimization (Section 2) and on flutter and post-flutter modeling (Section 3). These sections emphasize the key aspects relevant to aircraft design and are beneficial for readers not familiar with either of these topics. Next, we present numerical methods for flutter and post-flutter analysis, which are summarized in Fig. 1. These methods and examples of their application in aircraft design optimization problems are discussed in Sections 4 and 5. In reviewing previous work we focus on deterministic optimizations, but we also present few examples where constraints are formulated using failure probabilities (reliability-based design optimization) [20]. The paper concludes with remarks on the state of this field and the open challenges to be addressed for integrating flutter and post-flutter considerations into the optimal design of aircraft configurations. While the paper focuses on aircraft design optimization, the methods and challenges reviewed are more broadly applicable to the design optimization of engineering systems subject to stability and post-critical dynamic constraints.

2 Background on Multidisciplinary Design Optimization

Multidisciplinary design optimization couples the relevant disciplines of an engineering system and performing numerical optimization to aid the design of that system [21]. MDO considers several disciplines simultaneously such that their interactions can be leveraged, resulting in a better optimum than if each discipline were optimized sequentially [22]. Thus, considering MDO early in the design process allows engineers not only to improve the design, but also to minimize development time and cost of the overall design.

Performing MDO of aircraft configurations by describing its outer mold line (OML) and structural sizing with high fidelity requires a large number of design variables. Detailed aerodynamic optimization of wings requires hundreds of shape variables [23] and structural sizing based on a detailed finite-element wingbox model that is best utilized with an equally large number of sizing variables [24]. Gradient-based optimization methods are the feasible way to solve for high-dimensional problems within a reasonable computational time, especially when using high-fidelity analyses [25, 11]. Gradient-based methods require derivatives of the objective and constraint functions with respect to the design variables to help the optimization algorithm find the most promising search directions and establish rigorous optimality conditions.

While gradient-free algorithms are typically more robust and some of them explore the design space more widely, their cost is prohibitive when the number of design variables is large. Although gradient-based methods only guarantee convergence to a local optimum, this can be mitigated by using a multi-start technique [26]. Furthermore,
Flutter Analysis Methods (Section 4.1)

Eigenvalue Analysis Methods (Section 4.1.1)
Based on the flutter equation
\[
\left( \frac{U}{T} \right)^2 p^2 M_t + K_e - q_e Q_e(p) \right) q(p) = 0,
\]
with \( p = q + ik \) (non-dimensional Laplace variable) and GAF matrix approximated as transcendental function \( Q_e(ik) \). This leads to:

Iterative or Non-Iterative Methods
Flutter equation solved using iterative or non-iterative methods by assuming \( Q_e(p) \approx Q_e(ik) \).

Variables \( p \) and \( k \) treated as independent and \( \Re(p) = k \) enforced afterwards to select valid roots.

Rational Function Approximations (RFA)
GAF matrix approximated as
\[
Q_e(p) = A_1 + pA_2 + p^2A_2 + D(pI - R)^{-1}E_p
\]
Interpolative matrices obtained by fitting GAF data \( Q_e(ik) \).

Aeroelastic system recast as \( x = Ax \) by introducing additional aerodynamic states. Flutter analysis conducted as a standard eigenvalue analysis.

Direct Methods (Section 4.1.2)
Flutter point of \( x = f(x; U) \) computed directly by solving
\[
f(x_U; U) = 0
\]
\[
\begin{align*}
A_{uu} + \omega^2 u_t &= 0 \\
A_{uu} + \omega^2 u_t &= 0 \\
\varepsilon^2 u_t &= 0 \\
\varepsilon^2 u_t &= 1
\end{align*}
\]
with \( f(x_U; U) = 0 \) equilibrium equation and \( A = \partial f/\partial x \).

Post-flutter Analysis Methods (Section 5.1 and Section 5.2)

Model-Based Methods (Section 5.1)
Based on system representation \( x = f(x; U) \) and consist of:

Time-Domain Methods (Section 5.1.1)

Time-Marching Methods
LCO amplitude obtained by time-marching \( x = f(x; U) \) after perturbing the system equilibrium configuration at a specific flight condition.

Time-Periodic Methods

Cyclic Finite-Difference Methods
Finite-difference scheme applied to one LCO period. For each time interval
\[
x_t = \frac{\Delta t}{2} T \left[ f(x_U; U) + f(x_{t+1}; U) \right]
\]
LCO amplitude computed by solving the algebraic system obtained by assembling equations for all the time intervals, the cyclic boundary condition \( x_0 = x_M \), and a phase fixing condition.

Cyclic Spectral Element Methods
LCO period discretized by writing the solution vector over each time element as
\[
x_t(r) = \sum_{k=0}^{\infty} x^k(r)
\]
with \( x^k_r \) solution values at inner points. LCO amplitude computed as for cyclic finite differences.

Shooting Methods
System \( x = f(x; U) \) integrated over one LCO period by assuming first-guess \( x_0^{(0)} = x(0)^{(0)} \), procedure repeated by updating \( x_0^{(k)} \) to minimize
\[
R(x_0; T)^{(k)} \approx x(t_0; T)^{(k)} - x(t_0)^{(k)}
\]
Phase fixing condition also required.

Harmonic-Balance Methods (Section 5.1.2)
System \( x = f(x; U) \) assumed to have cyclic solution
\[
x(t) = x_0 + \sum_{k=1}^{M} x_k \cos(k\omega t) + x_k \sin(k\omega t)
\]
LCO amplitude computed by substituting assumed solution into \( x = f(x; U) \) and by setting total coefficients multiplying sine and cosine functions to zero. Phase fixing condition also required.

Nonlinear Perturbation Methods (Section 5.1.3)
Reduced-order representation of \( x = f(x; U) \) obtained around the flutter point and used to approximate LCO amplitude as a function of \( U \) (bifurcation diagram). Methods valid for small amplitudes and consist of the method of multiple scales, the center-manifold reduction, and the method of normal forms.

Numerical Continuation Methods (Section 5.1.4)
Given a nonlinear algebraic system
\[
g(q; U) = 0
\]
and known solution \((q_0; U_0)\) along with local tangent vector \( \tau_0 = [q^T; U_0^T] \), solution branch traced by computing new solution \( q \) at incremented \( U \). Methods consist of predictor-corrector and piecewise-linear methods. Typically applied to continue algebraic systems obtained by applying time-periodic or frequency-domain methods for one value of \( U \).

Model-Free Methods (Section 5.2)
Bifurcation diagram computed directly based on characteristics of system dynamic outputs (time responses). Possible approach is to write the time variation in the response envelope slope (recovery rate) as
\[
\lambda(t; U) = \frac{d\ln r}{dt} \approx p(t) + a_1(t)(U - U_F) + a_2(t)(U - U_F)^2
\]
Polynomial functions \( p(t) \), \( a_1(t) \) and \( a_2(t) \) obtained by fitting recovery rate values from pre-flutter time simulations in the \( \lambda - U \) plane for different \( t \). This gives values \( U \) that identify points \((\bar{t}, \tilde{U})\) on the bifurcation diagram.

Figure 1: Summary of flutter and post-flutter analysis methods.
recent studies failed to find multiple local minima (multimodality) in airfoil and wing shape design optimization [23, 11], except the case of planform optimizations where expected local minima were found related to choices such as upwards or downwards winglets [27].

The efficacy of gradient-based algorithms relies on accurate and efficient gradient computations. The gradient accuracy directly affects the ability to converge to the optimum with a specified tolerance and the order of convergence of the optimization. In the best case, inaccurate gradients increase the number of iterations required for convergence and in the worst case cause early stopping due to numerical issues. Efficiency gradient computation is also important because this computation is sometimes the bottleneck in the optimization cycle.

When it comes to methods for computing gradients, the finite-difference method is a popular choice because it is easy to implement and can always be used, even with black-box codes. The major drawbacks of the finite-difference method is that it is inaccurate and its computational cost scales poorly with the number of design variables. Unlike the finite-difference method, the complex-step method [28] is accurate, but its cost still scales unfavorably with the number of design variables, making it prohibitive for wing design applications. Furthermore, the complex-step method cannot be directly applied to programs that use complex numbers. Such programs need to be made complex-step compatible by modifying them so that the real and imaginary part of the complex number are represented as two real numbers. Another option for computing gradients is automatic differentiation (AD), which uses a software tool to parse the code of an analysis to produce a new code that computes derivatives of that analysis [29, 30]. Although AD can scale well with the number of variables, it does not handle iterative simulations efficiently in general. Finally, analytic methods are the most desirable because they are both accurate and efficient, especially for iterative simulations [31]. However, they require significant implementation effort. There are two main approaches within the analytic methods: the direct approach and the adjoint approach. The adjoint approach is especially attractive because the computational cost is dependent of the number of outputs of interest (objectives and constraints) but independent of the number of design variables [32, 31]. A coupled-adjoint approach can solve static aeroelastic problems [5, 6] and can be generalized to any multidisciplinary problem [33, 34].

In the context of aircraft design optimization subject to flutter or post-flutter constraints, most of the early efforts used gradient-based optimization with gradients computed with either finite differences or the direct analytic method. However, recent efforts implemented the more efficient adjoint approach, and some also used AD techniques. These applications are further discussed in Section 4.2 for flutter constraints and Section 5.3 for post-flutter constraints.

3 Background on Flutter and Post-Flutter Modeling

For aircraft designs to be useful and practical, the underlying models used in the flutter and post-flutter analysis need to capture the correct physics involved. However,
a simplification of the phenomena is often necessary to make problems tractable to solve. Therefore, the choice of model should balance the fidelity needed to obtain accurate predictions and the mathematical or computational tractability for design applications.

This section highlights the modeling aspects to be considered in flutter and post-flutter analysis, which are discussed in more detail in Sections 4 and 5 for obtaining meaningful results in a design optimization. By flutter, we mean the onset of divergent oscillations as the flight conditions of aircraft cross the critical stability boundary (flutter boundary).

Mathematically, flutter occurs at a Hopf bifurcation point [35] beyond which the system is in the post-flutter regime. Several post-flutter behaviors are possible, as discussed in detail by Dimitriadis [19] for a two-dimensional aeroelastic system with stiffness and damping nonlinearities. Among these behaviors, we are particularly interested in self-sustained oscillations with limited amplitude that remains constant in time, known as LCOs.

LCOs typically develop beyond the flutter boundary; however, for certain types of nonlinearities, they can also occur before reaching the flutter boundary [12]. Integrating post-flutter analyses into the design process can prevent this undesirable situation. This places additional challenges due to the difficulties in modeling nonlinear aeroelastic systems accurately and the inherent high cost of nonlinear aeroelastic computations.

### 3.1 Flutter Modeling

Flutter is defined as a self-exciting dynamic instability that is associated with the interaction of inertial, elastic, and aerodynamic forces [1]. At the onset of flutter, this aeroelastic instability can be physically described as an oscillation with a small amplitude that is constant in time triggered by a small-amplitude disturbance, as shown in Fig. 2. The flight condition in which the system damping vanishes, resulting in this self-sustained oscillation, represents the **flutter point** (or flutter boundary). For linear systems, the flutter point is defined as the minimum dynamic pressure at which at least one of the modes becomes unstable [1]. The dynamic pressure can be replaced by equivalent airspeed, and is a function of altitude and Mach number.

Past the flutter point, in the absence of restraining nonlinearities from the aerodynamics, the structure, or both, the amplitude of the oscillations grow exponentially. Fluid-structure interactions may also result in a static instability called divergence [1], which is not associated with oscillations. As for flutter, the structural response grows unbounded past the onset point, eventually reaching a limited-amplitude oscillation if restraining nonlinearities are present.

In the following discussion, we focus mainly on flutter phenomena, because for many practical configurations flutter occurs before divergence. However, accounting for divergence and the associated post-critical response in the design process shares many of the modeling and analysis aspects associated with flutter. Furthermore, some of the analysis methods and constraints discussed in Section 4 are applicable to divergence as well as flutter. Moreover, in this paper we focus on global wing or component flutter rather than localized effects such as panel flutter that typically occurs at supersonic
Figure 2: Aeroelastic system response before and past the flutter point. Prior to reaching the flutter point, the aeroelastic response is damped. At the flutter point, the system response is an oscillation with a small constant amplitude. Past the flutter point, a linear system response diverges, while a system with structural or aerodynamic nonlinearities develops a stable response with finite amplitude that remains constant in time, known as LCO.
Figure 3: $V - g$ diagrams for different types of flutter. Soft flutter is a gradual loss of damping with increased speed while hard flutter occurs abruptly and violently. A hump mode manifests itself as a damping decrease followed by an increase, which may result in considerably lower flutter speed.

flow conditions.

The possible flutter characteristics are illustrated in Fig. 3, which shows the variation of the modal damping with flight speed at a fixed altitude. This is known as $V - g$ diagram, which is a classical tool used in linear flutter analyses for determining the flutter point and interpreting the flutter characteristics. A similar representation can be obtained by varying dynamic pressure at a fixed Mach number.

Damping changes with flight speed in different ways among different designs, leading to different flutter behaviors. Soft flutter occurs when the damping decreases gradually with increasing flight speed, while hard flutter occurs when this decrease is abrupt. Another possibility is that there is a gradual decrease in damping with increasing flight speed, all the way to cross the zero value and beyond, followed by a damping increase. This phenomenon is known as a hump mode. These concepts are important when considering how to formulate a smooth and continuous flutter constraint for gradient-based optimization and are discussed in Section 4.

In flutter analysis, the physics described above is often represented by less expensive linear models due to the number of conditions to be considered for certification. However, nonlinear structural and aerodynamic effects or the interaction between elastic and rigid-body degrees of freedom (DOF), which become important in the presence of low structural vibration frequencies, may significantly impact the flutter point. Therefore, the flutter prediction accuracy depends on the appropriate modeling of nonlinear effects and boundary conditions.

Furthermore, nonlinear effects impact not only the models used in flutter analysis, but also the analysis methods themselves. For linear systems, flutter characteristics do not depend on the deformation state. Therefore, flutter is typically analyzed by considering the unloaded and undeformed structure. For nonlinear systems, stability characteristics vary with the deformation configuration. Therefore, flutter analysis
must be performed by computing the eigenvalues of the aeroelastic system linearized around equilibrium points for each flight condition to identify at what point the damping vanishes [12] as shown in Fig. 4. The eigenvalues can be computed by considering both the elastic and rigid-body DOFs (flutter in free flight) or by retaining only the elastic DOFs (traditional flutter) or the rigid-body DOFs (flight dynamic stability).

Many possible sources of nonlinearities can be present simultaneously in aeroelastic systems [14, 8, 19]. In this paper, we focus on aerodynamic nonlinearities due to transonic flow regimes and geometric structural nonlinearities due to large deflections, both of which are critical in the design of next-generation transport aircraft.

Aerodynamic nonlinearities due to shock waves and flow separation significantly impact the flutter speed. This decreases dramatically in the transonic regime, a phenomenon known as the *transonic dip* [36, 37, 38, 39, 40] illustrated in Fig. 5. Low-order, linear unsteady aerodynamic models commonly used in flutter analysis are in general accurate enough for subsonic and supersonic flows, but they severely overestimate the flutter speed for transonic conditions [41, 42, 43].

As shown in Fig. 5 for a hypothetical wing, linear theory is non-conservative when compared to nonlinear viscous models. Nonlinear inviscid models based on Euler equations can capture shock waves but they still fail to accurately predict the flutter boundary [44]. In many cases, the nonlinear inviscid theory predicts a highly conservative flutter speed at the dip, even though it is generally closer to viscous theory predictions. Depending on the severity of shocks, models based on Navier–Stokes equations (which include viscous and turbulence effects like boundary layer thickening, flow separation, and interactions between shocks and regions of separated flow) are necessary to obtain accurate flutter points [42]. Studies on various geometries demonstrated that taking into account viscous phenomena in the transonic flow regime improves the numerical prediction of transonic dip [45, 46, 47, 48].

A common approach to improve the accuracy of transonic flutter computations while minimizing the increase in computational cost is to use numerical or experimental corrections applied to potential-flow linear models [49]. However, the correction data may not be available for optimization, either because it requires high-fidelity computations that are too expensive or because it is obtained from wind-tunnel measurements. This problem can be addressed by analyzing flutter using time-accurate dynamic simulations and higher-fidelity aerodynamic models. On the other hand, flutter prediction based on time-accurate computational fluid dynamics (CFD) is a challenge even for just analyzing the final configuration and is currently prohibitive for design space exploration.

Methods exist that try to preserve the computational efficiency of lower-fidelity methods while retaining the nonlinear physics modeled by the higher-fidelity CFD methods. One possibility is to use time-linearized transonic small disturbance (TSD) equations. Linear small-disturbance theory is inadequate for capturing strong transonic shocks, but small-disturbance solutions about the steady nonlinear background flow computed using high-fidelity CFD can provide acceptable performance and accuracy [50]. Another possibility is to use the transonic equivalent strip (TES) theory [51] and a provided pressure distribution from either experimental data or a high-fidelity CFD code to compute the small-disturbance transonic aerodynamic loads for flutter.
Figure 4: Nonlinear flutter analysis process. The nonlinear equations of motion are linearized around the equilibrium solutions for each flight condition and the eigenvalues are computed by considering both the elastic and rigid-body DOFs (flutter in free flight), only the elastic DOFs (traditional flutter), or only the rigid-body DOFs (flight dynamic stability).
Figure 5: Characteristic transonic dip of a transport wing. Aerodynamic nonlinearities due to shock waves and flow separation have a significant impact on the flutter speed, which may decrease dramatically. Linear theory (e.g., DLM) is non-conservative when compared to nonlinear viscous theory (e.g., RANS). Nonlinear inviscid theory (e.g., Euler) predicts highly conservative flutter speed at the dip, but it is generally closer to viscous theory predictions.

analysis [52]. Furthermore, several efforts have applied the time-linearization directly to CFD solvers. This approach is appealing since they retain nonlinear effects, provide accurate results in the transonic regime, and intrinsically account for geometric properties of the body such as thickness and camber.

Kreiselmaier and Laschka [53] developed an unsteady method based on the small-disturbance Euler equations, which was later extended to small disturbance Navier–Stokes equations to include viscous effects [54]. The proposed method produced good results in the transonic flow regime [55, 56]. Thormann and Widhalm [57] developed a linear frequency domain (LFD) solver taking advantage of preconditioned Krylov GMRES [58] solution method. Later, Widhalm and Thormann [59] improved the algorithm and provided the analytic derivatives needed in the solution, improving the solver efficiency. The method was shown to be accurate when compared to full unsteady time-marching solutions.

These approaches consider unsteady aerodynamic models linearized about nonlinear equilibrium states and thus can capture the impact of static nonlinear effects on flutter. Moreover, they demonstrate computational savings well beyond an order of magnitude compared to fully unsteady time-marching solutions [53, 57]. However, computing derivatives of such methods for optimization is challenging due to the need for second-order derivative information.

Motivated by the interest in capturing key transonic flow physics with low computational cost, recent efforts also developed low-order unsteady transonic aerodynamic models suitable for integrating transonic flutter analyses into aircraft design.

Skujins and Cesnik [60] proposed a reduced-order unsteady aerodynamic model for
multiple Mach regimes based on linear convolution with a nonlinear static correction. The methodology included error estimation capabilities based on the newly developed method of segments, which represents a flexible wing as a collection of rigid spanwise segments subject to local angle of attack and Mach number conditions. The method of segments was also applied to transonic flutter analysis of a transport vehicle by Kitson and Cesnik [61].

Mallik et al. [62] developed a reduced-order model for HARW configurations by combining time-linearized Leishman indicial functions [63] with lift-curve and moment-curve slopes obtained by solving the RANS equations around airfoils at various Mach number and angle of attack conditions and for various thickness ratios. They obtained a state-space formulation for the airfoil unsteady aerodynamics to be used for eigenvalue analysis, which was extended to three-dimensional HARW discretized in spanwise strips by accounting for sweep correction. Flutter results were compared with wind-tunnel experiments for a truss-braced wing (TBW) configuration. The low-order model captured the transonic dip that was not predicted by potential-flow theory and presented good agreement with experiments at significantly lower computational cost compared to unsteady RANS simulations. These results showed the method suitability for conceptual HARW aircraft design including transonic flutter constraints.

Opgenoord et al. [64] developed a physics-based two-dimensional low-order model for transonic airfoils using the perturbations of the lowest-order volume-source and vorticity moments with respect to a known nonlinear background flow solution as the states. Evolution equations for these perturbations were derived and calibrated using data from high-fidelity Euler CFD simulations. A state-space unsteady aerodynamic model was obtained for airfoil flutter analysis which was later extended to three-dimensional HARW configurations [65] using strip theory and sweep correction, as done by Mallik et al. [62]. The method was applied in conceptual design and optimization problems including transonic flutter considerations [65, 66].

In addition to capturing transonic effects, a more recent flutter modeling need is to take into account geometric structural nonlinearities. These are particularly important in the analysis of HARW configurations, which achieve higher energy efficiency at the cost of increased structural flexibility and thus experience large deflections under normal operating load conditions. The changes in geometric shape and stiffness properties due to these deflections significantly affect the flutter boundary [13]. When deformations are large, traditional flutter analysis based on the vehicle undeformed shape does not capture the actual behavior of the aircraft during flight.

Studies on isolated HARWs [67], high-altitude long-endurance (HALE) flying-wing configurations [68, 12, 69], and commercial transport vehicles [61] pointed out the need to analyze very flexible aircraft in statically deformed configuration at trim, which varies with the flight condition. Including structural nonlinearities in flutter prediction is challenging for both analysis and design due to the high computational cost of nonlinear aeroelastic simulations and the flutter boundary dependency on the deformation state, which is not considered in linear approaches.

Finally, classical wing flutter analyses typically assume the vehicle to be clamped at the wing root. While this may be an acceptable simplification for some vehicles, it does not reflect the vehicle behavior in free flight [15, 16]. For some configurations,
simply including rigid-body DOFs influences the flutter solution substantially [12].
This occurs due to the coupling between rigid-body motion and structural dynamics
that arise in the presence of low natural vibration frequencies. These interactions
usually result in lower flutter points than the cantilevered configurations or different
flutter mechanisms like body-freedom flutter (BFF) [70]. Therefore, it is imperative to
understand the effect of boundary conditions and state variables on flutter prediction.

Mazidi et al. [71] investigated the effect of engine placement and roll maneuver
on flutter results. They observed that the roll maneuver has a destabilizing effect on
the flutter boundary dependent on the bank angle. Additionally, the location of the
engine or external store greatly affects the flutter boundary and the roll-induced effects.
Nearly all vehicles perform roll maneuvers during turns, making the inclusion of these
conditions relevant to the aircraft design process.

There has been further work on the effect of rigid-body DOFs on the flutter prob-
lem. Niblett [70] investigated the causes of BFF for a conventional wing-fuselage con-
figuration using a linear analytical flutter method. Su and Cesnik [12] investigated
the flutter behavior of a blended wing-body (BWB) aircraft both for cantilevered and
free-flying conditions using the University of Michigan’s Nonlinear Aeroelastic Sim-
ulation Toolbox (UM/NAST). They observed a reduction in the flutter speed when
the rigid body DOFs were included compared to the cantilever case. Moreover, the
flutter mode changed to include pitch and plunge motions, resulting in BFF. Similarly,
Jones and Cesnik [72] investigated the BFF characteristics of the X-56A experimental
aircraft, describing the entire modeling process used for the flutter prediction. Cesnik
and Su [73] analyzed the University of Michigan’s X-HALE very flexible aeroelastic
testbed [74] and observed that significant wing deformations can drive lateral BFF
due to the coupling of the Dutch roll and asymmetric wing bending modes. Su and
Cesnik [69] investigated the stability and dynamic response of a highly flexible flying
wing for different payload configurations and gust disturbances. They found that wing
deflections can lead to an unstable phugoid mode and an aperiodic short-period mode.
Similar behaviors were observed by Patil and Hodges [68] and Patil and Taylor [75].
Richards et al. [76] also analyzed the coupled flight dynamics and aeroelasticity of
flying wings. They noted that BFF occurred due to a coupling of the short period
pitching mode and the first elastic bending mode. They compared different inertial
configurations of the aircraft and noted that BFF depends largely on the inertia about
the pitch axis. They found a boundary value for pitch inertia that uncoupled the pitch
and bending modes, thus replacing BFF with a more conventional flutter. After pa-
rameter studies, they concluded that BFF was caused by low fuselage inertia, which
could be mitigated by redistributing the fuselage mass.

BFF is not exclusive to flying-wing configurations; conventional tube-and-wing air-
craft can also encounter this type of instability. Cavallaro et al. [77] investigated the
flutter behavior of a Prandtl Plane (boxed wing aircraft) including rigid-body DOFs
using MSC Nastran. Similarly to Richards et al. [76], they observed a dependence on
the fuselage mass in causing BFF: while the baseline configuration encountered flutter
without rigid-body contributions, increasing the fuselage mass resulted in BFF. There-
fore, capturing these phenomena in flutter analyses is advisable even for conventional
configurations and imperative when investigating non-conventional ones [15, 16].
Figure 6: Bifurcation diagrams for possible post-flutter scenarios. For supercritical bifurcation (a), LCOs occur only beyond the flutter point. For subcritical bifurcation (b), LCOs can occur even before the flutter point if disturbances are sufficiently large. In presence of higher-order stabilizing nonlinearities, both a stable (solid line) and an unstable (dashed line) solution branch exist for a range of flight conditions.

3.2 Post-Flutter Modeling

When aircraft flutter, disturbances perturbing an equilibrium state cause an oscillatory response with divergent amplitude. This may ultimately result in structural failure or, if the amplitude growth is gradual enough, the aircraft can be brought back at a stable operating point by pilot or control action. Dynamic nonlinearities in the aerodynamics or the structure, however, usually play a stabilizing role and restrain divergent motions, leading to a self-sustained oscillatory response of limited amplitude that remains constant in time, known as LCO. Unfortunately, nonlinear effects tend to be destabilizing and cause LCOs even before reaching the flutter point [14]. Furthermore, LCOs may or may not develop for a given flight condition depending on the amplitude of disturbances (initial conditions) and previous motion history (hysteresis) [78]. Therefore, when nonlinear effects become significant, post-flutter phenomena need to be considered in addition to the flutter boundary in the design process.

As mentioned in Section 1, in aircraft aeroelastic qualification, the term LCO is typically associated with control-surface free-play. In this paper, LCO has the more general meaning of post-flutter response leading to the development of self-sustained oscillations in vehicle components or in the entire structure. When nonlinear behavior is anticipated in aircraft configurations, optimization should include constraints on the post-flutter response, in addition to flutter constraints, to avoid the occurrence of undesirable LCOs at the operating conditions.

Post-flutter behaviors are typically classified using bifurcation diagrams. These diagrams represent the LCO amplitude variation with a chosen control parameter, called the bifurcation parameter. For aerelastic systems, this is usually the flight speed or the dynamic pressure [79]. Two qualitative post-flutter scenarios are possible, as illustrated in Fig. 6.

Figure 6a shows the bifurcation diagram for a system with supercritical post-flutter behavior. When the system is perturbed before the flutter point, the response al-
ways converges towards the initial equilibrium state regardless of the magnitude of the disturbance. When the perturbation is applied beyond flutter, the response grows unbounded until dynamic stabilizing nonlinearities lead to an LCO with limited amplitude that keeps constant in time. Each point on the bifurcation diagram gives the LCO amplitude as function of the flight parameter. In the case of supercritical behavior, the LCO amplitude increases smoothly with the flight parameter and with a decreasing slope as stabilizing nonlinearities become stronger [14].

The bifurcation diagram of a system with subcritical post-flutter behavior is illustrated in Fig. 6b. In this case, an unstable bifurcation branch (dashed line) exists below the flutter point. As a result, when disturbances are sufficiently large, the initial equilibrium configuration is lost. If higher-order stabilizing nonlinearities are present, both an unstable and a stable (solid line) solution branches exist for a range of flight conditions. In this case, the response grows until restraining nonlinearities lead to an LCO with an amplitude given by the upper stable bifurcation branch. Supercritical bifurcation is characterized by discontinuous behavior with respect to the amplitude of initial disturbance due to destabilizing nonlinear effects.

Supercritical LCOs are undesirable in aircraft, since they cause structural fatigue or failure, performance and ride comfort degradation, or loss of control. However, if the flutter boundary is satisfied with the required safety margins, LCOs never develop since the aircraft always operates below the flutter speed by certification. Additionally, enforcing flutter constraints during design can guarantee that flutter does not occur and therefore, supercritical LCOs are not encountered.

However, subcritical post-flutter behavior is a more serious problem, because instabilities may occur at flight conditions that are expected to be stable according to a conventional flutter analysis. Depending on the type of nonlinearities, these flight conditions could lie within the flight envelope. This considerably reduces the effectiveness of flutter constraints enforced during design.

When higher-order stabilizing nonlinearities are present, as shown in Fig. 6b, the system suddenly responds with a large-amplitude LCO when perturbations before the flutter point pass the amplitude limit given by the unstable solution branch. Similarly, the system suddenly responds with a large amplitude LCO when it is perturbed beyond the flutter point. Therefore, subcritical bifurcations are characterized by discontinuous behavior with respect to both the initial conditions and the flight parameters.

In contrast, LCO amplitude grows gradually and with much smaller amplitudes past the flutter point in supercritical systems (see Fig. 6a). Furthermore, subcritical post-flutter response is characterized by hysteresis. Figure 6b shows that once a stable LCO develops, the system does not recover the initial equilibrium state when the flight parameter is reduced below the critical (flutter) value.

To eliminate the LCO, it is necessary to reduce the flight parameter below the value at the folding point connecting the unstable and stable solution branches. Depending on the type of nonlinearities present in the system, this may be significantly lower than the flutter value. The above considerations make clear that when nonlinear effects become significant, design optimization should include post-flutter constraints for obtaining feasible optimal designs that do not experience subcritical LCOs.

The post-flutter behavior of wings with structural and aerodynamic dynamic non-
linearities was analyzed by various researchers [78, 80, 81, 82, 12]. This topic was also included in review papers [18, 14, 8]. However, only a few efforts integrated post-flutter analyses in design optimization as a constraint, as discussed in Section 5.4.

4 Flutter Analysis in Aircraft Design Optimization

Because flutter is a safety-critical phenomenon, analyses and experimental investigations are required for vehicle certification. Analyzing a configuration for flutter late in the development cycle is likely to result in an inefficient design solution or pose challenges mitigating unexpected instabilities, resulting in performance decrease, financial losses, or both. For this reason, flutter should be integrated into the design process early in the form of a constraint.

In this section, we review flutter prediction methods and previous research that addressed optimization subject to flutter constraints. Past work has focused primarily on flutter analyses using linear structures and linear aerodynamic models. Further work considered transport aircraft operating in the transonic regime and thus used linear structures and nonlinear aerodynamics. The recent trend towards more flexible aircraft has led to research on constraining flutter for geometrically nonlinear structures.

Due to the sheer number of design variables typically used in practical aircraft optimization, most of the previous work reviewed in this paper used gradient-based methods. However, studies on geometrically nonlinear configurations were limited to simple structures parametrized by few design variables. For this reason, efforts that considered gradient-free optimizations are also included.

4.1 Prediction Methods

Flutter computations are typically performed in the frequency domain by solving an eigenvalue problem. Well-established eigenvalue-analysis methods exist for flutter analysis of linear aeroelastic systems (e.g., k-, p-, pk- and g-method) [83, 84, 85, 86, 87]. These methods are also applicable to nonlinear systems by linearizing the equations of motion about the nonlinear equilibrium configuration at each flight condition for capturing static nonlinearities due to the structure (large deflections) or the aerodynamics (background transonic flow). Direct methods, based on the Hopf-bifurcation theory, can also be used to predict the flutter point of nonlinear aeroelastic systems directly in the frequency domain [18].

It is also possible to predict flutter in the time domain, but this incurs a much higher computational cost. When analyzing the stability of a system in the time domain, the flutter point is typically evaluated by perturbing equilibrium configurations at different flight conditions and by time-marching the equations of motion in order to verify the decay or growing of the response, or to extract damping values in a post-processing stage [88, 89, 90, 91, 92, 93, 94, 95].

The computational cost of time-domain flutter analysis based on transient simulations is currently prohibitive for optimization, particularly in the presence of aerodynamic or structural nonlinearities. The high computational cost is due to the large number of computations required to evaluate the flutter speed by means of flight speed
(or dynamic pressure) sweep or bisection. Additionally, ascertaining the stability close to the flutter point requires long integration times due to small damping values.

For gradient-based optimization, another challenge is the efficient computation of derivatives for time-marched systems. Adjoint methods are advantageous for optimizations with many design variables, but they are computationally expensive when applied to time marching solvers [96]. This is because the adjoint solution requires a reverse time integration following the forward integration for computing the system response [97], which results in large computational times and storage memory requirements [98]. Due to the above limitations, we focus on frequency-domain flutter prediction methods, but some fully coupled unsteady aeroelastic adjoint implementations do exist in literature [99, 100, 101, 102, 103, 104]. Eigenvalue-based methods are discussed in Section 4.1.1 and direct methods in Section 4.1.2.

4.1.1 Eigenvalue Analysis Methods

The discrete equations of motion for a generic linear aeroelastic system can be written as

\[
M\ddot{u}(t) + Ku(t) - F_{aero}(t) = 0,
\]

where \( N_s \) is the number of structural DOFs (nodal displacements and rotations), \( M, K \in \mathbb{R}^{N_s \times N_s} \) are the mass and stiffness matrices resulting from the finite-element discretization, \( u \in \mathbb{R}^{N_s} \) is the nodal displacement vector, and \( F_{aero} \in \mathbb{R}^{N_s} \) is the nodal aerodynamic load vector. With no loss in generality, we omit the structural damping in Eq. (1) along with other external loads acting on the structure. For the sake of conciseness, Eq. (1) omits the dependency of the structural matrices and aerodynamic loads on the design variables \( d \in \mathbb{R}^{N_d} \). Therefore, flutter and post-flutter analysis methods are presented for a fixed design and all the structural and aerodynamic quantities are updated at each optimization step based on the current values of the design variables.

Due to the sheer number of structural DOFs, it is common to reduce the computational effort by rewriting Eq. (1) in terms of a reduced set of \( N_r \) generalized coordinates (modal amplitudes) where \( N_r \ll N_s \). The displacement field can be then approximated by

\[
u(x, y, z; t) \approx \Phi_r(x, y, z)q(t),
\]

where \( q \in \mathbb{R}^{N_r} \) is the vector of retained generalized coordinates and \( \Phi_r \in \mathbb{R}^{N_s \times N_r} \) is a matrix whose columns contain the corresponding eigenvectors (mode shapes). Substituting Eq. (2) into Eq. (1) and pre-multiplying by \( \Phi_r^T \) yields

\[
M_r\ddot{q}(t) + K_rq(t) - F_{aero,r}(t) = 0,
\]

where \( F_{aero,r} = \Phi_r^T F_{aero} \in \mathbb{R}^{N_r} \) is the vector of generalized aerodynamic forces and, assuming that the eigenvectors are \( M \)-orthonormal, the generalized mass and stiffness matrices take the form

\[
M_r = \Phi_r^T M \Phi_r = I_r \in \mathbb{R}^{N_r \times N_r},
K_r = \Phi_r^T K \Phi_r = \text{diag}\{\omega_i^2\} \in \mathbb{R}^{N_r \times N_r},
\]

where \( \omega_i \) is the \( i \)-th natural angular frequency (\( i = 1, \ldots, N_r \)).
Assuming linear aerodynamics, the generalized aerodynamic forces are written in the Laplace domain as

\[ \hat{F}_{\text{aero},r}(s) = q_{\infty} \mathbf{Q}_r(s) \hat{q}(s) \] (5)

where \( s \) is the Laplace variable, \( q_{\infty} = \rho U^2 / 2 \) is the flight dynamic pressure corresponding to the density \( \rho \) and flight speed \( U \), \( \mathbf{Q}_r(s) \) is the generalized aerodynamic force (GAF) matrix, and \( \hat{q}(s) \) is the Laplace transform of \( q(t) \). Taking the Laplace transform of Eq. (3) yields

\[ [s^2 \mathbf{M}_r + \mathbf{K}_r - q_{\infty} \mathbf{Q}_r(s)] \hat{q}(s) = 0 , \] (6)

or

\[ \left[ \left( \frac{U}{b} \right)^2 p^2 \mathbf{M}_r + \mathbf{K}_r - q_{\infty} \mathbf{Q}_r(p) \right] \hat{q}(p) = 0 , \] (7)

where \( p = sb/U = g + ik \) is the nondimensional Laplace variable, \( b \) is the reference half chord, \( g \) is the nondimensional damping, and \( k \) is the reduced frequency [83].

Equation (7) requires computing the GAF in the Laplace domain. However, the GAF is typically given as a transcendental function, \( \mathbf{Q}_r(ik) \), of the reduced frequency. Two approaches are used for overcoming this problem [10]. One approach is to approximate \( \mathbf{Q}_r(p) \approx \mathbf{Q}_r(ik) \) and solve the flutter equation, Eq. (6), or an equivalent form by computing the GAF matrix in the reduced frequency domain while enforcing the condition \( \Im(p) = k \). An alternate approach is to obtain a rational function approximation (RFA) of \( \mathbf{Q}_r(ik) \) and use analytic continuation [105] to extend its domain from the imaginary axis (reduced frequency) to the entire complex plane (nondimensional Laplace variable). The aeroelastic system can be recast in state-space form by introducing additional aerodynamic states, such that flutter can be analyzed using a standard eigenvalue analysis.

When the GAF is represented as a transcendental function of \( k \), flutter analysis is performed using iterative or non-iterative methods that either compute the true damping only at the flutter point (\( k \)-method) or at all flight speeds or dynamic pressure values (root locus, \( p \)-, \( pk \)-, and \( g \)-methods) [106, 83, 84, 85, 87]. As an example, using the \( pk \)-method, assuming \( \mathbf{Q}_r(p) \approx \mathbf{Q}_r(ik) \), Eq. (7) is rewritten as [83]

\[ \left[ \left( \frac{U}{b} \right)^2 p^2 \mathbf{M}_r + \mathbf{K}_r - q_{\infty} \mathbf{Q}_r(ik) \right] \hat{q}(p) = 0 . \] (8)

This is a second-order nonlinear eigenvalue problem, where the nonlinearity stems from the dependency of \( \mathbf{Q}_r \) on the imaginary part of \( p \).

Several decompositions of the GAF \( \mathbf{Q}_r \) in Eq. (8) exist in the literature. Stanford [107] summarized these decompositions and showed that they predict the same flutter speed, but that the mode migration and characteristics may be very different. This directly impacts the optimization, resulting in different optimal designs.

As an example, a possible approach is to split the GAF into its real and imaginary parts as \( \mathbf{Q}_r = \mathbf{Q}_r^R + i\mathbf{Q}_r^I \). Assuming small damping, which gives \( p/(ik) \approx 1 \) [108],
Eq. (8) can be rewritten in first-order form as the generalized nonlinear eigenvalue problem [84, 109, 110]

\[
p \begin{bmatrix} I_r & 0 \\ 0 & \left( \frac{v}{k} \right)^2 M_r \end{bmatrix} \left\{ \hat{q} \right\} - \left[ -(K_r - q_\infty Q^R_r) \frac{q_\infty}{k} Q^l_r \right] \left\{ \hat{p} \hat{q} \right\} = 0 ,
\]

where \( I_r \in \mathbb{R}^{N_r \times N_r} \) is an identity matrix. For any given value of \( k \), Eq. (8) (or Eq. (9) with the above specific form of \( Q_r \)) can be solved using standard eigenvalue routines such as LAPACK [111] to compute the eigenvalues and eigenvectors. However, valid roots need to satisfy the equivalence \( \Im(p) = k \). One approach to this problem is to use an iterative solution algorithm as suggested by Hassig [83], similar to what is implemented in MSC Nastran [108]. As a disadvantage, iterative methods may converge to incorrect roots or not converge at all [112, 113].

In gradient-based optimization, robustness of analysis methods, continuity, and smoothness are of utmost importance. For this reason, noniterative methods [87, 113] that mitigate the aforementioned issues were more recently proposed and applied in optimization settings [109, 110, 114].

On the other hand, the GAF can be approximated by numerically fitting data in the reduced frequency domain using a RFA. This includes polynomial terms up to the quadratic term for representing aerodynamic stiffness, damping, and mass effects along with a rational term representing unsteady aerodynamic lag effects. Once a RFA is computed, its validity can be extended from the reduced frequency domain (imaginary axis of the complex plane) to the nondimensional Laplace domain (entire complex plane) using analytic continuation [105].

Several RFA approaches exist in the literature. Roger [115] proposed the following widely used form (given in nondimensional Laplace domain below),

\[
Q(p) = A_0 + p A_1 + p^2 A_2 + \sum_{l=3}^{N_l+3} \frac{p A_l}{p + \beta_l - 2} ,
\]

where \( N_l \) is the number of lag terms (aerodynamic roots) \( \beta_{l-2} \), which are specified by the user based on the reduced frequency range of interest. A common choice is to assume \( N_l = 3 \) [116]. The coefficient matrices \( A_0, \ldots, A_n \in \mathbb{R}^{N_r \times N_r} \) are determined using a least-square fitting of the GAF as a function of \( k \) [115], whose accuracy depends on the number of aerodynamic roots, as well as their values.

Karpel [116] reviewed early RFA methods and proposed a new technique based on the minimum number of aerodynamic states \( (N_m) \), corresponding to the approximation

\[
Q(p) = A_0 + p A_1 + p^2 A_2 + D [p I - R]^{-1} E p ,
\]

which includes Eq. (10) as a special case [117]. The interpolating matrices \( A_0, A_1, A_2 \in \mathbb{R}^{N_r \times N_r} \), \( D \in \mathbb{R}^{N_r \times N_m} \), \( R \in \mathbb{R}^{N_m \times N_m} \), and \( E \in \mathbb{R}^{N_m \times N_r} \) are again computed by fitting the GAF data at different reduced frequencies. However, the least-square fitting requires to solve a system of nonlinear algebraic equations due to unknown coefficients in the denominator, which results in an iterative process. Morino et al. [118] later proposed a simpler and easier approach to RFA based on an interpolative structure similar
to the one proposed by Karpel [116] but requiring the solution of a linear system. Furthermore, they also proposed higher-order RFA models. More recently, Ripepi and Mantegazza [119] proposed an improved RFA interpolation structure for state-space aeroelastic stability and gust response analysis based on three nonlinear least-squares identification techniques combined with a model-order reduction.

Eversman and Tewari [120] discussed methods to optimize the aerodynamic lag using gradient-free or gradient-based approaches for improving the fitting accuracy. They also argued that the $p$ in the numerator of the rational term in Eq. (10) can be omitted, resulting in the following alternate form

$$Q(p) = A_0 + p A_1 + p^2 A_2 + \sum_{l=3}^{N_l+3} \frac{A_l}{p + \beta_{l-2}}.$$  

They optimized the aerodynamic lag parameters and minimized the number of these parameters for a given level of accuracy using a gradient-free optimization method. Pasinetti and Mantegazza [121] also used optimization to obtain a state-space aeroelastic model reduced to a minimum number of states. More recently, Nissim [122, 123] formulated a minimum-state approximation as a nonlinear optimization problem considering the aerodynamic lag terms and the two matrices that operate on them as design variables.

Using a RFA of the GAF and analytic continuation, the aeroelastic system Eq. (7) can be rewritten in the state-space form $\dot{x} = Ax$. The state vector $x = [q^T_r, \dot{q}^T_r, a^T]^T \in \mathbb{R}^N$ includes the generalized coordinates $q_r$, their derivatives $\dot{q}_r$, and $N_a$ additional aerodynamic states $a$ (whose number varies depending on the type of RFA used), for a total number of $N = 2N_r + N_a$ states. Representing the aeroelastic system in terms of a linear time-invariant model with RFAs allows the aeroelastic eigenvalues to be computed using standard linear eigenvalue analysis techniques (root locus) and it facilitates the application of control theory for active flutter suppression [10]. Moreover, it enables state-space time-domain aeroservoelastic response analysis and facilitates synthesizing load alleviation control laws. RFAs have also been used in aeroservoelastic stability and response sensitivity computation [124, 125, 126, 127, 128].

The above treatments are typically applied to linear systems. For these systems, stability is analyzed about the undeformed configuration by describing the structure using a detailed linear finite element model (FEM) and by computing the GAF using linear potential-flow lifting-surface methods like the doublet-lattice method (DLM) [129] or panel methods [18, 8]. In the presence of nonlinear aerodynamic effects, GAF matrices are computed by considering small perturbations with respect to the nonlinear background flow solution for each flight condition.

One common approach to this problem is to use DLM [129] augmented with higher-fidelity numerical data or experimental measurements [49]. An alternate, more accurate approach is to use time-linearized CFD [57] to identify the GAF matrix while capturing static nonlinear aerodynamic effects directly. If structural nonlinearities are moderate, as in the case of modern transonic transport vehicles, this statically nonlinear aerodynamic description can be still combined with linear FEM structural models. However, when nonlinear effects in the structure become significant, the equations of motion
need to be linearized about the aeroelastic equilibrium configuration at each flight condition to retain the effects of large static deflections or background nonlinear steady flow. The eigenvalues of the linearized aeroelastic system are then computed for each flight condition to determine the flutter boundary, that is, the flight conditions where damping vanishes [12].

Due to the high computational cost of nonlinear aerostructural solutions and the number of flight conditions to be analyzed in the flutter search, structural nonlinearities are frequently captured with beam models in nonlinear flutter analyses [12, 69, 68, 130, 61, 7]. These models incur low computational cost, while still being able to capture the behavior of very flexible HARW vehicles. Low-fidelity aerodynamic formulations typically used in geometrically nonlinear flutter analyses are potential-flow theories with tip and compressibility corrections [12, 69, 68, 130, 7], vortex-lattice methods [131], or surrogate models [61].

The previous section was a brief introduction to eigenvalue-analysis methods. We now discuss the critical aspects to be considered when using these methods for constraining flutter in design optimization. As previously mentioned, we mainly refer to gradient-based optimization methods, which require the evaluation of the constraint values and their derivatives with respect to the design variables.

The process of solving for the eigenvalues and computing derivatives by differentiating the eigenvalue problem analytically is widely used [132, 133, 134, 124] and it is included in several commercial analysis and design software packages [135, 136, 137]. Depending on the form of the flutter equation, different levels of accuracy are considered when taking the derivative of the eigenvalue problem. If mode shapes are used to reduce the model order, the fixed-mode approximation [134] is often used, which neglects the derivatives of the mode shapes with respect to the design variables. Such practice is commonly applied for the sake of efficiency as these derivatives, depending on the number of structural DOFs and mode shapes, can be very expensive to compute. Examples of this approximation are found in recent literature for structural optimization (with no planform changes) [138, 109, 139] and topology optimization [114, 140]. However, ignoring the derivatives of the mode shapes can result in inaccurate gradients [134].

Another aspect that is critical to the efficacy of gradient-based optimization is the smoothness of the constraint function, which should be at least $C^1$ continuous. Several considerations are necessary to formulate a smooth flutter constraint function. A naive approach would be to specify the flutter point directly. However, this may introduce discontinuities in the constraint value between two consecutive design iterations, $d^i$ and $d^{i+1}$, due to mode switching or to a hump mode becoming active at a significantly lower speed [3]. Therefore, additional steps are required to ensure a continuous constraint.

Mode switching occurs when the mode that first becomes unstable (that is, at the lower speed) changes between two consecutive design iterations $d^i$ and $d^{i+1}$. An example of mode switching is shown in Fig. 7a, where the hypothetical damping of a system with two modes is plotted with respect to speed. Mode switching causes a $C^1$ discontinuity of the flutter point, which poses a challenge to gradient-based optimizers [141]. The imaginary part of the eigenvalue (frequency) also switches, and in many cases the frequencies coalesce, causing a mode to become unstable. A more serious problem is
Critical flutter mode switching with small changes in the design. This behavior causes a $C^1$ discontinuity in the flutter point value.

Hump mode becoming active with small changes in the design. This behavior causes a $C^0$ discontinuity in the flutter point value.

Figure 7: $V - g$ diagrams showing possible sources of discontinuities in the flutter constraint. Mode 1, which is the critical mode for design $d^i$, switches with mode 2 as the critical mode in design $d^{i+1}$ (adapted from Stanford et al. [138]).

when a hump mode is present in design $d^i$ and it becomes the critical mode in the new design, $d^{i+1}$, as shown in Fig. 7b. The constraint value experiences a $C^0$ discontinuity, which is even more challenging for gradient-based optimizers.

Techniques exist to mitigate these problems and they are summarized by Stanford et al. [138]. Frequency-separation constraints proposed by Langthjem and Sugiyama [142] and also by Odaka and Furuya [143] can prevent mode switching by enforcing the critical mode to remain the same. This approach is illustrated in Fig. 8 (left) for a hypothetical case. The disadvantage of this method is that $N_m - 2$ constraints are needed for a case with $N_m$ modes (with the expectation that two modes coalesce, and hence no constraint is needed for two modes). Furthermore, specifying the unstable mode could over-constrain the optimization process. A more serious flaw with this approach is that a hump mode is still possible.

Other techniques exist to handle both mode switching and hump modes. One approach is to enforce the real part of each eigenvalue to remain below a preset bounding curve. Such a bounding curve is proposed by Stanford et al. [144], which is a modified version of the one proposed by Ringertz [141]. The boundary (dash-dotted line
Figure 8: \( V - g \) diagrams showing two possible methods to prevent discontinuities in the flutter constraint: frequency-separation method (left) and damping boundary (right, solid black). Structural frequencies for a panel are often not as well separated as for a wing, causing the frequency separation constraints to be often active. (adapted from Stanford et al. [138]).

in Fig. 8) spans the operating conditions of interest from wind-off to some maximum speed. This approach mitigates the issues mentioned above and has further benefits: 1) No constraint is placed on the flutter point itself, so there is no need to compute it explicitly; 2) Modes that become unstable abruptly (hard flutter) and have steeper slope than the boundary are handled as well.

Although the above approach mitigates the discontinuity issues, it requires one to apply a constraint for each speed increment and mode. Thus, \( N_U \) speed increments and \( N_r \) modes results in \( N_r N_U \) constraints, dramatically increasing the total number of constraints. To mitigate the high cost of computing derivatives for so many constraints, the active set method can be used. This considers the full set of points, but reduces them to a smaller set before evaluating derivatives based on the constraint value. This approach was applied by Ringertz [141] and was also applied in the time domain by Kang et al. [145]. An alternate approach to reduce the number of constraints was proposed by Haftka [146], who suggested replacing parametric constraints by minimum-value constraints. This reduces the number of constraints to the total number of modes for the entire flight envelope. The approach can further benefit from the aggregation of the set of constraints for the individual modes into a single constraint using for example the Kreisselmeier–Steinhauser (KS) function [147, 148, 149].

Other classical constraint aggregation techniques can be used, such as the \( p \)-norm function or induced aggregates [150]. Kennedy and Hicken [150] compared the accuracy of classical aggregation functions with their proposed induced aggregates. They demonstrate that the induced aggregates provide better accuracy and properties compared to the classical aggregation functions (KS and \( p \)-norm). Lambe et al. [24] studied the accuracy and efficiency of the above aggregation methods for structural mass minimization of a wingbox subject to failure constraints. The induced exponential aggregation was shown to be more accurate compared to the classical constraint aggregation functions [24].
4.1.2 Direct Methods

Eigenvalue-analysis methods were originally introduced for linear flutter analysis and later applied to systems with structural or aerodynamic nonlinearities. Due to the growing interest in analyzing nonlinear aeroelastic systems, methods based on Hopf bifurcation theory that compute the flutter point directly in the frequency domain while enforcing the nonlinear aeroelastic equilibrium equations were more recently used. Approaches were also developed to compute derivatives for direct methods to use them for the formulation of constraints in gradient-based optimization (see Section 4.2).

Here, we introduce the formulation of the Hopf-bifurcation approach and highlight recent work that contributed to its development. In particular, we focus on efforts that compute the Hopf point directly [151]. This is in contrast to indirect methods that monitor eigenvalues, look for sign changes, and then perform a local search (using a Newton–Raphson, secant, bisection, or interpolation method) [19], similar to what was discussed in Section 4.1.1.

Consider a generic nonlinear aeroelastic system with the dynamics governing equations written as

\[
\dot{x} = f(x; U).
\]  

(13)

The state vector \(x = [q^T, \dot{q}^T]^T \in \mathbb{R}^N\) includes the generalized coordinates assumed to describe the system and their derivatives, for a total number of \(N\) states, \(f \in \mathbb{R}^N\) is a nonlinear vector field, and \(U\) is the flight speed (the bifurcation parameter). The system Jacobian matrix is \(A := \partial f / \partial x \in \mathbb{R}^{N \times N}\). A Hopf bifurcation point (flutter point) occurs when the flight speed reaches the critical value \(U_F\), which corresponds to an equilibrium state \(x_F\) such that \(f(x_F; U_F) = 0\). The Jacobian matrix evaluated at the Hopf point \(A(x_F; U_F)\) has a pair of purely imaginary eigenvalues \(\pm i\omega_F\), where \(\omega_F\) is the flutter angular frequency. At this point, they cross the imaginary axis, causing the Hopf bifurcation. As explained in Section 3.2, this leads to an LCO.

Let the right complex eigenvectors associated with the bifurcating eigenvalues be \(u = u_r + i u_i\), and \(\overline{u} = u_r - i u_i\), where \(u_r, u_i \in \mathbb{R}^N\). The critical eigenvectors are computed by solving

\[
(A - iI\omega_F)u = 0,
\]  

(14)

where \(I \in \mathbb{R}^{N \times N}\) is an identity matrix, and using the normalization condition \(c^T u = i\), where \(c \in \mathbb{R}^N\) is arbitrary and constant. The flutter point is given by the solution of the nonlinear algebraic system,

\[
\begin{cases}
    f(x_F; U_F) = 0 \\
    Au_r + \omega_F u_i = 0 \\
    Au_i - \omega_F u_r = 0 \\
    c^T u_r = 0 \\
    c^T u_i = 1
\end{cases},
\]  

(15)

where the \(3N + 2\) unknowns are the components of \(x_F, u_r,\) and \(u_i\), the flutter speed \(U_F\), and the angular frequency \(\omega_F\). This system can be re-written in the compact form as

\[
F(X) = 0, \quad \text{where} \quad X = [x_F^T \quad u_r^T \quad u_i^T \quad U_F \quad \omega_F]^T,
\]  

(16)
and solved iteratively via the Newton–Raphson method. Starting from a first guess $X^{(0)}$ for the solution vector, at the $n$th iteration this is updated by solving

$$J^{(n)} \Delta X^{(n)} = -F^{(n)}$$

(17)

where $J = \partial F / \partial X$ is the Jacobian of the extended system and $\Delta X^{(n)}$ is the Newton–Raphson update. Once the linear system, Eq. (17), is solved, the solution vector at the next iteration is updated as $X^{(n+1)} = X^{(n)} + \Delta X^{(n)}$ and the process is repeated until convergence is reached.

While the Newton–Raphson method usually exhibits fast convergence, it is not without some disadvantages. In particular, this method requires a suitable initial guess $X^{(0)}$ that is sufficiently close to the Hopf point in order to converge successfully [152]. Furthermore, this method may converge to other Hopf points, such as divergence [153], or to a zero velocity.

A common strategy to overcome these issues is to solve the eigenvalue problem using large $\Delta U$ steps and track when one eigenvalue crosses the real axis. Once the crossing has been identified, the direct Newton method is applied to identify the Hopf point, which typically converges in few iterations [154, 155].

Computing the Jacobian $J$ accurately is also a challenge. Various techniques can be employed to approximate the Jacobian when solving Eq. (17): approximating it entirely using finite differences, approximating or omitting individual entries at the expense of convergence, or reducing the size of the Jacobian by taking advantage of symmetries in the problem [152, 153, 155, 156]. For practical aircraft configurations modeled by CFD and FEM, the resulting Jacobian can be a large sparse system (which requires a preconditioning strategy) that is also ill-conditioned [153, 157].

Hui and Tobak [158] were among the first authors to apply the Hopf-bifurcation method to investigate the flutter behavior of nonlinear aeroelastic systems. They analyzed the stability of pitch motion about a large mean angle of attack and showed that the aerodynamic damping vanishes when the mean angle of attack exceeds a critical value.

Morton and Beran [152] studied a two-degree-of-freedom airfoil undergoing plunge and pitch in inviscid transonic flow. The aerodynamics was modeled with an Euler CFD solver, while the structural model consisted of linear and torsional springs for the plunge and pitch DOFs, respectively. They computed the flutter point directly using the Hopf-bifurcation method and validated it with time-accurate analyses. The Jacobian was computed using second-order finite differences and the system, Eq. (16), was solved using a Newton–Raphson method. They showed that the approach is computationally efficient compared to full time-accurate simulations, and that it produces accurate results for inviscid transonic flow.

Badcock et al. [159] extended their previous work [153] and applied a modified Newton–Raphson method to solve the Hopf-bifurcation problem. The method was demonstrated on the AGARD 445.6 aeroelastic wing [160] by computing the full flutter boundary in an inviscid transonic flow. For this configuration, the direct approach was up to two orders of magnitude more efficient than time marching.

Woodgate and Badcock [161] extended this work and analyzed three geometries: the Goland wing [162], a transonic wing-body configuration [163], and the AGARD 445.6
The direct method reduced the computational cost by one to three orders of magnitude compared to time-accurate simulations.

More recently, Badcock and Woodgate [157] further improved the approach for large-scale CFD-based systems coupled with FEM structures. They used a Schur-complement eigenvalue formulation to treat parts of the system Jacobian in a parallel computing environment. This procedure eliminated the process of solving large sparse systems, which are almost singular, further increasing the method efficiency. This method shows similarities to the eigenvalue methods applying mode tracking discussed in Section 4.1.1. The approach was applied to the Goland wing, a supersonic transport arrow rudder buzz in transonic flow [164], the wing-body configuration mentioned above [163], and a generic fighter configuration based on publicly available F-16 data [157, 165]. The computational cost of the proposed method varied from 5 to 60 times that of a steady-state solution.

In the context of gradient-based optimization, formulating a successful flutter constraint using the direct method needs careful consideration. Possible discontinuities due to competing flutter mechanisms in the design process, such as mode swapping or hump modes, described in Section 4.1.2, may confuse the gradient-based optimizer. However, techniques addressing these concerns in the context of direct methods are deficient or nonexistent.

### 4.2 Applications to Optimization

Despite early work by researchers such as Haftka [3, 146] and Hajela [166] optimization subject to flutter constraints is still not a standard design practice. More recently, several authors have integrated flutter constraints into design and investigated their effect on the optimal solutions. These efforts are summarized in Table 1 and reviewed below. They differ in the use of eigenvalue analysis (EV), direct (Hopf), or time-domain (TD) prediction methods for flutter analysis, the fidelity of structural and aerodynamic models, and the optimization problem formulation in terms of objective, type and number of design variables, and use of gradient-based (GB) or gradient-free (GF) solution algorithms. Efforts that used gradient-based algorithms also differ in the methods used for computing derivatives with respect to design variables. Finally, previous work shown in Table 1 differs in the types and number of design variables that were considered in the optimization problem. Most of the efforts included only structural sizing variables without considering changes in the aerodynamic shape. Fewer efforts optimized the airfoil or planform shapes. Including planform shape variables is challenging because changes in the mode shapes and the corresponding natural frequencies need to be considered when computing derivatives, incurring additional computational cost.
Table 1: Summary of efforts on optimization or derivative computation considering flutter.

<table>
<thead>
<tr>
<th>Effort</th>
<th>Method*</th>
<th>Aerodynamics†</th>
<th>Structure‡</th>
<th>Objective§</th>
<th>Design variables</th>
<th>Shape</th>
<th>N</th>
<th>Algorithm‖</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turner [167]</td>
<td>EV</td>
<td>Strip theory</td>
<td>Panel</td>
<td>Min mass</td>
<td>Structural</td>
<td>Airfoil</td>
<td>Planform</td>
<td>3</td>
</tr>
<tr>
<td>Bhatia and Rudisill [168]</td>
<td>EV</td>
<td>–</td>
<td>Beam</td>
<td>Min mass</td>
<td>12 GB-Analytic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rudisill and Bhatia [169]</td>
<td>EV</td>
<td>–</td>
<td>Beam</td>
<td>Min mass</td>
<td>12 GB-Analytic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gwin and Taylor [170]</td>
<td>EV</td>
<td>DLM</td>
<td>Beam</td>
<td>Min mass</td>
<td>33 GB-Analytic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ringertz [141]</td>
<td>EV</td>
<td>DLM</td>
<td>Composite</td>
<td>Min mass</td>
<td>9 GB-Analytic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mallik et al. [171]</td>
<td>EV</td>
<td>Strip theory</td>
<td>Beam</td>
<td>Min FB, Max TOGW</td>
<td>19 GF-Genetic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jonsson et al. [110]</td>
<td>EV</td>
<td>DLM</td>
<td>Shell</td>
<td>Max range</td>
<td>3 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stanford et al. [144]</td>
<td>EV</td>
<td>TSD+Euler</td>
<td>Shell</td>
<td>Min mass</td>
<td>92000 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen et al. [137]</td>
<td>EV</td>
<td>Euler</td>
<td>Shell</td>
<td>–</td>
<td>Derivative computation only</td>
<td></td>
<td></td>
<td>GB-Complex-step</td>
</tr>
<tr>
<td>Bartels and Stanford [139]</td>
<td>EV</td>
<td>RANS</td>
<td>Shell</td>
<td>Min mass</td>
<td>711 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Varyar et al. [7]</td>
<td>EV</td>
<td>Lifting line</td>
<td>NL beam</td>
<td>Min FB</td>
<td>12 GB-Finite difference</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xie et al. [131]</td>
<td>EV</td>
<td>DLM</td>
<td>NL beam</td>
<td>Min mass</td>
<td>44 GF-Direct</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhatia and Beran [172]</td>
<td>EV</td>
<td>Euler</td>
<td>NL beam</td>
<td>Min mass</td>
<td>8 GB-Analytic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lupp and Cesnik [130]</td>
<td>EV</td>
<td>Strip theory</td>
<td>NL beam</td>
<td>Min FB</td>
<td>5 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kennedy et al. [156]</td>
<td>Direct</td>
<td>Panel</td>
<td>Shell</td>
<td>–</td>
<td>Derivative computation only</td>
<td></td>
<td></td>
<td>GB-Adjoint</td>
</tr>
<tr>
<td>Beran et al. [173]</td>
<td>Direct</td>
<td>ONERA stall</td>
<td>NL beam</td>
<td>–</td>
<td>Derivative computation only</td>
<td></td>
<td></td>
<td>GB-Adjoint</td>
</tr>
<tr>
<td>Mani and Mavriplis [99]</td>
<td>TD</td>
<td>Euler</td>
<td>Mass-spring</td>
<td>Max fl. speed</td>
<td>32 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. [101]</td>
<td>TD</td>
<td>RANS</td>
<td>Mass-spring</td>
<td>Max fl. speed</td>
<td>48 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhang et al. [102]</td>
<td>TD</td>
<td>RANS</td>
<td>FEM</td>
<td>Max fl. speed</td>
<td>120 GB-Adjoint</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Methods: EV—Eigenvalue method, TD—Time-domain
† Aerodynamics: DLM—Doublet lattice method, TSD—Transonic small disturbance, RANS—Reynolds averaged Navier–Stokes
‡ Structures: NL—Nonlinear
§ Objective: FB—Fuel burn, TOGW—Takeoff gross weight
‖ Algorithm: GB—Gradient based, GF—Gradient free
In one of the early efforts to constrain flutter, Turner [167] formulated a mass minimization problem by considering distribution of material rather than the structure topology to meet a specified flutter speed. Later, Bhatia and Rudisill [168] developed a numerical procedure to minimize wing mass while satisfying a flutter constraint. They applied the procedure to a uniform cross-section box beam consisting of three bays in order to minimize the mass while maintaining the flutter speed. Rudisill and Bhatia [169] improved the rate of convergence of this procedure by computing second-order derivatives of the eigenvalues and of the flutter speed with respect to the design variables. Gwin and Taylor [170] developed the method of feasible directions for the mass minimization of a structure subject to a minimum flutter speed. They were able to handle up to 33 structural design variables in a supersonic problem.

Due to the limited computational capabilities of the time, these early efforts used simple aerodynamic and structural models and employed a similar strategy to enforce the flutter constraint. They all formulated the flutter problem in the frequency domain as an eigenvalue problem, which was then differentiated with respect to the design variables. The derivatives of the eigenvalues were obtained using the left and right eigenvectors and one of these efforts also considered second-order derivatives to better guide the optimization process and improve the rate of convergence [169].

Ringertz [141] applied the $k$-method to minimize the weight of a cantilevered wing in incompressible flow subject to flutter and divergence constraints. The structure was modeled as a composite FEM model, while the unsteady aerodynamic loads were computed using the DLM. The eigenvalue problem was analytically differentiated in the modal space. A continuous flutter constraint was formulated using a boundary to constrain the damping values, which resulted in a large number of constraints. The method was demonstrated on a rectangular wing and on a swept wing with taper, where the objective was to minimize weight with respect to element group thicknesses subject to flutter and divergence constraints. In both cases, a considerably lighter design was achieved.

Mallik et al. [171] investigated the impact of a previously developed flutter constraint [174] on the MDO of a TBW aircraft. The flutter speed was computed using the $k$-method applied to a linear pre-stressed structural model. The unsteady aerodynamic loads were computed using Theodorsen theory [175] with a Prandtl–Glauert compressibility correction. They implemented an iterative procedure to ensure consistency of the flutter speed, Mach number, and altitude, and optimized a representative TBW configuration for minimum takeoff weight and fuel burn using a genetic algorithm. The flutter constraint was formulated by enforcing a minimum flutter Mach number and was added to several other mission constraints. Comparing the optimization results with those obtained by removing the flutter constraint, they showed that this is necessary to obtain a flutter-free optimal solution.

Jonsson et al. [110] developed a flutter constraint suitable for high-fidelity gradient-based optimization including wing planform variables for the first time. They developed an efficient non-iterative root finding method to compute the flutter eigenvalues, based on the $pk$ flutter equation that includes a scheme to track the mode migration between subsequent reduced frequencies and designs.

The flutter-free flight envelope was implicitly defined using a constraint curve for
the damping values to prevent discontinuities in the flutter constraints and to allow the computation of derivatives. The differences between damping values and the constraint curve were aggregated into a single value using the KS function [147]. Accurate and efficient derivatives of the aggregated flutter constraint value were computed with respect to both structural sizing and wing planform variables. The derivatives were computed using a combination of analytic and automatic differentiation methods in reverse (adjoint) mode and validated against the complex-step method.

The flutter constraint was used to maximizing the range of an idealized rectangular wing (flat plate) by varying structural thickness and planform design variables. The optimal design had an increased aspect ratio and an improved range while remaining flutter and divergence free. This approach provides a way to include airfoil shape variables in the future to perform full aerostructural design optimization [176] subject to flutter constraints.

Several authors have focused on including aerodynamic nonlinearities in flutter analysis to optimize transonic configurations. Stanford et al. [144] evaluated six different novel tailoring schemes employed in mass minimization optimization. They analyzed the flutter characteristics using the pk method and the commercial software ZTRAN [50] to retain aerodynamic nonlinearities. The nonlinear higher-fidelity Euler code, ZEUS, was used to compute steady background flow at multiple transonic Mach numbers for a fixed cruise shape. Using these steady-state CFD solutions as an input, the linearized unsteady loads were computed for a range of reduced frequencies using time-linearized transonic small disturbance (TSD) analyses about the equilibrium solutions.

The system damping values were forced to be under a stability boundary, similarly to the approach by Ringertz [141]. The transonic aerodynamic loads were computed offline before the optimization and their derivatives were obtained by differentiating the eigenvalue problem [134]. Flutter, stress, and buckling constraints computed in this study were all aggregated using a KS function [147]. They considered the fixed-mode derivatives to improve computational efficiency, an approach that does not allow for shape changes.

They studied the undeflected Common Research Model (uCRM) wing [176] and obtained six different optimal wing structures corresponding to different tailoring schemes, all for the same operating condition and setup. The six tailoring schemes considered for the structural design were metallic thickness variations, functionally graded materials, balanced or unbalanced composite laminates, curvilinear tow steering, and distributed trailing edge control surfaces. While there was a structural wing mass reduction for every case, many of the lighter designs had an active flutter constraint, while the buckling constraint was active for the heavier cases.

Chen et al. [137] extended their previous work [136] by computing derivatives of flutter constraints with respect to shape variables using ZEUS coupled with a boundary layer code. The derivatives with respect to shape were computed using the complex-step approach [28, 31], which is numerically exact. The flutter constraint was formulated using the g-flutter method [87], which was differentiated analytically with respect to the design variables. The approach was verified for a cantilevered planform, similar to the F-5 geometry [177, 178], where a structure consisting of 10 spars, 10 ribs, and
upper and lower skins was modeled using MSC Nastran. While no optimization results were presented, the flutter derivatives were verified against exact results.

Bartels and Stanford [139] proposed an approach to enforce a CFD-based flutter constraint for gradient-based structural optimization of transonic vehicles. The flutter analysis was performed as a standard eigenvalue analysis on the state-space representation of the aeroelastic system obtained via RFA [115]. The GAF matrix of the baseline structure was identified from time-linearized unsteady RANS simulations about nonlinear steady-state solutions [179]. The GAF matrix of the updated design was then computed by projecting the updated modes onto the baseline modes. The methodology was used to minimize the uCRM mass subject to structural and aeroelastic constraints. The flutter was constrained by requiring that the real part of the system eigenvalues be below a bounding curve [166]. The optimization assumed a fixed-mode approximation. The authors compared the optimal solutions obtained using unsteady RANS and DLM aerodynamics in the optimization loop. They showed that the DLM-based solution was not conservative and had a significantly different thickness distribution compared to the CFD-based optimal design.

Opgenoord et al. [65] extended a low-order two-dimensional transonic flutter prediction model [64] to wings and implemented the model into a conceptual aircraft design tool to investigate the impact of geometric parameters and Mach number on the flutter boundary. Furthermore, they optimized the D8.0 configuration by minimizing the maximum take-off weight and fuel burn with and without a transonic flutter constraint. Enforcing the flutter constraints resulted in lower optimal aspect ratio and a weight penalty or lower fuel burn reduction compared to the case when the flutter constraint was omitted. Opgenoord et al. [66] also used the developed flutter model to optimize the D8.0 configuration by minimizing weight with and without enforcing a flutter constraint in addition to stress and buckling constraints. The optimal design including the flutter constraint showed only a slight mass increase thanks to the appropriate aerelastic tailoring of the lattice structure.

While several authors have performed flutter-constrained optimizations using nonlinear aerodynamic models, examples of flutter constraints considering nonlinear structures are more rare due to the more recent interest in optimizing very flexible aircraft. Variyar et al. [7] developed a framework for MDO of geometrically nonlinear aircraft subject to flutter constraints by coupling the SUAVE design tool [180] with the ASWING nonlinear aeroelastic solver [181]. They developed an interface to convert the arbitrary aircraft designs output by SUAVE into equivalent nonlinear beam models for ASWING to drive structural and aeroelastic analyses and to post-process the results. The flutter speed was obtained iteratively by computing the eigenvalues of the statically deformed aircraft at different flight conditions.

This MDO framework was used to optimize the Sugar VOLT strut-braced aircraft [182] for minimum fuel burn subject to mission, maneuver, gust, and flutter constraints. The flutter constraint was implemented by imposing a minimum flutter speed, and derivatives with respect to the design variables were obtained via finite differences. The authors performed three MDO cycles by adding the maneuver, gust, and flutter constraints to the mission constraints. Despite having better performance, the optimal solution achieved with only maneuver and gust constraints experienced
flutter within the flight envelope, highlighting the need for a flutter constraint in the
design process.

Lupp and Cesnik [130] studied the effect of a flutter constraint including geometric-
ical nonlinearities on the design of a BWB. They extended the University of Michigan’s
Nonlinear Aeroelastic Simulation Toolbox (UM/NAST) to utilize AD in reverse mode
to determine coupled aeroelastic derivatives including geometrical nonlinearities. They
proposed an algorithm to increase the computational efficiency of the gradient eval-
uation for a geometrically nonlinear aeroelastic analysis. The sample optimization
formulation included a geometrically nonlinear beam-based flutter constraint using a
KS aggregation [110] to obtain a scalar constraint for the entire flight envelope. The
authors ran three fuel burn minimization cases: with a strength constraint, with a
linear flutter constraint, and with a geometrically nonlinear flutter constraint with the
wing chord distribution, wing box size, and wing box thickness as design variables.
While the linear flutter constraint became active over the strength constraint, it was
not conservative compared to the geometrically nonlinear constraint. They concluded
that a geometrically nonlinear flutter constraint is needed for very flexible aircraft.

Xie et al. [131] used a flutter constraint to minimize the weight of a very flexible
wind tunnel model. The purpose of this constraint was to ensure flutter within the
wind tunnel speed range. The optimization problem coupled a geometrically nonlinear
beam solver with a vortex lattice code for the static aeroelastic analysis, while a doublet
lattice code was used to compute the unsteady loads. Since they use a gradient-free
algorithm, no derivatives were computed. They compared optimization results based
on linear and geometric nonlinear beam models subject to a flutter constraint. In
addition to flutter, tip displacement and torsion angle constraints were also enforced.
The linearly optimized configuration resulted in a wing lighter than the optimal so-
lution obtained with the nonlinear flutter constraint. Furthermore, the flutter and
displacements constraints were violated when the linear optimized configuration was
analyzed considering geometric nonlinear effects. The difference in the results high-
lighted the need for a flutter constraint when optimizing very flexible vehicles and the
importance of using nonlinear flutter prediction methods not only for analysis, but in
design optimization as well.

Bhatia and Beran [172] developed a framework to optimize thermally stressed non-
linear structures subject to transonic flutter constraints. They showed that including
aerothermoelastic static nonlinearities in the flutter analysis impacts the optimal de-
sign. These effects are particularly important for high-speed vehicles, which are subject
to significant thermoelastic stresses when flying through the transonic regime during
reentry [183]. The authors optimized a skin panel with respect to the thickness and
density distributions to minimize the mass subject to a flutter constraint. The struc-
ture was modeled using a Timoshenko beam finite element with nonlinear von Kármán
strain, while the transonic flow was solved via a finite-element discretization of the Eu-
er equations. The structure was linearized around the static thermoelastic response
and the aerodynamics was linearized around the background steady transonic flow past
the baseline geometry. Flutter was analyzed in the frequency domain as an iterative
$V-g$ solution [1] considered the system linearized around the nonlinear equilibrium
configuration for each operating condition. The flutter speed was constrained directly.
The optimal solution obtained by linearizing around the thermally stressed equilibrium configuration showed a mass lower than the result using an unstressed analysis.

The Hopf-bifurcation method has also been applied in design optimization. Kennedy et al. [156] proposed a variant of previous bifurcation approaches [159, 157] to optimize an aeroelastic system subject to a flutter constraint, which was formulated in terms of flight speed rather than damping. The proposed method had the benefit of not requiring a search of the flutter point (which may be located well outside of the flight envelope) at each design iteration. The KS function [147] was used to smooth the effect of mode switching in the constraint value by aggregating the less damped modes, which yielded smooth gradients. The authors presented preliminary results for a medium fidelity three-dimensional aerodynamic panel code coupled with the structural finite-element code TACS [184]. Although no detailed optimization was performed, they performed a preliminary study on the uCRM benchmark [176].

Beran et al. [173] developed a fast adjoint method to compute derivatives of flutter points computed via the Hopf-bifurcation method for gradient-based MDO. The approach was applied to the highly flexible cantilevered wing studied by Tang and Dowell [81, 82]. Both geometric nonlinearities due to the structure and aerodynamic effects described using the ONERA stall model [82] were considered when computing the flutter point and its derivatives with respect to aerodynamic and structural design variables. While no optimization study was presented, the authors outlined the future work required to apply the methodology: validate the derivatives, assess the computational cost compared to alternate time- and frequency-domain flutter prediction methods, and develop the handling of mode switching to avoid discontinuous flutter points.

Some rare examples of optimizations with flutter constraints based on time-accurate analyses are also found in the literature, but they have been restricted to simple problems. In an early work, Holden [185] applied a collocation method to constrain the aeroelastic response envelope for a wing optimization problem. More recently, Mani and Mavriplis [99] were among the first to present a fully coupled unsteady adjoint for aeroelastic optimization. They demonstrated their method successfully in a shape optimization of two-dimensional airfoil section to suppress flutter, using a total of 32 design variables in the form of Hicks–Henne bump functions [186]. Later, Mishra et al. [100] extended previous work and presented the fully coupled unsteady adjoint for three-dimensional aeroelastic problems, which was demonstrated in a shape optimization of a flexible rotorcraft configuration.

Zhang et al. [101] developed a coupled adjoint method for unsteady aerostructural problems solved via time simulations. The method was applied to an airfoil shape optimization problem with the goal of suppressing flutter. The aerodynamics was computed with an Euler CFD code coupled with a boundary layer code to account for viscous effects. Both the continuous and discrete coupled adjoint were developed for steady-state analyses. The discrete approach proved more promising and only this version was developed for unsteady cases. A damping objective function was proposed that used a Hilbert transform [92] of the nonlinear unsteady time-history. To achieve the required flutter boundary, the damping objective function was minimized to obtain a neutral response, indicating the flutter point. The authors applied the methodology
to the optimization of the two-dimensional (2D) Isogai airfoil [37, 38] to suppress flutter. Only derivatives with respect to shape variables were computed and the airfoil shape was parametrized using 48 Hicks–Henne bump functions [186]. A neutrally stable (zero-damping) configuration was obtained for a given flight condition.

Zhang et al. [102] extended their previous work [101], where the coupled adjoint was developed for time-marching simulations. Two objective functions were used: one that maximized the flutter boundary and another that matched a given flutter boundary. To maximize the flutter boundary, they minimized the squared and time averaged history of the lift coefficient. To match a given flutter boundary, the damping value of a given time history was minimized to obtain a neutral response. B-spline curves were chosen as a parametrization method due the large number of Hicks–Henne functions that were previously needed. The authors presented steady-state optimization results for a 2D airfoil optimized to match a given pressure distribution and for the 3D ONERA M6 case [187], where the objective was a composite function of lift and drag. Configurations considered in the unsteady optimization consisted of the 2D Isogai airfoil and the Goland wing as modeled by Kurdi et al. [188]. For the unsteady 2D airfoil, two optimization cases were considered: flutter margin maximization and a flutter boundary matching (i.e., a neutral response for the given flight condition). The design variables were the plunge and pitch stiffness values. For the Goland wing, they optimized the aerodynamic shape to maximize the flutter speed with respect to 120 shape variables. A structural optimization of the Goland wing was also performed to maximize the flutter speed with respect to the skin thickness. No optimization was performed using simultaneously structural sizing and aerodynamic shape variables.

Kiviaho et al. [189] developed a flutter constraint using their previously developed unsteady aeroelastic framework with adjoint sensitives [103, 104]. The flutter constraint is based on a matrix pencil method [90] applied to a time-history, which estimates the damping based on most critical aeroelastic modes. Two methods are proposed, a direct flutter point evaluation which finds the flutter point where the dynamic pressure lower bound is specified as the design flight condition, and a flutter margin or clearance approach where the dynamic pressure times some tolerance is fixed. The direct method was demonstrated in a single design variable optimization of an airfoil were the dynamic pressure is minimized subject to the flutter constraint in order to identify the flutter point.

### 4.3 Open Problems

Now that we reviewed methods for flutter prediction and their applications to optimization (see Table 1), we can summarize the open problems in goal of integrating flutter constraints into aircraft design optimization.

Gradient-based optimization is the preferred choice for optimizations with respect to large numbers of design variables. When enforcing a flutter constraint in a gradient-based optimization, a serious challenge is varying structural and aerodynamic design variables simultaneously to optimize the aircraft external shape, planform, and internal sizing. Most of previous work optimized only structural sizing, and only a few efforts included airfoil shape and wing planform design variables as well. Furthermore,
simplifying assumptions like the fixed-mode approximation were frequently used when computing derivatives to limit computational cost \([138, 109, 139]\). These assumptions are adequate for a structural optimization, but they can lead to inaccurate results when varying aerodynamic properties because this can cause significant changes in the mode shapes at each design iteration. Therefore, gradient-based optimizations with respect to structural, planform, and shape variables need approaches that take into account the derivatives of the mode shapes when computing the derivatives of the flutter constraints. Few examples of these approaches applied to simplified configurations or using lower-fidelity models are available in the literature \([110, 7, 130]\). However, they still have to be demonstrated on practical configurations parametrized by large number of structural and aerodynamic design variables.

A second major challenge is developing efficient flutter analysis models and methods applicable in the presence of aerodynamic or structural nonlinearities. In previous aircraft optimizations including linear flutter constraints, the natural choice was to analyze flutter in the frequency-domain due to the availability of well-established and computationally efficient eigenvalue-analysis methods. The computational cost of these methods is a big challenge when including aerodynamic or structural nonlinearities because the flutter point depends on the equilibrium state. For each flight condition considered in the flutter search, the steady background flow solution (in the case of aerodynamic nonlinearities) or the coupled aerostructural equilibrium (in the case of both aerodynamic and structural nonlinearities) needs to be determined first. Next, the linearized model about each equilibrium point must be identified for computing the aeroelastic eigenvalues and determine at which point modal damping vanished. This process must be repeated for each flight condition, while linear methods analyze flutter by considering the undeformed configuration at zero angle of attack for each flutter search point. Moreover, in the presence of nonlinear effects multiple equilibrium points may also exist for each flight condition, which further increases complexity and computational cost. Stability must be analyzed about all equilibrium points, otherwise critical constraint values may be missed.

For large high-fidelity models with both structural and aerodynamic nonlinearities, computing the aerostructural equilibrium points and the corresponding linearized systems may be computationally prohibitive. For moderately flexible configurations, structural nonlinearities can be neglected, so eliminating the need to solve a nonlinear static aeroelastic problem at each flight condition. However, transonic aerodynamic nonlinearities still require identifying a linearized unsteady aerodynamic model for each steady background flow solution. Computing and retaining aerodynamic nonlinearities to accurately predict the flutter point in the transonic regime remains a big challenge due to the large computational cost associated with CFD. To address this problem, some efforts tried to preserve the computational efficiency of lower-fidelity methods while retaining the nonlinear physics modeled by the higher-fidelity CFD methods \([52, 50]\). Other works proposed time-linearized CFD solvers \([53, 57]\) or reduced- or low-order models calibrated using CFD \([64, 62]\). Despite these progresses, optimizing aircraft subject to transonic flutter constraints is still an open problem, particularly when using gradient-based methods and when seeking both structural sizing and aerodynamic shape changes.
Due to these challenges, previous work mainly optimized linear aircraft configurations or included only aerodynamic nonlinearities while limiting to structural sizing. Aerostructural optimizations considering both aerodynamic and geometric structural nonlinearities used low-order models or optimized simple configurations to limit the computational cost [7, 130]. No previous work considered both structural and aerodynamic nonlinearities in a high-fidelity MDO setting. Additionally, rigid-body DOFs were never included in the flutter constraints, which may lead to unfeasible designs for configurations prone to coupled rigid-elastic instabilities.

Few studies used alternatives to frequency-domain eigenvalue analysis methods for nonlinear flutter prediction, like the Hopf-bifurcation method or time-domain simulations. Such studies are rare because the high computational cost of time-accurate analyses makes them prohibitive to optimize complex configurations. For this reason, these applications were limited to simple configurations and frequently two-dimensional problems.

5 Post-Flutter Analysis in Aircraft Design Optimization

As discussed in Section 3, flutter analysis is not sufficient to characterize the dynamic response of nonlinear aeroelastic systems. For these systems, destabilizing nonlinear effects may cause LCOs even before reaching the flutter boundary, if the disturbances are sufficiently large (this is the subcritical post-flutter scenario shown in Fig. 6). Preventing or mitigating this behavior through design optimization requires constraining not only flutter, but post-flutter characteristics as well.

Characterizing and constraining the LCO behavior of aircraft requires the computation of bifurcation diagrams like the ones shown in Fig. 6. Several methods exist for this purpose. Standard techniques to obtain bifurcation diagrams are dependent on the mathematical representation of the system dynamics in the form of Eq. (13). We refer to these techniques as model-based prediction methods. However, a mathematical representation of the system is not always available, particularly for complex aircraft configurations that would require too much theoretical development or computational effort. Therefore, model-free bifurcation prediction approaches that use only output data have been recently proposed and applied to aeroelastic systems.

In this section, we summarize the theory of model-based and model-free methods to predict post-flutter behavior and discuss examples of their application to optimization problems and open challenges in this field.

5.1 Model-Based Prediction Methods

Model-based post-flutter prediction methods can be broadly classified under four categories: time-domain methods, frequency-domain methods, nonlinear perturbation methods, and continuation methods [79]. Time- and frequency-domain methods compute limit cycles for a given flight speed (or other flight parameters). Each speed corresponds to a point in the bifurcation diagram, which must be constructed point-by-point by means of a sweep on speed. Nonlinear perturbation methods seek a reduced-order representation of Eq. (13) in a neighborhood of the Hopf bifurcation point, which cor-
responds to the flutter point, and use the reduced-order model to directly obtain a local approximation of the bifurcation diagram. These methods are accurate close to the flutter point, but they lose accuracy for large perturbations. Finally, continuation methods trace a branch of the bifurcation diagram starting from a known point obtained from a time- or frequency-domain analysis method.

5.1.1 Time-Domain Methods

Time-domain methods can be either time marching or time periodic. Both types discretize the time domain to compute the transient response. Time-marching methods compute the response to a perturbation about an equilibrium point, while time-periodic methods compute the fully developed LCO directly.

Time-marching methods are the most common approach to characterize post-flutter behavior. These methods integrate Eq. (13) from the initial time up to the development of a periodic solution. This can be done using a variety of explicit or implicit time integration schemes, although implicit schemes are preferred due to the better stability characteristics that allow larger time steps, and thus lower computational times.

The main advantage of time marching is the direct applicability to models of any fidelity, from analytic representations to high-fidelity models based on the coupling of FEM and CFD. Most early studies on LCOs for aeroelastic systems were based on time-domain simulations [80, 78, 67, 190, 12, 191, 192, 193]. There have also been studies that validated the numerical results against experiments [194, 81, 82]. The review papers cited previously also cover this topic [18, 8, 14].

Despite the insights obtained from these studies, they typically considered simple configurations such as 2D airfoils [192, 193, 190] or plates [191], cantilevered wings modeled using beam or shell structures [81, 82, 190, 78], or low-order models of full-vehicle configurations [80, 67, 12]. This is because brute-force integration of Eq. (13) has several drawbacks that makes it prohibitive for analyzing complex configurations and for optimizing even simple ones.

An obvious disadvantage of analyzing LCOs via transient simulations is that periodic solutions may require long integration times to develop fully. Furthermore, tracing a portion of the bifurcation diagram for a fixed design requires performing several transient simulations. For optimization, this procedure also needs to be repeated at each design iteration. Therefore, time-marching LCO prediction methods are computationally costly for design optimization of practical configurations.

Additionally, optimizing large high-fidelity models requires the use gradient-based approaches, which need derivatives, typically computed using adjoint methods. Despite their advantages for optimizations with respect to large numbers of design variables, adjoint methods are computationally costly when computing gradients for dynamic responses [96]. This is because a reverse time-integration is necessary to obtain the adjoint vector of a time-marched system [97]. This cannot be done simultaneously to the analysis computing the dynamic response and it requires large memory storage [98]. For low- and medium-fidelity models, this issue can be mitigated by computing analytic derivatives of unsteady response metrics directly.

Stanford and Beran [195] developed a medium-fidelity nonlinear aeroelastic frame-
work for implicit time marching with direct computation of analytic derivatives and used it to optimize a wing structure for minimizing the LCO amplitude. Kennedy and Martins [196] developed a fully coupled adjoint-based derivative computation method for time-dependent aeroelastic optimization of flexible aircraft, using a panel code and a finite element code [184]. Recent work has focused on developing analytic adjoint derivatives for unsteady aeroelastic computations involving CFD solutions and demonstrating their feasibility for optimization [99, 104, 189]. However, these approaches have not so far been used for constraining LCOs of nonlinear aeroelastic systems.

In contrast to time-marching methods, time-periodic approaches skip the initial transient phase and only compute fully developed periodic solutions. Time-periodic methods consist of cyclic finite-difference methods [79], cyclic spectral element methods [197, 198, 199], and shooting methods [19, 79]. A related class of methods is based on finite elements in time and on combined methods for simultaneous spatial and temporal model order reduction [200, 201].

Cyclic finite differences apply a time-marching scheme to a period $T$ of the LCO (between $t = t_0$ and $t = t_0 + T$), and enforce the solution periodicity at the time boundaries. The LCO period is discretized in $M$ non-dimensional time intervals defined by the points $\tau_0, \ldots, \tau_M = \tau_0 + 1$, where $\tau = t/T$ is the non-dimensional time variable. The governing equations, Eq. (13), can be approximated at the mid-point of the $i$th time interval using a central-difference scheme for the time derivative and the trapezoidal rule for the right-hand side (single-step Adams–Moulton method), yielding [79]

$$x_i - x_{i-1} = \frac{\Delta \tau}{2} T \left[ f(x_i; U) + f(x_{i-1}; U) \right], \quad (18)$$

where $\Delta \tau$ is the non-dimensional time step, and $i = 1, \ldots, M$, and the vector $x$ lists the $N$ system DOFs. Writing an equation like this for each time interval and adding the periodicity condition $x_0 = x_M$ yields a system of $NM$ nonlinear algebraic equations in the $NM + 1$ unknowns $x_0, \ldots, x_{M-1}$, and $T$. This system can be solved by applying a phase fixing condition [19]. Stanford et al. [202] presented a formulation of the cyclic finite-difference method for the Crank–Nicholson time scheme applied to a nonlinear beam in periodic motion.

The need for a phase fixing condition stems from the periodic nature of the solution, and it is common to all the time-periodic and frequency-domain LCO prediction methods. The periodicity condition is satisfied by assuming any time $t$ within the fully developed LCO as the reference initial time $t_0$. Therefore, an additional condition is necessary for obtaining a unique solution. Several techniques are available for phase fixing [19]. One approach is to fix one of the unknown states and determine the remaining ones along with the LCO period by applying the Newton–Raphson method to the remaining system of $NM$ equations. To achieve convergence, the assumed value of the fixed state must lie within the amplitude range of the periodic solution. An appropriate first guess for initializing the Newton–Raphson iterations is also necessary. Another approach to phase fixing is to solve the undetermined system resulting from the application of a time-cyclic (or frequency-domain) method using the Moore–Penrose pseudo-inverse [203]. Alternatively, one may enforce the integral orthogonality phase condition [204] or the Poincaré phase fixing [205]. In any case, cyclic finite differences
requires the solution of a system of $NM$ equations to determine the amplitudes of all the system states over the LCO period, which can result in prohibitive computational costs for large-dimensional models.

The cyclic spectral element method discretizes the time domain into $M$ elements that can have a variable time length. Transforming the time interval $t \in [t_{i-1}, t_i]$ spanned by the $i$th element (which corresponds to a time step) into the non-dimensional interval $\tau \in [-1, 1]$, the solution within the $i$th element can be approximated as a $K$th-order Lagrange polynomial

$$x_i(\tau) = \sum_{k=0}^{K} x_k^i \psi_k^i(\tau), \quad (19)$$

where the unknown coefficients $x_k^i$ are the solution vectors evaluated at the zeros of the $K$th-order Lobatto polynomial, while $\psi_k^i(\tau)$ is the $k$th-order Lagrange polynomial ($k = 1, \ldots, K$). Then, Eq. (19) is substituted into the equations of motion written for each spectral element and the system of nonlinear algebraic equations for the whole LCO period is assembled by enforcing the inter-element continuity and the periodicity boundary condition. An example of this procedure applied to a nonlinear beam in periodic motion is reported by Stanford et al. [202].

As for cyclic finite differences, spectral element methods result in an algebraic system of dimension equal to the product of the number of temporal and spatial DOFs (the additional unknown introduced by the LCO period is eliminated by phase fixing). However, since the solution vector is evaluated at multiple inner points within each time interval (that is, each spectral element), the number of temporal DOFs is generally higher than for cyclic finite differences. This results in larger algebraic systems and thus higher computational costs.

Finally, an alternative time-periodic approach for computing LCOs is to couple a shooting strategy with time integration performed over a single LCO period [206]. The shooting method starts with initial guesses for the initial condition, $x_0^{(0)} := x(t_0)^{(0)}$, and the LCO period, $T^{(0)}$, and numerically integrates Eq. (13) from $t = t_0$ to $t = t_0 + T^{(0)}$ using any time-marching scheme (explicit or implicit, with fixed or variable time step). Then, the residual

$$R(x_0, T)^{(0)} := x(t_0 + T)^{(0)} - x(t_0)^{(0)} \quad (20)$$

is evaluated and the procedure is repeated iteratively by updating the guesses for the initial condition and the period using the Newton–Raphson method until the norm of $R$ is below a desired threshold. Shooting methods result in $N + 1$ unknowns that can be solved after adding one phase fixing condition.

The main advantage of time-periodic methods (cyclic finite differences, cyclic spectral elements, and shooting) over time marching is that they skip the initial transient phase and only compute the fully developed LCO. This is accomplished by solving an algebraic system, which also allows to compute adjoint derivatives more efficiently than in the case of time-domain simulations [202]. On the other hand, solving for a periodic response requires an appropriate phase fixing condition. Furthermore, adequate initial guesses are needed for convergence when solving the algebraic system of equations, which requires previous knowledge of the post-flutter behavior. Another weakness of
cyclic finite differences and cyclic spectral elements is that the number of unknowns to be solved simultaneously may be very high, because it equals the product of temporal and spatial DOFs. This problem is mitigated when using shooting methods, for which the number of unknowns is equal to the number of spatial DOFs. However, these approaches require time integrations over one LCO period in an iterative process, which may be computationally expensive. Finally, time-periodic methods may experience convergence difficulties for highly nonlinear problems and dynamic responses with high harmonic content [202].

Reduced-order modeling (ROM) techniques, such as proper orthogonal decomposition (POD), can be used in combination with time-domain methods to limit the cost of computing adjoint derivatives of post-flutter behavior metrics [202]. However, these techniques are of limited use in design optimization because a ROM is valid only for a single set of design variables, and it must be updated at each optimization step. ROMs can be trained with respect to the design variables as well, but it is currently not feasible for large numbers of design variables.

5.1.2 Frequency-Domain Methods

Frequency-domain post-flutter prediction is based on the harmonic balance methods, which assume that a periodic solution of Eq. (13) exists for a given value of flight speed and seek it in the form of a truncated Fourier series expansion [207]

\[ x(t) = x_{c0} + \sum_{k=1}^{M} \left[ x_{ck} \cos(k\omega t) + x_{sk} \sin(k\omega t) \right] , \]  

where \( k \) is the harmonic order, \( M \) is the number of harmonics, \( \omega = 2\pi/T \) is the fundamental angular frequency of the response, and \( x_{c0}, x_{ck}, x_{sk} \in \mathbb{R}^{N} \) are the vectors of unknown coefficients. Assuming \( M = 1 \) yields the fundamental harmonic balance (HB) method, while higher-order harmonic balance (HOHB) methods are obtained for \( M > 1 \). Variations of the classical balancing method include the nonlinear frequency domain form [208] and the time spectral form [209], which are summarized by Hall [210].

To compute the limit cycle for a given condition, Eq. (21) is substituted into Eq. (13) and the total coefficient multiplying each sine and cosine function, along with the sum of the constant terms, is set to zero. This yields a nonlinear algebraic system of \( N \times (2M+1) \) equations with \( N \times (2M+1) + 1 \) unknowns, which can be written as

\[ g(x_{c0}, x_{ck}, x_{sk}, \omega; U) = 0 . \]  

For a given value of \( U \), this can be solved once an appropriate additional condition is enforced for phase fixing, as discussed for time-cyclic methods.

One advantage of HB and HOHB methods is their applicability to systems of generic complexity, from two-dimensional airfoils [211, 212, 213, 214, 215, 216, 217, 218, 219] to three-dimensional high-fidelity models [220, 210]. Since they use time-periodic schemes, HB methods result in an algebraic system of equations that may provide a starting point for numerical continuation in order to obtain the bifurcation diagram for multiple
values of the parameter $U$ [19]. One weakness of HB methods is that the number of harmonics describing the limit cycle is assumed. HB solutions with few harmonics are inaccurate for complex bifurcation diagrams having multiple folding points (for instance, Fig. 6b shows a case with one folding point), or when the response has a high harmonic content [221]. HOHB methods become computationally expensive as the number of unknowns in Eq. (22) increases. In general, truncating Eq. (21) up to an adequate number of terms requires previous knowledge on the post-flutter behavior of the system, which limits the method applicability for design optimization.

5.1.3 Nonlinear Perturbation Methods

Nonlinear perturbation methods seek a reduced-order representation of the system dynamics close to the bifurcation point, which is used to approximate the bifurcation diagram locally [222]. These methods lose accuracy for large perturbations, but they may still provide sufficient information to determine whether the post-flutter behavior is subcritical or supercritical. Approximate solutions from nonlinear perturbation analyses may be also used as first-guess for time-domain methods, to choose the number of retained harmonics when using frequency-domain methods, or to initialize numerical continuation schemes.

Nonlinear perturbation techniques for the analysis of nonlinear systems are the method of multiple scales, the center-manifold reduction, and the method of normal forms [223, 79, 222]. The method of multiple scales seeks an approximate solution as a series expansion function of different time scales, where the time scales are treated as independent variables. The center-manifold reduction projects the system dynamics onto the subspace defined by the bifurcating mode (the mode that causes flutter). The projected dynamic equations can be then manipulated for obtaining an analytical representation of the bifurcation diagram by means of a coordinate transformation. This latter procedure is known as the method of normal forms.

Previous efforts have applied the method of multiple scales [224, 225, 155] and center-manifold reduction [226, 227, 161] to study nonlinear aeroelastic systems. Nayfeh [228] also compared the performance of the method of multiple scales and the center-manifold reduction combined with the method of normal forms applied to nonlinear systems experiencing Hopf bifurcations. The two methods yielded equivalent normal forms of the bifurcation diagram, but the method of multiple scales required less manipulation of the mathematical model and resulted in a compact solution for the type and amplitude of LCO [79, 228].

Vio et al. [229] compared time-integration, frequency-domain, nonlinear perturbation, and continuation methods for predicting LCOs of a one-degree-of-freedom aeroelastic system. They concluded that nonlinear perturbation methods were less accurate than other methods for large amplitudes. Regardless, the method of multiple scales can be used for obtaining a representation of bifurcation diagrams locally around the flutter point suitable for enforcing post-flutter constraints in optimization [155] (see also Section 5.2). For this reason, we present the key steps in the method below.

Consider the following expansion in a neighborhood of the Hopf bifurcation point
\[(x_F; U_F)\]

\[
x = x_F + \epsilon \hat{x},
\]

\[
U = U_F + \epsilon^2 \hat{U},
\]  

(23)

where \(\epsilon > 0\) is a small non-dimensional quantity. Expanding Eq. (13) as a Taylor series about \((x_F; U_F)\) and using the equilibrium condition yields

\[
\dot{\hat{x}} = A \hat{x} + \epsilon^2 \hat{U} B \hat{x} + \epsilon C(\hat{x}, \hat{x}) + \epsilon^2 D(\hat{x}, \hat{x}, \hat{x}) + \ldots,
\]  

(24)

where \(B := \partial f / \partial U\), \(C\) is generated by a vector-valued symmetric bilinear form, and \(D\) by a vector-valued symmetric trilinear form. These three terms can be evaluated via finite differences [155] or any other method for computing derivatives [31]. The solution of Eq. (24) close to the Hopf bifurcation point is locally sought as

\[
\hat{x} = \hat{x}_1(T_0, T_2) + \epsilon \hat{x}_2(T_0, T_2) + \epsilon^2 \hat{x}_3(T_0, T_2) + \ldots,
\]  

(25)

where \(T_0 = t\) and \(T_2 = \epsilon^2 t\) are different time scales treated as independent variables. To find the solution, we substitute Eq. (25) into Eq. (24) and equate the coefficients of the like powers of \(\epsilon\). After manipulation, the LCO peak-to-peak amplitude \(a\) as function of the speed perturbation is given by

\[
a = \sqrt{\hat{U} \beta_{1r} / \beta_{2r}},
\]  

(26)

where

\[
\beta_1 := \beta_{1r} + i \beta_{1i} = v^T B u,
\]

\[
\beta_2 := \beta_{2r} + i \beta_{2i} = -v^T [2C(u, z_0) + C(u, z_2) + 3D(u, u, u)/4],
\]  

(27)

and the vectors \(u, v \in \mathbb{R}^N\) are the right and left eigenvectors of \(A\) associated with the bifurcating eigenvalue. Finally, the vectors \(z_0, z_2 \in \mathbb{R}^N\) are computed from

\[
A z_0 = -C(u, u)/2,
\]

\[
(2i \omega - A) z_2 = C(u, u)/2.
\]  

(28)

Equation (26) describes the type and amplitude of the LCO in a neighborhood of the bifurcation point. The bifurcation is subcritical if \(\beta_{1r}\) and \(\beta_{2r}\) have opposite sign, since there exists a real solution for the LCO amplitude for \(\hat{U} < 0\) (below the flutter speed). The bifurcation is supercritical if \(\beta_{1r}\) and \(\beta_{2r}\) have the same sign, and this has smaller LCO amplitude for larger \(\beta_{2r}\).

One advantage of the method of the multiple scales is that it gives a compact solution for the LCO amplitude as a function of the speed perturbation \(\hat{U}\) and of the system parameters described by \(\beta_{1,2}\). However, computing the matrices \(A, B, C, D\) is costly for large high-fidelity models and these matrices must be updated at each optimization step. Additionally, the method is not appropriate to constrain the bifurcation diagram for large amplitudes (far from the bifurcation point). For instance, the method may not be adequate to constrain the folding point of subcritical bifurcation diagrams, which may require the computation of the post-flutter dynamic response for large disturbances or for values of the bifurcation parameter much lower that the critical one.
5.1.4 Continuation Methods

Continuation methods are a class of numerical methods that determine how solutions of a nonlinear system vary with certain parameters [230]. In the context of bifurcation theory, these methods trace branches in bifurcation diagrams starting from a known point along the curve (for instance, the flutter point) and the unit vector tangent to the curve at that point. Continuation methods consists of predictor-corrector and piecewise-linear methods [79]. Predictor-corrector methods consist of a parametrization strategy for the branch of solutions to be continued, a predictor that gives an initial guess for the new point on the branch based on a previous known point, a corrector to adjust the prediction, and a step-control strategy for determining the increment in the continuation parameter as moving along the branch. On the other hand, piecewise-linear methods follow a piecewise-linear curve that approximates a branch of solutions.

In this section, we discuss only predictor-corrector methods because they are the most popular choice for tracing bifurcation diagrams using numerical continuation. Numerical continuation is frequently applied to trace Hopf bifurcation diagrams starting from a known point given by a HB solution [18]. The nonlinear system (22) is used below as a practical example. This system can be rewritten as the nonlinear algebraic system

\[
\mathbf{g}(\mathbf{q}; U) = 0, \tag{29}
\]

where \( \mathbf{q} \in \mathbb{R}^M \) is the vector of unknowns. Given a solution \((\mathbf{q}_0, U_0)\), numerical continuation computes how \( \mathbf{q} \) varies for small variations of \( U \) such that Eq. (29) is satisfied. The existence of a unique solution \( \mathbf{q}_1 = \mathbf{q}_0 + \Delta \mathbf{q} \) for \( U_1 = U_0 + \Delta U \) in a neighborhood of \((\mathbf{q}_0, U_0)\) is ensured by the Implicit Function Theorem (IFT) as long as

\[
\det \mathbf{J}|_{(\mathbf{q}_0, U_0)} \neq 0, \tag{30}
\]

where \( \mathbf{J} = \partial \mathbf{g}/\partial \mathbf{q} \) is the Jacobian of \( \mathbf{g} \).

Several predictor-corrector schemes exist with different parametrizations, predictor and corrector methods, and parameter increment strategies. These schemes can be classified under two broad categories: natural parameter and (pseudo) arc-length continuation schemes [231, 232, 233]. Natural parameter schemes perform the continuation by varying the control parameter (sequential continuation), which for the flutter problem would be \( U \). Arc-length and pseudo arc-length schemes perform the continuation with respect to the arc length \( s \) along the solution branch, while the control parameter is treated as an additional unknown. Both methods assume that a starting point along the branch \((\mathbf{q}_0, U_0)\) and local unit tangent vector \( \mathbf{t}_0 = [\mathbf{q}_0^T, U_0']^T \) are known, where the prime denotes the derivative with respect to \( s \).

Natural parameter continuation is the simplest continuation strategy. The range of \( U \) is discretized in \( M \) intervals defined by the values \( U_0, \ldots, U_M \), and the solution \( \mathbf{q}_i \) obtained for \( U_i \) is used as the zero-th order predictor for the solution \( \mathbf{q}_{i+1} \) corresponding to the parameter value \( U_{i+1} \). Then, the predictor is corrected using a Newton–Raphson iteration scheme. Therefore, natural parameter continuation requires that the system Jacobian matrix be nonsingular at any point along the solution path, which causes the method to fail at folds in the bifurcation diagram.
In (pseudo) arc-length continuation schemes, the solution path is parametrized by the arc-length, while $U$ is treated as an unknown quantity. Therefore, Eq. (29) becomes

$$g(q(s); U(s)) = 0 .$$  \hspace{1cm} (31)

Starting from a known solution $(q_0, U_0)$ for $s = s_0$, a tangent (first-order) predictor is used to approximate the solution for $s_1 = s_0 + \Delta s$ as

$$q_1^* = q_0 + q_0' \Delta s ,$$

$$U_1^* = U_0 + U_0' \Delta s .$$ \hspace{1cm} (32)

Arc-length schemes use a Newton–Raphson iteration to correct the predictor in Eq. (32), which still causes convergence problems at folding points in the solution branch. To overcome these issues, pseudo arc-length continuation schemes apply the Newton–Raphson correction the augmented system

$$\begin{cases} 
g(q_1; U_1) = 0 \\
h(q_1; U_1) = 0 \
\end{cases} \hspace{1cm} (33)$$

where $h(q_1; U_1) := n_0 \cdot t_0$ and

$$n_0 = \begin{bmatrix} q_1 - q_1^* \\
U_1 - U_1^* \end{bmatrix}$$ \hspace{1cm} (34)

is the vector from $(q_1^*, U_1^*)$ to the corrected solution $(q_1, U_1)$. The system, Eq. (33), typically has nonsingular Jacobian at folding points, such that the numerical continuation does not encounter convergence issues. This procedure corresponds to seeking the corrected solution as the intersection between $n_0$, which is the direction normal to $t_0$ passing through $(q_1^*, U_1^*)$, and the solution branch.

An advantage of numerical continuation schemes and of (pseudo) arc-length schemes in particular, is that they can compute large portions of bifurcation diagrams. This is particularly useful for characterizing subcritical bifurcation branches for large-amplitude disturbances. Shukla and Patil [219] presented an application of numerical continuation to nonlinear control optimization for suppressing subcritical LCOs (see Section 5.3). A disadvantage of continuation schemes is the need for a known starting point on the solution branch. A parameter sweep on speed or arc length is necessary to trace the branch, followed by a Newton–Raphson solution to obtain each single point. The procedure must then be repeated at each optimization step, incurring a high computational cost for large high-fidelity models. Finally, when numerical continuation is directly applied to the algebraic systems resulting from the application of an HB method, the predicted bifurcation diagram is also a function of the number of retained harmonics (see Section 5.1.2).

5.2 Model-Free Prediction Methods

The post-flutter prediction methods described in Section 5.1 are based on a mathematical model of the system in the form of Eq. (13). Starting from this representation, the
bifurcation diagram can be found using different approaches. However, accurate models of complex realistic systems are not always available; they may be inaccurate due to simplifying assumptions and uncertainties, or they may require too much theoretical development, computational effort, or both.

Model-free techniques that use only output data to forecast bifurcations have been proposed and applied to different types of systems [234]. These methods exploit signals of approaching bifurcations in the system response to predict the bifurcation point and diagram, or at least a portion of it. Although these techniques are naturally oriented towards forecasting bifurcations based on experimental measurements, they can also use data from numerical models. In this case, the model is treated as a black-box, where the model-free technique seeks to approximate the bifurcation diagram without operating on the mathematical representation of the system itself, but only on the output data it produces.

A model-free method for forecasting Hopf bifurcations in aeroelastic systems was proposed based on the critical slowing down (CSD) phenomenon [235]: when the system approaches the flutter point, transient responses to perturbations last longer before the original equilibrium is recovered. Measuring the recovery rate from perturbations in the pre-bifurcation regime allows the forecasting of both the flutter speed and a portion of the bifurcation diagram. The method, originally proposed by Lim and Epureanu [236] and further extended by Epureanu and Ghadami [234] to the application to aeroelastic systems, presents some advantages compared to standard model-based prediction techniques. The method is applicable to large disturbances, which allows the prediction of large portions of the bifurcation diagram. The disturbances may be of any type, including realistic perturbations experienced by the structure during operation. The approach is applicable to both low- and high-frequency oscillatory recoveries, and requires few time histories of the system post-flutter response obtained only in the pre-bifurcation regime. The model is applicable to both experimental and computational data. In this latter case, the model is treated as a black-box and its fidelity does not influence the application of the prediction method. This model-free technique has been applied to characterize the post-flutter behavior of a two-dimensional airfoil [234] and of a idealized cantilevered wing [237]. The main steps in the formulation of Ghadami and Epureanu [234] are summarized below and illustrated in Fig. 9.

Considering a nonlinear system described by one degree-of-freedom, \( r \), the change rate of response amplitude close to the bifurcation point can be expanded as

\[
\dot{r} = r[p(r) + \alpha_1(r)(U - U_F) + \alpha_2(r)(U - U_F)^2 + \ldots],
\]

where \( p(r) \), \( \alpha_1(r) \), and \( \alpha_2(r) \) are polynomial functions independent of \( U \), and \( U_F \) is the flutter speed. This expansion is with respect to the flight speed and not with respect to the amplitude, which is not assumed to be small. From Eq. (35), the recovery rate \( \lambda \) is evaluated as

\[
\lambda(r; U) := \frac{d \ln r}{dt} \approx p(r) + \alpha_1(r)(U - U_F) + \alpha_2(r)(U - U_F)^2,
\]

The order of the polynomial defines the minimum number of recoveries (time responses) from perturbations necessary for the bifurcation forecasting. Since Eq. (36) considers a
Figure 9: Bifurcation forecasting based on the CSD using the method (adapted from Ghadami et al. [237]).

second-order expansion of the recovery rate with respect to the bifurcation parameter, the forecasting requires at least three time responses. However, a higher number can be considered to improve the prediction accuracy. The perturbations can have different amplitudes and can also be of different types (e.g., gust loads, maneuver loads, control-surface inputs).

The steps in the bifurcation forecasting are described below by assuming that \( M \) time responses to perturbations have been computed for \( M \) parameter values \( U = U_1, \ldots, U_M \) in the pre-flutter regime. For a fixed amplitude \( \tilde{r} \), the recovery rates \( \lambda_m = \lambda(\tilde{r}; U_m) \) \( (m = 1, \ldots, M) \) are computed using different techniques depending on the type and frequency content of the recovery [234, 237]. The recovery rates \( \lambda_m = \lambda(\tilde{r}; U_m) \) can be then fitted in the \((U, \lambda)\) plane according to Eq. (36). The intersection between the fitted curve and the axis \( \lambda = 0 \) yields the flight speed \( \tilde{U} \) that corresponds to the point with amplitude \( \tilde{r} \) on the bifurcation diagram (see Fig. 9). Repeating the procedure for different values of \( \tilde{r} \) gives a portion of the bifurcation diagram and the flutter speed, where the flutter speeds corresponds to \( \tilde{r} = 0 \). The bifurcation diagram can be forecast up to the maximum response amplitude \( \tilde{r}_{\text{max}} \) observed during the recoveries, which is not assumed to be small.

Handling large-dimensional systems requires additional considerations because the recoveries may involve contributions from many modes. To accurately evaluate the recovery rate, the contribution from the bifurcating mode—the most affected by the CSD [234, 237]—must be isolated beforehand in the recoveries for the DOFs of interest. Therefore, both projection onto the space spanned by the bifurcating mode [234] and frequency-filtering [237] are used. However, the appropriate approach to isolate the contribution in the dynamic response that better shows the effect of the CSD needs
further investigation in order to apply the method to general aeroelastic systems described by many DOFs. This is particularly true when recoveries are measured far from the flutter point, because the system modes vary as the critical mode may switch or be a hump mode when approaching the bifurcation.

5.3 Applications to Optimization

Optimization studies with post-flutter constraints are rare in the literature. This is because modeling post-flutter efficiently is challenging and because contemporary aircraft configurations do not usually experience global LCOs within the flight envelope. Previous efforts are summarized in Table 2 and reviewed below. These efforts used time marching (TM), time cyclic (TC), harmonic balance (HB), nonlinear perturbation (NP), and numerical continuation (NC) methods for post-flutter prediction and mainly considered simple systems parametrized by few design variables. As mentioned previously in Table 1, optimizations including flutter constraints considered combinations of structural, planform, and airfoil shape design variables. In contrast, optimization including post-flutter constraints have considered either only structural sizing or only airfoil shape, but no planform changes.
Table 2: Summary of sample efforts that included a post-flutter analysis in a design optimization problem.

<table>
<thead>
<tr>
<th>Effort</th>
<th>Method*</th>
<th>Models</th>
<th>Design variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Aerodynamics†</td>
<td>Structure‡</td>
</tr>
<tr>
<td>Missoum et al. [238]</td>
<td>TM</td>
<td>Theodorsen</td>
<td>Mass-spring</td>
</tr>
<tr>
<td>Stanford et al. [202]</td>
<td>TM, TC</td>
<td>–</td>
<td>NL beam</td>
</tr>
<tr>
<td>Stanford and Beran [154]</td>
<td>TM</td>
<td>VLM</td>
<td>Shells</td>
</tr>
<tr>
<td>Stanford and Beran [239]</td>
<td>TC</td>
<td>Piston theory</td>
<td>Shells</td>
</tr>
<tr>
<td>Stanford and Beran [155]</td>
<td>NP</td>
<td>ONERA stall</td>
<td>NL beam</td>
</tr>
<tr>
<td>Thomas et al. [213, 240]</td>
<td>HB</td>
<td>RANS</td>
<td>Mass-spring</td>
</tr>
<tr>
<td>Shakla and Patil [219]</td>
<td>HB+NC</td>
<td>Theodorsen</td>
<td>Mass-spring</td>
</tr>
<tr>
<td>He et al. [241, 242]</td>
<td>TS</td>
<td>Euler</td>
<td>Mass-spring</td>
</tr>
</tbody>
</table>

* Methods: TM—Time marching, TC—Time cyclic, HB—Harmonic balance  
† Aerodynamics: VLM—Vortex lattice method, RANS—Reynolds average Navier–Stokes  
‡ Structures: NL—Nonlinear  
§ Objective: LCO—Limit cycle oscillation, fl.—flutter, displ.—displacement  
‖ Algorithm: GB—Gradient based, GF—Gradient free
Missoum et al. [238] proposed a two-step methodology for the design optimization of nonlinear aeroelastic systems under probabilistic flutter and subcritical post-flutter constraints. In the first step, explicit failure boundaries are constructed in the design space. In a second step, these boundaries are used in a reliability-based optimization. The failure boundaries are obtained from a support vector machine classifier [243]. This is a learning model that classifies samples as belonging to one or the other of two categories—feasible or unfeasible—based on the classification of an initial set of training samples. For flutter, the classification of the training samples is based on the system eigenvalues (frequency-domain approach) or on the time history of the mechanical energy during the recovery from a perturbation (time-domain approach). For subcritical post-flutter behavior, the classification was made by detecting discontinuities in the response amplitude using the $K$-means clustering technique [244]. The methodology was demonstrated for a two-dimensional airfoil with cubic plunge and pitch stiffnesses. The subcritical post-flutter failure boundary was used to minimize the sum of the cubic stiffnesses under a probabilistic constraint, where the stiffness values were uncertain design variables.

Stanford et al. [202] compared the performance of different time-domain methods for gradient-based optimization of nonlinear structures undergoing limit cycles. They analyzed implicit time-marching, cyclic finite differences, and cyclic spectral elements, each with and without a preliminary POD model order reduction. The test case was an actuated nonlinear beam discretized in finite elements, which was optimized with respect to the element cross-sectional properties to either maximize or minimize the tip deflection, where no constraints were enforced. The authors pointed out advantages and disadvantages of time-domain prediction methods (see Section 5.1.1) applied to optimization. The main conclusion was that time-periodic schemes enable more efficient adjoint derivative computation than time marching, but they require previous knowledge on the post-flutter response and suffer convergence issues for highly nonlinear problems or high harmonic content. In a later work, Stanford and Beran [154] performed an LCO-amplitude minimization of a cantilevered wing using a medium-fidelity nonlinear time-simulation framework with analytic derivatives. They pointed out that while computing LCOs via time-marching may be feasible for low- and medium-fidelity gradient-based optimization, the cost of this approach remains prohibitive for large high-fidelity models parametrized by many design variables.

Stanford and Beran [239] also used cyclic spectral elements combined with POD to perform a gradient-based structural optimization of a geometrically nonlinear cantilevered wing. The wing was modeled as a plate-type structure and discretized in finite elements, and was assumed to operate above its flutter speed in supersonic flow. The aerodynamics was modeled via piston theory. The optimization minimized the wing mass with respect to the thickness distribution for both a deterministic and a probabilistic constraint on the supercritical LCO amplitude. The reliability-based optimization starting from the deterministic optimal design increased the reliability index with a limited mass increment with a similar thickness distribution pattern. The authors discussed the scalability of the approach to high-fidelity multidisciplinary problems. Although the combination of a cyclic scheme and POD enabled efficient adjoint derivative computation, the methodology still relies on time marching. Computing
the POD basis at each optimization step and a first-guess for convergence of spectral element solutions requires transient simulations. Therefore, the approach is not well scalable to large-dimensional optimization problems.

Later, Stanford and Beran [155] performed a gradient-based structural optimization of the cantilevered wing studied by Tang and Dowell [81, 82], which showed subcritical post-flutter behavior. The wing was modeled with beam finite elements with leading-order geometric nonlinearities and was discretized with finite elements. The unsteady aerodynamics was modeled using the ONERA stall model [82]. The flutter speed was computed using a direct method [157] and the bifurcation diagram in a neighborhood of the flutter point was approximated using the method of multiple scales [79]. The optimization problem minimized the supercritical post-flutter amplitude with respect the beam stiffness and inertia distribution subject to constant mass and flutter speed constraints. The derivatives were computed by finite differences. The authors showed the effectiveness of the methodology in turning the subcritical bifurcation into a super-critical one and pointed out a contrast between the supercritical post-flutter response amplitude and the flutter speed. The optimal solution obtained using the method of multiple scales was verified by comparing the bifurcation diagram with the one obtained using the spectral element method, showing that the nonlinear perturbation analysis yielded a conservative result.

Thomas et al. [240] presented a discrete adjoint approach to compute derivatives with respect to geometric design variables for an aeroelastic system analyzed using the HB method (see Section 5.1.2). The work extended a discrete adjoint previously developed for a CFD code capable of solving the Reynolds-Averaged Navier–Stokes (RANS) equations for transonic flow conditions in the frequency domain [245]. Derivatives with respect to shape were obtained using AD and verified using finite differences. The proposed method was used to redesign the NLR 7301 airfoil for improved flutter speed and LCO characteristics. They showed that a small increase in thickness-to-chord ratio improved the flutter speed and LCO behavior, such that the reduced velocity was significantly increased while keeping the same baseline pitching amplitudes. In contrast, small variations in camber reduced the flutter speed and deteriorated the LCO characteristics, namely, the reduced velocity decreased for the same baseline pitching amplitudes.

Thomas and Dowell [246] applied their previously developed aeroelastic adjoint implementation [240] to optimize NLR-7301 airfoil for a transonic flow condition. They used a class-shape function transformation technique to parametrize the geometry [247], which considered 14 parameters for each surface for a total of 28 design variables. Three optimization cases were considered: one maximizing the flutter speed and two LCO cases maximizing the reduced velocity subject to a pitch amplitude constraint. The geometry optimized for the flutter speed demonstrated a subcritical LCO behavior, which may call into question the usefulness of the results. However, the airfoil optimized for the reduced velocity demonstrated improved supercritical LCO behavior when enforcing a large pitch amplitude (4 deg). The reduced velocity was improved around the specified pitch amplitude (2–5 deg), but was made worse for lower amplitudes (0–2 deg). Flutter and LCO characteristics were both improved when enforcing a medium pitch amplitude (2 deg).
Shukla and Patil [219] recently performed a control optimization to suppress the subcritical post-flutter behavior of a two-dimensional airfoil with nonlinear pitch stiffness. The bifurcation diagram was computed using the HB method and the Monroe–Penrose numerical continuation [230]. The optimization sought to modify the shape of the subcritical bifurcation diagram in order to minimize the cost of the nonlinear feedback control while enforcing a subcritical post-flutter behavior. Derivatives of the objective with respect to the control variables were computed analytically using the direct method. Using the proposed method, the authors synthesized a control law to suppress the subcritical post-flutter behavior of the open-loop system, resulting in a closed-loop system with a supercritical bifurcation diagram.

He et al. [241] proposed a block-Jacobian preconditioned coupled Jacobian–free Newton–Krylov (CNK) method based on the time spectral (TS) method [209] to compute flutter and post-flutter efficiently. In contrast to the method proposed by Thomas et al. [212], \( O(N_{CSD}) \) CFD evaluations can be avoided. Here, \( N_{CSD} \) is the structural DOFs multiplied by number of time instances. The state variables in the coupled system are thus the aerodynamic and structural states in addition to a given motion magnitude and phase. More recently, He et al. [242] proposed a coupled adjoint method to compute flutter or post-flutter constraint derivatives with respect to the design variables. Similar to the analysis method, \( O(N_{CSD}) \) CFD evaluations can be avoided by solving the coupled adjoint, which directly computes the derivatives of the whole aeroelastic system. Preliminary results were provided for a 2D airfoil section optimized to maximize the flutter speed index with respect to four aerodynamic shape variables.

### 5.4 Open Problems

Post-flutter behavior has been the subject of research from a long time, and several well-established theoretical and computational methods have been developed to study the post-flutter nonlinear dynamic response of aeroelastic systems. Time-domain simulations, periodic time schemes, harmonic balance, nonlinear perturbation methods, and numerical continuation were all applied to analyze nonlinear configurations. However, previous work mainly focused on simple systems like 2D airfoils or cantilevered wings. Furthermore, contemporary aircraft typically show linear behavior, so integrating post-flutter constraints into design has been more a research interest than a practical industrial need.

In order to include post-flutter constraint into aircraft design, one major open problem is predicting the post-flutter behavior of realistic full-scale aircraft with adequate fidelity and feasible computational cost. This is challenging due to difficulties in developing accurate dynamically nonlinear models and the computational effort of performing repeated post-flutter computations, which is costly even for simple systems. For these reasons, only few efforts included post-flutter constraints in an optimization. These optimizations considered systems modeled analytically or simple beam-like or plate-like finite element models and used a small number of design variables. The post-flutter response was typically analyzed using time-domain methods, either direct time-marching or time-cycling schemes, which are hardly scalable to large high-fidelity
models. Furthermore, no optimization subject to post-flutter constraints considered both aerodynamic shape and structural sizing variables.

The performance of different post-flutter prediction methods applied to optimization are still unclear, since limited comparisons and feasibility studies were performed. Therefore, identifying methods providing an adequate compromise of fidelity and computational cost is a critical issue. The methods must be scalable to large-dimensional systems and suited for repeated computations within an optimization cycle. Among the traditional post-flutter analysis methods, only the method of multiple scales and numerical continuation of harmonic-balance solutions seem scalable. However, these methods require computing several matrices for each design point, which incurs a high computational cost for large-dimensional models. Both methods also involve inner Newton–Raphson iterations and may require previous knowledge on the system behavior for reliable convergence, which is not available during optimization. More recent model-free methods are an appealing alternative for formulating post-flutter designs constraints. However, these methods need further development to demonstrate their suitability for practical wing or full-vehicle configurations.

Another critical problem is the choice of a suitable post-flutter constraint metric that does not compromise performance. This is particularly important for avoiding subcritical oscillations before reaching the flutter boundary. A possible approach is to constrain the folding point of the bifurcation diagram rather than the flutter point, but this may result in an overly conservative design solution or introduce discontinuities in the constraint between subsequent optimization iterations. Alternatively, the type of bifurcation can be constrained instead such that subcritical behavior is changed into supercritical one [155, 219]. However, this may have a detrimental effect on the flutter point [155]. Additional investigations are required to explore possible contrasts between flutter and subcritical post-flutter constraints and how to effectively integrate them simultaneously in the design process of configurations prone to nonlinear behavior. Optimizing realistic configurations using gradient-based methods also requires post-flutter constraints that are smooth with respect to the design variables in order to compute their derivatives.

Finally, uncertainties have a strong influence on post-flutter behavior: distinct, but nominally identical aircraft may experience different response at the same flight condition [14], and small changes in the natural frequencies can cause large variations in the response amplitude and in the bifurcation diagram [248]. Therefore, appropriate safety margins or probabilistic constraints are needed to ensure practical optimal solutions [238, 239]. While several methods exist to compute LCOs including uncertainties for a single configuration, design optimization requires effective approaches suited to repeated computations.

### 6 Conclusions

In this paper, we reviewed flutter and post-flutter prediction methods and past work that addressed their integration into optimization in the form of dynamic aeroelastic constraints. We pointed out advantages and disadvantages of existing prediction
approaches and outlined the open problems to be addressed for their application to aircraft MDO, with a focus on next-generation transport vehicles. Optimizing these configurations for reducing fuel burn is likely to give light-weight, high-aspect-ratio structures that are prone to flutter. Therefore, including flutter constraints in the optimization process is critical to obtain feasible designs and avoid major modifications at advanced design stages, which cause performance, time, and financial losses. Despite its potential to speed up and reduce the cost of aircraft production cycle, including flutter constraints in design optimization is challenging. Flutter analysis is computationally expensive due to the number of conditions that must be considered and the large design space required to parametrize realistic configurations. Optimizing the structural sizing and the aerodynamic shape simultaneously is challenging because the aerodynamic properties need to be recomputed at each design iteration. This is particularly critical for transonic flow conditions, due to the high computational cost of CFD analyses.

Next-generation high-aspect ratio aircraft are likely to experience new phenomena, such as couplings between rigid-body and elastic DOFs and geometrically nonlinear effects. These effects need to be considered in the flutter constraints because they significantly impact the flutter point. This further increases the cost of flutter analyses compared to traditional linear analyses based on the undeformed structure and frequently not including rigid-body DOFs. Moreover, nonlinear effects can cause subcritical LCOs even before reaching the flutter boundary. Therefore, additional constraints on the post-flutter response are also needed to avoid undesirable subcritical oscillations.

For the MDO to produce useful results, the above phenomena must be captured by the models used for dynamic aeroelastic prediction, which requires better model fidelity and accuracy. For dynamic aeroelastic computations to be integrated in MDO, these must be performed with acceptable computational costs, which demands fast and robust prediction methods and efficient optimization algorithms.

Starting from the review of prediction methods and previous applications to optimization, we offered a detailed overview on the state of the art in the integration of flutter and LCO into design. Moreover, we also described the open problems to be addressed to make MDO subject to dynamic aeroelastic constraints a standard practice.

Optimizations subject to flutter constraints have been addressed mainly using frequency-domain eigenvalue-analysis methods. Practical MDO of complex aircraft configurations requires to use gradient-based methods and compute derivatives using AD, adjoint methods, or a combination of the two to handle large numbers of design variables. Optimizing both the structural sizing and aerodynamic shape subject to flutter constraints is challenging because the aerodynamic properties need to be recomputed at each design iteration. Moreover, in presence of nonlinear effects the flutter boundary depends on the equilibrium state. Therefore, the structural and aerodynamic properties must be computed not only at each design iteration, but also for each flight condition for a given design. In the presence of aerodynamic or structural nonlinearities, identifying the linearized characteristics about nonlinear steady states at each optimization step is a challenge that has limited the scope of previous stud-
ies. Therefore, there is a need for new approaches to efficiently and robustly identify structural and aerodynamic properties at each optimization step.

Optimizations subject to post-flutter (LCO) constraints have been addressed only for very simple configurations and mostly using time-domain methods, which are currently not scalable to large-dimensional problems. Other techniques, like nonlinear perturbation methods and numerical continuation, were also explored. However, these all require significant computational effort for evaluating each design point and updating several quantities within the optimization cycle. Finally, the frequency-domain methods, the harmonic balance, and the time-spectral method have also been applied in optimization, but only for airfoil problems. Therefore, prediction methods for the efficient and robust LCO computation of practical configurations needs to be developed. Additionally, it is still unclear what metric of the nonlinear dynamic response is the most appropriate to constrain the bifurcation diagram in order to avoid subcritical behavior while not compromising the flutter boundary. Finally, this constraint needs to be formulated in a way that is suitable for gradient-based optimization.

Future research should address the above challenges to effectively apply design optimization to aircraft design subject to flutter constraints. Post-flutter and dynamic aeroelastic constraints will become increasingly important to address the current trend towards more flexible vehicles and increasing aircraft wing spans.

While this paper focuses on aircraft design optimization applications, the methodologies and challenges discussed here are broadly applicable to the optimization of other engineering systems subject to stability and post-critical constraints.

7 Acknowledgments

The material is based upon work supported by Airbus in the frame of the Airbus–Michigan Center for Aero-Servo-Elasticity of Very Flexible Aircraft. Special thanks to Bret Stanford (NASA Langley Research Center) and Eli Livne (University of Washington) for their helpful and comprehensive suggestions.

References


